TEACH YOURSELF ALGEBRA FOR ELECTRICAL CIRCUITS

This page intentionally left blank.

TEACH YOURSELF ALGEBRA FOR ELECTRIC CIRCUITS

K. W. JENKINS

McGRAW-HILL
New York Chicago San Francisco Lisbon London Madrid
Mexico City Milan New Delhi San Juan Seoul
Singapore Sydney Toronto

McGraw-Hill

A Division of The ***McGraw-Hill*** *Companies*

0-07-141471-1

The material in this eBook also appears in the print version of this title: 0-07-138182-1

DOI: 10.1036/0071414711

Want to learn more?

We hope you enjoy this McGraw-Hill eBook! If you'd like more information about this book, its author, or related books and websites, please **click here**.

PREFACE

Many books can be had on the subject of electric circuits. Some are elementary, requiring little mathematical skills, while others require a considerable knowledge of calculus.

This book can be considered a compromise, in that it uses no calculus but does make considerable use of algebra. This includes ordinary algebra and also the special algebras of logic and matrices. All are carefully explained in the text, along with interesting and important applications.

The manner in which the book is used will depend of course upon the individual. Some will wish to start on page 1 and continue on consecutively from that point. Others might want to pick and choose. For instance, on a first reading some might prefer to postpone study of Chapter 11 and jump directly from Chapter 10 to Chapters 12 and 13.

At any rate, I hope that you, as an individual, will find the book interesting and, in the long run, a valuable contribution to your professional advancement.

K. W. Jenkins

This page intentionally left blank.

CONTENTS

CHAPTER 1 **Electric Charge and Electric Field. Potential Difference** 1

- **1.1** Electrification and Electric Charge 1
- **1.2** Coulomb's Law and the Unit of Charge 8
- **1.3** Electric Field Strength 10
- **1.4** Potential difference; the Volt 12

CHAPTER 2 **Electric Current. Ohm's Law. Basic Circuit Configurations** 15

- **2.1** Electric Current 15
- **2.2** Electromotive Force 18
- **2.3** Electrical Resistance. Ohm's Law. Power 21
- **2.4** Some Notes on Temperature Effects 25
- **2.5** The Series Circuit 27
- **2.6** The Parallel Circuit 32
- **2.7** Series-Parallel Circuits 35

CHAPTER 3 **Determinants and Simultaneous Equations** 38

- **3.1** Introduction to Determinants 38
- **3.2** The Second-Order Determinant 39
- **3.3** Minors and Cofactors. Value of any nth-Order Determinant 41
- **3.4** Some Important Properties of Determinants 46
- **3.5** Determinant Solution of Linear Simultaneous Equations 52
- **3.6** Systems of Homogeneous Linear Equations 55

CHAPTER 4 **Basic Network Laws and Theorems** 58

- **4.1** Introduction 58
- **4.2** Kirchhoff's Current Law 58
- **4.3** Kirchhoff's Voltage Law 60
- **4.4** The Method of Loop Currents 62
- **4.5** Conductance. Millman's Theorem 66
- **4.6** Thevenin's Theorem 68
- **4.7** Norton's Theorem 70
- **4.8** The Method of Node Voltages 73

CHAPTER 5 **Sinusoidal Waves. rms Value. As Vector Quantities** 76

5.1 Introduction 76
5.2 The Sinusoidal Functions and the Tangent Function 77
5.3 Graphics. Extension beyond 90 Degrees, Positive and Negative 80
5.4 Choice of Waveform. Frequency. The Radian 88
5.5 Power; rms Value of a Sine Wave of Voltage or Current 93
5.6 Sinusoidal Voltages and Currents as Vectors 96
5.7 Power Calculations 105
5.8 Application of Loop Currents 108

CHAPTER 6 **Algebra of Complex Numbers** 114

6.1 Imaginary Numbers 114
6.2 Complex Numbers. Addition and Multiplication 119
6.3 Conjugates and Division of Complex Numbers 120
6.4 Graphical Representation of Complex Numbers 122
6.5 Exponential Form of a Complex Number 125
6.6 Operations in the Exponential and Polar Forms. De Moivre's Theorem 128
6.7 Powers and Roots of Complex Numbers 131
6.8 Complex Numbers as Vectors 134

CHAPTER 7 **Inductance and Capacitance** 136

7.1 Introduction 136
7.2 Introduction to Magnetism 137
7.3 Electromagnetism 138
7.4 Self-Inductance 140
7.5 The Unit of Inductance 142
7.6 Capacitors and Capacitance 144
7.7 Capacitors in Series and in Parallel 148

CHAPTER 8 **Reactance and Impedance. Algebra of ac Networks** 151

8.1 Inductive Reactance. Impedance 151
8.2 *RL* Networks 155
8.3 Capacitive Reactance. *RC* Networks 160
8.4 The General *RLC* Network. Admittance 165
8.5 Real and Apparent Power. Power Factor 169
8.6 Series Resonance 174
8.7 Parallel Resonance 180

CHAPTER 9 **Impedance Transformation. Electric Filters** 187

9.1 Impedance Transformation. The "L" Section 187
9.2 The "T" and "Pi" Equivalent Networks 190
9.3 Conversion of Pi to T and T to Pi 196

9.4 Impedance Transformation by T and Pi Networks 198
9.5 Frequency Response. The Basic *RC* and *RL* Filter Circuits 201
9.6 The Symmetrical T Network. Characteristic Impedance 213
9.7 Low-Pass Constant-*k* Filter 219
9.8 High-Pass Constant-*k* Filter 223

CHAPTER 10 **Magnetic Coupling. Transformers. Three-Phase Systems 227**

10.1 Introduction to Magnetic Coupling; the Transformer 227
10.2 Dot-Marked Terminals. Induced Voltage Drops 230
10.3 Sinusoidal Analysis of Magnetically Coupled Circuits 234
10.4 The "T" Equivalent of a Transformer 239
10.5 The Band-Pass Double-Tuned Transformer 241
10.6 The Ideal Iron-Core Transformer 250
10.7 The Three-Phase Power System. Introduction 255
10.8 Y-Connected Generator; Phase and Line Voltages 256
10.9 Current and Power in Balanced Three-Phase Loads 261
10.10 The Unbalanced Case; Symmetrical Components 265
10.11 Some Examples of Unbalanced Three-Phase Calculations 272

CHAPTER 11 **Matrix Algebra. Two-Port Networks 277**

11.1 Introduction to Matrix Algebra 277
11.2 Product of Two Matrices 281
11.3 The Inverse of a Square Matrix 286
11.4 Some Properties of the Unit Matrix 291
11.5 Algebraic Operations. Transpose of a Matrix 292
11.6 Matrix Equations for the Two-Port Network 294
11.7 Continuing Discussion of the Two-Port Network 299
11.8 Matrix Conversion Chart for the Two-Port Network 303
11.9 Matrix Operations for Interconnected Two-Ports 306
11.10 Notes Regarding the Interconnection Formulas 312
11.11 Some Basic Applications of the Formulas 316

CHAPTER 12 **Binary Arithmetic. Switching Algebra 324**

12.1 Analog and Digital Signals. Binary Arithmetic 324
12.2 Boolean or "Switching" Algebra. Truth Tables 338
12.3 Digital Logic Symbols and Networks 347

CHAPTER 13 **The Digital Processor. Digital Filters 357**

13.1 Bandwidth Requirements for Digital Transmission. Sampling Theorem. PAM and PCM 357
13.2 Analog Signal in Sampled Form. Unit Impulse Notation 364
13.3 The *z*-Transform 366
13.4 The Inverse *z*-Transform 373

Contents

13.5 The Discrete-Time Processor 377
13.6 The Form of, and Basic Equations for, a DT Processor 379
13.7 Stability and Instability. Poles and Zeros 383
13.8 Structure of DT Processors 389
13.9 Digital Filters; The Basic Algebra 393

Appendix 401

Note 1. Some Basic Algebra 401
Note 2. Fundamental Units 404
Note 3. Prefix Nomenclature 405
Note 4. Vectors 405
Note 5. Increment (Delta) Notation 409
Note 6. Similar Triangles. Proof of Eq. (98) 410
Note 7. Identity for $\sin(x + y)$ 411
Note 8. Often-Used Greek Letters 412
Note 9. Sinusoidal Waves of the Same Frequency 412
Note 10. Sinusoidal Waves as Vectors 413
Note 11. Rational and Irrational Numbers 414
Note 12. The Concept of Power Series 415
Note 13. Series *RL* Circuit. L/R Time Constant 416
Note 14. Series *RC* Circuit. *RC* Time Constant 417
Note 15. ωL is in Ohms 418
Note 16. $j\bar{Z} = \bar{Z}$ Rotated through 90 Degrees 419
Note 17. $1/\omega C$ is in Ohms 419
Note 18. Harmonic Frequencies. Fourier Series 419
Note 19. Logarithms. Decibels 421
Note 20. Phase (Time-Delay) Distortion 423
Note 21. Logarithmic Graph Paper 425
Note 22. $\text{Log } XY = \text{Log } X + \text{Log } Y$ 426
Note 23. Discussion of Eq. (344) 426
Note 24. Amplitude Modulation. Sidebands 427
Note 25. Trigonometric Identity for ($\sin x \sin y$) 429
Note 26. L Proportional to N^2 429
Note 27. Arrow and Double-Subscript Notation 430
Note 28. Square Root of 3 in Three-Phase Work 431
Note 29. Proof of Eq. (467) (True Power) 432
Note 30. The Transistor as Amplifier 432
Note 31. Shifting Theorem 434
Note 32. Unit Impulse 435
Note 33. Algebraic Long Division 437

Solutions to Problems 440

Index 551

TEACH YOURSELF ALGEBRA FOR ELECTRICAL CIRCUITS

This page intentionally left blank.

CHAPTER 1

Electric Charge and Electric Field. Potential Difference

1.1 Electrification and Electric Charge

It is an experimental fact that a glass rod, after being briskly rubbed with a silk cloth, has the ability to attract bits of paper, straw, and other light objects to it. A glass rod in such a condition is said to be *electrified* or *charged*, and to contain a kind of "electric fluid" we'll call *electric charge*.

Glass is not the only substance that can be electrified by friction (rubbing), as almost all substances have this property to a greater or less degree.

If a body is *not* electrified it is said to be in an electrically *neutral* condition. Thus, a glass rod that has not been rubbed by a cloth is in an electrically neutral condition.

Suppose we have a glass rod equipped with a rubber handle, as in Fig. 1. Let us suppose the glass rod has been charged by some means, as by rubbing with a silk cloth.

We will find that as long as we hold the assembly *by the rubber handle* the rod will *stay electrified*, that is, will continue to "hold its charge" for a long period of time. This is because

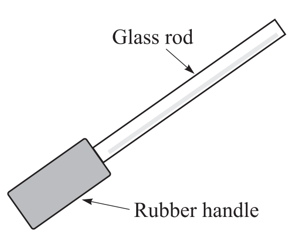


Fig. 1

the rubber handle is a *good electrical* INSULATOR, meaning that it does not allow the charge on the rod to leak off through it to the neutral earth.

Thus, an electrical "insulator" is any substance that offers *great opposition* to the movement or flow of electric charge through it. Rubber, porcelain, and dry wood are examples of good insulating materials.

On the other hand, almost all *metals* offer *very little opposition* to the movement or flow of electric charge through them, and are said to be *good* CONDUCTORS of electric charge. Silver, copper, and aluminum, for example, are examples of very good conductors of electric charge.

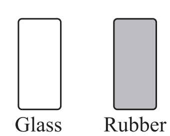
Of course, there is no such thing as a "perfect" insulator or conductor. A perfect insulator would allow *no* movement of charge through it, while a perfect conductor would offer *no opposition* to the flow of charge through it. For many practical purposes, however, substances like rubber, stone, quartz, and so on, can be considered to be perfect insulators, while substances like silver, copper, and gold can be considered to be perfect conductors of electric charge.

Now suppose, in Fig. 1, that the glass rod is replaced by a charged copper rod. If we hold the assembly by means of the rubber handle only, the copper rod will of course continue to hold its charge. If, however, the charged rod is touched to a metal stake driven a foot or so into the earth (down to where the soil is moist), tests will then show that the copper rod *has lost its electric charge*. The explanation is that the charge carried by the rod was "drained off" into the earth through the metal stake, thus putting the rod back into its original uncharged, neutral condition.

It should be pointed out that the earth is such a huge body that we are not able to change its state of charge to any noticeable degree; hence we will consider the earth to be, overall, an electrically neutral body at all times.

Since we mentioned "moist earth" above, it should be mentioned that chemically pure water is a poor conductor. However, most ordinary tap water contains traces of metallic salts, and so on, so that such water is a fairly good conductor of charge. This brings up the point that, when making an electrical connection to the earth, we should go deep enough to get into moist soil; thus, a metal pipe driven only a short distance into dry soil would not be effective in conducting electric charge to and from the earth.

As mentioned before, all substances can be electrified by friction (rubbing). We have already found that a glass rod becomes highly electrified when rubbed briskly with a silk cloth. In the same way, we find that a *hard rubber rod* becomes electrified when rubbed with a piece of cat's fur. Such a rubber rod, when electrified, will attract to it bits of paper and straw just as does an electrified glass rod. Experiment, however, shows there is some kind of *fundamental difference* between the charge that appears on the glass rod and the charge that appears on the rubber rod. To investigate further, let us denote glass and rubber rods as shown below.



We can now perform an experiment that will demonstrate that there are TWO KINDS *of electric charge*, one of which we will call "positive" and the other "negative." The procedure is as follows.

Let us charge two glass rods by rubbing with silk cloth, and two hard rubber rods by rubbing with cat's fur. Let us suppose the rods are then suspended from the ceiling by means of dry silk strings. The dry strings are insulators which will prevent the charges from leaking off the rods, and yet will allow the rods to swing freely. We now observe the three experimental results shown in Fig. 2.

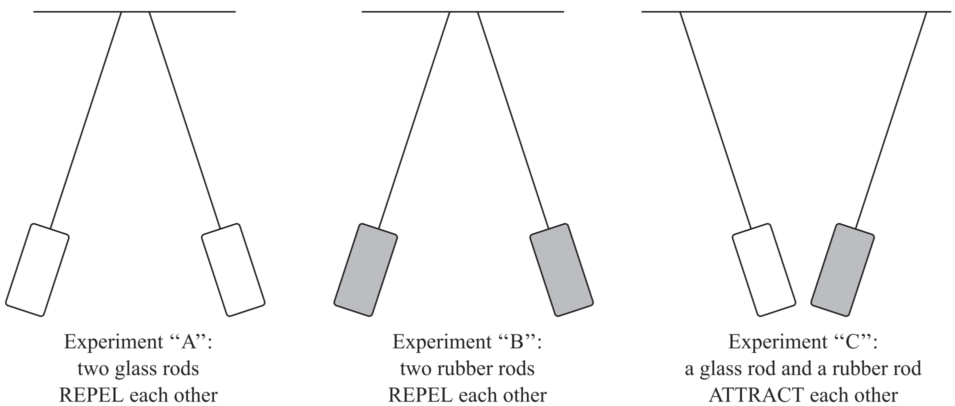


Fig. 2

Since both glass rods were charged by the same means (rubbing with silk cloth), it follows that both glass rods carry the same type of charge. Likewise, since both rubber rods were charged by the same means (rubbing with cat's fur), it follows that both rubber rods carry the same type of charge.

It follows, then, that *if* a glass rod carried the same kind of charge as a rubber rod, then a glass rod and a rubber rod would *repel* each other, but experiment C shows they *attract* each other. Therefore the type of charge on the glass rod must be *different* from the type of charge on the rubber rod.

> So far, then, experiments A, B, and C show there are *at least* TWO *different kinds of electric charge*. The kind appearing on the glass rod is called POSITIVE *electric charge*, and the kind appearing on the rubber rod is called NEGATIVE *electric charge*.

Now consider the following. As mentioned before, all substances can be charged by friction to a greater or less degree. Let us charge, by identical means, two rods *both* made of the same substance "x," which can be any material we wish to test. Since both rods are made of the same material, and both are charged by the same means, it follows that both rods will carry the same kind of charge. Experiment then shows that any two such rods that carry the same kind of charge *will always* REPEL *each other*. Such experiments establish the general rule that LIKE CHARGES ALWAYS REPEL EACH OTHER.

We next make a series of experiments to see what reaction there is between a charged rod of any material x and charged rods of glass and hard rubber. Here is what we find.

1. If a charged rod of any substance x *repels* a charged rod of glass, it will *attract* a charged rod of rubber; hence in this case the rod of substance x carries the *same kind of charge as the glass rod*, which is "positive" charge.
2. If, on the other hand, a charged rod of any substance x *attracts* a charged rod of glass, it will *repel* a charged rod of rubber; hence in this case the rod of substance x carries the *same kind of charge as the rubber rod*, which is "negative" charge.

Hence we can now summarize that

> As far as we can determine by experiment there are TWO KINDS *of electric charge*. For reference purposes, the type that appears on a glass rod rubbed with silk cloth is POSITIVE *charge*, and the type that appears on a hard rubber rod rubbed with cat's fur is NEGATIVE *charge*. Experiment verifies the general rule that LIKE CHARGES REPEL EACH OTHER *and* UNLIKE CHARGES ATTRACT EACH OTHER.

It should be pointed out that an electrically *neutral* body contains EQUAL AMOUNTS of positive and negative charges. If, however, some of the *negative charge* is *removed* from the body, then that body is left with *more positive charge than negative charge*, and therefore becomes a *positively charged body*. Or, if some of the *positive charge* is *removed* from a body, the body is left with an *excess of negative charge* and therefore becomes a *negatively charged body*.

Of course, if positive charge is added to a neutral body, then that body becomes a "positively charged body." Or, if negative charge is added to a neutral body, that body then becomes a "negatively charged body."

It should also be mentioned that, while electric charge can be transferred from one body to another, *it can never be destroyed*; this is a basic law of nature, and is known as THE PRINCIPLE OF CONSERVATION OF ELECTRIC CHARGE.

Let us next discuss *induced* electrical charges. Suppose we have a round ball of conducting material (aluminum, for instance), resting on a dry insulating stand, as shown in Fig. 3, where it's assumed the aluminum ball is in an electrically neutral state.

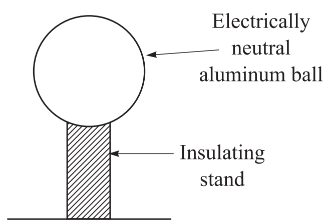


Fig. 3

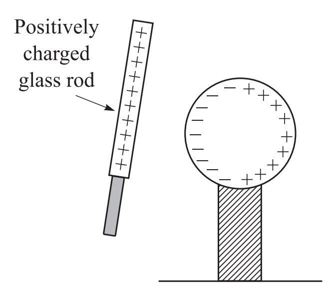


Fig. 4

Let us now bring a positively charged glass rod up near to (but not touching) the aluminum ball, as in Fig. 4. Remember that the ball is electrically neutral, that is, it contains equal amounts of positive and negative charge.

Now, since LIKE CHARGES REPEL and UNLIKE CHARGES ATTRACT, we will find that a portion of the positive charge in the ball will be *repelled* over to the right-side of the ball, and a portion of the negative charge in the ball will be *attracted* over to the left-side of the ball. This action will result in a concentration of *positive charge* on the right-side of the ball and a concentration of *negative charge* on the left-hand side of the ball, as illustrated in Fig. 4.

The concentrations of positive and negative charges on the ball in Fig. 4 are examples of *induced* electric charges. Thus, an "induced" charge is a concentration of positive or negative charge on a region of a body, due to the *nearness* of a charged body. In the experiment of Fig. 4 we are allowed to bring the glass rod as close to the aluminum ball as we wish, as long as the rod does not touch the ball. Of course, the closer we bring the charged glass rod to the ball, the greater is the degree of separation of the charges in the ball.

It should be noted that, in Fig. 4, the ball *considered as a whole* is still an electrically neutral body, even though there is localized separation of charges on the ball. If we were to completely *withdraw* the charged glass rod, the separated charges on the ball would come back together again, restoring the ball to the condition it was in Fig. 3.

Now suppose the charged glass rod, in Fig. 4, is allowed to *touch* the ball for a moment and then is pulled away, out of the vicinity of the ball. To understand what would happen in this case, remember that the ball, considered as a whole, is electrically neutral before it is touched by the rod. The glass rod, however, is *not* neutral; it carries more positive charge than negative charge.

Hence, when the rod *touches* the ball, part of the excess positive charge, on the rod, *will flow over to the ball*. Then, when the rod is pulled away, part of the excess charge will remain on the ball and part will remain on the rod. Just what proportion passes over to the ball, and what proportion remains on the rod, depends on several factors, such as relative areas of rod and ball, and so on. We can summarize what has been said about Figs. 3 and 4 so far, as follows.

> In Fig. 3 we start off with an insulated, electrically neutral metal ball, that is, the ball contains equal amounts of positive and negative charges.
>
> In Fig. 4, a positively charged rod is brought near the ball. If the charged rod is now withdrawn from the vicinity of the ball without touching it, then the ball returns to the original condition of Fig. 3. While the charged rod is near the ball, induced charges appear on the ball, as indicated in Fig. 4.
>
> If, however, the rod *touches* the ball, and then is withdrawn from the vicinity of the ball, *then the ball remains permanently charged*. (Actually, since there's no such thing as a perfect insulator, the charge will very slowly leak off to the neutral earth through the insulating stand.)

Thus we see that one way to charge an insulated body, such as the ball of Fig. 3, is to momentarily touch it with a charged body, such as the charged rod of Fig. 4.

Let's continue now with the idea of charge and movement of charge. We know that an electrically neutral body contains equal amounts of positive and negative charges. Suppose, now, that we wish an electrically neutral body to become *positively charged*. We can accomplish this by either *adding positive charge to the body*, or *removing negative charge from the body*.

Either way, the body, which was neutral to begin with, ends up a positively charged body. It is important to notice that, from an external, mathematical standpoint, it makes no difference whether we assume that positive charge flows *into* the body or negative charge flows *out* of the body.

At this point we might digress just a moment to say a word about "electrons." Very briefly, electrons are tiny charges of *negative* electric charge. The flow of charge, in a metallic conductor, such as a copper wire, is known to actually be a flow of negative charges (electrons). But electrons are not the only carriers of moving electric charge; *positive* charge carriers, in the form of positive "ions," are also important charge carriers, especially in liquids and gases.

To continue, suppose we want an electrically neutral body to become *negatively* charged. We can accomplish this by either *adding negative charge to the body*, or *removing positive charge from the body*.

Either way, the body, which was neutral to begin with, ends up as a *negatively* charged body. Again, it is important to note that, as far as the final result is concerned in any such single experiment, it makes no difference whether we assume that negative charge flows *into* the body or positive charge flows *out* of the body. However, in order that our mathematical equations be consistent, and that our notation always mean the same thing, we must select *one standard procedure* and then stick with that procedure or convention.

Hence, except for any special cases where we might say otherwise, let us now agree to use the following conventions when dealing with charged bodies and movement of charge.

1. A positively charged body is one having an EXCESS *of positive charge*.
2. A negatively charged body is one having a DEFICIENCY *of positive charge*.
3. *Only* POSITIVE CHARGE is free to move or flow.

As a first illustration of these conventions, consider the insulated, positively charged body A in Fig. 5. Notice that the switch (SW) is "open," which prevents any movement of charge along the copper wire.

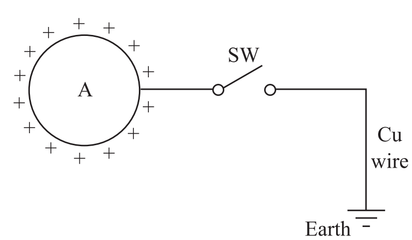


Fig. 5

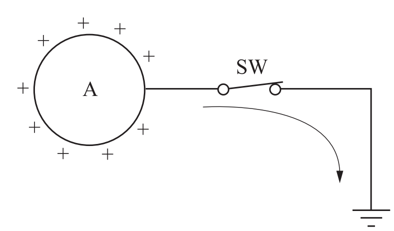


Fig. 6

If the switch is now *closed*, as in Fig. 6, positive charge commences to flow *from body A to the neural earth* as shown by the arrow in Fig. 6. Charge continues to flow until body A becomes electrically neutral with respect to the earth, at which time charge then ceases to flow.

Or, consider the insulated *negatively charged* body B in Fig. 7. If the switch is now closed (Fig. 8), positive charge commences to flow *from the earth to the body B* as shown by the arrow in the figure. Charge continues to flow until body B becomes electrically neutral, at which time charge ceases to flow. It should be remembered that the earth is an electrically neutral body containing, for all practical purposes, an unlimited supply of equal positive and negative charges.

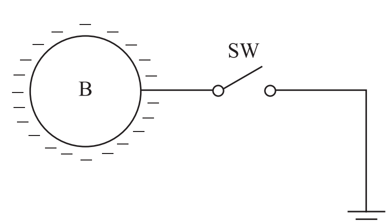


Fig. 7

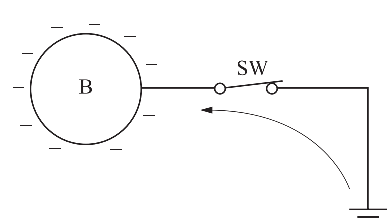


Fig. 8

As another example, consider Fig. 9, which shows a body A positively charged and a body B negatively charged, both bodies being insulated from the earth in this example.

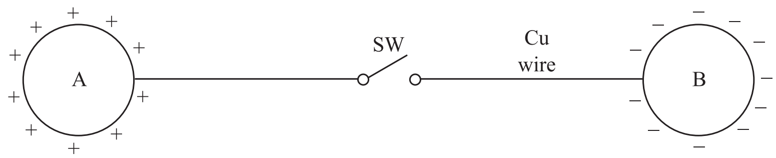


Fig. 9

For discussion purposes, suppose body A contains an *excess* of 100 units of positive charge and body B contains a *deficiency* of 20 units of positive charge. Notice that the two bodies together have a combined *excess* positive charge of 80 units.

If the switch is now closed, positive charge will flow *from body A to body B* until both bodies have an excess of positive charge. Thus, assuming A and B to be identical aluminum balls, charge will cease to flow through the copper wire when both balls have an excess positive charge of 40 units each.

As a final example, consider bodies A and B in Fig. 10. We'll assume they are identical aluminum balls.

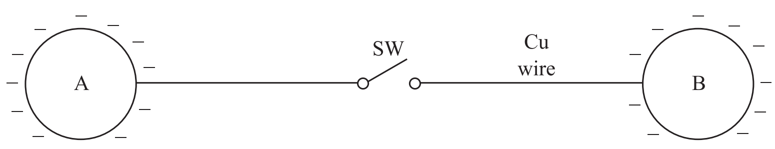


Fig. 10

Notice that *both* bodies are shown as negatively charged; that is, *both* bodies have a *deficiency of positive charge*. Just for discussion purposes, let's assume that

body A has a deficiency of 80 units of positive charge,

body B has a deficiency of 30 units of positive charge.

Note that the two bodies have a combined total deficiency of 110 units of positive charge.*

What happens when the switch in Fig. 10 is closed? To answer this, we must keep in mind that a negatively charged body simply does not have enough positive charge to completely neutralize the negative charge. For practical purposes, however, any large material body, such as a copper penny, a glass rod, and so on, has an *inexhaustible* or unlimited supply of both positive and negative charges (see footnote). All we can do is merely upset the *balance* of charge, positive or negative, either side of the neutral charge

---

* It may be helpful to understand that from a practical standpoint it is impossible for us to drain anywhere near *all* the positive or negative charges from a body of any ordinary size; any such body, for practical purposes, contains an *unlimited* supply of positive and negative charges. Take, for example, two ordinary copper pennies. IF we could withdraw *all* the positive charge from one of the pennies and *all* the negative charge from the other, the two pennies would then have unlike charges and would thus attract each other. Calculation shows that if the two pennies were ONE MILE APART the force of attraction between them would be over SIX BILLION TONS. The point we wish to make is that while bodies A and B above have less positive charge than negative charge, each *still* possesses an enormous amount of positive charge. It is only when we deal with individual atoms or molecules that we can have complete or nearly complete charge removal.

condition of a body. Therefore, when the switch in Fig. 10 is closed, *positive charge flows from body B to body A* until each body has an *equal deficiency* of 55 units of positive charge.

There are no problems here, but this section should be read and reread until you have all the facts firmly in mind.*

1.2 Coulomb's Law and the Unit of Charge

We have learned that two types of electric charge exist, one type being called *positive* and the other *negative*. If a body contains equal amounts of both types it is said to be in an *electrically neutral* condition. If it contains more positive charge than negative charge it is said to be *positively charged*, or if it contains more negative than positive charge it is said to be a *negatively charged body*.

The amount or quantity of excess electric charge carried by a body is denoted by $\pm q$ or $\pm Q$, the sign used depending on whether the excess charge is positive or negative. We recall that bodies carrying excess amounts of like charge REPEL each other, while bodies carrying excess amounts of unlike charge ATTRACT each other.

What is called an ELECTRIC FIELD always exists in the three-dimensional space surrounding an electric charge or group of electric charges. If the charges are *at rest* (that is, are "stationary" or "static" relative to our frame of reference), they are called *electrostatic charges*, and the fields produced by such charges at rest are called *electrostatic fields*. The behavior of charges at rest, that is, electrostatic charges, and the fields produced by them, is the subject of this and the next two sections.

The UNIT AMOUNT *of electric charge* is called the *coulomb* ("KOO lohm"), in honor of the French physicist Charles Coulomb. Coulomb, who published the results of his experiments in 1785, showed that the FORCE OF ATTRACTION OR REPULSION between two quantities of electric charge, q_1 and q_2, is *directly proportional to the product of the two charges* and *inversely proportional to the square of the distance between them*. This is known as "Coulomb's law," which takes the mathematical form

$$F = \frac{k}{K}\frac{q_1 q_2}{r^2} \tag{1}$$

where F is the magnitude of the force of attraction or repulsion between the two charges q_1 and q_2, and r is the distance between them.† The meaning of the constants k and K will be explained in the following discussion, but first let us discuss the meaning of, and the restrictions placed on, eq. (1).

In eq. (1), it is assumed that q_1 and q_2 are "point charges," that is, that the charges q_1 and q_2 are concentrated on bodies *whose dimensions are very small compared with the distance r between them*. Consider, for instance, the two charged spheres in Fig. 11.

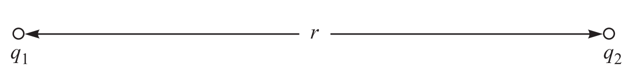


Fig. 11

For instance, if the spheres in Fig. 11 are 0.1 inches in diameter and are separated a distance of, say, 10 inches, they would, for all practical purposes, behave as two point charges for which $r = 10$ inches.

---

* Also see note 1 in Appendix.

† "q" will always denote "electric charge."

You may recall that Newton's third law states that to every force there is an equal but oppositely directed force. Thus the forces acting on the above point charges have *equal magnitudes* (given by eq. (1)), but *point in opposite directions* along the straight line drawn through the two charges. This is illustrated in Fig. 12, for the case of two like charges (which repel each other) and two unlike charges (which attract each other). We've considered force acting to the right to be "positive" and force acting to the left to be "negative."

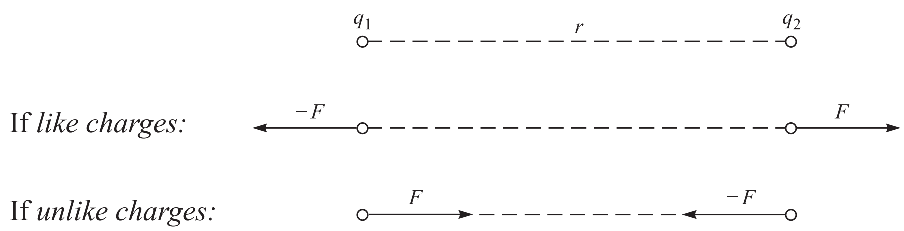


Fig. 12

Let's next discuss the meanings of the constants k and K in eq. (1). We begin by pointing out that the value of the force of attraction or repulsion between two charges depends not only on the values of the charges themselves and the distance between them, but also upon the *medium* that surrounds the charges. For instance, the force action between two charges immersed in say *mineral oil* (just as an example) is considerably different from what it would be if the same two charges were the same distance apart in air.

The medium surrounding the charges is called the DIELECTRIC, and the effect of the dielectric is taken into account, in eq. (1), by means of the *dielectric constant* K, the value of K depending upon the type of dielectric the charges are immersed in. The dielectric constant K is defined as the ratio of the force *in vacuum* to the force *in the given dielectric*. K is thus a dimensionless constant (the ratio of one force to another force), and is given the arbitrary value $K = 1$ *for vacuum* (also, $K = 1$ for air dielectric, for all practical purposes). Thus, for vacuum or air dielectric eq. (1) becomes

$$F = kq_1q_2/r^2 \tag{2}$$

Next, the value of k above will depend upon the *units* that we choose to measure force, distance, and charge. Since we'll use the more practical engineering meter-kilogram-second (mks) system,* force will be measured in *newtons*, distance in *meters*, and charge in *coulombs*.

For these units we find that k is approximately equal to 9×10^9, and thus, for mks units, eq. (2) becomes

$$F = \frac{(9 \times 10^9)q_1q_2}{r^2} \tag{3}$$

where F = force in newtons, the qs are electric charges in coulombs, r = distance in meters.

Let us set $q_1 = q_2 = 1$, and $r = 1$, in the above; doing this gives a force F of

$$F = 9 \times 10^9 \text{ newtons} = 1 \text{ million tons, approx.}$$

Thus, in Fig. 11, if q_1 were a positive charge of 1 coulomb and q_2 a negative charge of 1 coulomb, and $r = 1$ meter, the force of attraction between the two charges would be approximately *1 million tons*. From this, it's apparent that it's impossible, in the real world, to have large *separated* concentrations of electric charges. Here we emphasize

---

* See note 2 in Appendix.

the word ''separated.'' An ordinary copper penny, for example, contains about 130,000 coulombs of positive charge and 130,000 coulombs of negative charge, but the charges are not separated but are ''mixed together'' uniformly throughout the penny. Hence the penny is, overall, an electrically neutral body, with zero net force acting upon it.

Problem 1
Calculate the force of attraction between two unlike charges of 6 microcoulombs* each, separated a distance of one-fourth of a meter in air. Answer in pounds.

1.3 Electric Field Strength

In section 1.2 we pointed out that an ''electric field of force'' always exists in the three-dimensional space surrounding an electric charge or group of charges. If the charges are at rest they are called ''electrostatic charges'' and the fields produced by such charges are called ''electrostatic fields.''

Electrostatic fields are represented graphically by imaginary ''lines of electric force'' or ''field lines.'' A *field line* is any path, in the field, along which a small positive ''test charge'' would naturally be propelled if it were free to move in the field.

The simplest configuration of ''field lines'' exists in the space around a single isolated charge, such as around a positive charge $+q$, as illustrated in Fig. 13. In the figure, the charge $+q$ is assumed to be present on a small spherical surface. Figure 13 is thus a cross-sectional view in the three-dimensional space including the central charge $+q$.

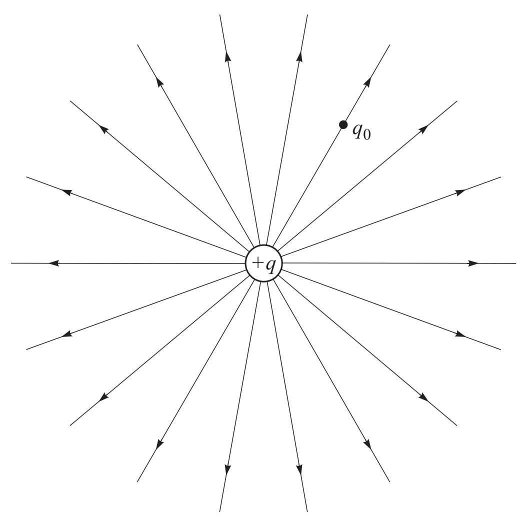


Fig. 13. (Note the small positive ''test charge'' q_0.)

Also shown in Fig. 13 is a *very small positive test charge*, as mentioned above, and denoted by ''q_0'' (q sub zero) in the figure. In this particular case the test charge q_0 would experience a *force of repulsion* away from the central positive charge $+q$, and therefore the ''direction arrowheads'' on the field lines point outward, as shown.

* See note 3 in Appendix.

In dealing with electrostatic fields, the small test charge q_0 is understood to always be a *positive* charge. Thus, *if* the central charge in Fig. 13 were a *negative* charge $-q$, the positive test charge q_0 would experience a force of *attraction* instead of repulsion, and the arrowheads on the field lines would point *inward* toward the central charge $-q$, and would end or "terminate" on the spherical surface in Fig. 13.

In connection with the last statement we have the following point to make. Since the direction of the field lines is defined as the direction in which a *positive* test charge would move, or tend to move, it follows that electrostatic lines of force *go from positively charged bodies to negatively charged bodies*; that is, electrostatic lines of force *originate on positively charged bodies* and *terminate on negatively charged bodies*. In Fig. 13 we cannot, of course, show the outward-going lines as terminating on a negative charge, because Fig. 13 illustrates the hypothetical case of a single isolated charge a very great distance from any other charge or charges. This fact, of the lines originating on positive surfaces and terminating on negative surfaces, will be evident later, when we sketch the field of closely spaced charges.

Next, *the* STRENGTH *of an electric field* at any point in the field is defined in terms of the *force* that a very small positive test charge would experience if placed at the point in question. Since force is a vector quantity* field strength is also a vector quantity.

To be specific, *the* ELECTRIC FIELD STRENGTH *at any point* is denoted by $\bar{E}$ and is defined as the ratio of the *force in newtons* to the *charge in coulombs* carried by a very small positive test charge placed at the point in question. Thus the concise definition of "electric field strength" at a point is

$$\bar{E} = \bar{F}/q_0 \tag{4}$$

where $\bar{F}$ is the *force in newtons* experienced by a very small positive test charge of q_0 *coulombs* when placed at the point. We can imagine that the test charge q_0 is allowed to become vanishingly small, so that its presence in the field does not in any way affect the charge distribution on the bodies that are producing the field.

Equation (4) shows that electric field strength is measured in *newtons per coulomb* which, as we'll show in the next section, is the same as "volts per meter."

With the preceding in mind, the equation for the field strength at any point in the electric field of an isolated charge q (Fig. 13) can be found as follows. First, in eq. (3) set $q_1 = q$, $q_2 = q_0$, $k = 9 \times 10^9$ and let us define that $\bar{u}$ is a *unit vector* (a vector of magnitude 1, having the same direction as the force vector $\bar{F}$ that acts on q_0). Taking these steps, eq. (3) becomes, for Fig. 13,

$$\bar{F} = k\bar{u}qq_0/r^2 \tag{5}$$

From eq. (4), however, $\bar{F} = q_0\bar{E}$, and thus, substituting $q_0\bar{E}$ in place of $\bar{F}$ in eq. (5), we have that the field strength at any point in the field of an isolated charge of q coulombs (Fig. 13) is equal to

$$\bar{E} = k\bar{u}q/r^2 \tag{6}$$

where $k = 9 \times 10^9$.

If more than one charge acts on the test charge q_0 we then have that

$$\bar{E} = \bar{E}_1 + \bar{E}_2 + \bar{E}_3 + \cdots \tag{7}$$

showing that the *total resultant field strength* $\bar{E}$ at a point is the vector sum of the field strengths due to the individual charges, the effect of each charge being considered by itself as if the others were absent.†

---

* See note 4 in Appendix.

† This illustrates the very important "principle of superposition."

It should be mentioned that in pictorial sketches of electric fields, the relative strength of field is indicated by the *density* of the lines of force. Thus, in regions of *high values of field strength* the lines are drawn *closer together*, while in regions of lower strength they are drawn farther apart. In Fig. 13, for example, the closer we get to the charge q, the greater is the field strength, a fact which is shown by the increased density of the lines as we move closer to the charge q. This is also illustrated in Fig. 14, which is a cross-sectional diagram of the field in the 3-dimensional space surrounding two charges equal in magnitude but opposite in sign.

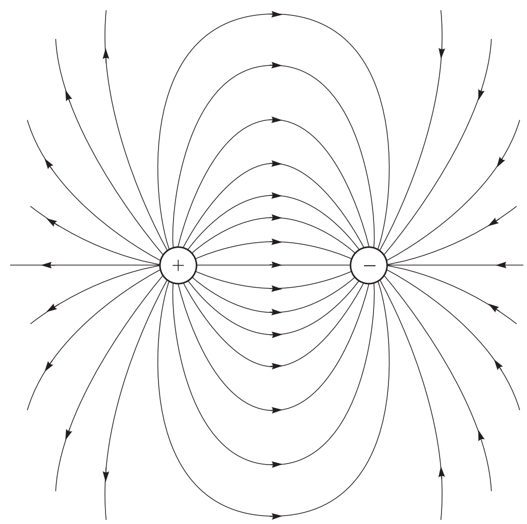


Fig. 14. Sketch showing some of the hypothetical "lines of electric force" in the electric field existing in the neighborhood of two charges $+q$ and $-q$. The lines are all closed, originating on $+q$ and terminating on $-q$; lack of space prevents us showing that all the lines are closed.

Problem 2

On the x, y coordinate plane (letting x and y be distance in meters) it is given that a positive charge of 3 microcoulombs is concentrated at the origin, and a negative charge of -2 microcoulombs is concentrated on the x axis at the point $x = 24$. Find the magnitude and direction of the field strength, relative to the x axis, at the point (15, 6). Air or vacuum dielectric assumed.*

(*Answer*: $228.917\underline{/-11.827^\circ}$ newtons/coulomb)

1.4 Potential Difference; the Volt

In section 1.3 we found that the electrostatic field is a *vector* field, the "field strength" at any point in the field being denoted by $\bar{E}$, where the vector $\bar{E}$ is measured in "newtons per coulomb" (newtons/coulomb). Note that "field strength" is a measurement at any particular POINT in the electric field.

---

* If the reader is not familiar with trigonometry, just postpone doing this problem until pp. 76–85 and eq. (110) in Chap. 5 have been read.

The same electric field can also be expressed in terms of a *scalar* quantity called POTENTIAL DIFFERENCE, which is a measurement involving any TWO POINTS of interest in the field.

The measurement of ''potential difference'' is based upon the fact that energy must be expended, that is, WORK *must be done* in order to move a positive charge q against an electric field. After such movement ceases, and the charge q has been pushed to a new point in the field, we find that all the work done is now *stored in the electric field*, and we might say that the charge q possesses ''potential energy of position.'' The situation is very much like raising, say, a 10-pound iron ball up to a point p against the force of gravity. At the end of the movement, all the work done is *stored in the gravitational field* of the earth–ball system, or, if you wish, in the ball as ''potential energy of position.''

In this regard, it should be noted that the electrostatic field (and also the gravitational field) is a ''friction-free'' system. That is energy (work) can be *stored* in the field, but no losses due to friction occur in moving a charge q through the field. Thus the work done in moving a charge q from a point p_1 to a point p_2 is the same *regardless of the path taken* in going from p_1 to p_2.

With the foregoing in mind, we now define that the *potential difference between two given points* in an electrostatic field is equal to the *work per unit charge* required to move positive charge from the one point to the other point against the field.

Let us denote potential difference by V. In the mks system, work is measured in *joules* and charge is measured in *coulombs*. Hence in the mks system potential difference is the ratio of joules to coulombs which is given the special name ''volts,'' in honor of the early Italian physicist Alessandro Volta.

Thus, if W is the work in joules required to move q coulombs of charge between two points in an electric field, then, by definition, the potential difference between the two points is

$$\text{pot. diff.} = V = \frac{W}{q} = \text{joules per coulomb} = \text{volts} \tag{8}$$

Since work and charge are both scalar quantities* it follows that potential difference, W/q, is also a scalar quantity.

Thus we now have two ways to specify the measurement of an electrostatic field. The first way is in terms of a *vector* quantity $\bar{E}$, the ''field strength'' at any particular point in the field. The second way is in terms of the scalar quantity V, the ''potential difference'' between any two points of interest in the field. It is potential difference that we will deal with most often in our work.

Let us close this section with a few more words about ''field strength,'' the magnitude of which is denoted by E. From section 1.3 we recall that E is basically measured in newtons per coulomb. Also in section 1.3 we mentioned that ''newtons per coulomb'' is the same as ''volts per meter.'' To show that this is true, manipulate the ''units'' like algebraic quantities, as follows, in which we recall that *work* (joules) = *force times distance* (newtons times meters),

$$\begin{aligned} E &= \frac{\text{newtons}}{\text{coulombs}} = \frac{\text{newtons} \times \text{meters}}{\text{coulombs} \times \text{meters}} = \left(\frac{\text{joules}}{\text{coulombs}}\right)\frac{1}{\text{meters}} = \frac{\text{volts}}{\text{meters}} \\ &= \text{volts per meter} \end{aligned}$$

* This is because no sense of *direction* is involved in finding the *sum* of different amounts of work (energy); for instance, 10 joules + 20 joules = 30 joules. Electric charge is likewise a scalar quantity.

Thus, if we wish, field strength can also be expressed in "volts per meter," which is dimensionally (that is, in terms of fundamental units) equal to "newtons per coulomb."

Problem 3
The product "qV" is in what units?

Problem 4
What is the potential difference between two points if 2.65 joules of work is done in moving 0.0078 coulombs of charge between the two points?

Problem 5
Let $\bar{E}_a$ and $\bar{E}_b$ denote the field strength at two different points, a and b, in an electric field. If the values of $\bar{E}_a$ and $\bar{E}_b$ are given, would this information alone be sufficient to allow the calculation of the potential difference V between the two points?

CHAPTER 2

Electric Current. Ohm's Law. Basic Circuit Configurations

2.1 Electric Current

Electric charge *in motion* is called "electric current"; that is, ELECTRIC CURRENT is simply ELECTRIC CHARGE IN MOTION. The concept of electric current is important because it is through the medium of electric current that practical use is made of the phenomenon of electricity.

Let us look again at Figs. 5 and 6 in Chap. 1. In Fig. 6, when the switch is closed the excess charge which flows to the earth through the copper wire constitutes an *electric charge* flowing in the wire. This "electric current" continues to flow until body A becomes electrically neutral, at which time the current ceases.

Electric current is measured in terms of the RATE OF FLOW *of electric charge*; thus, since charge is measured in *coulombs* and time is measured in *seconds*, we have the definition

> *Electric current* is measured in *coulombs per second* which is given the special name *amperes*.
>
> Electric current is represented by the letter i.
>
> $i =$ current in AMPERES = COULOMBS PER SECOND

The ampere, named in honor of the French physicist Ampère, is pronounced "AM peer" in English-speaking countries.

Let us now consider a cross section of a copper wire, or other conductor, through which electric current is flowing, as in Fig. 15. Let the current be flowing from left to right, as suggested by the lines with the arrowheads.

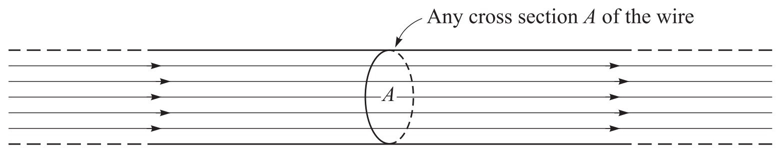


Fig. 15

If the wire in Fig. 15 is carrying a current of *one ampere* it means that electric charge is flowing across the area A at the rate of *one coulomb per second*. That is, one coulomb of charge passes through the area A every second.

Actually, the above statement is basically for the case where a STEADY, constant current of 1 ampere is flowing in the wire. If the current is *not* steady, but *changes with time*, that is, changes from instant to instant, then we must use the ''delta'' notation, thus,

$$\frac{\Delta q}{\Delta t} = i \tag{9}$$

In the above, Δq denotes a small amount of charge that passes through area A in a small interval of time Δt. Thus eq. (9) gives the average current over an interval of time Δt; the smaller Δt is, the smaller is Δq, and the closer eq. (9) comes to being ''instantaneous'' or ''exact'' current at a time t.*

Let us now continue, with a discussion of some details concerning electric charge and current.

Imagine that you are viewing a large cone-shaped pile of sand from a distance of say several hundred feet. Viewed from such a distance, the pile of sand appears to be a *continuous substance*, because you are not close enough to see the *individual grains* or particles that sand is actually made of.

Of course, as you approach the pile, and come right up to it, you begin to see that sand is *not* a continuous substance, but is actually composed of *discrete* (separate) particles or grains.

The same principle of ''graininess'' applies to all matter, be it gaseous, liquid, or solid, except that the ''grains'' are extremely small particles called *atoms* and *molecules*.

For instance, the water we see in a cup is not a ''continuous'' substance, but is composed of a vast number of tiny ''molecules'' of water. A single molecule of water is far too small to be seen under the most powerful microscope, but their existence has been proved by indirect means. We know that a single drop of water is composed of billions upon billions of individual water molecules. Water, as you know, is a compound of hydrogen and oxygen, each molecule of water being composed of two atoms of hydrogen and one atom of oxygen, the atoms being bound together by electrostatic forces.

All atoms and molecules are themselves composed of *electrons*, *protons*, and *neutrons*, as follows.

Electrons are tiny, basic units of *negative electric charge*; ALL *electrons* carry the *same amount of negative charge* which is often denoted by ''e'' where, approximately,

$$e = (1.602)10^{-19} \text{ coulomb of negative charge}$$

Protons are the tiny, basic units of *positive electric charge*; ALL *protons* carry the same amount of positive charge, which has the *same magnitude as that of the electron but of opposite sign*. The proton, however, has considerably *more mass* than the electron, the mass of the proton being about 1845 times that of the electron.

---

* See note 5 in Appendix.

The *neutron* is also one of the basic building blocks of which atoms are composed. Neutrons have the same mass as protons, but are electrically neutral particles.

Since electrons and protons carry the same magnitude of charge, but of opposite sign, it follows that an electrically *neutral* atom or molecule has equal numbers of electrons and protons.

We thus conceive that atoms of all materials are composed of electrons, protons, and neutrons. The relatively massive protons and neutrons are concentrated in the form of a ''nucleus'' in the center of the atom, while the electrons revolve or vibrate in different orbits or ''energy levels'' around the nucleus.

In the atoms of some substances, the electrons in the outer orbits, farther from the nucleus, are only loosely bound to the nucleus, and such atoms can readily gain or lose electrons. If a normally neutral atom or molecule has gained or lost electrons, it is said to be an *ion* (''eye on''), being a ''positive ion'' if it has *lost* electrons and a ''negative ion'' if it has *gained* electrons.

It is not, however, our intention or need to go into details of atomic structure here. All we wish to do, right now, is to point out that what we call ''electric current'' can be a flow of *electrons*, *ions*, or a *combination* of electrons and ions, depending on the substance we're dealing with.

In the case of *metals*, the electric current is largely a flow of ''free electrons'' that have become detached from the atoms of the metal. Thus, *good conductors*, such as silver and copper, are materials in which the electrons are easily detached from the atoms of the substance.

On the other hand, a *poor conductor* (good insulator) is a substance, such as rubber or porcelain, in which the electrons are tightly bound to the atoms and molecules and hence are not available for current flow.

In the cases of liquids and gases, the current flow is mainly by means of *ions*, which can be either positive or negative, or a combination of ions and electrons.

The foregoing naturally brings up the question of the *direction* in which electric current ''actually'' flows in a conductor. To answer that question we begin with a discussion of how we can *detect* the passage of electric current through a conductor.

First, as you would expect, it requires an *expenditure of energy*, that is, *work* has to be done to force the passage of electric charge through a conductor. This energy must, of course, come from some kind of source capable of doing work.

You might ask, ''What happens to the work that is supplied to force electric charge to flow in a conductor?'' The answer is that it may be transformed into *mechanical* energy (by means of a motor), or into *radiant* energy (as from a light bulb), or into *chemical* energy (in the formation of a battery), and so on, but at least a portion of the work will always be transformed into *heat* energy in the conductor, and this will of course cause the temperature of the conductor to rise. The point to be made here is simply that *one way* of detecting the passage of electric current through a conductor is to sense any *rise in temperature* of the conductor.

In addition to the temperature effect, we also find that electric current always establishes a *magnetic field* around any conductor through which it is flowing. This is a fact of very great importance, and one we will investigate in detail later on. Right now, however, we merely wish to point out that *another way* of detecting the passage of electric current through a conductor is to detect the presence of a *magnetic field* around the conductor.

Now consider the following. Imagine we have a long piece of rubber tubing, and we are told that the tube contains a conductor of electricity throughout its length. However, we cannot see inside the tube, and so we do not know whether it contains a solid *metallic* conductor (such as a copper wire), or some kind of a *liquid* or paste conductor (a dilute solution of any kind of acid, for instance).

If the tube contains a metal conductor, like a copper wire, the current will consist of a flow of *electrons*; but if the tube contains, instead, a liquid conductor of some type the current may consist of a flow of *positive ions*.

Now suppose we detect that the tube is getting *hot*, and also that a compass needle shows the presence of a magnetic field around the tube. These *external* effects tell us that an *electric current* is flowing in the conductor inside the tube. But we *cannot*, from these two external tests, tell whether the current is a flow of *electrons* or a flow of *positive ions*.

Since, in a given situation, electrons and positive ions flow in opposite directions, it follows that these two external tests (temperature and compass needle) will not tell us what the actual direction of current flow is. This means that, in the mathematical analysis of electric networks, it will *not be necessary* to take into account whether a given current is a flow of positive charge or negative charge, because the useful external effects are the same in either case. Hence, for the purposes of analysis, we can just as well assume that *all currents* consist of a flow of *positive charge*; therefore, for the sake of simplicity it will hereafter be assumed in this book that *all currents consist of a flow of* POSITIVE *charge*.*

Problem 6
1 coulomb of negative charge contains how many electrons?

Problem 7
If 9 million electrons cross area A in Fig. 15 in 0.1 microsecond, what is the average current in amperes during that time?

2.2 Electromotive Force

We must keep in mind that *energy has to be expended*, that is, WORK *must be done*, to force the carriers of electric charge to move in a conductor. With this in mind, we begin this section with a simple but very basic discussion, starting with Fig. 16.

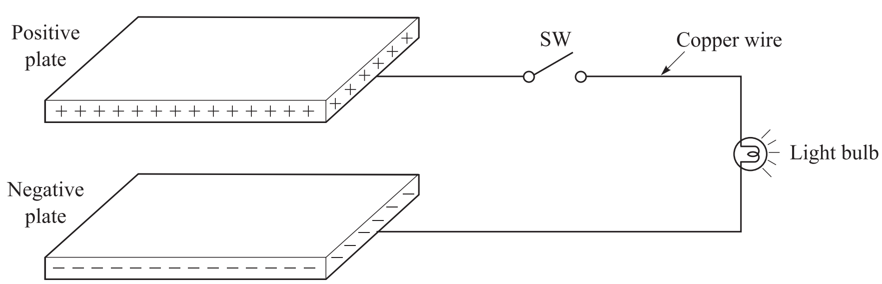


Fig. 16

In Fig. 16 we have two metal plates, the "top" plate being positively charged and the "bottom" plate negatively charged, as shown. A small positive test charge, if placed between the two plates, would experience a downward force, showing that an electric field exists in the region between the two plates.

We must remember that *work* had to be done to separate the positive and negative charges, and that the work, so done, is now stored as potential energy in the electric field

* Sometimes referred to as the flow of "conventional current."

between the plates. A certain potential difference, expressed in terms of V volts, exists between the two plates.

Let us now close the switch in Fig. 16. When this is done, there will be a brief, momentary flow of charge through the light bulb, the flow continuing until both plates are electrically neutral—that is, until the potential difference between them is reduced to zero.

In this action the light bulb will emit a brief flash of light, showing that the energy stored in the electric field is being converted into heat energy and radiant energy in the form of visible light.

Let us now suppose, in Fig. 16, that we are not satisfied with just a brief flash of light, but wish the light to burn *continuously* and uniformly.

To do this, we must maintain a *constant rate of flow of charge* of "q" coulombs per second through the bulb; that is, we must maintain a *constant current* of "i" amperes in the bulb. This, however, can be done *only if we* CONTINUOUSLY SUPPLY THE WORK REQUIRED TO MOVE THE POSITIVE CHARGES FROM THE NEGATIVE PLATE TO THE POSITIVE PLATE, against the *internal field* that exists between the positive and negative plates. This is illustrated in Fig. 17, in which just a "side view" of the positive and negative plates is shown.

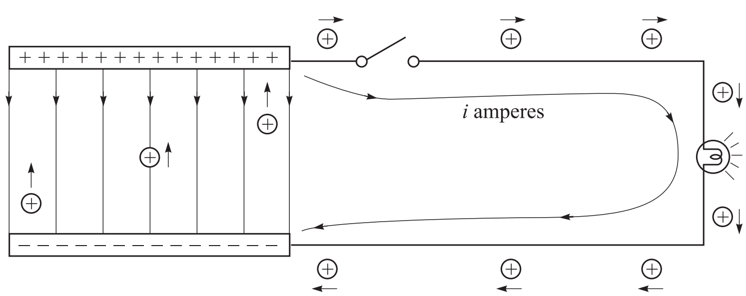


Fig. 17

In Fig. 17, the symbol $\oplus$ represents a few of the vast number of basic positive charges that, circulating around the circuit, constitute the current of i amperes. In the figure the "circuit" consists of the positive and negative metal plates, the electric field between them, the switch, the light bulb, and the connecting wires.

The situation is similar to that in which a mechanical pump forces water to flow through pipes connected to some kind of "water motor," which is a device having a rotor capable of converting the kinetic energy of the moving water into mechanical energy; such a "water circuit" is shown in Fig. 18, in which the water is being pumped around in the clockwise sense.

In the "water circuit" of Fig. 18, it should be understood that the pump, pipes, and motor are completely full of water; that is, the water is "continuous" at all points in the circuit. It then follows that "p" gallons of water *flows through every cross-section of the circuit every second*. Thus, if $p = 2$, this means that, all around the circuit, water is simultaneously flowing across *all cross-sections* (such as at a, b, and c in the figure, for example) at the rate of 2 gallons per second. This simply means that the *rate of flow of water* is the *same* all around the circuit.

Let us now return to Fig. 17. It should first be noted that (like the flow of water in Fig. 18) the *same amount of charge*, q coulombs/second, flows through *all cross-sections* in the circuit. Hence, like the rate of flow of water in Fig. 18, it follows that at any given instant the *current* "i" is the *same* all around the circuit.

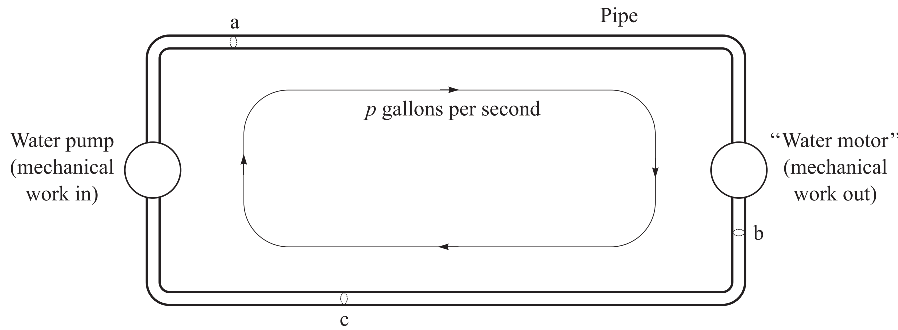


Fig. 18

Next, in Fig. 17, let us consider the charges $\oplus$ as they move from the negative plate "upward" toward the positive plate. Notice that the charges, while they are between the two plates, experience a force of *repulsion* due to the positive plate and a force of *attraction* due to the negative plate. Thus the charges between the plates experience a *force of repulsion/attraction*, due to the field between the plates, that tends to force them "downward" toward the negative plate. In order to overcome this force, and move the charges "upward" against the field toward the positive plate, *energy must be expended*, that is, WORK *must be done on the charges*. *Only* if this is done will it be possible to maintain the useful current i shown in Fig. 17, flowing from the positive plate, through the light bulb, to the negative plate.

Any device capable of exerting force on electric charges, and thus being able to do work on such charges and move them against an electric field, is called an "electric pump" or *electric generator*. An electric generator exerts what is called *electromotive force* (abbreviated "emf") on the charges it is "pushing through it," and is thus said to be a "seat" or "source" of "electromotive force" (emf).

It should be understood, of course, that a source of emf (an electric generator) does not "create" charge; it simply supplies the *energy* necessary to move the charges through it. A source of emf is thus like a water pump; the pump does not create water, but simply imparts kinetic energy to the water it is forcing through it.

There are two principal, practical types of electric generator. The first type is the *battery*, which depends for its operation upon the conversion of *chemical energy* into electrical energy. The second type of generator depends upon the phenomenon of "electromagnetic induction," in which *mechanical energy* is converted into electrical energy. We will make a detailed study of electromagnetic induction later on, but for the time being we'll assume our sources of emf to be batteries. This will have no effect on basic circuit theory, because that is independent of the manner in which the emf's are generated.

As already mentioned, a "battery" is a source of electromotive force in which chemical energy can be converted into electrical energy.

All batteries consist of individual "cells," in which each cell consists of a "positive electrode" and a "negative electrode," the two electrodes (also called "poles") being separated by a chemical compound called the "electrolyte," which can be in the form of a liquid or a paste.

For instance, the common "dry cell" consists of carbon and zinc electrodes, or poles, separated by a paste-type of electrolyte made of sawdust saturated with a solution of ammonium chloride. The carbon electrode is the "positive pole" and the zinc electrode is the "negative pole." The potential difference between the two poles is approximately 1.5 volts for a cell in good condition, when delivering current to an external load.

Basically, in a battery, the electric charges in the atoms and molecules have potential energy of position, due to certain electrostatic binding forces present. When the battery delivers current to an external load, chemical reactions occur in the battery in which the atoms and molecules are rearranged, and in the rearrangement the potential energy of the charges is reduced, being transformed into the kinetic energy required to move the charges against the internal field of the battery.

There are, as you may know, a number of different types of cell, each type having certain advantages and disadvantages. All cells, however, produce a relatively low value of potential difference between their electrodes, ranging in value from 1.2 to 2.2 volts, approximately. In order to obtain higher potential differences, which are often required in practical applications, it is generally necessary to connect two or more cells together to form a *battery* of cells, as illustrated in Figs. 19 and 20.

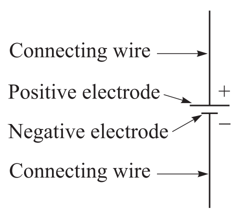


Fig. 19

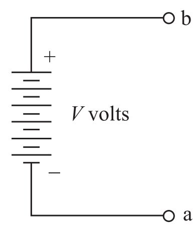


Fig. 20

Figure 19 is the symbol used, when drawing "schematic" circuit diagrams, to indicate the presence of a single cell. Notice that the longer horizontal line represents the *positive* electrode and the shorter horizontal line represents the *negative* electrode.

As mentioned above, in order to provide higher potential differences it is necessary to "stack" or "series-connect" a number of cells to form a "battery." This is indicated schematically in Fig. 20, where "a" is the *negative terminal* of the battery and "b" is the *positive terminal*; together, a and b are the positive and negative *output terminals* of the battery. Note that, going from a to b through the battery, the positive electrode of each cell connects to the negative electrode of the next cell. If "n" such cells are thus series-connected to form a battery, the potential difference between the battery output terminals will be *n times the potential difference of a single cell*. In Fig. 20, V is the potential difference between the output terminals a and b. Since potential difference is measured in *volts*, it is customary to call V the "battery voltage."

2.3 Electrical Resistance. Ohm's Law. Power

Let us now return to Fig. 17 and replace the two charged plates with a battery of V volts. Since the battery is a source of constant emf, it will be able to maintain a constant current of I amperes* flowing in the circuit, as in Fig. 21. (For reasons discussed in section 2.1, we will always assume current to consist of a flow of *positive* charges which flow *out* of the *positive terminal* of the battery, for the same reason that they flow out of the positive plate in Fig. 17.)

* Both "i" and "I" are used to represent electric current. i generally designates current that changes from instant to instant, while I designates a constant value of current. Thus in Figs. 21 and 22 the current would have a constant value of I amperes for given, fixed values of V and R.

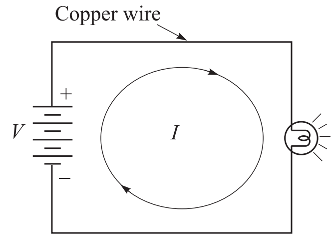


Fig. 21

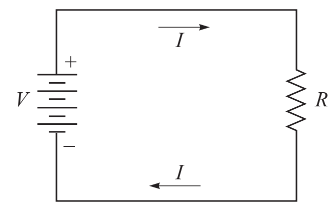


Fig. 22

In Fig. 21, the tungsten filament of the light bulb offers a considerable amount of *opposition*, or what is called ELECTRICAL RESISTANCE, to the passage of electric current through it. Because of the high resistance of the filament, the battery voltage V must be relatively high in order to produce the amount of current I required to heat the filament to incandescence. On the other hand, the copper wires used to connect the battery to the bulb have *very little resistance*; as a matter of fact, in almost all cases we can disregard the very small resistance of the wires used to connect the various parts of the circuit together, and assume that, for practical purposes, the connecting wires have *zero* resistance to the passage of current through them. We will always assume this to be the case, unless otherwise stated.

The amount of electrical resistance is denoted by R, and in electrical diagrams the presence of resistance is represented by the symbol –\/\/\/\/–. Using this symbol, we have redrawn Fig. 21 as Fig. 22, in which R denotes the "electrical resistance" of the tungsten filament in the light bulb.

We have already learned that substances that offer *little resistance* to the passage of current are called "conductors," while those that offer *great resistance* are called "insulators."

There are, of course, many grades of conductors (and insulators). Take, for example, two metals such as copper and tungsten. Both are classified as "conductors," but a copper wire is a *better conductor* than a tungsten wire of the same length and diameter; that is, the copper wire offers *less resistance* to the flow of current than does the tungsten wire.

Of course, a number of things determine which materials will be used as a conductor in a given case. In the design of an electric toaster, for instance, the heating element might consist of wire made of "Nichrome," which is a metal alloy having about 60 times the resistance of the same amount of copper wire. On the other hand, the "line cord" that connects the toaster to the wall plug will make use of low-resistance copper wire. (We should also remember that it will be necessary to make use of different *insulating materials*, such as mica, plastic, and rubber, in the construction of the toaster.)

The first comprehensive investigation into the nature and measurement of electrical resistance was made by the German physicist Ohm (as in "home") around the year 1826. After a lengthy series of experiments Ohm was able to report that

> The *current* in a conductor is *directly proportional to the potential difference between the terminals of the conductor*, and *inversely proportional to the resistance of the conductor*.

The above constitutes what is called OHM'S LAW. If we let

V = potential difference (emf) applied to the conductor,

I = current in the conductor,

R = resistance of the conductor,

then the algebraic form of Ohm's law becomes

$$I = \frac{kV}{R} \tag{10}$$

in which "k" is a constant of proportionality. Equation (10) says that *current* I is *directly proportional to potential difference* V and *inversely proportional to resistance* R. That is, the greater V the greater is I, and the greater R the less is I.

The *unit* of resistance is called the *ohm*; we define that a conductor has *1 ohm* of resistance if *1 ampere* of current flows when a potential difference of *1 volt* is applied to the conductor. Thus, if $V = 1$ and $R = 1$, then by definition $I = 1$, and eq. (10) becomes, $1 = k(1)/(1)$, which can be true only if $k = 1$. Therefore, if we express V in volts, I in amperes, and R in ohms, then eq. (10) becomes the basic OHM'S LAW

$$\boxed{I = \frac{V}{R}} \tag{11}$$

which states that *amperes is equal to volts divided by ohms*. It follows that Ohm's law can also be written in either of the forms

$$R = V/I \tag{12}$$

and

$$V = RI \tag{13}$$

Equation (12) says that *ohms* is equal to *volts divided by amperes*, while eq. (13) makes the equivalent statement that *volts* equals *ohms times amperes*. Equations (11), (12), and (13) are basic to electrical and electronic engineering, and should be committed to memory.

Now let's consider the POWER developed by the battery in Fig. 22. To do this, let us begin with a brief review of some basic concepts, as follows.

First, we have the idea of "energy," which is measured in terms of capacity to do *work*, which is measured in *joules*. In mechanics, when a force of F newtons acts through a distance of L meters, the agency supplying the force does an amount of work, W, equal to FL joules, that is, $W = FL$.

It is important, now, to notice that there is no *time* requirement in the definition $W = FL$. Thus, suppose in a certain case that "FL joules of work" must be done. Such a simple requirement is satisfied regardless of whether the work is done in 1 minute or in 10 minutes.

Actually, however, in plain language we know that a "more powerful" source of energy is required to do the work in 1 minute than in 10 minutes. For example, a small boy might, with the aid of a system of pulleys, raise a 100-pound weight 1 foot off the floor in, say, 60 seconds. An adult, however, might, without having to use pulleys, be able to do the same thing in, say, 6 seconds. The *same amount of work* (100 foot-pounds) is done in both cases, but the adult, while working, is delivering energy to the system 10 times as fast as the boy is capable of doing.

Thus, the *time rate* of doing work is important in practical engineering. In the mks system "time rate of doing work" is expressed in *joules per second*, which is given the special name *watts*, in honor of the early engineer James Watt. Thus we have the definition

$$\text{time rate of doing work} = \text{joules per second} = \text{watts} \tag{14}$$

Now let "P" denote the *power* developed by the battery in Fig. 22. We wish to show that the power, P *watts*, is equal to *the battery voltage times the current* I; that is, we wish to show that $P = VI$.

To do this, we make use of the basic definitions, volts = joules per coulomb and current = coulombs per second, and then manipulate the units as if they were ordinary algebraic quantities, thus

$$VI = \frac{\text{joules}}{\text{coulombs}}\,\frac{\text{coulombs}}{\text{seconds}} = \frac{\text{joules}}{\text{seconds}} = \text{joules per second} = \textit{watts}, \text{ thus,}$$

$$\boxed{P = VI} \tag{15}$$

which says that the *power in watts* produced by a battery of V *volts* when delivering a current of I *amperes* is equal to VI; that is, *power in watts* is equal to *volts times amperes*.

In Fig. 22, the power VI, produced by the battery, is delivered, by means of connecting copper wires, to the "load resistance" R. If R is, for example, the filament of a light bulb, then the power supplied to R will be converted into heat energy and into radiant energy in the form of visible light.

Or, if R is an electric motor, the majority of the battery output will, hopefully, be converted into useful mechanical energy, with the relatively small balance being lost in the form of heat energy.

The power output of the battery, given by eq. (15), is of course the same as the power delivered to and "consumed by" the load resistance R. We can therefore use eqs. (11) and (13) to write equations for the *power delivered to a resistance of* R *ohms*, as follows.

First, using eq. (11), eq. (15) becomes $P = V(V/R)$, so that

$$P = \frac{V^2}{R} \tag{16}$$

Or, using eq. (13), eq. (15) becomes $P = (RI)I$, so that

$$P = I^2R \tag{17}$$

Equations (11) through (17) should all be committed to memory, because they are of such fundamental importance. For our convenience, they are summarized below, where V is potential difference in *volts*, I is current in *amperes*, R is resistance in *ohms*, and power is in *watts*.

OHM'S LAW:	$V = RI$	$I = V/R$	$R = V/I$
POWER:	$P = VI$	$P = I^2R$	$P = V^2/R$

Problem 8
In Fig. 22, if $V = 48$ volts and $R = 6$ ohms, what current will flow?

Problem 9
In Problem 8, find the power output of the battery using eqs. (15), (16), and (17).

Problem 10
If the power input to a 75-ohm resistance is known to be 18 watts, what current is flowing?

2.4 Some Notes on Temperature Effects

In devices such as electric heaters, irons, and toasters, the basic purpose is simply to develop a required amount of *heat* in the resistance wire used in such devices.

In most applications, however, especially in electronics, resistance is *not* used in a circuit to develop heat, but is used for other purposes. The heat developed in resistance is, therefore, in most applications an *undesired* effect. The principal reasons why this is true are as follows.

1. The *resistance* of a given length of a given type of wire depends, to some extent, upon the *temperature* of the wire. Thus, as the temperature of a wire increases, due to increased heat input, its resistance also tends to increase, and this is generally an undesirable effect.
2. Excessive heat generation adversely affects the operation of other components in a circuit, and tends to cause physical deterioration of the resistor* itself.

Let us discuss items (1) and (2) in more detail. To begin, it should be pointed out that the four principal factors that determine the resistance of a wire conductor are

(a) the length L of the wire,
(b) the cross-sectional area A of the wire,
(c) the material of which the wire is made,
(d) the temperature T of the wire.

Let us deal with the first three items first. Experiment proves that the resistance R of a wire conductor is *directly proportional to the length* L and *inversely proportional to the cross-sectional area* A, a fact we show mathematically by writing

$$R = \rho \frac{L}{A} \tag{18}$$

where R is the resistance in ohms, L is the length in meters, A is the cross-sectional area in square meters, and where the proportional constant ρ (the Greek letter "rho") is called the *resistivity* ("ree sis TIV ity"), whose value depends upon the *material* the wire is made of and the *temperature* T of the wire. Note that, from eq. (18), we have

$$\rho = \frac{RA}{L} = \frac{(\text{ohms})(\text{meters})^2}{(\text{meters})} = (\text{ohms})(\text{meters})$$

thus showing that resistivity ρ has the dimensions "ohms times meters," or "ohm · m," as it is usually written.

As mentioned above, the value of ρ depends on the kind of metal the wire is made of, and the temperature T of the wire. It is found that the resistivity of metals increases linearly with temperature over a wide range of temperature, and this fact is expressed in the form

$$\rho = \rho_0[1 + \alpha_0(T - T_0)] \tag{19}$$

* A device made solely to introduce resistance into a circuit is called a *resistor* ("ree SIS tor"). Thus, a "100 ohm, wire-wound resistor" is a device constructed of wire having 100 ohms of resistance.

where

$\rho =$ resistivity of the given metal at any temperature T°C,

$\rho_0 =$ resistivity of the given metal at the standard reference temperature of $T_0 = 20$°C,

$\alpha_0 =$ the "temperature coefficient of resistance" of the metal at 20°C.

The values of ρ_0 and α_0 ("alpha sub zero") have been found experimentally, a short table of values being given below.

	ρ_0 (ohm · m)	α_0 (per °C)
Silver	$(1.59)10^{-8}$	$(3.75)10^{-3}$
Copper	$(1.75)10^{-8}$	$(3.80)10^{-3}$
Aluminum	$(2.83)10^{-8}$	$(4.03)10^{-3}$
Tungsten	$(5.50)10^{-8}$	$(4.70)10^{-3}$
Constantan	$(49.0)10^{-8}$	$(0.01)10^{-3}$

Note 1: "Constantan" is an alloy of 45% nickel and 55% copper having, as the table shows, a high value of resistivity and a very low value of temperature coefficient.
Note 2: In some wire tables a unit of length called the "mil" is used, where 1 mil = 0.001 inch. A "circular mil" is defined as the area of a circle 1 mil in diameter.

Let us next consider the *power rating* of a resistor, using, as a convenient example, a 100-ohm resistor.

Suppose, for example, that we are dealing with an application in which the resistor must carry a current of, say, 0.8 amperes. Then, by eq. (17), the power input to the resistor will be $P = I^2R = 64$ *watts*, which is 64 joules of work per second. Since 1 calorie = 4.186 joules,* we have $64/4.186 =$ *15.289 calories of heat will be developed in the resistor each second*. We thus have a problem in *heat transfer*, because if the heat generated in the resistor is not transferred away fast enough *the temperature of the resistor will continue to rise until it is destroyed.*

The ability of a resistor to dissipate heat depends greatly upon the amount of exposed surface area the resistor has. Thus, resistors that must dissipate relatively large amounts of heat must be made physically larger than resistors that must dissipate only a relatively small amount of heat. The amount of heat that a given resistor can safely dissipate also depends, of course, on the temperature of the surrounding (ambient) air, and whether the flow of air is by natural convection or is driven by a fan or blower.

Resistors can be purchased in values of resistance from less than 1 ohm to several megohms ("1 megohm" being 1 million ohms), and in power rating from $\frac{1}{4}$ watt to several hundred watts.

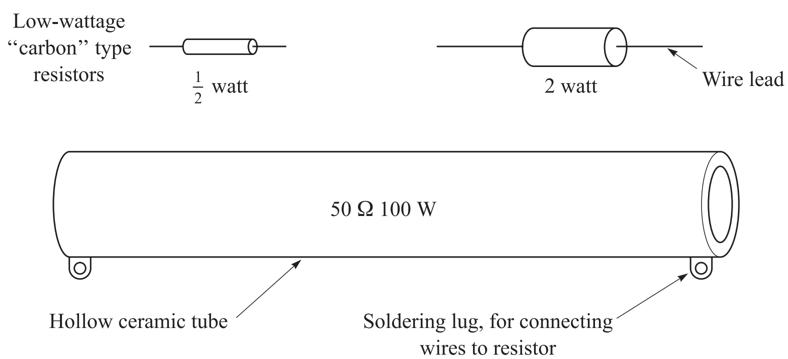
When specifying the "power rating" of a resistor, the manufacturer will also state the maximum temperature of ambient air for which the rating is valid. For example, a manufacturer might state that the power rating of a certain resistor is "5 watts at 30°C ambient," and an equipment designer must keep this in mind.

Resistors that must dissipate more than 2 or 3 watts are generally of the wire-wound type, consisting of resistance wire, of low temperature coefficient, wound on a ceramic tube.

---

* See note 2 in Appendix.

Lower wattage resistors are most usually of the "carbon" type, in which the resistance is formed from a compressed mixture of carbon with a suitable binder, and then encapsulated in a plastic case. In order to better visualize their construction and appearance, several resistors are sketched below in approximately their actual physical size.



Above is an example of appearance of a 50-ohm, 100-watt, wire-wound resistor, where the symbol "Ω" is the capital "omega" and is read "ohms." In the above type, the resistance wire is wound on the ceramic tube, after which the unit is enameled and baked.

Problem 11
If a 25-ohm resistor is carrying a current of 1.86 amperes, how many calories of heat must it be able to dissipate every second?

Problem 12
Noting that $°C = (5/9)(°F - 32)$, calculate the resistance of 450 feet of round aluminum wire of 1/2 inch diameter at 86°F. (*Answer*: 0.03188 ohms)

Problem 13
A certain length of copper wire is found to have 2.625 ohms of resistance at 30°C. What will be its resistance at 40°C? (*Answer*: 2.721 ohms)

Problem 14
The heating element of a heater is to be made to have 35 ohms of resistance at an operating temperature of 565°C. If Constantan resistance wire of diameter 1 millimeter (one-thousandth of a meter) is available, how many feet should be cut from a spool of wire at 20°C? Assume L and A are independent of temperature. (*Answer*: 183.056 feet)

2.5 The Series Circuit

A SERIES CIRCUIT or "series-connected circuit" is a circuit having JUST ONE CURRENT PATH. Thus, Fig. 23 is an example of a "series circuit" in which a battery of constant potential difference V volts, and three resistances, are all connected "in series."

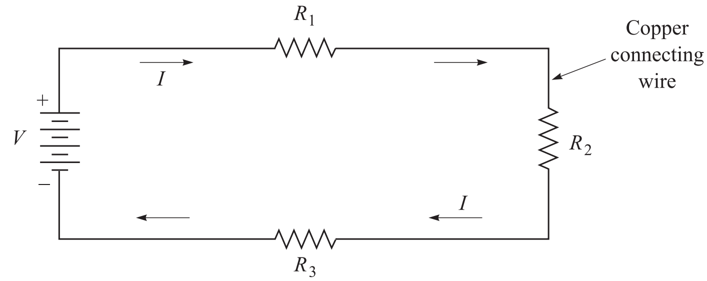


Fig. 23

Since a series circuit has just one current path, it follows *that all the components in a series circuit* CARRY THE SAME CURRENT I, a fact evident from inspection of Fig. 23.

As explained in section 2.1, the current I is assumed to be a flow of *positive charge*, and thus flows *out of the positive terminal of the battery* and around through the external circuit, reentering the battery at the negative terminal. This is indicated by the arrows in Fig. 23.

In a series circuit, the TOTAL *resistance*, R_T, that the battery sees is equal to *the* SUM *of the individual resistances*. Thus, in the particular case of Fig. 23 the battery sees a total resistance, $R_T = R_1 + R_2 + R_3$, while in the general case of "n" *resistances connected in series* the battery sees a total resistance of

$$R_T = R_1 + R_2 + R_3 + \cdots + R_n \tag{20}$$

By Ohm's law, eq. (11), it follows that the *current* I in a series circuit is equal to

$$I = \frac{V}{R_T} = \frac{V}{R_1 + R_2 + \cdots + R_n} \tag{21}$$

In the above, we're assuming the resistances of the copper connecting wires to be negligibly small in comparison with R_T, and this will be normally true in practical circuit work. (If such is not the case, then, in Fig. 23 for example, a fourth resistance would be added in series in the diagram, equal to the resistance of the connecting wires; but this will seldom be necessary.)

We have seen that a battery is a device capable of moving electric charge *against* the internal electric field that exists between its positive and negative terminals. As explained in section 2.2, a battery can do this because it is able to convert chemical energy into electrical energy. A battery is thus referred to as a *generator*, and is classified as an ACTIVE *device*, because it is a *source* of electrical energy.

Resistance, on the other hand, *consumes* electrical energy, removing it from the circuit in the form of heat. Since resistance does not produce or generate electrical energy, it is a non-active or PASSIVE type of circuit element.

A resistor, being a passive device, has no internal electric field until it is connected to a battery. When this is done, an internal electric field appears between the terminals of the resistor, a *potential difference* exists between the terminals, and current begins to flow.

The potential difference between the terminals of a resistor is called the VOLTAGE DROP *across the resistor*, and, by eq. (13), is equal to *the current I times the resistance R*; that is, the "voltage drop" across a resistance of *R ohms* carrying a current of *I amperes* is

IR volts. Note that, from eq. (21), we have the three relationships

$$V = IR_T \tag{22}$$

$$V = I(R_1 + R_2 + \cdots + R_n) \tag{23}$$

$$V = IR_1 + IR_2 + \cdots + IR_n \tag{24}$$

where V is the battery voltage, or "applied voltage" as it is generally called. From inspection of eq. (24) we have the important fact that

> In a series circuit, the *applied voltage* is equal to the *sum of the voltage drops*.

It should be pointed out that *the voltage drop across a resistor* is always *from plus to minus in the direction of the current flow*, a fact illustrated in Fig. 24.

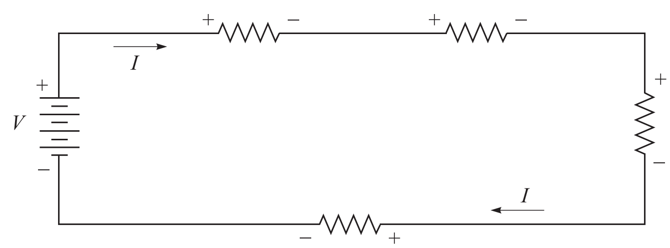


Fig. 24

It should be noted, in Fig. 24, that the battery voltage is *from minus to plus in the direction of the current*; thus *the battery voltage is exactly opposite to the sum of the voltage drops across the resistors*, which corresponds, in the electric circuit, to Newton's third law in mechanics, that is, an *applied force* is always balanced by an equal and opposite force.

Let us next consider the *power* relations in a series circuit. First, by eqs. (15), (16), and (17), a battery of *V volts* delivers a *total power output of P watts* given by any of the relationships

$$P = VI = V^2/R_T = I^2 R_T \tag{25}$$

where R_T and I are given by eqs. (20) and (21). Since power is a scalar quantity, it follows that the *total power P* is equal to *the sum of the powers* developed in the individual resistors, that is,

$$P = P_1 + P_2 + P_3 + \cdots + P_n \tag{26}$$

where the individual powers are found by applying eqs. (15), (16), and (17) to each individual resistor, using the voltage drop associated with each resistor.

To be more specific, let R_x be the value of any one of a number of resistors in series, such as in Figs. 23 and 24. If V_x is the *voltage drop across* R_x, then P_x, the *power input to* R_x, is

$$P_x = IV_x \qquad \text{(by eq. (15))}$$

$$P_x = V_x^2/R_x \qquad \text{(by eq. (16))}$$

$$P_x = I^2 R_x \qquad \text{(by eq. (17))}$$

Since the *voltage drop across a resistor* is the *current times the resistance* (by eq. (13)), we have that the voltage drop across a series resistor R_x is equal to

$$V_x = IR_x$$

and since, from eq. (21)

$$I = V/R_{\mathrm{T}}$$

we have that

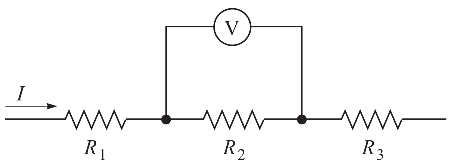
$$V_x = V(R_x/R_{\mathrm{T}}) \tag{27}$$

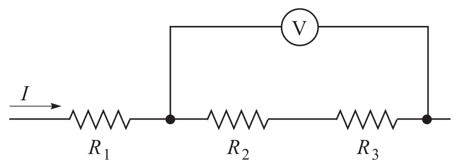
which is the value of V_x to be used in any of the above equations for P_x. R_{T} is the sum of all the series resistances, including R_x, and V is the applied battery voltage.

Let us now conclude this section with a discussion of several topics of importance in all circuit work, beginning with the *voltmeter* and the *ammeter*.

As the names imply, a "voltmeter" is an instrument for measuring *voltage* and an "ammeter" is an instrument for measuring "amperes," that is, *current*. It is not our purpose, at this time, to explain the inner workings of these devices, but only to describe how they are connected in a circuit to measure voltage or current.

A voltmeter is used to measure the voltage (potential difference) between any two points in a circuit, such as in the two figures below, where Ⓥ is the voltmeter.





In the left-hand figure the voltmeter is connected to read the voltage drop across R_2 only, while in the right-hand figure it reads the sum of the voltage drops across R_2 and R_3.

It should be pointed out that a voltmeter is constructed to have a *very high internal resistance*, so that it will have negligible effect on any circuit it is connected to.

On the other hand, an "ammeter," since it measures *current*, must be connected *in series* in the circuit, as in the figure below, where Ⓘ is the ammeter.

Since an ammeter is connected directly in the current path, an ammeter must be constructed to have a *very* LOW *internal resistance*, so that it will offer negligible resistance to the current flowing through it, and thus not cause any change in the current it is put in to measure.

Another point to be mentioned is that all practical sources of emf, including batteries, have *internal resistance* to a greater or less degree. For a battery, the internal resistance can be denoted by R_{b} and is included in the symbol for the battery, as shown in Fig. 25, where "+" and "−" are the *external* positive and negative terminals of the battery.

Internal resistance is *undesirable* in a battery or other type of generator for several reasons.

First of all, when the battery delivers current there is an *internal power loss* in the battery, equal to I^2R_{b}, which not only lowers the efficiency of the battery but may cause it to overheat and thus shorten its life.

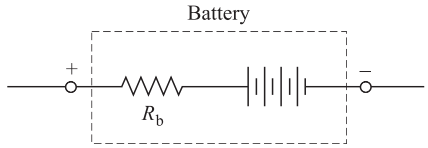


Fig. 25

Secondly, if the battery is delivering a current I, the presence of R_b causes an *internal voltage drop* in the battery, equal to IR_b, which *subtracts from the useful voltage output of the battery*, as explained in connection with Fig. 26, where x and y are the *external terminals* of the battery.

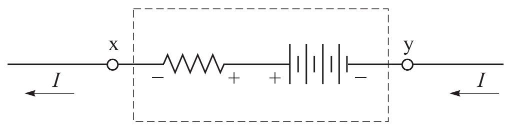


Fig. 26

Notice that the voltage drop across the internal resistance *opposes* or ''bucks'' the internal voltage generated by the battery, thereby *reducing the useful voltage available at the battery terminals*. This effect is proportional to the current I, so that the *more current* we attempt to make the battery produce, the *lower* is the output voltage available at the battery terminals. This is an undesirable effect because, ideally, we would like the battery voltage to remain *constant*, and not be affected by changes in current.

It follows, therefore, that a battery or other type of generator that must produce *large values of current* must be constructed to have *very low internal resistance*. If this is not done, the output voltage of the battery will vary widely with changes in output current, and the battery will be said to have ''poor voltage regulation.''

In our work here, we'll assume the batteries to have negligibly small values of internal resistance. In cases where this cannot be assumed, the internal resistance of the battery will be added in series with the battery, and its effect included in calculations of current.

The last point we wish to make is that circuits in which the *direction* of current flow does not change are called DIRECT-CURRENT *or ''dc'' circuits*. Since the direction of current flow through a battery does not change (unless it is being ''recharged''), a battery is an example of a dc generator, and thus the battery circuits in this and the next few sections are examples of dc circuits. Later in our work we'll take up the important case where the polarity of the generator periodically *reverses*, providing what is called *alternating* or ''ac'' circuits.

Problem 15

In Fig. 27, a battery of constant 48 volts is applied to five series-connected resistors, as shown, the resistance values being in ohms.

(a) What is the reading of meter M1?

(b) What is the reading of meter M2?

(c) Power output of the battery is __________ watts?

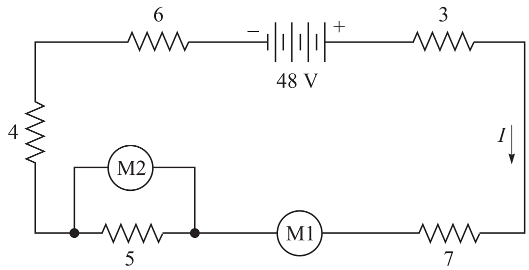


Fig. 27

Problem 16
In Fig. 27, calculate the power input to each resistor, then verify that eq. (26) gives the same answer as found in part (c) above.

2.6 The Parallel Circuit

A PARALLEL *circuit* is one in which the battery current *divides* into a number of "parallel paths." This is shown in Fig. 28, in which a battery, of constant emf V volts, delivers a current of I amperes to a load consisting of any number of n resistances connected "in parallel."

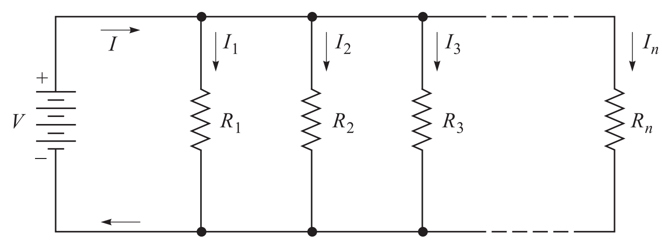


Fig. 28

The currents in the individual resistances are called the "branch currents," and the battery current I is often called the "line current." From inspection of Fig. 28 we see that, *in a parallel circuit*, the *battery current I is equal to the sum of the branch currents*, that is, in Fig. 28,

$$I = I_1 + I_2 + I_3 + \cdots + I_n \tag{28}$$

Next, from Fig. 28 we see that *the battery voltage V is applied equally to all n resistances*; that is, the *same voltage V* is applied to all the parallel branches. Hence, by Ohm's law (eq. (11)), the *individual branch currents* in Fig. 28 have the values

$$I_1 = V/R_1, \qquad I_2 = V/R_2, \ldots, I_n = V/R_n \tag{29}$$

Upon substituting these values into the right-hand side of eq. (28) we have

$$I = V\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots + \frac{1}{R_n}\right) \tag{30}$$

Now let R_T be the *total resistance as seen by the battery* in Fig. 28. Then, by Ohm's law, it has to be true that

$$I = \frac{V}{R_T} \tag{31}$$

Since the left-hand sides of the last two equations are equal, the two right-hand sides are also equal. Setting the two right-hand sides equal, then canceling the Vs, gives

$$\boxed{\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots + \frac{1}{R_n}} \tag{32}$$

where R_T is the *total effective resistance seen by the battery*. In words, eq. (32) says that

> If n resistances are connected in parallel, in the manner of Fig. 28, the *reciprocal of the total resistance* is equal to the *sum of the reciprocals of the individual resistances*.

The *power input* to *each resistance* in the parallel-connected circuit of Fig. 28 is (by eqs. (15), (16), (17)) found by any of the formulas

$$P_n = VI_n = V^2/R_n = I_n^2 R_n \tag{33}$$

where V is the battery voltage, and I_n is the current in resistor R_n. The *total power output P of the battery* is given by eqs. (25) and (26), where I is the battery current and where, now, R_T is found by means of eq. (32). The following example will be helpful.

Example

In Fig. 29, the battery voltage is $V = 65$ volts, and the values of the resistances, in ohms, are 38, 17, and 27, as shown.

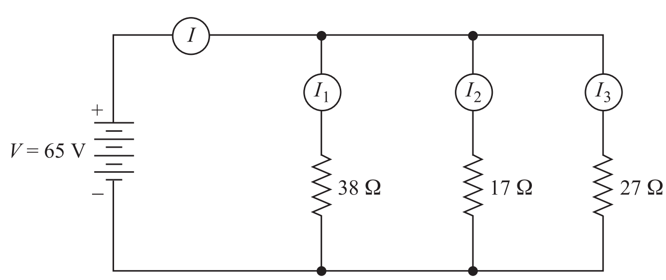


Fig. 29

In Fig. 29 we wish to find the following values:

(a) total resistance seen by the battery,
(b) current measured by the ammeters shown in the figure,
(c) power output of the battery,
(d) power input to each resistor.

Solutions

(a) Note that the three resistors are connected in parallel. Hence the resistance R_T seen by the generator (battery) is, by eq. (32),

$$\frac{1}{R_T} = \frac{1}{38} + \frac{1}{17} + \frac{1}{27} = 0.122\,176 \text{ reciprocal ohms,}$$

thus $R_T = 1/0.122176 = 8.18489$ ohms, *answer*.

(b) I = battery current $= 65/R_T = 65/8.18489 = 7.94146$ amperes, *answer*.
$I_1 = 65/R_1 = 65/38 = 1.710526$ amperes, *answer*.
$I_2 = 65/R_2 = 65/17 = 3.823529$ amperes, *answer*.
$I_3 = 65/R_3 = 65/27 = 2.407407$ amperes, *answer*.

(c) $P = VI = (65)(7.94146) = 516.195$ watts, *answer*.

(d) $P_1 = VI_1 = 65(1.710526) = 111.185$ watts, *answer*.
$P_2 = VI_2 = 65(3.823529) = 248.529$ watts, *answer*.
$P_3 = VI_3 = 65(2.407407) = 156.482$ watts, *answer*.

You should verify using your calculator, that the answers to parts (c) and (d) satisfy eq. (26).

Note: For the special case of two resistors in parallel, as in Fig. 30, eq. (32) gives the value $1/R_T = 1/R_1 + 1/R_2$ which, after combining the two fractions together over the common denominator R_1R_2, then inverting both sides, becomes

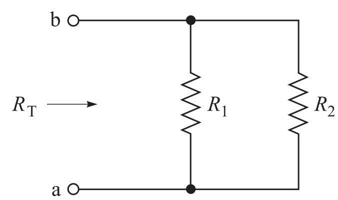


Fig. 30

$$\boxed{R_T = \frac{R_1R_2}{R_1 + R_2}} \tag{34}$$

Thus, the resistance seen looking into terminals a, b in Fig. 30 is given by eq. (34), being equal to the "product of the two resistors, over their sum." The combination of two resistors in parallel occurs so often in practical work that eq. (34) should be memorized.

Another special case sometimes encountered is that where the n resistors in Fig. 28 *all have the same value* of R ohms. In that case, eq. (32) becomes $1/R_T = n/R$, which, after inverting both sides, becomes

$$R_T = \frac{R}{n} \tag{35}$$

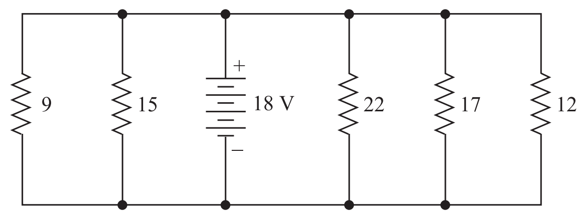
Thus, in Fig. 28, if the n parallel resistances *all have the same value of R ohms*, then the battery sees a total resistance of R/n ohms.

Problem 17

A battery having a constant terminal voltage of 18 volts is applied to five parallel-connected resistance loads, as shown below. Resistance values in ohms.
Find the following:

(a) resistance seen by battery,

(b) battery current,



(c) power output of battery,

(d) current in each resistor (check to see that eq. (28) is satisfied),

(e) power in each resistor (check to see that eq. (26) is satisfied).

Problem 18
A battery of constant voltage 15 volts is connected to two parallel-connected resistor loads of 25 ohms and 38 ohms. Find battery current and battery power.

Problem 19
A battery of constant 12 volts is applied to sixteen 25-ohm resistors connected in parallel. Find (a) battery current, (b) current in each resistor, (c) battery power output, (d) power delivered to each resistor.

2.7 Series-Parallel Circuits

Series-parallel circuits, also called "networks," consist of individual groups of series and parallel resistors. Such circuits, as long as they consist only of individual groups of series and parallel resistances, can always be reduced to *a single equivalent resistance*. Consider, as an example, the series-parallel circuit shown in Fig. 31, in which we wish to find the battery current I. It is given that the battery voltage is constant 45 volts, and the resistance values are in ohms.

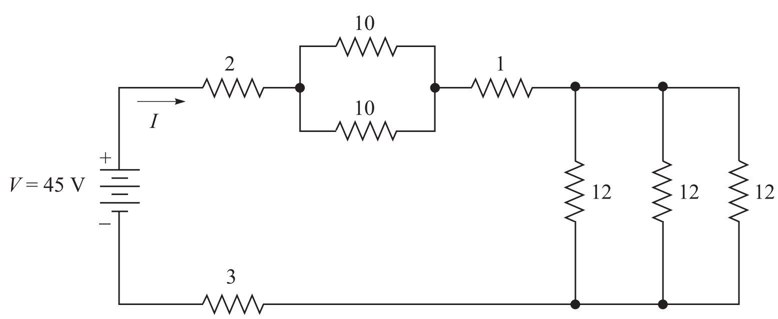


Fig. 31

Solution
First, by eq. (35), the two parallel 10-ohm resistors can be replaced with an equivalent single 5-ohm resistor, and the three parallel 12-ohm resistors can be replaced with a single 4-ohm resistor. When this is done, Fig. 31 becomes the simple *series circuit* shown in Fig. 32 (where "45 V" is read as "45 volts").

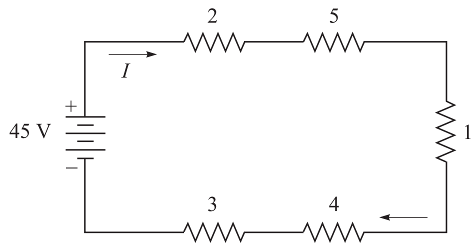


Fig. 32

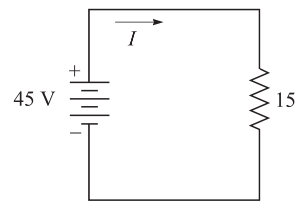


Fig. 33

Since the circuit of Fig. 32 is now a purely series circuit (section 2.5) we have that the original circuit of Fig. 31 reduces to a *single equivalent resistance of 15 ohms*, as shown in Fig. 33. Hence, by Ohm's law, eq. (11), the battery current I is equal to

$$I = V/R_T = 45/15 = 3 \text{ amperes, } \textit{answer}.$$

Problem 20
A battery of constant 36 volts is connected to the series-parallel circuit in Fig. 34. Resistance values are in ohms.

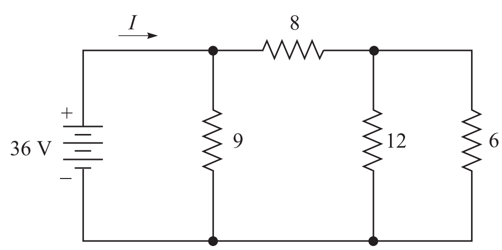
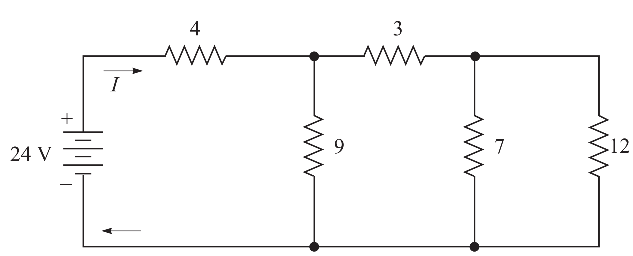


Fig. 34

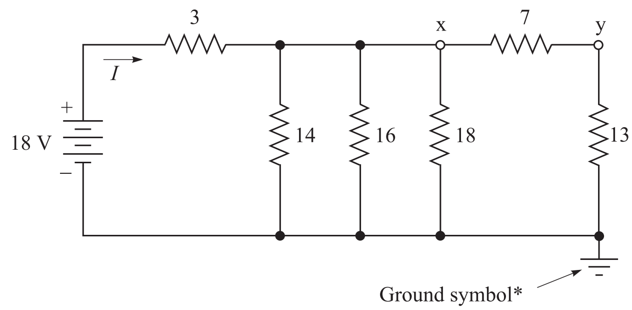
In the figure, let it be required to find:

(a) battery current I,
(b) power output of battery,
(c) current in 6-ohm resistor.

Problem 21
A battery of constant 24 volts is applied to the series-parallel network shown below. Resistance values in ohms. Find the battery current. (*Answer*: 2.97497 amps)



Problem 22
Given the series-parallel network as follows, resistance values in ohms, find:



(a) potential of point x with respect to ground,
(b) potential of point y with respect to ground.

Problem 23
What value of resistance must be connected in parallel with a 36-ohm resistor if the parallel combination is to be equivalent to a single 20-ohm resistor?

* The term "ground" is often used to denote a common reference point or reference line in a network. Such a "ground" may or may not be connected to an actual earth ground.

CHAPTER 3

Determinants and Simultaneous Equations

3.1 Introduction to Determinants

The purpose of this chapter is to prepare us for future work in the writing and solution of network equations. We'll find that network analysis produces systems of simultaneous equations, and such systems are most conveniently handled by making use of what are called *determinants*.

The study of determinants is not basically difficult, but it will call for close attention to details on your part. The results, however, will be well worth the time and effort you put into it. Let us begin with some definitions, as follows.

A ‘‘determinant’’ is a *square array of numbers*, or letters used to represent numbers, placed between two vertical bars. The numbers or letters are arranged in horizontal *rows* and vertical *columns*. An example of what a determinant looks like is shown in Fig. 35.

Notice that the *rows* are numbered from the top down, and the *columns* from left to right.

$$\begin{array}{rc}
 & \begin{array}{cccc}\text{1st column} & \text{2nd column} & \text{3rd column} & \text{4th column}\end{array} \\
\begin{array}{r}\text{1st row} \rightarrow \\ \text{2nd row} \rightarrow \\ \text{3rd row} \rightarrow \\ \text{4th row} \rightarrow \end{array} & \begin{vmatrix} 3 & 5 & -2 & 7 \\ 2 & 0 & 4 & -3 \\ 5 & 3 & 1 & 6 \\ -4 & 0 & -2 & 4 \end{vmatrix} = D
\end{array}$$

Fig. 35

Each number or letter in a determinant is called an *element* of the determinant. Since a determinant is always a *square* array of elements, the number of *rows* is always the same as the number of *columns*. Note that all rows and columns have the same number of elements.

A determinant is classified according to the number of rows (or columns) it has. The determinant in Fig. 35 is thus a "fourth-order" determinant, because it has four rows (and also, of course, four columns).

A determinant has a *value* equal to a *single number*. For instance, later on we'll be able to show that for the determinant of Fig. 35, $D = 312$.

The *location* of an element in a determinant will always be specified by giving FIRST *the number of the* ROW and THEN *the number of the* COLUMN it is located in.

For example, the location of the element −4 in Fig. 35 would be given as (4,1), meaning it is located at the intersection of the *fourth row* and the *first column*. As another example, the location of the element "6" would be specified as (3,4), meaning at the intersection of the *third row* and the *fourth column*.

Many of our discussions will be easier to follow if we represent the elements of a determinant by a letter with subscripts. The system of notation used is illustrated by the fourth-order determinant of Fig. 36. This figure illustrates how *subscripts* are used to identify the location of the elements in a determinant.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

Fig. 36

Notice that the subscript used with each element denotes *first the* ROW *and then the* COLUMN in which the element appears. This convention, of giving first the row and then the column, is always used. For example, the notation a_{23} denotes the element at the intersection of the second row and the third column (the notation a_{23} can be read as "a, two, three"). When it is deemed necessary, the row and column numbers can be separated by a comma, as, for example, $a_{16,11}$.

Problem 24

(a) How many elements in a seventh-order determinant?

(b) Using a with subscript, identify the element at the intersection of the fifth row and third column of a determinant of order five or higher.

(c) What is the difference between $a_{1,11}$ and $a_{11,1}$?

3.2 The Second-Order Determinant

By definition, a "second-order determinant" has two rows and two columns, and thus four elements. Using the standard notation of section 3.1, the general form of the second-order determinant is shown in Fig. 37.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Fig. 37

The second-order determinant is the *basic determinant*; later on we'll find that *all* determinants, regardless of order, can be expressed in terms of second-order determinants.

So far we've not given any meaning to the set of symbols in Fig. 37; that is, we've not *defined* what a determinant is to mean. For reasons that will become apparent to us later on, the *value*, D, of a *second-order determinant* is now defined to be as follows,

> The value of a second-order determinant is defined as being equal to the product $a_{11}a_{22}$ *minus* the product $a_{12}a_{21}$.

Thus, by definition, Fig. 37 has the value

$$D = \begin{matrix} a_{11} & & a_{12} \\ & \times & \\ a_{21} & & a_{22} \end{matrix} = a_{11}a_{22} - a_{12}a_{21} \tag{36}$$

In eq. (36) we used the two arrows to show the two multiplications, but such arrows are not, of course, normally shown.

Note that there is no "proof" of anything required here, because we are simply *defining* what a second-order determinant is. Later on you'll find out why it's convenient to define the meaning in this way.

We should also mention that there are no restrictions on what the numbers, represented by the letters, can be. Thus the elements of a determinant can be real numbers or complex numbers, or any combination of such numbers.

Example

Find the value of the determinant $\begin{vmatrix} 4 & 3 \\ 2 & 5 \end{vmatrix}$

Solution

This is a second-order determinant, and thus by eq. (36) we have,

$$(4)(5) - (3)(2) = 20 - 6 = 14, \quad \textit{answer}.$$

Find the values of the second-order determinants in problems 25 through 29.

Problem 25

(a) $\begin{vmatrix} 6 & 2 \\ 4 & 2 \end{vmatrix} =$ (b) $3\begin{vmatrix} 1 & -2 \\ 4 & 2 \end{vmatrix} =$

Problem 26

$$\begin{vmatrix} 4 & -2 \\ -3 & 6 \end{vmatrix} + \begin{vmatrix} 1 & -1 \\ 2 & -3 \end{vmatrix} + \begin{vmatrix} -4 & -2 \\ 3 & -1 \end{vmatrix} =$$

Problem 27

$$\frac{\begin{vmatrix} 5 & 10 \\ 6 & 4 \end{vmatrix}}{\begin{vmatrix} 4 & 4 \\ 7 & 5 \end{vmatrix}} =$$

Problem 28

$$\begin{vmatrix} 5x & y \\ -y & 2xy \end{vmatrix} =$$

Problem 29

$$\begin{vmatrix} (x+2) & (x-5) \\ 4 & (2-x) \end{vmatrix} =$$

3.3 Minors and Cofactors. Value of any *N*th-order Determinant

Let "N" be the *order* of any determinant, where N is any whole number greater than 1. Thus, if $N = 2$ we have a second-order determinant, if $N = 3$ we have a third-order determinant, and so on. An *Nth-order* determinant refers to a determinant of any order whatever, and is just the algebraic way of saying that we are talking about a determinant of any order in general.

Every element of a determinant has what is called the *minor* of that element. To find *the* MINOR *of any element*, *strike out the row and column in which the element is located*. The determinant that remains is defined to be the *minor* of that element. This is illustrated in the following example.

Example

Given the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

In accordance with the above definitions, write the minors of the elements a_{21}, a_{33}, and a_{22}.

Solutions

Minor of a_{21} is

$$\begin{vmatrix} \cancel{a_{11}} & a_{12} & a_{13} \\ \cancel{a_{21}} & \cancel{a_{22}} & \cancel{a_{23}} \\ \cancel{a_{31}} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} \qquad \textit{answer}.$$

Minor of a_{33} is

$$\begin{vmatrix} a_{11} & a_{12} & \cancel{a_{13}} \\ a_{21} & a_{22} & \cancel{a_{23}} \\ \cancel{a_{31}} & \cancel{a_{32}} & \cancel{a_{33}} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \qquad \textit{answer}.$$

Minor of a_{22} is

$$\begin{vmatrix} a_{11} & \cancel{a_{12}} & a_{13} \\ \cancel{a_{21}} & \cancel{a_{22}} & \cancel{a_{23}} \\ a_{31} & \cancel{a_{32}} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \qquad \textit{answer}.$$

Now, in any Nth-order determinant, let a_{ij} denote the element at *the intersection of any* ith *row and* jth *column* (“eye-th row and jay-th column”), as illustrated in Fig. 38.

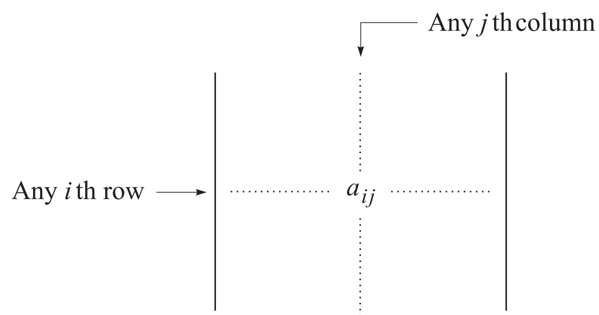


Fig. 38

As already defined, if we *strike out* the ith row and jth column, the determinant that remains is called the *minor* determinant of the element a_{ij}. Let us denote the minor of a_{ij} by M_{ij}, that is, let

$$M_{ij} = \text{the minor determinant of element } a_{ij}.$$

Now let A_{ij} denote what we’ll call the *cofactor* of element a_{ij}. The “cofactor” is defined as

$$\text{cofactor of } a_{ij} = A_{ij} = (-1)^{i+j} M_{ij} \tag{37}$$

The “cofactor” of a_{ij} is thus equal to the “minor” of a_{ij} multiplied by either $+1$ or -1, depending upon whether the sum, $i+j$, is an even or odd number, because $(-1)^{i+j} = +1$ if $i+j$ is *even*, but $(-1)^{i+j} = -1$ if $i+j$ is *odd.*

The *value of any Nth-order determinant* is now defined in terms of the *cofactors of any row or column*, as follows.

1. Select *any row* or *any column* of the determinant.
2. *Multiply each element* in that row (or column) by its cofactor.
3. The *value of the determinant* is the *sum* of the N products found in this way. Only one row, or one column, is used in this procedure.

You might think, offhand, that the above-defined procedure might conflict with the basic definition of eq. (36). This is not the case, however, because the *end result* of the above procedure is always the *sum of a number of second-order determinants*, the value of each such second-order determinant then being found by eq. (36). This will be illustrated in the following three examples, in which you'll notice that, in each case, the *final answer* is the *sum of several second-order determinants*.

Example 1
Find the value of the third-order determinant

$$\begin{vmatrix} 3 & 2 & 5 \\ 6 & 1 & 2 \\ 4 & -2 & 3 \end{vmatrix}$$

Solution
We'll use the three steps above, as follows.

1. We can select any row or any column we wish.* Let's suppose we decide to use the elements of the *first row* in this solution. Note that here the elements of the first row are $a_{11} = 3, a_{12} = 2$, and $a_{13} = 5$.
2. In this step we are to multiply each of the three elements in the first row by the *cofactor* of the element. Doing this, making use of eq. (37), gives the following three products:

$$a_{11}A_{11} = a_{11}(-1)^2 M_{11} = a_{11}M_{11} = 3\begin{vmatrix} 1 & 2 \\ -2 & 3 \end{vmatrix}$$

$$a_{12}A_{12} = a_{12}(-1)^3 M_{12} = -a_{12}M_{12} = -2\begin{vmatrix} 6 & 2 \\ 4 & 3 \end{vmatrix}$$

$$a_{13}A_{13} = a_{13}(-1)^4 M_{13} = a_{13}M_{13} = 5\begin{vmatrix} 6 & 1 \\ 4 & -2 \end{vmatrix}$$

3. By step 3, letting D be the value of the determinant, we have,

$$D = 3\begin{vmatrix} 1 & 2 \\ -2 & 3 \end{vmatrix} - 2\begin{vmatrix} 6 & 2 \\ 4 & 3 \end{vmatrix} + 5\begin{vmatrix} 6 & 1 \\ 4 & -2 \end{vmatrix}$$

Now, by eq. 36,

$$D = 3(3+4) - 2(18-8) + 5(-12-4) = -79, \quad \textit{answer}.$$

Example 2
Rework example 1, this time using the *second column*, instead of the first row.

Solution
We'll follow the three steps as in example 1, as follows.

1. The elements of the *second column* are (for the determinant of example 1)

$$a_{12} = 2, \qquad a_{22} = 1, \qquad a_{32} = -2$$

---

* The value, D, of a given determinant of any order is the *same* regardless of the row or column we decide to use. Take, for example, the general 3rd-order determinant shown at the beginning of this section. If you carefully "expand" the determinant in terms of the elements of ANY row or ANY column, your answer should reduce to the same value

$$D = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

in every case. The same principle applies to determinants of any order.

2. Same procedure as in example 1, except now we use the elements of the second column, as follows.

$$a_{12}A_{12} = a_{12}(-1)^3 M_{12} = -a_{12}M_{12} = -2\begin{vmatrix} 6 & 2 \\ 4 & 3 \end{vmatrix}$$

$$a_{22}A_{22} = a_{22}(-1)^4 M_{22} = a_{22}M_{22} = \begin{vmatrix} 3 & 5 \\ 4 & 3 \end{vmatrix}$$

$$a_{32}A_{32} = a_{32}(-1)^5 M_{32} = -a_{32}M_{32} = -(-2)\begin{vmatrix} 3 & 5 \\ 6 & 2 \end{vmatrix}$$

3. By step 3, letting D be the value of the determinant, we have

$$D = -2\begin{vmatrix} 6 & 2 \\ 4 & 3 \end{vmatrix} + \begin{vmatrix} 3 & 5 \\ 4 & 3 \end{vmatrix} + 2\begin{vmatrix} 3 & 5 \\ 6 & 2 \end{vmatrix}$$

and now, by eq. (36),

$D = -2(18 - 8) + (9 - 20) + 2(6 - 30) = -79,$ same *answer* as in example 1.

Example 3

Find the value of the fourth-order determinant

$$\begin{vmatrix} 4 & 3 & 1 & 3 \\ 2 & 1 & 4 & 2 \\ 6 & 0 & 0 & 5 \\ 1 & 3 & 7 & 2 \end{vmatrix}$$

Solution

Let us carry out the usual three-step solution, as follows.

1. In this particular case the *easiest* thing to do is to make use of the *third row*; this is because there are two *zeros* in the third row. We'll find that the presence of the zeros will considerably reduce the work required to get the solution. Note that the elements of the third row are

$$a_{31} = 6 \qquad a_{32} = 0 \qquad a_{33} = 0 \qquad a_{34} = 5$$

2.

$$a_{31}A_{31} = a_{31}(-1)^4 M_{31} = a_{31}M_{31} = 6\begin{vmatrix} 3 & 1 & 3 \\ 1 & 4 & 2 \\ 3 & 7 & 2 \end{vmatrix}$$

$$a_{32}A_{32} = 0$$

$$a_{33}A_{33} = 0$$

$$a_{34}A_{34} = a_{34}(-1)^7 M_{34} = -a_{34}M_{34} = -5\begin{vmatrix} 4 & 3 & 1 \\ 2 & 1 & 4 \\ 1 & 3 & 7 \end{vmatrix}$$

3. By step 3, letting D be the value of the determinant, we have

$$D = 6\begin{vmatrix} 3 & 1 & 3 \\ 1 & 4 & 2 \\ 3 & 7 & 2 \end{vmatrix} - 5\begin{vmatrix} 4 & 3 & 1 \\ 2 & 1 & 4 \\ 1 & 3 & 7 \end{vmatrix}$$

We thus now have to find the values of the two third-order determinants above. The results are shown below, in which both determinants have been expanded in terms of the elements of the first column.

$$D_1 = 6\left(3\begin{vmatrix} 4 & 2 \\ 7 & 2 \end{vmatrix} - \begin{vmatrix} 1 & 3 \\ 7 & 2 \end{vmatrix} + 3\begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix}\right) = (6)(-29) = -174 = \text{value of } \textit{first} \text{ determinant.}$$

$$D_2 = -5\left(4\begin{vmatrix} 1 & 4 \\ 3 & 7 \end{vmatrix} - 2\begin{vmatrix} 3 & 1 \\ 3 & 7 \end{vmatrix} + \begin{vmatrix} 3 & 1 \\ 1 & 4 \end{vmatrix}\right) = (-5)(-45) = 225 = \text{value of } \textit{second} \text{ determinant.}$$

The value of the given fourth-order determinant is, therefore,

$$D = D_1 + D_2 = -174 + 225 = +51, \quad \textit{final answer}.$$

Problem 30
Find the value of the following determinant, by expanding in terms of the elements of the first column.

$$\begin{vmatrix} 3 & 6 & 1 \\ -5 & 7 & -4 \\ 1 & -2 & 3 \end{vmatrix}$$

Problem 31
Repeat problem 30, this time expanding in terms of the elements of the third row (answer must be same as that in problem 30).

Problem 32
Find the value of the determinant

$$\begin{vmatrix} 6 & 1 & 0 \\ 2 & 4 & 0 \\ -3 & 5 & -4 \end{vmatrix}$$

Problem 33*
Find the value of the determinant

$$\begin{vmatrix} 2 & 2 & 5 & -4 \\ 1 & 0 & 2 & 2 \\ -3 & -1 & 3 & 0 \\ -1 & 0 & 6 & 3 \end{vmatrix}$$

* See footnote given with the solution to problem 33 (after Appendix).

Problem 34
Find the value of the determinant

$$\begin{vmatrix} 1 & 3 & 2 & 5 \\ 6 & 1 & -3 & 2 \\ 4 & -2 & 0 & 5 \\ 0 & 0 & 0 & -5 \end{vmatrix}$$

3.4 Some Important Properties of Determinants

There are some very useful properties of determinants that we should be familiar with. It won't be necessary that we stop, here, to give a formal proof of each property, but we'll explain the meaning of each one in detail.

To begin, let us state that the properties we'll study in this section are true for determinants of any order. Furthermore, any statement we will make about rows will *also* apply to columns, and vice versa. We'll also, when necessary, make use of the fact that the value D of a determinant is the same regardless of which rows or columns we happen to make use of in the process of finding the value of D; this was pointed out in section 3.3.

As we found in section 3.3, the value D of a determinant is equal to the sum of a number of terms, each term being a product of several different elements of the determinant. The first property of determinants (property 1) that we now wish to consider concerns the number and nature of such terms. We'll proceed in steps, as follows.

As already defined, an Nth-order determinant is a square array of N^2 elements arranged in an equal number of N rows and N columns. A determinant is equal to a single value, which we'll generally denote by D.

In section 3.2 we learned that the second-order determinant is the basic or "prototype" determinant, its structure and value being defined by eq. (36).

In section 3.3 we learned that any Nth-order determinant can be "expanded" in terms of the "minor" determinants of any row or column. Each such minor determinant can then be expanded in terms of its minor determinants, until finally, continuing on in this way, the original Nth-order determinant will be found to be equal to a *sum of a number of* SECOND-ORDER *determinants*, the value of each such second-order determinant being found by eq. (36). With this in mind, let us see what we can discover about the number and the nature of the terms whose sum equals the value D of the determinant.

To do this, let us begin with the third-order determinant ($N = 3$), shown in Fig. 39.

$$D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Fig. 39

Now expand Fig. 39 in terms of, say, the elements of the first row; thus

$$D_3 = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

and hence, by the basic eq. (36) we have that

$$D_3 = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \tag{38}$$

Inspection of eq. (38) shows that the value of a third-order determinant is equal to the *sum of 6 terms*, each term being the *product of 3 elements*. Note that each term has as factors one element, and only one, from each row and each column. For example, in the term "$a_{12}a_{23}a_{31}$,"

element a_{12} is in row 1 and column 2,
element a_{23} is in row 2 and column 3,
element a_{31} is in row 3 and column 1,

that is, every row and every column is represented once, and only once, in every term, as inspection of the subscripts in eq. (38) will show.

To continue, let us next consider a fourth-order determinant ($N = 4$) such as is shown in Fig. 40.

$$D_4 = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

Fig. 40

Now expand Fig. 40 in terms of the minors of, say, the first row, thus,

$$D_4 = a_{11}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14}\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix} \tag{39}$$

Now, from our study of Fig. 39 we know that a third-order determinant is equal to the *sum of 6 terms*; since eq. (39) (which applies to Fig. 40) is the sum of *four* such third-order determinants, it follows that the value of a fourth-order determinant is equal to the sum of $4 \times 6 = 24$ *terms*, each term consisting of *the product of 4 elements*. Again, every row and every column is represented just *once*, and only once, in each of the 24 terms (as examples, $a_{11}a_{22}a_{33}a_{44}, a_{12}a_{23}a_{31}a_{44}, a_{13}a_{22}a_{34}a_{41}$, and so on).

Now consider a *fifth-order determinant* ($N = 5$), and suppose the determinant is expanded in terms of the minors of any row or column we might choose. This will produce *five fourth-order determinants*, and since we know that each fourth-order determinant is equal to 24 terms, it follows that a fifth-order determinant is equal to the sum of $24 \times 5 = 120$ *terms*, each term consisting of the *product of 5 elements*. (Again, every row and every column will be represented just once in every one of the 120 terms.)

Next, the expansion of a sixth-order determinant will consist of the sum of $120 \times 6 = 720$ *terms* (each term the product of 6 elements), the expansion of a seventh-order determinant will consist of the sum of $720 \times 7 = 5040$ *terms* (each term the product of 7 elements), and so on and on in this manner.

In the above you may have noticed that $6 = 1 \cdot 2 \cdot 3$, $24 = 1 \cdot 2 \cdot 3 \cdot 4$, $120 = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$, and so on. Thus, defining the "factorial" notation, $N! = 1 \cdot 2 \cdot 3 \cdot 4 \ldots N$, we can summarize:

> **Property 1**. The expansion of an Nth-order determinant consists of the sum of $N!$ terms, each term the product of N elements, each term having as factors one element, and only one, from each row and each column.

As you can appreciate, finding the value of a high-order determinant using only paper and pencil would, practically speaking, be almost impossible. For example, the expansion of a tenth-order determinant would consist of $10! = 3{,}628{,}800$ terms, with 10 multiplications required to calculate the value of each term. Fortunately, however, this kind of work is exactly what the digital computer is extremely good at, doing millions of such calculations per second. Thus the solution of problems involving higher-order determinants is entirely practical.

Next, suppose the elements of any row or any column are all equal to *zero*; in such a case, "property 2" states:

> **Property 2**. If all the elements of any row or any column are equal to zero, the value of the determinant is equal to zero.

By property 1, every term in the expansion of such a determinant would contain zero as a factor, and thus the value of the determinant would be zero. Next we have

> **Property 3**. If any two rows or any two columns of a determinant are interchanged, the new determinant is equal in magnitude but opposite in sign to the original determinant.

As an aid to understanding why property 3 is true, let us use the 3rd-order determinant of Fig. 39 and eq. (38) as an example, and, for convenience, let us rewrite eq. (38) in the form

$$D_3 = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31})$$

Now, in Fig. 39, let us interchange, say, *rows 1 and 3*. In the last equation for D_3, above, this is done by changing, in the subscripts, *row 1 to row 3* and *row 3 to row 1*, to get

$$D_3 = (a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12}) - (a_{31}a_{23}a_{12} + a_{32}a_{21}a_{13} + a_{33}a_{22}a_{11})$$

Note that the second value for D_3 has the *same magnitude* as the first, but *opposite sign*. Continuing in this manner, we find that the $N!$ terms in the expansion of any Nth-order determinant always divide into two separate "plus and minus" groups, both groups having the same number of terms. If, now, any two rows, or any two columns, of the given determinant are interchanged, the same two groups of terms appear, but with opposite signs, in the new determinant. Next we have

> **Property 4**. If any two rows, or any two columns, of a determinant are identical, the value of the determinant is zero.

Let D be the value of a determinant having two identical rows. Now interchange the two rows; since they are identical rows, this does not change the value of D. However, by property 3, interchanging the two rows must produce $-D$; that is, we would have to have $D = -D$, which can only be true if $D = 0$. Now consider

> **Property 5**. If "m" is a factor of every element in any row or any column, then m may be removed from the elements of that row or column and placed in front of the determinant as a multiplier of the entire determinant.

Suppose that m is a factor of every element in a certain row or column of a determinant. By property 1, every term in the expansion of the determinant will contain m as a factor; but this produces the same result as first removing m, then expanding the determinant, and then multiplying the result by m. To illustrate,

$$\begin{vmatrix} ma_{11} & a_{12} & a_{13} \\ ma_{21} & a_{22} & a_{23} \\ ma_{31} & a_{32} & a_{33} \end{vmatrix} = m \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

From this we see that property 5 gives us the rule that "to multiply a determinant by a factor m, multiply every element of any *one row* or *one column* by m."

Let us next take up property 6, which is stated as follows:

> **Property 6**. If every element of any row or any column is the sum of two terms, then the determinant can be written as the sum of two determinants.

Using the following third-order determinant as an example, property 6 states that

$$\begin{vmatrix} (a_{11}+A) & a_{12} & a_{13} \\ (a_{21}+B) & a_{22} & a_{23} \\ (a_{31}+C) & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} A & a_{12} & a_{13} \\ B & a_{22} & a_{23} \\ C & a_{32} & a_{33} \end{vmatrix}$$

Note that in the original determinant (the left-hand side of the equation) each of the elements of the first column consists of the sum of two quantities. Each of these two quantities then appears separately as the first columns of the two determinants on the right-hand side of the equation.

The above equation, for the case of a 3rd-order determinant, can be verified by actually expanding the original determinant in terms of the elements of the first column. The extension to the case of any Nth-order determinant is then apparent. Next we have our final property, which is very useful:

Property 7. The value of a determinant is not changed if, to each element of any row, we add the corresponding element of any other row multiplied by the same number.

As usual, the statement applies to columns as well as rows.

To illustrate the meaning of property 7, let's begin with what we'll call the "original" determinant, shown on the left-hand side of eq. (40) (Fig. 41).

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} (a_{11}+ma_{13}) & a_{12} & a_{13} \\ (a_{21}+ma_{23}) & a_{22} & a_{23} \\ (a_{31}+ma_{33}) & a_{32} & a_{33} \end{vmatrix} \tag{40}$$

Fig. 41

In eq. (40), the right-hand side is the result of adding, to each element of the first column, the corresponding element of the third column multiplied by m.* We can show that the two sides of eq. (40) are equal as follows.

First apply, to the right-hand side, property 6 and then property 5; if we do this, the right-hand side becomes

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + m \begin{vmatrix} a_{13} & a_{12} & a_{13} \\ a_{23} & a_{22} & a_{23} \\ a_{33} & a_{32} & a_{33} \end{vmatrix}$$

Notice, now, that the second determinant above has *two identical columns*, and is thus equal to zero by property 4, proving that the two sides of eq. (40) are equal.

Property 7 can sometimes be used to greatly reduce the amount of work needed to find the value of a given determinant. This is accomplished by using property 7 to transform a given determinant into an equivalent determinant having *more zeros* as elements than the given determinant has. The following example will illustrate the procedure.

Example

Given the determinant

$$D = \begin{vmatrix} 5 & 1 & -2 \\ 3 & 2 & 1 \\ 2 & -4 & 3 \end{vmatrix}$$

(a) Find the value of D by expanding, in the usual way, the determinant in terms of, say, the elements of the second column.

Solution

$$D = -\begin{vmatrix} 3 & 1 \\ 2 & 3 \end{vmatrix} + 2\begin{vmatrix} 5 & -2 \\ 2 & 3 \end{vmatrix} + 4\begin{vmatrix} 5 & -2 \\ 3 & 1 \end{vmatrix}$$

$$= -(9-2) + 2(15+4) + 4(5+6) = -7 + 38 + 44 = 75, \quad \textit{answer}.$$

* Property 7 applies to any combination of two columns, or two rows, we might wish to use.

(b) Going back to the given determinant, find D by making use of property 7.

Solution
One way is as follows. To each element of row 2 add the corresponding element of row 1 multiplied by -2. Next, to each element of row 3 add the corresponding element of row 1 multiplied by 4; thus,

$$D = \begin{vmatrix} 5 & 1 & -2 \\ 3-10 & 2-2 & 1+4 \\ 2+20 & -4+4 & 3-8 \end{vmatrix} = \begin{vmatrix} 5 & 1 & -2 \\ -7 & 0 & 5 \\ 22 & 0 & -5 \end{vmatrix}$$

Notice now that all the elements in column 2 except one are *zeros*; this makes it easy to expand the determinant in terms of the elements of column 2, thus,

$$D = -\begin{vmatrix} -7 & 5 \\ 22 & -5 \end{vmatrix} = -(35 - 110) = 75, \ \textit{answer}, \ \text{as before.}$$

Problem 35
Given that

$$D = \begin{vmatrix} 6 & -3 & 5 \\ 24 & -16 & 120 \\ -12 & 1 & 25 \end{vmatrix}$$

use property 5 as an aid in finding the value of D.

Problem 36
Find the value of

$$\begin{vmatrix} 3 & 6 & -10 & 7 \\ 1 & 0 & 5 & -14 \\ -8 & 0 & 20 & 14 \\ 2 & 0 & -15 & 14 \end{vmatrix}$$

Problem 37
In a determinant, if the elements of any given row (or column) are added to, or subtracted from, the corresponding elements of any other row (or column), is the value of the determinant changed? (The given row or column remains, of course, unchanged in its original position.)

Problem 38
By inspection, verify that

$$\begin{vmatrix} 3 & -1 & 2 & 6 \\ 4 & 3 & 0 & 8 \\ 1 & -4 & 4 & 2 \\ 6 & 7 & 5 & 12 \end{vmatrix} = 0$$

Problem 39
Verify that

$$\begin{vmatrix} 2 & 0 & -1 & 5 & 0 \\ 0 & -1 & 0 & 2 & 4 \\ -1 & 3 & 2 & 0 & -1 \\ 3 & -2 & 0 & -1 & -1 \\ -2 & 4 & -3 & 2 & 0 \end{vmatrix} = 596$$

3.5 Determinant Solution of Linear Simultaneous Equations

A *linear* or "first degree" equation is one in which the unknowns are all raised to the first power. Let us take, as an example, the general form of a linear equation in three unknowns; thus,

$$ax + by + cz = k \tag{41}$$

where x, y, z denote the values of three unknown quantities, with a, b, c being the corresponding *constant coefficients* of the unknowns, and k denoting a single constant term on the right-hand side.

In our problems the value of the constants will be known, and we will be required to find the values of the unknown quantities.

As a general principle we know that the more complicated a problem is, *the greater is the amount of information needed* to solve the problem. For the case of linear simultaneous equations, this simply means that *the greater the number of unknown values* that must be found, the *greater is the number of equations* that will be required to find the values. This can be stated as follows.

> A problem involving linear equations in n unknowns requires, in general, n independent equations for its solution.*

Thus, if a problem involves linear equations in *two unknowns*, then *two equations* are required to find a solution; or, if a problem involves linear equations in *three unknowns*, then *three equations* are required to find a solution; and so on.

With the above in mind, we now wish to introduce a *general procedure* that can be used to find the solution to any system of n simultaneous linear equations.

To do this, let us take, as an example, the solving of a set of three simultaneous linear equations, in which we'll denote the values of the three unknowns by x, y, and z. Doing this will clearly show why the procedure is correct, and why it is valid for any set of n such equations. We proceed as follows.

Let the three simultaneous equations be designated as eq. (42), in which a through i represent constant coefficients and where P, Q, and R are constant terms on the right-hand sides of the equations, as shown. (To reduce the writing time we'll not use the usual "double-subscript" notation in this discussion.)

* Two equations are "independent" if neither one can be derived from the other. As a simple example, $x + y = 10$ and $2x + 2y = 20$ are not independent, because the second one was derived from the first by multiplying it by 2.

The first step, in using determinants to solve a system of simultaneous linear equations, is to set up a determinant *formed from the coefficients of the unknowns*, as shown in eq. (43).

$$\left.\begin{aligned} ax + by + cz &= P \\ dx + ey + fz &= Q \\ gx + hy + iz &= R \end{aligned}\right\} \tag{42}$$

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \tag{43}$$

Here we're following the custom of using the Greek capital letter delta, Δ, to denote the determinant formed from the coefficients of the unknowns.

Now multiply both sides of eq. (43) by x. Doing this, making use of property 5, we have that

$$x\Delta = \begin{vmatrix} ax & b & c \\ dx & e & f \\ gx & h & i \end{vmatrix} \tag{44}$$

Now apply property 7 to eq. (44) as follows. To each element of column 1 add the corresponding element of column 2 multiplied by y, so that eq. (44) becomes

$$x\Delta = \begin{vmatrix} (ax + by) & b & c \\ (dx + ey) & e & f \\ (gx + hy) & h & i \end{vmatrix} \tag{45}$$

Now apply property 7 to eq. (45) as follows. To each element of column 1 add the corresponding element of column 3 multiplied by z, so that eq. (45) becomes

$$x\Delta = \begin{vmatrix} (ax + by + cz) & b & c \\ (dx + ey + fz) & e & f \\ (gx + hy + iz) & h & i \end{vmatrix} \tag{46}$$

Now, in the above, compare each element in column 1 with the original three simultaneous equations of eq. (42). Doing this shows that *the elements of column 1* in eq. (46) *can be replaced by the constants P, Q, and R*. Doing this, then dividing both sides by Δ, we have that the *value of x* is equal to

$$x = \frac{\begin{vmatrix} P & b & c \\ Q & e & f \\ R & h & i \end{vmatrix}}{\Delta} \tag{47}$$

where Δ is *the determinant formed from the coefficients of the unknowns*, eq. (43).

Next, a formula for the value of y can be found as follows. First, multiply both sides of eq. (43) by y, then make use of property 5; thus,

$$y\Delta = \begin{vmatrix} a & by & c \\ d & ey & f \\ g & hy & i \end{vmatrix} \tag{48}$$

Now apply property 7, as follows. First, to each element of column 2 add the corresponding element of column 1 multiplied by x. Next, to each element of column 2 add the corresponding element of column 3 multiplied by z. Doing this, eq. (48) becomes

$$y\Delta = \begin{vmatrix} a & (ax+by+cz) & c \\ d & (dx+ey+fz) & f \\ g & (gx+hy+iz) & i \end{vmatrix} \tag{49}$$

Now, in the above, compare each element of column 2 with the original three simultaneous equations of eq. (42). Doing this shows that *the elements of column 2* in eq. (49) *can be replaced by the constants* P, Q, *and* R. Doing this, then dividing both sides by Δ, we have that the *value of* y is equal to

$$y = \frac{\begin{vmatrix} a & P & c \\ d & Q & f \\ g & R & i \end{vmatrix}}{\Delta} \tag{50}$$

where, as usual, Δ is *the determinant formed from the coefficients of the unknowns* (eq. (43)).

Next, using the same steps as those used to derive eqs. (47) and (50), we find that the *value of* z is equal to

$$z = \frac{\begin{vmatrix} a & b & P \\ d & e & Q \\ g & h & R \end{vmatrix}}{\Delta} \tag{51}$$

Thus the values of the three unknowns, x, y, and z, in eq. (42) can be found by direct use of eqs. (47), (50), and (51). It's also clear that the procedure used to derive these equations can be repeated *for any number of such simultaneous equations*. Thus we can now set down the RULE *for solving any set of n simultaneous linear equations in n unknowns*, known as "Cramer's rule," as follows.

Step 1. Write down the n equations neatly, in the standard form, as in eq. (42).

Step 2. Form a determinant using the *coefficients of the unknowns* as the elements. Let Δ be the value of this determinant, as in eq. (43).

Step 3. Select the unknown to be solved for. Now go to the determinant found in step 2 and form a new determinant by *replacing the coefficients of the desired unknown with the constant terms on the right-hand sides of the equations*. Call the value of this new determinant Δ' (delta prime).

Step 4. The value of the unknown is then equal to Δ' divided by Δ, that is, Δ'/Δ.*

The following problems are to be worked by applying, step-by-step, the above four-step rule.

* Except if $\Delta = 0$, because division by zero is not permitted.

Problem 40
Find the values of x and y that satisfy the two simultaneous equations

$$\begin{aligned} 5x + 2y &= -7 \\ -3x + 4y &= 25 \end{aligned}$$

Problem 41
Solve the following set of simultaneous linear equations for the unknown values of x, y, and z.

$$\begin{aligned} x + y + z &= 6 \\ x + y - z &= 0 \\ x - 2y - z &= 3 \end{aligned}$$

Problem 42
Same instructions as for problem 41.

$$\begin{aligned} x + y + z &= 4 \\ 3x + 4y - 2z &= -2 \\ -4y - 5z &= 1 \end{aligned}$$

Problem 43
Same instructions as for problem 41, now for w, x, y, and z, as follows.

$$\begin{aligned} w + x + y + z &= -4 \\ 3w - 2x + 4y + 4z &= 0 \\ -2w + 5x + 7y &= -12 \\ 3x + 2y - 3z &= 5 \end{aligned}$$

3.6 Systems of Homogeneous Linear Equations

There is an important class of systems of simultaneous linear equations in which the constant terms are all equal to *zero*. The general form of such a system of n linear equations in n unknown values can be indicated as follows.

Let $x_1, x_2, x_3, \ldots, x_n$ denote the unknown values, and let us denote the constant coefficients by as with subscripts, giving first the row, then the column; thus,

$$\left.\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n = 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n = 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n = 0 \end{array}\right\} \tag{52}$$

The term "linear homogeneous" is a suitable name for such a system because the xs are all raised to the same first power and the constant terms all have the same value of zero. Let us now consider how the unknown values of the xs can be found in such a case (the values of the constant "a" coefficients being given).

First of all, direct inspection of eq. (52) shows that *one solution* is that all the xs have the same value of *zero*, that is, that $x_1 = x_2 = \cdots = x_n = 0$. Our problem, however, is to find the really important values of x, in addition to the obvious, ''trivial,'' answer of zero.

To do this, let us apply the four steps of Cramer's rule. Doing this, we have no trouble with steps 1 and 2, but we find that *a difficulty arises with step 3*, because of the fact that the constant terms on the right-hand side are all equal to *zero*. To illustrate the problem, and show what is required to get a solution, let us work through the example of eq. (42) for the special case of P, Q, and R all equal to *zero*; thus,

$$\left.\begin{aligned} ax + by + cz &= 0 \\ dx + ey + fz &= 0 \\ gx + hy + iz &= 0 \end{aligned}\right\} \tag{53}$$

Now look back to the solutions we originally obtained for eq. (42). Notice that the value of Δ (eq. (43)) will remain unchanged, because Δ does not involve the values of P, Q, and R. Thus the value of Δ for eq. (53) will be the same as that given by eq. (43).

But now look back at the values of x, y, and z, given by eqs. (47), (50), and (51), for the case of eq. (42). Note that now, *for the homogeneous case of eq. (53)*, the equations will each contain a *column of zeros*, and thus, by property 2 of section 3.4, give the trivial value of *zero* for x, y, and z, as shown below.

$$x = \frac{\begin{vmatrix} 0 & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix}}{\Delta} = 0, \qquad y = \frac{\begin{vmatrix} a & 0 & c \\ d & 0 & f \\ g & 0 & i \end{vmatrix}}{\Delta} = 0, \qquad z = \frac{\begin{vmatrix} a & b & 0 \\ d & e & 0 \\ g & h & 0 \end{vmatrix}}{\Delta} = 0$$

Thus the difficulty is that, for the homogeneous case of eq. (53), step 3 of Cramer's rule gives the trivial answer that $x = y = z = 0$. One way, however, of resolving the difficulty is to convert the above expressions into the *indeterminate form* 0/0,* which can then, hopefully, be manipulated to yield non-zero values of the unknowns. To convert the above equations into the 0/0 form we see that Δ must be equal to zero; thus we have that

> A system of n homogeneous linear equations in n unknowns can have non-zero solutions if and only if *the determinant of the coefficients vanishes*, that is, only if $\Delta = 0$.

To continue with the example of eq. (53), this means that Δ, which is given by eq. (43), must be equal to *zero*; thus, expanding eq. (43) in terms of the elements of row 1, we must have that

$$\Delta = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} + b\begin{vmatrix} -d & f \\ -g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} = 0 \tag{54}$$

* In algebra, the division operation ''A divided by B equals C'' is defined to mean that

$$\frac{A}{B} = C \quad \text{if and only if} \quad A = BC$$

Thus consider the expression , $0/0 = C$. By definition this requires that $0 = 0C$, which is true for ANY value of C. Thus we say that ''0/0'' is ''indeterminate'' in value.

IF and only if eq. (54) is satisfied, can the system of eq. (53) have solutions other than the trivial $x = y = z = 0$. Now compare eq. (54) with the equation $ax + by + cz = 0$ from eq. (53); the comparison shows that if eq. (54) is satisfied, then eq. (53) is satisfied for the non-trivial values

$$x = \begin{vmatrix} e & f \\ h & i \end{vmatrix}, \qquad y = \begin{vmatrix} -d & f \\ -g & i \end{vmatrix}, \qquad z = \begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{55}$$

The procedure can be extended to homogeneous linear systems of any order.

Problem 44
Given the homogeneous linear system

$$\begin{aligned} 3x - 2y - 5z &= 0 \\ x - y - z &= 0 \\ 2x - y - 4z &= 0 \end{aligned}$$

Does the system possess non-trivial solutions? If so, find such a solution.

Problem 45
Given the homogeneous linear system

$$\begin{aligned} 4x - 18y - 7z &= 0 \\ 2x - 4py + pz &= 0 \\ px + 3y + 5z &= 0 \end{aligned}$$

where p is constant, find the values of p for which the system has non-trivial solutions, and find such a solution.

CHAPTER 4

Basic Network Laws and Theorems

4.1 Introduction

In this chapter we continue with the work we began in Chap. 2.

While the fundamental procedures of Chap. 2 are very useful they can, in some cases, become quite awkward to use. Also, there are some types of networks that cannot be separated into purely series and parallel groups of resistors, and in such cases the procedures cannot be used. The so-called "bridge networks" are of this type.

It is therefore necessary that we have available a more general method of circuit analysis than that used in Chap. 2. Fortunately such a procedure exists, resting upon what are generally called *Kirchhoff's current and voltage laws* (usually pronounced as "KIRK off"), which are the subject of this chapter.

4.2 Kirchhoff's Current Law

Kirchhoff's "current law" is based upon the fact that at any *connecting point* in a network the sum of the currents flowing *toward* the point is equal to the sum of the currents flowing *away from* the point. The law is illustrated in the examples in Figs. 42 and 43, where the arrows show the directions in which it is given that the currents are flowing. (The number alongside each arrow is the amount of current associated with that arrow.)

The example of Fig. 42 shows that if a current of 8 amperes is flowing *toward* a connecting point "p," then it is true that 8 amperes has to be flowing *away* from p. In the same way, Fig. 43 shows that if 11 amperes is flowing *toward* connection point p, then 11 amperes has to be flowing *away* from p.

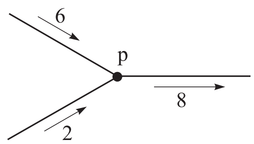


Fig. 42

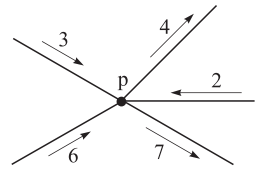


Fig. 43

Any connecting point, such as p in the above figures, is called a *node** (as in "load"), and the relationship at a node, or "nodal point," is summarized in *Kirchhoff's current law*:

> The sum of the currents flowing TO a node point equals the sum of the currents flowing FROM that point.

The currents flowing into and out of a node point are called "branch currents." Thus, in Fig. 42 the branch currents are 6, 2, and 8 amperes.

There is an important point to be made in regard to branch currents, which is explained with the aid of Figs. 44 and 45, in which the branch currents are denoted by I_1, I_2, and I_3, as follows.

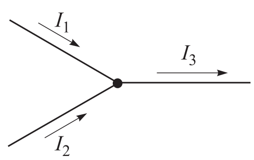


Fig. 44

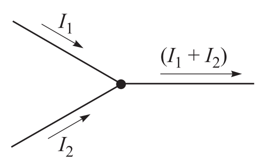


Fig. 45

In Fig. 44 note that separate notation is used to denote the value of each of the three branch currents. However, by Kirchhoff's current law, $I_3 = I_1 + I_2$, and thus, as shown in Fig. 45, we need to use only *two* current designations. In other words, if we know any two of the three currents in Fig. 44, we can then find the third current. In the same way, if there are, say, *four* branch currents entering and leaving a node point, and if we know any *three* of the currents, we can then find the fourth current, and so on.

It is important to note that Kirchhoff's current law can also be stated in terms of the "algebraic sum" of the currents at a junction or nodal point. This can be understood by referring to Fig. 44, as follows.

$$\text{In Fig. 44:} \quad I_1 + I_2 = I_3$$

$$\text{thus:} \quad I_1 + I_2 - I_3 = 0 \tag{56}$$

Equation 56 can also be arrived at by requiring that all current values at a junction point be put on one side of the current equation, and then requiring that currents flowing TO the point be listed as "positive" currents and currents flowing AWAY from the point be listed as "negative" currents. If this rule is understood to always apply, then Kirchhoff's current law can be stated in the form

> The algebraic sum of the currents at a node (junction point) is equal to zero.

* Also called a "junction" or "junction point."

4.3 Kirchhoff's Voltage Law

Note: before commencing this section you might wish to first review the meanings of the terms "active device" and "passive device" from section 2.5.

Consider, now, two points x and y. If we say that *x is at a* HIGHER VOLTAGE *than y*, we will mean that *x is* POSITIVE *with respect to y*.

Thus, going from a point y "up" to a more positive point x constitutes a RISE *in voltage*, while going from a point x "down" to a less positive point y constitutes a DROP *in voltage*.

In other words, going from *minus to plus* is a *voltage rise*, whereas going from *plus to minus* is a *voltage drop*.

We've already learned that the voltage drop across a resistor of *R ohms* carrying a current of I *amperes* is RI *volts* (eq. (13), Chap. 2), and that the polarity of the *voltage drop* across a resistor is always PLUS TO MINUS *in the direction of the current*, as in Fig. 46.

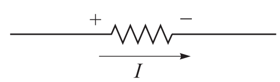


Fig. 46

Thus, if we go through a resistor *in the same direction as the current* we go from "plus to minus," which is a *voltage drop of* RI *volts*. But if we go through a resistor *against the current flow* we go from "minus to plus," which is a *voltage rise of* RI *volts*.

We likewise experience a *voltage rise* if we go through a battery from minus to plus, and a *voltage drop* if we go through a battery from plus to minus. All the foregoing facts can be summarized as follows.

> Moving through any circuit element, active or passive, *from* NEGATIVE TO POSITIVE *is a* VOLTAGE RISE, while moving *from* POSITIVE TO NEGATIVE *is a* VOLTAGE DROP. It should be remembered that the voltage across a resistance is always *positive to negative in the direction of the current*, as illustrated in Fig. 46.

Common sense tells us that if we go completely around a closed path in a circuit the *sum of the voltage rises* must equal the *sum of the voltage drops* in the path; that is, there can be no voltage "left over" in a closed path.

In network terminology any *closed path* is called a *loop*, and using this term the above fact, concerning voltage drops and rises, is summarized in *Kirchhoff's voltage law*:

> If we go in a specified direction completely around any loop (closed path) in any circuit, *the sum of the voltage drops* equals *the sum of the voltage rises* in the loop.

By "specified direction" we mean clockwise (cw) or counterclockwise (ccw). We can choose either direction but, having made a choice, we must keep that direction throughout the working of a given problem.

Let us now illustrate, with the aid of Fig. 47, how the foregoing definitions are actually applied. Note that Fig. 47 consists of two resistances and two batteries, all in series. Let

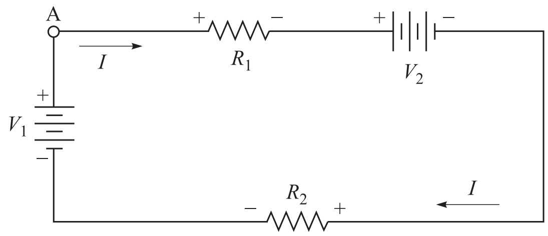


Fig. 47

us assume the current I is flowing in the cw sense around the loop, in which case the polarities (+ and −) appear across the resistances as shown.

Let us now write the equation for Fig. 47 in accordance with Kirchhoff's voltage law. To do this, we start at any point, such as A, and move completely around the circuit (we will assume in the cw sense here), listing the "voltage drops" and the "voltage rises" as we go. (In doing this, remember that we have defined that going from "minus to plus" constitutes a RISE in voltage and going from "plus to minus" constitutes a DROP in voltage.) Thus, if we agree to list all "voltage drops" on the left-hand sides of our equations and all the "voltage rises" on the right-hand sides, the Kirchhoff voltage equation for Fig. 47 is

$$R_1 I + V_2 + R_2 I = V_1$$

Note that V_2 appears as a voltage drop, because we go through that battery from plus to minus (+ to −). Or, putting all the battery voltages on the right-hand side, the above equation becomes

$$R_1 I + R_2 I = V_1 - V_2 \tag{57}$$

$$\text{hence} \quad I = \frac{V_1 - V_2}{R_1 + R_2}$$

Notice that if V_1 is greater than V_2 the current I will be *positive*, which means that the current *does* flow in the cw sense, as assumed in Fig. 47.

Note, however, that if V_2 is greater than V_1 then I will be *negative*, which means that the current I actually flows in the ccw sense, *opposite* to the direction assumed in Fig. 47. This will be true in general in all our work; thus, a *negative value of current* will mean that the current actually flows in a sense or direction *opposite* to what we assumed when we drew the current arrows in the circuit diagram.

Problem 46

In Fig. 48, the resistance values are in ohms and the battery emf's are in volts. Let the cw sense be the direction of positive current, as shown. Find I.

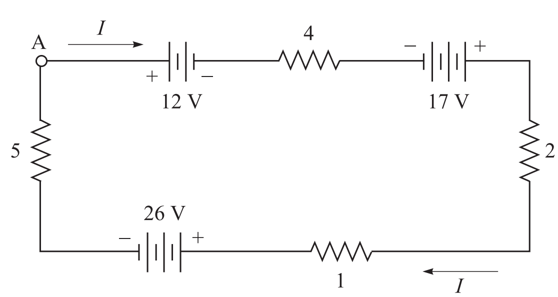


Fig. 48

4.4 The Method of Loop Currents

With all the foregoing in mind, let us now take up a widely used procedure in network analysis called the method of "loop currents." The principle is explained with the aid of Figs. 49 and 50.

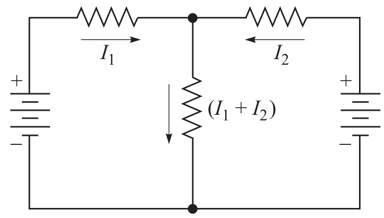


Fig. 49

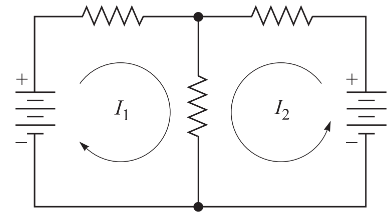


Fig. 50

In Fig. 49 we use the familiar method of showing a separate current arrow in each branch of the network. Note that the current flowing in the middle resistance is $(I_1 + I_2)$, in accordance with Figs. 44 and 45.

Notice, however, as shown in Fig. 50, that as far as circuit analysis is concerned we can assume that current I_1 *flows only around the left-hand loop*, and that current I_2 *flows only around the right-hand loop*; the current in the middle resistance is $(I_1 + I_2)$, just as in Fig. 49, and this fact is the basis of the loop-current method of network analysis.

The three steps in using the "loop current" method will now be summarized, followed by an example worked through in detail.

Step I. Draw and label the current flowing around each loop (closed path) in the network, with arrowhead indicating the direction the current is assumed to flow around the loop in each case. We must be sure not to "miss" any battery or resistor; that is, they must be all traversed by at least one loop current.

It makes no difference whether a current is assumed to flow cw or ccw, but, once labeled, the direction must not be changed during the working of the problem. If the assumed direction of current flow is opposite to the actual direction, the algebra will tell us this by producing a *negative value* for the particular current in question.

Step II. Write the voltage equation around each of the n loops of the network, thus generating *n linear equations in n unknown currents*. In writing the voltage equations around a loop, select any convenient "starting point" in the loop and go completely around the loop, returning to the starting point for that loop. In writing these equations we will, for the sake of uniformity, *always adhere to the following rules*, to put our equations in the form of eq. (57).

The voltages across resistors will be of the form "$\pm IR$," and will always be written on the *left-hand side of the equations*. They will be written "$+IR$" if we go through the resistor *in the direction of the current*, but "$-IR$" if in a direction *opposite to that of the current*.

All *battery voltages* will be written on the *right-hand sides of the equations*, and will be considered *positive* if we go through a battery *from minus to plus*, but *negative* if we go through the battery *from plus to minus*.

Step III. Solve the n equations, found in step II, for the required values of current. Since the equations are linear (of the first degree in I), it will generally be most convenient to use the method of determinants.

Example

Given: the network of Fig. 51, in which the three loop currents are all assumed to flow in the cw sense, as shown. The resistance values are in ohms. The PROBLEM is to find the potential V_a with respect to ground. (The "ground" or "earth" symbol was used before, in problem 22, Chap. 2, and has no particular significance here, but is just used as a convenient reference point for the network voltages.)

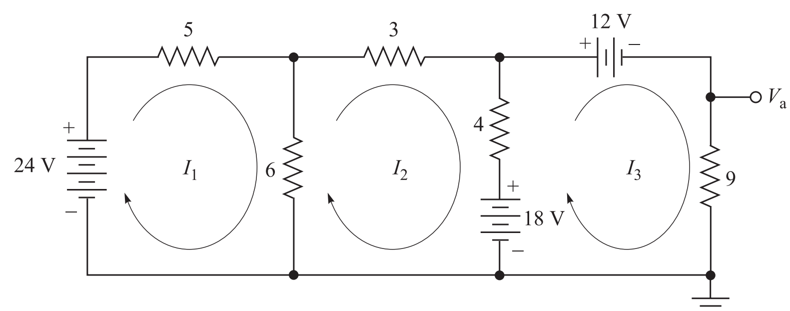


Fig. 51

Solution

Step I. In this problem, in order to use all the circuit elements, we must use three loop currents, which have been chosen as shown in the figure.

Step II. In accordance with the rules laid down, the three loop voltage equations are, from left to right in the figure (tracing around each loop in the direction indicated by the arrowhead), as follows:

$$\begin{aligned} \text{first loop:} &\quad 11I_1 - 6I_2 + 0I_3 = 24 \\ \text{second loop:} &\quad -6I_1 + 13I_2 - 4I_3 = -18 \\ \text{third loop:} &\quad 0I_1 - 4I_2 + 13I_3 = 6 \end{aligned}$$

In the above, we wrote $0I_1$ and $0I_3$ merely to keep the equations "lined up" in a convenient manner.

Step III. In this problem we are asked to find the voltage drop across the 9-ohm resistor, that is, the value of "$9I_3$"; hence, in this problem, we must *find the value of* I_3. The first step in doing this is to find the value "Δ" of the *denominator determinant*, which is the determinant *formed from the coefficients of the unknowns*; thus,

$$\Delta = \begin{vmatrix} 11 & -6 & 0 \\ -6 & 13 & -4 \\ 0 & -4 & 13 \end{vmatrix}$$

which, upon expanding in terms of the minors of the first column (and factoring the determinants where possible) is equal to

$$\Delta = 11\begin{vmatrix} 13 & -4 \\ -4 & 13 \end{vmatrix} - 12\begin{vmatrix} 3 & 0 \\ 2 & 13 \end{vmatrix} = 1215$$

To find the value of I_3 we use the same determinant as above, except now the *coefficients of* I_3 are replaced by the *constant terms on the right-hand sides* of the network equation found in step II. The value of I_3 is then equal to

$$I_3 = \frac{\begin{vmatrix} 11 & -6 & 24 \\ -6 & 13 & -18 \\ 0 & -4 & 6 \end{vmatrix}}{\Delta} = \frac{6\begin{vmatrix} 11 & -6 & 4 \\ -6 & 13 & -3 \\ 0 & -4 & 1 \end{vmatrix}}{1215} = \frac{(6)(71)}{1215} = 0.35062 \text{ amperes, approx.}$$

and therefore

$$V_a = 9I_3 = 3.15556 \text{ volts,} \quad \textit{answer}$$

Problem 47
In Fig. 52, the symbol "Ω" is the capital Greek letter omega, which is often used to denote "ohms." Using the method of loop currents, find the current in the 7-ohm resistance. (*Answer*: 1.24138 amperes)

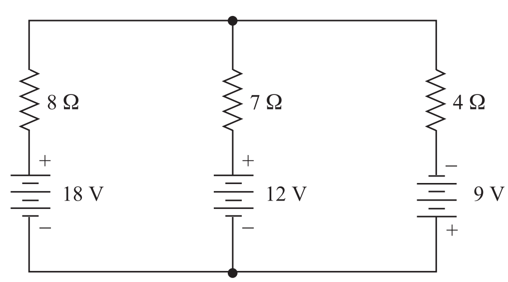


Fig. 52

Problem 48
In the "bridge-type" network in Fig. 53, the resistance values are in ohms. Find the voltage drop across the 4-ohm resistance. (*Answer*: 1.8462 volts)

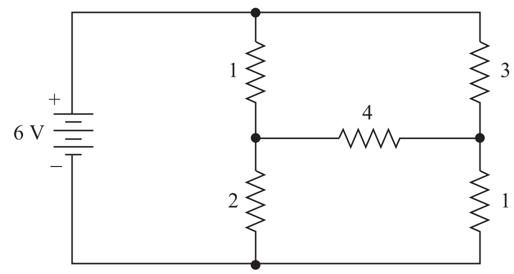


Fig. 53

Problem 49
In Fig. 54, find the potential at point "a" with respect to ground. Resistance values are in ohms. (*Answer*: 8.98294 volts)

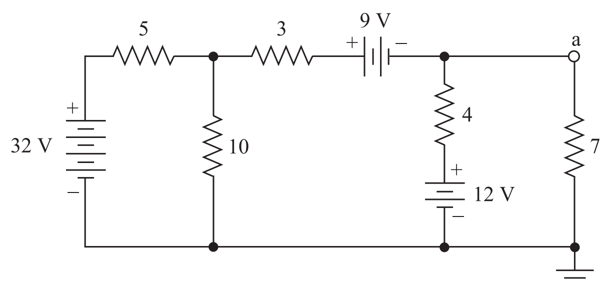


Fig. 54

Problem 50
This problem is included here as an example of the very important PRINCIPLE OF SUPERPOSITION, which we first met in section 1.3 (eq. (7)). As applied to electric networks, the principle of superposition can be stated as follows.

In a network composed of linear* elements and several generators, the current at any point in the network is the *sum of the currents due to* EACH GENERATOR CONSIDERED SEPARATELY, the other generators being replaced by their internal resistances. In our applications here, we'll assume the generators (batteries) to have zero internal resistance.

In problem 47, Fig. 52, we found the current in the 7-ohm resistor to be, to five decimal places, equal to 1.24138 amperes. Verify this answer by applying the principle of superposition to Fig. 52.

Problem 51
Explain why all actual resistances are "non-linear" to some degree.

Problem 52
In Fig. 53, suppose it is desired to replace the 3-ohm resistor with a resistor of R ohms, such that the current through the 4-ohm resistor will be zero. What must be the value of R? (*Answer*: $R = 0.5$ ohm)

* A "linear" resistance is one whose resistance value is independent of the amount of current flowing in the resistance. (A resistance is said to be "non-linear" if the resistance changes with changes in current.)

4.5 Conductance. Millman's Theorem

In certain types of network it is more convenient to work with the RECIPROCAL *of resistance* instead of directly with resistance itself. The reciprocal of resistance is called *conductance*, and is denoted by "G." Hence, by definition,

$$G = \frac{1}{R} = 1/R \tag{58}$$

Thus "conductance" is measured in units of "reciprocal ohms," called *mhos* ("mho" is "ohm" spelled backward). The term *siemen* is also used for reciprocal ohms and is the SI unit of electrical conductance. As eq. (58) shows, HIGH RESISTANCE means LOW CONDUCTANCE, and vice versa; for example,

$$\text{if} \quad R = 1000 \text{ ohms then } G = 1/1000 = 0.001 \text{ mho,}$$

$$\text{or if} \quad R = 0.001 \text{ ohm then } G = 1/0.001 = 1000 \text{ mhos.}$$

Thus the basic Ohm's law, $I = V/R$, becomes, in terms of conductance,

$$I = GV \tag{59}$$

showing that "amperes equals mhos times volts."

Next consider POWER, P. In section 2.3 we found that $P = V^2/R$, and thus

$$P = GV^2 \text{ watts} \tag{60}$$

Conductance is especially convenient to use when dealing with purely *parallel* networks, as the following will show.

In section 2.6 (eq. (32)), we found that the total resistance R_T of n *parallel-connected resistances* is found by means of the relationship

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots + \frac{1}{R_n}$$

which is not especially easy to use. Note, however, that by the definition of eq. (58) we have that the *total conductance* of n parallel-connected resistances is equal to the simple *sum* of the individual conductances; thus

$$G_T = G_1 + G_2 + G_3 + \cdots + G_n \tag{61}$$

The use of conductance is especially helpful in dealing with parallel networks in which each branch is composed of a conductance in series with a battery, as in Fig. 55.

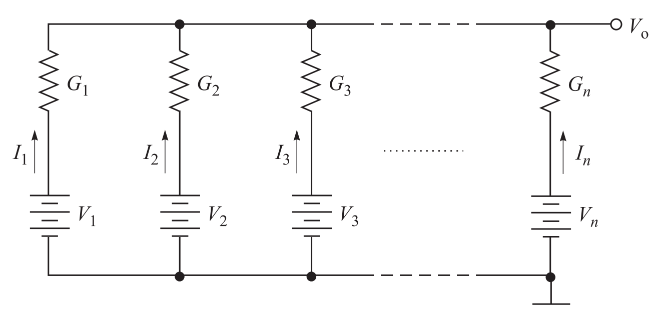


Fig. 55

In Fig. 55 let the values of the conductances and battery voltages be given, and let the object be to find a formula for calculating the output voltage V_o with respect to the "ground line." This can be done as follows.

First, it should be understood that no current flows "out of the circuit" at the point labeled V_o; that is, the voltage V_o looks, to the right, into an "open circuit." (An open circuit has "infinitely great resistance," that is, "zero conductance.")

Next, note that the upper horizontal line (which is at the potential V_o with respect to ground) is a "node" or "junction point" into which *all* the currents, I_1, I_2, and so on, flow. Hence, to satisfy Kirchhoff's current law, the ALGEBRAIC SUM of all these currents has to be *zero*; thus

$$I_1 + I_2 + I_3 + \cdots + I_n = 0 \tag{62}$$

Note that this can be true in general only if *both* negative and positive currents exist; that is, some of the currents in Fig. 55 will have to flow "downward" instead of "upward."

Now consider any one of the n parallel branches in Fig. 55; let us take, as an example, branch number 1 (that is, the branch to the far left side in Fig. 55). Now denote by V_{G1} the *voltage drop* across the conductance G_1. Since V_o is equal to the *battery voltage minus the voltage drop across* G_1, we have that

$$V_1 - V_{G1} = V_o$$

thus,

$$V_{G1} = V_1 - V_o$$

then multiplying through by G_1 gives:

$$G_1 V_{G1} = G_1 V_1 - G_1 V_o$$

But, by eq. (59), $G_1 V_{G1} = I_1 =$ the current in branch 1, and thus the last equation becomes

$$I_1 = G_1 V_1 - G_1 V_o$$

next, in the same way,

$$I_2 = G_2 V_2 - G_2 V_o$$

and so on

$$\vdots \qquad \vdots$$

$$I_n = G_n V_n - G_n V_o$$

Now note that, by eq. (62), the sum of all the right-hand sides is equal to *zero*; that is,

$$(G_1 V_1 + G_2 V_2 + \cdots + G_n V_n) - (G_1 + G_2 + \cdots + G_n) V_o = 0$$

and thus we have the desired result,

$$V_o = \frac{G_1 V_1 + G_2 V_2 + \cdots + G_n V_n}{G_1 + G_2 + \cdots + G_n} \tag{63}$$

Equation (63), which applies to the particular network of Fig. 55, is called *Millman's Theorem*, and is helpful in, for example, the study of transistor amplifier circuits and in other applications also. (If some of the branches do not contain batteries, the battery voltages for those branches are entered as "zero" in eq. (63).)

Problem 53

In Fig. 56, the circuit values are given in ohms and volts. Find V_o with respect to ground.

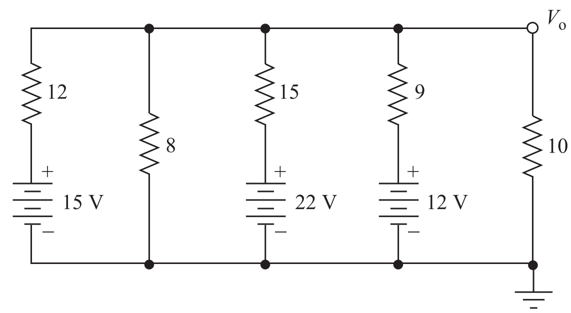


Fig. 56

Problem 54

In Fig. 56, suppose the 12-volt battery were to be reversed in polarity (turned "upside down"), everything else remaining unchanged. Find V_o.

(*Answer*: 2.84534 V)

4.6 Thevenin's Theorem

All practical sources of electrical energy (generators) have *internal resistance*, as was pointed out in section 2.5. The general case is illustrated in Fig. 57.

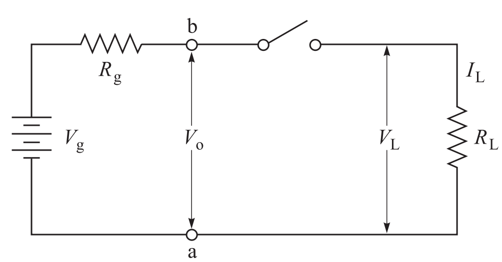


Fig. 57

In the figure, R_g denotes the *internal resistance* of the generator. We are here representing the generator by the battery symbol, but the generator can, of course, be any kind of source of direct-current (dc) voltage.

Note that R_L is an *external load resistance* that can be connected to the generator terminals, a and b, by closing the switch.

Also in the figure, V_g is the *total generated voltage* developed by the generator, while V_o is the voltage appearing at the generator terminals ab and V_L is the voltage across the load resistance. I_L is the current that flows when the switch is closed.

If the switch is open, then $V_o = V_g$ (because no current flows with the switch open, and thus there is no voltage drop across R_g), and of course $V_L = 0$ with the switch open.

This points out the fact that to *find the total generated voltage*, V_g, we must measure the voltage at the generator terminals *under open-circuited conditions*, that is, with the load R_L

disconnected from the generator. Also, to find R_g, the *internal resistance* of the generator, we must likewise measure the resistance between terminals ab with the *load disconnected*, that is, under "open-circuited" conditions, with all batteries (generators) being replaced with their internal resistances.

When the switch is closed, current I_L flows, producing a voltage drop across R_g, the internal resistance of the generator, thus causing the terminal voltage V_o to change in value. ($V_o = V_L$ when the switch is closed.)

With the above in mind, let "ab" be two terminals coming out of any network composed of generators and resistances, as indicated by the box in Fig. 58.

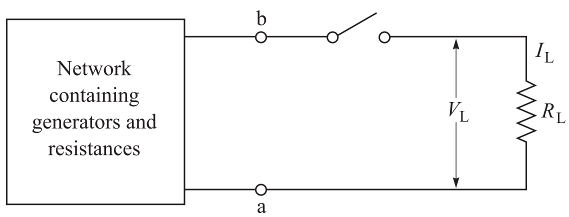


Fig. 58

In regard to Fig. 58, THEVENIN'S THEOREM (generally pronounced "THEV eh nin" in the United States) says that, as far as the voltage and current in any *external load resistance*, R_L, is concerned:

> The entire network, inside the box, can be replaced by a *single generator* whose generated voltage is equal to the open-circuit voltage appearing between a and b, and whose internal resistance is equal to the resistance seen looking back into the open-circuited terminals, with all generators removed and replaced with resistances equal to their internal resistances.

Now compare Fig. 58 with Fig. 57; Thevenin's theorem says that, whatever the *actual* network may be inside the box in Fig. 58, it can be replaced, as far as *external* results are concerned, by the equivalent single generator of Fig. 57, in which the values of V_g and R_g are determined by making two open-circuit measurements at terminals a and b, that is, with R_L disconnected.

Any such equivalent generator (Fig. 57) thus consists of a CONSTANT-VOLTAGE generator (generating a constant V_g volts) in series with a resistance of R_g ohms.

Thevenin's theorem is especially helpful when we desire to investigate changes in V_L and I_L when only the load resistance R_L is changed. This can be illustrated as follows.

Problem 55

In Fig. 56, suppose it is proposed to connect a load resistance of $R_L = 2$ ohms in parallel with the 10-ohm resistance. By making use of Thevenin's theorem,* find the current that would flow in the 2-ohm resistance. Repeat for $R_L = 3$ ohms and $R_L = 4$ ohms.

* In problems 55 and 56, assume the batteries to have zero internal resistance.

Problem 56
In Fig. 59, the resistance values are in ohms. Replace the network to the left of terminals a, b with the equivalent Thevenin generator.

(*Answer*: $V_g = 32.143$ V, $R_g = 4.179$ ohms)

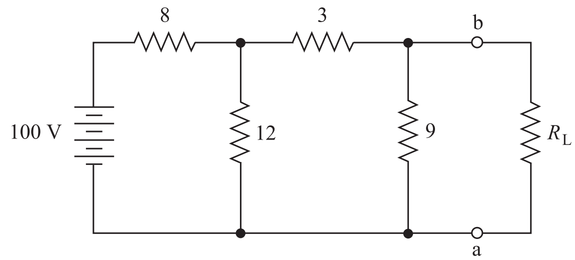


Fig. 59

4.7 Norton's Theorem

Another widely used network theorem, called NORTON'S THEOREM, makes use of a theoretical, but very useful, device called a CONSTANT-CURRENT GENERATOR. As the name says, a "constant-current generator" is a theoretical generator that delivers the SAME CONSTANT CURRENT TO ALL FINITE LOAD RESISTANCES *it is connected to*.

To understand what such a theoretical generator would have to be like, suppose, just for the sake of discussion, that, back in Fig. 57, the generated voltage of the generator was *100 billion volts* and that R_g, the internal resistance of the generator, was, let us say, *10 billion ohms*. Then the generator would deliver an output current I_L equal to

$$I_L = \frac{V_g}{R_g + R_L} = \frac{10^{11}}{10^{10} + R_L} \quad \text{amperes} \tag{64}$$

Now suppose the generator terminals a, b are "shorted together" (making $R_L = 0$). For this "short-circuit" condition of $R_L = 0$, the generator delivers a current of

$$I_L = \frac{10^{11}}{10^{10}} = 10 \text{ amperes}$$

Now suppose the short-circuit is removed, and a load resistance of *ten-thousand ohms* is connected between terminals a, b. Then the current delivered by the generator will be, by eq. (64),

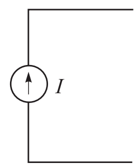
$$I_L = \frac{10^{11}}{10^{10} + 10^4} = \frac{10^{11}}{10^4(10^6 + 1)} = \frac{10{,}000{,}000}{1{,}000{,}001} = 9.99999 \text{ amperes, approx.}$$

Now suppose the load resistance is increased to say *one million ohms*. Then the current delivered by the generator is

$$I_L = \frac{10^{11}}{10^{10} + 10^6} = \frac{10^{11}}{10^6(10^4 + 1)} = \frac{100{,}000}{10{,}001} = 9.99900 \text{ amperes, approx.}$$

Notice that our generator delivers an *almost constant current of 10 amperes*, regardless of whether it works into a load of zero ohms or one million ohms. We thus see that a *true constant-current generator* is a theoretical device having *infinitely great generated voltage* but *infinitely great internal resistance*, the ratio of the two being equal to a finite constant current.

The symbol for a constant-current generator is shown below, where "I" is the value of the constant current, the arrow designating the direction of "positive current."



Let us now return to the two-terminal network inside the box of Fig. 58, to which a load resistance R_L can be connected, and take up the details of Norton's theorem.

Norton's theorem is expressed in terms of the short-circuit current delivered by the network, and in terms of conductances instead of resistances. This makes Norton's theorem especially useful in the study of parallel circuits. The statement of Norton's theorem is as follows, after which we'll give the proof of the statement.

> The current in any load conductance G_L, when connected to two terminals of a network, is the same as if G_L were connected to a constant-current generator whose constant current is equal to the current that flows between the two terminals when they are *short-circuited together*, this constant-current generator then being put in parallel with a conductance equal to the conductance seen looking back into the open-circuited terminals of the network. (In this last step, all generators are removed and replaced with conductances equal to their internal conductances.)

Norton's theorem is summarized graphically in Fig. 60, where I_{sc} is the short-circuit current that flows from the network when terminals a, b are "shorted" together. G_g is the conductance seen looking back into the network with the terminals open-circuited, that is, with the switch open. G_g is the reciprocal of R_g in Thevenin's theorem.

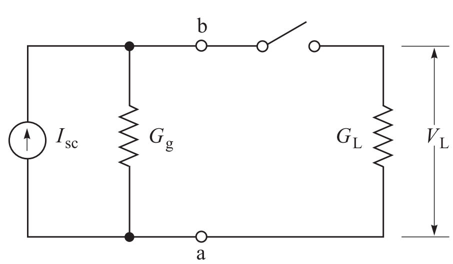


Fig. 60

The truth of Norton's theorem can be shown as follows. Let any two-terminal network be inside the box of Fig. 58. We know that, as far as the external load resistance R_L is concerned, the network inside the box can be replaced with the Thevenin equivalent generator of Fig. 57, where, by inspection of Fig. 57, we have

$$I_L = \frac{V_g}{R_g + R_L} \tag{65}$$

Now put a short-circuit (a copper wire) between terminals a, b in Fig. 57. The short-circuit current flowing between the terminals is then (since $R_L = 0$ for this condition)

$$I_{sc} = \frac{V_g}{R_g}$$

Now put this value of V_g, $V_g = I_{sc}R_g$, into eq. (65), then multiply both sides by R_L. Since $R_L I_L$ = the voltage across the load = V_L, we get

$$V_L = I_{sc}\left(\frac{R_g R_L}{R_g + R_L}\right)$$

Now multiply the numerator and denominator of the last fraction by $1/R_g R_L$. Then, by the definition of conductance (eq. (58)), the last equation becomes

$$V_L = \frac{I_{sc}}{G_g + G_L} \tag{66}$$

We now know that eq. (66) is the *correct equation* for the voltage V_L appearing across the load.

With this in mind, turn now to the proposed equivalent circuit of Fig. 60. Remembering that conductances in parallel add together like resistances in series (eq. (61)), and also remembering the basic relation, $I = GV$ (eq. (59)), we have, for Fig. 60, $I_{sc} = (G_g + G_L)V_L$, so that

$$V_L = \frac{I_{sc}}{G_g + G_L}$$

Since the last equation is the same as eq. (66), it follows that Figs. 57 and 60 produce *completely equal results* as far as any external load is concerned, and therefore either can be used.

TO SUMMARIZE:

Any two-terminal network consisting of generators and linear* bilateral* resistances can be replaced *as far as an external load connected to the two terminals is concerned* by either a Thevenin generator (Fig. 57) or a Norton generator (Fig. 60). If the external load consists of multiple parallel branches, it will generally be more convenient to use the Norton generator.

Problem 57

In Fig. 60, let I_L denote the value of the current that would flow in the load conductance G_L if the switch were closed. Now, by making use of the basic relationship $I = GV$ (eq. (59)), show that

$$I_L = \frac{G_L I_{sc}}{G_g + G_L} \tag{67}$$

* Recall that a resistance is *linear* if its value is independent of the amount of voltage applied to it or the amount of current flowing in it. It is *bilateral* if current is able to flow through it equally well in both directions.

Problem 58

(a) Going back to Fig. 59 in section 4.6, replace the network to the left of terminals a, b with the equivalent Norton generator.

(b) In Fig. 59, suppose $R_L = 10$ ohms. Show that the Thevenin and Norton generators produce the same external load current, $I_L = 2.267$ amperes, to three decimal places in each case.

Problem 59
In Fig. 60, if $I_{sc} = 6.155$ amperes and $G_g = 0.109$ mho, draw the Thevenin equivalent generator.

4.8 The Method of Node Voltages

The method of "node voltages" is a procedure for network analysis based upon Kirchhoff's current law (section 4.2). The procedure is as follows.

We begin by selecting one of the nodes in the network to be the *reference node*. The unknown voltages, at the other nodes in the network, are to be found *relative to the "zero voltage" at the reference node*.

Thus, in the "node method" of network analysis the NODE VOLTAGES ARE THE UNKNOWNS, instead of the currents as in the loop method. If a network has N nodes, then $N - 1$ node voltages will be present (because one of the N nodes is selected to be the "reference node," which is then taken to be at "zero reference potential").

We begin our discussion with Fig. 61, which shows a resistance of R ohms connected between two node points a and b. Let us assume a current of I amperes flowing through R *from node a to node b*, as shown.

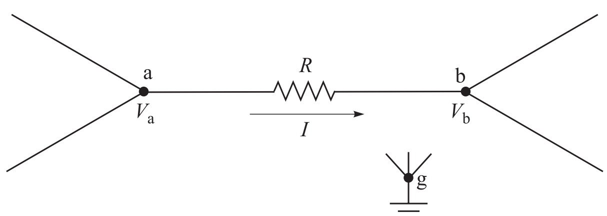


Fig. 61

In Fig. 61, g is the reference node. It will always be understood, unless definitely stated otherwise, that all node voltages in a network are given *relative to the reference node*, which is taken as being at *zero voltage*, that is, $V_g = 0$. Thus, in Fig. 61, the node voltages V_a and V_b are measured with respect to the reference node g.

In practical work it is often necessary, or at least desirable, to connect one side of a circuit to "earth" or "ground" by connection, for example, to an underground water pipe. This may be done to insure the safety of personnel, or to minimize noise pickup, and so on. Often, however, we'll use the ground symbol as a convenient symbol to designate the reference node, even though it may not actually be connected to an earth ground.

In Fig. 61 we are assuming that current is flowing from left to right, from node a to node b, as shown. This indicates that *node a is* POSITIVE *with respect to node b*, because conventional current* flows from a point of higher potential to a point of lower potential, that is, from positive to negative.

For example, if $V_a = +100$ volts and $V_b = +70$ volts (both measured with respect to g), then *node a* would be *30 volts positive with respect to node b*, so that current would flow from *node a to node b*, as in Fig. 61.

Since, by Ohm's law, current equals voltage divided by resistance ($I = V/R$), the situation in Fig. 61 is stated algebraically by writing

$$I = \frac{V_a - V_b}{R} \tag{68}$$

In applying the method of node voltages to a network we first designate the "reference node," which we'll generally do by use of the "ground" symbol as mentioned above. We then label the unknown voltages, at the different nodes, as V_a, V_b, V_c, and so on, all voltages being with respect to the "zero voltage" at the reference node. We next draw and label the "current arrows" I_1, I_2, I_3, and so on, at each node (see Fig. 44, section 4.2), and then write the current equation at each node (see eq. (56), section 4.2). Now, in each current equation, replace each current by its equivalent in terms of eq. (68); doing this gives us the required equations in terms of the unknown node voltages. If the battery voltages and resistance values are known, the resulting linear simultaneous equations can then be solved to find the node voltages.

Example

In Fig. 62, resistance values are in ohms. Using the node-voltage procedure find, to the third decimal place, V_b and V_c. Use the current arrows as given in the figure, or redraw them any way you wish.

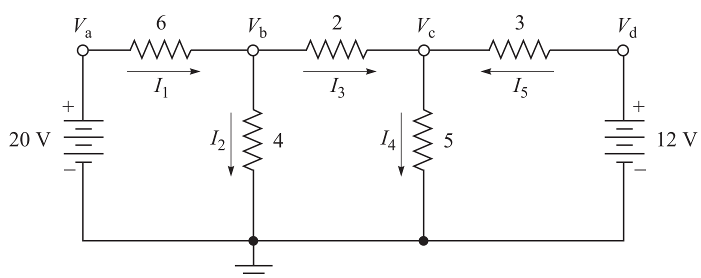


Fig. 62

Solution

First, for the node at V_b:

$$I_1 - I_2 - I_3 = 0 \tag{A}$$

and, for the node at V_c:

$$I_3 - I_4 + I_5 = 0 \tag{B}$$

* "Conventional current" is the flow of "positive charge" (section 2.1), which creates a "+ to −" voltage drop across a resistance in the direction of the current flow.

Next, by inspection we see that $V_a = 20$ volts and $V_d = 12$ volts; thus, applying eq. (68), we have

$$I_1 = \frac{20 - V_b}{6} \qquad I_3 = \frac{V_b - V_c}{2} \qquad I_5 = \frac{12 - V_c}{3}$$

$$I_2 = V_b/4 \qquad I_4 = V_c/5$$

Now, upon substituting these values into equations (A) and (B), you should find that

$$11V_b - 6V_c = 40$$
$$-15V_b + 31V_c = 120$$

and thus the *answers* are

$$V_b = 7.809 \text{ V} \quad \text{and} \quad V_c = 7.649 \text{ V}$$

Problem 60

In Fig. 63, use the node-voltage procedure to find the voltages at nodes 1 and 2. Resistance values are in ohms.

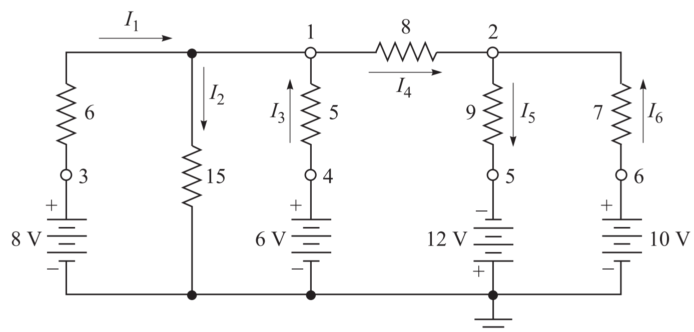


Fig. 63

CHAPTER 5

Sinusoidal Waves. rms Value. As Vector Quantities

5.1 Introduction

So far in our work we've dealt only with currents and voltages whose *senses of direction and polarity never change*. Such currents and voltages are called DIRECT *currents* and DIRECT *voltages*.

Thus, since the polarity of a battery does not change, we say that a battery is a source of "direct" voltage, producing a flow of "direct" current. The abbreviation "dc" is commonly used to identify such quantities; thus we have "dc voltage," "dc current," "dc power," and so on.

There is, however, another class of voltages and currents whose senses of both direction and polarity *continually* ALTERNATE *with time*, plus to minus, minus to plus, and so on, endlessly. The term "alternating" is used to denote such a voltage or current, and the abbreviation "ac" is used to denote such quantities. Thus we have "ac voltage," "ac current," "ac power," and so on.

Offhand, a person might expect that the algebra of "ac circuits" would be considerably different, and more difficult, than the algebra of "dc circuits." It is, of course, true that ac calculations can differ greatly from dc calculations. But such differences can be concisely expressed mathematically, and, when this is done, we'll find that the form of the algebraic statements (Kirchhoff's laws, loop currents, and so on) that we learned in dc work will carry over directly into our ac work. Let us now begin the study of this most interesting and useful subject.

5.2 The Sinusoidal Functions and the Tangent Function

Fundamental to the study of "alternating currents" are the SINUSOIDAL functions, the term "sinusoidal" ("sign u SOID al") denoting either the SINE *function* or the COSINE *function*. The "sine" and "cosine" are simple but remarkable functions, having properties unlike any others in the entire realm of mathematics.

We'll find that the two functions (sine and cosine) can be said to have identical "waveforms," the only difference being that the two are "shifted" with respect to each other on the horizontal or time axis. As a matter of fact, a person viewing a "sinusoidal" waveform on an oscilloscope, and having no other information, could not say whether the waveform represented a "sine" or a "cosine" function. Let us now proceed with some definitions.

We first meet the sine and cosine in the study of plane trigonometry, in connection with the geometry of a RIGHT TRIANGLE, using the terminology of Fig. 64.

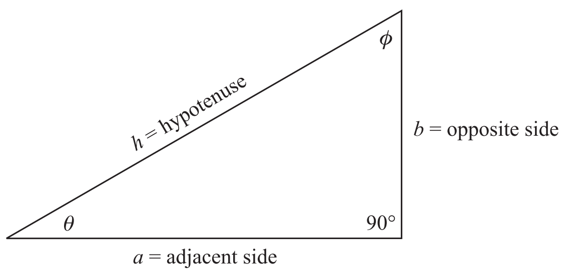


Fig. 64

In the triangle, the side h, opposite the 90° angle, is called the hypotenuse, as usual in a right triangle. We'll call the angle θ (theta) the "reference angle"; then the side opposite θ is called the OPPOSITE side and the side adjacent to θ is called the ADJACENT side, as shown. Remember that we always have a RIGHT triangle, so that one angle always remains fixed at 90°. Note that we've denoted the third angle by ϕ (phi or "fee"). Since the three angles of a plane triangle must add up to 180 degrees, it follows that $\phi = 90 - \theta$ degrees.

An important fact concerning Fig. 64 (or any triangle, for that matter) can be understood as follows. Imagine that we looked at Fig. 64 through a magnifying glass having a magnifying power of "k" times. Doing this would *change the apparent* SIZE of the triangle (each of the three sides would be multiplied by k), but it would *not change the* SHAPE of the triangle in any way; that is, it would *not change the* ANGLES in any way. This will serve to illustrate the important fact that

> In Fig. 64 the RATIO of any one side to either of the other two sides depends *not upon the* SIZE *of the triangle but only upon the* ANGLE θ. (The angle θ is automatically known if the reference angle θ is given, because $\phi = 90 - \theta$.)

In the study of alternating currents we deal mainly with *three different ratios* of the sides of a right triangle. In terms of an angle θ, the three ratios are called the SINE ("sign") of the angle θ, the COSINE ("KOH sign") of the angle θ, and the TANGENT of the angle θ. Then,

The expression *sine of* θ is abbreviated *sin* θ and read as "sine of theta,"

The expression *cosine of* θ is abbreviated *cos* θ and read as "cosine of theta,"

The expression *tangent of* θ is abbreviated *tan* θ and read as "tangent of theta."

Referring now to the "standard reference triangle" of Fig. 64, the above expressions are defined to mean that (and these definitions should be committed to memory)

$$\sin\theta = \frac{\text{opposite side}}{\text{hypotenuse}} = \frac{b}{h}$$

$$\cos\theta = \frac{\text{adjacent side}}{\text{hypotenuse}} = \frac{a}{h}$$

$$\tan\theta = \frac{\text{opposite side}}{\text{adjacent side}} = \frac{b}{a}$$

that is, in Fig. 64:

$$\boxed{\sin\theta = b/h} \qquad (69)$$

$$\boxed{\cos\theta = a/h} \qquad (70)$$

$$\boxed{\tan\theta = b/a} \qquad (71)$$

The quantities $\sin\theta$, $\cos\theta$, and $\tan\theta$ are the three principal "trigonometric functions." The values of the functions depend only upon the *angle* θ. In our work the angle θ will generally be regarded as the "independent variable." As mentioned previously, "$\sin\theta$" and "$\cos\theta$" are referred to as the *sinusoidal* functions.

Now let us see how the values of $\sin\theta$, $\cos\theta$, and $\tan\theta$ vary *as* θ *varies from* $\theta = 0^\circ$ *to* $\theta = 90^\circ$. To help us to do this, we've used our calculator to fill out a short four-place "table of values" as follows (which you can verify on your own calculator).

θ°	$\sin\theta$	$\cos\theta$	$\tan\theta$	θ°	$\sin\theta$	$\cos\theta$	$\tan\theta$
0	0.0000	1.0000	0.0000	35	0.5736	0.8192	0.7002
1	0.0175	0.9999	0.0175	40	0.6428	0.7660	0.8391
2	0.0349	0.9994	0.0349	45	0.7071	0.7071	1.0000
3	0.0523	0.9986	0.0524	50	0.7660	0.6428	1.1918
4	0.0698	0.9976	0.0699	55	0.8192	0.5736	1.4281
5	0.0872	0.9962	0.0875	60	0.8660	0.5000	1.7321
6	0.1045	0.9945	0.1051	65	0.9063	0.4226	2.1445
8	0.1392	0.9903	0.1405	70	0.9397	0.3420	2.7475
10	0.1737	0.9848	0.1763	75	0.9659	0.2588	3.7321
15	0.2588	0.9659	0.2680	80	0.9848	0.1737	5.6713
20	0.3420	0.9397	0.3640	85	0.9962	0.0872	11.4301
25	0.4226	0.9063	0.4663	88	0.9994	0.0349	28.6363
30	0.5000	0.8660	0.5774	90	1.0000	0.0000	"∞"

A discussion of the table follows, in which it will be helpful to refer to Fig. 65.

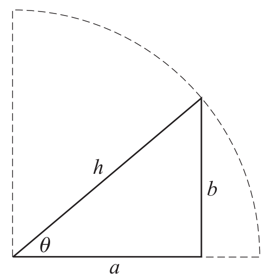


Fig. 65. Here we are holding *h* constant in length, as the angle θ is allowed to have any value from 0 degrees to 90 degrees. As the angle θ changes, the length of the sides *a* and *b* change, but *h* remains constant in length.

Let us begin our discussion of the foregoing table for the special cases of $\theta = 0^\circ$ and $\theta = 90^\circ$, by making use of eqs. (69), (70), and (71), and Fig. 65, as follows.

First, for $\theta = 0^\circ$ we have

$$\left.\begin{array}{lll} \sin\theta = b/h = 0/h = 0, & \text{that is,} & \sin 0^\circ = 0 \\ \cos\theta = a/h = h/h = 1, & \text{that is,} & \cos 0^\circ = 1 \\ \tan\theta = b/a = 0/h = 0, & \text{that is,} & \tan 0^\circ = 0 \end{array}\right\} \text{see table}$$

Next, for $\theta = 90^\circ$ we have

$$\left.\begin{array}{lll} \sin\theta = b/h = h/h = 1, & \text{that is,} & \sin 90^\circ = 1 \\ \cos\theta = a/h = 0/h = 0, & \text{that is,} & \cos 90^\circ = 0 \\ \tan\theta = b/a = h/0 = \infty, & \text{that is,} & \tan 90^\circ = \infty \end{array}\right\} \text{see table}$$

Let us now discuss, in more detail, the statement that "$\tan 90^\circ = \infty$." To begin, note that, as Fig. 65 shows, as θ comes CLOSER AND CLOSER to the value of 90 degrees the ratio b/a becomes GREATER AND GREATER in value. Mathematically we can say that, as θ becomes "vanishingly close" to the limiting value of 90 degrees, the value of the ration b/a "increases without bound," that is, becomes INFINITELY GREAT. The mathematical expression to indicate this situation is written as

$$\lim_{\theta \to 90^\circ} [\tan\theta] \to \infty \tag{72}$$

which can be read as "the tangent of θ becomes infinitely great as θ approaches the limiting value of 90 degrees."

It should be noted that "infinity" is not a specific number, but is *greater than any number you name*, however large. Since infinity is not a specific number, we say that $\tan\theta$ is *undefined* for $\theta = 90^\circ$. It is thus not correct to say that $\tan 90^\circ$ "equals" infinity, because infinity is not a specific value; in this case eq. (72) is really the proper statement to use. Nevertheless, it is common practice to abbreviate eq. (72) by simply writing that "$\tan 90^\circ = \infty$."

At this point we might mention, just briefly, how the values of $\sin\theta$, $\cos\theta$, and $\tan\theta$, listed in the foregoing "table of values," were originally found.

Originally, such tables were created by drawing, as carefully as possible, right triangles for different values of θ. Then, upon measuring the lengths of the sides as accurately as possible, the true values of the ratios b/h, a/h, b/a could be approximately determined for specific values of θ.

Later on, in the early part of the 18th century, following the invention of the calculus, the values could be directly calculated to as many decimal places as desired by the use of infinite series.

Problem 61
Using only the fact that an ‘‘equilateral’’ triangle has three equal sides and three equal angles, find the values of $\sin 60^\circ$, $\cos 60^\circ$, $\sin 30^\circ$, and $\cos 30^\circ$.

Problem 62
From inspection of the right triangle of Fig. 64, write the relationship between $\sin\theta$ and $\cos\theta$ in terms of θ only.

Problem 63
In angular measurement, ‘‘one degree’’ is equal to ‘‘60 minutes’’ (written $1^\circ = 60'$). If it is given that $\sin 62^\circ 38' = 0.88808$, the cosine of what angle is also equal to the same value, 0.88808?

Problem 64
As used in mathematics, ‘‘identity’’ denotes a relationship that is true for *all values* of an unknown quantity, while ‘‘equation’’ denotes a relationship that is true for only a *limited number* of values of the unknown. For example, $5(x-2) = 5x - 10$ is an ‘‘identity’’ because it is true for all values of x, but $5(x-2) = 0$ is an ‘‘equation’’ because it is true only for the value $x = 2$. Now, making use of the right triangle of Fig. 65 and eqs. (69), (70), (71), show that the following are valid ‘‘trigonometric identities’’:

$$\text{(a)*}\quad \sin^2\theta + \cos^2\theta = 1 \qquad\qquad \text{(b)}\quad \frac{\sin\theta}{\cos\theta} = \tan\theta$$

5.3 Graphics. Extension beyond 90 Degrees, Positive and Negative

If the values of a function are plotted as points on graph paper, and the points connected together, a curve called the ‘‘graphical representation’’ of the function is produced. In this section we’ll draw and investigate the graphical representations (that is, the ‘‘curves’’) of the three functions, $\sin\theta$, $\cos\theta$, and $\tan\theta$. Let us thus now plot, using rectangular coordinates, the three equations

$$y = \sin\theta \qquad y = \cos\theta \qquad y = \tan\theta$$

putting the dependent variable y on the vertical axis and the independent variable (the angle θ in degrees) on the horizontal axis. We’ll find that we can conveniently show the curves of $\sin\theta$ and $\cos\theta$ in the same figure, while using a separate figure to show the curve

* $\sin^2\theta = (\sin\theta)^2$, which is read as the ‘‘sine squared of theta’’ or as the ‘‘square of the sine of theta.’’ It should be carefully noted that $(\sin\theta)^2 = \sin^2\theta$ has an entirely different meaning from $\sin\theta^2$, which is the ‘‘sine of the quantity theta squared.’’ Also, if we mean to write $\sin^2\theta$, we must be careful not to drop the exponent down too far and write ‘‘$\sin 2\theta$,’’ because this would mean the ‘‘sine of two times theta.’’

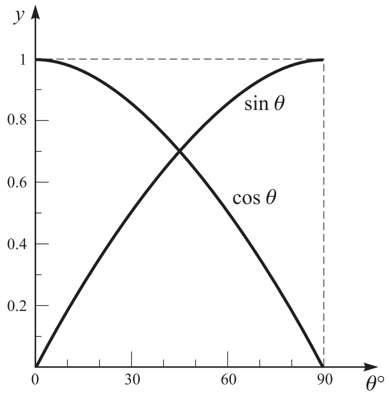


Fig. 66

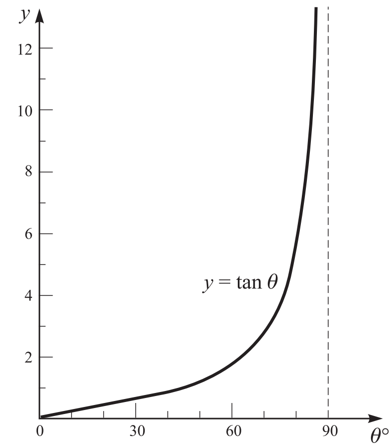


Fig. 67

of $\tan\theta$. Doing this, upon making use of the "table of values" given in section 5.2, produces Figs. 66 and 67.

In regard to Fig. 66 first, notice that "$\sin\theta$" has the value *zero* for $\theta = 0$, rising to the *maximum value of 1* for $\theta = 90^\circ$. Then note that "$\cos\theta$" is just the opposite, having the maximum value of 1 at *zero* degrees, decreasing to the value *zero* for $\theta = 90^\circ$.

Next, Fig. 67 depicts, graphically, the behavior of "$\tan\theta$" as the angle θ increases from $\theta = 0^\circ$ to $\theta = 90^\circ$. Note that $\tan\theta$ has the value *zero* for $\theta = 0^\circ$, then relatively slowly increases to the value 1 for $\theta = 45^\circ$, thereafter increasing in value faster and faster, finally becoming "infinitely great" in value as θ approaches 90 degrees (eq. (72)). The function $\tan\theta$ is simply "not defined" for $\theta = 90^\circ$, as was discussed in section 5.2.

So far we've defined the functions $\sin\theta$, $\cos\theta$, and $\tan\theta$ only for angles having values from $\theta = 0^\circ$ to $\theta = 90^\circ$. Actually, however, the functions are defined *for* ALL POSITIVE AND NEGATIVE VALUES *of the angle* θ, from "minus infinity" to "plus infinity." This is done as follows.

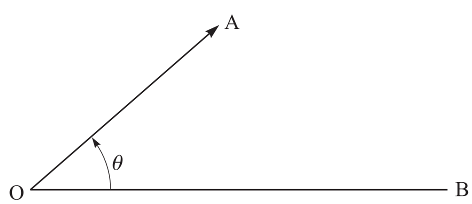


Fig. 68

In Fig. 68, the line OB ("oh, bee") is held *fixed* in position, and is called the *reference line*, or sometimes the "initial line." Next, we think of the line OA ("oh, aye") as *free to rotate around the point O*, thus generating the angle theta, θ. The line OA is sometimes called the "generating line" or, more often, the *radius vector*.

The generating line or "radius vector," OA, is free to revolve in either the *clockwise* (cw) or the *counterclockwise* (ccw) sense. It has, by general agreement, been decided to call the COUNTERCLOCKWISE *sense the* POSITIVE *sense of rotation*. Therefore the opposite sense, the CLOCKWISE direction or sense, is the NEGATIVE *sense of rotation*. These facts are illustrated in Figs. 69 and 70.

Thus, if OA rotates in the POSITIVE direction (ccw), it generates a POSITIVE angle, but if it rotates in the negative direction (cw), it generates a negative angle.

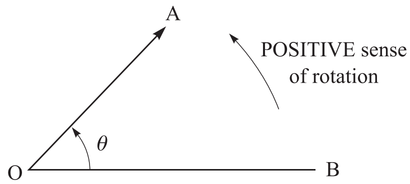


Fig. 69

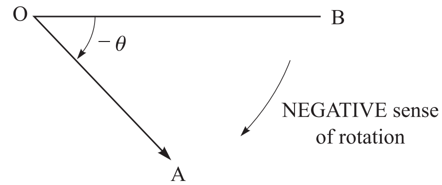


Fig. 70

Any given position of the radius vector OA can be expressed in terms of either a positive angle or a negative angle. An example of this is shown in Fig. 71, wherein the position of OA is the same in both cases.

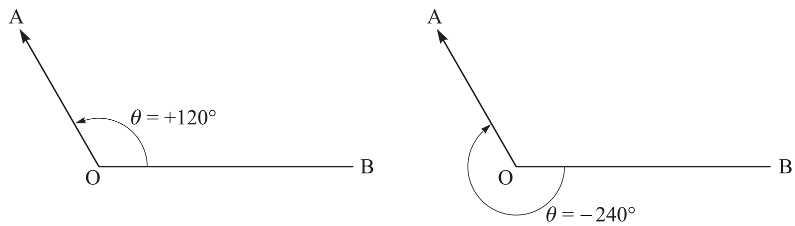


Fig. 71

As you can see in Fig. 71, OA ends up in the SAME POSITION whether we rotate it through a positive angle of 120 degrees or a negative angle of "minus 240" degrees. It thus follows that the POSITION of a radius vector OA is not changed if the angle θ is increased by any plus or minus integral multiple of 360 degrees; thus, in the example of Fig. 71, the position of OA is not changed if

$$\theta = 120° \pm (360n)°$$

where n is any integer (whole number).

In section 5.2, the functions $\sin\theta$, $\cos\theta$, $\tan\theta$, were defined only for values of θ lying in the interval from 0° to 90°. In this section, however, we've just been discussing values of θ outside the range of 0° to 90°, and this naturally brings up the question, "How are the functions $\sin\theta$, $\cos\theta$, $\tan\theta$ to be handled if the angle θ lies outside the range of zero degrees to 90 degrees?". For example, what meaning is to be given to terms such as sin 120°, cos 355°, tan 485°, and so on? In order to give meaning to such terms, the following system has been adopted, and is in universal use today.

We make use of the standard x and y rectangular coordinate system, with the four quadrants numbered as shown in Fig. 72. We then let the angle θ be generated by rotating the radius vector around the origin, as shown in Fig. 73. It should be emphasized that the four quadrants will always be numbered and referred to as shown in Fig. 72.

Next, in Fig. 73, A represents the LENGTH of the radius vector, it being defined that A is always given as a POSITIVE value. The *reference line*, from which all angles will be measured, will be the *positive side of the x-axis*. *Positive angles* are generated by rotating A in the ccw sense, and negative angles are generated when A is rotated in the cw sense (as in Figs. 69 and 70 above).

Now, with all the above in mind, the procedure for finding the values of the trigonometric functions of ANY ANGLE θ is as follows.

Begin by drawing a line from the tip of the radius vector *perpendicular to the x-axis*. Be careful to note that this line is always drawn from the tip of the radius vector to the *x-axis*, regardless of what quadrant the radius vector might be in. Drawing this perpendicular line

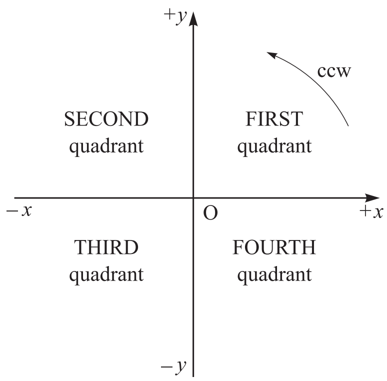


Fig. 72

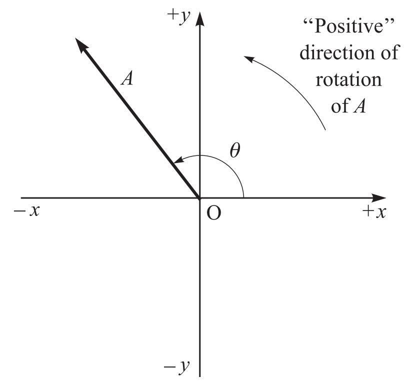


Fig. 73

will produce *a* RIGHT TRIANGLE *having an angle* ϕ, where ϕ will always be the angle *between A and the x-axis*. This is illustrated in Fig. 74 for the case where the independent variable (the angle θ) has a value such that A falls in the second quadrant.

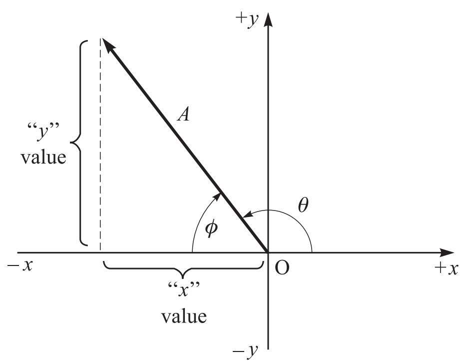


Fig. 74

Let us call ϕ the "subsidiary" angle. Note that, regardless of the quadrant that A falls in, ϕ itself will always be a simple angle in the range of 0° to 90°, which we'll define to always be a positive angle. Notice, however, that the "x values" and "y values" can be *either positive or negative*, depending upon the particular quadrant that A falls in (which, in turn, depends upon the value of the angle θ).

This combination, of positive values of ϕ from $\phi = 0^\circ$ to $\phi = 90^\circ$ and the positive and negative x and y values, is now utilized to define the values of $\sin\theta$, $\cos\theta$, and $\tan\theta$, *for any positive or negative value of* θ *of any magnitude*. This is done as follows.

To begin, inspection of Fig. 74 shows that

if A falls in the FIRST quadrant, x and y are BOTH POSITIVE,

if A falls in the SECOND quadrant, x is NEGATIVE and y is POSITIVE,

if A falls in the THIRD quadrant, x and y are BOTH NEGATIVE,

if A falls in the FOURTH quadrant, x is POSITIVE and y is NEGATIVE.

Next, given that A, the length of the radius vector, is always to be taken as positive, and also, recalling the basic definitions of the trigonometric functions from section 5.2, we now define that

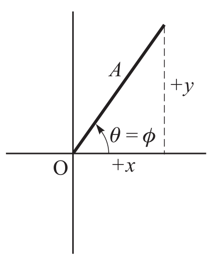
$$\sin\theta = \pm\sin\phi = \frac{y}{A} \tag{73}$$

$$\cos\theta = \pm\cos\phi = \frac{x}{A} \tag{74}$$

$$\tan\theta = \pm\tan\phi = \frac{y}{x} \tag{75}$$

where θ is a positive or negative angle of any value, and where ϕ is the subsidiary angle as previously described.

Since the signs of x and y depend upon the *quadrant* that A falls in, it follows that the signs ($+$ or $-$) of the functions $\sin\theta$, $\cos\theta$, and $\tan\theta$ will also depend upon the quadrant that A falls in. This is illustrated in Figs. 75 through 78. The relations between the functions follow from these figures and from eqs. (73), (74), and (75), and are so important that they should all be committed to memory.

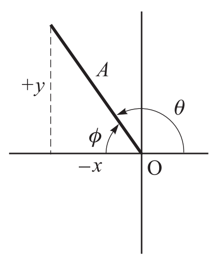


A in FIRST quad.
$\theta = 0°$ to $90°$
x = positive
y = positive

$\sin\theta = \sin\phi$
$\cos\theta = \cos\phi$
$\tan\theta = \tan\phi$

wherein: $\phi = \theta$

Fig. 75

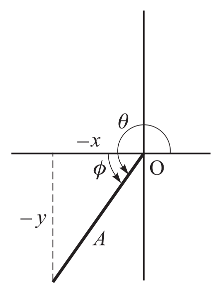


A in SECOND quad.
$\theta = 90°$ to $180°$
x = negative
y = positive

$\sin\theta = \sin\phi$
$\cos\theta = -\cos\phi$
$\tan\theta = -\tan\phi$

$\phi = 180 - \theta$

Fig. 76

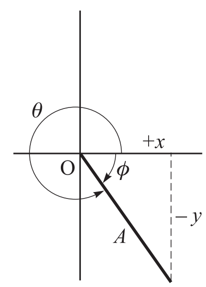


A in THIRD quad.
$\theta = 180°$ to $270°$
x = negative
y = negative

$\sin\theta = -\sin\phi$
$\cos\theta = -\cos\phi$
$\tan\theta = \tan\phi$

$\phi = \theta - 180$

Fig. 77



A in FOUTH quad.
$\theta = 270°$ to $360°$
x = positive
y = negative

$\sin\theta = -\sin\phi$
$\cos\theta = \cos\phi$
$\tan\theta = -\tan\phi$

$\phi = 360 - \theta$

Fig. 78

Before we leave this section there are some relationships, between the trigonometric functions for $+\theta$ and $-\theta$, that should be derived. This can be done with aid of Figs. 79 and 80.

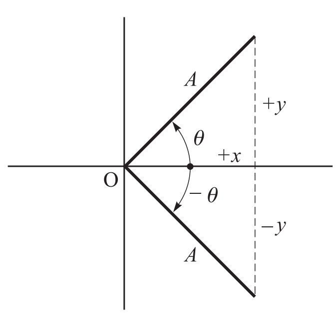


Fig. 79

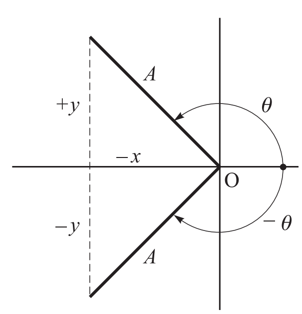


Fig. 80

From Fig. 79: thus:

$$\left.\begin{aligned}\sin\theta &= y/A\\ \sin(-\theta) &= -y/A\end{aligned}\right\}\ \sin(-\theta) = -\sin\theta$$

$$\left.\begin{aligned}\cos\theta &= x/A\\ \cos(-\theta) &= x/A\end{aligned}\right\}\ \cos(-\theta) = \cos\theta$$

$$\left.\begin{aligned}\tan\theta &= y/x\\ \tan(-\theta) &= -y/x\end{aligned}\right\}\ \tan(-\theta) = -\tan\theta$$

From Fig. 80: thus:

$$\left.\begin{aligned}\sin\theta &= y/A\\ \sin(-\theta) &= -y/A\end{aligned}\right\}\ \sin(-\theta) = -\sin\theta$$

$$\left.\begin{aligned}\cos\theta &= -x/A\\ \cos(-\theta) &= -x/A\end{aligned}\right\}\ \cos(-\theta) = \cos\theta$$

$$\left.\begin{aligned}\tan\theta &= y/-x = -y/x\\ \tan(-\theta) &= -y/-x = y/x\end{aligned}\right\}\ \tan(-\theta) = -\tan\theta$$

To summarize:

$$\boxed{\sin(-\theta) = -\sin\theta} \tag{76}$$

$$\boxed{\cos(-\theta) = \cos\theta} \tag{77}$$

$$\boxed{\tan(-\theta) = -\tan\theta} \tag{78}$$

Problem 65

Find the value of each of the following by making use only of the "table of values" given in section 5.2. (All angles are in degrees.)

(a) $\cos 115 =$ (c) $\tan 155 =$ (e) $\cos 95 =$ (g) $\sin 285 =$

(b) $\sin(-35) =$ (d) $\sin 255 =$ (f) $\tan(-285) =$ (h) $\sin(-188) =$

Thus, with aid of Figs. 75 through 78, the original "table of values," given in section 5.2, can now be extended to cover the full range of values from $\theta = 0^\circ$ to $\theta = 360^\circ$. A very short form of such a table follows, which let us now examine.

To do this, let us apply, to the following table of values, the same graphical procedure that we used at the beginning of this section. That is, let us now apply to the table the SAME PROCEDURE that gave us Figs. 66 and 67.

The results of doing this are shown in Figs. 81 and 82. Figure 81 shows, graphically, the behavior of the two "sinusoidal" functions, $y = \sin\theta$ and $y = \cos\theta$, over the complete range from $\theta = 0^\circ$ to $\theta = 360^\circ$. Figure 82 shows the behavior of $y = \tan\theta$ over the same range of values of θ. A brief discussion follows.

Consider Fig. 81 first. As the figure shows, the sinusoidal functions, $\sin\theta$ and $\cos\theta$, have a maximum value of 1 and a minimum value of -1. Note that the two waves are displaced from each other by 90 degrees; this is summarized by saying that the two waves are "90 degrees out of phase." Other than that, the two waveshapes are exactly the same.

$\theta°$	$\sin\theta$	$\cos\theta$	$\tan\theta$	$\theta°$	$\sin\theta$	$\cos\theta$	$\tan\theta$
0	0.0000	1.0000	0.0000	195	−0.2588	−0.9659	0.2680
15	0.2588	0.9659	0.2680	210	−0.5000	−0.8660	0.5774
30	0.5000	0.8660	0.5774	225	−0.7071	−0.7071	1.0000
45	0.7071	0.7071	1.0000	240	−0.8660	−0.5000	1.7321
60	0.8660	0.5000	1.7321	255	−0.9659	−0.2588	3.7321
75	0.9659	0.2588	3.7321	270	−1.0000	0.0000	∞
90	1.0000	0.0000	∞	285	−0.9659	0.2588	−3.7321
105	0.9659	−0.2588	−3.7321	300	−0.8660	0.5000	−1.7321
120	0.8660	−0.5000	−1.7321	315	−0.7071	0.7071	−1.0000
135	0.7071	−0.7071	−1.0000	330	−0.5000	0.8660	−0.5774
150	0.5000	−0.8660	−0.5774	345	−0.2588	0.9659	−0.2680
165	0.2588	−0.9659	−0.2680	360	0.0000	1.0000	0.0000
180	0.0000	−1.0000	0.0000				

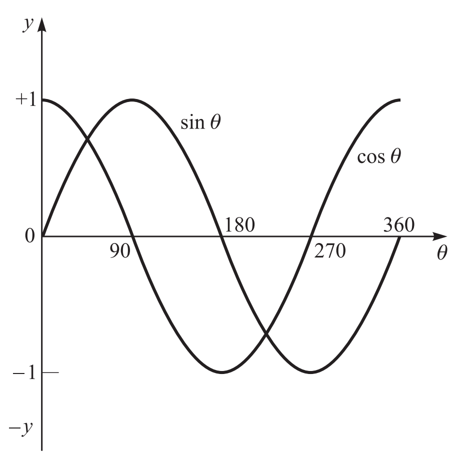


Fig. 81

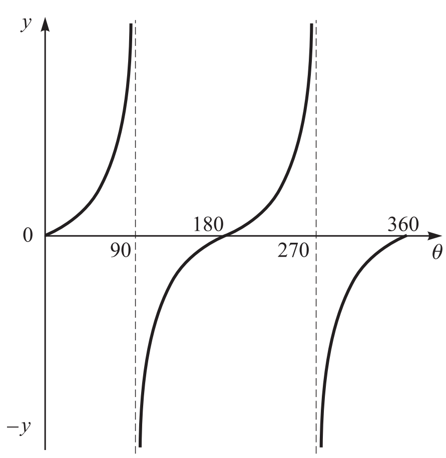


Fig. 82

The following facts concerning $\sin\theta$ and $\cos\theta$ (which are evident from inspection of Fig. 81) should be noted:

$\sin\theta = 0$ for $\theta = 0°$, $180°$, and $360°$,

$\sin\theta = 1$ for $\theta = 90°$, and $\sin\theta = -1$ for $\theta = 270°$;

$\cos\theta = 0$ for $\theta = 90°$ and $\theta = 270°$,

$\cos\theta = 1$ for $\theta = 0°$ and $360°$, and $\cos\theta = -1$ for $\theta = 180°$.

Now consider the function $y = \tan\theta$. From the above "table of values" we see that $\tan\theta$ becomes "infinitely great" for θ equal to either $90°$ or $270°$, and this condition is indicated in Fig. 82. The meaning of "infinitely great" was discussed in section 5.2 in connection with eq. (72). In Fig. 82, note that "$\tan\theta$" can become infinitely great in either the "positive" sense or the "negative" sense as θ approaches the value of either $90°$ or $270°$. Thus, if θ approaches the value of $90°$ from the *left* side of $90°$, then $\tan\theta$ becomes infinitely great in the "positive" sense, as shown in Fig. 82. If, however, θ approaches the value of $90°$ from the *right* side of $90°$, then $\tan\theta$ becomes infinitely great in the "negative" sense, as shown in Fig. 82. The function $\tan\theta$ is not defined for $\theta = 90^\circ$ or 270°, and is said to be "discontinuous" for these values of θ.

As previously mentioned, $\sin\theta$ and $\cos\theta$ are defined *for all positive and negative angles of any magnitude*. (Likewise for $\tan\theta$, except for values of θ for which $\tan\theta$ becomes infinitely great.)

To understand this, let us return briefly to Fig. 68, in which the radius vector OA can be in any given position, to which corresponds a given angle θ, with corresponding specific values of $\sin\theta$, $\cos\theta$, and $\tan\theta$.

Now note that if the radius vector OA makes any number of *full revolutions* of 360 degrees each, it simply returns to its *original given position*, thus reproducing the original values of $\sin\theta$, $\cos\theta$, and $\tan\theta$. Thus, if any given angle θ is increased or decreased by *any integral multiple of 360 degrees* the values of $\sin\theta$, $\cos\theta$, and $\tan\theta$ are unchanged; that is, the "new" values are the same as the "old" values. The mathematical statement is

$$\sin\theta^\circ = \sin(\theta \pm 360n)^\circ \tag{79}$$

$$\cos\theta^\circ = \cos(\theta \pm 360n)^\circ \tag{80}$$

where n is any integer (whole number).

The trigonometric functions are thus said to be PERIODIC or "repetitive" functions, because, for any given value of θ, the values of the functions are repeated over and over, endlessly, for each value of $\theta \pm 360n$ degrees, as stated in eqs. (79) and (80). This is illustrated in Fig. 83, which shows a few "cycles" of the graph or curve of the function $y = A\sin\theta$. Note that, since the maximum value of $\sin\theta$ is 1 (Fig. 81), it follows that the constant A is the maximum or "peak" value of the sine wave in Fig. 83.

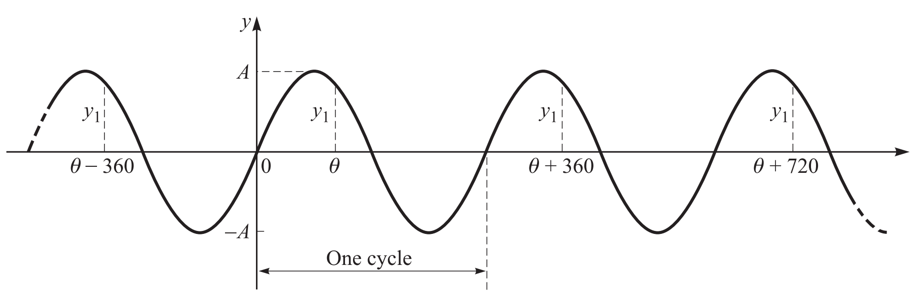


Fig. 83 $\boldsymbol{y = A\sin\theta}$.

In the figure, θ represents any value of angle we might select, and y_1 is the value of y for that particular value of θ. Notice that the same value, y_1, appears for every angle of $\theta \pm 360n$, as shown.

Problem 66

Letting θ and ϕ denote two angles of a right triangle, show that the following two identities are true (angles in degrees):

$$\sin\theta = \cos(90 - \theta)$$

$$\cos\theta = \sin(90 - \theta)$$

5.4 Choice of Waveform. Frequency. The Radian

In section 5.1 we defined the general difference between direct current and alternating current ("dc" and "ac").

Here, in the study of alternating currents in this book, let us now emphasize that, from now on, it will always be understood that we are talking about, and dealing with, SINUSOIDAL *waves of voltage and current*. There are a number of reasons why this is the practical thing to do.

First of all, it is highly desirable that the voltages and currents in large commercial power systems be as nearly sinusoidal as possible. This is because large generators, motors, transformers, and so on, operate much more smoothly and efficiently when sinusoidal voltages and currents are used, than when any other type of waveform is used.

As another example, in radio and television broadcasting the high-frequency "carrier" wave is sinusoidal in form. This is necessary to prevent excessive interference between stations. Also, sinusoidal waves are widely used in the testing and evaluation of electronic amplifiers, automatic control systems, and other electromechanical devices. Finally, and fortunately, the mathematical work is greatly reduced if sinusoidal waves are assumed. This is because of certain special characteristics possessed only by sinusoidal functions.*

With the foregoing in mind, let us return to Fig. 83, in which the independent variable is the angle θ, in degrees, and let us now bring *time*, t, into the picture.

To do this, let us think of the curve of Fig. 83 as being generated by a rotating "radius vector" of length A, as in Figs. 75 through 78. If A makes f REVOLUTIONS PER SECOND, then f is called the FREQUENCY of the wave, which is stated as so many CYCLES PER SECOND; hence, since each revolution of A generates 360 degrees, it follows that

$360ft =$ the total angle, θ, in degrees, generated in any total time of "t" seconds.

Thus, the equation for a *sine wave* of voltage v volts, having a frequency of f cycles per second (hertz),† can be written in the form

$$v = V \sin 360ft^\circ \tag{81}$$

where $v =$ instantaneous voltage at any time t seconds from $t = 0$, $V =$ maximum (peak) value of voltage, and $f =$ frequency in Hz (cycles/second).

In eq. (81) we can think of the instantaneous voltage, v, as being generated by a rotating radius vector of V volts. In this regard, since f is the number of revolutions per second, and since each revolution produces an advance of 360 degrees, it follows that

$360f =$ degrees advanced by the radius vector V per second,

and for this reason "$360f$" is sometimes called the "angular velocity" of the sine wave.

Let us now introduce a new term into our technical vocabulary, as follows.

---

* With respect to the variable θ, the RATE OF CHANGE of $\sin\theta$ is equal to the VALUE of $\cos\theta$, and the rate of change of $\cos\theta$ is equal to the value of $-\sin\theta$. The sine and cosine are the only periodic functions in all of mathematics that have such a uniquely useful property. We need not, however, dwell on this point at this time.

† The official term for "cycles per second" is "hertz" (abbreviated "Hz"), in honor of the great experimental physicist Heinrich Hertz.

So far in our work we've always measured angles in units of "degrees," a full circle being divided into 360 equal "degrees," as you know. There is, however, another unit of angular measurement called the "radian" that we should be familiar with. The "radian" is simply a unit of angular measurement, like the degree, but use of radians, instead of degrees, has certain advantages in theoretical work.

The radian unit of angular measurement is defined in references to a *circle*, just as the degree is. The radian will now be defined with the aid of Fig. 84.

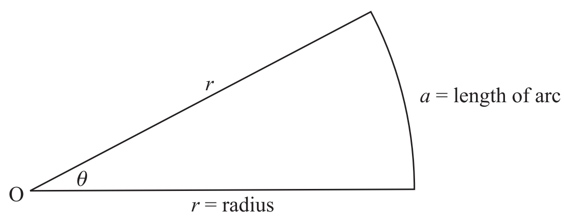


Fig. 84

In Fig. 84, let "O" denote the center of any circle of radius r, let "a" denote the "length of the arc" of the circle cut off by the two radii, and let "θ" denote the ANGLE "subtended" by the arc a, as shown. Then the *angle in* RADIANS is defined to be equal to the ratio of the *arc length to the radius*, that is,

$$\theta = \text{angle in radians} = \frac{\text{arc length}}{\text{radius}} = \frac{a}{r}$$

The ratio a/r does not, of course, depend upon the size of a particular circle we might draw, but only upon the value of the angle θ; that is, a is always directly proportional to r. Thus, using the standard notation of Fig. 84, we have the basic definition that the angle θ, *in radians*, is equal to

$$\boxed{\theta = \frac{a}{r}} \tag{82}$$

Figure 84 and eq. (82) should be committed to memory. Next, the relationship between RADIANS and DEGREES can be found by considering a FULL CIRCLE, as follows.

In a full circle we have $a =$ circumference of circle $= 2\pi r$; hence, by eq. 82, there are $2\pi r/r = 2\pi$ *radians in a full circle*. Since there are also 360 degrees in a full circle, we have that the ratio of *radians to degrees* is 2π *to 360* or π *to 180*, thus giving us the important relationship

$$\frac{\text{radians}}{\text{degrees}} = \frac{\pi}{180} \tag{83}$$

which should also be committed to memory. From eq. 83 we thus have the following two conversion formulas

$$\text{radians} = \left(\frac{\pi}{180}\right)(\text{degrees}) \tag{84}$$

$$\text{degrees} = \left(\frac{180}{\pi}\right)(\text{radians}) \tag{85}$$

Note that for *1 radian* eq. (85) gives the value, degrees $= (180/\pi)(1) = 57.295\,78$, that is, *one radian* equals *57.295 78 degrees* $= 57°17'45''$. Thus the "radian" is a much larger unit of angular measurement than the "degree." In any case, the conversion from one to the other is easily done with the aid of eqs. (84) and (85) and a calculator: for example, to convert 240 degrees to radians we use eq. (84); thus,

$$\text{radians} = (\pi/180)(240) = 4.188\,79, \quad \textit{answer}.$$

As previously mentioned, the radian is generally used in theoretical work instead of the degree. This fact leads us now to make the following IMPORTANT NOTE:

> From now on, *unless the "degree" symbol is shown*, all angles will be understood to be in *radians*.

Let us, therefore, now write eq. (81) in terms of *radians* instead of degrees; to do this, all we need do is substitute the total angle, $360ft$ degrees, into eq. (84); thus,

$$(\pi/180)(360ft) = 2\pi ft = \text{total } \textit{radians}$$

and therefore eq. (81) becomes, in terms of radian measurement of angles,

$$v = V \sin 2\pi ft \tag{86}$$

in which the "angular velocity" of the wave (see discussion following eq. (81)) is now $2\pi f$ radians per second.

We now have one final change in notation to make, as follows. It is universal practice to represent the quantity "$2\pi f$" by the small Greek letter "omega," written "ω"; thus it will always be understood that

$$\omega = 2\pi f = \text{angular velocity of radius vector in radians per second}$$

and thus, *t seconds* after starting at $t = 0$, the radius vector V will have covered a total "angular distance" of ωt radians, and thus eq. (86) now takes the standard form

$$v = V \sin \omega t \tag{87}$$

where v and V have the meanings defined in connection with eq. (81), and where $\omega = 2\pi f$, where f is the frequency in cycles per second (Hz). A corresponding sine wave of *current* will of course have the same basic form

$$i = I \sin \omega t \tag{88}$$

where now i is the instantaneous current and I is the maximum (peak) value of current.

For the next part of our discussion, let us begin by returning to the "table of values" listed in section 5.3. Note that the table covers the range of angular values from $\theta = 0°$ to $\theta = 360°$. What we now wish to do is to write the same table (omitting a few values here and there) in terms of *radians* instead of degrees. This can be done as follows. First, making use of eq. (84), you can verify that

$0° = 0$ rad.	$120° = 2\pi/3$ rad.	$210° = 7\pi/6$ rad.	$300° = 5\pi/3$ rad.
$30° = \pi/6$ rad.	$135° = 3\pi/4$ rad.	$225° = 5\pi/4$ rad.	$315° = 7\pi/4$ rad.
$45° = \pi/4$ rad.	$150° = 5\pi/6$ rad.	$240° = 4\pi/3$ rad.	$330° = 11\pi/6$ rad.
$60° = \pi/3$ rad.	$180° = \pi$ rad.	$270° = 3\pi/2$ rad.	$360° = 2\pi$ rad.
$90° = \pi/2$ rad.			

Let us now refer back to the table of values in section 5.3. Now, in that table, replace θ with ωt and degrees with their equivalent values from the above chart. This gives us the following table of values in terms of radians.

ωt	$\sin \omega t$	$\cos \omega t$	$\tan \omega t$	ωt	$\sin \omega t$	$\cos \omega t$	$\tan \omega t$
0	0.0000	1.0000	0.0000	$7\pi/6$	−0.5000	−0.8660	0.5774
$\pi/6$	0.5000	0.8660	0.5774	$5\pi/4$	−0.7071	−0.7071	1.0000
$\pi/4$	0.7071	0.7071	1.0000	$4\pi/3$	−0.8660	−0.5000	1.7321
$\pi/3$	0.8660	0.5000	1.7321	$3\pi/2$	−1.0000	0.0000	∞
$\pi/2$	1.0000	0.0000	∞	$5\pi/3$	−0.8660	0.5000	−1.7321
$2\pi/3$	0.8660	−0.5000	−1.7321	$7\pi/4$	−0.7071	0.7071	−1.0000
$3\pi/4$	0.7071	−0.7071	−1.0000	$11\pi/6$	−0.5000	0.8660	−0.5774
$5\pi/6$	0.5000	−0.8660	−0.5774	2π	0.0000	1.0000	0.0000
π	0.0000	−1.0000	0.0000				

We have previously (Figs. 81 and 82) sketched the curves of $\sin\theta$, $\cos\theta$, and $\tan\theta$ versus the angle θ in degrees. Now, in Fig. 85, we've used the table immediately above to sketch a couple of cycles of the functions, $y = A\sin\omega t$ and $y = A\cos\omega t$, versus the angle ωt in radians.

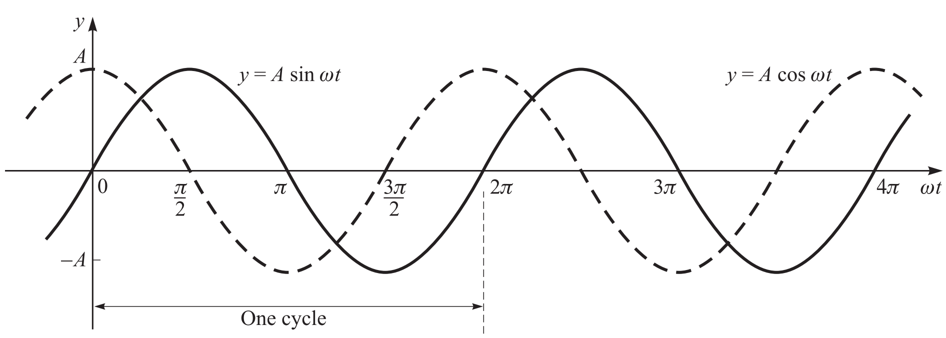


Fig. 85

In Fig. 85, the constant A denotes the maximum (peak) value of the sine and cosine functions. The independent variable is time t, in seconds, and, as always, $\omega = 2\pi f$, where f is constant frequency in cycles per second (Hz).

Also, in the figure, note that for $\omega t = 0$ the cosine function has the maximum value A, while at the same time the sine function has the value *zero*. Later in time, however, when $\omega t = \pi/2$, we see that the value of the cosine has fallen to *zero*, while the sine has risen to the maximum value of A. Thus, since the sine reaches its peak value at a *later time* than the cosine, we say that the sine function "lags" the cosine by $\pi/2$ radians (90 degrees). This is, of course, the same as saying that the cosine "leads" the sine by $\pi/2$ radians (90°).

As we know, the sinusoidal functions are PERIODIC functions, having a period of 2π radians. Any interval of 2π *radians* (360°) constitutes ONE CYCLE of a sinusoidal wave. In Fig. 85, for example, we indicate one particular cycle, in the interval from $\omega t = 0$ to $\omega t = 2\pi$.

For any given value of ωt, the value of a sinusoidal function is repeated over and over for each value of $\omega t \pm 2\pi n$, where n is any integer. Thus (corresponding to eqs. 79 and 80) we have that

$$\sin \omega t = \sin(\omega t \pm 2\pi n) \tag{89}$$

$$\cos \omega t = \cos(\omega t \pm 2\pi n) \tag{90}$$

where n is any integer.

Let us next find the relationship between the ''time of one cycle'' and the frequency f of a sinusoid.

The easiest way to do this is to make use of the particular cycle that begins at $\omega t = 0$ in Fig. 85 (which we've labeled ''one cycle'' in the figure). If we let ''large T'' denote the TIME OF ONE CYCLE, then, at the *end* of this particular cycle, when $t = T$, we see from the figure that $\omega T = 2\pi$, that is, $2\pi f T = 2\pi$, from which we get the desired relationship

$$\boxed{fT = 1} \tag{91}$$

in which f is frequency in cycles per second (Hz) and T is the time of one cycle in seconds.

We now conclude this section with a discussion of ''phase shift'' and ''phase angle'' as used in connection with sinusoidal waves.

We can begin by pointing out that the term ''phase,'' as used in electrical engineering, refers in general to an ''angular'' relationship of some kind. As applied to sinusoidal waves, the terms ''phase shift'' and ''phase angle'' refer to the amount of ANGULAR DISPLACEMENT of such waves; for instance, this might be the angular displacement with respect to the origin of the coordinate axes, or the angular displacement of one wave with respect to another wave of the same frequency.

For example, in the preceding discussion of Fig. 85 we noted that the cosine curve ''leads'' the sine curve by $\pi/2$ radians or 90°. In phase terminology we could say that the ''phase angle'' between the cosine and sine functions is $\pi/2$ radians, or that the cosine has a ''phase shift,'' or is ''phase shifted,'' in the amount of $+\pi/2$ radians with respect to the sine.

The phase angle between two sinusoidal waves of the same frequency is measured between any two successive, corresponding, points of the two waves. For instance, this can be the ''angular distance,'' in radians or degrees, between two successive ''peak values'' of the waves, or the angular distance between the points at which the curves are rising in the positive sense as they cross the horizontal axis. This is illustrated in Fig. 86, in which the phase angle, 45°, is the angle between two such consecutive ''crossover'' points, as shown. In this case, curve A can be said to ''lead'' curve B by 45 degrees, because A reaches its peak positive value 45° before B (as mentioned in the discussion of Fig. 85).

There is an item of interest in connection with Fig. 86 that should be mentioned. In the figure, the phase angle of 45° ($\pi/4$ radians) is angular displacement between the two waves themselves, as shown. The point we wish to make, however, is that this

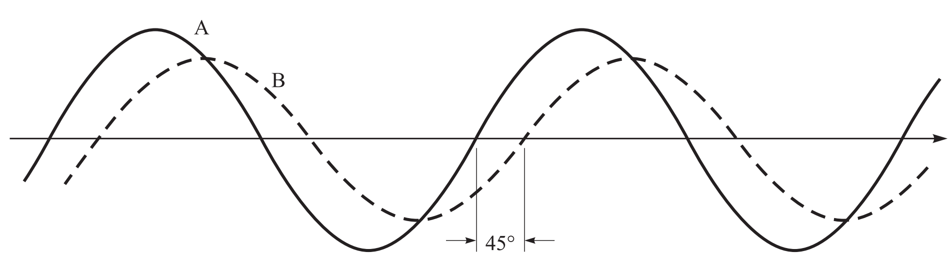


Fig. 86

information alone does not permit us to write the *equations* of the two waves, even if their peak values are given. This is because the location of the *origin* of the axes, relative to the waves, is not shown in the figure.

To illustrate this, suppose the peak value of A is 10 and the peak value of B is 7, and suppose it is given that *curve A passes through the origin* (in the manner of the sine wave in Fig. 85). With this information the equations of the two waves *can* now be written thus (in radians):

$$\text{for A:} \quad y = 10 \sin \omega t \tag{92}$$

$$\text{for B:} \quad y = 7 \sin(\omega t - \pi/4) \tag{93}$$

Equation (93) is the mathematical way of showing that sinusoid B "lags" sinusoid A by $\pi/4$ radians (45°).

Another point to be emphasized is that the curves A and B in Fig. 86 represent two sinusoidal functions having the SAME FREQUENCY. If two sinusoidal functions do *not* have the same frequency, then no fixed phase relationship exists between the two functions, and the term "phase shift" would have little meaning. This is illustrated in Fig. 87, in which A and B denote curves of two sinusoids having unequal frequencies.

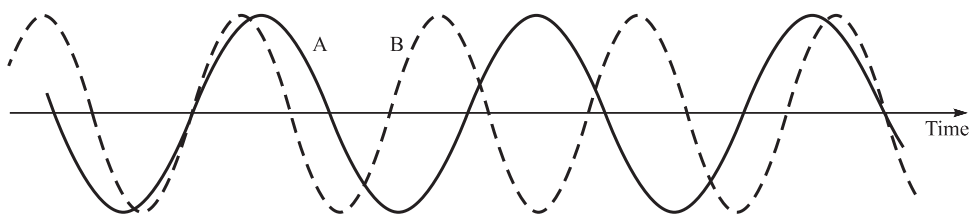


Fig. 87

Problem 67
Given that $v = 100 \sin 180\,000t^\circ$, find

(a) frequency in hertz,

(b) value of v at the instant $t = 0.15$ second.

Problem 68
If $\omega = 533\,850$ for a certain sine wave, find the time of one cycle in microseconds.

Problem 69
In Fig. 87, which wave, A or B, represents the higher frequency?

5.5 Power; rms Value of a Sine Wave of Voltage or Current

In section 2.3 we showed that in a dc (direct-current) circuit, the *power* P in watts, expended in a resistance of R ohms, is given by any of the formulas

$$P = VI, \qquad P = V^2/R, \qquad P = I^2R$$

where V is dc voltage and I is dc current.

Now let's consider the problem of how to calculate power in an ALTERNATING CURRENT (ac) circuit.

We at once see that there is a problem here, because the power in an ac circuit is not constant but changes from instant to instant throughout the cycle, because the voltage and current are both continually changing during the cycle.

This difficulty is resolved by defining that by "power" in an ac circuit we will always mean the AVERAGE POWER in the circuit. This definition leads to what is called the "effective" or "rms" value of an ac voltage or current. The development proceeds as follows.

To begin, let us make the following slight change in notation. In eqs. (87) and (88) we used V and I to denote the peak values of sine waves of voltage and current. Let us now, for convenience later on, change that notation and, hereafter, always denote peak values by V_{p} and I_{p}, instead of by plain V and I. Upon making this change in notation, eqs. (87) and (88) become

$$v = V_{\mathrm{p}} \sin \omega t \tag{94}$$

$$i = I_{\mathrm{p}} \sin \omega t \tag{95}$$

In the above equations, v and i denote *instantaneous* voltage and current at any time t seconds after we start "positive time" at $t = 0$. It then follows, from the basic considerations used to derive eq. (15) in Chap. 2, that *instantaneous power*, p, is equal to *instantaneous voltage times instantaneous current*, that is

$$p = vi \tag{96}$$

or, in terms of a load resistance of R ohms, eq. (96) can be written in the forms

$$p = v^2/R \qquad \text{and} \qquad p = i^2 R.$$

But, as already mentioned, we are not interested in "instantaneous" power; instead, we are interested in finding the AVERAGE POWER obtained over ONE COMPLETE CYCLE of the sine waves of eqs. (94) and (95). (The average power over any one complete cycle is the same for all cycles, and is the average power as long as the waves continue to exist.)*

To continue, let us now substitute, into eq. (96), the values of v and i from eqs. (94) and (95). Doing this, we have that the instantaneous power p in an ac circuit is equal to

$$p = V_{\mathrm{p}} I_{\mathrm{p}} (\sin \omega t)^2 = V_{\mathrm{p}} I_{\mathrm{p}} \sin^2 \omega t \tag{97}$$

in which we'll assume that the peak values, V_{p} and I_{p}, will remain constant in any given problem. From eq. (87), the angle ωt is in radians.

We recall that the function $\sin \omega t$ goes through one complete cycle in the period from $\omega t = 0$ to $\omega t = 2\pi$ (Fig. 85). Our problem, therefore, now is to *find the* AVERAGE VALUE *of eq. (97)* when the voltage and current waves, v and i (eqs. (94) and (95)), go through one complete cycle from $\omega t = 0$ to $\omega t = 2\pi$. One way this can be done is to make use of the trigonometric identity

$$\sin^2 \theta = \frac{1}{2}(1 - \cos 2\theta) \tag{98}^\dagger$$

* Power does not "accumulate" in a circuit, because power is a measure of the RATE at which work is being done. This could be compared to an automobile moving at a constant speed; if the speed is measured for a period of, say, 1 minute, the same speed is measured for all intervals of 1 minute, and is the speed for the entire trip.

† See note 6 in Appendix.

Now set $\theta = \omega t$ in eq. (98), then substitute the right-hand side into eq. (97) in place of $\sin^2 \omega t$. Doing this, eq. (97) becomes

$$p = \frac{V_p I_p}{2} - \frac{V_p I_p}{2} \cos 2\omega t \tag{99}$$

Now examine the right-hand side of eq. (99). The average value of the *first term* is $V_p I_p/2$, because it has the same constant value for the entire interval from $\omega t = 0$ to $\omega = 2\pi$. (If an automobile maintains a constant speed of 60 mph over a period of time, the average speed over the period is 60 mph.)

Note, however, that the average value of the *second term* is *zero* over the same interval from $\omega t = 0$ to $\omega t = 2\pi$. To understand why this is true, let us take a moment out to review the basic meaning of "average value," as follows.

Let $y_1, y_2, y_3, \ldots, y_n$ denote n different values of a variable y, measured over a certain range of values of whatever independent variable determines the values of y. Note that some of the values of y may be positive in value, and others negative in value. The "average value of y," over the particular interval chosen, is then defined to be equal to the *algebraic sum of all the n values, divided by n*, that is

$$\text{average value of } y = Y = \frac{y_1 + y_2 + y_3 + \cdots + y_n}{n} \tag{100}$$

Note that if the algebraic sum of the numerator values is *zero* the average value of y is *zero*, and this is exactly what happens in the case of the *second term* on the right-hand side of eq. (99). This is true because, in any number of complete cycles of a sinusoidal wave, there are as many positive values as negative values. Thus we have determined that the *average value of eq. (99)* (and also of eq. (97)) is equal to $V_p I_p/2$; that is, letting P denote *average power* produced in eqs. (99) and (97), we have that

$$P = \frac{V_p I_p}{2} = \frac{V_p}{\sqrt{2}} \cdot \frac{I_p}{\sqrt{2}} \tag{101}$$

where the "dot" means "times," and where V_p and I_p are the maximum (peak) values of an applied sinusoidal voltage and the resulting sinusoidal current.

Let us now denote the two quantities on the right-hand side of eq. (101) by V and I, without subscripts, thus

$$V = V_p/\sqrt{2} = 0.7071 V_p \tag{102}$$

$$I = I_p\sqrt{2} = 0.7071 I_p \tag{103}$$

The quantities represented by V and I above are called the "effective" values or the "rms" values* of sinusoidal voltages and currents of peak values V_p and I_p. Note that, using this notation, eq. (101) takes the form

$$\boxed{P = VI} \tag{104}$$

---

* "rms" stands for "*r*oot *m*ean *s*quare." Noting that "mean" is the same as "average," the origin of rms is as follows. First, instantaneous power, p, is proportional to the square of instantaneous current, i^2, as shown following eq. (96). Thus, over a period of time, AVERAGE POWER, P, is proportional to the *average value of* i^2, which, for n consecutive samples, let us denote by I^2; thus

$$I^2 = \frac{i_1^2 + i_2^2 + i_3^2 + \cdots + i_n^2}{n}$$

Thus I, as derived here, is the "square root of mean square" or rms value of current. Hence "rms" and "effective" have the same meaning, being the value of current or voltage used to calculate average power.

You'll recall that "power" is expressed in "watts" in electric circuit calculations (see eq. (15) in Chap. 2). The above equation states that the AVERAGE POWER P produced in a resistance of R ohms is equal to the RMS VOLTAGE TIMES THE RMS CURRENT.

If we now apply the foregoing procedure to the equations $p = v^2/R$ and $p = i^2R$, we find that

$$P = V^2/R \tag{105}$$

and

$$P = I^2R \tag{106}$$

and, comparing eqs. (106) and (104), we see that $I^2R = VI$, and thus we have OHM'S LAW for the sinusoidal "ac" case:

$$\boxed{I = \frac{V}{R}} \tag{107}$$

In eqs. (104) through (106), P is average power in watts, V and I are rms values of voltage and current, and R is load resistance in ohms.

It is important to note that eqs. (104) through (107), for the ac circuit, have exactly the same form, and are subject to the same algebraic manipulation, as the equations for dc circuits summarized following eq. (17) in section 2.3. This procedure can be used because, in ac circuit work, we are normally not interested in knowing *instantaneous* values of power, voltage, and current (given by eqs. (94) and (95) in the sinusoidal case), but only in average power and rms values of voltage and current, which are not functions of time.

Problem 70
Are eqs. (104) through (107) basically true for non-sinusoidal periodic waveforms of voltage and current, as well as sinusoidal?

Problem 71
It is given that ac voltmeters and ammeters are normally calibrated to read rms values. If the meter readings are 120 volts and 8.5 amperes, find the following values (sinusoidal conditions will always be assumed, unless definitely stated otherwise):

(a) average power input to the circuit,
(b) peak power input to the circuit.

5.6 Sinusoidal Voltages and Currents as Vectors

A generator of dc voltage is usually represented by the battery symbol, Fig. 88, while an ac generator is usually represented by the "slip rings" symbol of Fig. 89.*

In the dc case of Fig. 88, the polarity DOES NOT CHANGE WITH TIME; thus, in Fig. 88, one of the battery terminals will always be POSITIVE with respect to the other terminal. The situation can be indicated either by the use of "+" and "−" signs or by means of a "voltage arrow" placed alongside, with the understanding that the "head" of

* This originated as a symbol for the circular copper "slip rings" used to connect the rotating coils (armature) of an ac generator to outside stationary circuitry.

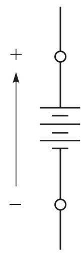


Fig. 88

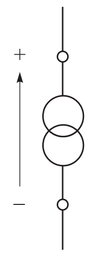
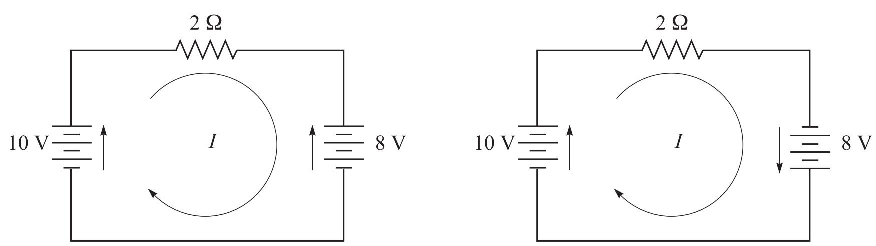


Fig. 89

the arrow is POSITIVE with respect to the "tail," as shown. That is, a voltage arrow will always point FROM THE NEGATIVE TERMINAL TO THE POSITIVE TERMINAL of a generator.

It's obviously not necessary to show both polarity marks and voltage arrow, and thus generally, from now on, we'll use only voltage arrows, having the meaning just described above.

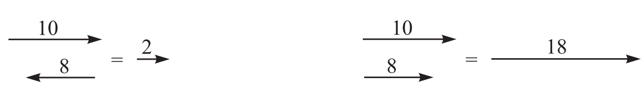
In our work with dc circuits, in Chap. 4, we found that it is absolutely necessary to indicate the POLARITIES of the dc generators in a network. The importance of this requirement is illustrated in the simple circuits shown in Fig. 90, in which batteries represent dc generators with polarities indicated by voltage arrows as shown. (The Greek letter Ω, capital "omega," denotes resistance in ohms.)



$I = \frac{V}{R} = \frac{10-8}{2} = 1\text{ A}$ $\qquad$ $I = \frac{V}{R} = \frac{10+8}{2} = 9\text{ A}$

Fig. 90

Notice that the two circuits give completely different values of current; this is because in the left-hand diagram the two dc generators are connected so as to OPPOSE each other, while in the right-hand diagram they are connected so as to AID each other. This can be understood by tracing around both circuits in the clockwise sense, remembering that going through a generator "with" the voltage arrow represents a *rise* in potential, while going through "against" the arrow represents a *decrease* in potential. The situation in Fig. 90 is represented graphically in "vector diagram" form below.



Left-hand circuit $\qquad$ Right-hand circuit

The above example illustrates the fact that two dc generators in the same circuit will either *totally* AID each other or *totally* OPPOSE each other, with no "in between" conditions possible. Thus, the voltages produced by two dc generators cannot be 45° apart, or 65° apart, or 125° apart, and so on. Likewise, the ‘‘phase angle’’ of the ‘‘voltage drops’’ across two resistances in a dc network can have only the relative value of 0° or 180°.

Now, however, consider the case of two series-connected AC GENERATORS of the same frequency. Here it’s possible to have two generators in which the phase angle between their voltage waves can be ANY POSITIVE OR NEGATIVE ANGLE FROM 0° TO 360°. Likewise, the voltage drop across a circuit component can have different phase angles relative to the voltage drops across other circuit elements and generator voltages. Thus, unlike dc circuit analysis, in ac work we must take the factor of PHASE into account. For this reason the algebra of ac circuits is more complicated, in its inner details, than the algebra of dc circuits. Broadly speaking, however, we’ll find that many of the basic procedures we learned in dc analysis will carry over directly to ac work.

With the above in mind, let us now turn our attention to the case where the voltage sources are AC GENERATORS instead of dc generators. We begin our discussion with Fig. 91, using the ac generator symbol of Fig. 89 with voltage arrows as previously mentioned. The circles labeled V_1, V_2, and V represent ac voltmeters, with I being an ac ammeter. It is given that ac meters are always calibrated to read rms values, unless definitely stated otherwise on the face of the meter. Let us now make a careful study of Fig. 91.

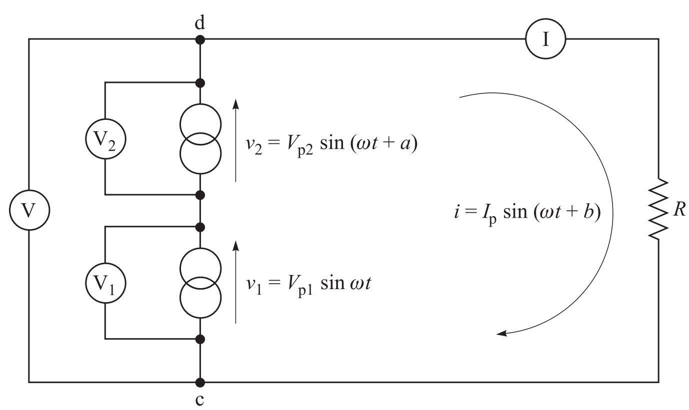


Fig. 91

Note, first, that the independent variable here is *time t*. The function ‘‘$\sin \omega t$’’ is, as we know, a ‘‘periodic’’ function of time, being alternatively positive and negative for equal time periods.*

In Fig. 91, v_1 and v_2 denote instantaneous values of voltage of two sine waves OF THE SAME FREQUENCY, with V_{p1} and V_{p2} being the ‘‘peak’’ values of the two waves.

Note that $v_1 = 0$ *when* $t = 0$; hence the lower sine wave passes through the *origin* of the axes, with the upper sine wave ‘‘leading’’ the lower wave by a radians (see discussion of ‘‘phase angle’’ in section 5.4). Inspection of Fig. 91 shows that

$$v_{cd} = v_1 + v_2 = V_{p1} \sin \omega t + V_{p2} \sin(\omega t + a) \tag{108}$$

* Each cycle of a sinusoidal wave is said to consist of a ‘‘positive alternation’’ and a ‘‘negative alternation.’’ For example, in Fig. 85, for the indicated ‘‘one cycle’’ of $A \sin \omega t$, the ‘‘positive alternation’’ exists from $\omega t = 0$ to $\omega t = \pi$, and the ‘‘negative alternation’’ from $\omega t = \pi$ to $\omega t = 2\pi$.

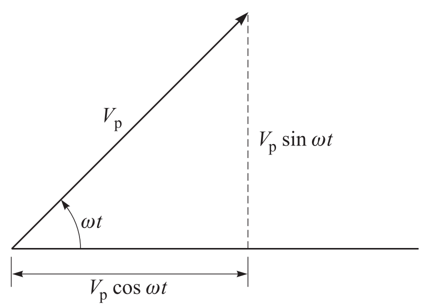
where v_{cd} = instantaneous voltage between terminals c and d, at any time t seconds after $t = 0$.

The voltage arrows, as used in the ‘‘instantaneous’’ picture of Fig. 91, have their usual meaning; in this case the arrows simply denote what the generator polarities would be during the times when the functions $\sin \omega t$ and $\sin(\omega t + a)$ are positive in value. The arrows would thus point ‘‘downward’’ during the times when $\sin \omega t$ and $\sin(\omega t + a)$ are negative in value.

The current arrow denotes the direction of flow of ‘‘positive’’ current, that is, the direction the current is assumed to flow during the times that $\sin(\omega t + b)$ is positive in value.

It should be noted that (108) is a simple addition of instantaneous voltages, and is *not* a vector relationship as it stands. Actually, however, in most practical work we are interested only in *rms values* of voltage and current, and *not* in instantaneous values. If we make use of this practical fact, it is possible to put the rms information contained in eq. (108) into a form to which vector methods *can* be applied; this extremely useful concept can be developed as follows.

As pointed out in section 5.3, it’s often convenient to represent the instantaneous values of a sinusoidal voltage or current as being generated by a *rotating ‘‘radius vector’’* (also called a ‘‘phasor’’), as illustrated in Fig. 92.



where V_p = peak value of sinusoidal wave
$V_p \cos \omega t$ = instantaneous ‘‘horizontal component’’ of V_p
$V_p \sin \omega t$ = instantaneous ‘‘vertical component’’ of V_p

Fig. 92

In Fig. 92, note that it is *not correct* to write that $V_p = V_p \cos \omega t + V_p \sin \omega t$, because the horizontal and vertical components must be added together *at right angles* to give V_p. Thus V_p in Fig. 92 can be regarded as being an instantaneous VECTOR quantity, having horizontal and vertical components as shown. It therefore follows that eq. (108) can be represented as being generated by TWO such rotating ‘‘phasors,’’ as illustrated in Fig. 93.

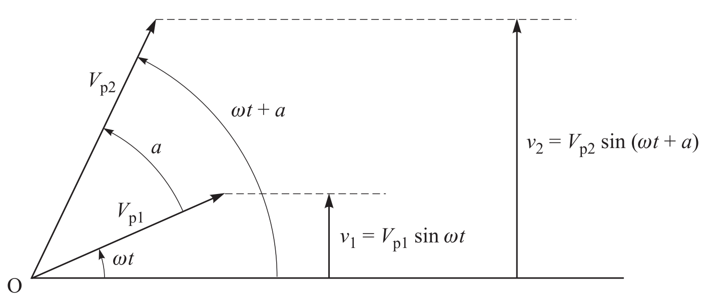


Fig. 93

Figure 93 is thus an instantaneous "picture" of eq. (108) at any time t seconds after $t = 0$. Note, however, that the phase angle a is *not* a function of time, but remains *constant* in value at all times. (It must be emphasized again that this is true only if v_1 and v_2 are waves of the SAME FREQUENCY.) Thus, as far as the relationship between the phase angle a and the constant peak values of the phasors are concerned, we can *disregard the rotary motion* in Fig. 93 and show only the constant vector relationship between the magnitudes of the phasors and the phase angle a, as in Fig. 94.

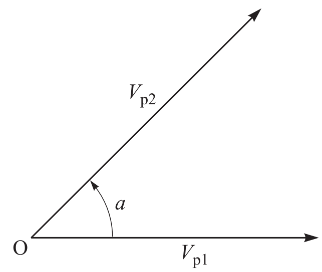


Fig. 94

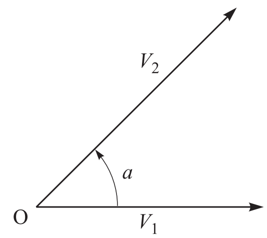


Fig. 95

In Fig. 94, let us now multiply V_{p1} and V_{p2} by 0.7071; this converts the peak voltages into *rms* voltages, giving us Fig. 95, in which V_1 and V_2 are given to be rms values of voltage. Since we'll normally be interested only in rms values of voltage and current, let us agree that, from now on, all such vector diagrams will represent *rms* values of voltage and current, as in the voltage case of Fig. 95.

In Fig. 95, V_1 is the rms value of the sine wave, $V_{p1} \sin \omega t$, which passes through the origin of the axes (as mentioned just prior to eq. (108)). Thus, in Fig. 95, it is reasonable to select V_1 to be the "reference vector," taken to be at an angle of "zero degrees," the angles of all other vectors being given *relative to the reference vector* V_1.*

In Fig. 95, the *total* RESULTANT *rms voltage*, which let us denote by V, is equal to the *vector sum* of V_1 and V_2; geometrically, V is equal to the diagonal of the parallelogram formed with V_1 and V_2 as the two sides. This is shown in Fig. 96, in which h is the angle between the resultant vector V and the reference vector V_1.

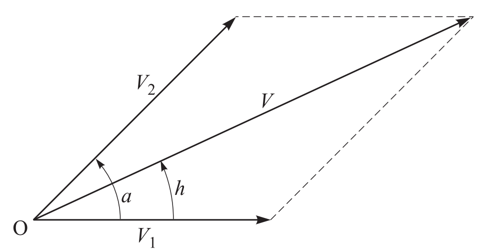


Fig. 96

Algebraically, one way of finding the magnitude and phase angle of V in Fig. 96, given the magnitudes of V_1 and V_2 and the phase angle a, is to first resolve V_1 and V_2 into their individual HORIZONTAL AND VERTICAL COMPONENTS; then the

$$\text{horizontal component of } V = \text{sum of horizontal components of } V_1 \text{ and } V_2$$
$$\text{vertical component of } V = \text{sum of vertical components of } V_1 \text{ and } V_2$$

* If you wish to study a review of vectors, see note 4 in Appendix.

The above rule is correct because the horizontal components all lie along the same straight line, and the vertical components also all lie along a straight line (perpendicular to the line containing the horizontal components).

Hence, upon using the above procedure, we now have the horizontal and vertical components of the resultant vector V, which let us denote by V_h and V_v respectively. Thus, using this notation, Fig. 96 becomes Fig. 97.

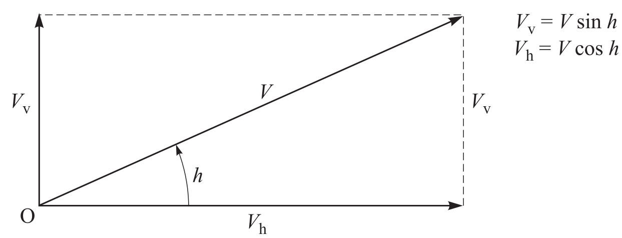


Fig. 97

Since Fig. 97 is a right triangle, the complete description of the resultant vector V, in both magnitude and phase angle, is given by the equations

$$|V| = \sqrt{V_h^2 + V_v^2} \tag{109}$$

$$h = \arctan(V_v/V_h) \qquad (110)*$$

In the same way, the resultant vector, V, of THREE or more vectors is equal to the vector sum of the horizontal components and the vertical components of the individual vectors.

Example 1

As previously mentioned, ac voltmeters and ammeters read MAGNITUDE OF RMS voltage and current. Going back to Fig. 91, suppose the two series-connected voltmeter readings are $V_1 = 58$ volts, $V_2 = 112$ volts, and suppose it is also known that V_1 and V_2 are 65 degrees out of phase. (It is understood, as always, that we're dealing with sinusoidal waves of the same frequency.)

(a) Taking V_1 to be the reference vector, with V_2 leading by 65°, find the magnitude and phase angle of the vector sum, V, of V_1 and V_2.
(b) What would be the reading of voltmeter "V" in the figure?

Solution

(a) First, the values of the horizontal and vertical components are as follows. For V_1, $V_h = 58$ volts and $V_v = 0$ volts. Next, for V_2, $V_h = 112 \cos 65^\circ = 47.333$ volts, and $V_v = 112 \sin 65^\circ = 101.507$ volts.

* In Fig. 97, by definition, $\tan h = V_v/V_h$; eq. (110) simply says that h is the ANGLE whose tangent is equal to V_v divided by V_h. For example, since

$$\tan 60^\circ = 1.7321$$

we have the inverse statement that

$$60^\circ = \arctan 1.7321$$

Hence the *horizontal* component of V is $58 + 47.333 = 105.333$ volts, and the *vertical* component of V is 101.507 volts. Thus,

$$\text{by eq. (109), } |V| = \sqrt{21\,398.71} = 146.283 \text{ volts, approx., and}$$
$$\text{by eq. (110), } h = \arctan(101.507/105.333) = \arctan 0.963\,68 = 43.94^\circ.$$

The above *answers* can be combined together in the convenient "polar" form, thus

$$\bar{V} = 146.283\underline{/43.94^\circ}$$

in which $\bar{V}$ denotes that V is a vector quantity, which can be read as "vector V." Thus the above answer can be read as "V is a vector quantity of magnitude 146.283 at angle of 43.94 degrees."

(b) 146.283 volts, *answer*, because ac meters read the magnitude of rms values.

Let us next consider certain details about the CURRENT and POWER that are produced by ac generators working into a purely RESISTIVE load. We begin our discussion with Fig. 98, in which a single ac generator, of peak voltage V_p, produces an ac current of peak value I_p amperes in a load resistance of R ohms, as shown.

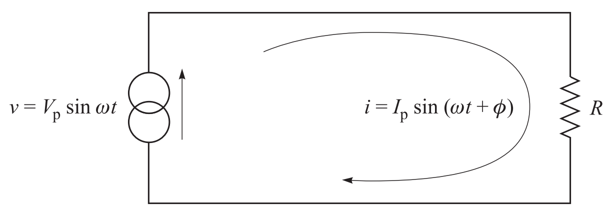


Fig. 98

Let us now SUPPOSE that a phase shift of ϕ radians exists between the current wave and the voltage wave, as shown in the figure. As we can see from the figure, the generator voltage v is at all times equal to the voltage drop across R, which is equal to Ri by Ohm's law. Thus at all times

$$v = Ri$$

$$V_p \sin \omega t = RI_p \sin(\omega t + \phi)$$

The above equation is true for all values of time, including $t = 0$, and upon setting $t = 0$ we have that (since $\sin 0 = 0$)

$$0 = RI_p \sin \phi$$

which can be true *only if* $\phi = 0$, which shows that there is ZERO PHASE SHIFT between the voltage and current waves in a purely *resistive* circuit. This fact is shown in phasor diagrams, for the case of Fig. 98, in Fig. 99, in which Fig. 99A and B are combined in the single diagram of Fig. 99C.

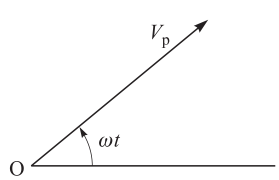


Fig. 99A

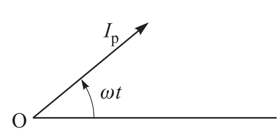


Fig. 99B

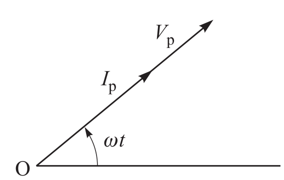


Fig. 99C

With the foregoing in mind, let us return to Fig. 91, and also to Fig. 93, which is the phasor diagram for Fig. 91.

In Fig. 91 we indicate a current $i = I_p \sin(\omega t + b)$; but notice that we do not show the phasor for this current in Fig. 93. Our object now, therefore, is to add, to the phasor diagram of Fig. 93, phasors that will account for the CURRENT, $i = I_p \sin(\omega t + b)$, that flows in Fig. 91. This can be done by calling upon the extremely useful "principle of superposition" (problem 50, section 4.4) as follows.

By the principle of superposition, the total effect of the two generators in Fig. 91 is the same as if each generator acted separately, producing its own separate component of current, the vector *sum* of the two components of current being equal to the total current.

Since, in Fig. 91, the load is a pure *resistance* of R ohms, we know that each of the two current components will be *in phase* with the generator voltage that produces it. Thus, if we let I_{p1} and I_{p2} be the peak values of currents produced separately by generator voltages V_{p1} and V_{p2}, then Fig. 93 can be redrawn to include the CURRENT PHASORS I_{p1} and I_{p2}, as shown in Fig. 100.

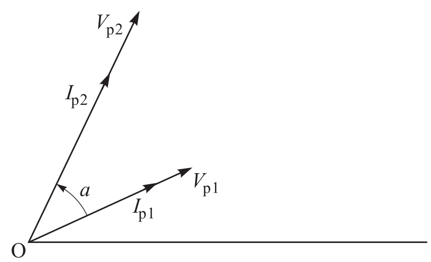


Fig. 100

Now let us multiply the magnitudes of all four vectors in Fig. 100 by 0.7071; this will not change the angle a in any way, but the lengths (magnitudes) of the vectors will now represent *rms* values instead of peak values. Let us denote the rms values of voltage by V_1 and V_2 (as we do in Figs. 95 and 96), and then let I_1 and I_2 denote the *rms currents* produced by the rms voltages V_1 and V_2.

Now let V denote the *vector sum* of the rms voltages V_1 and V_2, as shown in Fig. 96. Then let I denote the *vector sum* of the rms currents I_1 and I_2; it then follows that the *resultant rms vector I* lies in the *same direction as the resultant rms voltage vector V*,* as shown in Fig. 101. (Fig. 101 is thus the same as Fig. 96 except the current vectors have been added and the figure increased in size somewhat.)

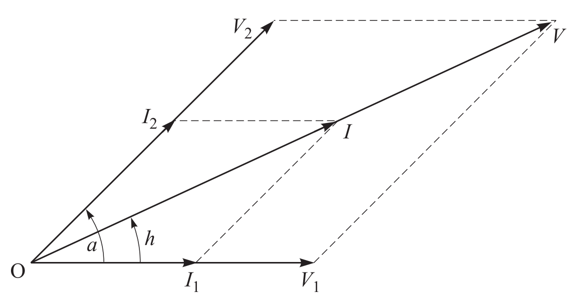


Fig. 101

* There must be "zero phase shift" between V and the current I it produces in a purely resistive circuit, as emphasized in Fig. 99.

Figure 101 is thus the COMPLETE VECTOR DIAGRAM for the purely resistive circuit of Fig. 91, in which the lengths of the vectors represent constant rms values of voltage and current. Note that V is the vector sum of generator voltages V_1 and V_2, while I is the vector sum of generator currents I_1 and I_2. Since Fig. 91 is a *resistive* circuit, note that I_1 is in phase with V_1, I_2 is in phase with V_2, and the *overall resultant current* I is in phase with the *overall resultant voltage* V (all as shown in the manner of Fig. 99C). We again emphasize that Fig. 101 does *not* represent an instantaneous "time" relationship, but depicts the constant phase relationships of the vectors representing the various constant rms values of current and voltage.

Also, comparison of Figs. 101 and 91 shows that the constant angle b, originally used in Fig. 91, is really the same as the angle h in Fig. 101; that is, h is the constant angle between the current wave i and the reference voltage wave v_1 in Fig. 91.

Relationships expressed in "vector diagram" form, such as in Fig. 101, are, of course, also expressible in "algebraic" form, using either the "bar" or "polar" notation (see example 1).

For example, Fig. 101 shows graphically that the total resultant rms voltage $\bar{V}$, appearing between terminals c and d in Fig. 91, is equal to the vector sum of voltages $\bar{V}_1$ and $\bar{V}_2$; this fact can be expressed algebraically by writing that, in Fig. 101, the general statement can be made that

$$\bar{V} = \bar{V}_1 + \bar{V}_2$$

or as

$$V\underline{/h} = V_1\underline{/0} + V_2\underline{/a}$$

in which the angles deboted by h and a will be expressed in either degrees or radians. In the last equation, $\bar{V}_1$ is taken to be the reference vector at "zero degrees or radians," with the total resultant voltage having a magnitude of V volts at an angle h relative to the $\bar{V}_1$ vector.

In any given case, the value of a resultant vector, $\bar{V} = V\underline{/\theta}$, is equal to the vector sum of all the horizontal and vertical components of all the individual vectors involved, as explained in example 1.

The above statements made concerning voltages apply, of course, to currents; thus

$$\bar{I} = \bar{I}_1 + \bar{I}_2 + \bar{I}_3 + \cdots$$

where $\bar{I}$ is the overall resultant current, equal to the vector sum of all the component vector currents $\bar{I}_1$, $\bar{I}_2$, $\bar{I}_3$, and so on, the magnitudes of all currents being in rms amperes, as usual.

Next, the basic OHM'S LAW, first stated for the dc case in section 2.3, now becomes, for the ac case, the vector relationship

$$\bar{I} = \frac{\bar{V}}{R} \tag{111}$$

where $\bar{I}$ = rms vector current, $\bar{V}$ = applied rms vector voltage, and R is resistance in ohms. R is a scalar quantity, there being no sense of direction associated with resistance.

It should be noted that eq. (111) applies to a circuit, such as Fig. 91, in which *a number of different generators*, all of the same frequency but having various phase relations with respect to one another, are applied to a common load resistance of R ohms. (Equation (107) in section 5.5 can be written in the simpler form $I = V/R$, because just *one* generator of voltage V is being considered in the discussion of this equation.)

Problem 72
Three series-connected ac generators, all of the same frequency, having rms vector voltages of $\bar{V}_1$, $\bar{V}_2$, $\bar{V}_3$, work into a load resistance of $R = 25$ ohms, as shown in Fig. 102.

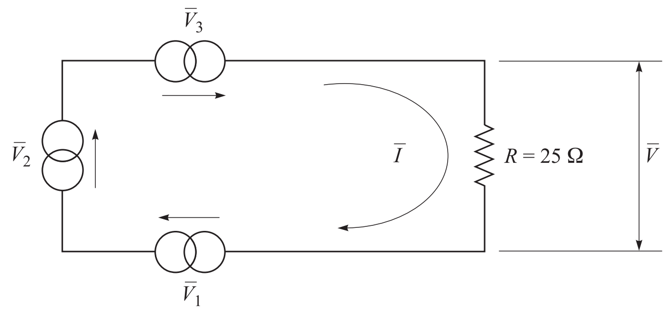


Fig. 102

It is given that, in rms volts,

$$\bar{V}_1 = 65\underline{/0^\circ} \qquad \bar{V}_2 = 90\underline{/60^\circ} \qquad \bar{V}_3 = 75\underline{/150^\circ}$$

(a) Using the given values, write the value of the output voltage $\bar{V}$ in "polar" form.
(b) An ac voltmeter placed across R would read __________ rms volts.
(c) For the given values, write the value of the current $\bar{I}$ in polar form. An ac ammeter, placed in series with R, would read __________ rms amperes.
(d) In the figure, is the current $\bar{I}$ "in phase" with the resultant output voltage $\bar{V}$?
(e) In the figure, is the current $\bar{I}$ in phase with any of the generator voltages?

5.7 Power Calculations

Let us begin by returning briefly to eqs. (94), (95), (96), and (97) in section 5.5.

First of all, eqs. (94) and (95) show that, in that section, we dealt ONLY with the case in which the "current wave" (eq. (95)) is exactly IN PHASE with the "voltage wave" (eq. (94)). Hence eq. (97), and therefore also eq. (104),

$$P = VI \text{ watts}$$

are true ONLY if the current wave is exactly IN PHASE with the voltage wave. Likewise, eqs. (105), (106), (107) are true only for the same condition. Hence, in this simple case we often don't bother to denote the rms vector values by $\bar{V}$ and $\bar{I}$, but just write V and I; thus, in section 5.5 we wrote $I = V/R$ (eq. (107)), and so on.

Now, however, let us consider the case where the voltage and current waves are NOT exactly in phase with each other; to indicate this, let us now write eqs. (94) and (95) as

$$v = V_\text{p} \sin \omega t \tag{112}$$

$$i = I_\text{p} \sin(\omega t - \theta) \tag{113}$$

which is the algebraic way of showing that the current wave "lags" the voltage wave by θ radians (ωt is the angular amount in radians).

Thus, for this condition, the original eq. (97) now becomes

$$p = V_p I_p (\sin \omega t) \sin(\omega t - \theta) \tag{114}$$

At this point it will be helpful to make use of the trigonometric identity*

$$\sin(x + y) = \sin x \cos y + \cos x \sin y$$

which, upon setting $x = \omega t$ and $y = -\theta$, and then making use of the identities $\cos(-h) = \cos h$ and $\sin(-h) = -\sin h$ (eqs. (77) and (76) in section 5.3), becomes

$$\sin(\omega t - \theta) = \sin \omega t \cos \theta - \cos \omega t \sin \theta$$

in which θ is a constant phase angle between the voltage and current waves (and thus $\cos\theta$ and $\sin\theta$ are also constant quantities in any given case). Now substitute the above identity for $\sin(\omega t - \theta)$ into eq. (114); doing this, eq. (114) becomes

$$p = V_p I_p [\sin^2 \omega t \cos \theta - \sin \omega t \cos \omega t \sin \theta] \tag{115}$$

Now, making use of eq. (98) in section 5.5, we have

$$\sin^2 \omega t = \tfrac{1}{2}(1 - \cos 2\omega t)$$

and, upon making this substitution into eq. (115), you should find that

$$p = V_p I_p [\tfrac{1}{2} \cos \theta - \tfrac{1}{2} \cos 2\omega t \cos \theta - \sin \omega t \cos \omega t \sin \theta] \tag{116}$$

in which p is "instantaneous" power. But, as explained in section 5.5, we are not interested in instantaneous power; instead, we are interested in finding the AVERAGE POWER produced by the generator voltage, $v = V_p \sin \omega t$, over ONE COMPLETE CYCLE, such as from $\omega t = 0$ to $\omega t = 2\pi$. Hence our discussion, from this point on, exactly parallels the discussion following eq. (99), and leads to the conclusion that the AVERAGE POWER P produced in eq. (116) is equal to

$$P = \frac{V_p I_p}{2} \cos \theta = \frac{V_p}{\sqrt{2}} \frac{I_p}{\sqrt{2}} \cos \theta$$

which, you'll notice, is the same as eq. (101) in section 5.5 EXCEPT that now we have the constant multiplier "$\cos\theta$." Thus, using the definitions of eqs. (102) and (103), we have that

$$\boxed{P = VI \cos \theta} \qquad \text{watts} \tag{117}$$

where P is average power in watts, V and I are rms values of voltage and current, and θ is the phase angle between the sinusoidal voltage and current waves. It makes no difference, of course, whether the current wave "leads" or "lags" the voltage wave, because $\cos(-\theta) = \cos \theta$.

Equation (117) says that the scalar quantity "power" is equal to the product of the magnitudes of the rms values $\bar{V}$ and $\bar{I}$ and the cosine of the phase angle between $\bar{V}$ and $\bar{I}$. This can be shown in connection with the vector diagram for $\bar{V}$ and $\bar{I}$ as shown in Fig. 103.

Thus the statement is made that POWER in an ac circuit is equal to the magnitude of voltage times the magnitude of the component of current that is "in phase" with the voltage vector.

---

* See note 7 in Appendix.

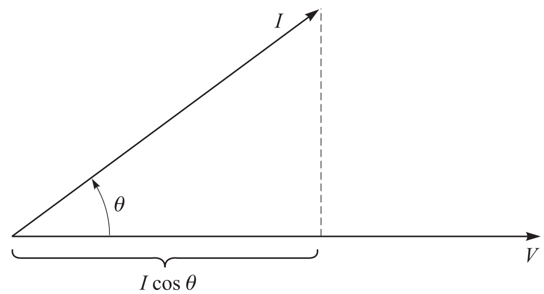


Fig. 103. $\boldsymbol{P = VI \cos \theta}$.

Equation (117) gives power in terms of V and I, that is, in terms of the magnitudes of the rms values of voltage and current and the angle θ between the $\bar{V}$ and $\bar{I}$ vectors.

Now suppose that V and I are the rms values of voltage across, and current through, a pure *resistance* of R ohms; in that case $\theta = 0$ (see discussion for Fig. 98) and, since $\cos 0 = 1$, eq. (117) becomes, if V is the voltage across and I the current through, a *resistance* of R ohms,

$$P = VI \tag{118}$$

or, since $V = RI$,

$$P = I^2 R \tag{119}$$

or, since $I = V/R$,

$$P = V^2/R \tag{120}$$

where, as before, P is the average power in watts.

Before going on let us pause, just briefly, to comment on vector diagram notation. In a vector diagram the *lengths* of the vectors represent the *magnitudes of the rms values* of voltage and current, and the "phase angles" are represented graphically by actually drawing the vectors at the specified angles with respect to each other and to the reference vector.

Thus, in the vector diagram of Fig. 103, V and I (not $\bar{V}$ and $\bar{I}$) denote the MAGNITUDES of the rms values of the vector voltage $\bar{V}$ and the vector current $\bar{I}$, while the *phase shift* is shown directly on the diagram by drawing the vector lengths at the required angle θ with respect to a reference vector or reference line. (The notation is further illustrated in Fig. 101.)

Problem 73

This is a continuation of problem 72, Fig. 102, using all the same values as given in that problem.

(a) The power produced in the 25-ohm resistive load is __________ watts.

(b) The power produced by each individual generator in Fig. 102 is as follows:

power produced by generator of voltage V_1 is __________ watts,
power produced by generator of voltage V_2 is __________ watts,
power produced by generator of voltage V_3 is __________ watts.

(c) The sum of the three answers found in part (b) must be equal to the answer found in part (a); check to see that your answers satisfy this requirement.

(d) Make a freehand sketch of the vector diagram showing the relationships in Fig. 102. Show and label the generator voltages, the output voltage, the current, and the various phase angles involved.

NOTE OF CAUTION

The "principle of superposition" states that, in a network composed of linear bilateral elements and several generators, the *total* CURRENT at any point in the network is equal to the sum of the currents due to each generator considered separately, the other generators being replaced by their internal resistances.*

The principle of superposition, however, cannot, in general, be applied in the same way to POWER calculations as it is in current calculations. This is because power is proportional to the SQUARE of current, $P = I^2 R$. Thus the *total power* produced in a resistance R, due to the presence of several different components of current in R, is *not* equal to the sum of the powers due to each current component considered separately as if the other components were absent. (This assumes all generators have the same frequency.)

For example, suppose a total current I is equal to the sum of two separate current components I_1 and I_2; that is, $I = I_1 + I_2$. The total power P_T produced in R is equal to

$$P_T = I^2 R = (I_1 + I_2)^2 R$$

thus

$$P_T = I_1^2 R + I_2^2 R + 2 I_1 I_2 R \tag{121}$$

Note that IF the principle of superposition applied to power calculations the above answer would be $P_T = I_1^2 R + I_2^2 R$, which, however, is *not* the same as the correct answer given by eq. (121). Hence the *total current* must be used when making power calculations. Thus in problem 73 total current of 4.957 amperes (4.957 A) is used in the power calculations.

5.8 Application of Loop Currents

The method of "loop currents," introduced in section 4.4, applies to ac networks as well as dc networks. All we need do is properly label the ac network, then apply the rules laid down in that section.

To illustrate the procedure, let us now work through the solution of an ac network problem using the loop current method. In doing this, we'll go into much more detail than a person normally would in working such a problem; we do this, of course, to emphasize the fundamental ideas involved.

Thus, for purposes of basic explanation, let us begin with our example network expressed in terms of the fundamental variable *time*, t, as shown in Fig. 104, in which resistance values are in ohms. The voltages and currents are sinusoidal, as always, with the voltage and current arrows having the meanings described in the discussion following eq. (108).

In the figure, note that only the (peak) generator voltages, including their polarity arrows and their phase angles, and the resistance values are given. The *unknowns* are thus the "peak values," I_{p1} and I_{p2}, of the current waves and their corresponding phase angles a and b, and the peak value V_p of the output voltage wave v_o and its phase angle c.† All phase angles are to be stated with reference to the voltage wave, $43 \sin \omega t$, which, since it passes through the origin of the horizontal time axis, will be said to have "zero" phase

* See problem 50 and also the footnote in section 4.7.

† Angle c, in this case, will be equal to angle b ($c = b$) because, as pointed out following Fig. 98, there is zero phase shift between the voltage across, and the current through, a resistance of R ohms.

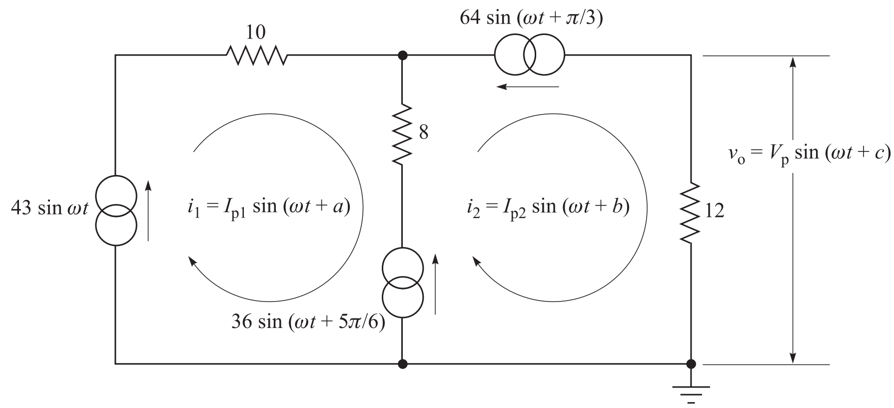


Fig. 104

shift. (Since the angular amount ωt is in *radians*, the phase angles a, b, c, as they appear in the figure, must also be expressed in radians.)

In the figure, let the voltage appearing across the 12-ohm resistor be the "output voltage" of the network. Thus v_o is the *instantaneous value* of the output voltage. As time continues to increase, v_o continually swings back and forth, sinusoidally, between the positive and negative peak values V_p and $-V_p$.

Now let the basic PROBLEM be to FIND THE MAGNITUDE OF THE RMS VOLTAGE that would appear across the 12-ohm load resistor (this being the value of voltage that would be read by an *ac voltmeter* connected across the 12-ohm resistance).

To solve the above problem by the method of loop currents, let us begin by returning to the "three-step" procedure outlined in section 4.4.

The three-step rule, as given in Chap. 4, was applied to dc circuits; it should be emphasized, however, that the rules apply AT ANY AND ALL INSTANTS OF TIME in any network. This is true because Kirchhoff's current and voltage laws (sections 4.2 and 4.3) are, and must be, satisfied *at any and all instants of time* in any network.

Now, in regard to our problem here and the above-mentioned "three steps," notice that we automatically satisfied "step I" when we drew and labeled the "current arrows" in Fig. 104.

Next, "step II" defines the rules to be used in writing the VOLTAGE EQUATIONS around each of the n loops in a network ($n = 2$ in Fig. 104). The *same rules*, given in the original statement of step II in section 4.4, will also apply to the ac circuit of Fig. 104 except that now the "plus and minus" battery polarity marks will be replaced by *voltage arrows* associated with each ac generator (see Fig. 89). Thus the original statement of step II would now, for the alternating-current (ac) case, be stated as follows.

All generator voltages will be written on the right-hand sides of the equations, and will be considered *positive* if we go through a generator *with the voltage arrow*, but *negative* if we go through the generator *against the voltage arrow*. The original statement about "voltage drops" across resistors will remain unchanged.

Thus, applying these step II rules to the ac network of Fig. 104, we have the following two voltage equations, which are valid at all instants of time, first around the left-hand loop (eq. (122)), then around the right-hand loop (eq. (123)).

$$18I_{p1}\sin(\omega t + a) - 8I_{p2}\sin(\omega t + b) = 43\sin\omega t - 36\sin(\omega t + 5\pi/6) \tag{122}$$

$$-8I_{p1}\sin(\omega t + a) + 20I_{p2}\sin(\omega t + b) = 36\sin(\omega t + 5\pi/6) - 64\sin(\omega t + \pi/3) \tag{123}$$

In the above equations the constant GIVEN values are, first, the resistance values in ohms (on the left-hand sides of the equations) and, second, the peak sinusoidal generator voltages and their phase angles (on the right-hand sides of the equations). The constant UNKNOWN values are the peak values of the sinusoidal current waves and their phase angles a and b. Note that the INDEPENDENT VARIABLE is *time*, t; the equations are thus *instantaneous* relationships, valid for all values of time, from $t = 0$ to any value whatever.

Hence (in the manner of Figs. 92 and 93) *each term* in the equations can be represented by a ''phasor,'' all phasors rotating at the same ''angular speed'' of ω radians per second. Thus the *total angle* generated by a phasor in t seconds is ωt radians, or, if a phase angle a must be taken into account, $\omega t + a$ radians (such as is illustrated in Fig. 93). It should be noted that although phasors are simply ''rotating line segments'' they can, nevertheless, be extremely helpful in understanding and visualizing the behavior of sinusoidal waves of voltage and current in ac networks.

Now, at this point in our work, we suggest that, before going on, the student should carefully REVIEW ALL THE DISCUSSIONS given in connection with eq. (108) and Figs. 92 through 97, including eqs. 109 and 110.

In conducting the above-proposed review, the reader may have noticed that we made use of two very important properties possessed by sinusoidal waves of the same frequency. Both properties concern the nature of the *sum* of two (or more) such waves. The properties were applied in section 5.6, perhaps without sufficient emphasis and explanation; let us, therefore, take the time now to briefly discuss them in more detail, beginning with the first property, which can be stated as follows:

> The SUM of two or more sinusoidal waves of the same frequency is equal to A SINGLE SINUSOIDAL WAVE of the same frequency.

Thus, if $A \sin \omega t$ is one sine wave and $B \sin(\omega t + a)$ is another sine wave (where $\omega = 2\pi f$, in which frequency f has the same value in both waves), the foregoing statement says that

$$A \sin \omega t + B \sin(\omega t + a) = C \sin(\omega t + b) \tag{124}$$

where A, B, and C are peak values and a and b are phase angles. A proof of eq. (124) is given in the Appendix.* It follows that the sum of *any number* of such waves is also equal to a single sinusoidal wave (because any two can be combined into a single wave, which can then be combined with a third into a single wave, and so on).

The second property of sinusoidal waves that we wish to emphasize is the fact that the ''peak'' values, and also the ''rms'' values, of different sinusoidal waves (all of the same frequency) can be added together vectorially AS IF THEY WERE ORDINARY VECTOR QUANTITIES (see Figs. 92 through 101, and the discussions given with these figures). Property two can thus be summarized in the statement:†

> The *rms value* of the SUM *of two or more sinusoidal waves of the* SAME FREQUENCY is equal to *the vector sum of the rms values of the individual sinusoids.*

* See note 9 in Appendix.
† See note 10 in Appendix.

In the above, the MAGNITUDE of each vector quantity represents the MAGNITUDE OF THE RMS VALUE of the corresponding voltage or current wave, while the ANGLE of the vector represents the relative PHASE SHIFT of the sinusoidal wave. (For a note on ''vector diagram'' notation, see paragraph just prior to problem 73.)

Let us now apply the foregoing to Fig. 104, in which the PROBLEM is to ''find the magnitude of the rms voltage across the 12-ohm resistor.'' Again, in doing this we'll go into much more detail than a person normally would for a problem such as this; we do this, of course, for the purpose of explanation.

Let us begin with the *time-dependent* eqs. (122) and (123), in which the terms can be thought of as being generated by rotating voltage and current phasors, all rotating at the same angular speed of ω rad/sec. Since, however, we're not interested in finding ''instantaneous'' values we can *disregard the rotary motion of the phasors* and concentrate on just the constant unknown peak values, I_{p1} and I_{p2}, and the constant phase angles a and b. Doing this *eliminates the variable time* and transforms eqs. (122) and (123) into the *vector equations*

$$18I_{p1}\underline{/a} - 8I_{p2}\underline{/b} = 43\underline{/0} - 36\underline{/5\pi/6} \tag{124a}$$

$$-8I_{p1}\underline{/a} + 20I_{p2}\underline{/b} = 36\underline{/5\pi/6} - 64\underline{/\pi/3} \tag{125}$$

in which the phase angles are in radians. As previously mentioned, in practical engineering work we normally use ''rms'' values of voltage and current instead of ''peak'' values. (This is reasonable, because rms values must be used in all the equations for calculating average POWER in ac circuits.)

Hereafter, therefore, we'll write our equations *in terms of rms values* instead of peak values; in terms of *notation* we'll indicate this by dropping the subscript ''p'' (which indicates ''peak''). Thus, since I_{p1} and I_{p2} represent peak magnitudes of current, then, dropping the p's, I_1 and I_2 would represent *rms* magnitudes of the currents. Thus, using eq. (103), we have that

$$I_{p1} = 1.4142I_1 \qquad \text{and} \qquad I_{p2} = 1.4142I_2$$

Let us therefore make the above substitutions into eqs. (124a) and (125); doing this, and also converting the phase angles to degrees (eq. (85)), allows us to rewrite eqs. (124a) and (125) as

$$25.456I_1\underline{/a} - 11.314I_2\underline{/b} = 43\underline{/0^\circ} - 36\underline{/150^\circ} \tag{126}$$

$$-11.314I_1\underline{/a} + 28.284I_2\underline{/b} = 36\underline{/150^\circ} - 64\underline{/60^\circ} \tag{127}$$

in which the peak voltage $43\underline{/0^\circ}$ is the reference vector, with peak voltages on the left-hand sides of the equations all being expressed in terms of *rms* currents I_1 and I_2. Next, as an aid in understanding the details which follow, let us adopt the notation below, which we'll lump together as ''eq. (128)''; thus

$$\left.\begin{aligned} I_1\underline{/a} &= \bar{I}_1 \qquad 43\underline{/0^\circ} = \bar{V}_1 \qquad 64\underline{/60^\circ} = \bar{V}_3 \\ I_2\underline{/b} &= \bar{I}_2 \qquad 36\underline{/150^\circ} = \bar{V}_2 \end{aligned}\right\} \tag{128}$$

in which, for convenience in this case only, we're using $\bar{V}_1$, $\bar{V}_2$, and $\bar{V}_3$, to denote peak values of voltage. Now make all the above substitutions into eqs. (126) and (127); doing this gives the following set of simultaneous equations:

$$25.456\bar{I}_1 - 11.314\bar{I}_2 = \bar{V}_1 - \bar{V}_2 \tag{129}$$

$$-11.314\bar{I}_1 + 28.284\bar{I}_2 = \bar{V}_2 - \bar{V}_3 \tag{130}$$

in which $\bar{I}_1$ and $\bar{I}_2$ are unknown rms vector currents and $\bar{V}_1$, $\bar{V}_2$, $\bar{V}_3$ are known (given) peak values of the given generator voltages.

At this point let us pause to note that the original Fig. 104, which is expressed in terms of the fundamental independent variable time, can now be redrawn in terms of the *vector quantities* of eqs. (129) and (130), as in Fig. 105.

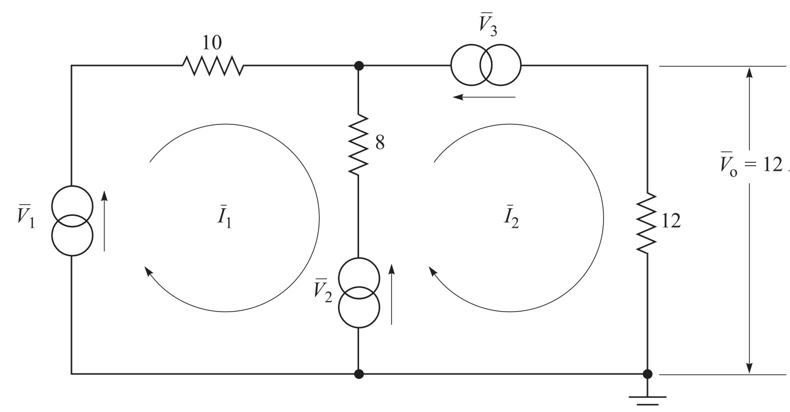


Fig. 105

In the figure, $\bar{I}_2$ is the vector *rms current* through the 12-ohm resistor; hence, $\bar{V}_o$ is the vector *rms output voltage*, and therefore, by Ohm's law,

$$\text{magnitude of rms output voltage} = |\bar{V}_o| = 12|\bar{I}_2| \tag{131}$$

Thus, to find the required ANSWER to our problem (find the voltmeter reading across the 12-ohm resistance), all we now need to do is to solve eqs. (129) and (130) for $\bar{I}_2$. The details of the solution of these equations for $\bar{I}_2$ by determinants are as follows, using the standard procedure of section 3.5.

$$\bar{I}_2 = \frac{\begin{vmatrix} 25.456 & (\bar{V}_1 - \bar{V}_2) \\ -11.314 & (\bar{V}_2 - \bar{V}_3) \end{vmatrix}}{\begin{vmatrix} 25.456 & -11.314 \\ -11.314 & 28.284 \end{vmatrix}} = \frac{11.314\bar{V}_1 + 14.142\bar{V}_2 - 25.456\bar{V}_3}{591.991} \tag{132}$$

The next step is to find the SUM OF THE VECTOR QUANTITIES in the numerator of the fraction to the right above.

This can be done by recalling that the horizontal component V_h and the vertical component V_v of the *resultant sum* of a number of vectors is equal, respectively, to the sum of the horizontal components and the sum of the vertical components of the individual vectors (see discussion with Figs. 96 and 97). Making use of this fact, and eqs. (109) and (110) (and (128)), we have, for the case of eq. (132) (note angle in 3rd gradient),

$$11.314\bar{V}_1 + 14.142\bar{V}_2 - 25.456\bar{V}_3 = 1388.719\underline{/236.38^\circ}$$

hence

$$|\bar{I}_2| = \frac{1388.719}{591.991} = 2.346 \text{ amperes}$$

thus

$$|\bar{V}_o| = (12)(2.346) = 28.150 \text{ volts}, \quad \textit{answer}.$$

Some comments regarding the foregoing problem are as follows.

Note that we began with Fig. 104, to which eqs. (122) and (123) apply, these being expressed in terms of the independent variable *time*, t.

Then, by requiring that only *rms values* need be calculated, we were able to *eliminate the variable time*, thus producing Fig. 105 and the corresponding vector eqs. (129) and (130). (It should be mentioned that an experienced engineer would not generally bother with Fig. 104 at all, but would start directly with Fig. 105 and eqs. (129) and (130).)

We might also comment once again on the sense of *direction* given to the voltage and current arrows in a network.

In Fig. 104, for example, $43\sin\omega t$ was chosen to be the REFERENCE VOLTAGE WAVEFORM, with a "voltage arrow" drawn beside the generator. The arrow is drawn with the meaning of Fig. 89, denoting the polarity "$-$ to $+$" during the times that $43\sin\omega t$ is positive in value. Next in Fig. 104 consider, for example, the voltage waveform having the peak value of 64 volts. It is GIVEN in the problem that this voltage waveform *leads* the reference voltage wave by $\pi/3$ radians in the sense "$-$ to $+$" indicated by the voltage arrow given alongside the generator.

Next, the CURRENT ARROWS define the sense in which "positive currents" flow in a network. In our work we'll usually draw all the current arrows around the loops in the same sense (let us say all clockwise), because this will tend to produce voltage equations that are somewhat more symmetrical in form. When a "negative current," $-I\sin(\omega t + a)$, occurs in a solution, it means that that particular current wave actually flows in a counter-clockwise sense during the times that $I\sin(\omega t + a)$ is positive in value.

Problem 74

In Fig. 105, find the magnitude of $\bar{V}_o$ if it had been given that the voltage arrow for $\bar{V}_3$ should point from left to right instead of from right to left.

(*Answer*: $|\bar{V}_o| = 37.992$ volts)

CHAPTER 6

Algebra of Complex Numbers

In this chapter we study the "algebra of complex numbers." This is a subject of great usefulness in all branches of engineering, being especially important in electrical and electronic engineering. We'll find that the algebra is not difficult and is, in itself, a most interesting study.

6.1 Imaginary Numbers

Let us begin with a bit of history. For a long time only ordinary positive and negative numbers were used, and the rules for working with these numbers were laid down and firmly established.

Later, positive and negative exponents were invented, and fitted into the established rules of mathematics. New rules for working with exponents had to be laid down, and these rules had to be in harmony with other rules previously established.

Then the "square root" of a positive number was defined. Thus, if the "square root of y" is equal to "x," this is denoted by the symbol

$$\sqrt{y} = x$$

which is defined to mean that

$$y = x^2$$

It was immediately noticed that the square root of a positive number must have TWO different values, equal in magnitude but opposite in sign. This had to be true because of a previous rule that had been laid down and established; this was the rule that the product of TWO POSITIVE NUMBERS or TWO NEGATIVE NUMBERS is always a POSITIVE number. Thus it had to be written that, for example,

$$\sqrt{1} = \pm 1$$

because

$$1 = (+1)^2$$

and

$$1 = (-1)^2$$

This naturally brought up the question, "Does the square root of a NEGATIVE number have a meaning?". What, for example, is the square root of *minus 1*? Thus, if we let x denote the value of the "square root of minus 1" we must write that

$$\sqrt{-1} = x$$

which, by definition, means that

$$-1 = x^2$$

Notice that x *cannot* be equal to $+1$ and it *cannot* be equal to -1, because neither of these numbers *when squared* is equal to -1. It was therefore understandably stated, by the early investigators, that the "square root of a negative number *does not exist*."

As time passed, however, it became clear that a fully unified system of algebra was not possible unless the square roots of negative numbers were accepted as truly being a *new kind of number*. Such numbers are called "imaginary" numbers, the name simply reflecting the historical fact that they were originally thought not to exist.

Thus we now have *two sets of numbers*, one being the set of ordinary positive and negative "real" numbers, the other being the set of positive and negative "imaginary" numbers.

In the set of real numbers *each individual number* is composed of a certain number of digits arranged in a certain sequence, producing a unique number, different from any other. Thus we have the positive real numbers, 0, 1, 2, 3, ... 25, 26, 27, ..., and so on, endlessly. Then, for each positive real number there is a corresponding NEGATIVE real number of the same magnitude.

In the real number system the number "1" is the *basic unit of measure or value*. Thus any other real number can be regarded as a multiple of 1; the number 23, for example, is equal to "23 times the value represented by the basic unit 1." We also note that

$$1 \times 1 = 1^2 = 1$$

Exactly the same procedure is used to denote the value of an imaginary number, except that now the *basic* IMAGINARY UNIT *of value* will be denoted by the letter "*j*," which is defined to be equal to *the* SQUARE ROOT OF MINUS ONE; that is, by definition,

$$\text{basic IMAGINARY UNIT} = j = \sqrt{-1} \tag{133}$$

Now let us see WHY the above definition makes sense. To do this, let us consider, again, the previously discussed equation

$$x^2 = -1 \tag{134}$$

Note that this equation has *no solution in the real number system*, because the SQUARE of any positive or negative real number is a POSITIVE real number. Thus x *cannot* be equal to *either* of the real numbers $+1$ or -1, because their *square* is equal to $+1$, and $+1$ is, of course, *not* equal to -1. Thus there is *no real value of* x that "satisfies" the requirement of eq. (134). A solution can, however, be obtained in terms of imaginary numbers, as follows.

By the definition of eq. (133)

$$j = \sqrt{-1}$$

Squaring to remove the radical sign

$$j^2 = -1$$

Hence, setting $x = j$ in eq. (134) gives the *true identity*, $-1 = -1$, thus proving that $x = j = \sqrt{-1}$ is a *valid solution* to eq. (134). This result is possible because we have accepted the existence of imaginary numbers and the basic definition of eq. (133).

Thus we have found that $x = j$ is a valid solution to eq. (134); it is important, however, to note that $x = -j$ is *also* a valid solution to eq. (134). Thus, setting $x = -j$ in eq. (134) gives

$$x^2 = (-j)^2 = (-1)^2 j^2 = j^2 = -1$$

which again gives the true identity, $-1 = -1$. Thus eq. (134) has TWO solutions (two "roots"), $x = j$ and $x = -j$. This is, of course, to be expected, because eq. (134) is a "second degree" algebraic equation. Thus to "find the value of x" in eq. (134) the procedure is to "take the square root of both sides," thus getting

$$x = \pm\sqrt{-1} = \pm j$$

Since j denotes the *unit* imaginary value, the general value of an imaginary number is indicated by writing "aj," where a can be any positive or negative real number. The real coefficient a thus expresses the value of the imaginary number relative to the basic imaginary unit j. Thus we have imaginary numbers such as $8j$, $150j$, and so on. It should also be mentioned that the product aj is "commutative," that is, that $aj = ja$.

We should also note that for each POSITIVE imaginary number there is a corresponding NEGATIVE imaginary number; thus, corresponding to $25j$ there exists $-25j$, and so on.

Since "1" is the unit value in the real number system, and "j" is the unit value in the imaginary number system, it's interesting to make the comparison that

$$1 \times 1 = 1^2 = 1$$

and

$$j \times j = j^2 = -1$$

Thus *j squared is equal to minus one*, a fact that we are of course already aware of, and one that we will often make use of in our future work.

Now let us continue on, and discover some very useful and surprising facts about imaginary numbers. To begin, we already know that

$$\boxed{\begin{aligned} j &= \sqrt{-1} && (135) \\ j^2 &= -1 && (136) \end{aligned}}$$

Let us remember that when "equal base numbers" are multiplied together the exponent of the product is equal to the *sum* of the exponents of the individual factors, that is

$$a^x a^y = a^{x+y}$$

Remembering the above fact, and also keeping eq. (136) in mind, we can now show that since $j^3 = j^2 j = -j$, we have

$$\boxed{j^3 = -j} \tag{137}$$

Next, since $j^4 = j^2 j^2 = (-1)(-1) = +1$, we have

$$\boxed{j^4 = +1} \tag{138}$$

To continue:

$$j^5 = j^4 j = j \qquad \text{(because } j^4 = 1\text{)}$$
$$j^6 = j^5 j = jj = j^2 = -1 \qquad \text{(because } j^5 = j\text{)}$$
$$j^7 = j^6 j = -j \qquad \text{(because } j^6 = -1\text{)}$$
$$j^8 = j^7 j = -jj = -j^2 = +1 \qquad \text{(because } j^7 = -j\text{)}$$

Let us now bring together all of the foregoing equations; thus

$$\begin{array}{ll} j^1 = j & j^5 = j \\ j^2 = -1 & j^6 = -1 \\ j^3 = -j & j^7 = -j \\ j^4 = +1 & j^8 = +1 \end{array}$$

and so on, endlessly.

Thus we see that the POWERS of the imaginary unit j consist of ONLY THE FOUR DIFFERENT VALUES, j, -1, $-j$, $+1$, and these four values are repeated over and over, in regular sequence, as shown above. Hence, basically all we need to remember are the *first four relationships*, thus

$$\boxed{\begin{array}{ll} j = j & j^3 = -j \\ j^2 = -1 & j^4 = +1 \end{array}}$$

Example 1

Remembering that $\sqrt{ab} = \sqrt{a}\sqrt{b}$, write each of the following in terms of the imaginary unit.

(a) $\sqrt{-36}$ (b) $\sqrt{-25x^6}$ (c) $\sqrt{-85y^{-2}}$

Solutions

(a) $\sqrt{-36} = \sqrt{(-1)(36)} = \sqrt{-1}\sqrt{36} = \pm j6$, *answer*.

(b) $\sqrt{-25x^6} = \sqrt{-1}\sqrt{25}\sqrt{x^6} = \pm j5x^3$, *answer*.

In getting the above answer we made use of fractional exponents and the basic relationship, $(x^a)^b = x^{ab}$; thus, $\sqrt{x^6} = (x^6)^{1/2} = x^3$.

(c) $\sqrt{-85y^{-2}} = \pm j(9.2196)(y^{-2})^{1/2} = \pm j\,9.2196\,y^{-1}$, *answer*, or
$= \pm j\,9.2196/y$, *answer*.

Example 2

Simplify each of the following.

(a) j^{20} (b) $-10j^{47}$ (c) $\dfrac{1}{j}$

Solutions

(a) The four possible values of the powers of j repeat themselves over and over, endlessly, in the order $j, -1, -j, 1$; since 4 goes into 20 exactly five times, it follows that

$$j^{20} = j^4 = 1, \quad \textit{answer}.$$

(b) Since 4 goes into 47 "11 times with 3 left over," we have that

$$-10\,j^{47} = -10\,j^3 = -10(-j) = 10j, \quad \textit{answer}.$$

(c) Multiply the numerator and denominator by j; thus

$$\frac{j}{jj} = \frac{j}{j^2} = \frac{j}{-1} = -j, \quad \textit{answer}.$$

Thus we have the very useful fact that

$$\boxed{\frac{1}{j} = -j} \tag{139}$$

The fact that "one over j is equal to minus j" is used so often that it should be committed to memory.

Problem 75
Express each of the following in terms of the imaginary unit j.

(a) $\sqrt{-144}$ (b) $\sqrt{-100x^4\,y^{-10}}$ (c) $\left(\dfrac{-x^2}{4y^2z^4}\right)^{1/2}$

Problem 76
Simplify each of the following as much as possible (the dot means "times").

(a) $-j \cdot j^2$ (d) $-j^{342}$ (f) $\dfrac{1}{j^{34}}$

(b) $j^5 \cdot j^8$ (e) $\dfrac{1}{j^3}$ (g) j^{-17}

(c) j^{31}

Problem 77
Solve the following for the unknown values of x.

(a) $x^2 + 16 = 0$ (c) $6x^2 + 94.8 = 0$

(b) $x^2 - 900 = 0$

6.2 Complex Numbers. Addition and Multiplication

We begin with the definition that a "complex number" is the algebraic SUM of a REAL number and an IMAGINARY number. Thus complex numbers have the general form

$$a + jb$$

where a is the real part and jb is the imaginary part of the complex number, with a and b both being real numbers. In working with complex numbers it's customary to write the real part first; thus we generally write $a + jb$, instead of $jb + a$.

The first rule, in the algebra of complex numbers, concerns the SUM of two or more such numbers, and can be stated in the following way.

REAL PART OF SUM = SUM OF THE REAL PARTS of the numbers
IMAGINARY PART OF SUM
= SUM OF THE IMAGINARY PARTS of the numbers

Example
Find the sum of the complex numbers $(6 + j3)$, $(10 - j7)$, and $(-4 + j9)$.

Solution

$$(6 + 10 - 4) + j(3 - 7 + 9) = 12 + j5, \quad \textit{answer}.$$

Next, the PRODUCT of two complex numbers is found in exactly the same way as in the ordinary algebra of real numbers, except we must remember that $j^2 = -1$.

Example 1
Find the product of the two complex numbers $(5 + j3)$ and $(2 - j5)$.

Solution
Using the ordinary "four-step rule" for multiplying two binomials, we have

$$\begin{aligned}(5 + j3)(2 - j5) &= 10 - j25 + j6 + 15 \\ &= (10 + 15) + j(-25 + 6) \\ &= 25 - j19, \quad \textit{answer}.\end{aligned}$$

Example 2
Find the value of $j(6 + j)(3 - j4)$.

Solution
Here we have to multiply three quantities together, the three quantities being

1. the unit imaginary number j
2. the complex number $(6 + j)$
3. the complex number $(3 - j4)$

The procedure in such a case is to FIRST find the product of ANY TWO of the three factors, then multiply that result by the remaining factor. In the above case let's first multiply factors (1) and (2) together, giving $j(6+j) = (-1+j6)$, which we must now multiply by factor (3), thus getting

$$\begin{aligned}(-1+j6)\,(3-j4) &= -3+j4+j18+24\\ &= 21+j22, \quad \textit{answer}.\end{aligned}$$

Problem 78

$$(6+j5)-(8+j4)+(4-j3)=$$

Problem 79

$$j5-7+j^3+1-j10-j^2 4+3+10j^{100}=$$

Problem 80

$$(2-j3)\,(6-j)=$$

Problem 81

$$(1+j)\,(1+j2)\,(-3+j5)=$$

Problem 82
Given that a, b, c, and d are real numbers, $(a+jb)\,(c+jd)=$

Problem 83

$$(6+j12)^2=$$

Problem 84

$$(1+j)^5=$$

6.3 Conjugates and Division of Complex Numbers

Two complex numbers that differ ONLY IN THE SIGNS OF THEIR IMAGINARY PARTS are called "conjugate complex numbers."

Thus $(a+jb)$ and $(a-jb)$ are conjugate complex numbers, each being the "conjugate" of the other.

The important fact that we are concerned with here is that the PRODUCT of two conjugate complex numbers is always a POSITIVE REAL NUMBER; thus

$$(a+jb)\,(a-jb)=a^2+b^2 \tag{140}$$

which you should verify by direct multiplication.

The relationship of eq. (140) is especially useful in finding the quotient of two complex numbers; this can be illustrated with the aid of eq. (141) below:

$$\frac{c+jd}{a+jb}=A+jB \tag{141}$$

in which all the letters (except j) represent positive or negative real numbers.

Eq. (141) says that the quotient of two complex numbers, $c + jd$ divided by $a + jb$, is equal to a SINGLE COMPLEX NUMBER, which we show as $A + jB$ in eq. (141). The PROBLEM is, given the complex numbers $c + jd$ and $a + jb$, *find the values of* A *and* B in eq. (141). Fortunately, this can be done by *multiplying the numerator and denominator of the fraction by the* CONJUGATE OF THE DENOMINATOR. To show why this is true, let us apply the rule to the fraction on the left-hand side of eq. (141); doing this, and making use of eq. (140), we find that eq. (141) becomes (on the left-hand side)

$$\frac{(c + jd)\,(a - jb)}{a^2 + b^2} = \frac{(ac + bd) + j(ad - bc)}{a^2 + b^2}$$

Notice that now the common denominator, $a^2 + b^2$, is a positive *real number,* which therefore *allows us to separate the fraction into its real and imaginary parts*; thus

$$\frac{ac + bd}{a^2 + b^2} + j\,\frac{ad - bc}{a^2 + b^2} = A + jB$$

hence showing that multiplying the numerator and denominator by the conjugate of the denominator *does* convert a given fraction into the form $A + jB$. Furthermore, in the above example we see that, by inspection,*

$$A = \frac{ac + bd}{a^2 + b^2} \quad \text{and} \quad B = \frac{ad - bc}{a^2 + b^2}$$

Example
Find the value of the quotient $\dfrac{5 - j4}{2 - j3}$.

Solution
By "find the value of" we mean "resolve the fraction into its real and imaginary parts." To do this we multiply the numerator and denominator by the "conjugate of the denominator"; thus

$$\frac{(5 - j4)\,(2 + j3)}{2^2 + 3^2} = \frac{10 + j15 - j8 + 12}{13} = \frac{22}{13} + \frac{j7}{13} = 1.6923 + j\,0.5385, \quad \textit{answer}.$$

The above procedure can be referred to as "rationalizing" the fraction, which enables us to simplify the ratio of two complex numbers into the form of a single complex number.

Problem 85
Find the values of each of the following:

$$\text{(a)} \quad \frac{3 + j4}{1 + j} \qquad\qquad \text{(b)} \quad \frac{14 - j25}{j5}$$

Problem 86
Find the value of

$$\frac{(4 + j)\,(2 - j)}{(1 + j)\,(3 - j2)}$$

* Two complex numbers are defined to be equal IF AND ONLY IF their real parts are equal and their imaginary parts are equal.

Problem 87
Simplify the expression

$$\frac{j12}{(1+j)^4} \qquad (\textit{Answer}:\ -j3)$$

To close this section, suppose a complex number is equal to *zero*; that is, suppose

$$a + jb = 0 + j0 = 0$$

Since two complex numbers can be equal only if the real parts are equal and the imaginary parts are equal, it follows that the above can be true only if $a = 0$ and $b = 0$; that is:

> A complex number can be equal to zero only if the real and imaginary parts are BOTH equal to zero.

6.4 Graphical Representation of Complex Numbers

We are familiar with the fact that real numbers can be represented as POINTS ON A STRAIGHT LINE. This is illustrated in Fig. 106, in which $X'X$ (line X prime, X) represents such a line. The point "0," called the "origin," represents the number *zero*.

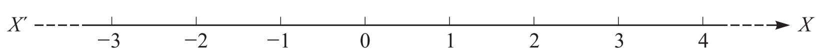


Fig. 106

As shown, all POSITIVE real numbers are represented as points to the RIGHT of the origin, while all NEGATIVE real numbers are represented by points to the LEFT of the origin.

We accept the statement that TO EVERY REAL NUMBER there corresponds ONE AND ONLY ONE POINT on line $X'X$. We accept the reciprocal statement that TO EVERY POINT on line $X'X$ there corresponds ONE AND ONLY ONE real number.

Fig. 106 is called the AXIS OF REAL NUMBERS, or simply the "axis of reals."

Now consider the corresponding representation of imaginary numbers. As we know, imaginary numbers have the form jb, where j is the imaginary unit (eq. (133)) and where b can be any positive or negative real number.

From the description of Fig. 106 it follows that the same idea can be applied to imaginary numbers; that is, IMAGINARY NUMBERS can *also* be represented as points on a straight line.

This is illustrated in Fig. 107, in which $Y'Y$ (line Y prime, Y), is the AXIS OF IMAGINARY NUMBERS (or "axis of imaginaries"), in the same way that Fig. 106 is the "axis of reals."

Fig. 107

Thus we can present EVERY REAL NUMBER as a point on the straight line of Fig. 106, and EVERY IMAGINARY NUMBER as a point on the straight line of Fig. 107.

But now consider the graphical representation of COMPLEX NUMBERS, $a + jb$. Take, for example, any two complex numbers, such as $2 + j2$ and $2 + j3$. Note that there is NO POINT on a straight line that can exclusively represent *both* of these numbers. Thus the infinite number of different complex numbers cannot possibly be represented by points on a one-dimensional straight line.

Instead, to graphically represent complex numbers a two-dimensional PLANE SURFACE is required. This requirement is met by positioning the "axis of imaginaries" perpendicular to the "axis of reals", the two axes intersecting at their common "0," thus creating the COMPLEX NUMBER PLANE, as shown in Fig. 108.

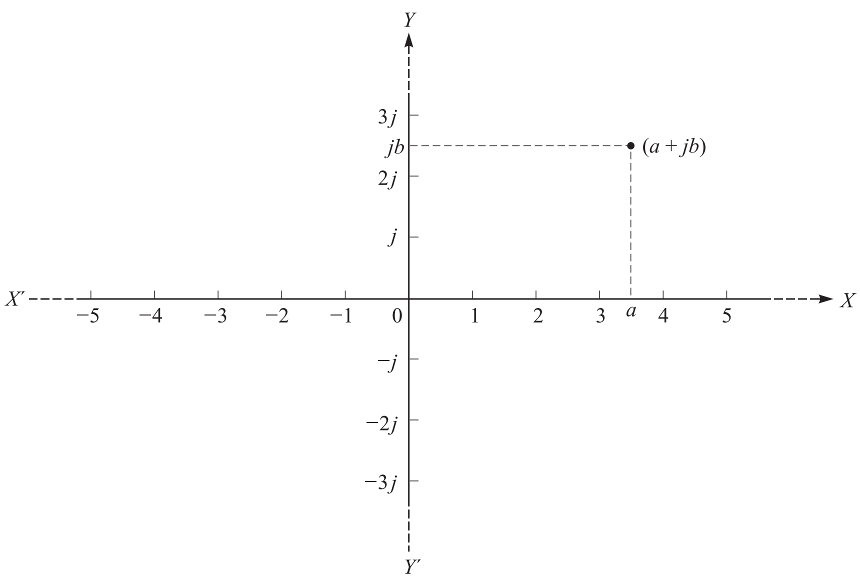


Fig. 108

Note that all negative and positive REAL numbers are represented by points on the "real" or x-axis (X', X), with all negative and positive IMAGINARY numbers represented by points on the "imaginary" or y-axis (Y', Y), while ALL COMPLEX NUMBERS are represented by points on the plane, such as the general complex number $(a + jb)$, as illustrated in the figure.

Thus, to locate the point that represents a given complex number $a + jb$, we locate the real part a on the x-axis, then locate the imaginary part b on the y-axis; the desired point is at the intersection of the horizontal and vertical lines drawn from b and a, as shown in the figure.

Thus each real, imaginary, and complex number is represented by a single unique point on the complex plane of Fig. 108.

Problem 88

In section 6.3 we saw that two complex numbers are equal only if their real parts are equal and their imaginary parts are equal. Is this fact evident from inspection of Fig. 108?

Thus far in our work we've written complex numbers in what is called the RECTANGULAR form, that is, in the form "$a + jb$."

It is also possible, and often highly desirable, to write complex numbers in what is called the POLAR or TRIGONOMETRIC form. In this form, a complex number is expressed in terms of *magnitude*, A, and *angle*, θ. The relationship between the "rectangular" and "polar" form of a complex number can be derived from inspection of Fig. 109.

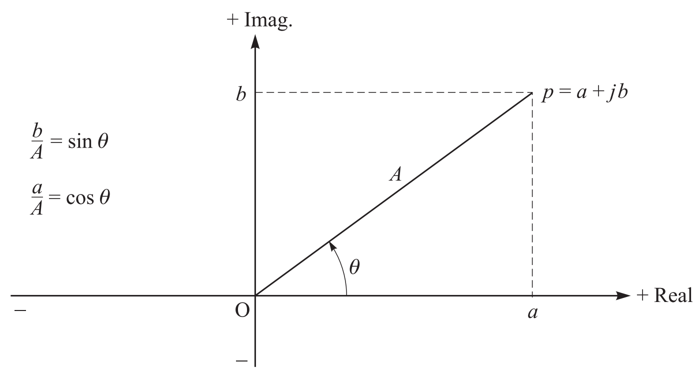


Fig. 109

Thus, in the figure,

$$\boxed{a + jb = A(\cos\theta + j\sin\theta)} \tag{142}$$

the right-hand side being the equivalent "trigonometric" form of $a + jb$."

The length A is called the "modulus" or "absolute" value of the complex number, and is always taken to be a POSITIVE value. The angle θ (theta) is called the "amplitude" of the complex number, with positive angles measured in the ccw (counterclockwise) sense. From inspection of Fig. 109,

$$A = \sqrt{a^2 + b^2} \tag{143}$$

$$\left.\begin{aligned} \tan\theta &= \frac{b}{a} \\ \theta &= \arctan(b/a) \end{aligned}\right\} \tag{144}$$

Complex numbers in the trigonometric form are often written in the abbreviated form $A\underline{/\theta}$, which is called the "polar" form.* Hence the rectangular, trigonometric, and polar forms all denote the same thing, a complex number; thus

$$a + jb = A(\cos\theta + j\sin\theta) = A\underline{/\theta} \tag{145}$$

In the following two problems round all calculator values off to 3 decimal places.

Problem 89

Write the following complex numbers in trigonometric form.

(a) $2.3 - j3.5 + 5.8 + j13.9$ (b) $\dfrac{7 - j2}{4 + j9}$

---

* The same notation is used in ordinary algebra when graphing purely real functions on real polar coordinate space. In our work, however, it will be used only in connection with the complex plane, as defined in eq. (145).

Problem 90
Write the following complex numbers in rectangular form, $a + jb$.

(a) $90\underline{/166^\circ}$ (b) $400\underline{/-126^\circ}$ (c) $17\underline{/45^\circ} - 22\underline{/265^\circ}$

6.5 Exponential Form of a Complex Number

In this section we introduce the "exponential" ("expo NEN shal") form of a complex number. We do this because certain operations can be greatly simplified if the complex numbers are written in this form.

It is necessary, however, in order to give a proper explanation of the ORIGIN of the exponential form, that we make use of certain procedures from more advanced mathematics. In our explanations we'll state the facts as clearly as we can, so that you can understand, in a good general way, the origin of the exponential form. It is important that we do this, because it will give you added confidence in handling complex numbers in this form, which, as you'll be pleased to find, is not at all hard to do. Let us proceed as follows.

The "ratio of the circumference of a circle to the diameter" is certainly one of the best known, and most used, numbers in mathematics. It is a constant ratio, universally represented by the Greek letter π (pi), being an irrational number* having the approximate value $\pi = 3.141\,592\,65\ldots$.

Another number, equal in importance to π, also exists. This number is denoted by the Greek letter ϵ (epsilon) and, like π, is an irrational number, having the approximate value $\epsilon = 2.718\,281\,828\,459\ldots$. The number represented by ϵ arises in the study of the "logarithmic function" and, for a very specific reason, is *defined* to be equal to

$$\epsilon = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = 2.718\,28\ldots \tag{146}$$

Using your calculator, you can verify, for example, the following approximate values of ϵ (rounded off to five decimal places):

$$\begin{array}{lll} n = 10 & \epsilon = (1.1)^{10} & = 2.593\,74 \\ n = 100 & \epsilon = (1.01)^{100} & = 2.704\,81 \\ n = 1000 & \epsilon = (1.001)^{1000} & = 2.716\,92 \\ n = 10\,000 & \epsilon = (1.0001)^{10\,000} & = 2.718\,15 \end{array}$$

Thus, as n becomes infinitely great ϵ does *not* become infinitely great, but becomes an infinite string of non-repeating decimals, having a limiting value of something less than 2.72.

As mentioned above, the number ϵ arises in the study of the logarithmic function. In regard to "logarithms," any system of logarithms has a "base number," which let us denote by b. Now let y be any positive number; the "logarithm of y" is then defined to be the POWER that the base number b must be raised to, to equal y. That is, the statement that

$$\log_b y = x \tag{147}$$

* See note 11 in Appendix.

means that

$$y = b^x \tag{148}$$

(where "$\log_b y$" is read as "the logarithm of y to the base b").

The *reason* the foregoing is important is that the expression in eq. (146) appears at a critical point in the development of the logarithmic and exponential functions in the calculus. It is found that the formulas of the calculus, in regard to logarithmic and exponential work, *are much simplifed if ϵ is used as the base number*. Hence, in advanced mathematics, it is always understood that ϵ is the logarithmic base number; thus, for this case eq. (148) becomes

$$y = \epsilon^x \tag{149}$$

which is the fundamental form of the "exponential function" as it appears in advanced mathematics.

Next, eq. (149) can be written in the equivalent form of a *power series in x*;* thus

$$\epsilon^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots + \frac{x^n}{n!} + \cdots \tag{150}$$

in which the exclamation mark "!" denotes the PRODUCT of all the positive integers *from 1 to n inclusive*; thus

$$n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \cdots \cdot n$$

hence, $2! = 1 \cdot 2 = 2$, $3! = 1 \cdot 2 \cdot 3 = 6$, $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$, and so on. The symbol ! can be read as "factorial"; thus we have "2 factorial," "3 factorial," and so on.

In the study of power series it is shown that the series form of eq. (150) is a valid representation of ϵ^x *for all positive and negative values of the variable x*.

Next, *the sine and cosine functions* can also be expressed in the power series form. Thus the following relationships *are valid for all positive and negative values of x*:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \tag{151}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \tag{152}$$

which, it should be noted, are valid in the above form only if angle x is measured in *radians*. (The series become awkward to write for x in degrees.) These are said to be "alternating" series, because the terms are alternately positive and negative in sign, as you can see.

Now let us return to eq. (150) and boldly take the step of replacing "x" with "jx." Assuming this is permissible (which it is), and upon taking careful note of the values of the "powers of j" from section 6.1, you should find that eq. (150) becomes

$$\epsilon^{jx} = \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\right) + j\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\right)$$

Now compare the last equation with eqs. (151) and (152); doing this, we see that

$$\boxed{\epsilon^{jx} = \cos x + j \sin x} \tag{153}$$

---

* See note 12 in Appendix.

which is called EULER'S FORMULA (pronounced "oiler"). This is one of the most important relationships in all of mathematics and the engineering sciences.*

In eq. (153) let us replace x with $-x$; doing this, and remembering that $\cos(-x) = \cos x$ and that $\sin(-x) = -\sin x$, we have the additional relationship

$$\boxed{\epsilon^{-jx} = \cos x - j\sin x} \tag{154}$$

Equations (153) and (154) are valid for all positive and negative values of the variable x. The left-hand sides of the equations are the "exponential forms" of the complex numbers on the right-hand sides.

Now let us return to Fig. 109 and eq. (142) in section 6.4. Comparison of eqs. (142) and (153), writing θ in place of x in eq. (153), shows that

$$a + jb = A\epsilon^{j\theta} \tag{155}$$

Thus the complex number $a + jb$ of Fig. 109 can also be represented in the form $A\epsilon^{j\theta}$, where A is the MAGNITUDE of the complex number and θ is the ANGLE of A on the complex plane.

Hence it is possible to represent a complex number in any of the following FOUR EQUIVALENT WAYS:

$$\text{rectangular form:} \quad a + jb \tag{156}$$

$$\text{polar form:} \quad A\underline{/\theta} \tag{157}$$

$$\text{trigonometric form:} \quad A(\cos\theta + j\sin\theta) \tag{158}$$

$$\text{exponential form:} \quad A\epsilon^{j\theta} \tag{159}$$

in which the MAGNITUDE A and ANGLE θ are given by eqs. (143) and (144) in section 6.4.†

Each of the four forms (eqs. (156) through (159)) has certain advantages and disadvantages, depending upon the type of operation (addition, subtraction, multiplication, or division) that is to be performed. Let us first take up the case of ADDITION AND SUBTRACTION as follows.

In section 6.2 we showed that the real and imaginary parts of the SUM OR DIFFERENCE of two or more complex numbers is, respectively, equal to the sum or difference of the REAL PARTS of the numbers and the sum or difference of the IMAGINARY PARTS of the numbers. Hence, in order to carry out the operation of addition or subtraction, the real and imaginary part of each complex number must be available *separately*; it thus follows that, to find the sum or difference of complex numbers, the numbers must first be expressed in the form of *either equation (156) or (158)*.

Problem 91

Find the algebraic sum of the following complex numbers (answer in the rectangular form of eq. (156)).

$$16\underline{/36^\circ} - 22\underline{/315^\circ} + (9.15 - j6.88) =$$

---

* Named for the great Swiss mathematician Leonard Euler (1707–1783).

† Theoretically the angle θ is in *radians*, because the forms of the sine and cosine series in eqs. (151) and (152) are derived for the case of the angle being measured in radians. In certain operations in the calculus it *is* necessary that the angles be in radians; however, in the ordinary operations of addition, subtraction, multiplication, and division, the angles can be expressed in degrees if we wish, and we will usually do so.

In several of the following problems we are asked to convert the rectangular form, $a + jb$, into the polar or exponential form; in doing this some caution is called for, as follows.

The conversion of $a + jb$ into the polar or exponential form requires the use of eqs. (143) and (144). While there is no difficulty in using eq. (143), care must be taken when using eq. (144), that is, *in finding the correct value of the angle* θ by using the equation $\theta = \arctan(b/a)$. This is because, for a given complex number $a + jb$, the correct value of θ depends upon the QUADRANT in the complex plane that the point representing $a + jb$ falls in. This is illustrated in Fig. 110, in which h is the angle given by eq. (144) using the magnitude of b/a only; that is

$$h = \arctan |b/a|$$

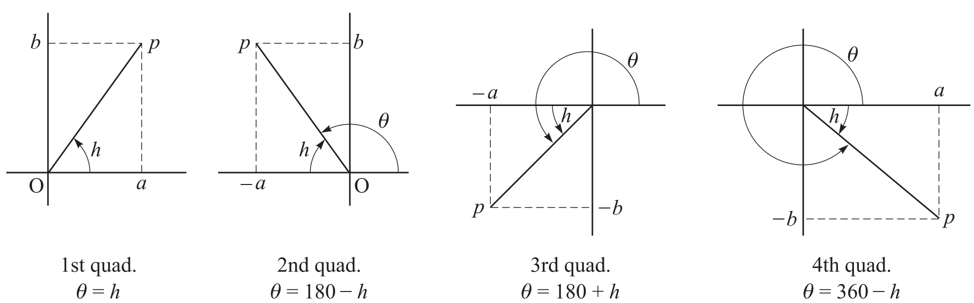


Fig. 110

Problem 92
Express each of the following complex numbers in exponential form.

(a) $3 + j4$ (c) $-3 - j4$

(b) $-3 + j4$ (d) $3 - j4$

Verify, by direct use of eq. (153), that your answers are correct.

Problem 93
Write the answer to the following sum in exponential form.

$$14\epsilon^{j112^\circ} + 8\epsilon^{j28^\circ} + 19\epsilon^{-j155^\circ} =$$

6.6 Operations in the Exponential and Polar Forms. De Moivre's Theorem

We have found that the algebraic ADDITION of complex numbers must be carried out in the "$a + jb$" (rectangular) form. This is because, in addition, the REAL PARTS and the IMAGINARY PARTS of the numbers must be separately added together to get the final resultant *sum* of the numbers (section 6.2).

It thus follows that complex numbers CANNOT BE DIRECTLY ADDED TOGETHER IN THE EXPONENTIAL FORM, because the real and imaginary parts are not shown separately in the exponential form (see problem 93).

However, while the exponential form is *not* suited to the addition and subtraction operation, it *is* very definitely suited to the MULTIPLICATION AND DIVISION operations. This is because in multiplication and division we can make use of the BASIC LAWS OF EXPONENTS, as the following will show.

Consider any two complex numbers

$$a + jb = A\epsilon^{jp}$$

$$c + jd = B\epsilon^{jq}$$

where A and B are the magnitudes of the numbers and p and q are the angular ''amplitudes'' of the numbers (eqs. (143) and (144), and Fig. 109).

Now let us consider the PRODUCT of the above two complex numbers. First, in the ''rectangular'' form we have (section 6.2)

$$(a + jb)(c + jd) = (ac - bd) + j(ad + bc) \tag{160}$$

Now consider the same multiplication if the two numbers are expressed in ''exponential'' form. Remembering that in multiplication EXPONENTS ARE ADDED, we have the result

$$(A\epsilon^{jp})(B\epsilon^{jq}) = AB\epsilon^{j(p+q)} \tag{161}$$

Thus, if complex numbers are expressed in EXPONENTIAL FORM, the PRODUCT of the numbers is a complex number whose MAGNITUDE IS THE PRODUCT OF THE MAGNITUDES and whose ANGLE is the SUM OF THE ANGLES of the individual numbers.

Comparison of eqs. (160) and (161) shows that multiplication in the exponential form is generally easier than multiplication in the rectangular form. Furthermore, the use of the exponential form can often simplify the mathematical work involved in theoretical investigations.

Now let us consider the DIVISION or ''quotient'' of the same two complex numbers. First, in the ''rectangular'' form we have (section 6.3)

$$\frac{a + jb}{c + jd} = \frac{(ac + bd) - j(ad - bc)}{c^2 + d^2} \tag{162}$$

Now consider the same division if the two numbers are expressed in ''exponential'' form. Remembering that in division the exponent of the denominator is SUBTRACTED from the exponent of the numerator, we have the result

$$\frac{A\epsilon^{jp}}{B\epsilon^{jq}} = \left(\frac{A}{B}\right)\epsilon^{j(p-q)} \tag{163}$$

Thus, if two complex numbers are expressed in exponential form, the QUOTIENT of the two numbers is a complex number whose MAGNITUDE is equal to the QUOTIENT OF THE TWO MAGNITUDES and whose ANGLE is equal to the angle of the numerator MINUS the angle of the denominator.

Comparison of eqs. (162) and (163) shows that division in the exponential form is generally easier than division in the rectangular form. Again, the use of the exponential form can often simplify the mathematical work involved in theoretical investigations.

It follows that the same information contained in eqs. (161) and (163) can also be expressed in POLAR form; thus

$$\text{product:} \quad (A\underline{/p})(B\underline{/q}) = AB\underline{/p+q} \tag{164}$$

$$\text{quotient:} \quad \frac{A\underline{/p}}{B\underline{/q}} = \frac{A}{B}\underline{/p-q} \tag{165}$$

It follows that eqs. (161) and (164) extend on to cover any number of factors; thus, for the case of *three* factors we take the product of the first two times the third, and so on for any number of factors.

Problem 94
Write, in rectangular form, the product of the three complex numbers

$$3\epsilon^{j112^\circ}, \qquad 4\epsilon^{-j62^\circ}, \qquad \text{and} \qquad 7\epsilon^{j165^\circ}$$

Problem 95
Write the product $(4\underline{/19^\circ})(3\underline{/39^\circ})$ in rectangular form.

Problem 96
Making use of eq. (159) and the laws of exponents, raise the complex number $(2 + j3)$ to the sixth power. Answer in rectangular form.

Problem 97

$$\frac{15\epsilon^{j62^\circ}}{36\epsilon^{j85^\circ}} = \qquad \text{(answer in rectangular form)}$$

Problem 98

$$\frac{16\underline{/102^\circ} - 9\underline{/390^\circ}}{7\underline{/75^\circ}} = \qquad \text{(answer in rectangular form)}$$

Now, to continue, let us begin by writing down Euler's formula (eq. (153)) thus

$$\epsilon^{j\theta} = \cos\theta + j\sin\theta \tag{``A''}$$

Or, replacing θ with $n\theta$, it is also true that

$$\epsilon^{jn\theta} = \cos n\theta + j\sin n\theta \tag{``B''}$$

Now raise both sides of ("A") to the power n; noting that $(\epsilon^{j\theta})^n = \epsilon^{jn\theta}$, we have

$$\epsilon^{jn\theta} = (\cos\theta + j\sin\theta)^n \tag{``C''}$$

Since the right-hand sides of ("B") and ("C") are both equal to the same thing, they are equal to each other, and thus we have the important result that

$$\boxed{(\cos\theta + j\sin\theta)^n = \cos n\theta + j\sin n\theta} \tag{166}$$

which is called *De Moivre's Theorem* ("dah MWAH vrah").* The theorem is true for *all positive and negative, integral and fractional*, values of n.

Problem 99
Two complex numbers are EQUAL only if their REAL PARTS are equal and their IMAGINARY PARTS are equal (see problem 88). Using this fact and De Moivre's theorem, find

(a) a trigonometric identity for $\cos 2x$,
(b) a trigonometric identity for $\sin 2x$.

Problem 100
Find the value of $2(\cos 48^\circ + j\sin 48^\circ)$ raised to the fifth power.

Problem 101
Find the value of

$$\frac{1}{4(\cos 17^\circ + j\sin 17^\circ)^3}.$$

Problem 102
Making use of eq. (161), and using the same procedure as in problem 99, find

(a) the trigonometric identity for $\cos(x+y)$,
(b) the trigonometric identity for $\sin(x+y)$.

6.7 Powers and Roots of Complex Numbers

A complex number can be written in the rectangular form using the notation of either eq. (156) or eq. (158); thus

$$(a+jb) = A(\cos\theta + j\sin\theta) \tag{167}$$

in which the MAGNITUDE A and ANGLE θ are given by eqs. (143) and (144) in section 6.4. Now let us raise both sides of the last equation to a power n; thus

$$(a+jb)^n = A^n(\cos\theta + j\sin\theta)^n$$

which, by virtue of eq. (166), can also be written in the form

$$\boxed{(a+jb)^n = A^n(\cos n\theta + j\sin n\theta)} \tag{168}$$

As in the case of eq. (166), eq. (168) is valid for *all positive and negative, integral and fractional, values of n*. Let us, however, first consider the case where n is any positive or negative INTEGER (whole number). The following two problems will illustrate the procedure for the case where n is a positive or negative integer.

Problem 103
Using eq. (168), show that $(3-j2)^7 = -4449.06 + j6553.97$, approximately.

* Named for the French mathematician Abraham De Moivre (1667–1754).

Problem 104
Using eq. (168), show that

$$\frac{6000}{(3+j4)^5} = -0.1457 + j1.9144, \text{ approximately.}$$

Now (in preparation for finding the *root* of a complex number) let us be reminded that the sine and cosine are PERIODIC functions having a period of 360 degrees (2π radians). This is expressed (in degrees) by eqs. (79) and (80), in section 5.3; thus

$$\sin\theta^\circ = \sin(\theta \pm 360k)^\circ \tag{169}$$

$$\cos\theta^\circ = \cos(\theta \pm 360k)^\circ \tag{170}$$

where $k = 0, 1, 2, 3, \ldots$, that is, where k is any positive INTEGER. If θ is measured in *radians* the corresponding equations are

$$\sin\theta = \sin(\theta \pm 2\pi k) \tag{171}$$

$$\cos\theta = \cos(\theta \pm 2\pi k) \tag{172}$$

Since k is an INTEGER* the above equations are *true for all values of* k; this is because if θ is increased or decreased by any INTEGRAL MULTIPLE of 360° (or 2π radians), *the angle simply returns to its original position*, as illustrated in Fig. 111.

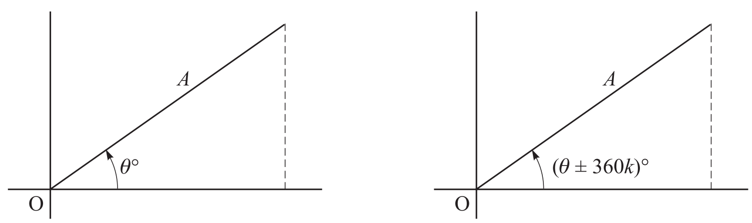


Fig. 111

Thus, since k is an INTEGER,

$$\cos\theta^\circ = \cos(\theta + 360k)^\circ \tag{173†}$$

and

$$\sin\theta^\circ = \sin(\theta + 360k)^\circ \tag{174}$$

and hence eq. (167) can be written in the form

$$(a + jb) = A[\cos(\theta + 360k)^\circ + j\sin(\theta + 360k)^\circ] \tag{175}$$

Now let "n" be any given positive INTEGER, and let us, using De Moivre's theorem, raise both sides of the last equation to the "$1/n$" power; thus,

$$(a + jb)^{1/n} = A^{1/n}\left[\cos\left(\frac{\theta}{n} + 360\frac{k}{n}\right)^\circ + j\sin\left(\frac{\theta}{n} + 360\frac{k}{n}\right)^\circ\right] \tag{176}$$

* k is a positive INTEGER, $k = 0, 1, 2, 3, \ldots$, in all our work in this section.

† From inspection of Fig. 111, note that $(\theta - 360k)^\circ = (\theta + 360k)^\circ$, because k is an integer.

It is important, now, to note that for all values of k *less than n* the quantity k/n is *not* equal to an integer, but is, instead, equal to a *fraction* of value less than 1. Thus, for all values of k *less than n* it is *not true* that

$$\sin\frac{\theta}{n} = \sin\left(\frac{\theta}{n} + 360\frac{k}{n}\right)^\circ$$

and thus we have that

> A complex number, $a + jb$, has *n roots*, given by eq. (176), for all values of k less than n, that is, for $k = 0, 1, 2, 3, \ldots, (n-1)$.*

Thus, in this section we've found that $(a + jb)$ raised to a positive or negative INTEGRAL power has just *one* value, given by eq. (168).

On the other hand, the "nth root" of $(a + jb)$—that is, $(a + jb)$ raised to the $1/n$ power—consists of *n distinct values*, given by eq. (176).

It should be noted that "inverse operations" tend to produce multiple answers, in the manner of eq. (176). For example, while 4^2 gives the single answer 16, note that

$$\sqrt{4} = +2 \quad \textit{and} \quad -2$$

or note that

$$\tan 45^\circ = 1.000$$

but

$$\arctan 1.000 = 45^\circ, 225^\circ, 405^\circ, 585^\circ, \text{and so on.}$$

In regard to eq. (176), it might be thought that such an equation, while interesting theoretically, would have little practical value. Actually, however, the equation has numerous engineering applications; it is, for example, central in the study of broad-band cascaded amplifier stages.

Problem 105
Find the four fourth roots of $(3 + j7)$. Show, on the complex plane, the points representing the four roots.

From our work in the foregoing problem, the procedure for finding the values (roots) of the complex expression $(a + jb)^{1/n}$ can be summarized as follows.

The GIVEN VALUES will be the values of a, b, and n. The first step, then, is to calculate the magnitude A and angle θ of the complex number $(a + jb)$, which is done by means of the formulas

$$A = \sqrt{a^2 + b^2} = (a^2 + b^2)^{1/2} \tag{177}$$

$$\theta = \arctan(b/a) \tag{178}$$

In eq. (177), the magnitude A is always taken to be a positive, real number. In finding the value of θ in eq. (178), we must take into account the quadrant that the point a, b lies in (see Fig. 110).

* The same values are merely repeated over and over for k greater than $(n-1)$.

As was pointed out in problem 105, the roots all have the *same magnitude*, $A^{1/n}$, which, by eq. (177), is equal to

$$A^{1/n} = (a^2 + b^2)^{1/2n} \tag{179}$$

Since the value of θ is known (by eq. (178)), and since n is a given value, we can now find the value of θ/n, which must now be substituted into eq. (176). We also know the n *different values of* k, $k = 0, 1, 2, \ldots, (n-1)$, which we must now substitute, in succession, into eq. (176). Doing this gives us the "n roots" of $(a + jb)^{1/n}$.

Problem 106
Find the fifth roots of the complex number $(19 - j33)$. For convenience, round off all numbers to two decimal places. Answers in polar and rectangular form.

Problem 107
Find the "cube roots of unity"; that is, find the values of $(1)^{1/3}$.

6.8 Complex Numbers as Vectors

Quantities that obey the PARALLELOGRAM LAW of addition and subtraction are said to be VECTOR quantities (see note 4 in the Appendix). Hence, by this definition, COMPLEX NUMBERS can be regarded as vector quantities, because complex numbers are basically added and subtracted in accordance with the parallelogram law. This can be shown as follows.

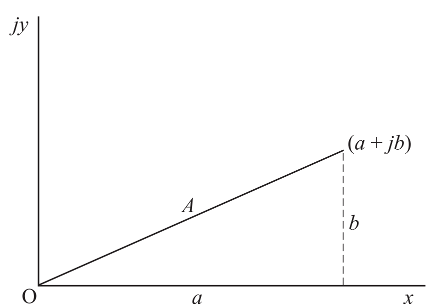


Fig. 112

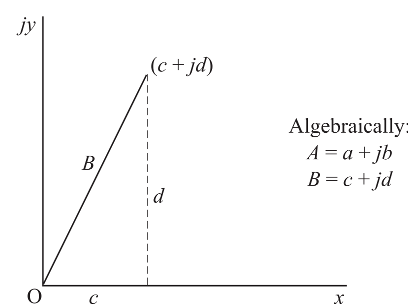


Fig. 113

Let A and B be two complex numbers, such as are illustrated geometrically in Figs. 112 and 113.

Let R be the SUM of the above two complex numbers; algebraically, from section 6.2, R is equal to

$$R = (a + c) + j(b + d) \tag{180}$$

Careful inspection of the above will show that R, *the SUM of the two complex numbers A and B* (given by eq. (180)), can be found *geometrically* by drawing B "off the end of A," as is done in Fig. 114.

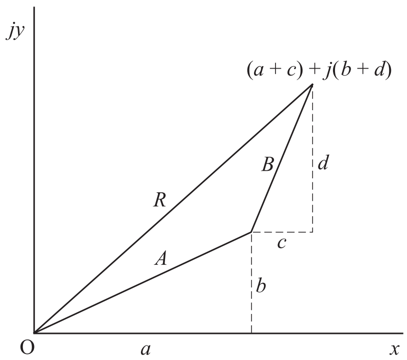


Fig. 114

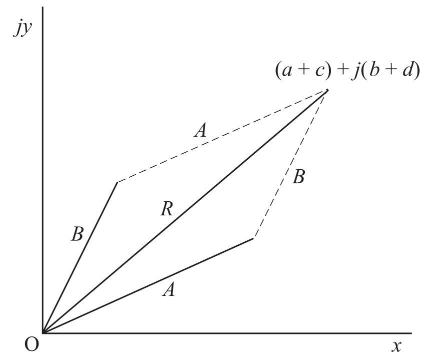


Fig. 115

By the basic eq. (180), we know that the *sum* of A and B is a complex number R, having a magnitude and angle equal to

$$|R| = \sqrt{(a+c)^2 + (b+d)^2}$$

$$\theta = \arctan(b+d)/(a+c)$$

and we see that the value represented by the line R in Fig. 114 exactly fits this requirement; thus R in Fig. 114 *does* represent the *sum* of A and B. Next, comparing Figs. 114 and 115, we see, in Fig. 115, that the sum R is also equal to the *diagonal of the parallelogram* formed with A and B as sides; thus, the sum R can be found, geometrically, by applying the PARALLELOGRAM LAW of addition. Hence complex numbers can be regarded as being *vector* quantities. We will make much use of this fact in our future work.

CHAPTER 7

Inductance and Capacitance

7.1 Introduction

Electric circuits consist of connections of ''active'' and ''passive'' devices (section 2.5). An active device is a *source of energy* in a circuit; a battery, for example, is an active device. An active device is also called a ''generator.'' As explained in section 2.2, a generator possesses what is called *electromotive force* (emf), which is measured in *volts*, denoted by V or v.

The passive elements in electric circuits are of three types, called RESISTANCE, INDUCTANCE, and CAPACITANCE.

Of these three, only *resistance*, R, *consumes electrical energy*, that is, *removes electrical energy from the circuit* by converting it into some other form of energy, such as heat energy or mechanical energy. The other two passive elements, inductance and capacitance, do not consume energy but only STORE ENERGY *momentarily*, in magnetic or electric fields, and then return it to the circuit.

We have already met the resistance parameter R in Chap. 2. We learned that resistance is measured in ''ohms,'' and that the ''voltage drop'' or ''counter emf'' developed across a resistor of R ohms by a current of I amperes is $V = RI$ volts. The *polarity* of the voltage drop across R is always ''plus to minus'' in the direction of the current flow through R. The *rate* at which energy is dissipated in a resistance of R ohms carrying a current of I amperes is $W = RI^2$ watts (joules per second).

Resistance can be in the form of an actual ''physical'' electrical resistance, such as the resistance wire in a toaster, or it can be in the form of an equivalent ''dynamic'' resistance, such as the mechanical load on a motor or the vibrating cone of a loudspeaker.

In the case of the toaster wire, the electrical energy is converted directly into heat energy in the circuit itself, and the wire, which is part of the physical circuit, is seen to get ''red hot.''

In the case of a ''dynamic'' resistance load, such as a motor driving a drill press, for example, most of the heat energy does not appear directly in the electrical circuit itself, but instead appears in the material being drilled. Or perhaps the motor lifts a load of bricks up to the top of a tower; in this case, most of the electrical energy is not converted into heat energy but, instead, remains stored in the load of bricks in the form of ''potential energy of position.'' Of course, as far as the *generator* is concerned, the effect is the same in any case; all the generator ''knows'' is that it ''sees'' a resistance load of R ohms.

It should be noted that resistance is sometimes present as an undesirable, but unavoidable, side effect in a circuit. On the other hand, resistance is often a necessary part of some electronic circuits, and so is purposely put into the circuit. Such resistances, called ''resistors,'' can be purchased in resistance values ranging from a small fraction of an ohm to many millions of ohms (megohms). This was discussed in section 2.4.

Now that we are familiar with the ''resistance parameter,'' the next step is to gain an equal understanding of the ''inductance'' and ''capacitance'' parameters. Let us begin with inductance, after which we'll take up capacitance. Since INDUCTANCE is associated with MAGNETIC fields, we begin with some basic details concerning the magnetic field.

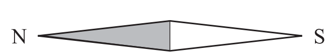
7.2 Introduction to Magnetism

We are all familiar with what is called a ''permanent magnet.'' A permanent magnet is simply a piece of steel having the ability to attract to it other pieces of steel and iron. Such a magnet is called ''permanent'' because it is capable of retaining its attractive ability for many years.

Of course, not all steels can be permanently magnetized. This is indeed fortunate, because the operation of many important electrical devices, such as transformers, depends upon the use of a steel that cannot be permanently magnetized. This type of steel (silicon steel) can be in a highly magnetized condition at one instant of time and then, almost instantly, lose all of its magnetization when the magnetizing force is removed.

A permanent magnet produces a ''magnetic field,'' which exists in the three-dimensional space surrounding the magnet. We can suppose that a magnetic field consists of ''lines of magnetic force'' in the space surrounding the magnet. It should be noted that this is the same concept that was used to describe ''lines of ELECTRIC force'' in section 1.3 (Figs. 13 and 14).

The ''direction'' of a magnetic field, at any point in the field, is defined according to the direction that a ''compass needle'' would point if placed at that point in the field. As you know, a compass needle has a ''north pole'' and a ''south pole,'' the north pole customarily being painted lightly, while the south pole is unpainted, as in the sketch below.



Let us define that the *direction* of a magnetic field, at any point in the field, is the direction in which the north pole of a small ''test compass'' would point if placed in the field at that point. Thus, in Fig. 116A, the direction of the field is from right to left, while in Fig. 116B the direction is from left to right.

In these figures, note that the ''lines of magnetic force'' are drawn closer together near the bottom of the figures than at the top; this is simply the graphical way of showing that the strength of the magnetic field is (in this case) greater in the region toward the bottom

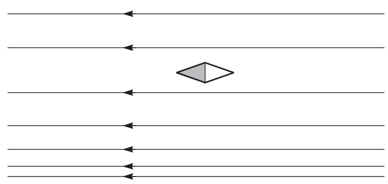


Fig. 116A

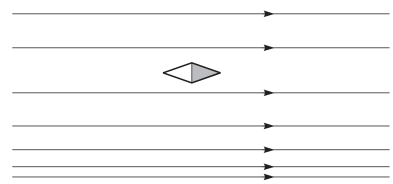


Fig. 116B

than near the top. (The same method was used in connection with the electric fields of Figs. 13 and 14 in Chapter 1.)

In this section we've considered the source of magnetic fields to be permanent magnets. More importantly, however, magnetic fields are also associated with *electric currents*, a phenomenon referred to as ''electromagnetism,'' which we introduce in the following section.

7.3 Electromagnetism

In the year 1820 the Danish physicist Oersted discovered the phenomenon of *electromagnetism*, that is, that A MAGNETIC FIELD EXISTS AROUND ANY CONDUCTOR CARRYING AN ELECTRIC CURRENT. Experimentation with a compass needle showed that the field existed at all points along a conductor, in the form of *concentric circles* around the conductor, as illustrated in Fig. 117. The arrow alongside the *i* indicates the direction of the current flow in the conductor.

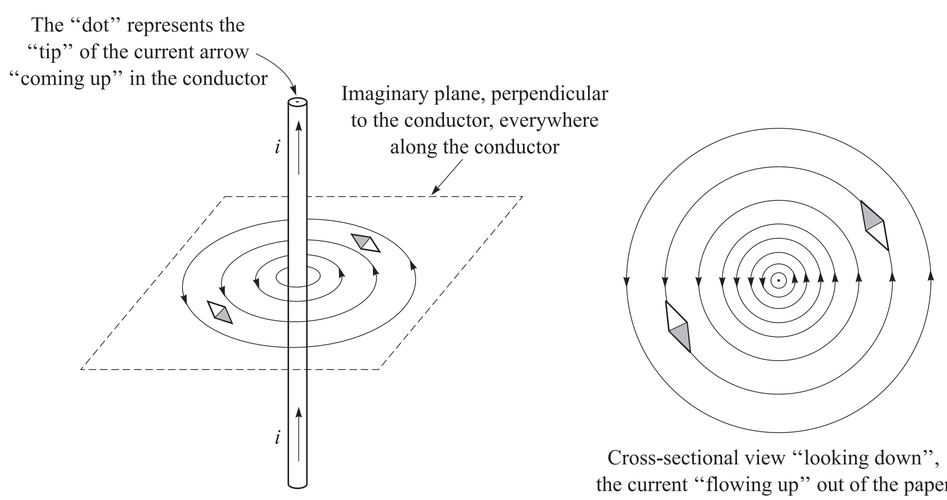


Fig. 117

In the figure, it should be understood that a similar plane can be drawn at every point along the conductor (we show just one such plane in the figure).

The closer we get to the wire conductor, the stronger is the magnetic effect. This fact is shown by drawing the lines of magnetic force closer together near the wire, and farther and farther apart as we move away from the wire, as shown in the figure.

In Fig. 117, the direction of the current i is given to be "upward" in the conductor, thus producing the magnetic field as shown. If, however, in the figure, the direction of the current i were *reversed*, the direction of the magnetic field would *also* be reversed; that is, *the direction of the magnetic field* depends upon *the direction of current flow*. The direction of the magnetic field can be found by using the "right-hand rule," as follows.

> Grasp the conductor with the RIGHT HAND, with the thumb pointing in the direction of conventional current flow. The fingers then curl around the conductor in the direction of the magnetic field produced by the current.

Note that the relationships shown in Fig. 117 are drawn in accordance with the right-hand rule. (Also note the compass alignment in the given field.)

It should be remembered that "lines of magnetic force" are imaginary lines that we draw to indicate the relative magnitude and direction of a magnetic field. Even though such lines are imaginary, they are very useful to us in visualizing and describing magnetic fields. The lines are also spoken of as *lines of magnetic* FLUX. Regions where the magnetic force is strong are said to be regions having a high density of "magnetic flux." Thus, in our discussion, the terms "lines of magnetic force" and "lines of magnetic flux" will be used interchangeably.

The magnetic effect produced by a current flowing in a wire can be increased by forming the wire into a circular COIL, as illustrated in Fig. 118. Let us suppose the coil consists of N turns of wire, wound on a cardboard tube, with a and b denoting the length and diameter of the coil, as shown. Let a current of i amperes be flowing in the coil, in the sense shown in the figure.

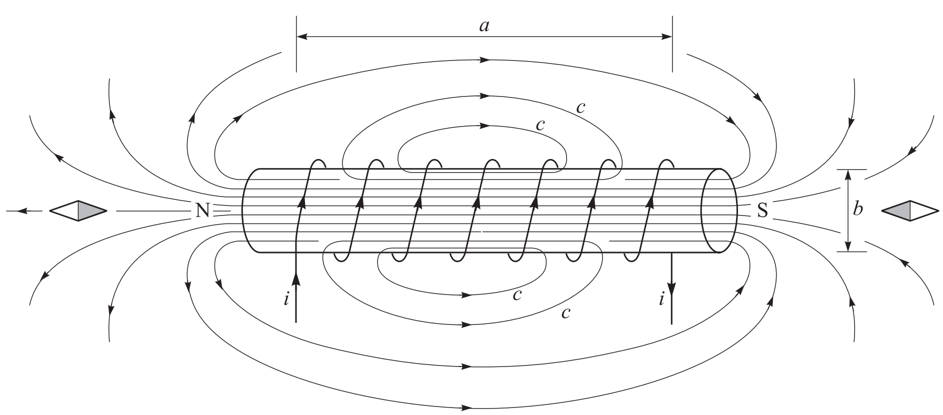


Fig. 118

First, in Fig. 118, note that the FLUX DENSITY *inside* the coil is greater, and more uniform, than it is at points *outside* the coil. Also note that *not all* of the generated flux passes through the *entire* interior of the coil; such lines are referred to as "leakage flux;" the lines labeled c represent such leakage flux. This effect can be reduced by winding the turns closer together.

In the figure, note that NORTH AND SOUTH MAGNETIC POLES exist at the ends of the coil. The nature of such poles (north or south) depends upon *the direction of the current*, and can be conveniently found by using the following "right-hand rule":

> If the fingers of the right hand curl around the coil in the direction of the current, the thumb points to the north pole.

Note that Fig. 118 is drawn in accordance with the above right-hand rule. Also note the compass alignment in the figure ("unlike poles attract each other").

The arrangement of Fig. 118 constitutes an "electromagnet." If the current i were reduced to zero ($i = 0$), the magnetic field would vanish.

It is important to emphasize that WORK *must be done* to create the magnetic field of Fig. 118. The work, so done, *then remains stored in the magnetic field* as "energy of the magnetic field." If, later on, the current is reduced to zero, the magnetic field "collapses" and the stored energy is returned to the circuit. Thus, since positive or negative *work* must be done, it is impossible to *instantly* change the state of a magnetic field (just as it is impossible to instantly change the state of motion of a physical body).

7.4 Self-Inductance

All conductors possess what is called "self-inductance." This is true regardless of whether the conductor is in the form of a straight wire, as in Fig. 117, or in the form of a coil of wire, as in Fig. 118 (although the effect is much greater in the case of the coiled form of Fig. 118 than in the straight wire of Fig. 117).

To understand the meaning of self-inductance, let us begin with the PRINCIPLE OF ELECTROMAGNETIC INDUCTION, discovered by Michael Faraday in 1831 (and, at almost the same time, by Joseph Henry in America). This famous principle states that

> A CHANGING magnetic field generates an electromotive force.

Notice that we emphasize the word CHANGING; if the magnetic field in a region of space is *constant*, that is, *not* changing, then NO *electromotive force* is generated or "induced" in that region of space.

But, as we have just learned in section 7.3, a MAGNETIC FIELD is generated by, and always accompanies, any ELECTRIC CURRENT. Thus a CHANGING ELECTRIC CURRENT generates a CHANGING MAGNETIC FIELD and therefore, by Faraday's principle, it follows that

> A changing electric current in a circuit induces an electromotive force in that same circuit.

This is the phenomenon of SELF-INDUCTANCE, mentioned at the beginning of this section. Actually, in practical work the term INDUCTANCE is usually used instead of the longer term "self-inductance." Thus, when we speak of the "inductance" of a coil, it will be understood that we mean "self-inductance."

It must be emphasized that inductance is a basic and very important component in electric circuit design. In practical work, required amounts of inductance are added to a circuit in the form of "inductance coils," called "inductors," as in Fig. 118. The greater the number of turns of wire, the greater is the amount of inductance possessed by such a coil.

If an extremely large amount of inductance is required, the wires are wound on an "iron core."

Next, the nature of the "self-induced voltage" that appears in an inductance coil can be summarized as follows.

> A self-induced voltage always appears at the terminals of an inductor coil whenever a CHANGE in the amount of current flowing in the coil occurs. The POLARITY of the self-induced voltage is always such as to OPPOSE THE CHANGE IN CURRENT in the coil.*

With the above in mind, let us now consider the following three possibilities concerning the state of the current i flowing in an inductor coil:

(a) a CONSTANT CURRENT is flowing in the coil;

(b) the current i is INCREASING in value;

(c) the current i is DECREASING in value.

A discussion of the above three cases, (a), (b), and (c), follows, in which the symbol represents an inductor coil.

(a) If there is NO CHANGE in the value of the current, then there is no change in the magnetic field and thus, by Faraday's principle, there is NO self-induced voltage induced into the coil. The energy stored in the magnetic field of the coil remains constant, and zero voltage appears between the terminals of the coil.

(b) Let it be given that a current i is flowing in an inductor coil in the direction of the arrow in Fig. 119, and let it be given that the current is *increasing* in value.

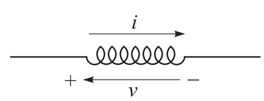


Fig. 119

In the figure, v is the *self-induced voltage* appearing between the coil terminals due to the increasing current. Note that the POLARITY of v is such that it OPPOSES the increasing current (this is in accordance with the principle stated above).

Thus, in this case v is like the voltage drop across a resistance of R ohms (section 2.5), EXCEPT that now *energy is not taken from the circuit* but is, instead, being *stored* in the magnetic field of the coil.

(c) Now let it be given that a current i is flowing in an inductor coil in the direction of the arrow in Fig. 120 (the same direction as in Fig. 119), and let it be given that the current is DECREASING in value.

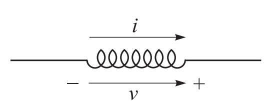


Fig. 120

* This is known as "Lenz's law."

Now notice that the direction of the self-induced voltage is *the same as the direction of the current* (the opposite of the situation in Fig. 119). This action satisfies the basic law that a self-induced voltage ALWAYS OPPOSES ANY CHANGE IN THE CURRENT THAT IS PRODUCING THE VOLTAGE (in Fig. 120 the current is attempting to decrease). This happens because, as the current in the coil begins to decrease, the "collapsing" magnetic field momentarily acts as a *generator*, thus returning the energy, stored in the magnetic field of the coil, back into the circuit.

As the figures show, the POLARITY MARKS ("+" and "−") at the ends of the coil depend (for a given direction of current) upon whether the current is increasing or decreasing in value. If the current is *constant*, then *no* self-induced emf appears across the coil.

7.5 The Unit of Inductance

It is a CHANGING MAGNETIC FIELD, produced by a CHANGING ELECTRIC CURRENT, that produces a self-induced voltage.

The *amount* of self-induced voltage produced depends upon HOW FAST *the magnetic field is changing*, and thus upon HOW FAST THE CURRENT IS CHANGING. This can be summarized in the statement that

The MAGNITUDE *of self-induced voltage* in a coil is proportional to the RATE OF CHANGE OF CURRENT in the coil.

Thus the amount of self-induced voltage produced by a current does *not* depend upon the *amount* of current flowing, but only upon the RATE OF CHANGE OF THE CURRENT, that is, upon *how fast* the current is changing. Now consider the following.

Mathematically, the *rate of change of current with respect to time* is denoted by the symbol*

$$\frac{\mathrm{d}i}{\mathrm{d}t} = \frac{\text{rate of change of current}}{\text{with respect to time}} = \text{amperes per second} = \text{amp/sec}$$

In the foregoing discussion it was pointed out that the *ratio* of the *self-induced voltage* v to the *rate of change of current* i has a *constant value*; the mathematical form of this statement is

$$v = k\,\frac{\mathrm{d}i}{\mathrm{d}t}$$

in which the *value* of the constant k depends upon the *particular coil* we are dealing with; that is, upon such things as the number of turns of wire, the spacing between the turns, the type of material the coil is wound on, and so on. Thus the value of the constant k depends upon the *physical characteristics of the coil itself*, and for this reason the constant k is called the INDUCTANCE *of the coil* (or, more fully, "self inductance," as mentioned in section 7.4).

---

* Read as "dee i, dee t." See note 5 in Appendix.

The inductance of a coil is almost universally denoted by "L." Upon using L instead of k, the above equation becomes

$$v = L\frac{\mathrm{d}i}{\mathrm{d}t} \tag{181}$$

thus,

$$L = \frac{v}{\mathrm{d}i/\mathrm{d}t} \tag{182}$$

showing that the INDUCTANCE, L, of a given coil is equal to the constant *ratio* of the *induced voltage* v to the RATE OF CHANGE *of current producing* v. The *unit* of inductance is called the HENRY, in honor of Joseph Henry; thus, if *1 volt* were induced when the current was changing at the rate of *1 ampere per second*, the inductance of that particular coil would be "1 henry."

A very great range of values of inductance is encountered in practical work, from millionths of a henry (microhenrys) to hundreds of henrys. The conversion formulas are

$$\text{microhenrys} = \text{henrys} \times 10^6$$
$$\text{henrys} = \text{microhenrys} \times 10^{-6}$$

Inductors can be broadly classified as being of either the "air core" type or the "iron core" type. In this regard, let us return briefly to the coil of Fig. 118, which, let us assume, is composed of a certain number of turns of wire, and in which, let us also assume, a constant current of I amperes is flowing.

The current I is the *magnetizing force* that produces the magnetic field, producing the "magnetic flux" that "links" the turns of the coil. The *amount* of such flux produced by a given amount of current I (in the given coil of Fig. 118) depends upon the type of *material* inside the cardboard tube the coil is wound on. In Fig. 118, if only *air* is inside the tube we have an "air core" type of inductor, in which the current I produces a certain amount of magnetic flux.

Now, in Fig. 118, suppose a rod composed of a compound or alloy of *iron* is inserted into the tube. We now have an "iron core" type of inductor, and we find that the amount of magnetic flux produced by the current I is *greatly increased* over the corresponding "air core" case. Thus, in applications requiring a large amount of inductance an "iron core" type of inductor is required.

In schematic diagrams, an "iron core" inductor is indicated by drawing a few parallel lines alongside the inductor symbol, as shown below.

Air core   Iron core

As mentioned above, the inductance of iron-core coils is much greater than that of air-core coils. Thus, while the inductance of an air-core coil is only a very small fraction of a henry, the inductance of an iron-core coil can be well in excess of 1 henry.

As a final comment, it should be noted that the *current* in an inductor coil cannot be INSTANTLY changed from one value to another value; that is, *time* is required to either increase or decrease the amount of current, flowing in an inductor coil, from one value to another value. This is related to the fact that time is required to change the amount of energy stored in the magnetic field of a coil.*

---

* See note 13 in Appendix.

Problem 108
The inductance of a certain coil is 0.62 henrys. Find the magnitude of the self-induced voltage during a 1-second time interval in which

(a) the current increases linearly from 1 ampere to 3 amperes,
(b) the current increases linearly from 28 amperes to 30 amperes.

Problem 109
Suppose, in problem 108, that the self-induced voltage is 5.52 volts at a certain instant of time. How fast is the current changing at that particular instant?

Problem 110
Suppose the current in a certain coil is increasing at the rate of 76 amp/sec at a certain instant of time. If the self-induced voltage is 0.048 volts at that instant, find the inductance of the coil in microhenrys.

7.6 Capacitors and Capacitance

We have learned that an INDUCTOR is a passive electrical circuit device that utilizes a *magnetic* field. A CAPACITOR, on the other hand, is a passive electrical circuit device that utilizes an *electric field* instead of a magnetic field.

Physically, a "capacitor" consists of two conducting surfaces called the PLATES, which are separated by an electrical *insulating material* called the DIELECTRIC.

In electrical drawings, capacitors are indicated by either of the symbols shown below, where the vertical lines represent the two "plates" of the capacitor and the horizontal lines represent the wires used to connect the capacitor to the rest of the circuit.

Usually the symbol using the curved line, on the right, is preferred, because the one on the left is the same as the symbol used to denote a set of relay contacts.

Capacitors have the ability to *store electric charge*, the amount of such charge depending upon the amount of voltage between the plates and a quantity called the CAPACITANCE of the capacitor, which we'll define a little later on.

The physical construction of a capacitor depends upon the amount of "capacitance" required, the amount of voltage that will appear between the plates, and the type of circuit it is to be used in. For example, in very-high-frequency work a simple arrangement of parallel aluminum or copper plates separated by air dielectric might be used, as in Fig. 121, where *s* is the separation between the plates.

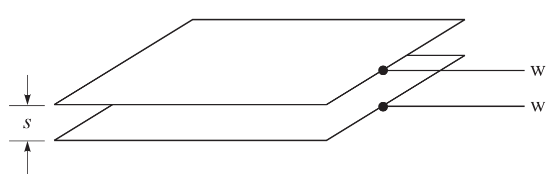


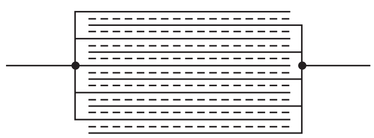
Fig. 121. (w, w are wires used to connect the capacitor into the circuit).

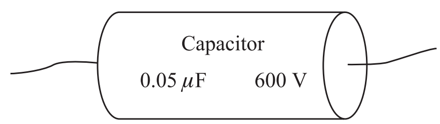
When a potential difference exists between the plates of a capacitor *electrical energy is stored in the electric field between the plates*, the amount of such stored energy being

1. directly proportional to the plate area A,
2. inversely proportional to the distance s between the plates,
3. directly proportional to the "dielectric constant K" of the material between the plates,
4. directly proportional to the square of the potential difference between the plates.

A brief discussion of these four items follows, beginning with the *plate area A* and the *plate separation s*.

From (1) and (2) it follows that, to get maximum energy storage, the plate area A should be *large* and the plate separation s should be *small*. To meet these requirements, practical capacitors are often made of alternate strips of aluminum foil and paper dielectric, as illustrated in the "side view" in the figure to the left below, where the dashed lines represent the edges of the strips of paper dielectric.





After the layers of aluminum foil and paper are in place, the whole assembly is then rolled tightly into the form of a cylinder, thus producing a capacitor having large plate area and small plate separation. Wire leads are then soldered to the aluminum plates, after which the unit is enclosed in a protective cardboard cover. The required technical information is then printed on the cover, producing a finished capacitor of the form shown in the figure to the right, above. Such capacitors are convenient to use and have a large amount of capacitance, relatively, in comparison with the amount of space they occupy.

Next, concerning item (3), the *dielectric constant K* of a material is defined as the ratio of the capacitance of a capacitor *with the material between the plates* to the capacitance *with vacuum between the plates*, where $K = 1$ for vacuum (also, $K = 1$ for air, for all practical purposes). In the definition of K, it's assumed that the dielectric material completely fills the space between the plates.

Finally, in connection with item (4) we have the problem of "voltage breakdown" of the dielectric material, which limits the amount of potential difference (volts) that can be safely applied between the plates of a given capacitor. Let us define that *the* DIELECTRIC STRENGTH *of a material* is the *maximum value of* ELECTRIC FIELD STRENGTH that the material can withstand without breaking down and permitting the passage of current. From the definition of "electric field strength" (volts per meter), it follows that the *field strength, E, between the plates of a capacitor* is equal to the *potential difference in volts, divided by the plate separation in meters*; that is

$$E = v/s \quad \text{volts per meter} \tag{183}$$

where v is the potential difference in volts between the plates, and s is the plate separation in meters.

For example, suppose $v = 100$ volts. If $s = 0.01$ meter (1 centimeter), then $E = 100/0.01 = 10{,}000$ volts per meter, but if $s = 0.001$ meter (1 millimeter), then $E = 100/0.001 = 100{,}000$ volts per meter, and so on. Equation (183) is simply a numerical way of expressing the fact that the *smaller* the separation between the plates, the *greater* is the tendency for a given voltage v to cause breakdown of the dielectric material between the plates.

Thus, when choosing capacitor dielectric material, the values of the dielectric constant K and the maximum permissible value of E must both be taken into account. A table of values for a few dielectric materials is given below, where small k stands for ''kilo'' (meaning ''thousand''). Thus, 2 kV = 2 kilovolts = 2000 volts, and ''kV/mm'' means ''kilovolts per millimeter'' (1 millimeter = 0.001 meter).

Material	Dielectric constant K	Maximum allowable E (kV/mm)
dry air	1.0	3.0
paraffined paper	2.5	150
plastic film	6 to 12	10 to 50
mica	6.5	175

Problem 111
If, in Fig. 121, the plate separation is 0.00065 meters and the dielectric material is dry air, what is the maximum voltage that should be applied to the capacitor?

In the discussion prior to Fig. 121, the term ''capacitance'' was introduced as being a measure of the ability of a capacitor to *store electric charge*. In this regard, let us suppose a current of i amperes is flowing into, and out of, a certain capacitor in the direction shown in Fig. 122.

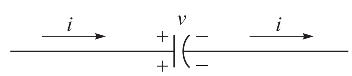


Fig. 122

Since we have agreed to regard ''electric current'' as a ''flow of positive electric charge,'' it follows that, in Fig. 122, an *excess* of positive charge is accumulating on the left-hand plate, with a corresponding *deficiency* of positive charge on the right-hand plate.

You will recall that electric charge is denoted by q and is measured in ''coulombs.'' The situation in Fig. 122 is that positive charge, accumulating on the left-hand plate, *repels* an equal amount of positive charge out of the right-hand plate, leaving the right-hand plate *negatively charged*. Thus if, at any instant of time, one plate of a capacitor has a charge of $+q$ coulombs, the other plate has an equal but opposite charge of $-q$ coulombs.

We should of course note that, while the same current i flows *into* a capacitor as flows *out*, no current actually flows *through* the dielectric material between the plates.* What happens is that, as the current continues to flow, ENERGY *is being stored in the electric field between the plates*, with a POTENTIAL DIFFERENCE building up between the two plates. In Fig. 122, for example, as the current continues to flow in the direction shown, a potential difference of v volts, having the polarity as shown, builds up across the capacitor.

With the foregoing in mind, let us return to the term ''capacitance'' which, as previously stated, is to be a measure of the ability of a capacitor to ''store electric charge.''

* For convenience, however, we do sometimes speak of the current ''through'' a capacitor.

In this regard, it's apparent that, for any given capacitor, the *amount* of stored charge depends *first* upon the amount of voltage, v, and *second* upon items (1), (2), and (3), listed following Fig. 121, that is, upon the physical *construction* of the capacitor in question. All these factors are taken into account by defining that the "CAPACITANCE of a capacitor" is equal to the *ratio* of the *magnitude of the stored charge* to the *potential difference* between the plates; thus, by definition,

$$C = \frac{q}{v} \tag{184}$$

where $q =$ magnitude of charge, in coulombs, stored on either plate, $v =$ potential difference, in volts, between the plates, $C =$ "capacitance" of the capacitor, in FARADS (for Michael Faraday).

Capacitance is thus measured in "coulombs per volt," which is called "farads." The farad is a very large unit of capacitance; in almost all practical work we deal with "microfarads" (millionths of a farad) and "picofarads" (millionths of a microfarad). Letting "F" denote farads, "µF"* (or sometimes "mfd") microfarads, and "pF" picofarads, the conversion factors are

$$\begin{aligned} \mathrm{F} \times 10^{6} &= \mu\mathrm{F} \qquad \text{(and thus),} \qquad & \mathrm{F} &= \mu\mathrm{F} \times 10^{-6} \\ \mathrm{F} \times 10^{12} &= \mathrm{pF} & \mathrm{F} &= \mathrm{pF} \times 10^{-12} \\ \mu\mathrm{F} \times 10^{6} &= \mathrm{pF} & \mu\mathrm{F} &= \mathrm{pF} \times 10^{-6} \end{aligned}$$

Problem 112

If a dc voltage of 290 volts is applied to a capacitor having 0.015 µF of capacitance, what magnitude of charge is stored on either plate?

We have noted that ENERGY is stored in the electric field between the plates of a capacitor. A formula, giving the amount of such energy, can be found as follows.

To begin, let us recall that, basically, the "potential difference" in volts between the two plates is equal to the *work*, W, in joules, required to move one coulomb of charge against the field from the negative plate to the positive plate. Thus "volts" is basically equal to "joules divided by coulombs," $v = W/q$.

Now suppose we have a capacitor with *zero volts* potential difference between the two plates, and then suppose we begin to move very small amounts of positive charge from one plate to the other plate. Suppose we continue to do this until we have transferred a total of q coulombs of charge, thus producing a potential difference of v volts between the plates.

Doing the above is equivalent to moving an *average* amount of charge of $q/2$ coulombs through a potential difference of v volts, and thus the *total work done* is (from above, work = volts × coulombs)

$$W = v(q/2) = vq/2 \text{ joules}$$

which, since there are no losses due to friction, is now all stored as POTENTIAL ENERGY in the electric field between the plates of the capacitor. Or, by eq. (184), writing "Cv" in place of "q," the above equation becomes

$$W = \tfrac{1}{2} C v^{2} \tag{185}$$

where $C =$ capacitance of the capacitor, in farads, $v =$ potential difference in volts between the two plates, and $W =$ energy, in joules, stored in the electric field.

* The Greek letter "µ" (mu) indicates multiplication by 10^{-6}.

It is important to note that the amount of energy stored is proportional to the capacitance C and the *square* of the voltage v. Further notes, regarding resistance and capacitance, appear in the Appendix.*

7.7 Capacitors in Series and in Parallel

In practical work it is sometimes necessary to use both series and parallel connections of capacitors, Let us first investigate the *series connection* with the aid of Figs. 123 and 124. In Fig. 123, the C's denote the capacitance, in farads, of each of the individual series-connected capacitors.

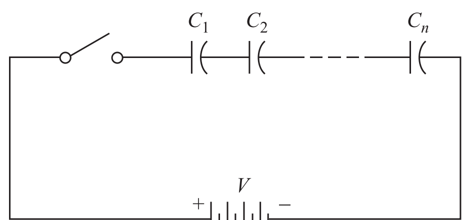


Fig. 123

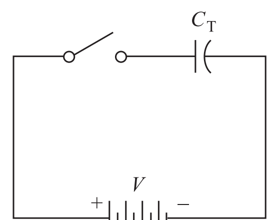


Fig. 124

In Fig. 124, C_T denotes the capacitance of a *single capacitor* that would have the *same capacitance* as the series connection of the n individual capacitors in Fig. 123. This means that theoretically, for purposes of analysis, the n series-connected capacitors of Fig. 123 can be replaced with the single equivalent capacitor of C_T farads of Fig. 124.

A formula for finding the value of C_T can be found by making use of the fact that the *magnitude of charge* is the same on both plates of a capacitor. This is because an amount of positive charge, flowing *into* one plate, repels the same amount of positive charge *out* of the other plate (see discussion following Fig. 122).

With this in mind, consider Fig. 123. When the switch is closed, a charge q flows *into* the left-hand plate of C_1, thus forcing the same amount of charge *out* of the right-hand plate of C_1 *into* the left-hand plate of C_2. This forces the same amount of charge *out* of the right-hand plate of C_2 *into* the left-hand plate of the next capacitor, and so on down the line, with the result that *all the series-connected capacitors* in Fig. 123 have the *same magnitude of charge* of q coulombs on their plates. Note that this satisfies the basic requirement that, at all times, charge flowing *out* of the positive terminal of the battery must be equal to the charge flowing *into* the negative terminal.

With the above in mind, now make use of the equation $v = q/C$ (eq. (184)). Using this equation, and remembering that all capacitors in Fig. 123 have the *same charge* q, we have, for Fig. 123,

$$\begin{aligned}
V_1 &= q/C_1 \quad (\text{where } V_1 = \text{voltage on capacitor } C_1) \\
V_2 &= q/C_2 \quad (\text{where } V_2 = \text{voltage on capacitor } C_2) \\
&\ \ \vdots \\
V_n &= q/C_n \quad (\text{where } V_n = \text{voltage on capacitor } C_n)
\end{aligned}$$

* See note 14 in Appendix.

Since the sum of the left-hand sides of the above equations is equal to the sum of the right-hand sides, we have that

$$V_1 + V_2 + \cdots + V_n = q(1/C_1 + 1/C_2 + \cdots + 1/C_n)$$

Or, since $V_1 + V_2 + \cdots + V_n =$ the battery voltage V, the last equation becomes

$$V = q\left(\frac{1}{C_1} + \frac{1}{C_2} + \cdots + \frac{1}{C_n}\right) \tag{186}$$

Now let us apply the same battery voltage V to the equivalent capacitance C_T in Fig. 124. Since C_T is to be equivalent to the circuit of Fig. 123, it must carry the same charge q, and hence, by eq. (184), it must be true that

$$V = \frac{q}{C_T} \tag{187}$$

Since the left-hand sides of the last two equations are equal, their right-hand sides are also equal, and upon making use of this fact we get the desired relationship

$$\frac{1}{C_T} = \frac{1}{C_1} + \frac{1}{C_2} + \cdots + \frac{1}{C_n} \tag{188}$$

Or, if we wish, we can invert both sides of the last equation and write that

$$C_T = \frac{1}{1/C_1 + 1/C_2 + \cdots + 1/C_n} \tag{189}$$

Thus either equation, (188) or (189), allows us to calculate the equivalent capacitance, C_T, of a *series connection* of n capacitors, where $C_1, C_2, \ldots, C_n$ are the capacitances of the individual capacitors.

When using series capacitors we must be able to calculate the *voltage* that will appear across *each capacitor* when the series string is connected to a battery of V volts, as in Fig. 123. This is important, because excessively high voltage across one of the capacitors could cause "voltage breakdown" of that capacitor, with subsequent failure of the whole series string. A formula that will allow us to calculate such a voltage can be found as follows.

Let C_x be the capacitance of any one of the series capacitors in Fig. 123, and let V_x be the voltage on that capacitor. Then, from the general equation $q = Cv$, we have that, for this capacitor,

$$q = C_x V_x$$

but also, by eq. (187),

$$q = VC_T$$

where V is the battery voltage. From inspection of the above two equations we see that

$$C_x V_x = VC_T$$

giving us the important result

$$V_x = V\,\frac{C_T}{C_x} \tag{190}$$

Next, let us consider the problem of finding the equivalent capacitance of a *parallel connection* of n capacitors. Such a parallel connection is shown in Fig. 125, with the equivalent single capacitor of capacitance C_T shown in Fig. 126.

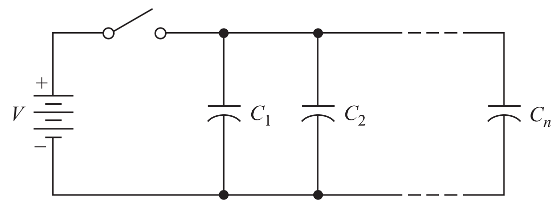


Fig. 125

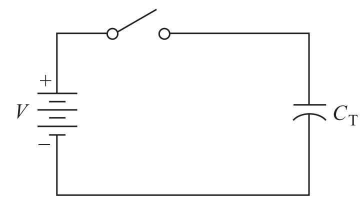


Fig. 126

In Fig. 125 it's evident that the *total charge q* delivered by the battery is the *sum of the charges* delivered to the individual capacitors, which is (making use of eq. (184)) equal to

$$q = V(C_1 + C_2 + \cdots + C_n)$$

The same charge in Fig. 126 is

$$q = VC_T$$

Comparison of this equation with the preceding equation shows that

$$C_T = C_1 + C_2 + \cdots + C_n \tag{191}$$

Equation (191) is thus the formula for calculating the equivalent capacitance of a PARALLEL connection of n individual capacitors.

Problem 113
Given four capacitors, of capacitances 0.09 μF, 0.17 μF, 0.12 μF, and 0.55 μF,

(a) Find the equivalent capacitance if the four are connected in series.
(b) Find the equivalent capacitance if the four are connected in parallel.

Problem 114
A dc voltage of 450 volts is applied to three series-connected capacitors having capacitances of 0.15 μF, 0.06 μF, and 0.48 μF. Is it theoretically sufficient, in this case, to specify that all three capacitors have a voltage rating of at least 300 volts?

Problem 115
In problem 114, find the total amount of energy stored in the three capacitors.

CHAPTER 8

Reactance and Impedance. Algebra of ac Networks

In this chapter we begin the study of the algebra of networks containing INDUCTANCE AND CAPACITANCE for SINUSOIDAL applied voltages. (The reasons why sinusoidal waves are so important are noted in the introduction to section 5.4.)

8.1 Inductive Reactance. Impedance

We begin with the basic case of inductance in series with resistance, to which a sine wave of peak voltage V_p is applied, as in Fig. 127.

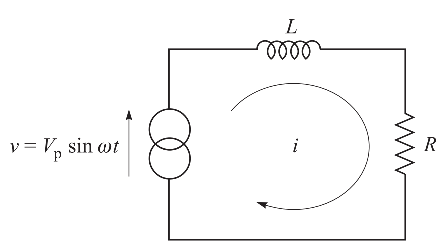


Fig. 127

In Fig. 127, L is inductance in henrys, and R is resistance in ohms. Also in the figure, v and i denote instantaneous values of voltage and current, with the voltage and current arrows having their usual meaning.*

* Before continuing, it will be wise to first review all of section 5.6.

In Fig. 127 the independent variable is *time*, t, in which $t = 0$ at the instant the sine wave of voltage is first applied to the circuit. Analysis shows, and experiment verifies, that the resulting current consists of the SUM OF TWO TERMS, one term being the TRANSIENT component of the current and the other term being the STEADY-STATE component of the current. As a matter of fact, analysis shows and experiment verifies that the current in Fig. 127 has the mathematical form

$$i = \underbrace{I_0 \epsilon^{-Rt/L}}_{\substack{\text{"transient"}\\ \text{component}}} + \underbrace{I_\mathrm{p} \sin(\omega t - \phi)}_{\substack{\text{"steady-state"}\\ \text{component}}} \tag{192}$$

In the equation, note, first, that the *transient* component is a NEGATIVE EXPONENTIAL FUNCTION OF TIME which *rapidly decreases in value* as time increases (see Fig. 18-A in note 13 in the Appendix). Thus the "transient" component *vanishes* very quickly, leaving only the permanent, sinusoidal, "steady-state" component.

As a matter of fact, in much practical work the effect of the transient component *can be completely ignored*; that is, only the PERMANENT SINUSOIDAL STEADY-STATE RESPONSE is of interest in most practical work. Let us therefore examine, in more detail, the steady-state component in eq. (192).*

First, taking the applied voltage, $V_\mathrm{p} \sin \omega t$, as the "reference wave," notice that the steady-state current wave is also a sine wave, but one that LAGS the reference voltage wave by ϕ radians (ωt is in radians). (At this point it should be noted that the current will ALWAYS "lag" the voltage in an "inductive" circuit; this fact will be made evident in the discussion which follows.)

To begin, let us suppose, in eq. (192), that the transient component of current has died out, so that only the final STEADY-STATE SINUSOIDAL CURRENT remains. It is this "sinusoidal steady-state condition," for Fig. 127, that we now wish to study in detail.

To begin, note that there are two "voltage drops" in Fig. 127, one across R, the other across L. These two voltage drops are depicted graphically, with respect to the current wave, in Fig. 128, where I_p is the peak value of the sinusoidal current.

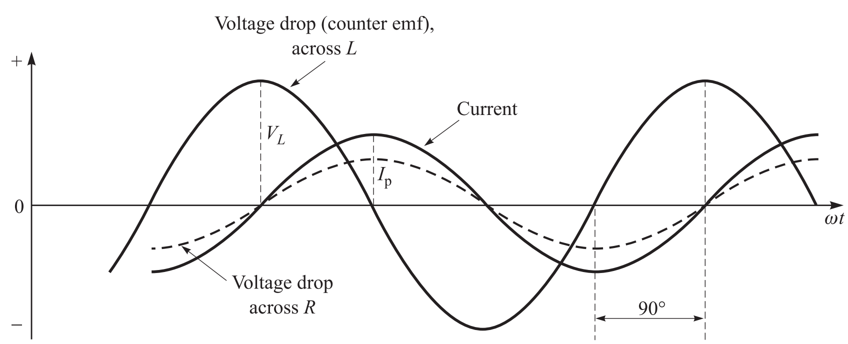


Fig. 128

First, as you know, the wave of voltage drop across a pure *resistance* of R ohms is exactly "in phase" with the wave of current through the resistance, a fact indicated graphically in Fig. 128.

Next, in the figure, note that the wave of voltage drop, the counter-emf, across the inductor LEADS the current wave through the inductor by 90 degrees. This is in

* I_p is the peak value of the steady-state sinusoidal current.

accordance with the basic principles outlined in section 7.4, in which we learned that the voltage induced in an inductor depends NOT upon the AMOUNT of current but only upon the RATE OF CHANGE of the current. In the case of the above figure, note that the RATE OF CHANGE of the current i is MAXIMUM when $i = 0$ and ZERO when $i = I_p$. This is the natural phenomenon that causes the voltage across an inductor of zero resistance to lead the inductor current by 90°. Thus, in the sinusoidal steady state, where (by eq. (192)) $i = I_p \sin(\omega t - \phi)$, the *voltage drop across L*, v_L, at any instant will be of the form

$$v_L = V_L \sin(\omega t - \phi + 90^\circ) \tag{193}$$

because v_L leads $I_p \sin(\omega t - \phi)$ by 90 degrees. Note that V_L (capital V, sub L) denotes the PEAK VOLTAGE DROP ACROSS L.

Next, let's consider the VOLTAGE EQUATION for Fig. 127 in the sinusoidal steady state. Such an equation must show that at *all instants of time* the *sum* of the voltage drops across R and L is equal to the applied generator voltage. The equation can be arrived at as follows.

First, the instantaneous voltage drop *across* R is $v_R = Ri = RI_p \sin(\omega t - \phi)$, in which we see that the PEAK VOLTAGE DROP ACROSS R is RI_p.

Next, V_L, the PEAK VOLTAGE DROP ACROSS L, depends, first of all, upon the magnitude of the MAXIMUM RATE OF CHANGE OF CURRENT, which occurs only at times where $i = 0$. This fact can be seen from examination of Fig. 128, in which it can be seen that the current curve is steepest (maximum "amperes per second"), when $i = 0$. Now, if we carefully inspect Fig. 128 we will see that the MAGNITUDE OF THE STEEPNESS of the current at *$i = 0$ increases* if (1) the PEAK VALUE of current, I_p, is increased, and (2) if the FREQUENCY, ω, is increased (for example, note the effect, on degree of steepness, if the number of current waves in Fig. 128 were doubled). These points, plus the fact that the peak value V_L is also proportional to the *value of L itself*, leads us, correctly, to the conclusion that the PEAK VALUE OF THE VOLTAGE DROP ACROSS L is $V_L = \omega L I_p$, and upon substituting this value into eq. (193), and remembering that the peak voltage drop across R is RI_p, we have that the SINUSOIDAL STEADY-STATE VOLTAGE EQUATION for Fig. 127 is

$$V_p \sin \omega t = RI_p \sin(\omega t - \phi) + \omega L I_p \sin(\omega t - \phi + 90^\circ) \tag{194}$$

where $\phi =$ phase angle of *current* with respect to the reference generator voltage.

In the above equation the quantity "ωL" is measured in *ohms** and is called INDUCTIVE REACTANCE. "Inductive reactance" is denoted by X_L; thus, $X_L = \omega L = 2\pi f L$ ohms; thus, by Ohm's law, $X_L I_p = \omega L I_p =$ peak voltage drop *across the inductor*, as shown in eq. (194).

Now, as pointed out in connection with Figs. 93 through 95 in section 5.6, the three rotating voltage components in eq. (194) can be regarded as three *stationary vector components*, as shown in Fig. 129, where V is now the *rms generator voltage*, taken as being the reference vector, and I is the rms current vector. Note that RI, the rms voltage drop across the resistance R, is "in phase" with the current vector I, while ωLI, the rms voltage drop across the inductor, *leads* the current vector I by 90 degrees.

In the figure, note that, in accordance with the basic Kirchhoff voltage law, the *vector sum* of the two voltage drops RI and ωLI is equal to the applied reference voltage V. Or, using the "overscore" or "bar" notation to indicate "vector quantity" (note 4 in the Appendix), the *algebraic statement* for Fig. 129 is

$$R\bar{I} + \omega L(\bar{I} + 90^\circ) = \bar{V} \tag{195}$$

* See note 15 in Appendix.

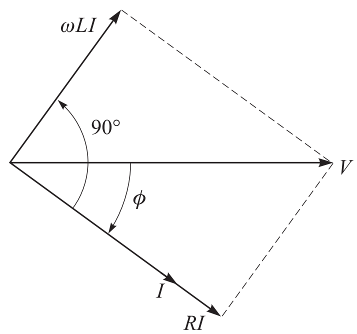


Fig. 129

where the notation $\omega L(\bar{I}+90^\circ)$ is used to indicate that the voltage drop across the inductor *leads* the current by 90 degrees.

We learned, in section 6.8, that COMPLEX NUMBERS obey the same parallelogram law as do vector quantities. Hence, as far as ALGEBRAIC OPERATIONS are concerned, we can regard $\bar{V}$ and $\bar{I}$, in eq. (195), as being *complex numbers* of the forms $\bar{V} = V' + jV''$ and $\bar{I} = I' + jI''$, where the "primes" indicate the "real" and "imaginary" parts of $\bar{V}$ and $\bar{I}$. However, before applying this concept to eq. (195), it should be noted that *multiplying a complex number by "j"* has the SOLE EFFECT of *rotating the complex number through an angle of +90 degrees*; thus, if $\bar{Z}$ is a complex number, then $j\bar{Z}$ has the same magnitude as $\bar{Z}$ but is rotated through $+90^\circ$.*

Hence, in eq. (195), let $\bar{I}$ and $\bar{V}$ be represented as *complex numbers* in the complex plane; thus, to indicate that the angular position of $\omega L\bar{I}$ is to be increased by 90 degrees, all we need do is write $j\omega L\bar{I}$ in place of $\omega L(\bar{I}+90^\circ)$; thus

$$R\bar{I} + j\omega L\bar{I} = \bar{V} \tag{196}$$

and thus, upon solving for $\bar{I}$, we have

$$\bar{I} = \frac{\bar{V}}{R + j\omega L} \tag{197}$$

in which $\bar{I}$ is the rms vector in the BASIC SERIES "RL" CIRCUIT of Fig. 127, where $\bar{V}$ is the applied sinusoidal voltage vector. In our work we'll generally take $\bar{V}$ to be the "reference" vector, in which case $\bar{V} = V\underline{/0^\circ} = V$, a real number (the magnitude of $\bar{V}$).

In the above equation, the *denominator* is called the IMPEDANCE of the circuit, which, for the case of Fig. 127, is a measure of the *combined effect* of the resistance R and the inductive reactance ωL.

Impedance is denoted by "$\bar{Z}$"; thus, the "impedance" of the basic series RL circuit of Fig. 127 is

$$\bar{Z} = R + j\omega L \tag{198}$$

showing that "impedance" is a complex number; thus eq. (197) becomes

$$\bar{I} = \frac{\bar{V}}{\bar{Z}} \tag{199}$$

which is OHM's LAW, in complex form, for the sinusoidal steady-state condition.

---

* See note 16 in Appendix.

In regard to the above, we will sometimes wish to deal only with the *magnitudes* of the complex numbers. To do this, all we need remember is that*

(a) the MAGNITUDE of the PRODUCT of complex numbers is equal to the PRODUCTS OF THE MAGNITUDES, and
(b) the MAGNITUDE of the QUOTIENT of two complex numbers is equal to the QUOTIENT OF THE TWO MAGNITUDES.

Hence, applying these rules to eq. (199) we have that

$$|\bar{I}| = \frac{|\bar{V}|}{|\bar{Z}|} \qquad \text{and} \qquad |\bar{I}|\,|\bar{Z}| = |\bar{V}| \tag{200}$$

Next, the *vector diagram* representation for $\bar{Z} = R + j\omega L$ (eq. (198)) is shown in Fig. 130.

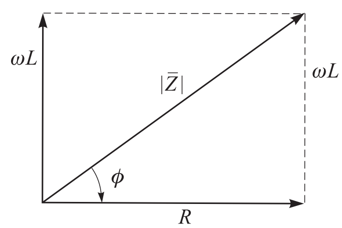


Fig. 130

Comparison of Fig. 130 with Fig. 129 shows that "ϕ" in Fig. 130, is the phase angle *between the voltage and current vectors* shown in Fig. 129. Figure 130 is spoken of as an "impedance triangle" and shows, for the series RL circuit of Fig. 127, that

$$\bar{Z} = |\bar{Z}|\underline{/\phi} \tag{201}$$

where, from Fig. 130,

$$|\bar{Z}| = \sqrt{R^2 + \omega^2 L^2} \tag{202}$$

and

$$\phi = \arctan \frac{\omega L}{R} \text{ (lagging)} \tag{203}$$

Problem 116
In the series RL circuit of Fig. 127, let the sinusoidal reference voltage be 115 volts rms, the resistance be 28 ohms, and the inductance be 0.12 henry. If the frequency is 60 Hz, find

(a) magnitude of rms current, (*Answer*: 2.1615 amperes)
(b) phase angle of current, (*Answer*: 58.245° lagging)
(c) reading of voltmeter placed across L.

8.2 *RL* Networks

In section 8.1 we introduced the basic series "RL" circuit of Fig. 127, to which is applied a sinusoidal voltage of V volts rms. We found that, by writing the *combined effects* of R and

* See eqs. (161) and (163) in section 6.6.

L in the form of a *complex number*, $\bar{Z} = R + j\omega L$ (called the “impedance”), then applying the “algebra of complex numbers,” we were able to find the magnitude and phase angle of the *rms current* flowing in the circuit.

Actually, the procedure of section 8.1 is not limited to the simple series circuit of Fig. 127 but applies, as well, to ANY type of series, parallel, or series-parallel connection of R and L components. The procedure is as follows.

First, in drawing circuit diagrams, let us, for convenience, agree to represent the series impedance, $\bar{Z} = R + j\omega L$, by the single symbol “,” which we’ll label $\bar{Z}$, as shown below,

where $\bar{Z} = R + j\omega L$.

For the case where $L = 0$, the symbol then represents the real number R (a pure resistance) or, if $R = 0$, the symbol then represents the imaginary number $j\omega L$ (a pure inductive reactance).

The simplest case consists of a *series connection* of n such impedances, as shown in Fig. 131.

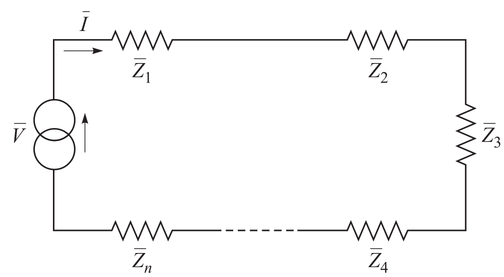


Fig. 131

In such a case the generator sees a *total impedance*, $\bar{Z}_{\mathrm{T}}$, equal to the *sum* of all the impedances; thus, by Ohm’s law (eq. (199)) we have, for Fig. 131, that

$$\bar{I} = \frac{\bar{V}}{\bar{Z}_{\mathrm{T}}} = \frac{\bar{V}}{\bar{Z}_1 + \bar{Z}_2 + \cdots + \bar{Z}_n} \tag{204}$$

in which each impedance has the general form $\bar{Z} = R + j\omega L$. Or, by the law of “addition of complex numbers” (section 6.2), the above equation can also be written in the form

$$\bar{I} = \frac{\bar{V}}{\bar{Z}_{\mathrm{T}}} = \frac{\bar{V}}{(R_1 + R_2 + \cdots + R_n) + j\omega(L_1 + L_2 + \cdots + L_n)} \tag{205}$$

Problem 117

A certain series circuit consists of three resistances of 8, 10, and 12 ohms, and two inductor coils of 12 and 25 millihenrys (mH), where 1 mH = 0.001 henry. If the applied sinusoidal reference voltage is 95 volts rms, and $\omega = 2\pi f = 1000$ radians/second, find

(a) magnitude of rms current,
(b) phase angle of current,
(c) voltage drop across the 25 mH coil.

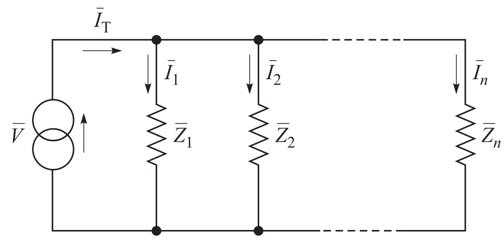


Fig. 132

Next consider the case of a purely PARALLEL network, as shown in Fig. 132.

Now let $\bar{I}_T$ = generator current $= \bar{I}_1 + \bar{I}_2 + \cdots + \bar{I}_n$, and $\bar{Z}_T$ = total impedance seen by the generator.

By Ohm's law:

$$\bar{I}_T = \frac{\bar{V}}{\bar{Z}_T} = \bar{I}_1 + \bar{I}_2 + \cdots + \bar{I}_n \tag{206}$$

Or, since the *same voltage* $\bar{V}$ is applied to all the branches, eq. (206) becomes, again by Ohm's law,

$$\frac{\bar{V}}{\bar{Z}_T} = \frac{\bar{V}}{\bar{Z}_1} + \frac{\bar{V}}{\bar{Z}_2} + \cdots + \frac{\bar{V}}{\bar{Z}_n}$$

that is,

$$\bar{V}\left(\frac{1}{\bar{Z}_T}\right) = \bar{V}\left(\frac{1}{\bar{Z}_1} + \frac{1}{\bar{Z}_2} + \cdots + \frac{1}{\bar{Z}_n}\right)$$

thus,

$$\frac{1}{\bar{Z}_T} = \frac{1}{\bar{Z}_1} + \frac{1}{\bar{Z}_2} + \cdots + \frac{1}{\bar{Z}_n} \tag{207}$$

Thus, in words, eq. (207) shows that "in a purely PARALLEL network, the RECIPROCAL *of the total impedance* seen by the generator is equal to the SUM OF THE RECIPROCALS *of the impedances of the individual branches*." Or, inverting both sides of eq. (207), we have the equivalent result that

$$\bar{Z}_T = \frac{1}{1/\bar{Z}_1 + 1/\bar{Z}_2 + \cdots + 1/\bar{Z}_n} \tag{208}$$

In words, eq. 208 says that, in a PARALLEL circuit, the *total impedance* seen by the generator is equal to the RECIPROCAL of the SUM OF THE RECIPROCALS of the individual impedances.

If (as often happens in practical work) the network consists of just *two* parallel impedances, then eq. (207) becomes

$$\frac{1}{\bar{Z}_T} = \frac{1}{\bar{Z}_1} + \frac{1}{\bar{Z}_2}$$

which (after multiplying both sides by $\bar{Z}_1\bar{Z}_2$ and then inverting) gives the special formula

$$\bar{Z}_T = \frac{\bar{Z}_1\bar{Z}_2}{\bar{Z}_1 + \bar{Z}_2} \tag{209}$$

for the impedance looking into a parallel connection of *two* impedances.

In using the foregoing equations, all we need do is write each impedance in the form of a complex number $R + j\omega L$, and then apply the algebra of complex numbers, remembering $j^2 = -1$.

Problem 118
In Fig. 133, the resistance and inductance values are in ohms and henrys.

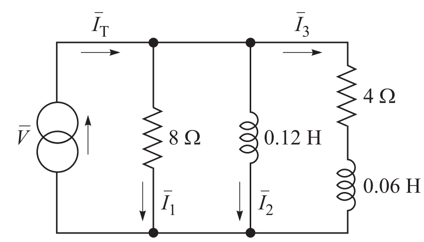


Fig. 133

Given that the applied reference voltage $\bar{V}$ is 60 volts rms, and that $\omega = 2\pi f = 100$ rad/sec, find*

(a) impedance $\bar{Z}_T$ seen by generator,
(b) generator current $\bar{I}_T$,
(c) phase angle ϕ between generator voltage and generator current,
(d) $\bar{I}_1 =$
(e) $\bar{I}_2 =$
(f) $\bar{I}_3 =$
(g) verify that the sum of the answers to (d), (e), (f) equals the answer to (b).
(h) using $\bar{V} = V\underline{/0^\circ} = 60$ volts rms as the reference vector, make a rough sketch of the vector diagram showing the answers to (d) through (g).

Problem 119
The load on a generator of $V\underline{/0^\circ}$ volts rms consists of a resistance of R ohms in parallel with a coil of inductance L henrys, as shown in Fig. 134.

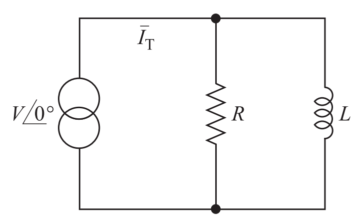


Fig. 134

Using eq. (209), write the equation for the generator current $\bar{I}_T$. Write the final answer in the rectangular form $I' + jI''$.

* We are here assuming there is NO MAGNETIC COUPLING between the two coils; that is, the magnetic field of each coil is confined to that coil only. The important case where this is not true is studied in Chap. 10.

Let us pause here, just a moment, to again point out the close correspondence between the procedures used in dc circuit analysis and steady-state ac circuit analysis.

Consider, for example, eqs. (32) and (34) in section 2.6; note that, to convert these two equations to the "ac" case (eqs. (207) and (209)), all we need do is replace the Rs with $\bar{Z}$s (in which $\bar{Z} = R + j\omega L$) and apply the "algebra of complex numbers," remembering that $j^2 = -1$.

In the same way, the treatment of series-parallel ac networks exactly parallels the treatment of dc networks in section 2.7, with the Rs replaced by $\bar{Z}$s.

Likewise, the basic dc procedure of "loop current" analysis, explained in section 4.4, applies also to ac circuit analysis, remembering, of course, that $j^2 = -1$. In this regard, consider the example of Fig. 135, containing two ac generators of known voltages $\bar{V}_1$ and $\bar{V}_2$ and three unknown "loop currents," as shown.

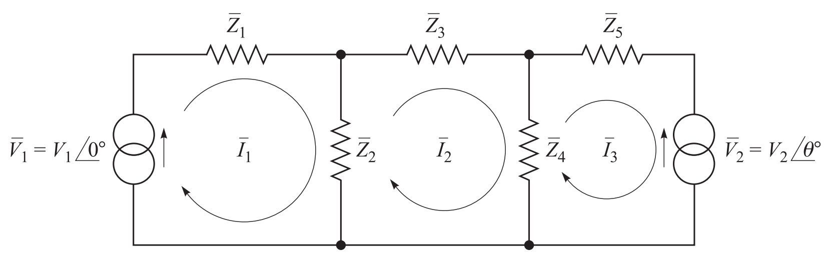


Fig. 135

All the quantities in Fig. 135 denote *complex numbers* representing both the *rms vector values* of voltage and current and the passive network components (the $\bar{Z}$s).

In this regard, using eq. (158) in Chap. 6, the two generator voltages can be put in the complex "rectangular" form; thus

$$\bar{V}_1 = V_1\underline{/0°} = V_1(\cos\ 0 + j\sin\ 0) = V_1(1 + j0) = V_1$$

and

$$\bar{V}_2 = V_2\underline{/\theta°} = V_2(\cos\ \theta + j\sin\theta) = V_2' + jV_2''$$

In the example of Fig. 135, let us assume that the voltages and impedances are *given*, and that the problem is to *find* the unknown values of the currents. In doing this, the voltage and current *arrows* shown in the figure are used in conjunction with KIRCHHOFF's VOLTAGE LAW, following the *same basic procedure* outlined for *dc networks* in section 4.4, except now we're dealing with complex (vector) quantities instead of scalar quantities. Also, the VOLTAGE DROPS will now be of the form $\pm\bar{Z}\bar{I}$ instead of $\pm RI$ as in the dc case. Thus, upon applying the rules of section 4.4, we have that the equations for the ac network of Fig. 135 are

$$\begin{aligned}
(\bar{Z}_1 + \bar{Z}_2)\,\bar{I}_1 \qquad & -\bar{Z}_2\,\bar{I}_2 & +0\,\bar{I}_3 = V_1 \\
-\bar{Z}_2\,\bar{I}_1 + (\bar{Z}_2 + \bar{Z}_3 & + \bar{Z}_4)\,\bar{I}_2 & -\bar{Z}_4\,\bar{I}_3 = 0 \\
0\,\bar{I}_1 \qquad & \bar{Z}_4\,\bar{I}_2 + (\bar{Z}_4 & + \bar{Z}_5)\,\bar{I}_3 = -\bar{V}_2
\end{aligned}$$

In regard to Fig. 135, it should be noted that the value of voltage $\bar{V}_2$ must be given with respect to the "reference voltage" $\bar{V}_1 = V_1\underline{/0°}$. Fundamentally, this is done by remembering that the voltages are sinusoidal waves of the same frequency, the wave of V_2 being θ degrees "out of phase" with V_1. (See discussion given with eq. (108) in section 5.6.)

Lastly, to solve the resulting simultaneous equations for any particular value of current it will generally be easiest to use the method of determinants.

Problem 120
Repeat problem 118 for the value of generator current, this time using the method of loop currents.

8.3 Capacitive Reactance. *RC* Networks

We begin with the basic case of a capacitor in series with a resistance, to which a sine wave of peak voltage V_p is applied, as in Fig. 136.

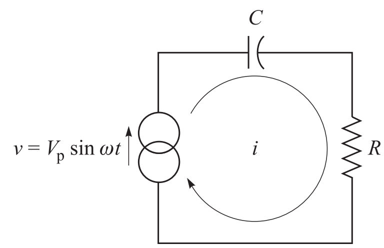


Fig. 136

In the figure, C is a capacitor of C farads and R is resistance in ohms. Also (as in Fig. 127) v and i denote instantaneous values of voltage and current. Again, as in Fig. 127, we'll be interested only in the permanent SINUSOIDAL STEADY STATE. In this case, the author feels that it's best not to get sidetracked by too many details; hence, let's commence by stating (without bothering about details to begin with) that, corresponding with eq. (194) in section 8.1, the steady-state VOLTAGE EQUATION for Fig. 136 is

$$V_p \sin \omega t = \underbrace{RI_p \sin(\omega t + \phi)}_{\substack{\text{voltage drop} \\ \text{across } R}} + \underbrace{\left(\frac{1}{\omega C}\right) I_p \sin(\omega t + \phi - 90^\circ)}_{\text{voltage drop across } C} \tag{210}$$

in which the applied voltage, $V_p \sin \omega t$, is taken as the "reference wave." Now let's compare eq. (210) with eq. (194), as follows. ($\phi =$ phase angle of *current*, with respect to $V_p \sin \omega t$, in both equations.)

Note, first, that the angular quantities ϕ and 90° have OPPOSITE SIGNS from what they have in eq. (194). This is due to the fact that the CURRENT in a capacitor LEADS THE VOLTAGE across the capacitor by 90°. This happens because FIRST, the amount of electric charge q stored in a capacitor is *proportional to capacitor voltage* ($q = vC$), and SECOND, capacitor *current* is the *rate of flow of charge* (coulombs per second). Now consider the following.

In Fig. 128 (section 8.1), suppose the voltage drop were across a capacitor instead of an inductor. Then, since $q = vC$, the curve of *charge q* would be drawn exactly "in phase" with the voltage curve. If these changes were made in Fig. 128 it would then be apparent that the *curve of the "rate of change of q"* (the current) would "lead" the curve of voltage drop by 90°.

Thus, in Fig. 136 we have that *the current* LEADS *the voltage drop across C by 90°* but is, as always, IN PHASE *with the voltage drop across R*. It thus follows that *the current wave will lead the applied voltage wave* by some intermediate angle ϕ, where ϕ will be an angle between 0° and 90°. The actual value of ϕ will depend upon the relative values of R, C, and ω in any given case.

Next, from inspection of eq. (210),

$$RI_p = \text{peak voltage drop across } R$$

and

$$(1/\omega C)I_p = \text{peak voltage drop across } C$$

Now recall, from section 5.8, that the peak value of the SUM of two sinusoidal waves of the same frequency is equal to the vector sum of the peak values of the individual sinusoids. Thus, the relationships of the peak values of the voltages in eq. (210) can be stated in the vector polar form,

$$V_p\underline{/0^\circ} = RI_p\underline{/\phi^\circ} + (1/\omega C)I_p\underline{/(\phi - 90)^\circ} \tag{211}$$

which shows that, geometrically, the two peak voltages on the right-hand side can be considered to be the adjacent and opposite sides of a right-angled triangle having V_p as the hypotenuse with ϕ the angle between RI_p and V_p (as will be shown later in connection with Fig. 138). Thus, we have that the steady-state PEAK CURRENT is equal to

$$I_p = V_p/\sqrt{R^2 + (1/\omega C)^2} \tag{212}$$

where also

$$\phi = \arctan(1/\omega RC) \tag{213}$$

Next, in Fig. 136 the voltage drop across R is, as always, *in phase* with the current; thus, since the peak voltage drop across R is equal to RI_p, and since the current LEADS the applied voltage V_p by an angle ϕ, we have the vector diagram as in Fig. 137.

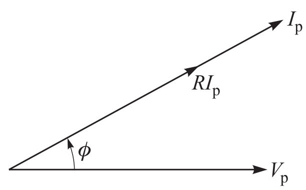


Fig. 137

Next, returning to eq. (210), the quantity $1/\omega C$ is called CAPACITIVE REACTANCE and is measured in *ohms*.* Capacitive reactance is denoted by X_C; thus, $X_C = 1/\omega C = 1/2\pi fC$ ohms.

Hence, by Ohm's law, $X_C I_p = (1/\omega C)I_p =$ peak voltage drop *across the capacitor*, which let us next consider as follows.

By the basic Kirchhoff voltage law, the *vector sum* of the *voltage across* R, RI_p, and the *voltage across the capacitor*, $(1/\omega C)I_p$, must be equal to the *applied voltage vector* V_p. This requirement, and the required relationship between R and $1/\omega C$ in eq. (212), will be satisfied only if the *voltage drop across the capacitor*, $(1/\omega C)I_p$, *lags* the current vector I_p by 90°, as shown in Fig. 138.

Now, in Fig. 138, multiply all the vector magnitudes by 0.7071, thus converting the peak values to *rms* values. Doing this, and using the usual "bar" notation to denote vector

* See note 17 in Appendix.

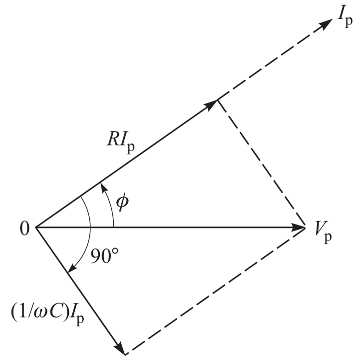


Fig. 138

quantities, we have (all rms values)

$$\bar{V} = \text{applied voltage, the reference vector}$$
$$\bar{I} = \text{current vector}$$
$$R\bar{I} = \text{vector voltage drop across } R$$
$$-j(1/\omega C)\bar{I} = \text{vector voltage drop across capacitor } C^*$$

Thus, by Kirchhoff's voltage law, the *basic vector equation* for the *series resistive-capacitive circuit* of Fig. 136 is

$$R\bar{I} - j(1/\omega C)\bar{I} = \bar{V} \tag{214}$$

or, since $-j = \dfrac{1}{j}$,

$$R\bar{I} + (1/j\omega C)\bar{I} = \bar{V} \tag{215}$$

and thus

$$\bar{I} = \frac{\bar{V}}{R - j(1/\omega C)} = \frac{\bar{V}}{R + \dfrac{1}{j\omega C}} \tag{216}$$

Equation (216) is the sinusoidal steady-state vector equation for the basic series "*RC*" circuit of Fig. 136, the vectors being represented by complex numbers on the complex plane. In our work we'll generally take $\bar{V}$ to be the "reference" vector, in which case $\bar{V} = V\underline{/0^\circ} = V + j0 = V$, a real number on the real axis of the complex plane.

In eq. (216) the *denominator* is called the IMPEDANCE of the circuit, which for the case of Fig. 136 is a measure of the combined effect of the resistance R and the capacitive reactance $1/\omega C$. As mentioned in connection with eq. (198), "impedance" is a complex number, denoted by $\bar{Z}$, and thus, by eq. (216), the impedance of the basic series RC circuit of Fig. 136 can be expressed in either of the forms

$$\bar{Z} = R - j(1/\omega C) = R + 1/j\omega C \tag{217}$$

Thus eq. (216) becomes

$$\bar{I} = \frac{\bar{V}}{\bar{Z}} \tag{218}$$

* $-j\bar{I}/\omega C$ *lags* $\bar{I}$ by 90°, as required by Fig. 138, now transferred to the complex plane.

which, as we've already seen in connection with eq. (199), is Ohm's law in complex notation for the sinusoidal steady-state condition in Fig. 136.

As we see from eq. (217), $1/\omega C$ is at right angles to R in the complex plane; hence the MAGNITUDE of the impedance in Fig. 136 is equal to

$$|\bar{Z}| = \sqrt{R^2 + (1/\omega C)^2} \tag{219}$$

Thus, in terms of magnitude, eq. (218) becomes, by eq. (200),

$$|\bar{I}| = \frac{|\bar{V}|}{|\bar{Z}|} = \frac{|\bar{V}|}{\sqrt{R^2 + (1/\omega C)^2}} \tag{220}$$

Equations (217) and (219) can be expressed in the form of the "impedance triangle" shown in Fig. 139.

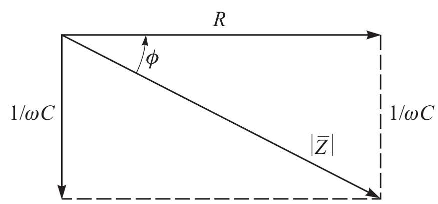


Fig. 139

Comparison of the above with Fig. 138 shows that ϕ in the above figure is the phase angle between the voltage and current vectors shown in Fig. 138, where

$$\tan\phi = (1/\omega C)/R$$

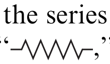
thus

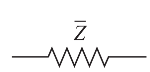
$$\phi = \arctan(1/\omega RC) \tag{221}$$

the current LEADING the applied voltage in a capacitive circuit.

Problem 121
In the series circuit of Fig. 136, let the sinusoidal reference voltage be 95 volts rms, the capacitance be 0.5 μF (microfarad), and the resistance 100 ohms. If the frequency is 4000 Hz, find

(a) magnitude of rms current,
(b) phase angle of current,
(c) reading of voltmeter placed across R,
(d) reading of voltmeter placed across C,
(e) vector sum of the voltages in (c) and (d).

The foregoing applies to ANY type of series, parallel, or series-parallel connection of R and C components. In doing this, the series circuit of Fig. 136 is the basic building block, represented by the single symbol "," which we'll label $\bar{Z}$, as shown below.



where $$\bar{Z} = R - \frac{j}{\omega C} = R + \frac{1}{j\omega C} \qquad \text{(as in eq. (217))}$$

In the above, either expression for $\bar{Z}$ can be used, but we'll more often use the form $\bar{Z} = R - j/\omega C = R - jX_C$, where $X_C = 1/\omega C$.

If, in a given case, ONLY RESISTANCE is present, then $\bar{Z} = R$, a real number; or, if ONLY CAPACITANCE is present, then $\bar{Z} = -j/\omega C = 1/j\omega C$, an imaginary number.

The simplest case consists of a *series connection* of n such impedances (see Fig. 131). In such a case the current is given by eq. (204) in section 8.2, in which each impedance now has the general form $\bar{Z} = R - j/\omega C$ (or $R + 1/j\omega C$). Hence, for the case of capacitance (instead of inductance), eq. (205) in section 8.2 would become

$$\bar{I}_T = \frac{\bar{V}}{\bar{Z}_T} = \frac{\bar{V}}{(R_1 + R_2 + \cdots + R_n) - \dfrac{j}{\omega}\left(\dfrac{1}{C_1} + \dfrac{1}{C_2} + \cdots + \dfrac{1}{C_n}\right)} \tag{222}$$

where, to summarize the notation, $\bar{V}$ and $\bar{I}$ are vector volts and amperes (being represented as complex numbers), R and C are in ohms and farads, and $\omega = 2\pi f$ is frequency in radians per second (f being frequency in cycles per second).

Problem 122

A certain series circuit consists of two resistances, both of 18 ohms, and three capacitors, each of 0.12 μF. If the applied sinusoidal reference voltage is 75 volts rms and $\omega = 500{,}000$ rad/sec, find

(a) magnitude of rms current,

(b) phase angle of current,

(c) magnitude of voltage drop across each capacitor.

Now consider the case of a purely PARALLEL connection of n such impedances (see Fig. 132). In this case the total impedance $\bar{Z}_T$ seen by the generator is given by eq. (207) or (208), each individual impedance now having the general form $\bar{Z} = R - j/\omega C = R + 1/j\omega C$. (For the special case of *two* impedances in parallel, eqs. (207) and (208) reduce to the convenient form of eq. (209).) After $\bar{Z}_T$ is found, the value of the total (generator) current $\bar{I}_T$ is then, by Ohm's law, equal to

$$\bar{I}_T = \frac{\bar{V}}{\bar{Z}_T}$$

Problem 123

In Fig. 140, the reference generator voltage is 60 volts rms, as shown. If the frequency is 100,000 rad/sec, find

(a) magnitude of generator current,

(b) phase angle of current $\bar{I}_T$.

The symbol Ω (capital "omega") denotes "ohms."

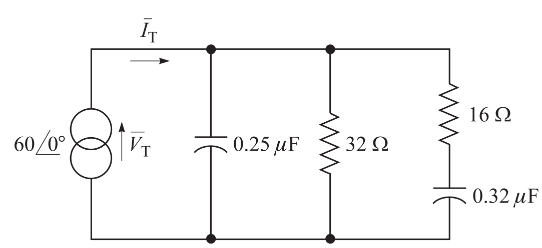


Fig. 140

8.4 The General *RLC* Network. Admittance

In practical work we often encounter networks containing not only resistance but also both inductive and capacitive reactances.

To find the "sinusoidal-steady-state" response of such a network, we simply replace the inductors and capacitors with the imaginary quantities, $jX_L = j\omega L$ and $-jX_C = -j/\omega C = 1/j\omega C$, where $\omega = 2\pi f$ radians per second. If the network is complicated we'll generally use the method of "loop currents," applying the Kirchhoff voltage and current laws and the algebra of complex numbers, as discussed in connection with Fig. 135.

Problem 124

In the series circuit of Fig. 141, the R, L, and C values are in ohms, microhenrys, and microfarads. Given that $f = 28$ kHz (kilohertz), find the vector voltage at point "a" with respect to ground.*

(*Answer* (in polar coord.), $\bar{V}_a = 22.43\underline{/143.95^\circ}$ volts)

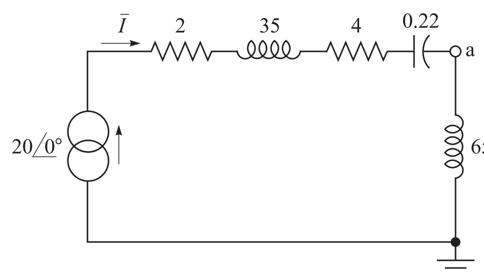


Fig. 141

Problem 125

In Fig. 142, the R, L, and C values are in ohms, microhenrys, and microfarads. As always, the generator voltage is in rms vector volts. Given that $\omega = 10^6$ radians/second, find the vector voltage at point "y" with respect to ground.

(*Answer* (in polar coord.), $\bar{V}_y = 6.19221\underline{/-26.566^\circ}$ volts)

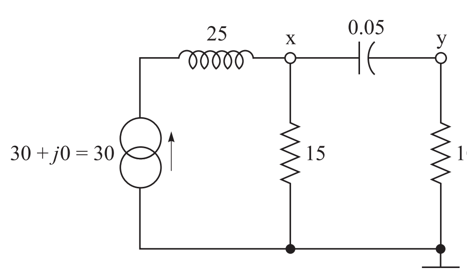


Fig. 142

Problem 126

Making use of the work already done in problem 125, find, in Fig. 142, the vector voltage at point "x" with respect to ground.

* See footnote with problem 22, Chap. 2.

Problem 127

Here we wish to apply the fundamental principle of Thevenin's theorem (section 4.6) to the circuit of Fig. 143, in which we'll assume the internal resistance of the generator is either negligibly small or is included in the resistance R. Find the voltage $\bar{V}'$ and internal impedance $\bar{Z}'$ of the equivalent Thevenin generator, as indicated in Fig. 144. (Final answers in terms of R, C, and ω.)

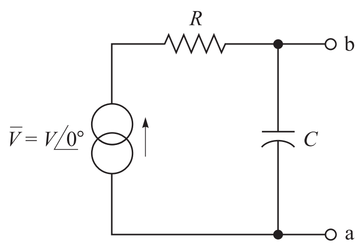


Fig. 143

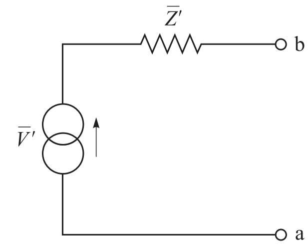


Fig. 144

Problem 128

In Fig. 145, the values are in ohms, microfarads, and microhenrys. If $\omega = 10^7$ rad/sec, find (a) vector voltage at point x with respect to ground, (b) voltmeter reading at point x with respect to ground.

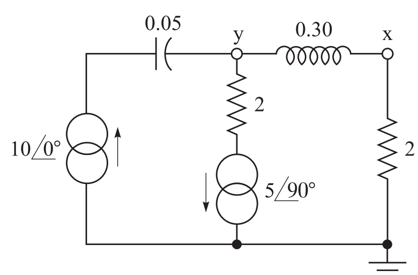


Fig. 145

Problem 129

Making use of the work already done in problem 128, find, in Fig. 145, the voltmeter reading at point "y" with respect to ground. (*Answer*: 3.5355 V, approx.)

It is sometimes advantageous to write the network equations in terms of the Kirchhoff CURRENT LAW instead of the voltage law. This involves the method of "node voltages," discussed in section 4.8, for dc circuits. For ac circuits, the current law states that the VECTOR *sum* of the rms currents flowing *to* a node (junction) point is equal to the vector sum of the currents flowing *away* from the point.

The procedure for ac circuits is basically the same as for the dc case illustrated in Fig. 61 in Chapter 4, except that for the ac case the voltages and currents are rms *vector* values of sinusoidal currents and voltages. Thus, for an ac case, Fig. 61 might be such as is illustrated in Fig. 146.

In Fig. 146,

$$\bar{I} = \frac{\bar{V}_a - \bar{V}_b}{\bar{Z}} \tag{223}$$

where $\bar{I}_1 + \bar{I}_2 = \bar{I} = \bar{I}_3 + \bar{I}_4 + \bar{I}_5$.

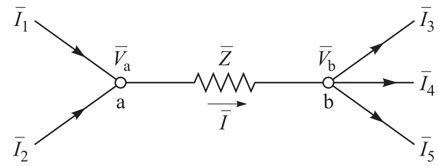


Fig. 146

Problem 130
In Fig. 147, the values of the circuit components are given in ohms, microhenrys, and microfarads, the frequency being 100,000 radians/second. Using the Kirchhoff current law with eq. (223), find, with respect to ground, the unknown voltages at nodes "x" and "y."

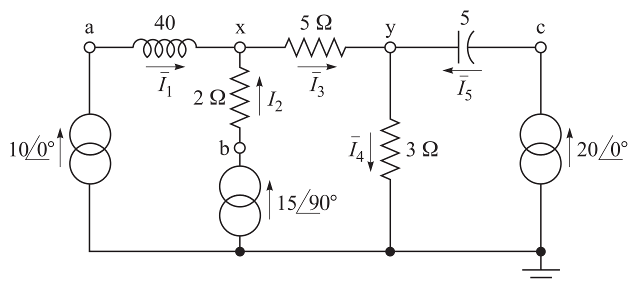


Fig. 147

Note: The "current arrows," which represent the rms vector values of the unknown currents, need not be drawn in the same directions as shown in Fig. 147. However, once selected, the directions must not be changed during the working of a given problem. The arrows insure that each current equation, written at each node point, is consistent with the current equations written at the other nodes.

Problem 131
Find the reading of an ac voltmeter connected between points x and y in Fig. 147.

In section 4.5 we found that it's sometimes an advantage, in certain types of dc network, to work with the RECIPROCAL *of resistance* instead of directly with resistance. We called the reciprocal of resistance "conductance," which we denoted by G; that is, $G = 1/R$. Conductance is thus measured in "reciprocal ohms," which we called "mhos."

In the same way, it's sometimes an advantage, in certain types of ac network, to work with the *reciprocal of* IMPEDANCE instead of directly with impedance.

The reciprocal of impedance is called "admittance," which is denoted by "$\bar{Y}$"; that is, $\bar{Y} = 1/\bar{Z}$. Since $\bar{Z}$ is, in general, a complex number, it follows that $\bar{Y}$ is also, in general, a complex number. Since impedance is measured in ohms, admittance, $\bar{Y} = 1/\bar{Z}$, is measured in reciprocal ohms or mhos.

As we found in section 4.5, it is especially convenient to work in terms of conductance when dealing with purely PARALLEL dc networks in the form of Fig. 55 in Chap. 4. In such a case, the output voltage V_0 is given by eq. (63), which is "Millman's theorem" for dc networks in the form of Fig. 55. For the steady-state *ac* case, Fig. 55 becomes Fig. 148.

In the figure, note that there is just *one* unknown node voltage, $\bar{V}_0$. Hence the VECTOR *sum* of all the currents flowing to this single point must be *zero*; that is, in Fig. 148

$$\bar{I}_1 + \bar{I}_2 + \bar{I}_3 + \cdots + \bar{I}_n = 0 \tag{224}$$

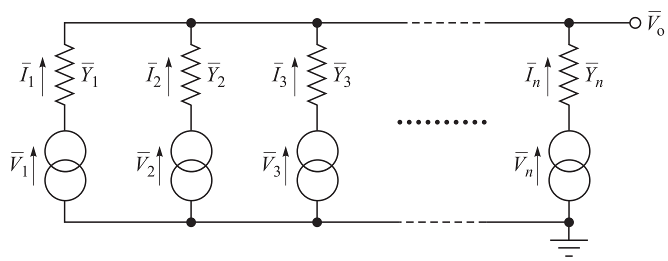
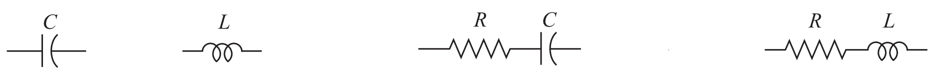


Fig. 148

Equation (224) corresponds to the *dc* case of eq. (62) for Fig. 55 in section 4.5. If we now apply to Fig. 148 the same basic procedure and steps as we applied to Fig. 55, we get the *ac equivalent of eq. (63)*, thus

$$\bar{V}_o = \frac{\bar{Y}_1\bar{V}_1 + \bar{Y}_2\bar{V}_2 + \cdots + \bar{Y}_n\bar{V}_n}{\bar{Y}_1 + \bar{Y}_2 + \cdots + \bar{Y}_n} \tag{225}$$

which is Millman's theorem for the ac case of Fig. 148, where $\bar{Y} = 1/\bar{Z}$. As examples,



$\bar{Z} = -jX_C$ $\quad$ $\bar{Z} = jX_L$ $\quad$ $\bar{Z} = R - jX_C$ $\quad$ $\bar{Z} = R + jX_L$

$\bar{Y} = \dfrac{1}{-jX_C}$ $\quad$ $\bar{Y} = \dfrac{1}{jX_L}$ $\quad$ $\bar{Y} = \dfrac{1}{R - jX_C} = \dfrac{R + jX_C}{R^2 + X_C^2}$ $\quad$ $\bar{Y} = \dfrac{1}{R + jX_L} = \dfrac{R - jX_L}{R^2 + X_L^2}$

$\bar{Y} = j/X_C = j\omega C$ $\quad$ $\bar{Y} = -j/X_L = -j/\omega L$

$(X_L = \omega L, X_C = 1/\omega C)$

Problem 132

In Fig. 149, the values of the network components are given in ohms, microhenrys, and microfarads. The frequency is 100,000 rad/sec. Find the ac voltmeter reading between point "a" and ground. (We've numbered the branches from left to right, 1 through 5, as shown.) Use Millman's theorem.

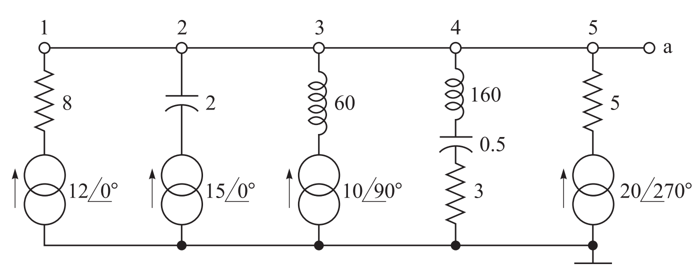


Fig. 149

8.5 Real and Apparent Power. Power Factor

Electrical POWER is the RATE at which energy is being expended in a circuit, and is measured in WATTS.

In section 5.5 we found that the AVERAGE POWER, P, produced in a purely RESISTIVE ac circuit is equal to

$$P = VI \text{ watts} \tag{226}$$

where V and I are the *rms values* of voltage and current.

It must be understood, however, that eq. (226) is correct only if the load is a pure RESISTANCE of R ohms. This is because in a purely resistive circuit the current and voltage waves are completely IN PHASE with each other, so that the current never reverses direction before the voltage does, and vice versa. If the current wave is *not* completely in phase with the voltage wave, then the average power P is *less* than the value given by eq. (226) and is then given by eq. (117) in Chap. 5, which is repeated below as eq. (227) (using ϕ instead of θ); thus

$$P = VI\cos\phi \text{ watts} \qquad (227)^*$$

in which ϕ is the PHASE ANGLE between the current and voltage waves, where V and I are the magnitudes of the rms values. The situation is represented in vector form in Fig. 150.

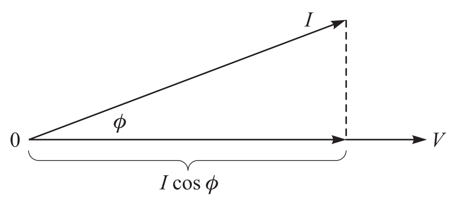


Fig. 150

In the figure, note that "$I\cos\phi$" is the component of $\bar{I}$ that is IN PHASE WITH THE VOLTAGE $\bar{V}$, and thus, in accordance with eq. (227), we see that TRUE POWER P, is equal to the voltage V times the component of current that is IN PHASE with $\bar{V}$. Thus we have the important fact that, in alternating-current work.

> TRUE POWER is equal to the RMS VOLTAGE times the rms component of current that is IN PHASE with the voltage.

Consider first the case of an ideal inductor or capacitor.† In either case, the voltage drop across, and the current through, are 90° out of phase, and thus the true power expended in an ideal reactor would be *zero*, since by eq. (227)

$$P = VI\cos 90^\circ = 0$$

* "$\cos\phi$" is called the "power factor."

† An "ideal" inductor or capacitor would have *zero resistance* and thus zero losses. If an actual coil or capacitor does have appreciable resistance, such resistance can be taken into account by assuming an equal resistance to be in series with an ideal coil or capacitor.

This is because inductors and capacitors store energy in their magnetic and electric fields during one-half of the cycle, then return the stored energy to the circuit during the other half of the cycle. Thus an ideal inductor or capacitor would be a totally lossless (wattless) device (see section 7.1). The difference in the instantaneous power relationships in a pure *resistance* and in a pure *reactance*, such as a capacitor, is illustrated graphically in Figs. 151 and 152 and the following discussion.

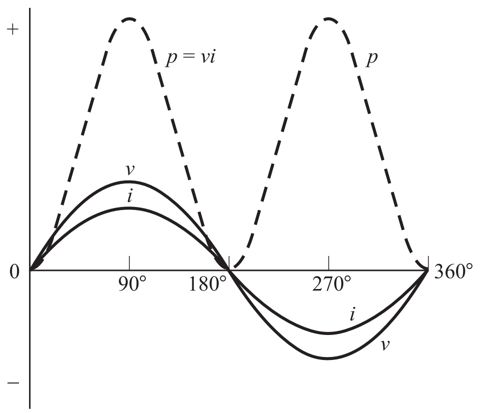


Fig. 151. *i* and *v* IN PHASE.

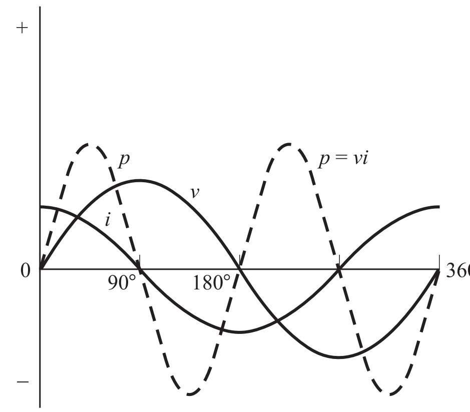


Fig. 152. *i* and *v* 90° OUT OF PHASE.

In the figures $v =$ instantaneous voltage, $i =$ instantaneous current, thus $p = vi =$ instantaneous power.

FIRST consider Fig. 151. Here v is the voltage applied to a pure *resistance* of R ohms. From the figure, note that the direction of the current is always the *same* as the direction of the voltage. Thus, in Fig. 151 we have that

from 0° to 180°, $p = (+v)(+i) = +p$, that is, "positive" power,
from 180° to 360°, $p = (-v)(-i) = +p$, that is, "positive" power.

Thus, in this case the generator *at all times* delivers power to the resistance R. By eq. (227), the power is equal to $P = VI \cos 0° = VI$ watts.

Next consider Fig. 152. Here v is a sinusoidal generator voltage applied to a pure capacitance. Note that the direction of the current is *not* always the same as the direction of the voltage; thus, in Fig. 152 we have that

from 0° to 90°, $p = (+v)(+i) = +p$, that is, "positive" power,
from 90° to 180°, $p = (+v)(-i) = -p$, that is, "negative" power,
from 180° to 270°, $p = (-v)(-i) = +p$, that is, "positive" power,
from 270° to 360°, $p = (-v)(+i) = -p$, that is, "negative" power.

Thus, in this theoretically ideal case (zero resistance), the net power output of the generator is *zero*; half the time the generator is storing energy in the capacitor, while during the other half the capacitor discharges, tending to run the generator as a motor.

Now consider a non-ideal case, having both resistance and reactance, such as the series *RC* circuit shown in Fig. 153 where, let's assume, the current leads the voltage by, say, 45°, as shown in Fig. 154.

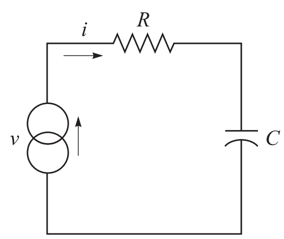


Fig. 153

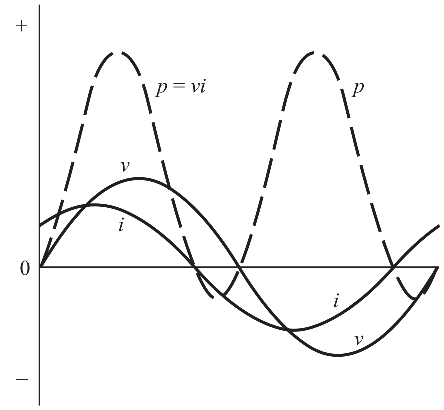


Fig. 154

Since we're dealing with alternating current, the generator voltage arrow in Fig. 153 alternately points ''up'' and ''down'' as the generator alternates in polarity, the arrow always pointing from the negative terminal of the generator to its positive terminal. ''Positive'' power is produced by the generator when the current flows through the generator in the direction of the generator voltage arrow. ''Negative'' power is produced when current flows through the generator ''against'' the generator voltage arrow, thus momentarily running the generator as a motor. This is similar to working with a storage battery; the battery *produces* power when current flows *out* of the positive terminal of the battery, but the battery *absorbs* energy (becomes the load) when current flows *into* the positive terminal (while the battery is being recharged).

For the situation in Fig. 154, note that most of the time the current flows through the generator in the same sense as the generator voltage, so that in this case ($\phi = 45^\circ$) positive power exceeds negative power. Thus, in this particular case, the *true* or ''positive'' power output of the generator is, by eq. (227), equal to

$$P = VI \cos 45^\circ = 0.7071\ VI$$

The above discussion of the capacitive circuit of Fig. 153 also applies, of course, to inductive-type loads. In the inductive case energy is stored in the magnetic field of the inductor coil, the magnetic field of the coil alternately being ''charged'' and ''discharged''; the inductive action causes the current to ''lag'' the applied voltage instead of ''leading'' as in the capacitive case illustrated in Fig. 154.

Another point to mention is as follows. Inspection of Figs. 151, 152, and 154 shows that instantaneous POWER in an ac circuit always pulsates sinusoidally at a frequency equal to *twice* the frequency of the applied voltage. This is an important factor that must be taken into account in certain practical design problems.

Another point of importance concerns the quantity cos ϕ, which we've called the circuit ''power factor'' (see eq. (227)). The power factor, cos ϕ, can also be expressed in terms of impedance and power, as follows.

In direct-current (dc) work, *power*, P, is given by the simple relationship $P = VI$. In the *ac* case, however, average power is given by eq. (227), which let us rewrite as eq. (228), thus

$$P = VI \cos \phi \tag{228}$$

where ϕ is the phase angle between the current and voltage waves, and V and I are magnitudes of rms values of voltage and current. Thus, in the ac case, the product VI may or may not be equal to the ''true power'' P. For this reason it's appropriate to call the product VI the "apparent power," because it must be multiplied by the "power factor," cos ϕ, to get the "true power" P. This can be expressed by writing eq. (228) in the form

$$(\text{true power}) = (\text{apparent power})(\cos\ \phi)$$

that is,

$$P_t = P_a\ \cos\ \phi = VI \cos\phi \tag{229}$$

where P_t = true power, and P_a = apparent power = VI.

Now recall that in an ac circuit ENERGY is actually expended ONLY *in the* RESISTIVE *component* of the circuit impedance. Thus TRUE POWER, in an ac circuit, is always equal to the "square of the rms current, times the resistance R"; that is

$$P_t = I^2 R$$

hence eq. (229) becomes

$$I^2 R = VI \cos\phi \tag{230}$$

It is also true that $V = IZ$ (eq. (200)), and thus, substituting this value of V in eq. (230), we have the important fact that

$$\cos\phi = R/Z \tag{231}$$

where R is the circuit resistance and Z is the magnitude of the circuit impedance (as can also be seen from inspection of Figs. 130 and 139). Hence another expression for the "power factor" is R/Z, and thus eq. (229) is extended to the form

$$P_t = VI \cos\phi = VI(R/Z) \tag{232}$$

To continue the discussion, let us first redraw Fig. 130 as Fig. 155, where $\omega L = X$ and $|\bar{Z}| = Z$.

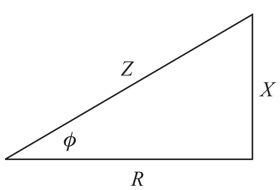


Fig. 155

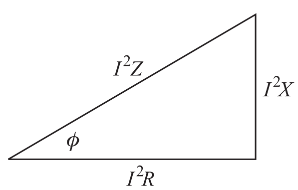


Fig. 156

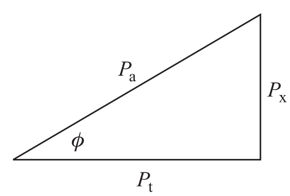


Fig. 157

Next, multiply all three sides of the triangle in Fig. 155 by the *square* of the magnitude of the rms current, I^2. Doing this preserves the angle ϕ, and Fig. 155 becomes Fig. 156 where, from our work above, $I^2R = P_t$ and (since $V = IZ$) $VI = I^2Z = P_a$, as shown in Fig. 157.

The meaning of P_x is as follows. Since reactance, X, is measured in *ohms*, it follows that I^2X is measured in *watts*; that is, I^2X represents what is called "reactive" power. Thus P_x represents power that does *not* represent energy that is transformed into heat, light, or mechanical energy, but only energy that is momentarily stored in electric and/or magnetic fields and then returned to the source. From inspection of Fig. 157, note that

$$P_a^2 = P_t^2 + P_x^2 \tag{233}$$

To summarize, the "true power" P_t in an ac circuit is given by eq. (228), in which cos ϕ is the "power factor," where

$$\cos\phi = R/Z = P_t/P_a \tag{234}$$

It should be noted that the above conclusions are valid for any configuration of series, parallel, or series-parallel network. Thus, if a generator of $\bar{V}$ volts delivers a current of $\bar{I}$ amperes into a certain network, and if the generator sees an impedance $\bar{Z} = R + jX$ ohms looking into the network, then the true power produced by the generator is equal to $P_t = VI \cos\phi$, where $\cos\ \phi = R/Z$.

Problem 133
In problem 116, find (a) the apparent power, (b) the true power, (c) the reactive power.

Problem 134
In problem 119 (Fig. 134), find the true power produced by the generator if $V = 28$ volts, $R = 12$ ohms, and $\omega L = 16$ ohms. (*Answer*: 65.333 W, approx.)

Problem 135
In Fig. 158, the circuit values are in ohms and henrys, the generator voltage being 32 V rms, as shown. Given that $\omega = 1000$ rad/sec, find the true power produced by the generator. (*Answer*: 226.98 W, approx.)

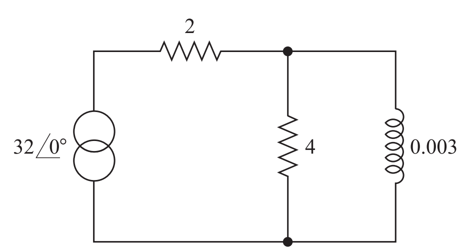


Fig. 158

Problem 136
Rework problem 135, this time using the method of "loop currents." Remembering that the power delivered to a pure resistance of R ohms is always equal to the "square of the magnitude of the rms current, times R" ($P = I^2R$), find

(a) power delivered to the 2-ohm resistance,
(b) power delivered to the 4-ohm resistance,
(c) check to verify that the sum of the answers in (a) and (b) is the same as the answer found in problem 135.

Problem 137
Find (a) the apparent power, (b) the true power, produced by the generator in problem 124 (Fig. 141).

Problem 138
Find the true power produced by the generator in Fig. 142, making use of the value of $\bar{I}_1$ found in problem 126.

Problem 139
In Fig. 142, find (a) power to the 15-ohm resistance, (b) power to the 10-ohm resistance. Verify that the sum of (a) and (b) is the same as the answer found in problem 138.

8.6 Series Resonance

The subject of this section is the basic *series RLC circuit* of Fig. 159. A brief review of the meaning of the symbols in Fig. 159 follows.

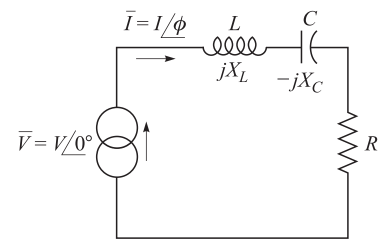


Fig. 159

First, V and I denote *rms* values of sinusoidal voltage and current of the same frequency, with ϕ being the phase angle of the current wave with respect to the reference voltage wave.

However (from section 5.6), rms values of sinusoidal waves of the same frequency can be manipulated as if they were *vector* quantities, and thus we write $\bar{V}$ and $\bar{I}$. Then, since *complex numbers* can *also* be manipulated as if they were vector quantities, we can write and manipulate sinusoidal voltages and currents in the forms $\bar{V} = V' + jV''$ and $\bar{I} = I' + jI''$, where V' and I' are the real components and V'' and I'' are the imaginary components of $\bar{V}$ and $\bar{I}$.

Next, in linear circuits* it is a fact that if the applied *voltages* are sinusoidal, then the *currents* are *also* sinusoidal, and the *voltage drops* across the passive R, L, and C components are also sinusoidal.† In regard to the voltage drops, remember that the voltage drop across *resistance* is "in phase" with the current, and, in the steady-state, the voltage drop across an *inductor* 'leads" the current by 90° and the voltage drop across a capacitor "lags" the current by 90°. Thus we have that

$$\begin{aligned} R\bar{I} &= \text{vector voltage drop across a resistance of } R \text{ ohms,} \\ j\omega L\bar{I} = jX_L\bar{I} &= \text{vector voltage drop across an inductor of } L \text{ henrys,} \\ -j(1/\omega C)\bar{I} = -jX_C\bar{I} &= \text{vector voltage drop across a capacitor of } C \text{ farads,} \end{aligned}$$

where the $\pm j$ factors shift the voltage vectors $X_L\bar{I}$ and $X_C\bar{I}$ the required $\pm 90°$ with respect to $\bar{I}$.

In Fig. 159 the circuit impedance is $\bar{Z} = R + j(X_L - X_C)$, and thus, by Ohm's law, the current $\bar{I}$ is equal to

$$\bar{I} = \frac{\bar{V}}{\bar{Z}} = \frac{\bar{V}}{R + j(X_L - X_C)} \tag{235}$$

* A circuit is "linear" if the *values* of R, L, and C are not dependent upon the *amount* of voltage or the *amount* of current.

† It should be noted that in a linear circuit containing L or C, to which a repetitive voltage is applied, the *current* will have the SAME WAVESHAPE as the applied voltage ONLY if the applied voltage is sinusoidal; this is a unique and valuable property of the sinusoidal function.

where $X_L = \omega L$ is the "inductive reactance" and $X_C = 1/\omega C$ is the "capacitive reactance," and where, since $\bar{V}$ is taken as the reference vector, we can write that

$$\bar{V} = V\underline{/0^\circ} = V$$

In magnitude,

$$|\bar{I}| = \frac{V}{\sqrt{R^2 + (X_L - X_C)^2}} \tag{236}$$

Now, in Fig. 159, let the rms value of the applied reference voltage be held *constant* at V volts. Also, for the time being, let us hold the values of L and C *constant* (at any particular values we might be interested in). Let us, however, be able to CHANGE THE FREQUENCY, ω, of the applied voltage V at will; in other words, let the FREQUENCY be the VARIABLE, with everything else, for the time being, *held constant*.

With this in mind, let's pause a moment to manipulate eq. (235) a bit, as follows ($\bar{V} = V$)

$$\bar{I} = \frac{V}{R + j(X_L - X_C)} = \frac{V[R - j(X_L - X_C)]}{R^2 + (X_L - X_C)^2} = \frac{V}{D}\,[R + j(X_C - X_L)]$$

where $D = R^2 + (X_L - X_C)^2$ (thus D is always a positive number). Hence the PHASE ANGLE ϕ of the current vector $\bar{I}$ with respect to the reference voltage V is

$$\phi = \arctan\frac{(X_C - X_L)}{R} \tag{237}$$

Thus, if Fig. 159 is an *inductive* circuit ($X_L > X_C$), then $(X_C - X_L)$ has a *negative* value and thus ϕ is a *negative* angle, showing that the current "lags" the voltage in an inductive circuit. Or, if in Fig. 159 the circuit is *capacitive* ($X_C > X_L$), then $(X_C - X_L)$ has a *positive* value and thus ϕ is a *positive* angle, showing that the current "leads" the voltage in a capacitive circuit.

Now consider the condition called SERIES RESONANCE, which is the condition where $X_L = X_C$, that is, where $(X_L - X_C) = 0$. We immediately see (by direct inspection of eqs. (235) and (236)) that at "series resonance" the generator in Fig. 159 sees a *pure resistance* of R ohms, the current at resonance thus being *in phase with* V, having the *maximum value* of $I = V/R$ amp.

The FREQUENCY at which $X_L = X_C$ is called the "resonant frequency" and is denoted by ω_0 if we're dealing in "radians/second," or f_0 if we're using "cycles/second" (hertz). ω_0 and f_0 can be read as "omega sub zero" and "f sub zero."

Thus the condition for "series resonance" is that $\omega_0 L = 1/\omega_0 C$, from which we have that

$$\omega_0 = \frac{1}{\sqrt{LC}}\ \text{radians/sec} \tag{238}$$

or

$$f_0 = \frac{1}{2\pi\sqrt{LC}}\ \text{cycles/sec (Hz)} \tag{239}$$

where L is inductance in henrys and C is capacitance in farads.

As we found above, the *current* in a series circuit has its MAXIMUM VALUE of $I = V/R$ amperes at resonance. The key to the behavior of the series circuit is contained in the *denominator* of eqs. (235) and (236). We see that the denominator has its *least value* at resonance, because ONLY at resonance is it true that $(X_L - X_C) = 0$. Note that the greater the *deviation* of the frequency away from the resonant frequency, the greater is the magnitude of $(X_L - X_C)$, thus the greater is the magnitude of the denominator and the *less* is the magnitude of current. The result is illustrated graphically in Fig. 160, the general form of a plot of eq. (236), where $X_L = \omega L$ and $X_C = 1/\omega C$.

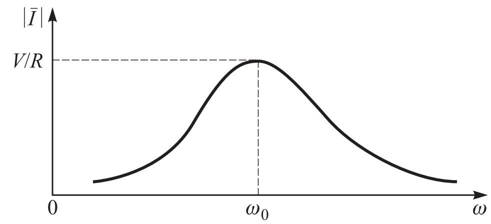
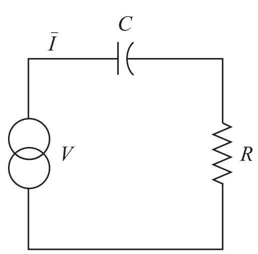


Fig. 160

Let us now summarize some facts about the basic series circuit of Fig. 159. To begin, let X denote the "net reactance" in Fig. 159 where, from inspection of eqs. (235) and (236), we see that

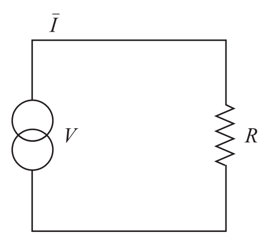
$$X = (X_L - X_C) = \left(\omega L - \frac{1}{\omega C}\right)$$

showing that inductive and capacitive reactances *tend to cancel each other out* in a series circuit. This is because their voltage drops are *180 degrees out of phase with each other*, being equal to $+jX_L\bar{I}$ and $-jX_C\bar{I}$. Thus, in Fig. 159, if ωL is *less than* $1/\omega C$ (below resonance), the generator sees a *capacitive circuit*, but if ωL is *greater than* $1/\omega C$ (above resonance), the generator sees an *inductive circuit*. Of course, if $X_L = X_C$ (the condition of resonance, $\omega = \omega_0$), the generator sees a pure resistance. These three possible conditions are illustrated in Figs. 161, 162, and 163.



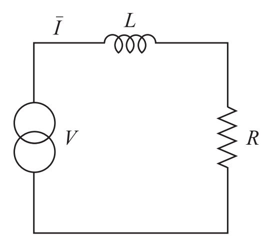
BELOW resonance:
$\omega < \omega_0 \quad X_C > X_L$
Current leads voltage

Fig. 161



Resonance:
$\omega = \omega_0 \quad X_L = X_C$
Current and voltage in phase

Fig. 162



ABOVE resonance:
$\omega > \omega_0 \quad X_L > X_C$
Current lags voltage

Fig. 163

A vector diagram for Fig. 159 for the condition of *resonance*, $\omega = \omega_0$, is given in Fig. 164, where V is generator voltage and I_0 is current at resonance $(I_0 = V/R)$.

Note that the voltage drops across L and C are equal in magnitude but 180° out of phase with each other. Note, also, that the magnitudes of the voltage drops across L and C can be MANY TIMES GREATER than the generator voltage V. This is possible because *at resonance* V_L and V_C exactly cancel each other out, leaving only the voltage drop RI_0 in the circuit $(V = RI_0)$.

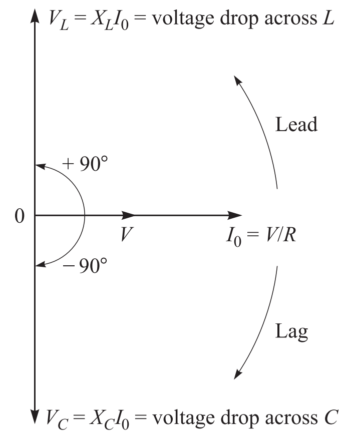


Fig. 164

Now suppose the frequency were to become (for example) GREATER than the resonant frequency ($\omega > \omega_0$). In this case the current $\bar{I}$ would LAG the generator voltage by some angle ϕ (eq. (237)), as shown in Fig. 165. (For convenience, Figs. 164 and 165 are not drawn to the same scale.) A brief discussion follows.

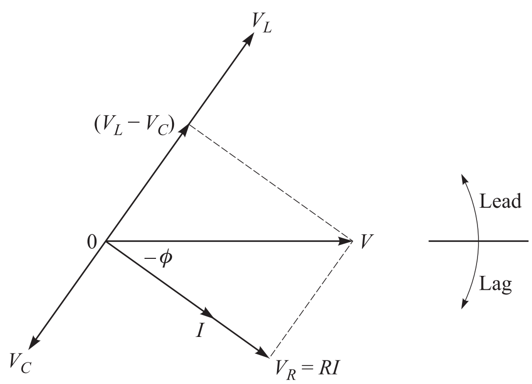


Fig. 165

Since we are no longer at resonance (here we're assuming a frequency above resonance), the magnitude of $\bar{I}$ would now be *less* than I_0 (by eq. (236), and seen in Fig. 160). The voltage drops V_L and V_C will still be at right angles to the current vector $\bar{I}$ and will still be 180° out of phase with each other, but now V_L will be *greater* than V_C, and thus V_L and V_C will *no longer completely cancel each other out*, but, instead, a *net voltage drop* of $(V_L - V_C)$ will appear between L and C, as shown in Fig. 165. The *vector sum* of the voltage $(V_L - V_C)$ and the voltage drop *RI across the resistance R* must and will be equal to the *generator voltage V* as shown in the figure.

For frequencies *below resonance* X_C will be greater than X_L, and the current $\bar{I}$ will *lead* the generator voltage V. The ''net voltage drop'' between C and L will be $(V_C - V_L)$, and the vector sum of this voltage and the voltage drop RI across the resistance R must again be equal to the generator voltage V.

Problem 140
Find the resonant frequency of a series circuit in which $L = 4$ microhenrys, $C = 0.0025$ microfarads, and $R = 0.65$ ohm.

Problem 141
In a certain series RLC circuit, $L = 400$ microhenrys. Find the value of C if the circuit must resonate at 500 kilohertz (500 kHz). (*Answer*: 253.3 pF (picofarads))

Problem 142
If, in Fig. 159, $L = 1$ microhenry, $C = 0.0025$ microfarad, $R = 5$ ohms, and if the generator voltage is 20 volts rms, find the following values:

(a) power output of generator at the resonant frequency,
(b) voltage drop across C at resonance.

If, now, the generator frequency is made equal to 10^7 rad/sec (all else unchanged), find

(c) magnitude of voltage drop across C,
(d) phase angle ϕ of current vector with respect to generator voltage,
(e) power output of generator.

Practically speaking, the phenomenon of "series resonance" is especially important because it can be used to select or "tune in" a desired signal, while rejecting all others. To investigate this most interesting and useful matter, let us begin with Fig. 166, in which a generator of reference voltage $\bar{V} = V\underline{/0^\circ} = V$ is applied to a series RLC circuit, as shown.

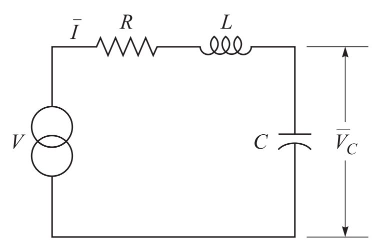


Fig. 166

The value of the current $\bar{I}$ is (from eq. (235)) equal to

$$\bar{I} = \frac{V}{R\left[1 + \dfrac{j(X_L - X_C)}{R}\right]} = \frac{V}{R}\,\frac{1}{1 + \dfrac{j}{R}\left(\omega L - \dfrac{1}{\omega C}\right)} \tag{240}$$

It is important, now, that we somehow get the *resonant frequency* ω_0 into eq. (240). This is an interesting exercise in algebraic manipulation, and can be done by first writing eq. (240) in the form

$$\bar{I} = \frac{V}{R}\,\frac{1}{1 + \dfrac{j\omega_0 L}{R\omega_0 L}\left(\omega L - \dfrac{1}{\omega C}\right)} \tag{241}$$

Equation (241) is, of course, the same as eq. (240), because $\omega_0 L/\omega_0 L = 1$. The next step is not so obvious, but after some study we realize that eq. (241) can also be written in the form

$$\bar{I} = \frac{V}{R} \frac{1}{1 + J\left(\dfrac{\omega_0 L}{R}\right)\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)} \tag{242}$$

which is true because multiplying by $1/\omega_0 L$ is the same as multiplying by $\omega_0 C$ (because $1/\omega_0 L = \omega_0 C$). Next, it is universal practice to represent the *ratio* of the reactance of the coil *at resonance* to the circuit resistance by "Q"; that is

$$\frac{\omega_0 L}{R} = Q \tag{243}$$

thus eq. (242) becomes

$$\bar{I} = \frac{V}{R} \frac{1}{1 + jQ\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)} \tag{244}$$

Now, in the following discussion, let it be given that, while the *amplitude* of the generator voltage will always remain *constant* at V volts, its *frequency* ω can be set to any value we might be interested in. Also, let the "output voltage" of the system be the voltage drop across the capacitor, $\bar{V}_C$, as shown in Fig. 166. Then, since

$$\text{voltage drop across } C = (\text{current})(\text{reactance of } C) = \bar{I}(-j/\omega C)$$

we have, using the value of $\bar{I}$ from eq. (244), that for Fig. 166 eq. (244) becomes

$$\frac{\bar{V}_C}{V} = \frac{\dfrac{-j}{R\omega C}}{1 + jQ\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)} \tag{245}$$

Let us now work on the *numerator* in the above, as follows:

$$-j\left(\frac{1}{R\omega C}\right) = -j\left(\frac{\omega_0}{\omega R\omega_0 C}\right) = -j\left(\frac{\omega_0}{\omega}\right)\left(\frac{1}{R\omega_0 C}\right) = -j\left(\frac{\omega_0}{\omega}\right)Q$$

in which we made use of the fact that $\dfrac{1}{\omega_0 C} = \omega_0 L$, then applied the definition of eq. (243); thus eq. (245) becomes

$$\frac{\bar{V}_C}{V} = \frac{-jQ\left(\dfrac{\omega_0}{\omega}\right)}{1 + jQ\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)} \tag{246}$$

Equation (246) is said to be in "dimensionless" form, because it requires only the *ratios* of like quantities and is thus valid for all systems of measurement. The equation is also said to be in "normalized" form, because it gives the value of V_C relative to the reference voltage V.

As previously mentioned, series resonance is of great practical importance because it can be used to select or "tune in" a signal of any desired frequency while rejecting all others. Thus we are especially interested in the behavior of eq. (246) IN THE IMMEDIATE VICINITY

OF THE RESONANT FREQUENCY ω_0. To aid in the study of eq. (246) in the close vicinity of ω_0, let us define that

$$d = \frac{\omega}{\omega_0} \tag{247}$$

that is, "d" is the ratio of ANY FREQUENCY ω, to the RESONANT FREQUENCY ω_0. Then eq. (246) becomes

$$\frac{\bar{V}_C}{V} = \frac{-j\dfrac{Q}{d}}{1 + jQ\left(d - \dfrac{1}{d}\right)} = \frac{-jQ}{d + jQ(d^2 - 1)}$$

In magnitude,

$$\left|\frac{\bar{V}_C}{V}\right| = \frac{Q}{\sqrt{d^2 + Q^2(d^2 - 1)^2}} \quad (d = \omega/\omega_0) \tag{248}$$

Problem 143
Letting $A = |\bar{V}_C/V|$, fill in the following table of values for Fig. 166. (Round final calculator values off to two decimal places.)

d	A, for $Q = 10$	A, for $Q = 20$	d	A, for $Q = 10$	A, for $Q = 20$
0.80			1.01		
0.85			1.03		
0.90			1.05		
0.93			1.07		
0.95			1.10		
0.97			1.15		
0.99			1.20		
1.00					

Use the results to sketch approximate curves of A versus d, for $Q = 10$ and $Q = 20$.

8.7 Parallel Resonance

In this section we propose to analyze the parallel *RLC* network of Fig. 167, where $\bar{Z}_p$ is the "input impedance" to the network, that is, the impedance a generator would see if connected to terminals a, b.

The circuit of Fig. 167 is important because it finds wide use as a "tuned load" at higher frequencies, in both receiver and transmitter work.

Note that *resistance* is assumed to exist only in the inductive branch of the circuit. This is because inductors (coils), being wound of wire, inherently have much more power loss* than capacitors (it being easy to obtain practically lossless capacitors).

* The resistance R in Fig. 167 is the sum of the actual resistance of the coil itself and any resistance "coupled" into L due to transformer action (when L is the primary of a transformer).

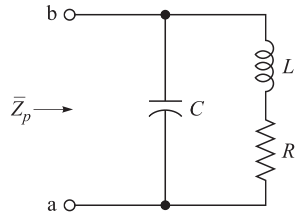


Fig. 167

Now, applying the standard procedure for two impedances in parallel (product of the two, over the sum), we have that, in Fig. 167,

$$\bar{Z}_\mathrm{p} = \frac{-jX_C(R + jX_L)}{R + j(X_L - X_C)} \tag{249}$$

Now "rationalize" the above fraction; that is, multiply the numerator and denominator by $R - j(X_L - X_C)$. Doing this, then setting $X_L = \omega L$ and $X_C = 1/\omega C$, you can verify that eq. (249) becomes

$$\bar{Z}_\mathrm{p} = \frac{1}{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}\left[\frac{R}{\omega^2 C^2} + j\left(-\frac{\omega^2 L^2}{\omega C} + \frac{\omega L}{\omega^2 C^2} - \frac{R^2}{\omega C}\right)\right] \tag{250}$$

We now obviously have a considerably more complicated condition than we had for the series case of section 8.6. Let us therefore try to put eq. (250) in a somewhat better form, especially in regard to the quantity $(\omega L - 1/\omega C)$. In an effort to do this, let us first write the *imaginary* component, inside the brackets, in the equivalent form

$$j\left(-\frac{\omega^2 L^2 \omega C}{\omega^2 C^2} + \frac{\omega LC}{\omega^2 C^2 C} - \frac{R^2 \omega C}{\omega^2 C^2}\right) = j\,\frac{\omega C}{\omega^2 C^2}\left(-\omega^2 L^2 + \frac{L}{C} - R^2\right)$$

Thus, $1/\omega^2 C^2$ now factors out of the entire quantity inside the brackets in eq. (250); doing this, and remembering the algebraic fact that, $A^2 B^2 = (AB)^2$, you can verify that eq. (250) becomes

$$\bar{Z}_\mathrm{p} = \frac{R}{R^2\omega^2 C^2 + (\omega^2 LC - 1)^2} + j\,\frac{\omega C\left(-\omega^2 L^2 + \dfrac{L}{C} - R^2\right)}{R^2\omega^2 C^2 + (\omega^2 LC - 1)^2} \tag{251}$$

To continue, let us now DEFINE that the "resonant frequency," for the parallel case of Fig. 167, is the frequency at which $\bar{Z}_\mathrm{p}$ becomes a PURE RESISTANCE. Let us denote the resonant frequency by "ω_0"; then, by definition, ω_0 is the frequency at which the imaginary or "reactive" component of eq. (251) becomes equal to *zero*. Note that this requirement will be satisfied if the *numerator* of the imaginary part of eq. (251) is equal to *zero*; that is, if

$$\omega C = 0$$

or if

$$\left(-\omega^2 L^2 + \frac{L}{C} - R^2\right) = 0$$

The first possibility, $\omega_0 C = 0$, is true only if $\omega = 0$ or if $C = 0$, and hence is of no practical importance. However, setting $\omega = \omega_0$ in the second possibility gives, as you can verify, the meaningful, correct answer

$$\omega_0 = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}} \tag{252}$$

in which only the positive value of the square root is to be taken (because "negative frequency" does not exist in the real world).

Now let ω' ("omega prime") be the frequency at which the *reactances* of the coil and capacitor are *equal*; that is, let ω' be the frequency at which $\omega L = 1/\omega C$. Thus, $\omega' L = 1/\omega' C$; hence

$$\omega'^2 = 1/LC$$

and, putting this value in place of $1/LC$, eq. (252) becomes

$$\omega_0 = \sqrt{\omega'^2 - \frac{R^2}{L^2}} \tag{253}$$

Thus, as eq. (253) shows, in the parallel circuit of Fig. 167 the frequency at which $\bar{Z}_p$ is a *pure resistance* is *not the same* as the frequency at which $X_L = X_C$.* Thus $\omega_0 = \omega'$ only for the theoretically ideal case of $R = 0$.

We have defined that the "resonant frequency" for the parallel case is the frequency, ω_0, for which $\bar{Z}_p$ in Fig. 167 becomes a *pure resistance*, which let us now denote by "R_0." Hence, if we set $\omega = \omega_0$ in eq. (251), the imaginary part vanishes and $\bar{Z}_p$ becomes equal to R_0; thus

$$R_0 = \frac{R}{R^2\omega_0^2 C^2 + (\omega_0^2 LC - 1)^2}$$

Now, in the above equation, replace ω_0 with the right-hand side of eq. (252); doing this, you can verify the important fact that

$$R_0 = \frac{L}{RC} \text{ ohms} \tag{254}$$

where R_0 is the pure resistance a generator sees looking into terminals a, b in Fig. 167, if the generator is operating at the frequency defined by eq. (252). From inspection of eq. (254), note that the *smaller* the value of R, the *larger* is the value of R_0; this is one reason why Fig. 167 is especially useful in certain practical applications (as will be commented on in the solution to the following problem).

Problem 144

In Fig. 167, suppose $L = 100$ microhenrys, $C = 100$ picofarads (100 pF), and $R = 50$ ohms. What value of $\bar{Z}_p$ would a generator, operating at a frequency given by eq. (252), see if connected to terminals a, b?

To continue our study of Fig. 167, it will be instructive to deal with the dimensionless *ratio* of the general value of $\bar{Z}_p$ (given by eq. (251)) to its value at resonance, R_0 (given by

* This is different from the *series* case of Fig. 159, in which the generator sees a pure resistance at the same frequency at which $X_L = X_C$. (See discussion just prior to eq. (238) in section 8.6.)

eq. (254)). Thus, upon multiplying both sides of eq. (251) by $1/R_0$, we have that

$$\frac{\bar{Z}_p}{R_0} = \frac{\dfrac{R}{R_0} - j\,\dfrac{\omega C}{R_0}\left(\omega^2 L^2 - \dfrac{L}{C} + R^2\right)}{R^2\omega^2 C^2 + (\omega^2 LC - 1)^2} \tag{255}$$

Actually, the above equation can be expressed in much better form if certain practical approximations are made. To show how this can be done, let us first denote the ratio of the *coil reactance at resonance* to its *resistance* by "Q"; thus

$$\frac{\omega_0 L}{R} = Q \tag{256}$$

Next *square both sides* of eq. (252), thus getting

$$\omega_0^2 + \frac{R^2}{L^2} = \frac{1}{LC} \tag{257}$$

But note that

$$\frac{R^2}{L^2} = \frac{\omega_0^2 R^2}{\omega_0^2 L^2} = \omega_0^2\left(\frac{R}{\omega_0 L}\right)^2 = \omega_0^2/Q^2$$

and upon making this substitution into eq. (257) we have that

$$\omega_0^2\left(1 + \frac{1}{Q^2}\right) = \frac{1}{LC} \tag{258}$$

At this point it should be noted that all of the foregoing equations concerning Fig. 167, are *exact* equations; that is, no simplifying assumptions have been made. Let us now, however, take into account the fact that *in most practical applications* of Fig. 167 the value of Q, as defined by eq. (256), will be EQUAL TO OR GREATER THAN 10; that is, in most practical work it will be true that $Q \geq 10$.

Let us therefore base the rest of our discussion of Fig. 167 on the assumption that Q will be equal to or greater than 10. Thus, for practical purposes we can write that

$$1 + \frac{1}{Q^2} = 1$$

and hence, for practical purposes, eq. (258) becomes

$$\omega_0^2 = \frac{1}{LC} \tag{259}$$

thus

$$\omega_0 L = 1/\omega_0 C \tag{260}$$

showing that, for practical purposes, in Fig. 167 it can be taken that $X_L = X_C$ at the same frequency, ω_0, that $\bar{Z}_p = R_0$.

Our goal, now, is to express eq. (255) in dimensionless form (similar to eqs. (246) and (248) in section 8.6), making use of the assumption that $(1 + 1/Q^2) = 1$. To do this requires a certain amount of trial and error; let us suppose, after a few trials, we try writing the *denominator* of eq. (255) in the form

$$R^2\omega_0^2(\omega/\omega_0)^2 C^2 + [(\omega/\omega_0)^2\omega_0^2 LC - 1]^2 \tag{261}$$

Now let (see eq. (247) in section 8.6)

$$d = \frac{\omega}{\omega_0} \tag{262}$$

Using this notation, and also noting, by eq. (259), that $\omega_0^2 LC = 1$, eq. (261) becomes

$$R^2 d^2 (\omega_0 C)^2 + (d^2 - 1)^2 \tag{263}$$

Next, by eq. (260), $\omega_0 C = 1/\omega_0 L$, and using this relationship, and also the definition of eq. (256), we can now write eq. (255) with a new, dimensionless denominator; thus

$$\frac{\bar{Z}_\mathrm{p}}{R_0} = \frac{\dfrac{R}{R_0} - j\dfrac{\omega C}{R_0}\left(\omega^2 L^2 - \dfrac{L}{C} + R^2\right)}{\dfrac{d^2}{Q^2} + (d^2 - 1)^2} \tag{264}$$

Now let us work on the *numerator* on the right-hand side of the above equation, as follows. First, making use of eqs. (254), (260), and (256), we have that

$$\frac{R}{R_0} = \frac{R^2 C}{L} = \frac{R^2 \omega_0 C}{\omega_0 L} = \left(\frac{R}{\omega_0 L}\right)^2 = \frac{1}{Q^2} \tag{265}$$

Next note that the quantity *inside the parentheses* in the imaginary part of eq. (264) can be written as

$$\begin{aligned}\left(\omega^2 L^2 - \frac{L}{C} + R^2\right) &= \frac{\omega^2 \omega_0^2 L^2}{\omega_0^2} - \frac{\omega_0 L}{\omega_0 C} + \frac{\omega_0^2 L^2}{Q^2} \\ &= (\omega_0 L)^2 (d^2 - 1 + 1/Q^2)\end{aligned} \tag{266}$$

where we used the relationships $\omega_0 L = 1/\omega_0 C$ and $R = \omega_0 L/Q$. Thus, substituting the results of eqs. (265) and (266) into eq. (264), we have that

$$\frac{\bar{Z}_\mathrm{p}}{R_0} = \frac{\dfrac{1}{Q^2} - \dfrac{j(\omega C)(\omega_0 L)^2}{R_0}\left(d^2 - 1 + \dfrac{1}{Q^2}\right)}{\dfrac{d^2}{Q^2} + (d^2 - 1)^2} \tag{267}$$

Now, for the last step (remembering that $\omega_0 L = 1/\omega_0 C$), note that

$$\begin{aligned}\frac{(\omega C)(\omega_0 L)^2}{R_0} = \frac{(\omega C)(\omega_0 L)^2 RC}{L} &= \frac{\omega(\omega_0 C)(\omega_0 L)^2 R(\omega_0 C)}{\omega_0(\omega_0 L)} \\ &= \frac{\omega}{\omega_0}\frac{R}{\omega_0 L} = \frac{d}{Q}\end{aligned}$$

and upon making this substitution into eq. (267), then multiplying the numerator and denominator by Q^2, we get the final desired result

$$\frac{\bar{Z}_\mathrm{p}}{R_0} = \frac{1 - jdQ(d^2 - 1 + 1/Q^2)}{d^2 + Q^2(d^2 - 1)^2} \tag{268}$$

also, therefore,

$$\left|\frac{\bar{Z}_\mathrm{p}}{R_0}\right| = \frac{\sqrt{1 + d^2 Q^2 (d^2 - 1 + 1/Q^2)^2}}{d^2 + Q^2(d^2 - 1)^2} \tag{269}$$

Problem 145
Letting $A = |\bar{Z}_p/R_0|$, fill in the following table of values for Fig. 167, for $Q = 20$ (round off calculator values to two decimal places). Sketch the curve of A versus d.

d	A	d	A	d	A	d	A
0.80		0.95		1.01		1.10	
0.85		0.97		1.03		1.15	
0.90		0.99		1.05		1.20	
0.93		1.00		1.07			

Problem 146
This problem (and problems 147 and 148) deals with Fig. 167 and, specifically, with the circuit of Fig. 168, in which $V\underline{/0^\circ} = V = 90$ volts rms, the sinusoidal reference voltage of the generator, and $\bar{I}_g$ is the generator current, as shown. The values of C, L, and R are in microfarads, microhenrys, and ohms.

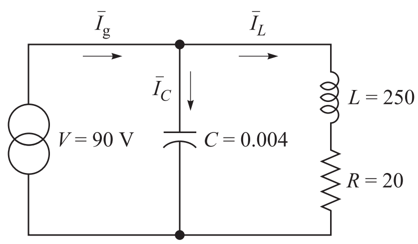


Fig. 168

(a) Find the APPROXIMATE value of ω_0 based upon eq. (259).

(b) Find the EXACT value of ω_0 based upon eq. (252). For practical purposes, will it be reasonable to use the approximate equations in the case of Fig. 168?

Problem 147
At the resonant frequency ω_0 in Fig. 168, find

(a) load seen by generator,

(b) generator current,

(c) power output of generator,

(d) capacitor current,

(e) inductor current,

(f) Q of circuit,

(g) power to R, from answer to (e). Check with answer to (c).

Problem 148
In Fig. 168, suppose the generator frequency is changed to the value $\omega = (1.05)10^6$ radians/second, the values of V, C, L, and R remaining unchanged. Find, approximately, the following values:

(a) value of $\bar{I}_g$, the generator current, (*Answer*: $0.025972 + j0.037122$ amp.)
(b) phase angle of $\bar{I}_g$ relative to generator voltage, (*Answer*: 55.022°)
(c) power output of generator. (*Answer*: 2.338 watts)

In closing, it should be noted that the conditions of series and parallel resonance both have important practical applications. As we have found, SERIES resonance is characterized by the condition of LOW RESISTANCE AND HIGH CURRENT, while in PARALLEL resonance we have the condition of HIGH RESISTANCE AND LOW CURRENT.

CHAPTER 9

Impedance Transformation. Electric Filters

In this chapter we'll consider impedance-matching by means of "L" sections and also by means of "T" and "pi" networks. We then define the decibel, and take up the algebra of some basic low-pass and high-pass filter networks.

9.1 Impedance Transformation. The "L" Section

In practical work it is often necessary that a given generator work into a certain specific value of load impedance. Often, however, the actual load impedance will be fixed at a value different from the value we desire the generator to see. In such a case, the actual load impedance can be transformed into the desired value by inserting an IMPEDANCE TRANSFORMING NETWORK between the generator and the load. This is illustrated, in a general way, in Fig. 169, where the internal impedance of the generator and the load impedance are taken to be pure resistances, R_g and R, as shown.

The impedance-transforming network is assumed to be inside the box, and to be connected to generator and load by means of four leads, as shown. The terminals labeled 1, 1 are the INPUT TERMINALS or "input leads" to the network, and those labeled 2, 2 are the OUTPUT TERMINALS or "output leads" that connect to the load resistance R.

The network inside the box may be of the "L," "T," or "pi" (π), type, or it may consist of a single coupled-circuit transformer. Each has certain advantages and disadvantages, and the type used, in any given case, will depend upon the particular problem being dealt

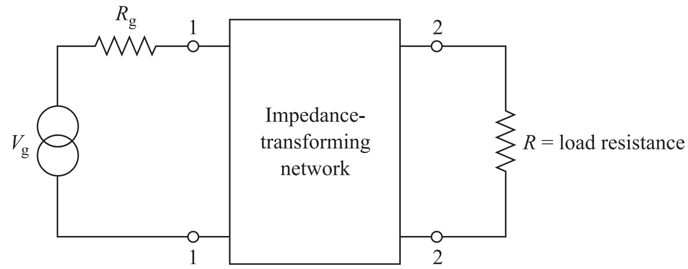


Fig. 169

with. In our work we'll assume the passive elements, inside the box, to be composed of PURE REACTANCES only, so that no energy will be lost in the impedance-matching network itself.

We begin with the "L" type of matching network, which consists of the $L - C$ arrangement shown inside the box outlined by the dashed lines in Fig. 170.

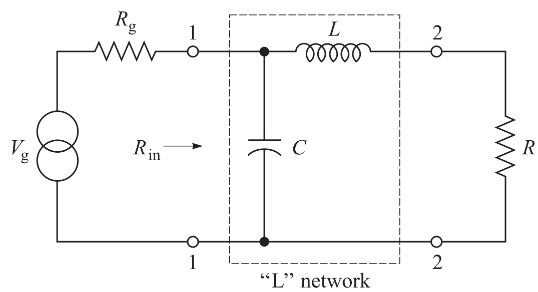


Fig. 170

Note that the capacitor and inductor form an "upside-down L," from which the network gets its name. Next note that the network to the right of terminals (1, 1) in Fig. 170 is *exactly the same* as the network looking into terminals (a, b) in Fig. 167 in section 8.7. Therefore, from our study of Fig. 167, *the resonant frequency*, ω_0, looking into terminals (1, 1) is given by eq. (252); thus

$$\omega_0 = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}} \quad \text{rad/sec} \tag{270}$$

Also, therefore, *at the resonant frequency*, the value of R_{in} in Fig. 170 is the same as the value of R_0 given by eq. (254); thus, *at resonance*,

$$R_{in} = \frac{L}{RC} \quad \text{ohms} \tag{271}$$

Thus, at resonance, the "L" network of Fig. 170 is capable of transforming a given value of load resistance R into a different value of resistance, R_{in}, as seen looking into terminals (1, 1).*

Next, from eq. (271), we have that, at resonance, $L = R_{in}RC$, and $C = L/R_{in}R$, and upon substituting these values of C and L successively into eq. (270), you should find that

* True if $R_{in} > R$, as discussed following eq. (273).

the values L and C, at resonance, in Fig. 170, are equal to

$$L = \frac{R}{\omega_0}\sqrt{\frac{R_{\text{in}}}{R} - 1} \quad \text{henrys} \tag{272}$$

$$C = \frac{1}{\omega_0 R_{\text{in}}}\sqrt{\frac{R_{\text{in}}}{R} - 1} \quad \text{farads} \tag{273}$$

Note, however, that the last two equations are valid only if $R_{\text{in}} > R$, because, if this is not true, the quantity under the square root signs will be negative, which would call for imaginary values of L and C. This case (where $R_{\text{in}} < R$) can, however, be handled by making use of the "reverse L" network shown in Fig. 171.

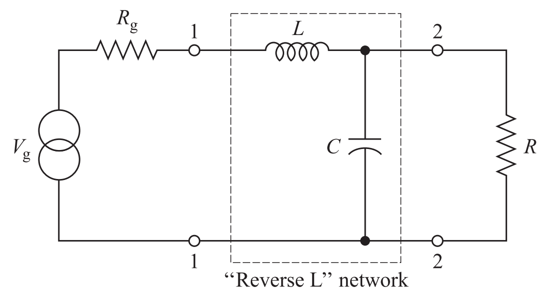


Fig. 171

The equations for the "reverse L" network are as follows. First, the *resonant* frequency ω_0 (the frequency at which the impedance, looking into terminals (1, 1), is a pure resistance) is equal to

$$\omega_0 = \sqrt{\frac{1}{LC} - \frac{1}{R^2C^2}} \quad \text{rad/sec} \tag{274}$$

Next, the value of the pure resistance, looking into terminals (1, 1), at the resonant frequency, is equal to

$$R_{\text{in}} = \frac{L}{RC} \quad \text{ohms} \tag{275}$$

Thus, from the above equation, we have that, at resonance, $L = R_{\text{in}}RC$ and $C = L/R_{\text{in}}R$. Upon substituting these values of C and L into eq. (274), you should find that the values of L and C, at resonance, in Fig. 171, are equal to

$$L = \frac{R_{\text{in}}}{\omega_0}\sqrt{\frac{R}{R_{\text{in}}} - 1} \quad \text{henrys} \tag{276}$$

$$C = \frac{1}{\omega_0 R}\sqrt{\frac{R}{R_{\text{in}}} - 1} \quad \text{farads} \tag{277}$$

To summarize, the principal advantage of the L and reverse L is simplicity, since only one capacitor and one coil are needed. It must be remembered that, strictly speaking, the generator sees a pure resistance only at the one frequency selected, but, practically speaking, also for a narrow band of frequencies centered at the resonant frequency, the range of frequencies depending upon the effective Q of the circuit.

Also, depending upon the frequency of operation and upon the values of R and R_{in}, the values of L and C may come out to be inconveniently large or small, from a practical standpoint. Hence, in some cases it will be necessary or desirable to go to a somewhat more complicated network, such as a "T" or "pi" (π) type.

Problem 149
The matching networks of Figs. 170 and 171 are both referred to, in general terms, as "L-type" networks. If the constants of such a network are known to be $R = 135$ ohms, $L = 28.5\,\mu$H,* and $C = 0.0036\,\mu$F, find the resonant frequency of the network. (*Answer*: 2.348 megahertz)

Problem 150
A load resistance of 16 ohms is to be transformed into 75 ohms by means of an L-type network. If the resonant frequency is to be 10^6 rad/sec, find the required values of C and L.

Problem 151
A load resistance of 125 ohms is to be transformed into 85 ohms by means of an L-type network. If the resonant frequency is to be 360 kHz (kilohertz), find the required values of C and L.

Problem 152
Following the same line of reasoning as in section 8.7, prove that eq. (274) is true.

Problem 153
Prove that eq. (275) is correct.

Problem 154
Derive eqs. (276) and (277).

9.2 The "T" and "Pi" Equivalent Networks

It is often helpful, in the analysis of complicated networks, to replace an actual network with a simpler network that is the EQUIVALENT of the actual network.

A second network is said to be equivalent to a first network if, when the first is replaced by the second, there is NO CHANGE in the values of the voltages and currents appearing AT THE INPUT AND OUTPUT TERMINALS) (1, 1) and (2, 2). Consider now, Figs. 172 and 173.

Let the box in Fig. 172 contain the *actual network*, which can be any linear, bilateral network we might be interested in. (For "linear" and "bilateral," see footnote in section 4.7).

Let $\bar{V}_g$ be the generated voltage of the generator, $\bar{Z}_g$ be the internal impedance of the generator, and $\bar{Z}_L$, be the external load impedance. Also let $\bar{V}_1$, $\bar{I}_1$, $\bar{V}_2$, and $\bar{I}_2$ be the voltages and currents at the input and output terminals of the actual network, as shown in Fig. 172. Since we're dealing with sinusoidal steady-state analysis, these voltages and

* The Greek letter "μ" (mu) is here read as "micro," and indicates multiplication by 10^{-6}. Thus, $28.5\,\mu\text{H} = (28.5)10^{-6}$ H (henrys).

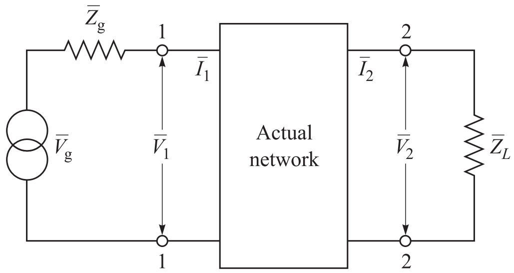


Fig. 172

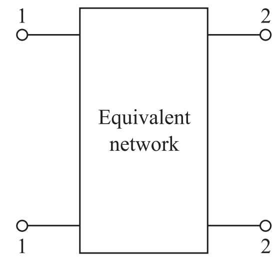


Fig. 173*

currents will be expressed, as usual, in the form of COMPLEX NUMBERS, each such number requiring the specification of TWO SEPARATE VALUES, the "magnitude," and the "phase angle."

Now, in Fig. 172, let us *remove the actual network* and replace it with the proposed EQUIVALENT NETWORK inside the box in Fig. 173. If the network inside the box in Fig. 173 is a true equivalent network, there will be, at the operating frequency, NO CHANGE in either magnitude or phase angle in any of the four complex numbers $\bar{V}_1$, $\bar{I}_1$, $\bar{V}_2$, or $\bar{I}_2$.

Let us now consider what the circuit of an "equivalent network" would have to be like. We begin with the reasonable assumption that the equivalent network should be as simple as possible. The simplest form of an equivalent network can then be arrived at as follows.

First, we have the fact that a complex number is composed of two independent parts, the magnitude and the phase angle. Since we are dealing here with *four* complex numbers ($\bar{V}_1$, $\bar{I}_1$, $\bar{V}_2$, $\bar{I}_2$), it follows that there are *eight* separate quantities (four magnitudes and four phase angles) involved in setting up an equivalent network. Note, however, from Fig. 172, that by Ohm's law

$$\bar{V}_2 = \bar{I}_2\bar{Z}_L$$

and thus, if the *external load impedance* $\bar{Z}_L$ is taken into account, the four complex quantities become

$$\bar{V}_1, \bar{I}_1, \bar{I}_2\bar{Z}_L, \qquad \text{and} \qquad \bar{I}_2$$

Thus only *six* of the quantities (three magnitudes and three phase angles) need be determined for the network itself, the other two (one magnitude and one phase angle) being supplied by the known value of the load impedance $\bar{Z}_L$. Hence the simplest equivalent network need have only *three* independent, adjustable impedances. Some thought will show that there are only *two ways* that three such impedances can be arranged, one arrangement being called the "T" network, the other being called the "pi" (π), as in Figs. 174 and 175.

Notes: In all our discussions of T and pi networks we'll use the same standard notation shown in Figs. 174 and 175. Also, in these figures, each $\bar{Z}$ can represent any simple series, parallel, or series-parallel connection of impedances.

Now let us take up the problem of *how to find the values* of the impedances in the equivalent networks above, beginning with the T network.

* Fig. 173 represents a "four-terminal" or "two-port" network. Terminals (1, 1) constitute the "input port" and (2, 2) the "output port."

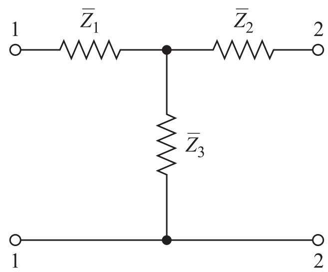


Fig. 174. "T" network.

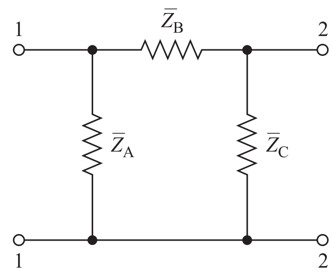


Fig. 175. "Pi" (π) network.

The values of the three unknown impedances in Fig. 174 have to be found by making either impedance calculations or actual measurements *at the terminals of the actual network* it is desired to replace. Since we must find the values of THREE unknown impedances, we must make at least three different, independent measurements on the actual network. The most convenient measurements to make are called the OPEN-CIRCUIT measurements and the SHORT-CIRCUIT measurements—these measurements are made on the *actual network*, as follows.

First, *disconnect* the generator and the load impedance $\bar{Z}_L$ from the actual network. Then make the following measurements or calculations at the terminals of the actual network.

$\bar{Z}_{1O}$ = impedance looking into terminals 1,1 with terminals 2,2 OPEN-CIRCUITED.
$\bar{Z}_{1S}$ = impedance looking into terminals 1,1 with terminals 2,2 SHORT-CIRCUITED.
$\bar{Z}_{2O}$ = impedance looking into terminals 2,2 with terminals 1,1 OPEN-CIRCUITED.
$\bar{Z}_{2S}$ = impedance looking into terminals 2,2 with terminals 1,1 SHORT-CIRCUITED.

The relationship of the above measurements (which, remember, are to be made on the actual network) to the values of the three elements of the hypothetical "equivalent T network" can be found with the aid of Fig. 176.

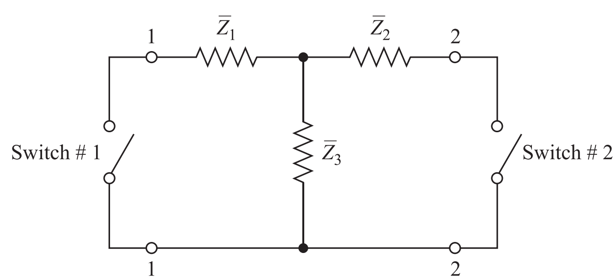


Fig. 176

Referring to Fig. 176, note that operation of the switches will provide us with the following information:

Switch #1	*Switch* #2		
open	open	$\bar{Z}_{1O} = \bar{Z}_1 + \bar{Z}_3$	(278)
open	closed	$\bar{Z}_{1S} = \bar{Z}_1 + \dfrac{\bar{Z}_2\bar{Z}_3}{\bar{Z}_2 + \bar{Z}_3}$	(279)
open	open	$\bar{Z}_{2O} = \bar{Z}_2 + \bar{Z}_3$	(280)
closed	open	$\bar{Z}_{2S} = \bar{Z}_2 + \dfrac{\bar{Z}_1\bar{Z}_3}{\bar{Z}_1 + \bar{Z}_3}$	(281)

Now subtract eq. (279) from (278): $\bar{Z}_{1\text{O}} - \bar{Z}_{1\text{S}} = \bar{Z}_3 - \dfrac{\bar{Z}_2\bar{Z}_3}{\bar{Z}_2 + \bar{Z}_3}$. Now multiply both sides of this equation by $(\bar{Z}_2 + \bar{Z}_3)$, then make use of eq. (280); this should give you

$$\bar{Z}_3 = +\sqrt{\bar{Z}_{2\text{O}}(\bar{Z}_{1\text{O}} - \bar{Z}_{1\text{S}})}$$

in which, as is shown, it's customary to use just the positive value of the square root. Thus, making use of this equation and also eqs. (280) and (278), we get the following relationships

$$\bar{Z}_3 = \sqrt{\bar{Z}_{2\text{O}}(\bar{Z}_{1\text{O}} - \bar{Z}_{1\text{S}})} \tag{282}$$

$$\bar{Z}_2 = \bar{Z}_{2\text{O}} - \bar{Z}_3 \tag{283}$$

$$\bar{Z}_1 = \bar{Z}_{1\text{O}} - \bar{Z}_3 \tag{284}$$

These are the three equations required to convert any actual four-terminal network into an EQUIVALENT T NETWORK. Note that the values of $\bar{Z}_{1\text{O}}$, $\bar{Z}_{1\text{S}}$, and $\bar{Z}_{2\text{O}}$ are found by making three calculations or measurements at the terminals of the ACTUAL NETWORK.

It should be understood that if the three equations are used to calculate definite values of R, L, and C at a given frequency, then the T network, so found, is equivalent to the actual network *only at that one frequency*. Practically speaking, however, such an equivalent network can satisfactorily replace the actual network over a band of frequencies usually extending a few percent above and below the "center frequency."

Now let's turn our attention to the *equivalent "pi" network* of Fig. 175. The equations for finding the values of an equivalent pi network are found in the same general way as those for the equivalent T network, and are as follows:

$$\bar{Z}_\text{A} = \frac{\bar{Z}'}{\bar{Z}_{2\text{O}} - \bar{Z}''} \tag{285}$$

$$\bar{Z}_\text{B} = \bar{Z}'/\bar{Z}'' \tag{286}$$

$$\bar{Z}_\text{C} = \frac{\bar{Z}'}{\bar{Z}_{1\text{O}} - \bar{Z}''} \tag{287}$$

where "Z prime" and "Z double prime" have the following values:

$$\bar{Z}' = \bar{Z}_{2\text{O}}\bar{Z}_{1\text{S}} \tag{288}$$

$$\bar{Z}'' = \sqrt{\bar{Z}_{2\text{O}}(\bar{Z}_{1\text{O}} - \bar{Z}_{1\text{S}})} \tag{289}$$

where $\bar{Z}_{1\text{O}}$ through $\bar{Z}_{2\text{S}}$ are found in the same way as for the equivalent T, as defined just prior to Fig. 176. In this regard, let us now replace the T network in Fig. 176 with the pi network of Fig. 175. Upon doing this, operation of the two switches will now produce the following results, as you should verify:

Switch #1	*Switch #2*		
open	open	$\bar{Z}_{1O} = \dfrac{\bar{Z}_A(\bar{Z}_B + \bar{Z}_C)}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C}$	(290)
open	closed	$\bar{Z}_{1S} = \dfrac{\bar{Z}_A\bar{Z}_B}{\bar{Z}_A + \bar{Z}_B}$	(291)
open	open	$\bar{Z}_{2O} = \dfrac{\bar{Z}_C(\bar{Z}_A + \bar{Z}_B)}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C}$	(292)
closed	open	$\bar{Z}_{2S} = \dfrac{\bar{Z}_B\bar{Z}_C}{\bar{Z}_B + \bar{Z}_C}$	(293)

Now, as an exercise in algebraic manipulation, let's verify that the values of $\bar{Z}_A$ and $\bar{Z}_B$ given by eqs. (285) and (286) are correct; one way to do this is as follows.

First, by eq. (291),

$$\bar{Z}_B = \frac{\bar{Z}_A\bar{Z}_{1S}}{\bar{Z}_A - \bar{Z}_{1S}} \tag{294}$$

Next, using the relationships found in the above switching table, we have that

$$\bar{Z}_{1S}\bar{Z}_{2O} = \frac{\bar{Z}_A\bar{Z}_B\bar{Z}_C}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} \tag{295}$$

Next, again making use of the relationships in the switching table, you can verify that

$$\bar{Z}_{2O}(\bar{Z}_{1O} - \bar{Z}_{1S}) = \frac{\bar{Z}_A^2\bar{Z}_C^2}{(\bar{Z}_A + \bar{Z}_B + \bar{Z}_C)^2} \tag{296}$$

Now *invert* both sides of the last equation, then take the square root of both sides, then multiply, respectively, the left-hand and right-hand sides of the result by the left-hand and right-hand sides of eq. (295); doing this, you should find that

$$\frac{\bar{Z}_{1S}\bar{Z}_{2O}}{\sqrt{\bar{Z}_{2O}(\bar{Z}_{1O} - \bar{Z}_{1S})}} = \bar{Z}_B = \frac{\bar{Z}'}{\bar{Z}''}$$

thus verifying that eq. (286) *is correct*. Now, in the above, substitute the right-hand side of eq. (294), in place of $\bar{Z}_B$, then solve for $\bar{Z}_A$ to verify that eq. (285) *is also correct*.

Next, the values of $\bar{Z}_A$ and $\bar{Z}_B$ (from eqs. (285) and (286)) can be substituted into, for example, eq. (292), the result then being solved for the value of $\bar{Z}_C$, which will prove that eq. (287) *is correct*.

The above results can be summarized in the statement that any linear, bilateral network, containing no internal generators, can be represented, at a single frequency, by a T or pi network.

Problem 155

The network in Fig. 177 is composed of pure resistances having values in ohms, as shown. Find the equivalent T network. Is the answer valid at all frequencies?

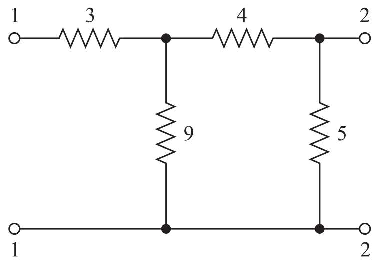


Fig. 177

Problem 156

In Fig. 178, it is given that $C = 2\,\mu\text{F}$ and $L = 150\,\mu\text{H}$, the resistance values being in ohms, as shown.

Draw the diagram of the equivalent T network, showing the required values of inductance, capacitance, and resistance, for operation at 10^5 radians/second.

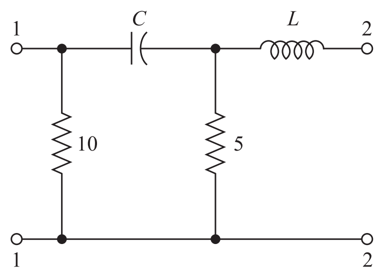


Fig. 178

Problem 157

Draw the T network equivalent of the purely resistive "bridge-type" network shown in Fig. 179. Resistance values are in ohms.

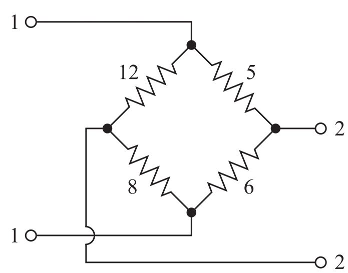


Fig. 179

Problem 158

Suppose the following measurements are made on a certain network at a particular frequency of interest:

$$\bar{Z}_{1\text{O}} = 18 + j12\,\text{ohms},$$

$$\bar{Z}_{1\text{S}} = 8 + j18\,\text{ohms},$$

$$\bar{Z}_{2\text{O}} = 20 - j12\,\text{ohms}.$$

Find the values of the equivalent pi representation of the network, at the frequency of interest.

9.3 Conversion of Pi to T and T to Pi

In practical work it's sometimes helpful to convert a given "pi" network into an equivalent "T" network, or to convert a given T network into an equivalent pi network. This can be done as follows, in which we'll continue to use the standard notation of Figs. 174 and 175.

In section 9.2 we defined the quantities $\bar{Z}_{1O}$, $\bar{Z}_{1S}$, $\bar{Z}_{2O}$, and $\bar{Z}_{2S}$ as being the values of *external measurements* made at the input and output terminals of a network. It follows that if two networks are to be *equivalent*, then the values of these external measurements must be the *same* for both networks. Algebraically, this means that the right-hand side of eq. (278) must be equal to the right-hand side of eq. (290), the right-hand side of eq. (279) must be equal to the right-hand side of eq. (291), the right-hand side of eq. (280) must be equal to the right-hand side of eq. (292), and likewise for eqs. (281) and (293); thus the following system of equations must be satisfied:

$$\bar{Z}_1 + \bar{Z}_3 = \frac{\bar{Z}_A(\bar{Z}_B + \bar{Z}_C)}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} \tag{297}$$

$$\bar{Z}_1 + \frac{\bar{Z}_2\bar{Z}_3}{\bar{Z}_2 + \bar{Z}_3} = \frac{\bar{Z}_A\bar{Z}_B}{\bar{Z}_A + \bar{Z}_B} \tag{298}$$

$$\bar{Z}_2 + \bar{Z}_3 = \frac{(\bar{Z}_A + \bar{Z}_B)\bar{Z}_C}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} \tag{299}$$

$$\bar{Z}_2 + \frac{\bar{Z}_1\bar{Z}_3}{\bar{Z}_1 + \bar{Z}_3} = \frac{\bar{Z}_B\bar{Z}_C}{\bar{Z}_B + \bar{Z}_C} \tag{300}$$

Equations (297) through (300) express the relationships that must always exist between two equivalent T and pi networks. Making use of these relationships, we can derive equations that will allow us to convert from one type of network to the other.

Suppose, for example, that $\bar{Z}_A$, $\bar{Z}_B$, and $\bar{Z}_C$ are *known*, and we wish to find the equations for calculating the equivalent T network. After several false starts, we find that the following procedure will work.

First, subtract eq. (298) from eq. (297) to get

$$\bar{Z}_3 - \frac{\bar{Z}_2\bar{Z}_3}{\bar{Z}_2 + \bar{Z}_3} = \frac{\bar{Z}_A(\bar{Z}_B + \bar{Z}_C)}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} - \frac{\bar{Z}_A\bar{Z}_B}{\bar{Z}_A + \bar{Z}_B}$$

or, after putting the left side over its common denominator and the right side over its common denominator, we have

$$\frac{\bar{Z}_3^2}{\bar{Z}_2 + \bar{Z}_3} = \frac{\bar{Z}_A^2\bar{Z}_C}{(\bar{Z}_A + \bar{Z}_B + \bar{Z}_C)(\bar{Z}_A + \bar{Z}_B)}$$

We've so far made use of eqs. (297) and (298); we can now make use of eq. (299), as follows. Multiply both sides of the last equation by "$\bar{Z}_2 + \bar{Z}_3$," then replace $\bar{Z}_2 + \bar{Z}_3$ by the right-hand side of eq. (299). Doing this, then solving for $\bar{Z}_3$, you should find that

$$\bar{Z}_3 = \frac{\bar{Z}_A\bar{Z}_C}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} \tag{301}$$

which allows us to find the value of $\bar{Z}_3$. Equation (301) is thus the *first* of the three equations we require; equations for finding the values of $\bar{Z}_1$ and $\bar{Z}_2$ can now be found as follows.

Substitute, into the left-hand side of eq. (297), the value of $\bar{Z}_3$ just found in eq. (301); doing this gives us the value of $\bar{Z}_1$; thus

$$\bar{Z}_1 = \frac{\bar{Z}_A \bar{Z}_B}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} \tag{302}$$

which is the *second* of the three equations we require. Next, substitute the value of $\bar{Z}_3$, from eq. (301), into the left-hand side of eq. (299) to get

$$\bar{Z}_2 = \frac{\bar{Z}_B \bar{Z}_C}{\bar{Z}_A + \bar{Z}_B + \bar{Z}_C} \tag{303}$$

which is the *third* and final equation that we require. Thus eqs. (301), (302), and (303) are the three equations needed to convert *a given pi network into an equivalent T network*. Again, it must be noted that, if the three equations are used to calculate definite values of R, L, and C *at a specific frequency*, then the T network, thus found, is equivalent to the given pi network *only at that specific frequency* (see discussion following eq. (284) in section 9.2).

Now consider the opposite problem; that is, suppose we must convert a *given T network* into its equivalent pi network. In such a case it can be shown that the required three equations are as follows:

$$\bar{Z}_A = \frac{\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3}{\bar{Z}_2} \tag{304}$$

$$\bar{Z}_B = \frac{\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3}{\bar{Z}_3} \tag{305}$$

$$\bar{Z}_C = \frac{\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3}{\bar{Z}_1} \tag{306}$$

Problem 159
Have a try at proving that eqs. (304), (305), and (306) are correct.

Problem 160
Given the pi network of Fig. 180, find the R, L, and C values for the equivalent T network for operation at 500 kilohertz ("Ω" denotes "ohms").

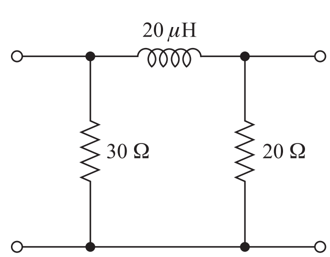


Fig. 180

Problem 161
The T network of Fig. 181 is composed of pure reactances having the values as shown for the frequency of operation. Find the values of the reactances required for the equivalent pi network.

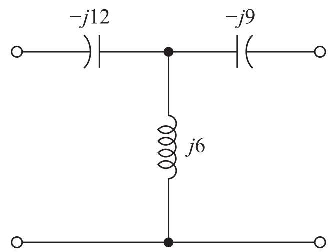


Fig. 181

9.4 Impedance Transformation by T and Pi Networks

In section 9.1 we studied the use of the "L" network as an impedance-transforming device. At the conclusion of that section we mentioned that it will often be necessary to use a T or a pi network in place of the simpler L network. This is especially true in certain applications where the suppression of harmonic frequencies* is important, such as in the output stage of a radio transmitter.

With this in mind, let us now look into the possibility of using a T or a pi network as an impedance-changing device; this can be done with the aid of Fig. 182 for the T case.

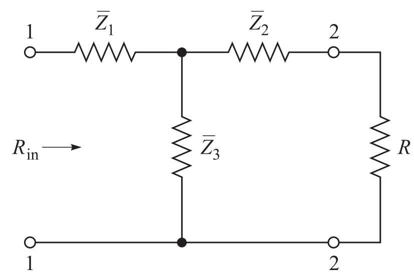


Fig. 182

In this section we'll assume the actual load to be a pure resistance of R ohms, as shown in the figure. We will also assume that we wish to see a pure resistance of R_{in} ohms looking into the input terminals (1, 1) as shown in the figure. This requirement can be stated mathematically for Fig. 182 by noting that, looking to the right into terminals (1, 1), we see that $\bar{Z}_1$ is *in series* with the combination of $\bar{Z}_3$ in parallel with $\bar{Z}_2 + R$; thus, looking into terminals (1, 1) we have that

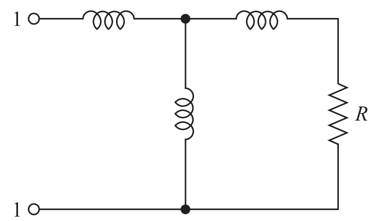
$$R_{in} = \bar{Z}_1 + \frac{(\bar{Z}_2 + R)\bar{Z}_3}{\bar{Z}_2 + \bar{Z}_3 + R} \tag{307}$$

Or, upon multiplying both sides of the above equation by $(\bar{Z}_2 + \bar{Z}_3 + R)$, you should verify that

$$RR_{in} + R_{in}(\bar{Z}_2 + \bar{Z}_3) = \bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3 + R(\bar{Z}_1 + \bar{Z}_3) \tag{308}$$

* See note 18 in Appendix.

Since we desire that no energy be lost in the T network itself, it follows that the three elements of the network will have to be composed of *pure reactances only*. As a first possibility, suppose we made all three elements of the T network inductive reactances (coils), as in the figure below.



In such a case it's apparent that we could never see a pure resistance looking into (1, 1) because there is no capacitance present to cancel out the inductive reactance.

In the same way, if all three elements were *capacitors* we could not possibly see a pure resistance looking into terminals (1, 1). Thus, for the particular conditions of this section, the T network will have to contain *both* inductive reactance and capacitive reactance—only then can there be a total cancellation of reactances, making it possible to see a pure resistance, R_{in}, when looking into terminals (1, 1). Thus we must have either two inductors and one capacitor, or two capacitors and one inductor, as shown in Figs. 183 and 184.

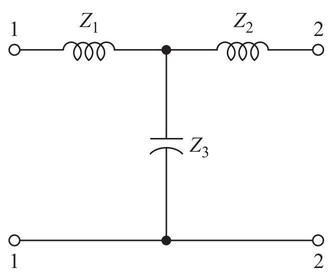


Fig. 183

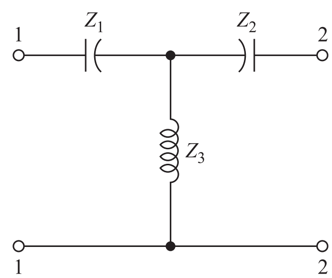


Fig. 184

In Fig. 183 note that $Z_1 = jX_1$, $Z_2 = jX_2$, $Z_3 = -jX_3$, while, in Fig. 184, $Z_1 = -jX_1$, $Z_2 = -jX_2$, $Z_3 = jX_3$. Putting these values into eq. (108) we have, for Fig. 183:

$$RR_{in} + jR_{in}(X_2 - X_3) = (-X_1X_2 + X_1X_3 + X_2X_3) + jR(X_1 - X_3)$$

for Fig. 184:

$$RR_{in} + jR_{in}(-X_2 + X_3) = (-X_1X_2 + X_1X_3 + X_2X_3) + jR(-X_1 + X_3)$$

Recall that two complex numbers can be equal only if the two real parts are equal and the two imaginary parts are equal; hence, inspection of the last two equations shows that for *both* Figs. 183 and 184:

$$RR_{in} = (-X_1X_2 + X_1X_3 + X_2X_3) \tag{309}$$

for Fig. 183:

$$R_{in}(X_2 - X_3) = R(X_1 - X_3) \tag{310}$$

for Fig. 184:

$$R_{\text{in}}(-X_2 + X_3) = R(-X_1 + X_3) \tag{311}$$

In a practical problem the value of the *actual* load resistance R (Fig. 182) and the *desired* value of R_{in} would be known. This would leave us with three unknown reactances and, for either of the above two figures, just two simultaneous equations. In such a case we could select any reasonable value for one of the three reactances, say X_3, and then use the two equations to calculate the required values of the other two reactances.

Actually, however, in most practical work a "balanced T" (or "balanced pi" if we're using a pi network) network would be used. A "balanced" network is one in which the *magnitudes* of the three reactances are all equal in value; thus, setting $X_1 = X_2 = X_3 = X$ in eq. (309), we find that, for a balanced T network, the *common magnitude of reactance*, for either Fig. 183 or Fig. 184, is equal to

$$X = \sqrt{RR_{\text{in}}} \tag{312}$$

In addition to the balanced T network, the "balanced pi" network, shown in Fig. 185, finds wide use as the output stage of a radio transmitter.

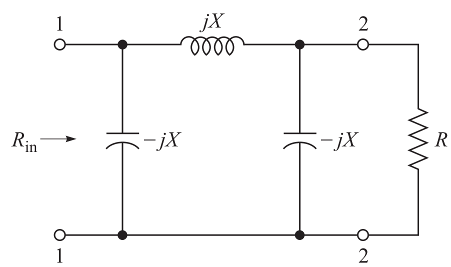


Fig. 185

In terms of standard pi-network notation (Fig. 175 in section 9.2) we have, for the balanced pi network of Fig. 185,

$$\bar{Z}_A = -jX, \quad \bar{Z}_B = jX, \quad \bar{Z}_C = -jX$$

We can, if we wish, convert the above pi network into an *equivalent T network*, thus

$$\text{by eq. (302),} \qquad \bar{Z}_1 = \frac{(-jX)(jX)}{-jX} = jX$$

$$\text{by eq. (303),} \qquad \bar{Z}_2 = \frac{(jX)(-jX)}{-jX} = jX$$

$$\text{by eq. (301),} \qquad \bar{Z}_3 = \frac{(-jX)(-jX)}{-jX} = -jX$$

We can now make use of the equations for the T network, and thus, setting $X_1 = X_2 = X_3 = X$ in eq. (309), we have, *for the balanced pi network* of Fig. 185, that

$$X = \sqrt{RR_{\text{in}}} \tag{313}$$

as for the balanced T network (eq. (312)).

Problem 162

A generator, having an internal resistance of 36 ohms, generates, on open circuit, 90 volts rms at a frequency of 175 kHz (kilohertz). It is necessary that the generator deliver maximum power to a load resistance of 115 ohms.* It is desired to use, as an impedance-matching device, a balanced T network of the type shown in Fig. 183. Find the following:

(a) values of L and C required,
(b) generator current,
(c) voltage at input terminals of the T network,
(d) load current and load voltage.

Problem 163
T and pi networks are widely used in the output stages of radio transmitters, one purpose being to transform a given load resistance into a more suitable value of resistance, the other purpose being to reduce the transmission of undesired harmonic energy. The following will illustrate this feature.

A generator, having 36 ohms of internal resistance, generates 100 volts rms at a frequency of 300,000 radians per second, and 20 volts at the third harmonic frequency of 900,000 rad/sec.

It is desired to deliver maximum possible power to a 100-ohm load at the fundamental frequency of 300,000 rad/sec and, in addition, to reduce the percentage of the undesired third harmonic voltage appearing across the 100-ohm load. To do this, it is proposed to use a balanced pi network of the form of Fig. 185. Assuming the "principle of superposition" applies, find

(a) values of L and C required, (*Answer*: 200 μH, 0.0555 μF)
(b) magnitude of the desired fundamental-frequency voltage across the 100-ohm load, (*Answer*: 83.34 volts)
(c) magnitude of the undesired third-harmonic voltage across the 100-ohm load. (*Answer*: 1.305 volts)

9.5 Frequency Response. The Basic *RC* and *RL* Filter Circuits

In this section we introduce the important concept of the "sinusoidal steady-state FREQUENCY RESPONSE" of linear† electrical networks.

To find the steady-state "frequency response" of a given system, we feed *into* the system a sine wave of *constant amplitude* (constant peak value, and thus constant rms value), but whose *frequency* can be adjusted to any value we desire. Since the system is linear, the *output wave* will be a pure sinusoid of the same frequency as the input wave, but will have, in general, a *different amplitude and phase* from the input signal. This is illustrated in Fig. 186.

---

* It is a fact that if the internal resistance R_g of a generator is FIXED in value, but the load resistance R_L is ADJUSTABLE in value, then MAXIMUM POWER will be produced in the load resistance when the load resistance is made EQUAL to the fixed internal resistance of the generator, that is, when, $R_L = R_g$.

It should be noted, however, that, for this condition ($R_L = R_g$) *half* the total power produced is dissipated internally in the generator. Hence this condition is generally used only in low-power applications, where the heat produced in the generator can be carried away fast enough to prevent the generator from being destroyed by overheating.

† Note that "frequency" is not connected with the property of "linearity" (see first footnote in section 8.6).

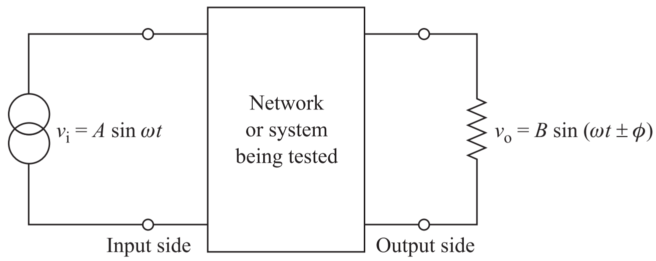


Fig. 186

In Fig. 186,

A = peak value (amplitude) of the input sine wave. The value of A is held *constant*, while we are free to vary the frequency in any way we wish.

B = amplitude of output signal. Generally speaking, the value of B will vary with frequency, depending upon the network or system inside the box.

ϕ = the *phase angle* or "phase shift" of the output sinusoid *relative to the input sine wave*. The angle ϕ is basically in radians (ωt = radians) and can be either positive or negative, depending on the nature of the network inside the box.

The sinusoidal waveforms associated with Fig. 186 are illustrated in Figs. 187 and 188.

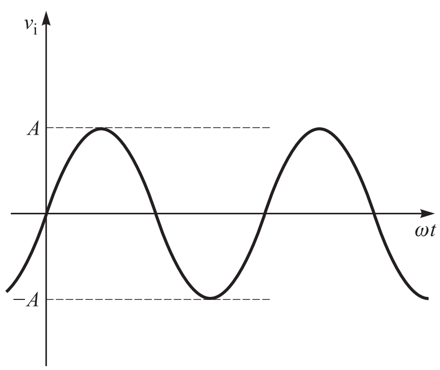


Fig. 187. INPUT sine wave in Fig. 186.

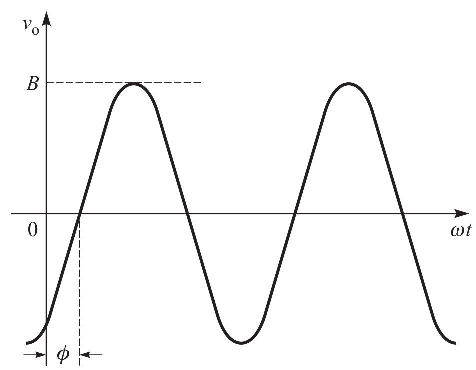


Fig. 188. OUTPUT sine wave in Fig. 186.

Note 1: The output wave in Fig. 188 happens to be shown as "lagging" the input wave by ϕ radians; hence, in this case, the equation shown on the output side of Fig. 186 would be written as $v_o = B\sin(\omega t - \phi)$.

Note 2: The ratio B/A is the VOLTAGE GAIN of the system (or voltage loss, if $B < A$).

It should be noted, however, that it's often more important to specify the POWER GAIN of a system instead of the voltage gain. The power gain, P, of a system is defined as the ratio of the output power to the input power (output watts to input watts); thus

$$P = \frac{P_{out}}{P_{in}} \tag{314}$$

Actually, however, it's generally more meaningful to deal with the *logarithm* of P instead of directly with P (this is because of certain characteristics of human response to changes in power levels). The unit most often used to measure power ratios is called the *decibel* (abbreviated ''dB''), and is defined by the equation

$$\text{dB} = 10 \log P \tag{315}$$

where ''$\log P$'' is the logarithm of P to the base 10.*

Problem 164

(a) If the input power to an amplifier system is 8 watts and the output power is 16 watts, find the power gain in decibels.
(b) If the input power to a network is 1.65 watts and the output power is 0.26 watt, express the loss in decibels.
(c) If the input power to a system is 0.75 watt, and the system is known to have a power gain of 18 decibels, find the output power.

To continue, let's next consider the condition of FREQUENCY DISTORTION or ''frequency discrimination.''

To begin, *frequency distortion* (or ''discrimination'') is produced by the *unequal treatment of the different frequency components* (harmonics) in a given signal. Frequency distortion is thus produced by the presence of L or C, or both, in a circuit; this is because X_L and X_C are both *functions of frequency* (ωL and $1/\omega C$).

Thus, suppose a certain non-sinusoidal signal wave is applied to the input terminals of a network containing inductance L. Since the value of X_L is *different* for each harmonic component, it follows that the relative values of the different harmonic frequencies appearing at the output side of the network will be *different* from their relative values at the input side of the network, and thus the output signal will not have exactly the same waveshape as the input signal. Thus the output wave will be *distorted* to some extent, relative to the given input signal wave.

It should be noted that ''frequency discrimination'' is sometimes an undesirable condition and sometimes a necessary condition. Thus, while such discrimination is undesirable in, say, an audio amplifier system, it is a necessary condition in the operation of frequency ''filter'' networks.

To complete the above discussion, it should be noted that there are really *two factors* to be considered when evaluating the effect of frequency distortion. The first concerns the amount of AMPLITUDE distortion produced, and the second concerns what is called TIME-DELAY or PHASE distortion.

Thus, the various harmonic frequencies at the input side of a system have certain amplitudes and time relationships with one another, and if either the relative amplitudes or the relative time relationships change, the resultant waveshape at the output side of the system will not be exactly the same as the input waveshape.†

The presence of frequency distortion can be found experimentally by applying a *constant-amplitude* sine wave to the input side of a system and measuring the resulting amplitude and phase angle at the output side as the frequency of the input sine wave is slowly varied over any desired frequency range (keeping the amplitude of the input sine wave constant). The output amplitudes and phase angles, thus found, are then plotted on graph paper. Actually, in most practical work, only the variations in output *amplitudes* are recorded and plotted, since if a system has acceptable amplitude response it will generally

* See note 19 in Appendix.

† See note 20 in Appendix.

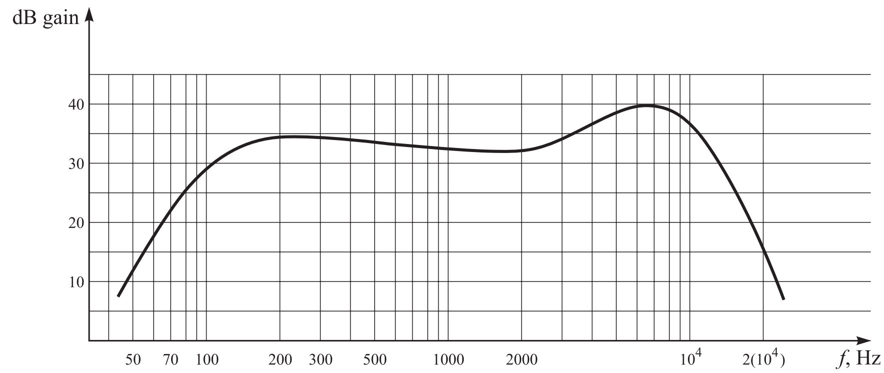


Fig. 189

also have acceptable phase response. A typical frequency response curve, showing decibel gain versus frequency, drawn on semi-log paper,* is shown in Fig. 189.

With the foregoing in mind, let's now apply the algebra of sinusoidal steady-state circuit analysis to some basic circuits, as follows.

CASE I: BASIC LOW-PASS *RC* NETWORK

This is a series *RC* circuit in which *the output voltage appears across the capacitor*, as shown in Fig. 190.

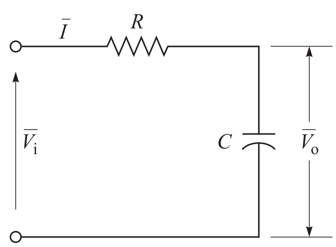


Fig. 190

In the figure, since R and C remain constant in value, and since the reactance of C decreases with increasing frequency, it follows that *as the* FREQUENCY INCREASES *the* OUTPUT VOLTAGE ACROSS C DECREASES; in other words, the higher the frequency, the lower is the output voltage (we are, of course, assuming the amplitude of $\bar{V}_i$ remains constant as the frequency increases). Thus the circuit of Fig. 190 discriminates against the higher frequencies, and is therefore called a "low-pass" type of network.

In Fig. 190 let the input voltage be the reference vector; thus $\bar{V}_i = V_i\underline{/0^\circ} = V_i$. Then, in the figure, let us denote the *ratio* of $\bar{V}_o$ to V_i by "$\bar{G}$" ($\bar{G} = \bar{V}_o/V_i$), which we'll call the "gain" of the circuit.

We now wish to find the manner in which $\bar{G}$ varies *in magnitude and phase* as the frequency of V_i changes (the amplitude of V_i always remaining constant). To do this we first note that, by Ohm's law, in Fig. 190,

$$\bar{I} = \frac{V_i}{R - jX_C}$$

* See note 21 in Appendix.

thus, again by Ohm's law,

$$\bar{V}_o = \bar{I}(-jX_C) = \frac{V_i(-jX_C)}{R - jX_C}$$

Now, dividing both sides by V_i and noting that $-jX_C = \frac{-j}{\omega C} = \frac{1}{j\omega C}$, you should find that

$$\bar{G} = \frac{1}{1 + j\omega RC} \tag{316}$$

where $\bar{G}$ is the SINUSOIDAL STEADY-STATE VOLTAGE GAIN of the basic RC low-pass filter of Fig. 190, where ω is any frequency in radians per second ($\omega = 2\pi f$).

In eq. (316), let us regard the product RC as having a constant value in any given case, with the *variable* being the frequency ω. We now wish to develop certain important relationships that exist between $\bar{G}$ and ω. This can be done in an interesting way algebraically, as follows.

Let us begin by creating, by definition, the equation

$$RC = 1/\omega_1 \tag{317}$$

where, since RC is constant, the frequency ω_1 is also constant. It is permissible to write such an equation, because the unit of measurement for both RC and $1/\omega$ is "seconds" (eq. (91) in Chap. 5, and note 14 in the Appendix). Thus, replacing RC with the right-hand side of eq. (317), eq. (316) becomes

$$\bar{G} = \frac{1}{1 + j(\omega/\omega_1)} \tag{318}$$

The advantage of eq. (318) over (316) is that we no longer have to work with actual absolute values of ω (such as $\omega = 508$ rad/sec or $\omega = 10{,}750$ rad/sec, and so on), but only with the simple *ratio* of ω to ω_1; the frequency is now said to be "normalized" with respect to the reference frequency ω_1.

Eq. (318) contains all the information, concerning both the amplitude and phase response, of the network of Fig. 190. Let us first investigate the amplitude response, as follows.

The AMPLITUDE response is determined by the MAGNITUDE of eq. (318); thus the amplitude response of Fig. 190 is equal to

$$|\bar{G}| = \frac{1}{\sqrt{1 + (\omega/\omega_1)^2}} = [1 + (\omega/\omega_1)^2]^{-1/2} \tag{319}$$

Now, as noted in note 19 in the Appendix, because $\bar{G}$ is a voltage ratio, the decibel relationship would be written

$$\text{dB} = 20 \log |\bar{G}| \tag{320}$$

and thus, using the value of $|\bar{G}|$ from eq. (319) and remembering that $\log A^n = n \log A$, eq. (320) gives the value

$$\text{dB} = -10 \log[1 + (\omega/\omega_1)^2] \tag{321}$$

The minus sign appears because the "voltage gain" in Fig. 190 is less than 1 for all values of ω greater than zero, and the logarithm of a number less than 1 is negative.

The procedure now is to plot eq. (321) on semilog paper, putting decibel gain on the linear vertical axis and the frequency ratio ω/ω_1 on the logarithmic horizontal axis.

To draw the curve, we first calculate a "table of values" of decibel gain for various values of the independent variable ω/ω_1. Thus, the following table was obtained by using eq. (321) to calculate dB gain for each of the following given values of ω/ω_1.

ω/ω_1	dB gain	ω/ω_1	dB gain	ω/ω_1	dB gain
0.2	−0.17	2.0	−6.99	10	−20.04
0.4	−0.65	4.0	−12.31	20	−26.03
0.6	−1.34	6.0	−15.68	30	−29.55
0.8	−2.15	8.0	−18.13	40	−32.04
1.0	−3.01			50	−33.98

Plotting decibel gain versus frequency ratio, we get the curve shown in Fig. 191.

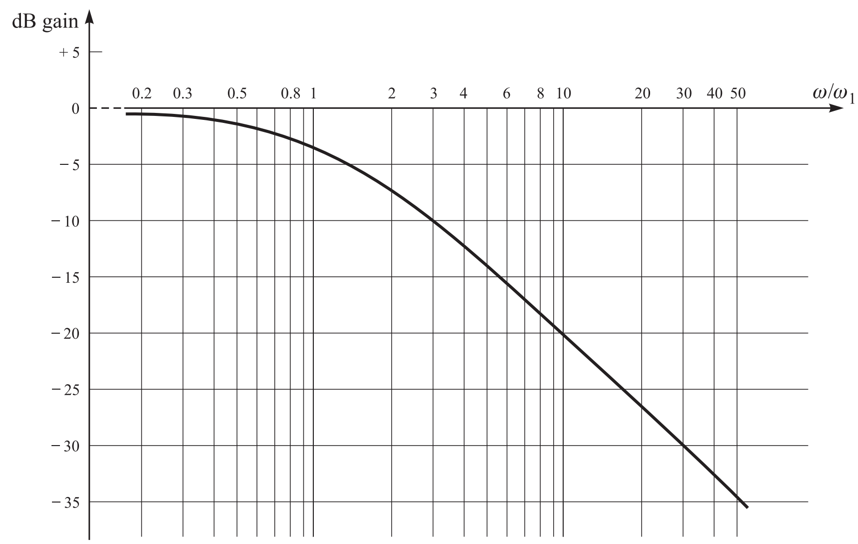


Fig. 191

Remember that negative decibel gain means signal attenuation; that is, the output voltage is less than the input voltage. Figure 191 shows that Fig. 190 is a low-pass network, because the higher the frequency ω, the greater is the attenuation of the output signal relative to the constant amplitude input signal.

This is an appropriate time to introduce another term, called the HALF-POWER frequency, that is widely used in the evaluation of the frequency response of networks and systems. The definition is as follows.

A "half-power" frequency is a frequency at which the OUTPUT POWER of a network is reduced to ONE HALF its maximum value under the condition of constant amplitude of input signal. Thus, setting $P = 1/2 = 0.5$ in eq. (315), we have that

$$\text{dB} = 10 \log 0.5 = -3 \text{ decibels, very nearly}$$

showing that a half-power frequency is a frequency at which the power gain of a network is *down 3 decibels* from its maximum or reference power. Note that, for the PARTICULAR

CASE of the low-pass RC circuit of Fig. 190, upon setting $\omega/\omega_1 = 1$ in eq. (321), we have that

$$\text{dB} = -10 \log 2 = -3 \text{ decibels}$$

hence the half-power frequency for Fig. 190 is at $\omega/\omega_1 = 1$, that is, for $\omega = \omega_1 = 1/RC$ (making use of eq. (317)).

Problem 165
Suppose, in Fig. 190, that $C = 0.05\,\mu\text{F}$. Find the required value of R if the half-power frequency is to be 7.2 kHz.

Problem 166
In problem 165, at what frequency, in kHz, will the power gain be -6 decibels?

So far we've concentrated our attention on the *amplitude* response of Fig. 190 (given by eq. (319)). Let us now complete our work with an examination of the PHASE response of the network. To do this, we can begin with the basic relationship

$$\bar{V}_o = \bar{G} V_i \tag{322}$$

where V_i is the reference input rms voltage. Thus, by eq. (318),

$$\bar{V}_o = \frac{V_i}{1 + j(\omega/\omega_1)} = \frac{V_i}{1 + jh} \tag{323}$$

where, for convenience, we've temporarily set $(\omega/\omega_1) = h$. Now let's "rationalize" the last fraction; that is, let us multiply the numerator and denominator by the "conjugate" of the denominator, $(1 - jh)$,* so that eq. (323) becomes

$$\bar{V}_o = \left[\frac{1}{1+h^2} + j\,\frac{-h}{1+h^2}\right] V_i = (A + jB) V_i \tag{324}$$

where

$$A = 1/(1 + h^2)$$

and

$$B = -h/(1 + h^2)$$

Equation (324) shows that component A is IN PHASE with the reference vector V_i, while component B LEADS V_i by 90 degrees, the *vector sum* of A and B being IN PHASE with the output voltage vector $\bar{V}_o$, as shown in Fig. 192, in which ϕ (phi or "fee") is the PHASE ANGLE (phase shift) between the output voltage $\bar{V}_o$ and the input reference vector V_i.

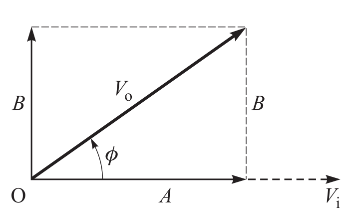


Fig. 192. Here, $\tan\phi = \text{opp}/\text{adj} = B/A$; thus, $\phi = \arctan(B/A)$.

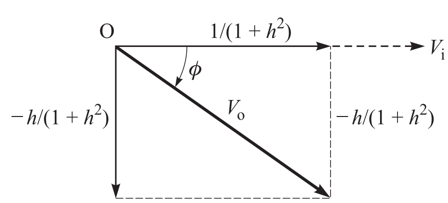


Fig. 193. Here, $\tan\phi = \text{opp}/\text{adj} = -h$; thus, $\phi = \arctan(-h)$; that is, $\phi = \arctan(-\omega/\omega_1)$.

* Doing this will show, separately, the real and imaginary components of $\bar{V}_o$.

If, now, in Fig. 192 we replace A and B with their values defined in connection with eq. (324) then Fig. 192 becomes Fig. 193, which shows that the output voltage vector $\bar{V}_o$ LAGS the input reference voltage by the angular amount of

$$\phi = \arctan(-\omega/\omega_1) = -\arctan(\omega/\omega_1) \tag{325}$$

Your calculator, applied to the above equation, should produce the following table of values.

ω/ω_1	ϕ°	ω/ω_1	ϕ°	ω/ω_1	ϕ°
0.01	−0.6	0.2	−11.3	4	−76.0
0.02	−1.2	0.4	−21.8	6	−80.5
0.04	−2.3	0.6	−31.0	8	−82.9
0.06	−3.4	0.8	−38.7	10	−84.3
0.08	−4.6	1.0	−45.0	50	−88.9
0.10	−5.7	2.0	−63.4	100	−89.4

Inspection of the table brings out the following facts concerning the PHASE-SHIFT characteristics of the basic low-pass network of Fig. 190.

1. The phase angle ϕ is negative, meaning that the output voltage LAGS the input voltage. For this reason, Fig. 190 is called a ‘‘lag’’ network in control system terminology.
2. It has been shown that, if a network is to produce *no phase distortion*, ϕ must be proportional to ω; that is, the *ratio* of ϕ to ω must be *constant*, $\phi/\omega = k$ (note 20 in the Appendix). The above table of values shows that this requirement is very nearly satisfied, in the case of Fig. 190, *for low frequencies*. To see that this is true, let us set $\omega_1 = 1$ and, using the above table of values, construct the following table, where ϕ is in degrees and the ratio ϕ/ω is rounded off to the nearest whole number.

ϕ	ω	ϕ/ω	ϕ	ω	ϕ/ω
−0.6	0.01	−60	−31.0	0.60	−52
−1.2	0.02	−60	−38.7	0.80	−48
−2.3	0.04	−58	−45.0	1.00	−45
−3.4	0.06	−57	−63.4	2.00	−32
−4.6	0.08	−58	−76.0	4.00	−19
−5.7	0.10	−57	−80.5	6.00	−13
−11.3	0.20	−57	−84.3	10.00	−8
−21.8	0.40	−55	−88.9	50.00	−2

The table plainly shows that, *for practical purposes*, the ratio ϕ/ω is constant for all *low frequencies*, up to $\omega = 0.04$; that is, since we're using $\omega_1 = 1$ here, up to the value of $\omega/\omega_1 = 0.04$. Actually, for *most* practical purposes, we can say that ϕ/ω is constant up to the value $\omega/\omega_1 = 1$; that is, to $\omega = \omega_1$, where ω_1 is the *half-power* frequency. Thus our CONCLUSIONS regarding Fig. 190 can be summarized as follows.

Figure 190 is a low-pass network having, for practical purposes, no amplitude or phase discrimination for frequencies LESS than the half-power frequency, but increasing amounts of such discrimination as the frequency increases beyond the half-power frequency.

Problem 167

In Fig. 190, given that $R = 10{,}000$ ohms and $C = 0.25\,\mu\text{F}$, if $\bar{V}_i = V_i = 10$ volts rms, find the amplitude and phase of the output voltage $\bar{V}_o$ if (a) the frequency is 30 Hz, (b) the frequency is increased to 300 Hz.

Problem 168

In the above problem (same values of R and C), suppose an ideal square wave of voltage, having a steady frequency of 30 square waves per second, is applied to the input. Explain why the output voltage waveform will not also be an ideal square wave.

CASE II: BASIC HIGH-PASS *RC* NETWORK

This is a series RC circuit in which *the output voltage appears across the resistor* R, as shown in Fig. 194.

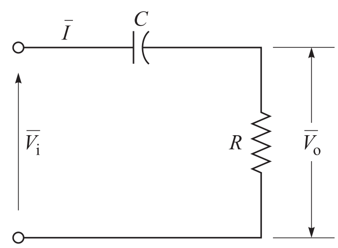


Fig. 194

In the figure, it is given that the AMPLITUDE of the reference input signal, $\bar{V}_i = V_i\underline{/0^\circ} = V_i$, is to remain constant, while the FREQUENCY is allowed to increase from a very low value to progressively higher values. From the figure, note that as the frequency of V_i *increases* the reactance of C *decreases*, and thus *as the* FREQUENCY INCREASES *the* OUTPUT VOLTAGE ACROSS R INCREASES; in other words, the higher the frequency, the higher is the output voltage. Thus the circuit of Fig. 194 discriminates against the lower frequencies and is therefore a "high-pass" type of network. The algebra for Fig. 194 parallels that for Fig. 190, as follows.

First, by Ohm's law,

$$\bar{V}_o = R\bar{I} = \frac{RV_i}{R - jX_C}$$

Now set $-jX_C = 1/j\omega C$; doing this, then multiplying the numerator and denominator by $j\omega C$, we have

$$\frac{\bar{V}_o}{V_i} = \bar{G} = \frac{j\omega RC}{1 + j\omega RC} \tag{326}$$

which corresponds to eq. (316). Now, again setting $RC = 1/\omega_1$, as in eq. (317), the last equation becomes

$$\bar{G} = \frac{j(\omega/\omega_1)}{1 + j(\omega/\omega_1)} = \frac{jh}{1 + jh} \tag{327}$$

where $h = (\omega/\omega_1)$.

Equation (327) contains all the information concerning both the amplitude and phase response of the network of Fig. 194. We'll first investigate the amplitude response, as follows.

The AMPLITUDE response is determined by the MAGNITUDE of eq. (327); thus, recalling that if A and B are complex numbers, $|A/B| = |A|/|B|$, we have that

$$|\bar{G}| = \frac{h}{\sqrt{1+h^2}} = h(1+h^2)^{-1/2} \tag{328}$$

which corresponds to eq. (319). Now, to express the amplitude response in *decibels*, let us, in the manner of eq. (320), write that

$$\mathrm{dB} = 20\log[h(1+h^2)^{-1/2}]$$

Now make use of the fact that $\log XY = \log X + \log Y$,* and that $\log X^n = n \log X$.† Doing this, the last equation becomes

$$\mathrm{dB} = 20\log(\omega/\omega_1) - 10\log[1+(\omega/\omega_1)^2] \tag{329}$$

which expresses the ratio of $\bar{V}_o/V_i$ in Fig. 194 in terms of decibels. Using your calculator, you can quickly verify that the following table of values is correct for eq. (329).

ω/ω_1	dB gain	ω/ω_1	dB gain	ω/ω_1	dB gain
0.02	−34.0	0.2	−14.2	2.0	−0.97
0.04	−28.0	0.4	−8.6	4.0	−0.26
0.06	−24.5	0.6	−5.8	6.0	−0.12
0.08	−22.0	0.8	−4.1	8.0	−0.07
0.10	−20.0	1.0	−3.0	10.0	−0.04

A plot of the above data gives Fig. 195, the curve of dB gain versus frequency ratio.

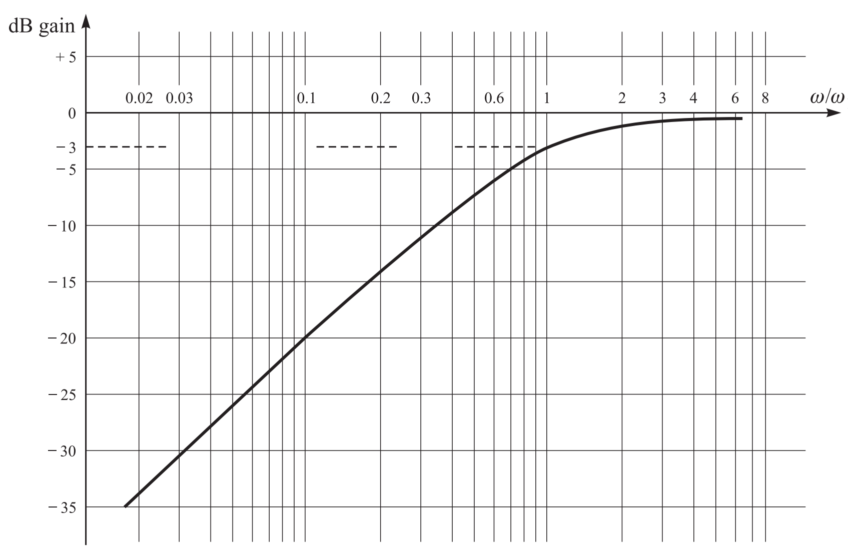


Fig. 195

* See note 22 in Appendix.

† See eq. (9-A) in note 19 in Appendix.

Figure 195 verifies that the circuit of Fig. 194 is a *high-pass network* because the higher the frequency, the less is the attenuation of the input signal as it appears at the output terminals.

In the discussion following Fig. 191 the term "half-power frequency" was introduced as being the frequency at which the power gain of a network is *down 3 decibels* from its maximum or reference value. Note, now, that setting $\omega = \omega_1$ in eq. (329) gives the value

$$\text{dB} = 20\log\ 1 - 10\log 2 = -3 \text{ decibels}$$

showing that *the value of* ω at the half-power frequency in Fig. 194 is $\omega = \omega_1 = 1/RC$ (by eq. (317)).

Problem 169
Suppose, in Fig. 194, that $R = 1200$ ohms. Find the required value of C if the half-power frequency is to be 2.2 kHz.

Problem 170
(a) In problem 169, at what frequency, in hertz, will the power gain be -6 decibels? (*Answer*: 1274 Hz)
(b) At what frequency will the power gain be -2 decibels? (*Answer*: 2877 Hz)

Next, to investigate the PHASE RESPONSE of Fig. 194 we'll basically repeat the procedure for Fig. 190, as follows. First, by eq. (322), $\bar{V}_o = \bar{G}V_i$, which, after rationalizing the value of $\bar{G}$ given by eq. (327), becomes

$$\bar{V}_o = \left[\frac{h^2}{1+h^2} + j\,\frac{h}{1+h^2}\right]V_i = (A + jB)V_i \tag{330}$$

showing that $\bar{V}_o$ is the *vector sum* of the voltage drops AV_i and jBV_i, as shown in Fig. 196.

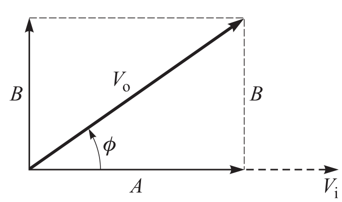


Fig. 196

For convenience in the above, let $d = (1 + h^2)$. Then, by inspection, we have

$$\tan\phi = B/A = (h/d)/(h^2/d) = 1/h$$

hence,

$$\phi = \arctan(1/h)$$

or, since $h = \omega/\omega_1$ and $\omega_1 = 1/RC$,

$$\phi = \arctan(\omega_1/\omega) = \arctan(1/\omega RC) \tag{331}$$

or, since $\omega_1 = 1/RC$ is constant (in any given case), let us, for convenience here, set $\omega_1 = 1$. Then eq. (331) becomes

$$\phi = \arctan(1/\omega) \tag{332}$$

We have said, correctly, that if a network produces phase shift, then, in order that there be zero phase-shift distortion, *the phase shift must be proportional to frequency* ω. Investigation of eq. (332) shows, however, that phase shift is *not* proportional to frequency, neither at low frequencies nor at high frequencies. Offhand, this would seem to mean that the high-pass filter of Fig. 194 would produce an unacceptable amount of phase distortion, even at the desired higher frequencies.

The point to notice, however, is that *for high frequencies* the phase shift approaches *zero* degrees, and thus the time delay approaches *zero* as we go to the higher frequencies (eq. (13-A), note 20, Appendix). Thus the criterion for zero phase-shift distortion is that either the network should produce a negligibly small amount of phase shift or, if a sizeable amount of phase shift is produced, then such phase shift must be proportional to frequency.

CASES III AND IV: BASIC LOW-PASS AND HIGH-PASS *RL* FILTERS

The two basic filters are shown in Figs. 197 and 198.

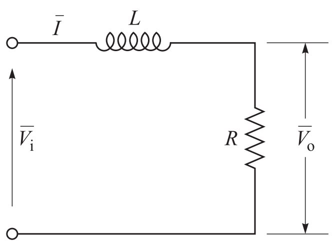


Fig. 197. Basic low-pass filter.

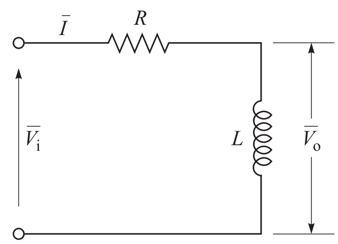


Fig. 198. Basic high-pass filter.

In the figures, R is resistance in ohms and L is inductance in henrys, so that $X_L = \omega L$ is the inductive reactance in ohms. We'll take $\bar{V}_i$ to be the input voltage reference vector and write that $\bar{V}_i = V_i\underline{/0^\circ} = V_i$. The *amplitude* of $\bar{V}_i$ is to remain fixed in value while we allow its frequency ω to increase slowly from a very low value to a very high value. The vector output voltage is denoted by $\bar{V}_o$, as shown. As previously, we'll let $\bar{G}$ denote the ratio of output to input voltage, $\bar{G} = \bar{V}_o/V_i$.

The operation of both circuits depends upon the fact that the value of the reactance ωL, increases as the frequency increases, and thus (since V_i is constant) the voltage drop across R decreases as the frequency increases. Thus, in Fig. 197 the *higher the frequency* the *lower* the value of $\bar{V}_o$, while in Fig. 198 the higher the frequency the *greater* the value of $\bar{V}_o$.

Note that, in both cases, we have, by Ohm's law, that

$$\bar{I} = \frac{V_i}{R + j\omega L} \tag{333}$$

and upon making use of this fact we find, for the LOW-PASS case of Fig. 197, that

$$|\bar{G}| = \frac{R}{\sqrt{R^2 + (\omega L)^2}} \tag{334}$$

and

$$\phi = -\arctan(\omega L/R) \tag{335}$$

and for the HIGH-PASS case of Fig. 198, that

$$|\bar{G}| = \frac{\omega L}{\sqrt{R^2 + (\omega L)^2}} \tag{336}$$

and

$$\phi = \arctan(R/\omega L) \tag{337}$$

Problem 171
Prove that eqs. (334) and (335) are correct.

Problem 172
Prove that eqs. (336) and (337) are correct.

Problem 173
Show that the "half-power" frequency in Fig. 197 is R/L radians/second.

9.6 The Symmetrical T Network. Characteristic Impedance

The *RC* and *RL* filters of section 9.5 have the advantage of simplicity, but they have the two disadvantages of having *high loss of signal power* and *poor frequency response characteristics*. This is illustrated in Fig. 199, which shows the general form of the curve of the relative magnitude of $\bar{G}$ versus frequency for the basic *RC* and *RL* networks of Figs. 190 and 197 in section 9.5. As already defined, $|\bar{G}| = |\bar{V}_o/V_i|$, and ω_0 is the half-power frequency, that is, the frequency at which the voltage ratio $|\bar{G}|$ is reduced to 70.7% of its maximum value* (here the maximum value of $|\bar{G}|$ is at $\omega = 0$).

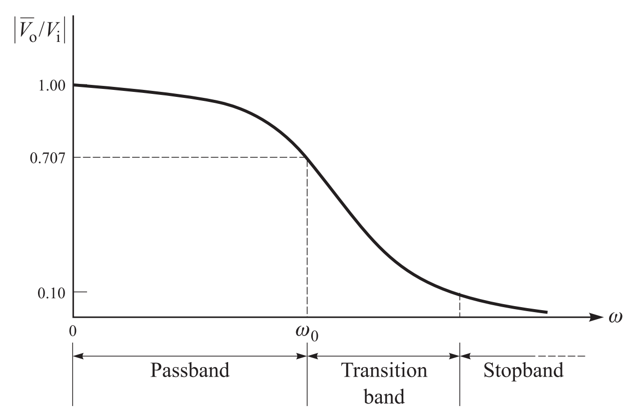


Fig. 199

* From section 9.5, a "half-power" frequency is a frequency at which the output power of a network is reduced to one-half its maximum value, being equivalent to −3 decibels. Or, in terms of VOLTAGE RATIO, by eq. (320), $\text{dB} = 20 \log 0.707 = -3$ decibels.

Inspection of the figure shows that $|\bar{V}_o| = V_i$ for $\omega = 0$; that is, $|\bar{V}_o/V_i| = 1$ for $\omega = 0$. Then, as ω is increased in value (V_i always remaining constant), the output $|\bar{V}_o|$ decreases in value until the "end of the passband" is reached, which is generally taken to be the frequency at which $|\bar{V}_o/V_i| = 0.707$; this is the frequency at which the output is "down 3 decibels" from its maximum value.

Next, as ω is further increased in value, we enter the "transition band," as shown in the figure. The high-frequency end of the transition band can be taken as the frequency at which $|\bar{V}_o/V_i| = 0.1$, as shown above, which is the frequency at which the output is "down 20 decibels" from its maximum value. Beyond the transition band we enter the "stopband," as shown. We assume that, for practical purposes, the output of the filter can be considered to be negligibly small for all frequencies in the stopband.

We must remember that a "low-pass" filter is used in applications where it is desired to pass all signal components *below* a specified frequency and *reject* all components above that frequency. It therefore follows that the *width of the transition band* is especially critical, and should be *as narrow as possible* so as to prevent, as much as possible, the passage of unwanted higher frequency components through the filter.

As you would expect, it's possible to design low-pass filters having frequency response curves GREATLY SUPERIOR to those of the simple *RC* and *RL* filters illustrated in Fig. 199. To do this requires the use of both inductance and capacitance (*L* and *C*), arranged in various ways, making use of the phenomenon of series and parallel resonance. One such arrangement is known as the "constant-*k*" type of filter, and is the subject of the next two sections. Before getting into these sections, however, we must first take up the "symmetrical T" network, as follows.

Let us begin by referring back to the standard "T" network notation of Fig. 174 in section 9.2. Let us then define that a "symmetrical" T network is a T network in which $Z_1 = Z_2$. Thus, Fig. 174 could be redrawn to represent a symmetrical T as shown in Fig. 200.

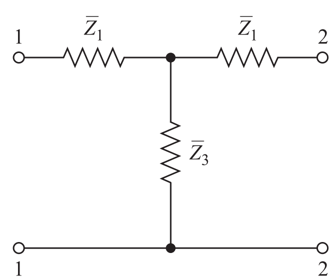


Fig. 200

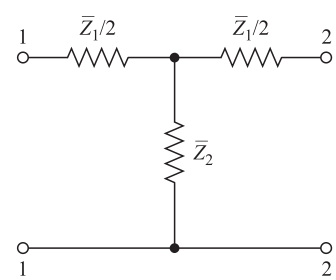


Fig. 201

Actually, however, the notation of Fig. 200 is *not* used in the SPECIAL CASE of the symmetrical T network; instead, it has become the custom to denote each of the series elements by $\bar{Z}_1/2$ and the shunt element by $\bar{Z}_2$, as shown in Fig. 201. This will cause no difficulties so long as we remember that the special notation in Fig. 201 is to be used ONLY for the case of the symmetrical T network.

In Fig. 201, terminals (1, 1) are given to be the INPUT terminals and (2, 2) the OUTPUT terminals. Now suppose a load impedance $\bar{Z}_L$ is connected to the output terminals in Fig. 201; doing this will cause a certain value of INPUT IMPEDANCE, $\bar{Z}_{in}$, to appear at the input terminals as shown in Fig. 202.

Note that, looking into terminals (1, 1) in Fig. 202, we see a series-parallel circuit consisting of $\bar{Z}_1/2$ *in series* with the parallel combination of $\bar{Z}_2$ and $(\bar{Z}_1/2 + \bar{Z}_L)$. Thus

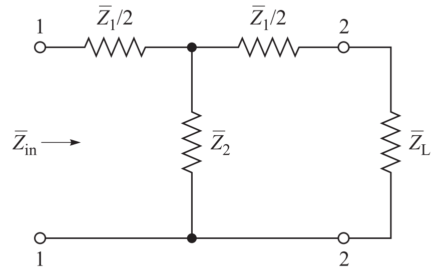


Fig. 202

the *input impedance* seen looking into terminals (1, 1) in Fig. 202 is

$$\bar{Z}_{in} = \frac{\bar{Z}_1}{2} + \frac{\bar{Z}_2\left(\dfrac{\bar{Z}_1}{2} + \bar{Z}_L\right)}{\bar{Z}_2 + \dfrac{\bar{Z}_1}{2} + \bar{Z}_L} \tag{338}$$

Now suppose that $\bar{Z}_1$ and $\bar{Z}_2$ are chosen to have the particular values that will make the INPUT IMPEDANCE EQUAL TO THE LOAD IMPEDANCE. To *find* the required values of $\bar{Z}_1$ and $\bar{Z}_2$ needed to accomplish this, all we need do is replace $\bar{Z}_{in}$ with $\bar{Z}_L$ in eq. (338) and then solve for $\bar{Z}_L$. In doing this, however, it has become customary to replace both $\bar{Z}_{in}$ and $\bar{Z}_L$ with the single symbol "$\bar{Z}_0$" ; thus, setting $\bar{Z}_{in} = \bar{Z}_L = \bar{Z}_0$ in eq. (338) we have that

$$\bar{Z}_0 = \frac{\bar{Z}_1}{2} + \frac{\bar{Z}_2\left(\dfrac{\bar{Z}_1}{2} + \bar{Z}_0\right)}{\bar{Z}_2 + \dfrac{\bar{Z}_1}{2} + \bar{Z}_0} \tag{339}$$

Thus, for the SPECIAL CONDITION represented by eq. (339), $\bar{Z}_{in} = \bar{Z}_L = \bar{Z}_0$, Fig. 202 becomes Fig. 203.

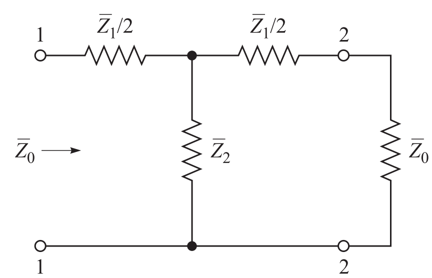


Fig. 203

Having the above condition (input impedance equal to load impedance) has important advantages that we'll point out later on. To be able to produce this condition, however, we must find the relationship that must exist between $\bar{Z}_0$ and the two impedances $\bar{Z}_1$ and $\bar{Z}_2$, which can be done by *solving eq. (339) for* $\bar{Z}_0$, as follows.

First, multiplying both sides of eq. (339) by the *denominator* of the right-hand fraction puts the equation in the form

$$Z_0\left(Z_2 + \frac{Z_1}{2} + Z_0\right) = \frac{Z_1}{2}\left(Z_2 + \frac{Z_1}{2} + Z_0\right) + Z_2\left(\frac{Z_1}{2} + Z_0\right)$$

Now multiplying as indicated, then collecting like terms, you should find that

$$\bar{Z}_0^2 = \bar{Z}_1\bar{Z}_2 + \frac{\bar{Z}_1^2}{4}$$

thus we have

$$\bar{Z}_0 = \sqrt{\bar{Z}_1\bar{Z}_2 + \bar{Z}_1^2/4} \tag{340}$$

The impedance $\bar{Z}_0$ given by eq. (340) is called the "characteristic impedance" of a symmetrical T network. Or, if such a network already exists, the value of $\bar{Z}_0$ can be found experimentally by making "open circuit" and "short circuit" measurements, as indicated in Figs. 204 and 205.

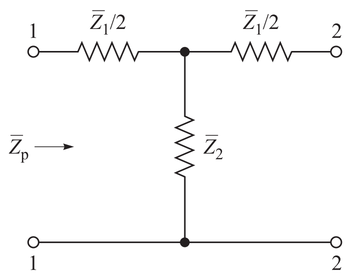


Fig. 204

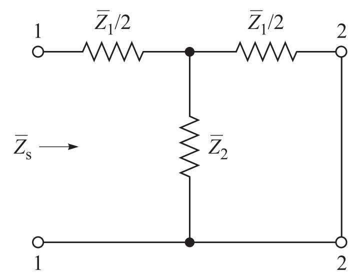


Fig. 205

In Fig. 204, $\bar{Z}_p =$ impedance looking into $(1, 1)$ with $(2, 2)$ OPEN-CIRCUITED, whereas in Fig. 205, $\bar{Z}_s =$ impedance looking into $(1, 1)$ with $(2, 2)$ SHORT-CIRCUITED.

By inspection, note that

$$\bar{Z}_p = \bar{Z}_1/2 + \bar{Z}_2$$

and

$$\bar{Z}_s = \frac{\bar{Z}_1}{2} + \frac{\bar{Z}_1\bar{Z}_2/2}{\bar{Z}_1/2 + \bar{Z}_2} = \frac{Z_1}{2} + \frac{\bar{Z}_1\bar{Z}_2}{\bar{Z}_1 + 2\bar{Z}_2}$$

Now, using these values of $\bar{Z}_p$ and $\bar{Z}_s$, a careful multiplication will show that

$$\bar{Z}_p\bar{Z}_s = \frac{\bar{Z}_1^2}{4} + \bar{Z}_1\bar{Z}_2$$

hence, upon making use of the relationship given just prior to eq. (340), we have that

$$\bar{Z}_0 = \sqrt{\bar{Z}_p\bar{Z}_s} \tag{341}$$

Equation (341) is important because it provides a way to find, by actual laboratory measurement, the characteristic impedance of a network known to be of the symmetrical T form.

In using eqs. (340) and (341) we must remember that the *addition and subtraction* of complex numbers can be performed only in the *rectangular* form. On the other hand, to raise a complex number to a *fractional power* the number must be expressed in the "trigonometric," "polar," or "exponential" form (sections 6.6 and 6.7).

Problem 174

Find the characteristic impedance of a low-pass symmetrical T network in which

$\bar{Z}_1 = 2 + j5$ and $\bar{Z}_2 = 4 - j3$. (*Answer*: $(4.68 + j2.03)$ ohms, approx.)

Problem 175
If the open-circuit and short-circuit measurements on a certain symmetrical T network give magnitudes of 15 and 25 ohms and angular displacements of 35 and −75 degrees, find the characteristic impedance of the network.

In the foregoing discussion we should remember that the value of an impedance $\bar{Z}$ (assuming given values of R, L, and C) depends upon the value of the *frequency* ω. Thus the value of $\bar{Z}_0$ given by eq. (340) depends upon the value of ω, and is different for each value of ω.

The difference between Figs. 202 and 203 is that Fig. 203 is for the particular case for which

$$\bar{Z}_L = \bar{Z}_0 = \sqrt{\bar{Z}_1\bar{Z}_2 + \bar{Z}_1^2/4}$$

which can be exactly true, in general, for only one frequency, because $\bar{Z}_L$ and $\bar{Z}_0$ will be represented by different mathematical equations. This effect will be considered later on in our treatment of the constant-k filter.

Problem 176
In Fig. 206, the values are in henrys, farads, and ohms. Find the frequency, in rad/sec, at which $Z_{in} = 2$ ohms. (*Answer*: 173.21 rad/sec)

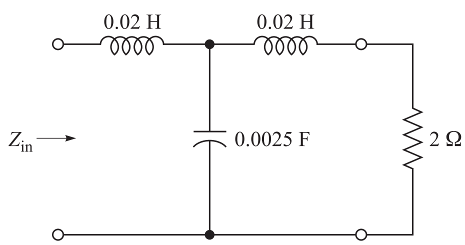


Fig. 206

Let us continue with the symmetrical T of Fig. 203, which depicts the particular condition in which $\bar{Z}_{in} = \bar{Z}_L = \bar{Z}_0$. As we know, eq. (340) applies to this condition and, in addition to eq. (340), several other useful relationships exist. As an aid in finding these relations let us begin by redrawing Fig. 203 as Fig. 207.

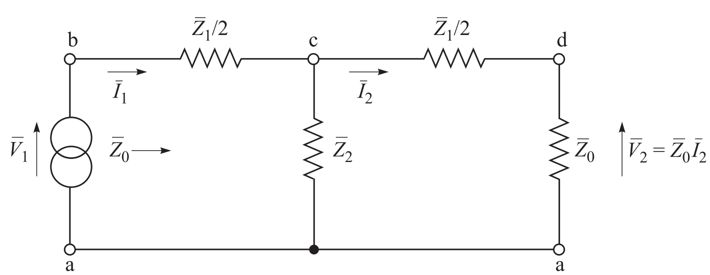


Fig. 207

By Kirchhoff's voltage law, the sum of the *voltage drops* AROUND ANY CLOSED PATH IN A NETWORK is equal to the sum of the *generator voltages* in that path. Thus, in Fig. 207, around the closed path abcda we have that

$$\bar{V}_1 = \bar{Z}_1\bar{I}_1/2 + \bar{Z}_1\bar{I}_2/2 + \bar{Z}_0\bar{I}_2$$

or, since $\bar{V}_1 = \bar{I}_1\bar{Z}_0$,

$$\bar{Z}_0\bar{I}_1 = \bar{Z}_1\bar{I}_1/2 + \bar{Z}_1\bar{I}_2/2 + \bar{Z}_0\bar{I}_2$$

from which we have the important relationship

$$\frac{\bar{I}_1}{\bar{I}_2} = \frac{\bar{Z}_0 + \dfrac{\bar{Z}_1}{2}}{\bar{Z}_0 - \dfrac{\bar{Z}_1}{2}} \tag{342}$$

Problem 177
Find the ratio of $\bar{I}_1$ to $\bar{I}_2$ in Fig. 206 for the frequency at which $\bar{Z}_{\text{in}} = 2$ ohms.
(*Answer*: $-0.500 + j0.866$)

As you can see from the above, eq. (342) was derived for the particular frequency for which $\bar{Z}_{\text{in}} = \bar{Z}_{\text{L}} = \bar{Z}_0$. For any frequency in general, eq. (342) becomes

$$\frac{\bar{I}_1}{\bar{I}_2} = \frac{\bar{Z}_{\text{L}} + \bar{Z}_1/2}{\bar{Z}_{\text{in}} - \bar{Z}_1/2} \tag{343}$$

in which $\bar{Z}_{\text{L}}$ will be a given load impedance and $\bar{Z}_{\text{in}}$ can be found by eq. (338).

Problem 178
Derive eq. (343), using the same procedure as in deriving eq. (342).

Problem 179
In Fig. 206, find $\bar{I}_1/\bar{I}_2$ for $\omega = 200$ rad/sec. (*Answer*: $-1 + j$)

Problem 180
If the T network in problem 174 is terminated in its characteristic impedance, find the magnitude and phase angle of $\bar{I}_1$ relative to $\bar{I}_2$.

It's important to note, here, that the equation for the current ratio can also be written in the following form, in which $\bar{Z}_{\text{L}}$ denotes ANY VALUE of load impedance:

$$\frac{\bar{I}_1}{\bar{I}_2} = 1 + \frac{\bar{Z}_1}{2\bar{Z}_2} + \frac{\bar{Z}_{\text{L}}}{\bar{Z}_2} \tag{344*}$$

Or, setting $\bar{Z}_{\text{L}} = \bar{Z}_0$, then making use of eq. (340), and noting that

$$\frac{1}{\bar{Z}_2}\sqrt{\bar{Z}_1 Z_2 + \frac{\bar{Z}_1^2}{4}} = \sqrt{\frac{\bar{Z}_1}{\bar{Z}_2} + \frac{\bar{Z}_1^2}{4\bar{Z}_2^2}}$$

eq. (344) becomes, for the special case in which $\bar{Z}_{\text{L}} = \bar{Z}_0$,

$$\frac{\bar{I}_1}{\bar{I}_2} = 1 + \frac{\bar{Z}_1}{2\bar{Z}_2} + \sqrt{\frac{\bar{Z}_1}{\bar{Z}_2} + \left[\frac{\bar{Z}_1}{2\bar{Z}_2}\right]^2} \tag{345}$$

* See note 23 in Appendix.

You'll note that the last four equations give the value of the CURRENT ratio $\bar{I}_1/\bar{I}_2$. If, however, we wish to work in terms of VOLTAGE ratio, this can be done by noting that $\bar{V}_1 = \bar{I}_1\bar{Z}_{in}$ and $\bar{V}_2 = \bar{I}_2\bar{Z}_L$; hence for any value of $\bar{Z}_L$,

$$\frac{\bar{V}_1}{\bar{V}_2} = \frac{\bar{Z}_{in}}{\bar{Z}_L}\frac{\bar{I}_1}{\bar{I}_2} \tag{346}$$

or, for $\bar{Z}_{in} = \bar{Z}_L = \bar{Z}_0$,

$$\frac{\bar{V}_1}{\bar{V}_2} = \frac{\bar{I}_1}{\bar{I}_2} \tag{347}$$

Thus, as eq. (347) shows, at the particular frequency for which a symmetrical T network is terminated in its characteristic impedance, the voltage ratio is EQUAL to the current ratio.

9.7 Low-Pass Constant-*k* Filter

In this section we study a form of symmetrical T network in which $\bar{Z}_1$ and $\bar{Z}_2$ will be given to be PURE REACTANCES of opposite sign, where (using the terminology of Fig. 202)

$$\bar{Z}_1 = jX_L = j\omega L \quad \text{and} \quad \bar{Z}_2 = -jX_C = 1/j\omega C$$

and where the network is terminated in a "pure resistance" of $\bar{Z}_L = R_L$, as shown in Fig. 208.

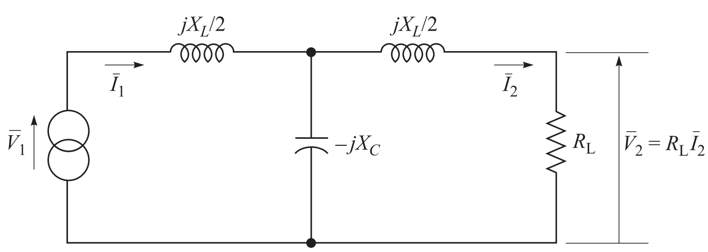


Fig. 208

Note that at VERY LOW frequencies X_L is VERY LOW while X_C is VERY GREAT, so that $\bar{V}_2$ is practically equal to $\bar{V}_1$ at such frequencies ($\bar{V}_2$ is actually EQUAL to $\bar{V}_1$ at $\omega = 0$).

On the other hand, at very HIGH frequencies X_L is VERY GREAT and X_C is very LOW, so that $\bar{V}_2$ is practically equal to zero at such frequencies. Thus, in just a general way, we see that Fig. 208 constitutes a "low-pass" type of network. Such a general observation is, of course, not sufficient for engineering purposes; to get specific information let us now apply the algebra of complex numbers to Fig. 208, as follows (using the notation of Fig. 202).

From note 23 in the Appendix,

$$\bar{I}_2 = \frac{\bar{V}_1\bar{Z}_2}{\Delta} = \frac{\bar{V}_1\bar{Z}_2}{\dfrac{\bar{Z}_1^2}{4} + \bar{Z}_1\bar{Z}_2 + \left(\dfrac{\bar{Z}_1}{2} + \bar{Z}_2\right)\bar{Z}_L}$$

hence, by Ohm's law,

$$\bar{V}_2 = \bar{I}_2\bar{Z}_\mathrm{L} = \frac{\bar{V}_1\bar{Z}_2\bar{Z}_\mathrm{L}}{\dfrac{\bar{Z}_1^2}{4} + \bar{Z}_1\bar{Z}_2 + \left(\dfrac{\bar{Z}_1}{2} + \bar{Z}_2\right)\bar{Z}_\mathrm{L}}$$

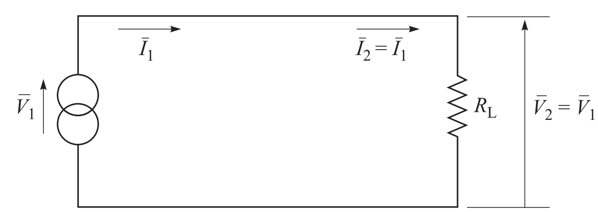
Now make the substitutions $\bar{Z}_1 = jX_L$, $\bar{Z}_2 = -jX_C$, and $\bar{Z}_\mathrm{L} = R_\mathrm{L}$; upon doing this, then multiplying the numerator and denominator by j, you should find that (taking $\bar{V}_1 = V_1\underline{/0^\circ} = V_1$ as reference vector),

$$\frac{\bar{V}_2}{V_1} = \frac{X_C R_\mathrm{L}}{\left(X_C - \dfrac{X_L}{2}\right)R_\mathrm{L} + jX_L\left(X_C - \dfrac{X_L}{4}\right)}$$

Now, on the right-hand of the above equation, multiply the numerator and denominator by $1/X_C R_\mathrm{L}$; doing this, then making the substitutions $X_L = \omega L$ and $X_C = 1/\omega C$, you should find that

$$\frac{\bar{V}_2}{V_1} = \frac{1}{\left(1 - \dfrac{\omega^2 LC}{2}\right) + j\,\dfrac{\omega L}{R_\mathrm{L}}\left(1 - \dfrac{\omega^2 LC}{4}\right)} \tag{348}$$

At this point let's pause and try to decide upon a reasonable value for R_L in Fig. 208. To do this, let us begin by noting that at ZERO frequency ($\omega = 0$) Fig. 208 would become



Thus, for $\omega = 0$ the generator would see a pure resistance of R_L ohms, with $\bar{V}_2/\bar{V}_1 = 1$ and with zero phase shift between $\bar{V}_2$ and $\bar{V}_1$, *which would be the desired condition here* because Fig. 208 is to be a *low-pass* filter passing, as uniformly as possible, all frequencies from $\omega = 0$ to whatever the cut-off frequency is to be. With this in mind, and upon setting $\bar{Z}_1 = j\omega L$ and $\bar{Z}_2 = 1/j\omega C$ in eq. (340), we have that the CHARACTERISTIC IMPEDANCE of the T-network in Fig. 208 is equal to

$$\bar{Z}_0 = \sqrt{\frac{L}{C} - \frac{\omega^2 L^2}{4}} \tag{349}$$

We have agreed, however, that the generator will see the actual value of R_L for $\omega = 0$; thus, upon setting $\omega = 0$ in eq. 349 we have that

$$\bar{Z}_0 = R_\mathrm{L} = \sqrt{\frac{L}{C}} \tag{350}$$

which is the resistance the generator will see only at $\omega = 0$, and is thus the actual value of R_L that will be used in Fig. 208.

Problem 181

Suppose, in Fig. 208, that the inductance of each coil is 800 microhenrys and that $C = 0.04$ microfarads.

(a) What value of R_L should be used?
(b) At what frequency will the characteristic impedance be equal to zero?
(c) What value of impedance will the generator see at the frequency at which $Z_0 = 0$?
(d) Why doesn't the generator see zero impedance in part (c)?

Thus, since $\dfrac{1}{R_L} = \sqrt{\dfrac{C}{L}}$, we have that

$$\frac{\omega L}{R_L} = \omega L\sqrt{\frac{C}{L}} = \omega\sqrt{\frac{L^2C}{L}} = \omega\sqrt{LC}$$

hence eq. (348) becomes

$$\frac{\bar{V}_2}{V_1} = \frac{1}{(1-\omega^2 LC/2) + j\omega\sqrt{LC}(1-\omega^2 LC/4)} \tag{351}$$

which has the form

$$\frac{\bar{V}_2}{V_1} = \frac{1}{A+jB}$$

In the above the variable is frequency, ω radians/second, in which ω denotes ANY FREQUENCY in general. It will be helpful, however, in understanding what eq. (351) says, if the equation is written in terms of the *ratio* of ω to *some* PARTICULAR VALUE *of* ω that we'll denote by ω_0. In doing this, ω_0 can be any particular value we choose, but in the case of eq. (351) one convenient way would be to set

$$\omega_0^2 = \frac{4}{LC}$$

that is,

$$LC = \frac{4}{\omega_0^2} \tag{352}$$

and if you now substitute the above value of LC into eq. (351) you should find that

$$\frac{\bar{V}_2}{V_1} = \frac{1}{[1-2(\omega/\omega_0)^2] + j2(\omega/\omega_0)[1-(\omega/\omega_0)^2]} \tag{353}$$

so that the ratio $\bar{V}_2/V_1$ is now expressed in terms of the value of ω relative to the fixed value $\omega_0 = 2/\sqrt{LC}$, and we say that the equation is ''normalized'' relative to ω_0. The advantage of doing this is that now we can investigate eq. (351) in general terms without having to carry along the values of L and C. As a further step let us make the substitution

$$h = (\omega/\omega_0) \tag{354}$$

thus eq. (352) becomes

$$\frac{\bar{V}_2}{V_1} = \frac{1}{(1-2h^2) + j2h(1-h^2)} = \frac{1}{A+jB} \tag{355}$$

Now let it be given that we wish to investigate only the manner in which the *magnitude* of $\bar{V}_2/V_1$ varies with h; in which case you can verify that eq. (355) becomes

$$\left|\frac{\bar{V}_2}{V_1}\right| = \frac{1}{\sqrt{A^2 + B^2}} = \frac{1}{\sqrt{1 + 4(h^6 - h^4)}} \tag{356}$$

thus,

$$\left|\frac{\bar{V}_2}{V_1}\right| = [1 + 4(h^6 - h^4)]^{-1/2} \tag{357}$$

which in terms of *decibels* becomes (see eqs. (319), (320), and (321))

$$\text{dB} = -10\log[1 + 4(h^6 - h^4)] \tag{358}$$

Now, using your calculator, you can verify that the following "table of values" is correct for eq. (358), in which we've rounded off dB values to two decimal places.

h	dB	h	dB
0.1	0.00	1.0	0.00
0.2	0.03	1.2	−6.67
0.4	0.39	1.4	−11.97
0.6	1.75	1.6	−16.22
0.8	3.87	1.8	−19.73
0.9	3.00	2.0	−22.86
		2.2	−25.57

A plot of the above results on semi-log paper is given in Fig. 209, with a brief discussion following.

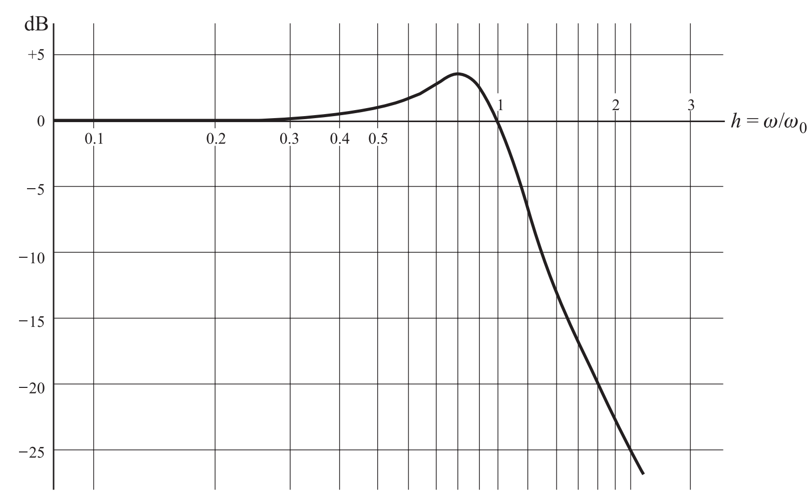


Fig. 209

Thus the gain of the low-pass network of Fig. 208 is practically constant from $h = 0$ to $h = 0.3$; then, as h increases, the gain rises to a maximum value of 3.87 dB for $h = 0.8$, after which the gain rapidly decreases as h increases in value. The rise in gain from $h = 0.3$ to

$h = 1.0$ is due to a resonant condition that comes into play between L and C, the effect peaking, very approximately, at $h = 0.8$. Then, as h increases in value beyond $h = 1.0$, the condition of resonance is lost and the gain begins to rapidly decrease. It is thus the presence of *both* inductance and capacitance, L and C, having the possibility of resonance, that causes the curve of Fig. 209 to be so much improved over the simple curve of Fig. 191.

As inspection of Fig. 209 shows that the gain falls off rapidly for values of h greater than 1, and for that reason $h = 1$ is sometimes taken to be the "cut-off" condition for Fig. 208. Since $h = \omega/\omega_0$, it follows that $h = 1$ when $\omega = \omega_0$; thus the cut-off frequency, ω_c, can be taken to be equal to ω_0 and hence, by eq. (352),

$$\omega_c = \frac{2}{\sqrt{LC}} \tag{359}$$

Lastly, we should mention that the term "constant-k" is applied to Fig. 208 because the values of $\bar{Z}_1$ and $\bar{Z}_2$ are such that their product has a constant value, independent of frequency; thus

$$\bar{Z}_1\bar{Z}_2 = (j\omega L)(1/j\omega C) = L/C$$

Problem 182
In Fig. 208, suppose $C = 0.015\,\mu\text{F}$. If a cut-off frequency of 100,000 Hz is desired, what must be the inductance of each coil? (*Answer*: 337.7 μH)

Problem 183
Using eq. (355), show that the phase shift of $\bar{V}_2$ with respect to V_1 is equal to

$$\phi = -\arctan[2h(1 - h^2)/(1 - 2h^2)]$$

9.8 High-Pass Constant-*k* Filter

Here again we study a form of symmetrical T network in which $\bar{Z}_1$ and $\bar{Z}_2$ are pure reactances of opposite sign where, using the terminology of Fig. 202,

$$\bar{Z}_1 = -jX_C = -j/\omega C \quad \text{and} \quad \bar{Z}_2 = jX_L = j\omega L$$

and where the network is terminated in a pure resistance of $\bar{Z}_L = R_L$ ohms, as shown in Fig. 210.

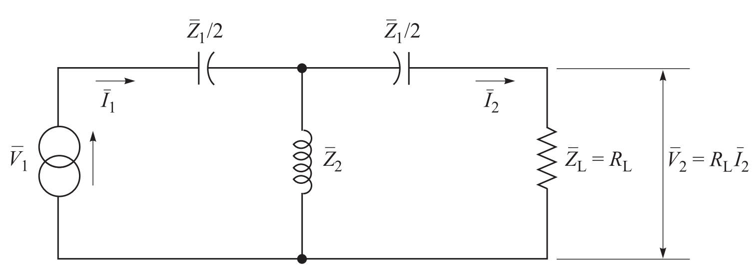


Fig. 210

Note that at very LOW frequencies the reactances of the series capacitors are very HIGH and the reactance of the shunt inductor is very LOW, so that $\bar{V}_2$ is very LOW at such frequencies. On the other hand, at very HIGH frequencies the reactances of the series capacitors are very LOW and the reactance of the shunt inductor is very HIGH, so that $\bar{V}_2$ is nearly EQUAL to $\bar{V}_1$ at such frequencies. Thus (in just a general way) we see that Fig. 210 constitutes a "high-pass" type of filter. But now let us get down to specific details.

To do this, we begin by noting that the two equations following Fig. 208 apply equally well to Fig. 210; thus, if you will now substitute into the second equation following Fig. 208 the values

$$\bar{Z}_1 = -j/\omega C \quad \text{and} \quad \bar{Z}_2 = j\omega L \quad (\text{and } \bar{Z}_L = R_L)$$

and then, after doing this, multiply the numerator and denominator by $-j/\omega L R_L$, you should find that

$$\frac{\bar{V}_2}{V_1} = \frac{1}{\left(1 - \dfrac{1}{2\omega^2 LC}\right) + j\dfrac{1}{\omega L R_L}\left(\dfrac{1}{4\omega^2 C^2} - \dfrac{L}{C}\right)} \tag{360}$$

Now let's pause and try to decide upon a reasonable value for R_L. To do this, we note that at very HIGH frequencies Fig. 210 would, for all practical purposes, become as shown in Fig. 211.

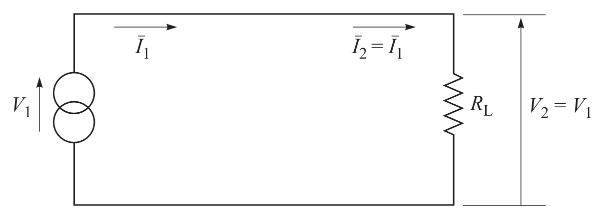


Fig. 211

Thus, for very HIGH values of ω the generator would see a pure resistance of R_L ohms, in which $\bar{V}_2/V_1 = 1$, with zero phase shift between $\bar{V}_2$ and V_1. *This would be the desired condition* here, because Fig. 210 is to be a HIGH-PASS filter.

With this in mind, and upon setting $\bar{Z}_1 = -j/\omega C$ and $\bar{Z}_2 = j\omega L$ in eq. (340), we have that the "characteristic impedance" of the T-network of Fig. 210 is equal to

$$\bar{Z}_0 = \sqrt{\frac{L}{C} - \frac{1}{4\omega^2 C^2}}$$

in which, as you'll note, the value of the term $1/4\omega^2 C^2$ *decreases rapidly in value as ω increases*; thus, at the preferred HIGH frequencies the value of $\bar{Z}_0$ becomes, for practical purposes, equal to $\sqrt{L/C}$ ohms. It thus makes sense to let

$$\bar{Z}_0 = R_L = \sqrt{\frac{L}{C}} \tag{361}$$

because this will cause Fig. 210 to become equal to the desired condition of Fig. 211 at high frequencies. Next, let us try to express eq. (360) in terms of a dimensionless ratio ω/ω_0, as we did for the case of the low-pass filter (eq. (353)). To do this, let us begin by noting that the imaginary term in the denominator of eq. (360) can be written as

$$\begin{aligned} \frac{j}{\omega L C R_L}\left(\frac{1}{4\omega^2 C} - L\right) &= \frac{j}{\omega C}\frac{1}{R_L}\left(\frac{1}{4\omega^2 LC} - 1\right) \\ &= \frac{j}{\omega\sqrt{LC}}\left(\frac{1}{4\omega^2 LC} - 1\right) \end{aligned} \tag{362}$$

because, by eq. (361),

$$\frac{j}{\omega C}\frac{1}{R_L}=\frac{j}{\omega}\sqrt{\frac{C}{LC^2}}=\frac{j}{\omega\sqrt{LC}}$$

Thus, upon substituting the j term (eq. (362)) into eq. (360), we have the desired form

$$\frac{\bar{V}_2}{V_1}=\frac{1}{\left(1-\dfrac{1}{2\omega^2 LC}\right)+j\dfrac{1}{\omega\sqrt{LC}}\left(\dfrac{1}{4\omega^2 LC}-1\right)} \tag{363}$$

The above form is especially useful because it can readily be expressed in terms of the ratio of any frequency ω to a fixed reference frequency ω_0. This can be done by defining that the reference frequency be equal to

$$\omega_0=\frac{1}{2\sqrt{LC}} \tag{364}$$

Thus, $\sqrt{LC}=1/2\omega_0$ and $LC=1/4\omega_0^2$, and upon making these substitutions into eq. (363), then making the substitution

$$h=\omega/\omega_0 \tag{365}$$

(that is, $\omega_0/\omega=1/h$), and then multiplying the numerator and denominator by h^3, you should find that eq. (363) becomes

$$\frac{\bar{V}_2}{V_1}=\frac{h^3}{h(h^2-2)+j2(1-h^2)} \tag{366}$$

If, now, we wish to investigate only the manner in which the *magnitude* of $\bar{V}_2/V_1$ varies with frequency, then eq. (366) becomes

$$\left|\frac{\bar{V}_2}{V_1}\right|=\frac{h^3}{\sqrt{h^2(h^2-2)^2+4(1-h^2)^2}}=\frac{h^3}{\sqrt{4+(h^6-4h^2)}}$$

or, since $1/\sqrt{X}=1/X^{1/2}=X^{-1/2}$, we can write the above in the form

$$\left|\frac{\bar{V}_2}{V_1}\right|=h^3[4+(h^6-4h^2)]^{-1/2}$$

or, in *decibels* (see notes 19 and 22 in the Appendix and eqs. (319), (320), and (321)), the above equation becomes

$$\mathrm{dB}=60\log h-10\log[4+(h^6-4h^2)] \tag{367}$$

Using your calculator, you can verify that the following "table of values" is correct for eq. (367), in which we've rounded off the dB values to two decimal places.

h	dB	h	dB	h	dB
0.2	−47.78	1.2	3.87	1.8	1.33
0.4	−29.15	1.3	3.68	1.9	1.09
0.6	−17.47	1.4	3.10	2.0	0.90
0.8	−8.13	1.5	2.51	2.5	0.39
1.0	0.00	1.6	2.02	3.0	0.20
1.1	2.79	1.7	1.63	4.0	0.06

A plot of the above results on semi-log paper is given in Fig. 212, with brief discussion following.

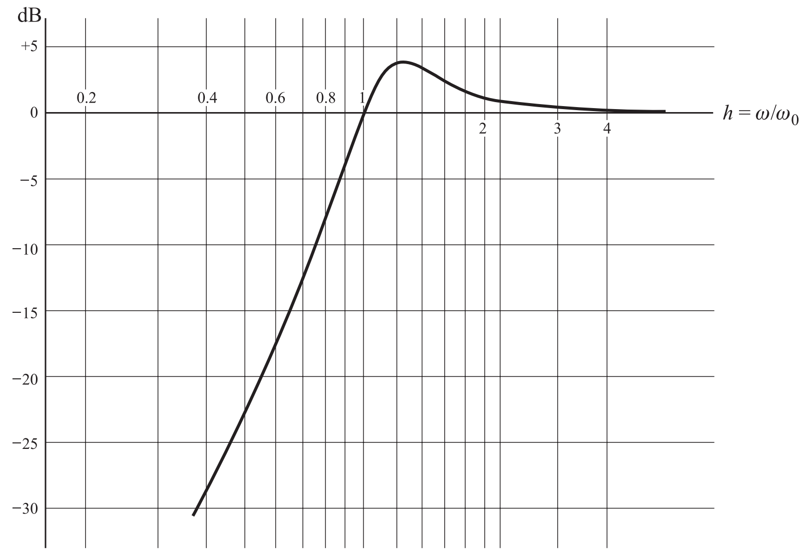


Fig. 212

Note that gain decreases rapidly for values of h less than 1; hence, if we wish, $h = 1$ can be taken to be the cut-off condition for Fig. 210. Thus, if ω_c is the cut-off frequency, then $\omega_c/\omega_0 = 1$; that is, $\omega_c = \omega_0$. Hence, by eq. (364)

$$\omega_c = \frac{1}{2\sqrt{LC}} \quad \text{rad/sec}$$

or

$$f_c = \frac{1}{4\pi\sqrt{LC}} \quad \text{hertz}$$

CHAPTER 10

Magnetic Coupling. Transformers. Three-Phase Systems

In this chapter we continue the study of the sinusoidal steady-state analysis of networks. This will include the theory and calculation of magnetically coupled circuits (transformers) and three-phase power calculations, with an introduction to the theory of "symmetrical components" as applied to three-phase circuits. These are all interesting applications of the algebra of the complex plane to the electric circuit.

10.1 Introduction to Magnetic Coupling; the Transformer

We begin with the suggestion that a careful rereading of sections 7.2 through 7.5 should be made at this time.

Suppose we have a coil of inductance L henrys, carrying a current of i amperes. If the coil current *changes*, then the amount of magnetic flux produced by the current also changes, thus causing, as we know, a *self-induced voltage* to appear in the coil.

Now suppose a *second coil* is brought up close to the first coil. Then some of the lines of flux, generated by the current in the first coil, *will link with some of the turns of the second coil*. It thus follows that a changing current in the first coil will produce a changing amount of flux in the second coil, *thereby causing a voltage to be induced into the second coil*. This is the principle of ELECTROMAGNETIC COUPLING between two coils, and is the basis of the highly important "electrical transformer."

A "transformer" thus consists of two coils placed relatively close together, so that at least part of the flux generated by each coil also links with the other coil. We call one of the coils the PRIMARY COIL and the other the SECONDARY COIL. We'll usually let L_1 be the inductance of the primary coil, and L_2 be the inductance of the secondary coil.

In the case of transformers designed to operate at the ''power line'' frequency of 60 Hz, the two coils are wound on what is called an ''iron core,'' which is constructed of thin sheets of silicon steel stacked and bolted together to form an assembly such as is illustrated in Fig. 213. The primary and secondary coils can be wound on opposite ''legs'' of the iron core, as shown in Fig. 214. The coils must be wound using insulated wire, to prevent adjacent turns from ''shorting'' together and to prevent the coils from making electrical contact with the iron core. On electrical diagrams, a transformer of this type is represented by the symbol shown in Fig. 215. The vertical lines drawn between L_1 and L_2 in Fig. 215 tell us that the coils are wound on an iron core.

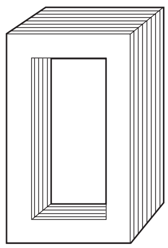


Fig. 213

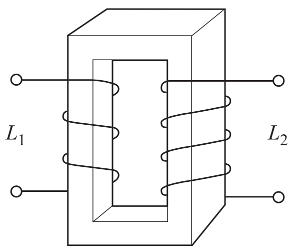


Fig. 214

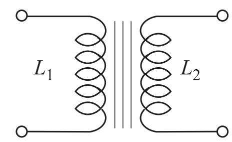


Fig. 215

The iron core thus serves a double purpose: it serves as a rigid coil form on which the primary and secondary coils are wound, and, since it is a ferromagnetic material, it allows a relatively small current to generate a large amount of magnetic flux. Transformers using iron cores are used mainly in power system work (60 Hz) and in audio frequency work (20 Hz to 16,000 Hz). At the very high frequencies used in radio, television, and radar, for example, transformer coils often consist of only a few turns of wire, wound on a ceramic or plastic tube. Since no ferromagnetic material is used, they are often referred to as ''air core'' transformers. The schematic symbol used to represent this type of transformer is the same as that shown in Fig. 215, except that no vertical lines are drawn between the primary and secondary coils.

The transformer is a most useful device, used to change voltage levels, match impedances, separate ac and dc currents and voltages, and so on. Let us begin our analysis with Fig. 216, which shows a transformer ''T'' in which i_1 and i_2 are the primary and secondary currents flowing at any instant of time t.

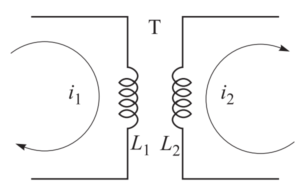


Fig. 216. L_1 **= inductance of primary coil;**
L_2 **= inductance of secondary coil.**

In the following discussion we'll make use of the ''principle of superposition,'' that is, that the COMBINED EFFECT *of both currents* in Fig. 216 is equal to the SUM OF THE EFFECTS *of each current considered separately*, as if the other current were absent.

With this in mind, let us, in Fig. 216, begin by disregarding the effect of current i_2 and considering only the effect of current i_1; to do this, let us (as in section 7.5) use the notation

$$\frac{di_1}{dt} = \text{rate of change of primary current}$$

and

$$v_2 = \text{voltage induced into } L_2 \text{ due to changing primary current}$$

Notice that we now have exactly the same situation that we had in section 7.5, that is, we are dealing with *an induced voltage caused by a changing current*. Thus the mathematical description of the situation will have *exactly the same form* as eq. (181) in that section, except that now, since we're dealing with the ratio of *current change in one coil* to the *voltage induced in a second coil*, we'll write "M" instead of "L" to signify this; thus

$$v_2 = M\frac{di_1}{dt} \tag{368}$$

in which M is measured in henrys and is called the MUTUAL INDUCTANCE between the two coils.

Or, if we disregard the effect of i_1 and consider only the effect of i_2, the preceding equation would become

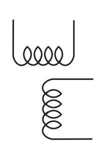
$$v_1 = M\frac{di_2}{dt} \tag{369}$$

where di_2/dt = rate of change of secondary current, and v_1 = voltage induced into primary coil L_1 due to changing secondary current.

In the above, it's important to note that the *same value of mutual inductance M* is used in both equations. This is a significant fact that can be proved to be true, both theoretically and by direct experiment.

The value of the mutual inductance M, for a given transformer, depends upon the following factors.

1. The *self-inductances* of the primary and secondary coils, that is, the values of L_1 and L_2.
2. The presence or absence of *ferromagnetic material*.
3. The *spacing* between the two coils; the farther apart they are, the less is the mutual inductance.
4. The *orientation* of L_1 with respect to L_2; for example, the mutual inductance if the two coils are at right angles to each other (as shown below) would be very small as compared to what it would be for the orientation of Fig. 216.



Another transformer parameter, called the COEFFICIENT OF COUPLING, denoted by "k," is widely used in practical work. It is defined as follows (where, as usual, "inductance" means "self-inductance"). In a transformer, let

L_1 = inductance of the primary coil,

L_2 = inductance of the secondary coil,

M = mutual inductance of the transformer.

Then the ‘‘coefficient of coupling’’ k between the two coils is defined by the equation

$$k = \frac{M}{\sqrt{L_1 L_2}} \tag{370}$$

thus,

$$M = k\sqrt{L_1 L_2} \tag{371}$$

The ‘‘coefficient of coupling’’ is a measure of the *closeness of coupling* that exists between the primary and secondary coils. k has a *maximum possible value of 1*; this value, $k = 1$, represents a theoretical case where ALL the flux generated by the primary current links ALL the turns of the secondary coil, and ALL the flux generated by the secondary current links ALL the turns of the primary coil. This situation can never be completely realized in any actual transformer because, for practical reasons, the coils cannot be constructed and positioned so that ALL the flux generated in one coil will link with every turn of the other coil.

Hence, *k is always less than 1* for an actual transformer. The value of k will usually run from 1 to 10% for ‘‘air core’’ transformers and up to 99.5% for high-quality iron-cored transformers. The ‘‘best’’ value of k will depend upon the particular application we are interested in; for air-core transformers, used at very high frequencies, the best value of k is usually from 1 to 10%, while for iron-cored transformers, used at lower frequencies, the value of k should usually be as close to 100% as possible.

Problem 184

Suppose a transformer has a primary coil of inductance 0.9 henry, a secondary coil of inductance 15 henrys, and coefficient of coupling 85%. What voltage is induced into the secondary coil if the primary current is

(a) a steady current of 25 amperes,

(b) changing at a rate of 25 amperes per second?

Problem 185

(a) Show that, for any given transformer, the SAME RATE OF CHANGE OF CURRENT, IN EITHER COIL, INDUCES EQUAL VOLTAGE IN THE OTHER COIL.

(b) As applied to problem 184, this says that if the secondary current were changing at the rate of 25 amperes/second, this would induce a voltage of 78.0775 volts in the primary coil. Verify, by calculation, that this would be true.

10.2 Dot-Marked Terminals. Induced Voltage Drops

Networks containing transformers are solved using the same procedures as always, except that if transformers are present we must take into account the voltage drops represented by eqs. (368) and (369).

It is important that the *correct signs* be used with these equations. To insure that the correct signs are used, each transformer in the network must be marked with what are called POLARITY DOTS. *Two* such dots are required for each transformer, one being placed at one of the terminals on the primary side and the other at one of the terminals on the secondary side. The dots are used to indicate the sense in which the primary and secondary coils are wound with respect to each other; the dots are placed so that

> Currents *flowing into the dot-marked terminals* produce magnetic flux *in the same direction in the magnetic path*.

This is illustrated in Figs. 217 and 218 below, where we're using "coil forms" so that we can more easily see the senses in which the coils are wound. Note that the primary coils are wound in the *same sense* in both figures, but the secondary coils are wound in *opposite senses* in the two figures.

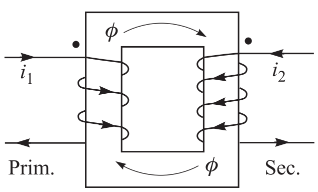


Fig. 217

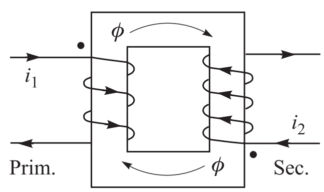


Fig. 218

Now, using the right-hand rule (with fingers curled in the direction of current flow, the thumb points in the direction of magnetic flux), you can verify that *in both figures* currents *flowing into the dot-marked terminals* produce *magnetic flux ϕ in the same direction in the magnetic path*. This means that current flowing into *either* dot-marked terminal has the SAME EFFECT MAGNETICALLY as current flowing into the *other* dot-marked terminal.

With this in mind, let us redraw Fig. 216 as Fig. 219, in which we'll assume that the winding senses are such that the "dot markings" are correct as shown.

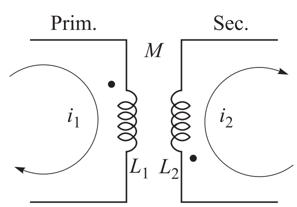


Fig. 219

In Fig. 219, let i_1 and i_2 denote instantaneous values of primary and secondary current which, when flowing in the directions shown, will be considered to be "positive" values of current.

Next, note that there are *three inductance values* to be considered; thus

L_1 = self-inductance of primary coil,

L_2 = self-inductance of secondary coil,

M = mutual inductance between primary and secondary coils.

Now notice that there are FOUR VOLTAGE DROPS present in Fig. 219, as follows.

Voltage drops in primary coil

$$\begin{cases} L_1 \dfrac{\mathrm{d}i_1}{\mathrm{d}t} = \text{voltage drop due to changing primary current (section 7.5)} \\ M \dfrac{\mathrm{d}i_2}{\mathrm{d}t} = \text{voltage drop due to changing secondary current (eq. (369))} \end{cases}$$

Voltage drops in secondary coil

$$\begin{cases} L_2 \dfrac{\mathrm{d}i_1}{\mathrm{d}t} = \text{voltage drop due to changing secondary current (section 7.5)} \\ M \dfrac{\mathrm{d}i_2}{\mathrm{d}t} = \text{voltage drop due to changing primary current (eq. (368))} \end{cases}$$

Now let

v_{pri} = instantaneous voltage drop in primary coil, and
v_{sec} = instantaneous voltage drop in secondary coil, and thus

$$v_{\text{pri}} = L_1 \frac{\mathrm{d}i_1}{\mathrm{d}t} + M \frac{\mathrm{d}i_2}{\mathrm{d}t} \tag{372}$$

$$v_{\text{sec}} = M \frac{\mathrm{d}i_1}{\mathrm{d}t} + L_2 \frac{\mathrm{d}i_2}{\mathrm{d}t} \tag{373}$$

in which all positive signs are used because of the way that i_1 and i_2 are associated with the dot-marked terminals in Fig. 219 (i_2 has the same effect, magnetically, as i_1, and vice versa).

Now assume the *sinusoidal steady-state condition*, so that the instantaneous currents are given by the equations

$$i_1 = I_{\text{p1}} \sin(\omega t - \phi_1) \tag{374}$$

$$i_2 = I_{\text{p2}} \sin(\omega t - \phi_2) \tag{375}$$

where I_{p1} and I_{p2} are peak (maximum) values of current, and ϕ_1 and ϕ_2 are the current phase angles with respect to the sinusoidal generator voltage. Then the *maximum* RATES OF CHANGE of the currents are equal to the maximum values of $\mathrm{d}i_1/\mathrm{d}t$ and $\mathrm{d}i_2/\mathrm{d}t$ (see discussion in connection with eq. (194) in Chap. 8); thus

$$\left[\frac{\mathrm{d}i_1}{\mathrm{d}t}\right]_{\max} = \omega I_{\text{p1}} \quad \text{and} \quad \left[\frac{\mathrm{d}i_2}{\mathrm{d}t}\right]_{\max} = \omega I_{\text{p2}}$$

and thus, by the basic eq. (181) in Chap. 7, the PEAK VALUES of the FOUR SINUSOIDAL SELF-INDUCED VOLTAGES are

$$\omega L_1 I_{\text{p1}} \quad \omega M I_{\text{p2}} \quad \omega M I_{\text{p1}} \quad \omega L_2 I_{\text{p2}}$$

and since, in the sinusoidal steady state, an *induced voltage* always *leads* the inducing current rent by 90 degrees, eqs. (372) and (373) become, for the sinusoidal steady-state case,

$$v_{\text{pri}} = \omega L_1 I_{\text{p1}} \sin(\omega t - \phi_1 + 90^\circ) + \omega M I_{\text{p2}} \sin(\omega t - \phi_2 + 90^\circ) \tag{376}$$

$$v_{\text{sec}} = \omega M I_{\text{p1}} \sin(\omega t - \phi_1 + 90^\circ) + \omega L_2 I_{\text{p2}} \sin(\omega t - \phi_2 + 90^\circ) \tag{377}$$

Note: Although we've written $-\phi_1$ and $-\phi_2$, these angles can be either positive or negative, depending upon the particular circuit the transformer is connected into.

Now, as we know (see discussion following eq. (194) in Chap. 8), the four rotating voltage components in the above two equations can be regarded as *four stationary vector* *components*. In doing this, we'll use ''rms'' values of voltage and current instead of ''peak'' values, and thus, letting

$\bar{V}_{\text{pri}} =$ rms vector value of induced voltage appearing in PRIMARY coil, and
$\bar{V}_{\text{sec}} =$ rms vector value of induced voltage appearing in SECONDARY coil,

in the primary coil,

$$\bar{V}_{\text{pri}} = \omega L_1(\bar{I}_1 + 90^\circ) + \omega M(\bar{I}_2 + 90^\circ)$$

in the secondary coil,

$$\bar{V}_{\text{sec}} = \omega M(\bar{I}_1 + 90^\circ) + \omega L_2(\bar{I}_2 + 90^\circ)$$

where notation of the form $X(\bar{I} + 90^\circ)$ is used to denote that a vector voltage drop $X\bar{I}$ *leads* the current vector $\bar{I}$ by 90°. However, since we're representing vector quantities by complex numbers, all we need to do, to rotate a vector through 90°, is to multiply the vector by j (as discussed in connection with eq. (196) in Chap. 8). Thus the preceding two equations become

$$\bar{V}_{\text{pri}} = j\omega L_1\bar{I}_1 + j\omega M\bar{I}_2 \tag{378}$$

$$\bar{V}_{\text{sec}} = j\omega M\bar{I}_1 + j\omega L_2\bar{I}_2 \tag{379}$$

in which ωL_1 and ωL_2 are the inductive reactances of the primary and secondary coils, each considered separately (as if they were not part of the transformer), and where the quantity ωM is measured in ohms and called the ''mutual reactance.''

Equations (378) and (379) are expressions connecting the four inductive voltage drops present in a transformer. We'll make use of these equations in section 10.3, to derive the basic equations for the sinusoidal steady-state analysis of coupled circuits.

Before going on to the next section, however, there are some more points we should mention regarding the ''polarity dots'' associated with a transformer.

First, it should be noted that, in some simple problems involving only a single transformer, in which we need to find only the *magnitude* of the secondary current, it will not be necessary to be concerned about polarity dots at all. In such a case we can simply assume that the primary and secondary currents flow into like-marked terminals.

However, in more complicated cases, in which, for example, we have two or more transformers having a common secondary current, then we *must* take the ''dot-marks'' into account. This can be a problem, as it may not be possible to actually *see* the senses in which the primary and secondary coils of a transformer are wound (because the windings may be covered with a strong tape, used both to insulate the windings and to keep them firmly in place). If the manufacturer has not dot-marked such a transformer, this can be done by experimental means, by the user, in several ways. One such way is explained with the aid of Figs. 220 and 221, as follows, in which V is a constant value of ac voltage and I is the reading of an ac ammeter.

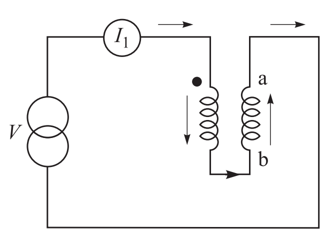


Fig. 220

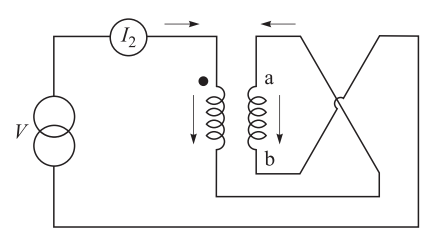


Fig. 221

Note that in the test setup the primary and secondary coils are connected *in series* in both figures, and hence, in each figure, the primary current is equal to the secondary current. Now, in the figures, let the arrows denote the direction of current at a particular moment. Notice that the direction of the primary current is the *same* in both figures, but the secondary current flows in *opposite* directions. This means that in *one* of the figures the mutual magnetic effects of the primary and secondary currents will be ADDITIVE, while in the other figure they will be SUBTRACTIVE. Since the *total reactance* seen by the generator is greater in the additive case than in the subtractive case, it follows that the ammeter will read a *smaller* value of current in the additive case than in the subtractive case. With these points in mind, the actual test procedure is as follows.

First, to one of the primary leads attach a tag having a ''dot,'' then attach tags to the secondary leads, one labeled ''a'' and the other ''b,'' so that we have the condition shown in the figures. Now apply the ac voltage V first to the setup of Fig. 220, and then to that of Fig. 221. Upon doing this, suppose it is found that I_1 is *smaller* than I_2; this would mean that current flowing into the ''b'' terminal has the same effect, magnetically, as current flowing into the dot-marked terminal, and thus *terminal b should be marked with a dot*. Or, if the opposite is found to be true (I_2 less than I_1), then terminal ''a'' would be marked with a dot.

10.3 Sinusoidal Analysis of Magnetically Coupled Circuits

As we know, sinusoidal rms values of voltages and currents of the same frequency can be regarded as vector quantities, and thus can be represented and manipulated in the form of complex numbers.

With this in mind, let a sinusoidal voltage of $\bar{V}$ volts rms be applied to the basic transformer-coupled circuit of Fig. 222, thus producing primary and secondary currents of $\bar{I}_1$ and $\bar{I}_2$ amperes rms, as shown. We now wish to derive formulas that will, given the circuit constants, enable us to calculate the values of $\bar{I}_1$ and $\bar{I}_2$.

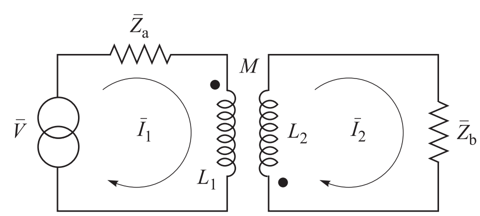


Fig. 222

In Fig. 222, L_1 is the inductance of the primary coil considered by itself, L_2 is the inductance of the secondary coil considered by itself, and M is the mutual inductance between the two coils. Also, $\bar{Z}_a$ = partial impedance of primary circuit. This includes any internal impedance of the generator and any resistance L_1 may have, but does *not* include the inductive reactance ωL_1 of the primary coil, as can be seen from Fig. 222. Similarly, $\bar{Z}_b$ = partial impedance of secondary circuit, including any resistance the secondary coil may have. Thus $\bar{Z}_b$ includes everything around the secondary loop *except* the inductive reactance ωL_2 of the secondary coil.

Let us now write the voltage equations for Fig. 222, beginning with the voltage equation around the PRIMARY circuit. To do this, note that there are *three voltage drops* around the primary loop, the first being $\bar{Z}_a\bar{I}_1$, with the other two being induced voltage drops given by eq. (378). Since the vector sum of these three voltage drops is equal to the applied voltage $\bar{V}$, the voltage equation around the primary loop is

$$(\bar{Z}_a + j\omega L_1)\bar{I}_1 + j\omega M\bar{I}_2 = \bar{V} \tag{380}$$

Next, there are likewise three voltage drops around the SECONDARY loop, one being $\bar{Z}_b\bar{I}_2$, the other two being induced voltage drops given by eq. (379). Since there is no independent generator included in the secondary loop, the vector sum of these three drops has to be equal to *zero*, and thus the voltage equation around the secondary loop is

$$j\omega M\bar{I}_1 + (\bar{Z}_b + j\omega L_2)\bar{I}_2 = 0 \tag{381}$$

Now let

$$\bar{Z}_1 = \bar{Z}_a + j\omega L_1 \tag{382}$$

$$\bar{Z}_2 = \bar{Z}_b + j\omega L_2 \tag{383}$$

where it should be very carefully noted that $\bar{Z}_1 =$ the total series impedance AROUND THE COMPLETE PRIMARY CIRCUIT, considered by itself, just as if the secondary circuit did not exist; similarly, $\bar{Z}_2 =$ the total series impedance AROUND THE COMPLETE SECONDARY CIRCUIT, considered by itself, just as if the primary circuit did not exist.

Now, in eqs. (380) and (381), replace the coefficients of the unknown currents with the left-hand sides of eqs. (382) and (383). If we do this, eqs. (380) and (381) become our FUNDAMENTAL EQUATIONS FOR THE BASIC TRANSFORMER-COUPLED CIRCUIT OF FIG. 222; thus

$$\bar{Z}_1\bar{I}_1 + j\omega M\bar{I}_2 = \bar{V} \tag{384}$$

$$j\omega M\bar{I}_1 + Z_2\bar{I}_2 = 0 \tag{385}$$

We can now solve the above two equations for the primary current $\bar{I}_1$ as follows:

$$\bar{I}_1 = \frac{\begin{vmatrix} \bar{V} & j\omega M \\ 0 & \bar{Z}_2 \end{vmatrix}}{\begin{vmatrix} \bar{Z}_1 & j\omega M \\ j\omega M & \bar{Z}_2 \end{vmatrix}} = \frac{\bar{V}\bar{Z}_2}{\bar{Z}_1\bar{Z}_2 + \omega^2 M^2}$$

Or, upon dividing the numerator and denominator of the last fraction by $\bar{Z}_2$, the equation for the PRIMARY CURRENT takes the form

$$\bar{I}_1 = \frac{\bar{V}}{\bar{Z}_1 + \dfrac{\omega^2 M^2}{\bar{Z}_2}} \tag{386}$$

Note that the denominator in the above equation is the total impedance seen by the generator in Fig. 222. Hence, since $\bar{Z}_1$ is the impedance of the primary circuit considered by itself, it follows that the *second term in the denominator* is the impedance "coupled" or "reflected" *from the secondary circuit into the primary circuit*. Thus, letting $\bar{Z}_{\text{ref}}$ denote this "reflected impedance," we have

$$\bar{Z}_{\text{ref}} = \frac{\omega^2 M^2}{\bar{Z}_2} \tag{387}$$

where $\bar{Z}_{ref}$ = impedance coupled or "reflected" *into the primary circuit from the secondary circuit* (note that $\bar{Z}_{ref}$ appears *in series with the primary coil*); and $\bar{Z}_2$ = total series impedance of the secondary circuit, considered by itself, as defined following eq. (383).

Hence, as far as the generator is concerned, Fig. 222 can be redrawn as in Fig. 223.

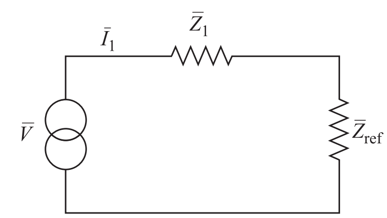


Fig. 223

Thus, from Fig. 223,

$$\bar{I}_1 = \frac{\bar{V}}{\bar{Z}_1 + \bar{Z}_{ref}} \tag{388}$$

Another useful relationship is found by solving eq. (385) for I_2; thus

$$\bar{I}_2 = \frac{-j\omega M \bar{I}_1}{\bar{Z}_2} \tag{389}$$

Problem 186

In Fig. 222 both current arrows are drawn in the clockwise sense. What change, if any, would appear in eq. (389) if

(a) both arrows were drawn in the counterclockwise sense?

(b) if the $\bar{I}_1$ arrow remained in the cw sense but the $\bar{I}_2$ arrow were drawn in the ccw sense?

Problem 187

In Fig. 224, the primary and secondary coils have equal inductances of 45 millihenrys. (Any resistance the coils may have is included in the 5-ohm and 2-ohm resistance values.) Find the magnitude of generator current, given that the frequency is 200 rad/sec. (*Answer*: 5.249 amperes)

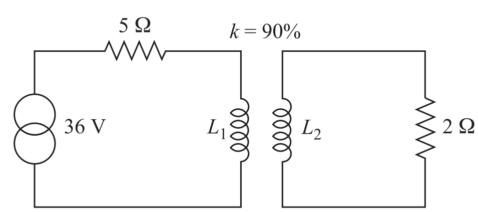


Fig. 224

(See "discussion note" given with the solution to the above problem.)

Problem 188

In Fig. 225, a generator of zero internal impedance produces 28 volts at 400,000 rad/sec. The primary and secondary coils have equal inductances of 60 microhenrys.

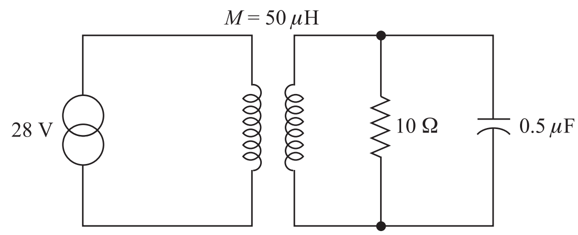


Fig. 225

Find,

(a) magnitude of generator current, (*Answer*: 6.030 amperes)

(b) magnitude of current in secondary coil. (*Answer*: 6.000 amperes)

Problem 189

In Fig. 226, the generator frequency is $100{,}000/2\pi$ hertz (Hz), the values of the circuit parameters being in ohms, microhenrys, and microfarads. It is also given that $L_1 = 120\,\mu\text{H}$ and $L_2 = 200\,\mu\text{H}$.

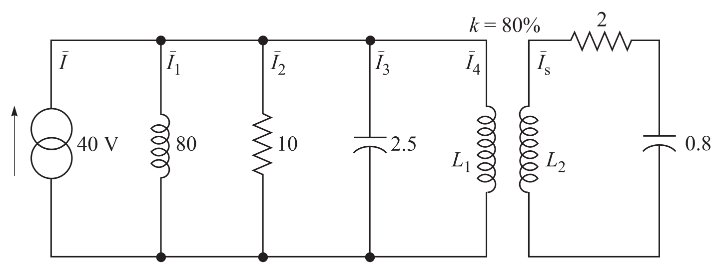


Fig. 226

(a) Find the magnitude of generator current. (*Answer*: 10.966 amperes)

(b) Find the magnitude of secondary current. (*Answer*: 7.292 amperes)

Parts (c) through (e), which follow, will serve as a review of some important points concerning the calculation of POWER, as developed in section 8.5.

(c) Using only the details found in part (a) and the principle developed in connection with Fig. 150, find the TRUE POWER produced by the generator in Fig. 226.

(d) Repeat part (c), now making use only of eq. (228) and Fig. 155.

(e) Repeat part (c), now making use of the fact that the power produced in a pure resistance is equal to the ''square of the magnitude of current, times resistance.''

In problems 187, 188, and 189 it should be noted that we did not need to ''dot-mark'' the transformers in order to find the required answers. This is because our analysis was *based upon Fig. 222*, which led to the basic pair of eqs. (384) and (385). Let us consider this in more detail, as follows.

Suppose, in Fig. 222, that the position of the dots on the secondary side had been reversed (everything else remaining the same). In that case eqs. (384) and (385) would have read

$$\bar{Z}_1\bar{I}_1 - j\omega M\bar{I}_2 = \bar{V}$$
$$-j\omega M\bar{I}_1 + Z_2\bar{I}_2 = 0$$

which, you'll notice, will give the *same value* of $\bar{I}_1$ as given by eq. (386) (and the same value of reflected impedance as given by eq. (387)). As a matter of fact, changing the position of the dots in Fig. 222 would only have the effect of changing the phase angle of the secondary current by 180 degrees (the same as changing the direction of $\bar{I}_2$, as discussed in problem 186).

It must be emphasized, however, that in some types of circuit it *is necessary* to take the placement of the dots into account. These are cases in which the primary and secondary currents are not totally separated, as in Fig. 222, but, instead, share a common circuit element. In such cases we must be careful to assign the proper algebraic sign to all *mutually induced voltages* of the form $\pm j\omega M\bar{I}$. This is illustrated in the following problems.

Problem 190

Figures 227 and 228 depict two circuits that are identical in all respects EXCEPT for the placement of the reference dots. Write the equation for the current $\bar{I}$ for (a) Fig. 227, (b) Fig. 228.

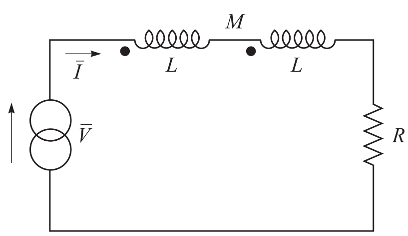


Fig. 227

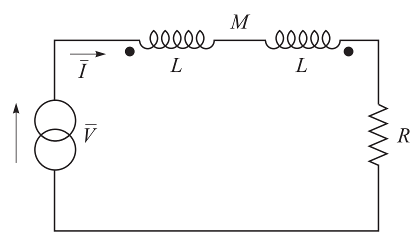


Fig. 228

Problem 191

In Fig. 229, the generator produces $\bar{V} = 50\underline{/0^\circ}$ volts at a frequency of 12,500 rad/sec. It is given that $L_1 = 0.0020$ henry and $L_2 = 0.0045$ henry, the resistance values being in ohms. Find the magnitude of voltage $\bar{V}_o$ relative to ground.

(*Answer*: 23.54 V)

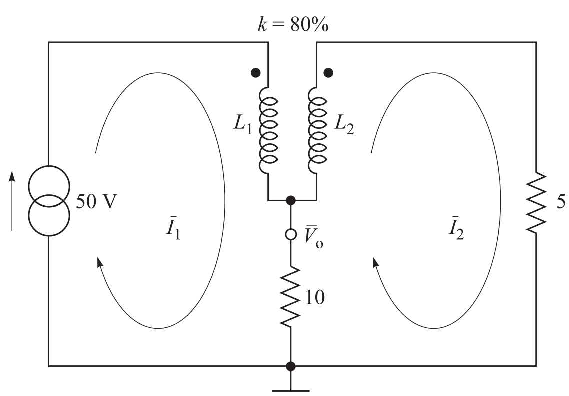


Fig. 229

Problem 192
Rework problem 191 with everything remaining the same EXCEPT let it be given that the secondary winding is now wound in the opposite sense from that assumed in Fig. 229. (*Answer*: 26.94 V)

Problem 193
In problem 191 show that $\bar{V}_o = 23.54\underline{/-64.29^\circ}$.

Problem 194
In problem 192 show that $\bar{V}_o = 26.94\underline{/-19.85^\circ}$.

10.4 The "T" Equivalent of a Transformer

In the analysis of transformer-coupled circuits it's often helpful to replace a transformer with its T-NETWORK EQUIVALENT. This can be done as follows, beginning with Fig. 230.

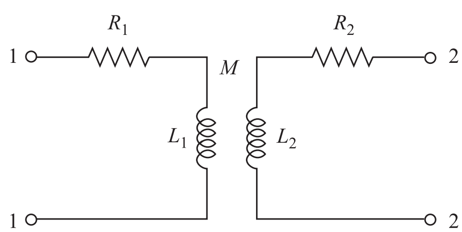


Fig. 230

In the figure,

R_1 = resistance of the primary coil,
R_2 = resistance of the secondary coil,
L_1 = inductance of the primary coil,
L_2 = inductance of the secondary coil,
M = mutual inductance between L_1 and L_2.

The conversion of a network into its T equivalent is done by using the procedure of section 9.2. Thus, for Fig. 230, we begin with the following relationships:

$$\bar{Z}_{1O} = R_1 + jX_1$$

(because, with terminals 2, 2 open, the secondary circuit has no effect when we look into terminals 1, 1);

$$\bar{Z}_{2O} = R_2 + jX_2$$

(because, with terminals 1, 1 open, the primary circuit has no effect when we look into terminals 2, 2)

$$\bar{Z}_{1S} = R_1 + jX_1 + \frac{\omega^2 M^2}{R_2 + jX_2}$$

(here, looking into 1, 1 with 2, 2 shorted, we're making use of eq. (387)).

Now, upon substituting the above values into eq. (282), you should find that

$$\bar{Z}_3 = \sqrt{-\omega^2 M^2} = j\omega M = jX_{\mathrm{m}} \tag{390*}$$

Then, by eq. (283)

$$\bar{Z}_2 = \bar{Z}_{2\mathrm{O}} - \bar{Z}_3 = R_2 + jX_2 - jX_{\mathrm{m}}$$

thus,

$$\bar{Z}_2 = R_2 + j\omega(L_2 - M) \tag{391}$$

and, by eq. (284)

$$\bar{Z}_1 = \bar{Z}_{1\mathrm{O}} - \bar{Z}_3 = R_1 + jX_1 - jX_{\mathrm{m}}$$

thus,

$$\bar{Z}_1 = R_1 + j\omega(L_1 - M) \tag{392}$$

Hence, by equations (390) through (392), we have that the T EQUIVALENT of the transformer of Fig. 230 is as shown in Fig. 231.

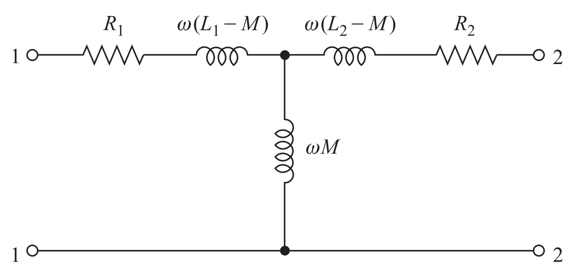


Fig. 231

Thus, while Figs. 230 and 231 are very different in appearance, they are equal in performance as far as *alternating current* (ac) is concerned, at a given frequency.

Problem 195
Laboratory measurements on a certain transformer show that the primary winding has 14 ohms of resistance and 0.09 henry of inductance, while the secondary winding has 4 ohms of resistance and 0.04 henry of inductance. The coefficient of coupling is found to be 30%. Draw and label the "T equivalent" of the transformer, giving the correct values of resistance and inductance.

Problem 196
Let a generator, producing 100 volts rms at a frequency of 500 radians/second, be connected to the primary terminals of the transformer in the preceding problem. The generator has negligible internal impedance. Now let a load of 2 ohms be connected to the secondary terminals. Making use of the T equivalent for the transformer found in the preceding problem, find the magnitude of the voltage across the 2-ohm load. (*Answer*: 1.961 volts)

Problem 197
Rework problem 196, this time using the basic coupled circuit formulas derived in section 10.3.

* $X_{\mathrm{m}} = \omega M$ is called the "mutual reactance," where M is the mutual inductance.

In closing, it should be noted that, in Fig. 231, the value of either $(L_1 - M)$ or $(L_2 - M)$ may come out to have a *negative* value. In such a case Fig. 231 is still, for purposes of analysis, a perfectly valid representation of the transformer, even though "negative inductance" does not physically exist. Thus, even though such an equivalent T could not be physically constructed, it would still be a perfectly valid representation of the actual transformer at a given frequency.

10.5 The Band-Pass Double-Tuned Transformer

THE TRANSMISSION OF INFORMATION THROUGH SPACE, that is, by "wireless," is accomplished by impressing the information upon a "carrier wave" whose frequency is *much higher* than the highest frequency present in the information to be transmitted.

The process of transferring information onto a high-frequency carrier wave is called "modulation," and the carrier wave is said to be "modulated" by the information.

An *unmodulated* carrier wave consists of a SINGLE-FREQUENCY sinusoidal wave occupying just ONE POINT in the frequency spectrum. However, when a carrier wave is modulated *new frequencies*, above and below the carrier frequency, are created. These new frequencies are called "side-band" frequencies, and appear as a cluster of frequencies with the carrier frequency in the center.* It is for this reason that a circuit designed to handle a *modulated carrier wave* must be a BAND-PASS type of network. One network that is useful in this regard is the double-tuned transformer, which let us now investigate with the aid of Fig. 232.

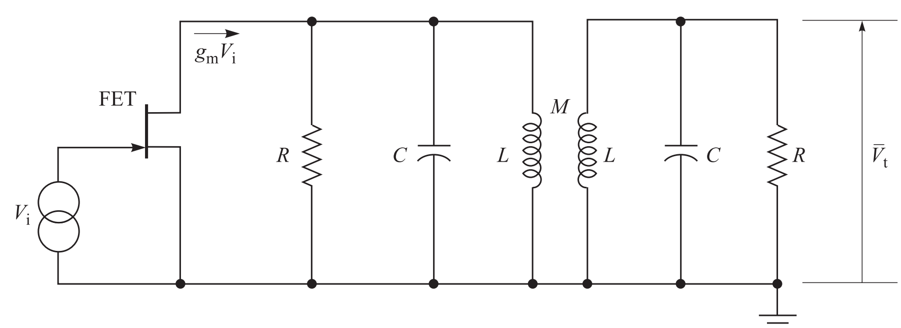


Fig. 232

To begin, the component labeled FET is a solid-state device called a "field-effect" transistor. The INPUT signal voltage is denoted by V_i, which we'll take as the reference vector. The OUTPUT voltage is denoted by $\bar{V}_t$; thus the VOLTAGE GAIN of the stage is

$$\bar{G} = \bar{V}_t / V_i \tag{393}$$

A field-effect transistor has very high internal gain but very high internal resistance, and is therefore a CONSTANT-CURRENT type of generator (section 4.7). Thus the output current of a FET, for given V_i, remains very nearly *constant* as the value of the load

* See note 24 in Appendix, also note 25.

impedance changes; this is true for all values of load impedances normally encountered in practical work. In Fig. 232 the constant-current output of the FET has the value $g_m V_i$, where g_m is a constant transistor parameter, called the "transconductance," whose value depends upon the particular transistor being used.

One way to begin the analysis of Fig. 232 is to convert the constant-current generator into an equivalent *constant-voltage* generator; one convenient way to do this is to start with Fig. 233.

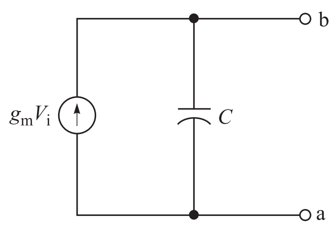


Fig. 233

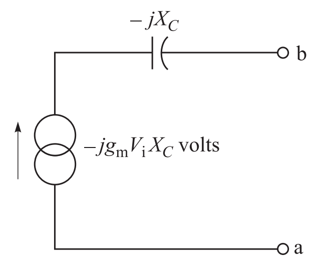


Fig. 234

In Fig. 233, note that we've detached the constant-current generator and the capacitor of C farads from the primary side of Fig. 232. We now wish to convert Fig. 233 into an equivalent constant-voltage generator; this can be done by making use of Thevenin's theorem (section 4.6) as follows.

First, in Fig. 233, note that the open-circuit voltage between terminals (a, b) is equal to the current $g_m V_i$ times the reactance of the capacitor C; thus the *voltage* of the equivalent generator is equal to $-jg_m V_i X_C$, as shown in Fig. 234.

Next, the *internal impedance* of the equivalent generator is equal to the impedance seen looking into terminals (a, b) in Fig. 233 with the FET replaced by its internal impedance. Since a FET has an extremely high internal resistance or impedance, it follows that the impedance, looking into terminals (a, b) in Fig. 233 is merely equal to the reactance of capacitor C, that is, $-jX_C$. Thus, by Thevenin's theorem, Fig. 234 is the *constant-voltage* equivalent of Fig. 233.

Next let's consider the case of parallel R and L, as shown in Fig. 235.

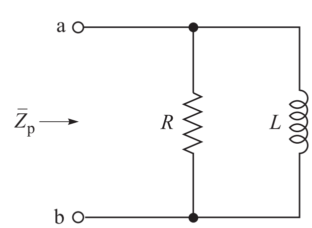


Fig. 235

Our object now is to convert the *parallel* circuit of Fig. 235 into an approximately equivalent *series* circuit. To do this, we begin by noting that the input impedance looking into terminals (a, b) in Fig. 235 is equal to (product of the two, over the sum)

$$\bar{Z}_p = \frac{jRX_L}{R + jX_L} = \frac{jRX_L(R - jX_L)}{R^2 + X_L^2} = \frac{RX_L^2 + jR^2 X_L}{R^2 + X_L^2} \tag{394}$$

In the PARTICULAR APPLICATION HERE, however, the above equation can, for practical purposes, be considerably simplified.

To do this, we must look back to Fig. 232 and note that L and C here constitute a PARALLEL resonant circuit with $X_L = X_C$ at the carrier frequency (the center frequency of the passband).

As we found in section 8.7, the input impedance to a parallel circuit is *high* at and near the resonant frequency, even though the *individual reactance values*, X_L and X_C, will have quite *low* values at the same frequencies. At the same time, the value of the shunt resistance R must be *much higher* than either X_L or X_C, in order to prevent R from ''swamping out'' the effect of the high-impedance parallel LC circuit at resonance, which would, among other things, cause the gain of the stage to be excessively low at and near the resonant frequency.

Thus, in practice, the value of R^2 will be *much greater* than the value of either X_L^2 or X_C^2, and hence, for practical purposes, the *denominator* of eq. 394 can be written as R^2 instead of $R^2 + X_L^2$; thus eq. 394 becomes,

$$\bar{Z}_\mathrm{p} = \frac{RX_L^2 + jR^2X_L}{R^2} = \frac{X_L^2}{R} + jX_L \tag{395}$$

that is,

$$\bar{Z}_\mathrm{p} = r + jX_L \tag{396}$$

where $r = X_L^2/R$.

Note that eq. (396) represents a resistance of r ohms *in series* with the reactance jX_L. Thus the *parallel* circuit of Fig. 235 can be replaced by the *series* circuit of Fig. 236.

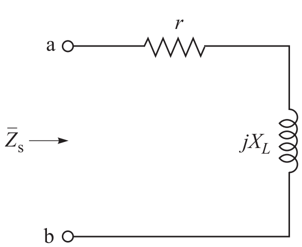


Fig. 236

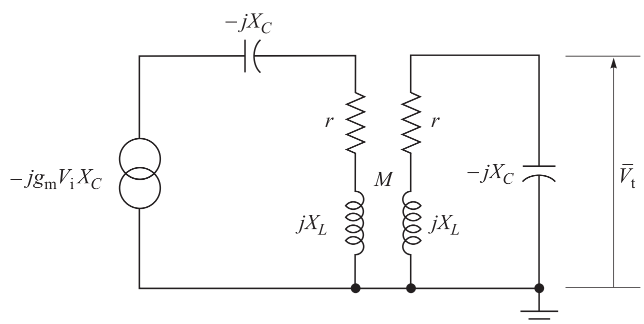


Fig. 237

With all the foregoing in mind, note that the original circuit of Fig. 232 can now be drawn in the form of Fig. 237. In this figure, r can include any resistance the primary and secondary windings may have.

Now, for the final step, replace the transformer with its T equivalent (Fig. 231). Doing this, and also adding the two loop currents $\bar{I}_1$ and $\bar{I}_2$, Fig. 237 becomes Fig. 238.

We now wish to find the value of $\bar{G}$ in eq. (393). Since, by inspection of Fig. 238, $\bar{V}_\mathrm{t} = (-jX_C)\bar{I}_2$, eq. (393) becomes

$$\bar{G} = -jX_C\bar{I}_2/V_\mathrm{i} \tag{397}$$

Thus, to find $\bar{G}$ we must find the value of $\bar{I}_2$. To do this, let us begin by writing the two loop-voltage equations for Fig. 238, which, as you can verify, are

$$[r + j(X_L - X_C)]\bar{I}_1 - jX_\mathrm{m}\bar{I}_2 = -jg_\mathrm{m}V_\mathrm{i}X_C \tag{398}$$

$$-jX_\mathrm{m}\bar{I}_1 + [r + j(X_L - X_C)]\bar{I}_2 = 0 \tag{399}$$

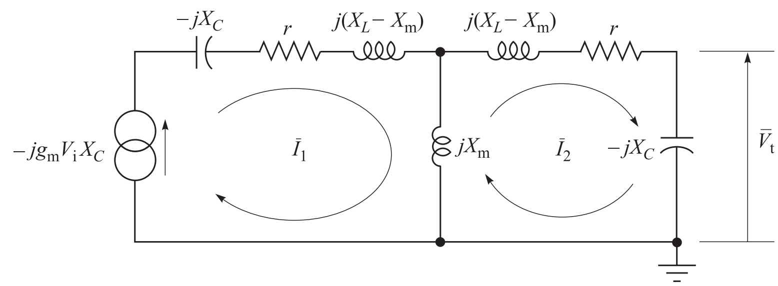


Fig. 238

One way to solve the above two equations for $\bar{I}_2$ is to use the method of determinants. To do this, as you'll recall, the first step is to find the value of Δ (delta), where Δ, here, is the value of the determinant formed from the coefficients of the two unknown currents, which, as you should verify, gives the value

$$\Delta = r^2 + X_m^2 - (X_L - X_C)^2 + j2r(X_L - X_C)$$

and thus, continuing with determinant procedure, we have

$$\bar{I}_2 = \frac{g_m V_i X_m X_C}{\Delta}$$

and hence, by eq. (397),

$$\bar{G} = \frac{-jg_m X_m X_C^2}{\Delta} = \frac{-jg_m X_m X_C^2}{r^2 + X_m^2 - (X_L - X_C)^2 + j2r(X_L - X_C)} \tag{400}$$

The difficulty now is that it's hard to "get a handle" on eq. (400), to see what it really means. Let us therefore work on the equation and try to get it in a different form, better suited to our needs. This can be done by a combination of algebraic manipulation and making use of certain relationships that are known to exist in practical applications of Fig. 232. One such procedure is as follows.

We begin by defining that, in our work here, the "resonant" frequency will be the frequency at which $X_L = X_C$.

In regard to Fig. 232, since the primary and secondary circuits separately consist of equal parallel values of R, L, and C, it follows that the primary and secondary circuits will separately have the same value of resonant frequency, which we'll denote by "ω_0" radians/second. (Our equations will appear less cluttered up if we use "rad/sec" instead of "cycles/sec," where, as always, $\omega = 2\pi f$.) Thus, to begin, we define that, in Fig. 232, the resonant frequency is defined by the equation

$$\omega_0 L = \frac{1}{\omega_0 C} \tag{401}$$

For frequencies other than the resonant frequency we write "ω" instead of "ω_0." Thus, $X_L = \omega L$ and $X_C = 1/\omega C$ will denote reactance values at any frequency ω, while the notations $X_{0L} = \omega_0 L$ and $X_{0C} = 1/\omega_0 C$ will denote reactance values at the resonant frequency.

Next, after a certain amount of trial and error, we find that the following algebraic manipulations produce some very useful relationships. Let us begin by writing

$$X_L - X_C = \omega L - \frac{1}{\omega C} = \frac{\omega\omega_0 L}{\omega_0} - \frac{\omega_0}{\omega\omega_0 C}$$

and thus, by eq. (401), we have that

$$(X_L - X_C) = \omega_0 L\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right) \tag{402}$$

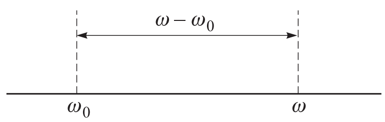
Next note that

$$\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} = \frac{\omega^2 - \omega_0^2}{\omega_0 \omega} = \left(\frac{\omega - \omega_0}{\omega_0}\right)\left(\frac{\omega + \omega_0}{\omega}\right) \tag{403}$$

In the above equation let

$$d = \frac{\omega - \omega_0}{\omega_0} \tag{404}$$

in which $\omega - \omega_0$ is the frequency difference between any frequency ω and the resonant frequency ω_0, as illustrated in the figure below.



Thus, in eq. (404) we see that d is the FRACTIONAL DEVIATION FROM RESONANCE for any given frequency ω. (For instance, if $\omega = 1.02\omega_0$, then $d = 0.02$; that is, ω is two percent greater than ω_0.)

In our work here, ω_0 will be the frequency of the "carrier wave," while ω will be the highest side-band frequency of importance in the modulated wave. Thus, to satisfactorily amplify a given modulated wave, the circuit of Fig. 232 should pass all frequencies in the range $\omega_0 \pm \omega$.

Returning now to eqs. (402) through (404), we see that

$$(X_L - X_C) = \omega_0 L d\left(\frac{\omega + \omega_0}{\omega}\right)$$

or, since (by eq. (404)) $\omega = d\omega_0 + \omega_0$, the above equation becomes

$$(X_L - X_C) = \omega_0 L d\left(\frac{d+2}{d+1}\right) \tag{405}$$

At this point let's pause to consider what *actual values* of d we might expect to encounter in practical work. In doing this, it should be noted that the double-tuned circuit of Fig. 232 is especially suited for use as an "intermediate frequency" (IF) amplifier in AM and FM receivers. In this regard consider, for example, the standard broadcast-band FM receiver. Here f_0 (the IF) is generally selected to be 10,700 kHz (10.7 MHz), with side bands extending to 100 kHz either side of f_0. Thus in this case the value of d is, by eq. (404),* equal to

$$d = (10{,}800 - 10{,}700)/10{,}700 = 0.01, \text{ approx.}$$

This illustrates that the value of d will normally be very much less than 1, and thus in practical engineering work eq. (405) can be written as

$$(X_L - X_C) = 2\omega_0 L d \tag{406}$$

* In "cycles/second" eq. 404 would be $d = (f - f_0)/f_0$.

Now, going back to eq. (400), replace $(X_L - X_C)$ with the right-hand side of the above equation. Next, in eq. (400), set $X_{\rm m} = \omega M = k\omega L$ (by eq. (371)) and also set $X_C = 1/\omega C$. Doing this, eq. (400) becomes

$$\bar{G} = \frac{-jg_{\rm m}kL/\omega C^2}{r^2 + k^2\omega^2L^2 - 4\omega_0^2L^2d^2 + j4r\omega_0 Ld} \tag{407}$$

Next, what is called the "Q" factor of an inductor coil is universally defined as the *ratio* of the REACTANCE of the coil to its RESISTANCE. If the coil appears in a resonant circuit, then we'll define the Q in terms of the resonant frequency of the circuit; thus

$$Q = \frac{\omega_0 L}{r} \tag{408}$$

where r is the total resistance in series with the coil, including any resistance the coil windings may have.

One reason it's convenient to work in terms of Q is because it's easy to find both the inductance of a coil and its Q by use of a standard piece of laboratory equipment called a "Q-meter."

Let us therefore write our equations in terms of Q; thus, noting that $r = \omega_0 L/Q$, eq. (407) becomes

$$\bar{G} = \frac{-jg_{\rm m}kL/\omega C^2}{\dfrac{\omega_0^2L^2}{Q^2} + k^2\omega^2L^2 - 4\omega_0^2L^2d^2 + \dfrac{j4\omega_0^2L^2d}{Q}} \tag{409}$$

We next might note that the second term in the denominator of the above equation can be written as

$$k^2\omega^2L^2 = k^2\omega^2\omega_0^2L^2/\omega_0^2 = k^2\omega_0^2L^2\left(\frac{\omega}{\omega_0}\right)^2$$

and, upon making this change, note that $\omega_0^2L^2$ factors from the denominator. Doing this, then multiplying the numerator and denominator by Q^2, eq. (409) becomes

$$\bar{G} = \frac{-jg_{\rm m}LkQ^2/\omega C^2}{\omega_0^2L^2[1 + k^2Q^2(\omega/\omega_0)^2 - 4Q^2d^2 + j4Qd]} \tag{410}$$

Now multiply the *denominator* by $1/\omega_0^2L^2$ and the *numerator* by $\omega_0^2C^2$ (which is permissible, because, by eq. (401), $1/\omega_0 L = \omega_0 C$). Doing this, eq. (410) becomes

$$\bar{G} = \frac{-jg_{\rm m}\omega_0 LQ^2(\omega_0/\omega)k}{1 + k^2Q^2(\omega/\omega_0)^2 - 4Q^2d^2 + j4Qd} \tag{411}$$

At this point let's pause to consider some relationships that are known to exist in practical applications of Fig. 232. To make the first point, note that, by eq. (404),

$$\frac{\omega}{\omega_0} = 1 + d$$

and therefore (see discussion following eq. (405)) it follows that, for almost all practical applications, it's perfectly permissible to write that $(\omega/\omega_0) = (\omega_0/\omega) = 1$; hence, for practical purposes eq. (411) becomes

$$\bar{G} = \frac{-jg_{\rm m}\omega_0 LkQ^2}{1 + k^2Q^2 - 4d^2Q^2 + j4dQ} \tag{412}$$

In the above it might seem, at first glance, that the entire denominator could be reduced to 1, since, in practical work, d and k are both very much less than 1. This, however, will not be the case because in practical work Q will generally be much greater than 1 (a value of $Q = 100$ is entirely possible). Thus products of the forms "dQ" and "kQ" can have significant values, and hence must not be dropped from the denominator.

In the above equation the voltage gain $\bar{G}$ is expressed in complex form. However, in most work it will be sufficient to know only how the *magnitude* of $\bar{G}$ varies with frequency. With this in mind, let us write eq. (412) in the form

$$|\bar{G}| = G = (g_m\omega_0 L)\frac{kQ^2}{\sqrt{(1 + k^2Q^2 - 4d^2Q^2)^2 + 16d^2Q^2}} \tag{413}$$

In the above, remember that d is the "fractional deviation of any frequency ω from the resonant frequency ω_0" (thus $d = 0$ for $\omega = \omega_0$). What we now wish to investigate is how the *value of k* affects the way in which G *varies* relative to the *value of d*. That is, how the *value of k* affects the shape of the curve of G versus d.

As an illustration of how G varies with d, let's first consider a specific value, after which we'll point out some practical, general conclusions regarding Fig. 232.

Let us take, as our example, the case for $Q = 50$ (a reasonable value). Then, for this value of Q (and omitting the constant multiplier $g_m\omega_0 L$), eq. (413) becomes

$$G = \frac{2500k}{\sqrt{(1 + 2500k^2 - 10{,}000d^2)^2 + 40{,}000d^2}} \tag{414}$$

in which we'll take d as the independent variable and k, the coefficient of coupling, as a parameter whose effect we wish to investigate.

In regard to k, it should be noted that what is called "critical coupling" is defined as being equal to $1/Q$. Thus, if "k_c" denotes critical coupling, we have that $k_c = 1/Q$. Hence, in the present example we have that $k_c = 1/50 = 0.02 = 2\%$.

In the above example it will be interesting to plot the curves of G *versus* d for several different values of the parameter k; let us select the values $k = 0.01$, $k = k_c = 0.02$, $k = 0.03$. To do this, we successively substitute into eq. (414) the chosen values of k, with the results shown in the table below, with final calculator values rounded to one decimal place.

d	Value of G for $k = 0.01$	Value of G for $k = 0.02$	Value of G for $k = 0.03$
0.000	20.0	25.0	23.1
±0.005	17.7	24.8	23.7
±0.010	12.4	22.4	24.9
±0.020	5.2	11.2	18.4
±0.030	2.6	5.4	9.0
±0.040	1.5	3.1	5.0

The above results are plotted in Figs. 239 and 240, in which d, on the horizontal axis, is understood to be multiplied by 10^{-2}.

In regard to the above, let's return to Fig. 232 and first suppose that the two coils are physically far apart, so that only very "loose coupling" exists between the coils. As we would expect, the voltage gain would be quite low in such a case. This would, for example, be the condition for the case of $k = 0.01$ in Fig. 239.

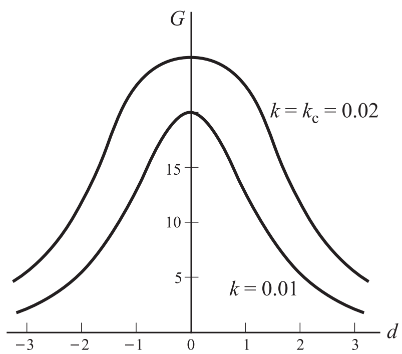


Fig. 239

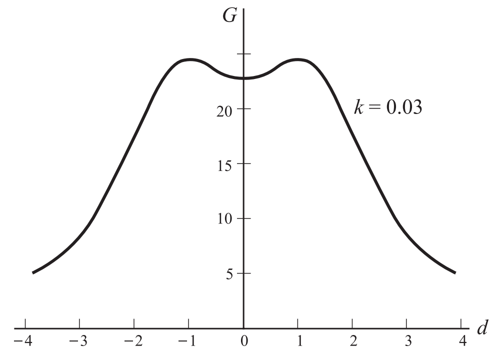


Fig. 240

Now, in order to increase the gain, we would naturally move the two coils closer to each other, thus increasing the value of k. However, as we continue to move the coils closer and closer together, a seemingly peculiar effect takes place, as follows.

At first, as the value of k increases, the value of the gain at resonance *also increases*, as we would expect. This trend, however, continues only until we reach the condition of CRITICAL COUPLING, at which point the gain at resonance has reached its *maximum possible value*; this would be the condition for the example case of $k = k_c = 0.02$ in Fig. 239. If, now, the value of k is increased BEYOND the value of k_c, we find that the gain at the normal center frequency *decreases*, while the ''gain versus frequency'' curve begins to show TWO SEPARATE RESONANT PEAKS, one below $d = 0$ and one above $d = 0$; an example of this condition is illustrated for the case of $k = 0.03$ in Fig. 240. In this state the circuit is said to be ''overcoupled.'' We thus find that, as the value of k increases (beyond the value of k_c), the gain at $d = 0$ continues to *decrease* while the separation between the two high peaks continues to *increase*. Hence for large values of k the ''G versus d'' curve would become highly distorted, so that the circuit of Fig. 232 would become useless.

Having said this, however, it should be noted that the use of a relatively SMALL AMOUNT of overcoupling (as in Fig. 240) can be used as a practical way to produce a relatively good type of ''band-pass'' circuit.

In regard to the last statement, compare the curves for $k = 0.02$ and $k = 0.03$ in Figs. 239 and 240. Note that, compared with $k = 0.02$, the somewhat overcoupled case of $k = 0.03$ produces (for most practical purposes) a nearly flat-topped gain curve from approximately $d = -1$ to $d = +1$. Also note that the gain falls off quite rapidly for frequencies beyond the region of the two peaks. Hence, by a proper choice of k, Fig. 232 can be made to serve as a reasonably good band-pass type of circuit.

This makes the circuit of Fig. 232 especially useful as an amplifier of high-frequency modulated carrier waves. This is because such a circuit will pass, with almost uniform gain, both the carrier wave and the necessary side-bands either side of the carrier, while effectively discriminating against possible nearby interfering signals.

In closing, the explanation, in words, for the existence of a double-peaked gain curve can be summarized, very briefly, as follows.

Taking Fig. 240 as an example, consider, first, the condition of the circuit in the neighborhood of $d = -1$. Since the frequency here is *less* than ω_0, it follows that X_C is greater than X_L in both the primary and secondary circuits. Hence the reactance reflected into the primary coil is *inductive* in nature,* and in an amount sufficient to increase the

* See ''discussion note'' given with solution to problem 187.

inductive reactance of the primary coil to a point where it resonates with the primary capacitor at a frequency *less* than ω_0. This causes increased voltage to appear across the primary load, and ultimately across the output capacitor on the secondary side at a frequency less than ω_0.

Now consider the circuit in the neighborhood of $d = +1$. Since the frequency here is greater than ω_0, it follows that X_L is greater than X_C in both the primary and secondary circuits. Hence the reactance reflected into the primary coil is *capacitive* in nature, and in an amount sufficient to decrease the inductive reactance of the primary coil to a point where it resonates with the primary capacitor at a frequency greater than ω_0. This again causes increased voltage to appear across the primary load, and ultimately across the output capacitor on the secondary side at a frequency greater than ω_0.

Lastly, consider the condition for $d = 0$ in Fig. 240. Here the frequency is ω_0, which is actually the basic resonant frequency of the circuit. This is because ω_0 is the only frequency at which the primary and secondary circuits are simultaneously resonant.

We have discovered, however, that for larger values of k or M the gain at ω_0 can be *less* than the gain at the other two resonant frequencies above and below ω_0. To see why this can be true recall, from eq. (387), that

$$\bar{Z}_{\text{ref}} = \omega^2 M^2 / \bar{Z}_2$$

where $\bar{Z}_2$ is the series impedance of the secondary circuit considered by itself. Since $\bar{Z}_2$ has its *minimum value* at resonance ($\bar{Z}_2 = r$ at ω_0), it follows that a comparatively large value of resistance can be reflected into the primary coil at resonance, so that the primary current, for larger values of M, could be less at ω_0 than at the other two resonant frequencies. Thus the voltage induced into the secondary circuit at ω_0 could be less than at the other two resonant frequencies, causing a reduced value of secondary current, with the final result of reduced voltage drop across the output capacitor.

Problem 198

Suppose two coils are wound in the same sense on a cylindrical coil form, as in Fig. 241. Let it be given that an instrument for the measurement of inductance is available, and that the coils have been found to have inductances of L_1 and L_2 henrys, as shown.

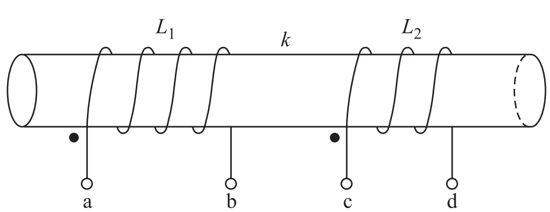


Fig. 241

In the above, it's also possible to find the value of the coefficient of coupling k by making two additional inductance measurements. The first measurement is made with the coils connected so that their magnetic fields AID each other, and the second is made with the coils connected so that their magnetic fields OPPOSE each other; let us denote these results by L_{aid} and L_{opp}.

Find, now, the equation for k in terms of all of the above measurements. (Note: in doing this, it will be convenient to imagine that a generator of V volts is applied to each of the two different conditions.)

10.6 The Ideal Iron-Core Transformer

In the following we'll be dealing, as usual, with rms sinusoidal voltages and currents.

In order for a transformer to operate effectively, the primary current must be able to induce an ADEQUATE MAGNITUDE OF VOLTAGE into the secondary coil. Thus, if $\bar{I}_1$ is the primary current, then (see eq. (379)) the magnitude of voltage induced into the secondary coil is equal to

$$V_2 = \omega M I_1$$

or, by eq. (371),

$$V_2 = \omega k \sqrt{L_1 L_2}\, I_1 \tag{415}$$

Fundamentally, the magnitude of V_2 is proportional to the RATE OF CHANGE of primary current and, for a given situation, the rate of change of primary current *increases* as the frequency increases. (See discussion following Fig. 128 in section 8.1.) Thus, at the higher frequencies (let us say above the audio range), it's not difficult to obtain an adequate value of V_2. Let us, therefore, now consider the situation at the low frequencies such as, for example, at the power-line frequency of 60 Hz.

Inspection of eq. (415) shows that, in order to produce an adequate value of V_2, we could take one or more of the following steps:

1. increase the primary current I_1;
2. increase the value of the coefficient of coupling k;
3. increase the inductances of the primary and secondary coils.

Possibility (1) must be avoided if at all possible, one reason being to prevent excessive power loss in the primary coil.

Next, concerning possibility (2), we should strive to make the value of the coefficient of coupling, k, as close as possible to its *maximum theoretical value of 1*. This is done by winding the primary and secondary coils as close together as possible.

Finally, and this is where the *iron core* comes into the picture, we must utilize possibility (3) and make the *inductances* of the primary and secondary coils as large as possible.

From a practical standpoint, however, it would be difficult to get the large inductance values required for operation at low frequencies without using an iron core. This is because, in order to get the large number of flux linkages required for large L, we would have to use very large coils containing perhaps many hundreds of turns of wire, which would, in itself, introduce other problems and practical difficulties.

These difficulties are avoided by winding the primary and secondary coils on a common, closed "iron core," as illustrated in Fig. 242.

The "iron core" of Fig. 242 is actually constructed of thin sheets or "laminations" of silicon steel bolted tightly together. (The laminated construction greatly reduces energy losses generated in the iron core by the rapidly changing magnetic flux.) In electrical diagrams, the presence of an iron core is indicated by drawing several vertical lines between the primary and secondary coils, as shown in Fig. 243. The use of the iron core with its high value of permeability allows the production of a large amount of flux with only a relatively few turns of wire. Hence, we can greatly reduce the number of turns and still have enough flux linkages to produce the large values of L_1 and L_2 required for operation at the low frequencies.

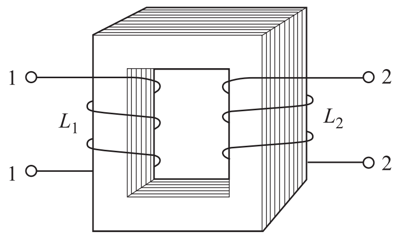


Fig. 242

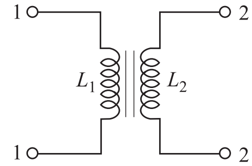


Fig. 243. "Schematic" symbol for Fig. 242.

With these points in mind, the IDEAL IRON-CORE TRANSFORMER is defined as a theoretical transformer having the following characteristics.

(a) ZERO ENERGY LOSSES, which means that the primary and secondary coils have zero resistance, and no energy loss in the iron core.

(b) UNITY COEFFICIENT OF COUPLING, that is, $k = 1$; hence, $M = k\sqrt{L_1 L_2} = \sqrt{L_1 L_2}$.

(c) The *inductive reactances* X_1 and X_2 of the primary and secondary coils are "infinitely great" in value, but, for any given transformer, the RATIO of X_1 to X_2 is a constant finite number a; thus

$$a = X_1/X_2$$

Now let a *finite load impedance* of $\bar{Z}_L = R + jX$ ohms be connected to the output terminals of an "ideal" transformer, as in Fig. 244.

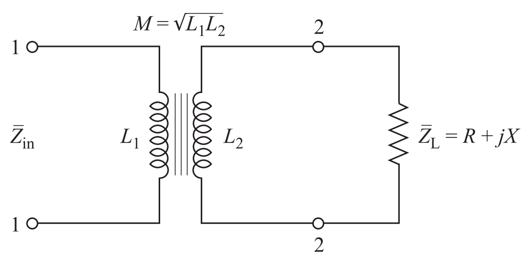


Fig. 244

Thus we have

$$\omega L_1 = X_1 = \text{ inductive reactance of primary coil}$$
$$\omega L_2 = X_2 = \text{inductive reactance of secondary coil}$$
$$\omega M = \omega\sqrt{L_1 L_2} = \sqrt{\omega L_1 \omega L_2} = \sqrt{X_1 X_2}$$

Hence, by eq. (387) in section 10.3, the impedance $\bar{Z}_{in}$ seen looking into terminals $(1, 1)$ in Fig. 244 is equal to

$$\bar{Z}_{in} = jX_1 + \frac{X_1 X_2}{R + j(X_2 + X)} = jaX_2 + \frac{aX_2^2}{R + j(X_2 + X)} \quad (\text{since } X_1 = aX_2)$$

Now, upon rationalizing and collecting like terms, you should find that the above relationship becomes

$$\bar{Z}_{\text{in}} = aR\left[\frac{X_2^2}{R^2 + (X_2 + X)^2}\right] + ja\left[\frac{R^2 X_2 + XX_2^2 + X^2 X_2}{R^2 + (X_2 + X)^2}\right] \tag{416}$$

Now, in accordance with condition (C), we allow X_2 to become infinitely great in value. Doing this, and holding everything else constant, notice that both the real and imaginary parts of eq. (416) take the form "infinity over infinity," ∞/∞. Since "infinitely great" is not a specific value, ∞/∞ is said to be an "indeterminant" form.

This, however, does *not* mean that a definite, specific answer can never be found in such a case. If an answer does exist, it can sometimes be found by first merely changing the *form* of the given expression and *then* allowing the variable to become infinitely great. This is true for the case of eq. (416) as follows.

First divide the numerators and denominators in both of the fractions by X_2^2. Doing this, and noting that, algebraically,

$$(X_2 + X)^2/X_2^2 = (1 + X/X_2)^2$$

eq. (416) becomes

$$\bar{Z}_{\text{in}} = aR\left[\frac{1}{R^2/X_2^2 + (1 + X/X_2)^2}\right] + ja\left[\frac{R^2/X_2 + X + X^2/X_2}{R^2/X_2^2 + (1 + X/X_2)^2}\right]$$

Now let X_2 become infinitely great, $X_2 \to \infty$. When this happens, note that $1/X_2$ approaches the value *zero*, that is,

$$\text{when } X_2 \to \infty, \text{then } 1/X_2 \to 0$$

and therefore (remembering that R and X, in Fig. 244, have only finite values) we see that, for the IDEAL CASE of infinitely great X_2, $\bar{Z}_{\text{in}}$ becomes equal to

$$\bar{Z}_{\text{in}} = a(R + jX) = a\bar{Z}_{\text{L}} \tag{417}$$

Since a is a real number, eq. (417) shows that the ideal transformer is an impedance-matching device that changes only the MAGNITUDE of the load impedance; that is, the phase angle of $\bar{Z}_{\text{in}}$ is the same as the phase angle of $\bar{Z}_{\text{L}}$.

While it's impossible, of course, to build a true ideal transformer, a well-designed iron-core transformer will come very close to being ideal at low frequencies. For example, a high-quality audio transformer, designed for use in the amplifier of a high-fidelity sound system, may well have both an efficiency and a coefficient of coupling in excess of 99%. So, in many cases an actual transformer can be considered to be ideal, for practical engineering purposes.

Let's continue now, first with a brief mention of what is called the "magnetizing current" of an iron-core transformer.

In doing this, we'll refer to Fig. 244 and assume that, for practical purposes, the transformer can be considered to be "ideal." As we have seen, this means that the values of the reactances X_1 and X_2 must be very great. Since the primary and secondary coils are to have only relatively small numbers of turns of wire, this means that, to satisfy the requirement of large values of X_1 and X_2, the iron core must possess a very high value of relative permeability.*

---

* Consider a coil of N turns. If L_{iron} is the inductance of the coil with an *iron* core and L_{air} is the inductance with *air* as core, we'll define "relative permeability"' of iron to air to be the ratio of L_{iron} to L_{air}. This ratio can be in the order of 1000 to 1 for silicon steel. Thus, for a given number of N turns, the reactance of an iron-core coil could be 1000 times the reactance of the corresponding air-core coil.

Now suppose, in Fig. 244, that a sinusoidal voltage of V reference volts is applied to the input terminals with the *secondary side open-circuited*. Then $\bar{Z}_{\text{in}} = jX_1$, and the input current that flows in this condition is called the transformer "magnetizing current," which we'll denote as $\bar{I}_{\text{m}}$. By Ohm's law,

$$\bar{I}_{\text{m}} = V/jX_1 = -j(V/X_1)$$

showing that, since X_1 is very large, the magnetizing current of a high-quality iron-core transformer is a VERY SMALL CURRENT, LAGGING THE APPLIED VOLTAGE V BY 90 DEGREES. Recall that true power P in an ac circuit is, by eq. (227) of Chap. 8, equal to $P = VI \cos \phi$, where ϕ is the phase angle between V and I. Since for the above condition $\phi = 90^\circ$, and since $\cos 90^\circ = 0$, we see that the magnetizing current consumes NO ENERGY, and is thus spoken of as the "wattless magnetizing current."

In this regard, let us note that in a high-quality iron-core transformer the amount of *magnetic flux* in the iron core, and the amount of magnetizing current, both remain very nearly *constant* in value, independent of the amount of load current drawn by Z_{L}. This is because when alternating current flows in the secondary winding it tends to oppose or "buck" the alternating flux produced in the core by the primary current (this is in accordance with Lenz's law, section 7.4). This reduction of alternating flux causes a lower counter emf to be induced into the primary coil, which at once allows more current to flow into the coil, bringing the flux back up to its previous level.

To continue on, let us note that the important eq. (417) is expressed in terms of the constant a, defined as

$$a = \frac{X_1}{X_2} = \frac{\omega L_1}{\omega L_2} = \frac{L_1}{L_2}$$

which shows that eq. (417) is basically expressed in terms of the ratio of the *inductances* of the primary and secondary coils.

Actually, however, in practical work (dealing with iron-core transformers only) it's much more convenient to deal with TURNS RATIO than inductance ratio. The turns ratio T of a transformer is defined as

$$T = \frac{\text{number of turns of wire on primary coil}}{\text{number of turns of wire on secondary coil}} = \frac{N_1}{N_2}$$

It thus follows that, in order to express eq. (417) in terms of *turns ratio* we must know the relationship that exists between the *inductance* L of a coil and the *number of turns* N the coil has. Fortunately, for the ideal case of $k = 100\%$ it is known that *the* INDUCTANCE *of an ideal coil is proportional to the* SQUARE *of the number of turns*, that is,

$$L = k'N^2$$

where k' is a constant of proportionality.* Thus, since the two similarly constructed coils of Fig. 244 would have the same value of k', it would be true that

$$L_1 = k'N_1^2 \quad \text{and} \quad L_2 = k'N_2^2$$

and hence, upon substituting these values in the above equation for a, we have that

$$a = N_1^2/N_2^2 = (N_1/N_2)^2 = T^2$$

* See note 26 in Appendix.

and thus eq. (417) can be written in the more practical form

$$\bar{Z}_{\text{in}} = T^2(R + jX) = T^2\bar{Z}_{\text{L}} \tag{418}$$

where T is the ratio of PRIMARY TURNS TO SECONDARY TURNS.

Another important fact can be deduced as follows. Since $k = 100\%$ in the ideal case, this means that the *same value of flux*, created by the magnetizing current, links *every turn* of the transformer. Hence the "volts induced per turn" is the same on both the primary and secondary sides. Thus, if V_1 is the voltage across the N_1 primary turns and V_2 is the voltage across the N_2 secondary turns, then

$$\frac{V_1}{N_1} = \frac{V_2}{N_2}$$

that is,

$$\frac{V_1}{V_2} = \frac{N_1}{N_2} = T \tag{419}$$

showing that, in an ideal transformer, the VOLTAGE RATIO is equal to the TURNS RATIO.

It is also true that the "power input" and the "power output" of any type of transformer are equal to $P_{\text{in}} = V_1 I_1 \cos\phi_1$ and $P_{\text{out}} = V_2 I_2 \cos\phi_2$. In an ideal transformer, however, $\phi_1 = \phi_2$ and hence, in an ideal transformer,

$$V_1 I_1 = V_2 I_2$$

that is,

$$\frac{V_1}{V_2} = \frac{I_2}{I_1} \tag{420}$$

Problem 199

A load impedance of $(3 - j5)$ ohms is connected to the secondary terminals of an ideal transformer of turns ratio 4 to 1, primary to secondary. A constant value of 240 volts rms is applied to the primary terminals. Find the following:

(a) impedance seen looking into primary terminals,
(b) secondary current,
(c) power to load.

Problem 200

A transformer has a primary inductance of L_1 henrys and a secondary inductance of L_2 henrys, with coefficient of coupling k. Assuming negligible winding resistance, show that, if the secondary terminals are shorted together, the impedance seen looking into the primary terminals would be equal to

$$\bar{Z}_{\text{in}} = j\omega L_1(1 - k^2) \text{ ohms}$$

Problem 201

An iron-core power transformer has primary inductance of 4 henrys and secondary inductance of 3 henrys, with 99.5% coefficient of coupling. The windings have negligible resistance. If the transformer is connected to a 120 volt, 60 Hz power line, how much line current would theoretically flow if the secondary terminals were accidentally shorted together?

10.7 The Three-Phase Power System. Introduction

Let us begin with the ordinary single-generator, two-wire, ac circuit of Fig. 245.

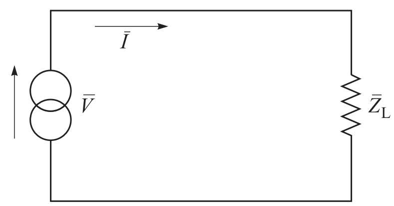


Fig. 245

We'll refer to this as a "single-phase" circuit, in which the basic equations for current and power are

$$\bar{I} = \bar{V}/\bar{Z}_L \quad \text{and} \quad P = VI\cos\phi$$

where V and I denote the magnitudes of the rms vector quantities $\bar{V}$ and $\bar{I}$, and where ϕ is the phase angle between the $\bar{V}$ and $\bar{I}$ vectors.

The single-phase circuit of Fig. 245 is, of course, very basic and much used in low-power applications. It is, however, not well-suited for the generation and transmission of large amounts of power, nor for the operation of large industrial-type ac motors. Thus, instead of the simple single-phase circuit of Fig. 245, almost all commercial electric power is generated and transmitted using what is called the "three-phase" system.*

There are important reasons for this. One is that the overall generation and transmission efficiency of three-phase systems is considerably higher than that of single-phase systems.

Another reason (as we'll show later on) is the fact that the INSTANTANEOUS POWER in a balanced three-phase system is *constant*, which is completely unlike the pulsating form of power in a single-phase system. This is an important advantage in the operation of high-horsepower ac motors. Also, in regard to the much-used single-phase system of Fig. 245, there are *three* such single-phase circuits available in a three-phase circuit.

Three-phase power is produced by a "three-phase generator," which can basically be described as follows.

A three-phase generator fundamentally consists of three separate but identical SINGLE-PHASE GENERATORS rigidly attached to a common shaft. The three generators produce EQUAL MAGNITUDES OF RMS VOLTAGE of the same frequency, but the three sets of windings are positioned on the shaft so that there is a PHASE DISPLACEMENT OF 120 DEGREES between the three single-phase voltage waves. The three separate generators are then connected together to form ONE COMPLETE, SYMMETRICAL, SINUSOIDAL THREE-PHASE GENERATOR.

In regard to the last statement, let us note that the three component generators will be connected together to form either a "Y-connected" generator or "Δ-connected" (delta-connected) generator. These two basic generator connections are shown in Figs. 246 and

* FREQUENCY is understood to be 60 cycles/second (60 Hz), $\omega = 2\pi f = 377$ rad/sec.

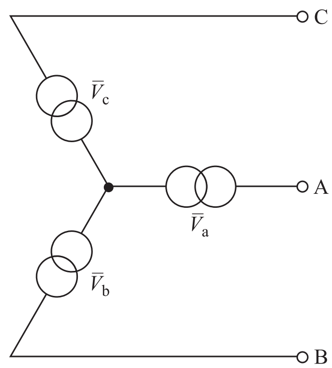


Fig. 246. "Y" connection.

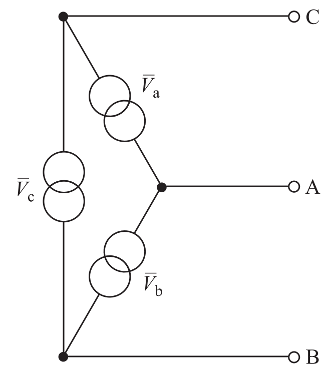


Fig. 247. "Δ" connection.

247, in which $\bar{V}_a$, $\bar{V}_b$, and $\bar{V}_c$ are the three single-phase voltages which, from now on, will be called the PHASE VOLTAGES.

In the above, A, B, and C denote the OUTPUT TERMINALS of the three-phase generators. These three terminals will be connected, by means of a three-wire transmission line, to a three-phase load. A three-phase system is thus basically a *three-wire* system, driven by three interconnected single-phase generators of the same frequency and same rms voltage but with phase differences of 120 degrees. A three-phase generator satisfying these conditions is said to be a BALANCED generator.

10.8 Y-Connected Generator; Phase and Line Voltages

In all of our work we'll assume a Y-connected type of generator, since this is normally the connection used in three-phase power generation, and we'll assume the generator to be completely "balanced," unless specifically stated otherwise. (The load impedance, however, can be of either the Y-type or the Δ-type.)

This is illustrated in Fig. 248, in which a balanced Y-connected generator is connected to a balanced three-phase load. (The load is said to be "balanced" because all three

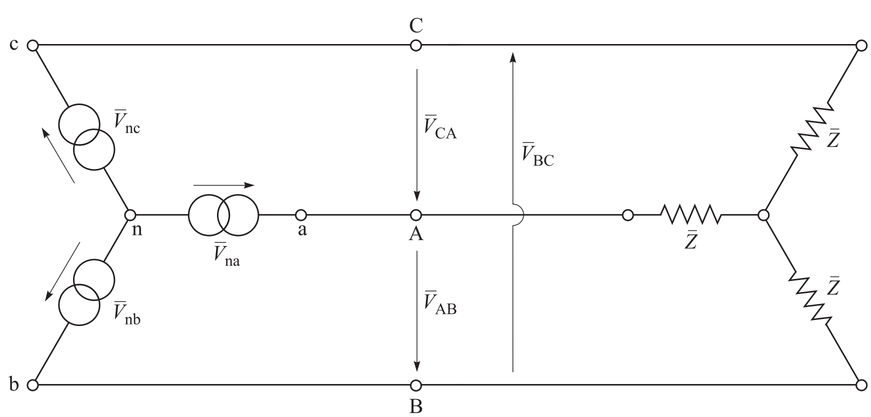


Fig. 248

impedances are given to have the same value of $\bar{Z}$ ohms.) In the figure we've used, for the purpose of comparison, both the standard "arrow" notation and the "double-subscript" notation to denote the positive direction of the rms voltage and current vectors relative to some particular "reference vector."*

In this section we wish to find the relationships between the PHASE VOLTAGES and the LINE VOLTAGES, the "line voltage" being the voltage between any two output lines (shown as $\bar{V}_{AB}$, $\bar{V}_{BC}$, and $\bar{V}_{CA}$ in the figure).

With this in mind, let us concentrate our attention on the generator end of the figure, paying special attention to the subscript notation, in which SMALL subscript letters denote PHASE voltages and LARGE subscript letters denote LINE voltages.

Let us take the junction point "n" as the common "reference point" in the system. The three individual vector phase voltages are then given with respect to the point n. Thus, if we take the phase voltage $\bar{V}_{na}$ to be the "reference vector," then let us agree that, by definition, we have

$$\bar{V}_{na} = V_{na}\underline{/0^\circ} \qquad \bar{V}_{nb} = V_{nb}\underline{/-120^\circ} \qquad \bar{V}_{nc} = V_{nc}\underline{/-240^\circ} \tag{421}$$

where

$$V_{na} = V_{nb} = V_{nc} = V_p$$

because it is given that the phase voltages all have EQUAL MAGNITUDES. Thus the relationships in eq. (421) can be written as

$$\bar{V}_{na} = V_p\underline{/0^\circ} \qquad \bar{V}_{nb} = V_p\underline{/-120^\circ} \qquad \bar{V}_{nc} = V_p\underline{/-240^\circ} \tag{422}$$

where V_p is the MAGNITUDE of the phase voltages.

Now, in the figure, imagine the junction point n to be the origin of the complex plane, and that V_{na} lies on the positive x-axis. Then the quantities in eq. (422) could be written in the complex rectangular form $V_p(\cos\theta + j\sin\theta)$; thus, using degrees, and remembering that $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$,

$$\bar{V}_{na} = V_p(\cos 0 + j\sin 0) = V_p \tag{423}$$

$$\begin{aligned}\bar{V}_{nb} &= V_p[\cos(-120) + j\sin(-120)]\\ &= V_p(\cos 120 - j\sin 120) = (-0.5 - j0.8660)V_p\end{aligned} \tag{424}$$

$$\begin{aligned}\bar{V}_{nc} &= V_p[\cos(-240) + j\sin(-240)]\\ &= V_p(\cos 240 - j\sin 240) = (-0.5 + j0.8660)V_p\end{aligned} \tag{425}$$

Let us now look VERY CAREFULLY at Fig. 248, beginning with the two lines A and B, shown again in Fig. 249.

In Fig. 249, note that $\bar{V}_{AB}$ is the voltage drop from line A to line B. If, now, we choose to start at A and trace around the loop in the cw sense (following the usual rule of setting the vector sum of the voltage drops equal to the vector sum of the generator voltages), we have that

$$\bar{V}_{AB} = -\bar{V}_{nb} + \bar{V}_{na}$$

Then, upon making use of eqs. (424) and (423), you should find that

$$\bar{V}_{AB} = (1.5 + j0.8660)V_p \tag{426}$$

* See note 27 in Appendix.

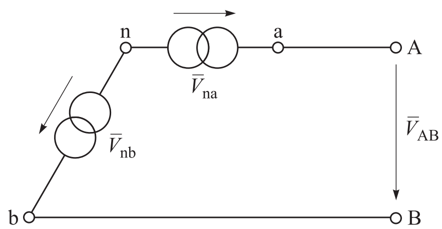


Fig. 249

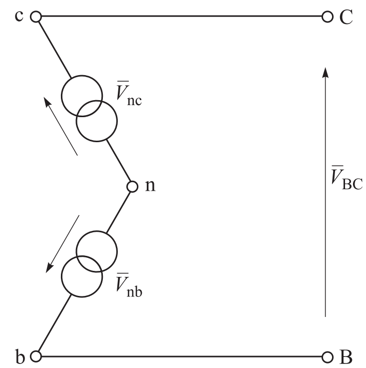


Fig. 250

or

$$\bar{V}_{AB} = 1.732 V_p \underline{/30^\circ} \tag{427}$$

in which let us note that $1.732 = \sqrt{3}$.*

Next, let's consider the loop formed by lines B and C, as shown in Fig. 250.

Note that $\bar{V}_{BC}$ is the voltage drop from line B to line C, as shown. If, now, we start at B and trace around the loop in the ccw sense, we have that

$$\bar{V}_{BC} = -\bar{V}_{nc} + \bar{V}_{nb}$$

and thus, upon making use of eqs. (425) and (424), we find that

$$\bar{V}_{BC} = -j1.732 V_p \tag{428}$$

or

$$\bar{V}_{BC} = 1.732 V_p \underline{/270^\circ} \tag{429}$$

Lastly, let's consider the loop formed by lines C and A, as shown in Fig. 251.

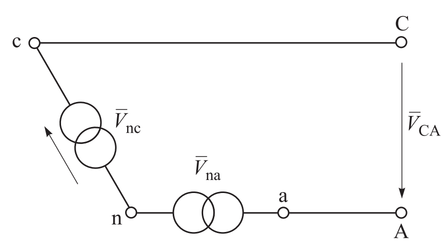


Fig. 251

Note that $\bar{V}_{CA}$ is the voltage drop from line C to line A; if, now, we start at C and trace around the loop in the cw sense, we have that

$$\bar{V}_{CA} = -\bar{V}_{na} + \bar{V}_{nc}$$

and hence, upon making use of eqs. (423) and (425), we have that

$$\bar{V}_{CA} = (-1.5 + j0.8660) V_p \tag{430}$$

* See note 28 in Appendix.

or (V_{CA} being in the 2nd quadrant)

$$\bar{V}_{CA} = 1.732 V_p \underline{/150^\circ} \tag{431}$$

Now let us summarize what our algebraic work has revealed about the Y-connected generator of Fig. 248. To begin, let's bring together the equations for the three LINE VOLTAGES; thus

$$\text{by eq. (427):} \quad \bar{V}_{AB} = 1.732 V_p \underline{/30^\circ} \tag{432}$$

$$\text{by eq. (429):} \quad \bar{V}_{BC} = 1.732 V_p \underline{/270^\circ} \tag{433}$$

$$\text{by eq. (431):} \quad \bar{V}_{CA} = 1.732 V_p \underline{/150^\circ} \tag{434}$$

The first point we wish to note is that inspection of the above three equations shows that the magnitude of LINE VOLTAGE produced by a balanced Y-connected generator is equal to 1.732 times the magnitude of the PHASE VOLTAGE; that is

$$V_L = 1.732 V_p \tag{435}$$

where V_L is the magnitude of the line voltage; thus

$$|\bar{V}_{AB}| = |\bar{V}_{BC}| = |\bar{V}_{CA}| = V_L \tag{436}$$

The second thing we wish to find is the complete VECTOR DIAGRAM showing the relationships among the various voltages in Fig. 248. In doing this, let us remember that the phase voltage $\bar{V}_{na}$ is the *reference vector* in Fig. 248.

Let us therefore begin with the vector diagram for the PHASE VOLTAGES, which is the vector diagram representation of eq. (421), as shown in Fig. 252.

Now, to complete our diagram, all we need do is add the LINE VOLTAGE vectors to Fig. 252. This can be done by noting the following facts.

First, by eq. (432), $\bar{V}_{AB}$ "leads" the reference vector $\bar{V}_{na}$ by 30°.

Next, noting that $\underline{/270^\circ} = \underline{/-90^\circ}$, eq. (433) shows that $\bar{V}_{BC}$ "lags" $\bar{V}_{na}$ by 90°.

Lastly, noting that $\underline{/150^\circ} = \underline{/-210^\circ}$, eq. (434) shows that $\bar{V}_{CA}$ "lags" $\bar{V}_{na}$ by 210°.

Combining these facts with Fig. 252 gives the COMPLETE voltage vector diagram for Fig. 248, as shown in Fig. 253.

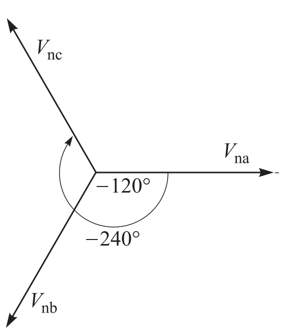


Fig. 252

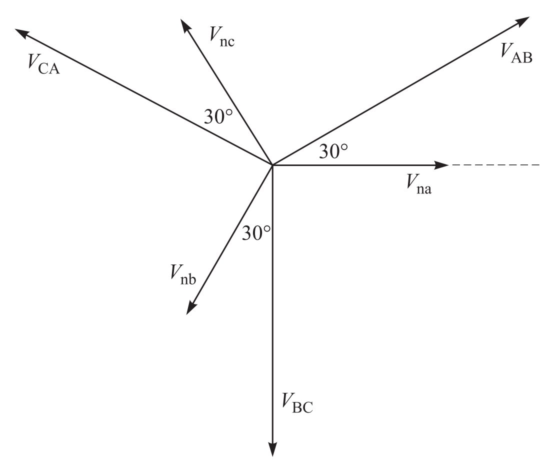


Fig. 253

Thus, as Fig. 253 shows, in a balanced Y-connected generator the line-voltage vectors "lead" the phase-voltage vectors by 30 degrees.

Let us note that the transmission of large blocks of power requires that transmission-line voltage be as high as possible. This is necessary to prevent excessive power loss in the line. Thus, commercial power-line voltages in the order of 120,000 volts rms are commonly used.

For several reasons, however, it's not practical to build power generators having such high output voltages. Thus, in the generation of large amounts of power, the generator will not usually be connected directly to the outgoing transmission line (as shown in Fig. 248). Instead, a relatively low value of generator voltage is used, which is then ''stepped up'' by a three-phase transformer to the desired high voltage for the transmission line. This is illustrated in Fig. 254, in which a balanced Y-connected generator is coupled to a transmission line through a ''delta-to-Y'' (Δ–Y) step-up transformer.

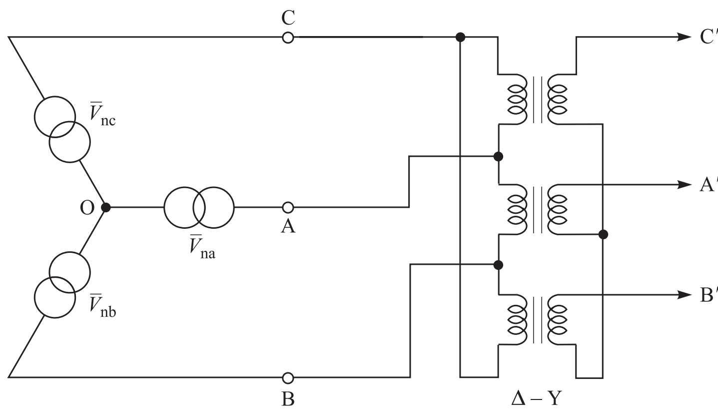


Fig. 254

In Fig. 254, A′, B′, and C′ denote the three wires of the outgoing transmission line. Actually, in diagrams such as the above, in which operation is at 60 Hz, it's understood that it will be necessary to use iron-core transformers. Hence, in practical drawings the iron-core symbol is omitted, and the above Δ–Y transformer would be drawn as shown in Fig. 255.

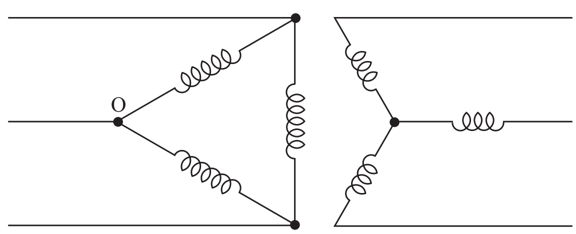


Fig. 255

In Figs. 254 and 255, note that the full generator output (the line voltage) is applied to each of the three Δ-connected primary coils. This voltage, after being stepped up by each individual transformer, then becomes the ''phase voltage'' on the Y-connected secondary side.

Problem 202

Given that there is no external load on the Δ-connected generator of Fig. 247, explain why there is no current flow around the closed-loop circuit formed by the three generators.

Problem 203

(a) In Fig. 248, find the phase voltage if the line voltage is 3300 volts.

(b) Suppose, in Fig. 254, that the transmission-line voltage is required to be 66,000 volts. If each of the three transformers has a turns ratio of 1-to-12 (primary turns to secondary turns), what value of generator phase voltage is required?

10.9 Current and Power in Balanced Three-Phase Loads

Here we take up the case in which a balanced Y-connected generator feeds a BALANCED three-phase load of $\bar{Z}$ ohms per phase, taking, first, the case of a balanced Y-connected load. Let us begin by redrawing Fig. 248, now adding some additional notation, as shown in Fig. 256.

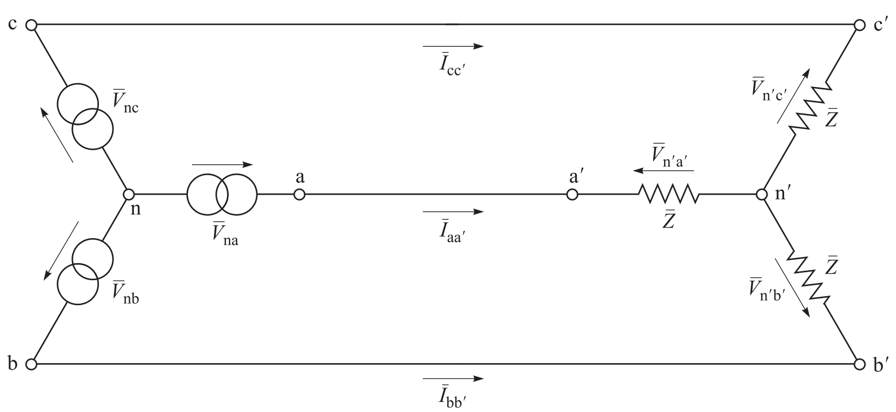


Fig. 256

It's apparent, by inspection, that the above can be a totally symmetrical, balanced system only if the *generator phase voltages* are equal to the *corresponding voltage drops in the load*; that is, only if

$$\bar{V}_{na} = \bar{V}_{n'a'} \qquad \bar{V}_{nb} = \bar{V}_{n'b'} \qquad \bar{V}_{nc} = \bar{V}_{n'c'}$$

Also, in the figure, the three "line currents" are denoted by $\bar{I}_{cc'}$, $\bar{I}_{aa'}$, and $\bar{I}_{bb'}$ as shown. Also note, from direct inspection of the figure, that the three line currents are actually equal to the "phase currents" (this is true only for a balanced Y-connected load). Since we're dealing with a balanced system, it follows that the line currents (and also the phase currents in this case) all have *equal magnitudes*; that is

$$|\bar{I}_{cc'}| = |\bar{I}_{aa'}| = |\bar{I}_{bb'}| = I_L = I_p \tag{437}$$

Next, the POWER, P, produced in the above balanced Y-connected load can be found as follows. First, as before, let V_p be the equal magnitudes of the three phase voltages. Then, since V_p and I_p denote the magnitudes of the rms voltages and currents *in each of the* *three load impedances*, it follows (from section 8.5) that the POWER P_p produced *in each of the three impedances* is equal to

$$P_p = V_p I_p \cos\phi \tag{438}$$

and thus the TOTAL POWER P_T produced in *all three* impedances in Fig. 256 is equal to

$$P_T = 3V_p I_p \cos\phi \tag{439}$$

where $\cos\phi$ is the same "power factor" of each of the three equal impedances. Or since, by eq. 435,

$$V_p = V_L/\sqrt{3}$$

and also since,

$$I_p = I_L$$

eq. (439) can also be written as

$$P_T = \sqrt{3}V_L I_L \cos\phi \tag{440}$$

which gives the total power produced in the balanced Y-connected system of Fig. 256 in terms of *line voltage and line current*.

Problem 204
In Fig. 256, suppose the generator phase voltage is 330 volts and $\bar{Z} = 15 + j9$ ohms. Find the total power output of the generator. (*Answer*: 16,014.03 watts)

Problem 205
In problem 204, show that the line currents lag the line voltages by approximately 61°.

Next, suppose the load in Fig. 256 were *delta-connected* instead of Y-connected. In such a case the situation at the load-end of the line would be as shown in Fig. 257, where $\bar{V}_{AB}$, $\bar{V}_{BC}$, and V_{CA} denote the three *line voltages* (as in Fig. 248). Also, let us denote the three *line currents* by $\bar{I}_A$, $\bar{I}_B$, and $\bar{I}_C$, as shown.

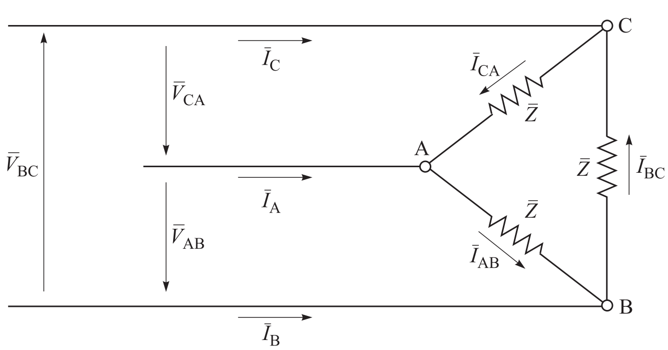


Fig. 257

Let $\bar{I}_{AB}$, $\bar{I}_{BC}$, and $\bar{I}_{CA}$ denote the three *phase currents*. Also let $\bar{V}_{AB}$ be the reference vector, and let ϕ be the phase angle between the phase voltages and phase currents, as shown in Fig. 258.

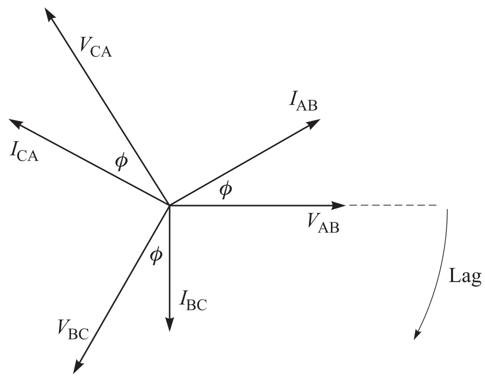


Fig. 258

Now, in regard to Fig. 257, the power PER PHASE is equal to

$$P_{\mathrm{p}} = V_{\mathrm{p}} I_{\mathrm{p}} \cos\phi$$

hence,

$$\text{TOTAL POWER} = P_{\mathrm{T}} = 3V_{\mathrm{p}} I_{\mathrm{p}} \cos\phi \tag{441}$$

as in eq. (439), where V_{p} and I_{p} are the magnitudes of the phase voltages and phase currents and $\cos\phi$ is the "power factor" of each of the three equal impedances $\bar{Z}$.

Inspection of Fig. 257 shows, however, that the *line voltage* is EQUAL to the *phase voltage* in a Δ-connected load. Thus eq. (441) can be written as

$$P_{\mathrm{T}} = 3V_{\mathrm{L}} I_{\mathrm{p}} \cos\phi \tag{442}$$

where V_{L} is the magnitude of line voltage. In the equation, however, we'd like also to have the current expressed in terms of *line* current; this can be done as follows.

Consider (for example) junction point A in Fig. 257; by Kirchhoff's current law, the current equation at A is equal to

$$\bar{I}_{\mathrm{A}} + \bar{I}_{\mathrm{CA}} - \bar{I}_{\mathrm{AB}} = 0$$

thus,

$$\bar{I}_{\mathrm{A}} = \bar{I}_{\mathrm{AB}} - \bar{I}_{\mathrm{CA}} \tag{443}$$

Now, for simplicity, let's consider $\bar{I}_{\mathrm{AB}}$ as the reference vector (this will have no effect on the relative magnitudes of the phase and line currents). Then, since $\bar{I}_{\mathrm{CA}}$ lags $\bar{I}_{\mathrm{AB}}$ by 240°, and since the phase currents all have equal magnitudes, eq. (443) becomes (angles in degrees)

$$\begin{aligned}\bar{I}_{\mathrm{A}} &= (I_{\mathrm{p}}\underline{/0} - I_{\mathrm{p}}\underline{/-240}) = (I_{\mathrm{p}}\underline{/0} - I_{\mathrm{p}}\underline{/120}) \\ &= (\cos 0 + j\sin 0 - \cos 120 - j\sin 120) I_{\mathrm{p}}\end{aligned}$$

thus,

$$\bar{I}_{\mathrm{A}} = (1.5 - j0.8660) I_{\mathrm{p}}$$

or since, from inspection of Fig. 257,

$$|I_{\mathrm{A}}| = |I_{\mathrm{L}}|$$

the last equation becomes

$$|I_A| = I_L = \sqrt{3} I_p$$

hence, in magnitudes,

$$I_p = I_L/\sqrt{3}$$

thus eq. 442 becomes

$$P_T = \sqrt{3} V_L I_L \cos\phi \tag{444}$$

where

$$|I_A| = |I_B| = |I_C| = I_L$$

Thus, comparison of eqs. (440) and (444) shows that P_T is calculated the same way for either a Y-connected or a Δ-connected load.

In the above, be reminded that V_L and I_L are the rms values of line voltage and line current, and thus P_T is the total average power produced in a balanced three-phase load.

In section 10.7 we mentioned that the *total* INSTANTANEOUS POWER in a balanced three-phase system is *constant*. This interesting and important fact can be proved as follows. In Fig. 257, let v_a, v_b, and v_c denote the *instantaneous* values of the three sinusoidal line voltages; then, letting V' denote the three equal peak voltages, the instantaneous values of the three voltage waves are

$$\begin{aligned} v_a &= V' \sin\omega t \\ v_b &= V' \sin(\omega t - 120^\circ)* \\ v_c &= V' \sin(\omega t - 240^\circ) = V' \sin(\omega t + 120^\circ) \end{aligned}$$

Then, since the above voltages work into identical loads, the corresponding instantaneous currents would be, letting I' denote the three equal peak currents,

$$\begin{aligned} i_a &= I' \sin(\omega t + \phi) \\ i_b &= I' \sin(\omega t + \phi - 120^\circ) \\ i_c &= I' \sin(\omega t + \phi + 120^\circ) \end{aligned}$$

where ϕ is the phase angle between the voltage and current waves. Then the *total* INSTANTANEOUS POWER p is equal to

$$p = v_a i_a + v_b i_b + v_c i_c$$

which, upon making the above substitutions, becomes

$$\begin{aligned} p = V'I'((\sin\omega t)\sin(\omega t + \phi) &+ [\sin(\omega t - 120^\circ)]\sin(\omega t + \phi - 120^\circ) \\ &+ [\sin(\omega t + 120^\circ)]\sin(\omega t + \phi + 120^\circ)) \end{aligned}$$

Note that the above result seemingly says that p is a *function of time t*, that is, that p *varies* from instant-to-instant, thus contradicting the statement we made that p remains

* Since it's always understood that $\omega t = 2\pi f t$ radians, it's basically incorrect to write $\sin(\omega t - 120^\circ)$; instead, since $120^\circ = 120(\pi/180) = 2\pi/3$ rad., we should really write $\sin(\omega t - 2\pi/3)$. However, the particular operation above is such that there is no harm in writing 120° instead of $2\pi/3$ radians if we wish to do so.

constant, independent of time. If, however, you have the patience to carefully apply the following trigonometrical identities

$$\sin x \sin y = \tfrac{1}{2}[\cos(x-y) - \cos(x+y)] \quad \text{(note 25 in Appendix)}$$

and

$$\cos(x+y) = \cos x \cos y - \sin x \sin y \quad \text{(note 6 in Appendix)}$$

to the foregoing equation for p, you'll find that the expression actually reduces to

$$p = 1.5V'I'\cos\phi \tag{445}$$

which, since V', I', and ϕ are all *constants* in any given case, shows that p is *also constant* in any given case in any balanced three-phase system.

Problem 206
Show that p, in eq. (445), is EQUAL to P_T in eqs. (440) and (444).

10.10 The Unbalanced Case; Symmetrical Components

Let us begin by observing that a SINGLE plane vector is defined in terms of *two independent variables*, its MAGNITUDE and its ANGULAR POSITION relative to an agreed-upon reference axis.

The independent variables are also referred to as "degrees of freedom"; thus, a single plane vector is said to have "two degrees of freedom." Such a vector, $\bar{A} = A\underline{/h}$, is illustrated in Fig. 259, where A and h are the two degrees of freedom. (As always, "positive angles" are measured in the ccw direction from the reference axis.)

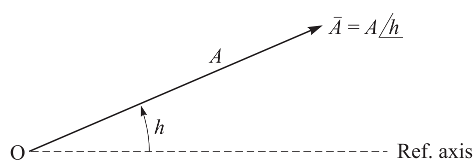


Fig. 259

Next consider a *balanced* set of three plane vectors. As we know, this is any set of three vectors having EQUAL MAGNITUDES and EQUAL PHASE DISPLACEMENTS.

Now let $\bar{A}_1$, $\bar{B}_1$, and $\bar{C}_1$ be such a balanced set, in which the equal phase displacement is 120°, and let us take $\bar{A}_1$ as the "reference vector," displaced an angle h from the reference axis, as illustrated in Fig. 260, where $|\bar{A}_1| = |\bar{B}_1| = |\bar{C}_1| = A_1 = B_1 = C_1$.

We are already familiar with the fact that the *vector sum* of such a balanced set of vectors is equal to *zero*.

In this regard, note that a balanced set of plane vectors has just TWO degrees of freedom, these being the common magnitudes of the vectors and the angular displacement h of the reference vector from the reference axis. Thus the common magnitude, $A_1 = B_1 = C_1$, and the reference angle h are the "two degrees of freedom" in Fig. 260.

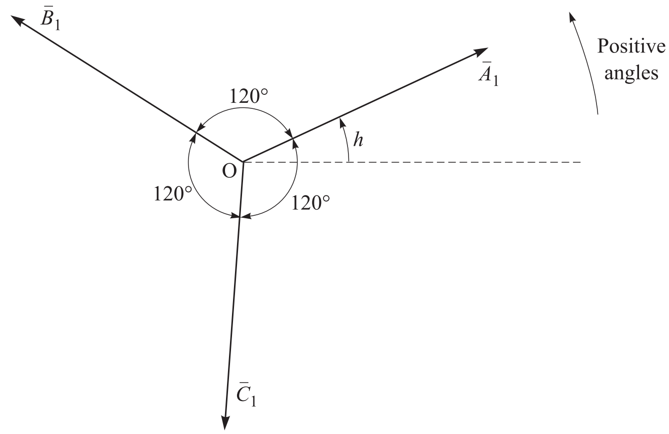


Fig. 260

Next, in Fig. 260, assuming the diagram to be drawn on the complex plane, note that*

$$\bar{B}_1 = \bar{A}_1 \epsilon^{j120} \quad \text{and} \quad \bar{C}_1 = \bar{A}_1 \epsilon^{j240}$$

where "120" and "240" are understood to be angles in degrees† (120° and 240°), and thus that general form of the algebraic equation for the balanced case of Fig. 260 can be written as

$$\bar{A}_1 + \bar{A}_1 \epsilon^{j120} + \bar{A}_1 \epsilon^{j240} = 0 \tag{446}$$

the right-hand side reflecting the fact that the vector sum of such a balanced set of vectors is zero.

Now consider an *unbalanced* set of three plane vectors $\bar{A}$, $\bar{B}$, and $\bar{C}$, such as is illustrated in Fig. 261, in which let

$$\bar{S} = \bar{A} + \bar{B} + \bar{C} \tag{447}$$

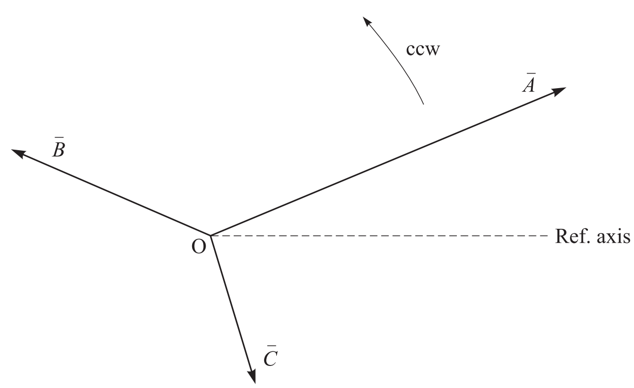


Fig. 261

* Let $\bar{V} = V\epsilon^{ja}$ be a vector quantity on the complex plane. Now multiply $\bar{V}$ by ϵ^{jb}, thus

$$\bar{V}\epsilon^{jb} = V\epsilon^{ja}\epsilon^{jb} = V\epsilon^{j(a+b)}$$

showing that *multiplying a vector* $\bar{V}$ *by* ϵ^{jb} rotates $\bar{V}$ through the angle b but does not change the magnitude of $\bar{V}$.

† See footnote in connection with eq. (159) in Chap. 6.

where $\bar{S}$ is the vector sum of the three vectors, in which $\bar{S}$ MAY or MAY NOT be equal to zero, depending upon the particular circumstances. It follows that such an unbalanced set of three vectors will, in general, have SIX degrees of freedom (two for each of the individual vectors).

As you would expect, unbalanced conditions sometimes do occur in practical three-phase work. Fortunately, the solution of such problems can be expedited by means of what is called "symmetrical components." This is an algebraic procedure based upon the fact an UNBALANCED set of three vectors can be expressed as *the sum of* THREE BALANCED SETS *of three vectors each*. The procedure is important because it allows the solution of a more difficult *unbalanced* problem in terms of the superposition of three easier *balanced* problems. In this regard, let us remark that the procedure is not only of great practical value but is also an interesting example of the application of the algebra of the complex plane to electric circuit problems. Let us begin our explanations as follows.

First, we've seen that a *balanced* set of plane vectors possesses just *two* degrees of freedom, while an *unbalanced* set of three plane vectors possesses, in general, *six* degrees of freedom.

Now, in regard to physical systems, it is a fundamental fact that the number of degrees of freedom must remain the *same* in any valid equivalent description of a system. It thus follows that it will, in general, require the sum of *three* balanced sets to replace *one* unbalanced set. OUR PROBLEM, therefore, is to *find three balanced sets* that are vectorially equivalent to A GIVEN UNBALANCED SET of three vectors.

In the method of symmetrical components the problem is solved by resolving the given unbalanced set into three balanced sets called the "positive sequence" set, the "negative sequence" set, and the "zero sequence" set. Let us first consider the positive sequence and negative sequence sets, which we'll define in connection with Figs. 262 and 263.

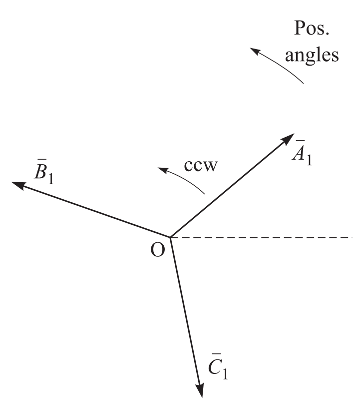


Fig. 262. POSITIVE SEQUENCE set.

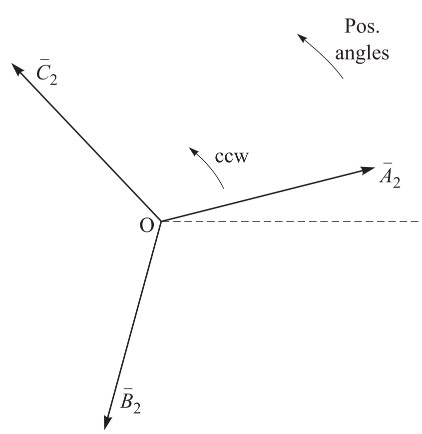


Fig. 263. NEGATIVE SEQUENCE set.

In the figures, note that both of the sets are *balanced*, meaning that $\bar{A}_1$, $\bar{B}_1$, $\bar{C}_1$ have equal magnitudes, also $\bar{A}_2$, $\bar{B}_2$, $\bar{C}_2$ have equal magnitudes, with the phase displacements between vectors being 120° in both sets.

As usual, the vectors can represent either rms values of sinusoidal voltages and currents or impedances. The vectors themselves, in any given case, always remain fixed in position, with "positive angles" measured in the ccw sense, as shown in the figures. Note that both sets are specified relative to the same common origin and reference line.

In the figures, note that the positive sequence set is identified by the subscript "1," while the negative sequence set is identified by the subscript "2." We'll use this method of identification IN ALL OF THE WORK that follows.

Since both sets are balanced, it follows that the vector sum in each case is zero; that is

$$\left.\begin{aligned} \bar{S}_1 &= \bar{A}_1 + \bar{B}_1 + \bar{C}_1 = 0 \\ \bar{S}_2 &= \bar{A}_2 + \bar{B}_2 + \bar{C}_2 = 0 \end{aligned}\right\} \tag{448}$$

Next, in regard to Figs. 262 and 263, the term "sequence" refers to the *order* in which the *letters* appear in the diagram in the *ccw sense*, as follows.

If (in going around the diagram in the ccw sense) the order of the letters is "ABC" we are said to have a "positive sequence" of vectors, but if the order is "ACB" we have a "negative sequence." Thus, in accordance with this definition, Fig. 262 is a positive sequence of vectors and Fig. 263 is a negative sequence set. The concept of "sequence" is important for the following reason.

First note that, by eq (448), the SUM of the two sequences can be written in the form

$$\bar{S}_1 + \bar{S}_2 = (\bar{A}_1 + \bar{A}_2) + (\bar{B}_1 + \bar{B}_2) + (\bar{C}_1 + \bar{C}_2) \tag{449}$$

in which the *left-hand side* is the sum of *two* BALANCED *sets* of three vectors each, and (as we'll show in problem 208) the *right-hand side* represents an UNBALANCED *set of three vectors*; thus eq. (449) shows that it's possible to represent an *unbalanced* set of three vectors (the right-hand side) as the sum of two *balanced* sets of three vectors each. This is the basic principle behind the method of "symmetrical components." The following two problems will clarify this point.

Problem 207

Let $\bar{A}_1$, $\bar{B}_1$, $\bar{C}_1$ and $\bar{A}_1'$, $\bar{B}_1'$, $\bar{C}_1'$ be two sets of "positive sequence" vectors (note the "1" subscripts). Using eqs. (446) and (449), show that the vector sum of the two sets is equivalent to a single *balanced* set of three vectors.

Problem 208

Let $\bar{A}_1$, $\bar{B}_1$, $\bar{C}_1$ be a positive sequence set of vectors, and $\bar{A}_2$, $\bar{C}_2$, $\bar{B}_2$ be a negative sequence set. Show that the vector sum of the two sets is a single *unbalanced* set of three vectors.

The above two problems show that an *unbalanced* set of three vectors can be represented as the sum of *two balanced sets* of three vectors each ONLY if one of the sets is a positive sequence set and the other a negative sequence set, where "positive sequence" and "negative sequence" *are defined in connection with Figs. 262 and 263* where, algebraically,

$$\bar{B}_1 = \bar{A}_1\epsilon^{j120} \quad \text{and} \quad \bar{B}_2 = \bar{A}_2\epsilon^{j240}$$
$$\bar{C}_1 = \bar{A}_1\epsilon^{j240} \quad \text{and} \quad \bar{C}_2 = \bar{A}_2\epsilon^{j120}$$

To bring out another important point let us begin by writing eq. (449) in the form

$$\left.\begin{aligned} \bar{A}_1 + \bar{A}_2 &= \bar{A}' \\ \bar{B}_1 + \bar{B}_2 &= \bar{B}' \\ \bar{C}_1 + \bar{C}_2 &= \bar{C}' \end{aligned}\right\} \tag{450}$$

in which $\bar{A}'$, $\bar{B}'$, and $\bar{C}'$ are the three components of the *unbalanced* set of vectors.

We must not conclude from the foregoing, however, that ALL unbalanced sets of three vectors can be expressed as the sum of just two balanced sets of vectors (one positive sequence set and one negative sequence set). This is because inspection of eqs. (448) and (449) shows that eq. (450) is valid only if

$$\bar{A}' + \bar{B}' + \bar{C}' = 0$$

that is, only if the sum of the vectors in the unbalanced set is equal to *zero*, a condition which MAY or MAY NOT be true in practical work. This indicates that a more general form of eq. (450) is needed, to cover cases in which the sum of the unbalanced vectors is *not* equal to zero. In this regard, a more general form of eq. (450) can be arrived at by thinking in terms of, "degrees of freedom," as follows.

We recall that, in general, an unbalanced set of three vectors, such as in Fig. 261, has "six degrees of freedom." So far, however, in Figs. 262 and 263 we have only *four* degrees of freedom. Hence, in addition to Figs. 262 and 263 we must, in order to include the most general unbalanced condition, add *one more set of three balanced vectors* to bring the degrees of freedom up to six.

This is done by defining what is called a "zero sequence" set of vectors which consists of *three* IDENTICAL *vectors*, meaning that the three vectors all have the *same magnitude of amplitude* at the *same angle k* with respect to the reference axis, as shown in Fig. 264.

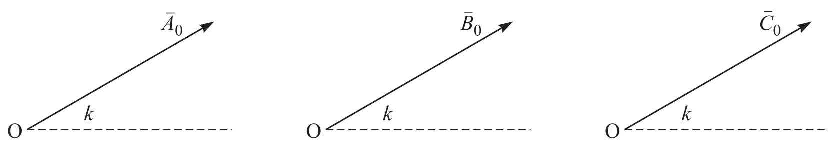


Fig. 264

Thus a "zero sequence" set of three vectors has "two degrees of freedom" and, using the subscript "0," is defined by writing that

$$\bar{A}_0 = \bar{B}_0 = \bar{C}_0 \tag{451}$$

Now, while a zero-sequence set of vectors is a balanced set (in accordance with the definition following Fig. 259), note that the *vector sum is* NOT *zero* (as it is when the vectors are 120° apart); instead, for a zero-sequence set we have that

$$\bar{A}_0 + \bar{B}_0 + \bar{C}_0 = 3\bar{A}_0 = 3\bar{B}_0 = 3\bar{C}_0 \tag{452}$$

Thus, in the three equations that comprise eq. (450), we now add $\bar{A}_0$ to both sides of the first equation, $\bar{B}_0$ to both sides of the second equation, and $\bar{C}_0$ to both sides of the third equation. Then, letting $\bar{A}' + \bar{A}_0 = \bar{A}$, and so on, eq. (450) becomes

$$\left.\begin{aligned} \bar{A}_1 + \bar{A}_2 + \bar{A}_0 &= \bar{A} \\ \bar{B}_1 + \bar{B}_2 + \bar{B}_0 &= \bar{B} \\ \bar{C}_1 + \bar{C}_2 + \bar{C}_0 &= \bar{C} \end{aligned}\right\} \tag{453}$$

where $\bar{A}$, $\bar{B}$, and $\bar{C}$, without subscripts, represent three components of an equivalent *unbalanced* set of three vectors having the required six degrees of freedom. Thus we've now expressed an *unbalanced* set of vectors, $\bar{A}$, $\bar{B}$, $\bar{C}$, in terms of the components of three *balanced* sets of vectors.

Now suppose the components of an *unbalanced* set are *known*, and we wish to *find* the values of the three equivalent *balanced* sets. That is, let the PROBLEM be: GIVEN the values of an unbalanced set, $\bar{A}$, $\bar{B}$, and $\bar{C}$, *find* the values of the components of the three *balanced* sets whose vector sum is equal to the given unbalanced set.

This would seem, offhand, to be a most difficult problem, but fortunately, because of the symmetry of balanced sets, it turns out not to be so hard after all. Let us proceed as follows.

First, carefully note, again, the set of equations given just prior to eq. (450); doing this, and also keeping eq. (451) in mind, note that eq. (453) becomes (using degrees)

$$\bar{A}_1 + \bar{A}_2 + \bar{A}_0 = \bar{A}$$

$$\bar{A}_1\epsilon^{j120} + \bar{A}_2\epsilon^{j240} + \bar{A}_0 = \bar{B}$$

$$\bar{A}_1\epsilon^{j240} + \bar{A}_2\epsilon^{j120} + \bar{A}_0 = \bar{C}$$

Now, for convenience, let

$$\bar{a} = \epsilon^{j120} \tag{454}$$*

then, also†

$$\bar{a}^2 = \epsilon^{j240}$$

and hence the foregoing three equations can be written in the easier-to-handle forms

$$\bar{A}_1 + \bar{A}_2 + \bar{A}_0 = \bar{A} \tag{455}$$

$$\bar{a}\bar{A}_1 + \bar{a}^2\bar{A}_2 + \bar{A}_0 = \bar{B} \tag{456}$$

$$\bar{a}^2\bar{A}_1 + \bar{a}\bar{A}_2 + \bar{A}_0 = \bar{C} \tag{457}$$

Note that we now have three simultaneous equations in *three unknowns*, $\bar{A}_1$, $\bar{A}_2$, and $\bar{A}_0$, the values of which can be found in several ways, including the method of elimination. Let us, however, use the more straightforward method of determinants, as follows.

You'll recall that the first step in the procedure is to find the value of the determinant "formed from the coefficients of the unknowns" which, as you should now verify from inspection of the above three equations, is equal to

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ \bar{a} & \bar{a}^2 & 1 \\ \bar{a}^2 & \bar{a} & 1 \end{vmatrix} = -\bar{a}(\bar{a}^3 - 3\bar{a} + 2)$$

Or, since $\bar{a} = \epsilon^{j120}$, then $\bar{a}^3 = \epsilon^{j360} = 1$, the above reduces to

$$\Delta = 3\bar{a}(\bar{a} - 1) \tag{458}$$

To continue on, in our solution of eqs. (455) through (457), let us next find the value of $\bar{A}_1$; thus

$$\bar{A}_1 = \frac{\begin{vmatrix} \bar{A} & 1 & 1 \\ \bar{B} & \bar{a}^2 & 1 \\ \bar{C} & \bar{a} & 1 \end{vmatrix}}{\Delta} = \frac{\bar{A}(\bar{a}^2 - \bar{a}) - \bar{B}(1 - \bar{a}) + \bar{C}(1 - \bar{a}^2)}{3\bar{a}(\bar{a} - 1)}$$

* Thus, $\bar{a} = \epsilon^{j120} = (\cos 120 + j \sin 120) = (-0.5 + j0.8660)$, in rectangular form.

† Using the basic relationship $(x^y)^n = x^{ny}$.

thus

$$\bar{A}_1 = \tfrac{1}{3}[\bar{A} + \bar{a}^{-1}\bar{B} - (1 + \bar{a}^{-1})\bar{C}]$$

Note, however, since

$$\bar{a} = \epsilon^{j120}, \text{then } \bar{a}^{-1} = \epsilon^{j120} = \epsilon^{j240} = \bar{a}^2 \quad (-120^\circ = +240^\circ)$$

thus

$$\bar{A}_1 = \tfrac{1}{3}[\bar{A} + \bar{a}^2\bar{B} - (1 + \bar{a}^2)\bar{C}] \tag{459}$$

then, since

$$\begin{aligned} -(1 + \bar{a}^2) &= -(1 + \epsilon^{j240}) = -(1 + \cos 240 + j\sin 240) \\ &= -0.5 + j0.8660 = \epsilon^{j120} = \bar{a} \end{aligned}$$

eq. (459) becomes

$$\bar{A}_1 = \tfrac{1}{3}(\bar{A} + \bar{a}^2\bar{B} + \bar{a}\bar{C}) \tag{460}$$

hence

$$\bar{B}_1 = \bar{a}\bar{A}_1 \tag{461}$$

and

$$\bar{C}_1 = \bar{a}^2\bar{A}_1 \tag{462}$$

the above three equations being the required components of the POSITIVE SEQUENCE *set of vectors* of Fig. 262.

Next, if you very carefully again apply the same procedures to eqs. (455) through (457), you should find that first

$$\bar{A}_2 = \tfrac{1}{3}(\bar{A} + \bar{a}\bar{B} + \bar{a}^2\bar{C}) \tag{463}$$

then

$$\bar{B}_2 = \bar{a}^2\bar{A}_2 \tag{464}$$

and

$$\bar{C}_2 = \bar{a}\bar{A}_2 \tag{465}$$

these being the components of the NEGATIVE SEQUENCE *vectors* of Fig. 263.

All that remains now is to find the value of $\bar{A}_0$, which can easily be done as follows. Let us, in eqs. (455) through (457), add up all the vectors BY COLUMNS; doing this, and noting that the first and second columns are both balanced vectors, we have that

$$0 + 0 + 3\bar{A}_0 = \bar{A} + \bar{B} + \bar{C}$$

and thus

$$\bar{A}_0 = \bar{B}_0 = \bar{C}_0 = \tfrac{1}{3}(\bar{A} + \bar{B} + \bar{C}) \tag{466}$$

which are the three equal components of the ZERO SEQUENCE set of Fig. 264.

Thus, GIVEN the components $\bar{A}$, $\bar{B}$, $\bar{C}$, of an UNBALANCED set of three plane vectors, eqs. (460) through (466) allow us to find the three balanced sets of vectors equivalent to the given unbalanced set.

Problem 209
Given the following unbalanced set of three plane vectors (in the same ccw sense of $\bar{A}$ to $\bar{B}$ to $\bar{C}$ as in Fig. 261)

$$\bar{S} = 15\underline{/0^\circ} + 9\underline{/100^\circ} + 24\underline{/215^\circ}$$

show that the given unbalanced set can be expressed as the sum of the following three sets of balanced vectors:

$$\text{positive sequence:} \quad \bar{S}_1 = 15.701(\epsilon^{-j16.301} + \epsilon^{j103.699} + \epsilon^{j223.699})$$
$$\text{negative sequence:} \quad \bar{S}_2 = 6.365(\epsilon^{j71.639} + \epsilon^{j311.639} + \epsilon^{j191.639})$$
$$\text{zero sequence:} \quad \bar{S}_0 = 7.920\epsilon^{j218.233}$$

10.11 Some Examples of Unbalanced Three-Phase Calculations

In a COMPLETELY BALANCED three-phase system the three line voltages have equal magnitudes, displaced from each other by 120°, and the three load impedances have equal values of $\bar{Z}$ ohms.

If these conditions do not exist, the three-phase system is said to be UNBALANCED. Thus, any one of the following three conditions of unbalance may be encountered.

1. The line voltages are unequal, either in magnitudes, or phase angles, or both.
2. The three load impedances are not all equal.
3. A combination of (1) and (2) is present.

To solve such problems by the method of symmetrical components, the basic procedure is to first resolve the *unbalanced* system into the sum of three *balanced* systems, then separately find the solution to each of the balanced systems. The final answer is then the SUM OF THE SEPARATE SOLUTIONS, in accordance with the principle of superposition (as in problem 50, Chap. 4, for example). (This assumes that, for practical purposes, the circuit is a "linear" system; see footnote in section 8.6.)

The following problems, in which Y-connected generators will feed Y-connected loads, will serve to illustrate the basic procedures.

Such problems can be worked either in terms of "line voltage" or "phase voltage," but in the problems here we've elected to specify phase voltages instead of line voltages (the relationships between the two were developed in section 10.8 and are summarized in Fig. 253, which shows the relation between phase and line voltages in a balanced Y-connected generator).

Problem 210
In Fig. 265, a zero-sequence voltage of $\bar{V}_0$ volts is applied, through a three-wire line, to a Y-connected load as shown.

Explain, from inspection of the figure, why no zero-sequence current can flow in any such three-wire Y-connected load.

Fig. 265

Problem 211

An unbalanced Y-connected three-phase generator is connected, by means of a three-wire line, to a balanced Y-connected resistive load of 12 ohms per phase. It is given that the generator phase voltages are, in volts, equal to

$$\bar{A} = 90\underline{/0^\circ} \quad \bar{B} = 72\underline{/120^\circ} \quad \bar{C} = 54\underline{/240^\circ}$$

in the same ccw sense $\bar{A}$ to $\bar{B}$ to $\bar{C}$ as in Fig. 261, with $\bar{A}$ the reference voltage. This is shown in schematic diagram form in Fig. 266.

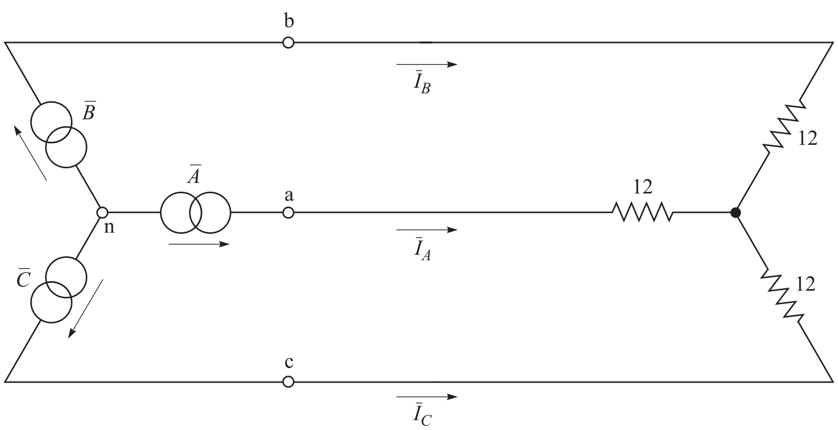


Fig. 266. The given three-phase system.

Using the method of symmetrical components, show that the magnitudes of the line currents, in amperes, are equal to

$$|\bar{I}_\mathrm{A}| = 6.764 \ (answer) \quad |\bar{I}_B| = 6.062 \ (answer) \quad |\bar{I}_C| = 5.268 \ (answer)$$

Problem 212

Let us, in problem 211 (Fig. 266), denote the line voltages by $\bar{V}_{\mathrm{ab}}$, $\bar{V}_{\mathrm{bc}}$, and $\bar{V}_{\mathrm{ca}}$. Show that, in volts, $|\bar{V}_{\mathrm{ab}}| = 140.58$ $|\bar{V}_{\mathrm{bc}}| = 109.49$ $|\bar{V}_{\mathrm{ca}}| = 126.00$.

Problem 213

One formula for calculating average power is $P = RI^2$, where I is magnitude of rms current. Using this formula, find the total power P produced by the generator in problem 211.

Problem 214
It can be shown that the true average sinusoidal power output P of a generator is equal to the REAL PART (r.p.) of the PRODUCT of the generator voltage $\bar{V}$ and the CONJUGATE of the generator current $\bar{I}$, that is

$$P = \text{r.p.}[\bar{V}\bar{\bar{I}}] \tag{467*}$$

where $\bar{\bar{I}}$ ("double overscore") denotes the conjugate of $\bar{I}$. Verify that the use of eq. (467) gives the same answer as found in problem 213.

The preceding problems dealt with Fig. 266, which is a case in which an UNBALANCED GENERATOR feeds a BALANCED LOAD. Now let us consider the opposite case, in which a BALANCED GENERATOR feeds an UNBALANCED LOAD, as illustrated in Fig. 267, where voltage $\bar{V}$ is the reference phase voltage with reference to the junction point "O." Also, $\bar{V}_\text{N}$ is the voltage at the load junction point, also with respect to the point O.

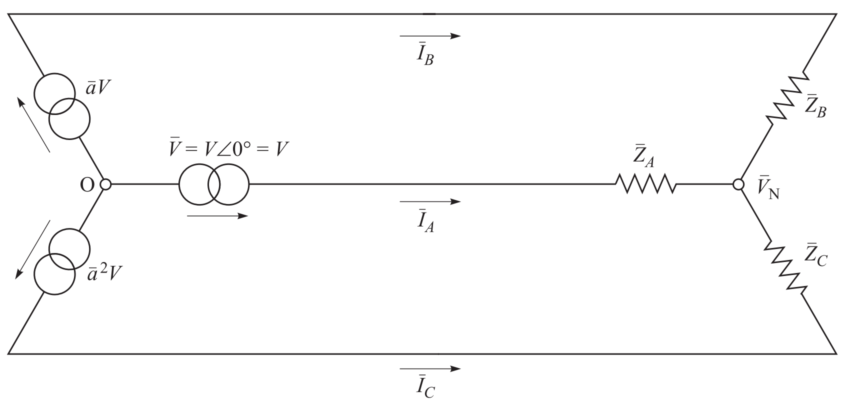


Fig. 267

Keep in mind that $\bar{V}_\text{N}$ denotes the voltage drop from the junction point in the load to the junction point at O. Therefore, since the generator voltage is equal to the sum of the voltage drops in any closed path, we have the following three voltage equations in Fig. 267:

$$V = \bar{Z}_A\bar{I}_A + \bar{V}_\text{N}$$
$$\bar{a}V = \bar{Z}_B\bar{I}_B + \bar{V}_\text{N}$$
$$\bar{a}^2 V = \bar{Z}_C\bar{I}_C + \bar{V}_\text{N}$$

Here we're dealing with the general case of unequal load impedances; thus the three line (and phase) currents will, in general, be unequal, and hence can be resolved into the sum of positive and negative sequences; thus

$$\bar{I}_A = \bar{I}_{A1} + \bar{I}_{A2}$$
$$\bar{I}_B = \bar{I}_{B1} + \bar{I}_{B2}$$
$$\bar{I}_C = \bar{I}_{C1} + \bar{I}_{C2}$$

* See note 29 in Appendix.

which, upon making use of eqs. (461), (462), (464), and (465), can also be written as

$$\bar{I}_A = \bar{I}_{A1} + \bar{I}_{A2}$$
$$\bar{I}_B = \bar{a}\bar{I}_{A1} + \bar{a}^2\bar{I}_{A2}$$
$$\bar{I}_C = \bar{a}^2\bar{I}_{A1} + \bar{a}\bar{I}_{A2}$$

and thus, upon substituting these values of $\bar{I}_A$, $\bar{I}_B$, and $\bar{I}_C$ into the first set of equations following Fig. 267, you should find that

$$\bar{Z}_A\bar{I}_{A1} + \bar{Z}_A\bar{I}_{A2} + \bar{V}_\mathrm{N} = V \tag{468}$$
$$\bar{a}\bar{Z}_B\bar{I}_{A1} + \bar{a}^2\bar{Z}_B\bar{I}_{A2} + \bar{V}_\mathrm{N} = \bar{a}V \tag{469}$$
$$\bar{a}^2\bar{Z}_C\bar{I}_{A1} + \bar{a}\bar{Z}_C\bar{I}_{A2} + \bar{V}_\mathrm{N} = \bar{a}^2V \tag{470}$$

Problem 215

Here you are asked to complete the foregoing discussion concerning Fig. 267 as follows. Making use of eqs. (460) and (463), and the fact that, by Kirchhoff's current law, $\bar{I}_\mathrm{A} + \bar{I}_\mathrm{B} + \bar{I}_\mathrm{C} = 0$, show that the values of the line currents in Fig. 267 are given by the equations

$$\bar{I}_A = \frac{1.732(\bar{Y}_A\bar{Y}_C\epsilon^{j30} + \bar{Y}_A\bar{Y}_B\epsilon^{-j30})V}{\bar{Y}_A + \bar{Y}_B + \bar{Y}_C} \tag{471}$$

$$\bar{I}_B = \frac{1.732(\bar{Y}_B\bar{Y}_C\epsilon^{j90} - \bar{Y}_A\bar{Y}_B\epsilon^{-j30})V}{\bar{Y}_A + \bar{Y}_B + \bar{Y}_C} \tag{472}$$

$$\bar{I}_C = \frac{-1.732(\bar{Y}_A\bar{Y}_C\epsilon^{j30} + \bar{Y}_B\bar{Y}_C\epsilon^{j90})V}{\bar{Y}_A + \bar{Y}_B + \bar{Y}_C} \tag{473}$$

where, in terms of the admittances

$$\bar{Y}_A = 1/\bar{Z}_A \quad \bar{Y}_B = 1/\bar{Z}_B \quad \bar{Y}_C = 1/\bar{Z}_C$$

in which the "reciprocal ohms" are called "mhos."

Problem 216

In Fig. 267 let $V = 125$ volts, $\bar{Z}_A = (3 + j4)$ ohms, $\bar{Z}_B = 8$ ohms, and $\bar{Z}_C = 5$ ohms. Using symmetrical components, verify that the magnitudes of the line currents, in amperes, are equal to

$$|\bar{I}_A| = 25.99 \quad |\bar{I}_B| = 22.85 \quad |\bar{I}_C| = 17.37$$

Problem 217

In three-phase work we often deal in terms of "line voltage" instead of "phase voltage." Thus, in Fig. 267, suppose that the reference voltage is taken to be the line voltage, $\bar{V}_{AB} = V_{AB}\underline{/0^\circ} = V_\mathrm{AB}$, this being the voltage from the center wire to the top wire in the diagram.

What changes would be required to express eqs. (471) through (473) in terms of line voltage instead of phase voltage?

It should be noted that the three-phase problems we've solved here, using the method of symmetrical components, could also have been solved by the ordinary method of loop currents. It should, however, also be noted that other types of three-phase problems exist in which application of the method of symmetrical components provides the only practical way of obtaining exact and rigorous solutions.

CHAPTER 11

Matrix Algebra. Two-Port Networks

Here we take up the subject of "matrix algebra," which has important applications in the study of electric networks. It should be noted, however, that matrix algebra finds wide use in many fields of endeavor, from economics to computer graphics, for example. Your time will therefore be well spent in mastering this interesting and useful subject.

11.1 Introduction to Matrix Algebra

Let us begin by defining that a *matrix* ("MAY triks") is a *rectangular array of elements*, the elements being arranged in a definite order in horizontal *rows* and vertical *columns*.

The *location* of any element in a matrix is always specified by giving *first the* ROW and *then the* COLUMN that the element is located in. Thus, using subscripts, a notation such as a_{23} denotes the element at the intersection of the *second row* and the *third column* (a symbol such as a_{23} can be read as "a, two, three"). When it is deemed necessary, the row and column subscripts are separated by a comma (for example, $a_{16,11}$).

A matrix is usually identified as such by enclosure in square brackets. Figure 268 is an example of a "3 by 4" matrix, meaning it has *3 rows* and *4 columns*.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

Fig. 268

In the above, note that the subscripts used with each element denote *first the row and then the column* in which the element appears. This convention, of giving first the row number and then the column number, is always used.

A matrix having *m rows* and *n columns* is said to be an "*m* by *n*" matrix. An *m* by *n* matrix therefore consists of *mn* elements. Thus, the "3 by 4" matrix above consists of a total of 12 elements.

The presence of an *m* by *n* matrix is often denoted by the symbol $(m \times n)$, where the cross is read as "by." Figure 268 represents a (3×4) matrix.

A matrix consisting of only a *single row* of elements is called a "row matrix," while a matrix consisting of only a *single column* of elements is called a "column matrix." Thus Fig. 269 is an example of a "1 by 3" row matrix and Fig. 270 is a "3 by 1" column matrix.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix}$$

Fig. 269

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}$$

Fig. 270

A matrix having the same number of rows as columns, that is, an "*m* by *m*" matrix, is called a SQUARE *matrix*. The general example of a (3×3) square matrix is shown below in Fig. 271.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Fig. 271

You will recall, from Chap. 3, that a *determinant* is also a square array of elements. Let us emphasize, however, that a square matrix and a determinant are two entirely different things. A matrix, including a square matrix, is simply an ordered array of elements; it is a mathematical symbol and, taken as a whole entity, it has no numerical value. A *determinant*, on the other hand, represents *a single number* or value, which can be found by expanding the determinant according to the rules laid down in Chap. 3. It is true that, in certain circumstances, a determinant is formed from a square matrix, but this is a result of a special operation, as we'll learn later on.

The "main diagonal" of a square matrix consists of all the elements lying on the diagonal line drawn from the upper left-hand element down to the lower right-hand element. Thus the "main diagonal" of the 3 by 3 square matrix above consists of the elements a_{11}, a_{22}, and a_{33}.

If the elements in the main diagonal of a square matrix are *all ones*, (all "1"s), and all the other elements are *zeros*, the square matrix is then called a *unit* or *identity* matrix. Figure 272 is a "unit matrix" of order 5.

The unit matrix, which may of course be of any order n ($n = 5$ in the above), is usually denoted by the symbol **I**, and will be useful in some of our later work.

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

Fig. 272

Consider next the general matrix of *m rows and n columns*, that is, the general "m by n" matrix. The standard notation associated with the general $(m \times n)$ matrix is shown in Fig. 273.

$$
\begin{array}{rc}
 & \begin{array}{ccccccc} \text{Col. 1} & \text{Col. 2} & & \text{Any } j\text{th column} & & \text{Last } (n\text{th}) \text{ column} \end{array} \\
\begin{array}{r} \text{First row} \rightarrow \\ \text{Second row} \rightarrow \\ \\ \text{Any } i\text{th row} \rightarrow \\ \\ \text{Last } (m\text{th}) \text{ row} \rightarrow \end{array} &
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\
\vdots & \vdots & & \vdots & & \vdots \\
a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\
\vdots & \vdots & & \vdots & & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn}
\end{bmatrix}
\end{array}
$$

Fig. 273

In the above, note that a_{ij} denotes the *general element* of the matrix at the intersection of any ith *row and* jth *column* ("eye-th" row and "jay-th" column).

With the foregoing in mind, we're now free to define some of the operations of matrix algebra. We begin as follows, where it should be noted that the plural of matrix is "matrices" ("MAY trah seez").

Let **A** and **B** denote two matrices. We define that two such matrices can be *equal*, that is, $\mathbf{A} = \mathbf{B}$, only if

(a) **A** and **B** have the same number of m rows and the same number of n columns; that is, only if both are $(m \times n)$ matrices, and

(b) all corresponding elements of **A** and **B** are equal.

For example, if **A** and **B** are both (2×3) matrices, thus,

$$
\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}
$$

then $\mathbf{A} = \mathbf{B}$ only if $a_{11} = b_{11}$, $a_{12} = b_{12}$, $\ldots$, $a_{23} = b_{23}$.

Next, two matrices can be *added or subtracted* only if they have the same number of m rows and the same number of n columns. If this requirement is met, then we define that the *sum or difference* of two matrices is obtained by adding or subtracting *corresponding pairs of elements* in the two matrices. For example, the sum or difference of the two (2×3)

matrices given above is equal to

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}(a_{11}+b_{11}) & (a_{12}+b_{12}) & (a_{13}+b_{13})\\ (a_{21}+b_{21}) & (a_{22}+b_{22}) & (a_{23}+b_{23})\end{bmatrix}$$

$$\mathbf{A}-\mathbf{B}=\begin{bmatrix}(a_{11}-b_{11}) & (a_{12}-b_{12}) & (a_{13}-b_{13})\\ (a_{21}-b_{21}) & (a_{22}-b_{22}) & (a_{23}-b_{23})\end{bmatrix}$$

Since, for example, $(a_{11}+b_{11})=(b_{11}+a_{11})$, it follows, from inspection of the above, that the addition and subtraction of matrices can be done in any order we wish—that is, matrix addition and subtraction are *commutative* operations; thus

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$

and

$$\mathbf{A}-\mathbf{B}=-\mathbf{B}+\mathbf{A}$$

Suppose, now, that the above two matrices happened to be EQUAL matrices, $\mathbf{A}=\mathbf{B}$. By the definition of equality already laid down, this means that $a_{11}=b_{11}$, $a_{12}=b_{12}$, and so on. Therefore, if $\mathbf{A}=\mathbf{B}$, the sum of $\mathbf{A}$ and $\mathbf{B}$ at the top of the page becomes

$$\mathbf{A}+\mathbf{B}=2\mathbf{A}=\begin{bmatrix}2a_{11} & 2a_{12} & 2a_{13}\\ 2a_{21} & 2a_{22} & 2a_{23}\end{bmatrix}$$

Likewise, if we were dealing with the sum of say *three* equal matrices, then all the above "2" coefficients would be "3" coefficients, and so on, for the sum of any number of equal matrices. Therefore, to be consistent, we must define that *a constant times a matrix* is obtained by *multiplying* EVERY ELEMENT *of the matrix by the constant*.

Thus, if k is any constant, and $\mathbf{A}$ is (for example) a 2×3 matrix, then k *times* $\mathbf{A}$ is equal to

$$k\mathbf{A}=k\begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{bmatrix}=\begin{bmatrix}ka_{11} & ka_{12} & ka_{13}\\ ka_{21} & ka_{22} & ka_{23}\end{bmatrix}$$

Fig. 274

and so on, in the same way, for the product of k and any matrix. (Note that this rule is different from that for *determinants*, in which a constant k multiplies only the elements of any one row or any one column of the determinant.)

Problem 218

A "5×9" matrix is a rectangular array of ____________ elements arranged in ____________ rows and ____________ columns. The notation $a_{4,6}$ denotes the element at the intersection of ____________ four and ____________ six. A "unit matrix" is always a ____________ matrix. If $\mathbf{A}$ is a 6×5 matrix, and if $\mathbf{A}=\mathbf{B}$, then $\mathbf{B}$ is a ____________ matrix.

Problem 219

If $\mathbf{A}$ is a 3×4 matrix and $\mathbf{B}$ is a 4×3 matrix, does the sum $\mathbf{A}+\mathbf{B}$ exist?

Problem 220

$$\text{If } \mathbf{A}=\begin{bmatrix}3 & 2 & -4\\ 1 & -7 & 5\end{bmatrix} \text{ and } \mathbf{B}=\begin{bmatrix}0 & 2 & 1\\ 1 & 2 & 4\end{bmatrix}, \text{ then, } \mathbf{A}+\mathbf{B}=$$

Problem 221

$$\begin{bmatrix} 0 & 4 \\ 3 & -2 \end{bmatrix} - \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} =$$

Problem 222

(a) $$\begin{bmatrix} 6 & 2 & -1 \\ 4 & 0 & 0 \\ 3 & 0 & 5 \end{bmatrix} + \begin{bmatrix} 2 & -2 & -3 \\ 4 & -3 & 0 \\ 1 & 1 & -6 \end{bmatrix} =$$

(b) $$\begin{bmatrix} 1 & 3 \\ 4 & -6 \end{bmatrix} + \begin{bmatrix} 2 & -9 \\ 3 & 6 \end{bmatrix} - \begin{bmatrix} 7 & -4 \\ 0 & 10 \end{bmatrix} =$$

Problem 223
Given that

$$\begin{bmatrix} 3a & 2b & c \\ 4e & 5f & 2g \end{bmatrix} = \begin{bmatrix} 18 & -6 & 4 \\ 20 & 15 & -12 \end{bmatrix}$$

what values do the letters represent?

11.2 Product of Two Matrices

Let us now define the *matrix product* **AB**, that is, "matrix **A** times matrix **B**." First of all we'll find that, in matrix multiplication, the ORDER in which the factors are written is important. Thus, in general, the matrix products **AB** and **BA** will NOT be equal. The reason for the seemingly peculiar way that matrix multiplication is defined will become clear to us later on in our work.

The first requirement, in the definition of matrix multiplication, is that

> A matrix product, in the order **AB**, exists only if the number of *columns* of the first matrix **A** is equal to the number of *rows* of the second matrix **B**.
>
> Thus if **A** is an $(m \times n)$ matrix and **B** is an $(n \times p)$ matrix, then the product of the two in the order **AB** *does exist*, because **A** has n columns and **B** has the same number of n rows. If this requirement is *not* satisfied, then the product in the order **AB** *cannot be taken*.

Any two such matrices, in which the number of columns of **A** is equal to the number of rows of **B**, are said to be *conformable* in the order **AB**. Thus the matrix product **AB** exists only of **A** and **B** are "conformable matrices" in the order **AB**.

If **A** and **B** are two matrices conformable in the order **AB**, then *the product* **AB** is itself a *matrix* **C**, whose elements are found according to the following rule:

> If $\mathbf{AB} = \mathbf{C}$, the *element* c_{ij}, at the intersection of the ith row and the jth column of $\mathbf{C}$, is equal to the *sum of the products of corresponding pairs* of the elements of the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$.

The above definition can be expressed as a general formula as follows.

Let $\mathbf{A}$ be an $(m \times n)$ and $\mathbf{B}$ an $(n \times p)$ matrix. Note that $\mathbf{A}$ has n *columns* and $\mathbf{B}$ has n *rows*, so that the product in the order $\mathbf{AB} = \mathbf{C}$ does exist. The procedure for finding the value of any element c_{ij} in the product $\mathbf{C}$ as stated in the above rule, is illustrated in Fig. 275.

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \cdots & b_{1j} & \cdots \\ \cdots & b_{2j} & \cdots \\ & \vdots & \\ \cdots & b_{nj} & \cdots \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{bmatrix}$$

Any ith row of $\mathbf{A}$, where $i = 1, 2, 3, ..., m$, because $\mathbf{A}$ has m rows	Any jth column of $\mathbf{B}$, where $j = 1, 2, ..., p$, because $\mathbf{B}$ has p columns	Element c_{ij} in matrix $\mathbf{C}$

Fig. 275

It follows from the definition and from inspection of Fig. 275 that the SUM OF THE PRODUCTS OF CORRESPONDING PAIRS of the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$ gives

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \tag{474}$$

Since $\mathbf{A}$ is an "m by n" matrix we have that $i = 1, 2, 3, \ldots, m$, and since $\mathbf{B}$ is an "n by p" matrix we have that $j = 1, 2, 3, \ldots, p$, and therefore the product matrix $\mathbf{C}$ will consist of m *rows and p columns* of elements; that is

> If $\mathbf{A}$ is an $(m \times n)$ matrix and $\mathbf{B}$ is an $(n \times p)$ matrix, the product $\mathbf{AB}$ is an $(m \times p)$ matrix $\mathbf{C}$, as illustrated in Fig. 276.

$$\mathbf{AB} = \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ c_{31} & c_{32} & \cdots & c_{3p} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

Fig. 276

The first step in finding the above product, $\mathbf{AB} = \mathbf{C}$, is to note the values of m, n, and p, for the given problem. The second step is then to *calculate the value of each of the elements in the matrix* $\mathbf{C}$, which is done by making use of eq. (474). Consider now the following.

Example

Find the matrix product $\mathbf{AB} = \mathbf{C}$ if

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 4 & 2 \\ -1 & 5 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 2 & -3 \\ 4 & 1 \end{bmatrix}$$

Solution

Since the first factor $\mathbf{A}$ has *2 columns* and the second factor $\mathbf{B}$ has *2 rows*, a product in the order $\mathbf{AB}$ does exist. Note that since $\mathbf{A}$ is a (3×2) matrix and $\mathbf{B}$ is a (2×2) matrix, the *product* $\mathbf{C}$ will be a (3×2) matrix; that is, $\mathbf{C}$ will have 3 rows and 2 columns of elements. Hence the solution, $\mathbf{AB} = \mathbf{C}$, is the form

$$\begin{bmatrix} 2 & 3 \\ 4 & 2 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

Let us begin by finding the value of element c_{11}. By definition, c_{11} is equal to the *sum* of the *products of corresponding pairs* or *row 1 of* $\mathbf{A}$ and *column 1 of* $\mathbf{B}$, and hence

$$c_{11} = (2)(2) + (3)(4) = 16$$

Let us next find the value of element c_{21}. By definition, this is equal to the sum of the products of corresponding pairs of *row 2 of* $\mathbf{A}$ and *column 1 of* $\mathbf{B}$, and hence

$$c_{21} = (4)(2) + (2)(4) = 16$$

In the same way, the value of c_{31} is the sum of the products of corresponding pairs of *row 3 of* $\mathbf{A}$ and *column 1 of* $\mathbf{B}$; thus

$$c_{31} = (-1)(2) + (5)(4) = 18$$

We've now found the values of the elements of the *first column of* $\mathbf{C}$; the next step is to find the values of the elements of the *second column of* $\mathbf{C}$, beginning with element c_{12}. Again, by definition, the value of c_{12} is equal to the *sum* of the *products of corresponding pairs of row 1 of* $\mathbf{A}$ *and column 2 of* $\mathbf{B}$, and hence

$$c_{12} = (2)(-3) + (3)(1) = -3$$

Next, the value of c_{22} is the sum of the products of corresponding pairs of *row 2 of* $\mathbf{A}$ and *column 2 of* $\mathbf{B}$; thus

$$c_{22} = (4)(-3) + (2)(1) = -10$$

Finally, in the same way, the value of element c_{32} is the sum of the products by pairs of *row 3 of* $\mathbf{A}$ *and column 2 of* $\mathbf{B}$, and hence

$$c_{32} = (-1)(-3) + (5)(1) = 8$$

and thus we have

$$\begin{bmatrix} 2 & 3 \\ 4 & 2 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 16 & -3 \\ 16 & -10 \\ 18 & 8 \end{bmatrix} \quad (\textit{Answer})$$

In the above answer, note, for example, that we *cannot* factor a "2" out of the first column; such "factoring out" can be done only if a number factors out of *all* the elements of a matrix, as discussed at the end of section 11.1.

Important Note: Suppose, for example, that **A** is a "3 by 2" matrix and **B** is a "2 by 4" matrix. Then the product of the two *can be taken in the order* **AB** because the first factor **A** has *2 columns* and the second factor **B** has *2 rows*; that is, in this case

$$\mathbf{A}_{3,2}\mathbf{B}_{2,4} = \mathbf{C}_{3,4}$$

But notice that *multiplication in the order* **BA** *cannot be done*, because if we attempt the product **BA** we have

$$\mathbf{B}_{2,4}\mathbf{A}_{3,2}$$

where now the first factor **B** has *4 columns* and the second factor **A** has *3 rows*, and so they are not conformable in the order **BA**. This illustrates the fact that, in general,

Matrix multiplication is not commutative; that is, in general, **AB** *does not equal* **BA**.

This is true even if **A** and **B** are conformable in both **AB** and **BA** form, as the following example illustrates.

Let

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$$

Note that the products **AB** and **BA** both exist, because the number of columns of the first factor equals the number of rows of the second factor regardless of whether we write **AB** or **BA**. But note that

$$\mathbf{AB} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} (6+6) & (15+2) \\ (2+12) & (5+4) \end{bmatrix} = \begin{bmatrix} 12 & 17 \\ 14 & 9 \end{bmatrix}$$

whereas

$$\mathbf{BA} = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} (6+5) & (4+20) \\ (9+1) & (6+4) \end{bmatrix} = \begin{bmatrix} 11 & 24 \\ 10 & 10 \end{bmatrix}$$

which, since the answers are not equal, shows that in general **AB** *and* **BA** *do not represent equal matrix products*, even if **A** and **B** are conformable in either order of multiplication.

Problem 224
Find matrix **C**, given that

$$\begin{bmatrix} 2 & 4 \\ 6 & -4 \end{bmatrix} \begin{bmatrix} 12 \\ 8 \end{bmatrix} = \mathbf{C}$$

Problem 225
Given

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -2 \\ 3 & 0 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 2 & -1 \\ 3 & 4 \\ -5 & 2 \end{bmatrix}$$

find the matrix product $\mathbf{AB} = \mathbf{C}$

Problem 226
Find **C**, where

$$\begin{bmatrix} 0 & 2 \\ 0 & 4 \\ 6 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 \\ 6 & 0 \end{bmatrix} + \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} + \begin{bmatrix} 4 & -2 \\ 3 & 6 \end{bmatrix} \right) = \mathbf{C}$$

Problem 227
Find **C**, if

$$\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} -7 & 6 & 0 \\ -2 & 1 & 4 \end{bmatrix} = \mathbf{C}$$

Problem 228
Find the matrix product

$$\begin{bmatrix} 1 & 2 & -2 & 3 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 7 & 2 \\ 4 & 2 & -5 & 10 \end{bmatrix} \begin{bmatrix} 6 \\ -2 \\ 0 \\ 9 \end{bmatrix} = \mathbf{C}$$

Problem 229
Find the matrix **C**, where

$$\begin{bmatrix} 1 & 1 & 4 \\ 0 & -3 & 2 \\ 9 & 6 & -5 \end{bmatrix}^2 = \mathbf{C}$$

Problem 230
If

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 & -3 \\ 1 & 2 & 0 & 4 \\ 3 & 2 & -6 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & -2 \\ 3 & 0 \\ 2 & 5 \\ -6 & 1 \end{bmatrix}$$

find the matrix product **AB**.

11.3 The Inverse of a Square Matrix

Let us now consider the important case where two sets of unknowns, which we'll denote by xs and ys, are related by means of n simultaneous linear equations. Let us take, as an example, the case for $n = 3$, as shown below, where the as are constant coefficients. Note that the subscript of each a coefficient gives the location, row, and column of that particular a.

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = y_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = y_3 \end{aligned}\right\} \tag{475}$$

Now notice, as a result of the definition of multiplication of matrices laid down in section 11.2, that the above set of equations can be written in the form of a *matrix product*; thus

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \tag{476}$$

or, in abbreviated form

$$[\mathbf{A}][\mathbf{X}] = [\mathbf{Y}]$$

or, more simply, we may merely write

$$\mathbf{AX} = \mathbf{Y} \tag{477}$$

where, in eq. (477), $\mathbf{A}$ represents *the square matrix formed from the constant "a" coefficients*, and $\mathbf{X}$ and $\mathbf{Y}$ represent the column matrices formed from the values of the unknowns in eq. (476).

Now suppose we wish to have the $\mathbf{X}$ matrix alone, by itself, on the left-hand side of eq. (477). To indicate this we change matrix (477) into the form

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y} \tag{478}$$

where the matrix $\mathbf{A}^{-1}$ is called *the* INVERSE *of the square matrix* $\mathbf{A}$, or simply as "the inverse of matrix $\mathbf{A}$," since only a square matrix can have an inverse.

Let us now find the actual form that $\mathbf{A}^{-1}$ must have in order to legitimately transform (477) into (478). To do this, let us go back and solve the given set of simultaneous equations, eq. (475), for the x values, using the standard procedure of determinants from Chap. 3. Doing this, we have

$$x_1 = \frac{\begin{vmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{vmatrix}}{\Delta}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & y_1 & a_{13} \\ a_{21} & y_2 & a_{23} \\ a_{31} & y_3 & a_{33} \end{vmatrix}}{\Delta}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \\ a_{31} & a_{32} & y_3 \end{vmatrix}}{\Delta}$$

where, as usual, Δ (delta) is the value of the determinant formed from the constant a coefficients.

Now recall, from Chap. 3, that the value of any determinant of order 3 or more can be found by expanding the determinant in terms of the *minors* of any row or any column of the determinant. Since we have to deal with determinants when finding the inverse matrix, let us review, just briefly, some details from Chap. 3.

A determinant of order n is a square array of elements having n rows and n columns, there thus being n elements in each row and each column. Now let a_{ij} denote the element at the intersection of the ith row and jth column. If we then strike out the row and column in which a_{ij} appears, the determinant that remains is of order $n-1$ and is called the *minor determinant* of element a_{ij}.

In Chap. 3 we showed that the *value of a determinant* is equal to the sum of the products of the elements of any row or column and their corresponding minor determinants, each such product being multiplied by $(-1)^{i+j}$. Thus, if we expand the third-order determinant in the last expression for x_1 above, using minors of the first column, we have that

$$x_1 = \left(y_1 \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - y_2 \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + y_3 \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \right) \frac{1}{\Delta}$$

Likewise, if we expand the third-order determinant in the last expression for x_2, in terms of the minors of the second column, we have that

$$x_2 = \left(-y_1 \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + y_2 \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - y_3 \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \right) \frac{1}{\Delta}$$

and lastly, if we expand the third-order determinant in the last expression for x_3, in terms of the minors of the third column, we have that

$$x_3 = \left(y_1 \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - y_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + y_3 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right) \frac{1}{\Delta}$$

Now, in the last three equations above, let us denote the value of each second-order minor determinant, *including the sign factor* $(-1)^{i+j}$, by the notation A_{ij}, where i and j are the numbers of the row and column struck out to form the minor determinant.* Using this notation, the last three equations above become

$$\begin{aligned} x_1 &= \frac{A_{11}}{\Delta} y_1 + \frac{A_{21}}{\Delta} y_2 + \frac{A_{31}}{\Delta} y_3 \\ x_2 &= \frac{A_{12}}{\Delta} y_1 + \frac{A_{22}}{\Delta} y_2 + \frac{A_{32}}{\Delta} y_3 \\ x_3 &= \frac{A_{13}}{\Delta} y_1 + \frac{A_{23}}{\Delta} y_2 + \frac{A_{33}}{\Delta} y_3 \end{aligned}$$

which in matrix notation becomes

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \tag{479}$$

Equation (479) above is the *inverse form* of eqs. (476) and (477); that is, (479) is of the form

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y} \tag{480}$$

* A_{ij} is called the *cofactor* of element a_{ij}. If M_{ij} is the minor determinant of any element a_{ij}, the "cofactor" of a_{ij} is the *minor determinant with proper sign included*; thus

$$A_{ij} = (-1)^{i+j} M_{ij}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are the x and y column matrices. Comparison of eqs. (479) and (480) shows that $\mathbf{A}^{-1}$, *the* INVERSE OF THE SQUARE MATRIX $\mathbf{A}$ *of eqs. (476) and (477)*, is equal to

$$\mathbf{A}^{-1} = \frac{1}{\Delta}\begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \tag{481}$$

where Δ = value of the determinant formed from the constant a coefficients of the original three simultaneous linear equations (eq. (475)), and where $\mathbf{A}_{ij}$ = value of the cofactor formed by deleting the row and column of element a_{ij} in the original determinant formed from the a coefficients. It should be understood that all subscripts, everywhere, refer to the ith row and jth column of the *original simultaneous equations* (475). Thus, in (481), A_{21} is actually the cofactor of a_{21} in the *second row and first column* in the original eq. (475), even though A_{21} appears in the first row, second column position in (481). The procedure will be clear from the discussion that follows eq. (482) below.

It's apparent that the foregoing work can be extended to finding the inverse of any nth order square matrix $\mathbf{A}$. Thus, *given any square matrix* $\mathbf{A}$ *of order* n, as in eq. (482) below,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \tag{482}$$

The procedure for finding the inverse matrix $\mathbf{A}^{-1}$, necessary to satisfy the relationships

$$\mathbf{AX} = \mathbf{Y}$$

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y}$$

can be summarized in the following steps.

Step 1
Find Δ, the value of the nth-order determinant formed from the elements of the given square matrix $\mathbf{A}$ of eq. (482).

Step 2
Replace each element in the given matrix $\mathbf{A}$ by its cofactor to get a matrix which we'll call $\mathbf{A}_0$ ("A sub zero"); thus

$$\mathbf{A}_0 = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & A_{n3} & \cdots & A_{nn} \end{bmatrix} \tag{483}$$

Now *interchange the rows and columns* in (483) (that is, let the first row become the first column, the second row become the second column, and so on)* and then multiply by $1/\Delta$. The result is the *inverse matrix* $\mathbf{A}^{-1}$ *of the given matrix* $\mathbf{A}$; that is, the matrix capable of transforming eq. (477) into (478).

* This generates the "transpose" of matrix $\mathbf{A}_0$, as we'll define in section 11.5.

Example

$$\text{Given the matrix } \mathbf{A} = \begin{bmatrix} 2 & 1 & -4 \\ 1 & 5 & 3 \\ -2 & 0 & 1 \end{bmatrix}, \text{ find } \mathbf{A}^{-1}$$

Solution

Step 1

Note that **A** is a square matrix of order 3. Putting the matrix in the form of a third-order determinant we have (where, for convenience, we've expanded the determinant in terms of the minors of the second column)

$$\Delta = \begin{vmatrix} 2 & 1 & -4 \\ 1 & 5 & 3 \\ -2 & 0 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 3 \\ -2 & 1 \end{vmatrix} + 5\begin{vmatrix} 2 & -4 \\ -2 & 1 \end{vmatrix} = -37$$

Step 2

From the definition of "cofactor" (see first footnote in this section) and then from eq. (483) above, we have that the value of $\mathbf{A}_0$ is

$$\mathbf{A}_0 = \begin{bmatrix} +\begin{vmatrix} 5 & 3 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ -2 & 1 \end{vmatrix} & +\begin{vmatrix} 1 & 5 \\ -2 & 0 \end{vmatrix} \\ -\begin{vmatrix} 1 & -4 \\ 0 & 1 \end{vmatrix} & +\begin{vmatrix} 2 & -4 \\ -2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ -2 & 0 \end{vmatrix} \\ +\begin{vmatrix} 1 & -4 \\ 5 & 3 \end{vmatrix} & -\begin{vmatrix} 2 & -4 \\ 1 & 3 \end{vmatrix} & +\begin{vmatrix} 2 & 1 \\ 1 & 5 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 5 & -7 & 10 \\ -1 & -6 & -2 \\ 23 & -10 & 9 \end{bmatrix}$$

Now *interchange the rows and columns* in the last matrix; that is, let the first row become the first column, the second row become the second column, and so on (that is, take the "transpose" of $\mathbf{A}_0$, as described below eq. (483)). Thus, upon switching corresponding rows and columns in this manner, and remembering to multiply the whole by $1/\Delta$, we have

$$\mathbf{A}^{-1} = -\frac{1}{37}\begin{bmatrix} 5 & -1 & 23 \\ -7 & -6 & -10 \\ 10 & -2 & 9 \end{bmatrix} \qquad (\textit{final answer})$$

Note: since $-1/37 = (-1)(1/37)$, we can, if we wish, change $-1/37$ to $1/37$, provided we multiply *every element* in the matrix by -1, in accordance with the discussion given at the end of section 11.1. We can likewise move the "1/37" factor inside the matrix, provided we multiply every element in the matrix by $1/37$.

It is important to note that if $\Delta = 0$, then $1/\Delta$ has the indeterminant form $1/0$; thus the inverse matrix $\mathbf{A}^{-1}$ can exist *only if the determinant of matrix* **A** *is not zero*. This leads to the definition that a *non-singular* matrix is a square matrix whose determinant value Δ is *not zero*. All other matrices are called *singular matrices*. Thus, only a "non-singular" matrix has an inverse.

We give, without proof at this time, the following two relationships involving the inverse operation

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A} \tag{484}$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \tag{485}$$

It's apparent that finding the inverse of a matrix of higher order, using only paper and pencil, is a really time-consuming operation. Fortunately, however, this is exactly the type of work that the digital computer is exceedingly good at. Digital computer programs are available for finding the inverse of any nth order matrix, a fact largely responsible for the increased use of matrix methods in engineering and scientific work.

Your problems here are as follows.

Problem 231
Find the inverse of the third-order matrix

$$\begin{bmatrix} 2 & 0 & 4 \\ 5 & 6 & 0 \\ -3 & -1 & 2 \end{bmatrix}$$

Problem 232
When finding the inverse of a "2×2" matrix, it should be noted that the "minor" of each element is a "1×1" determinant. For example, if

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the "minor" of element b is found, in the usual way, by striking out the row and column in which b appears; thus

$$M_b = \begin{vmatrix} \cancel{a} & \cancel{b} \\ c & \cancel{d} \end{vmatrix} = c$$

so that the "cofactor" of element b is $A_b = (-1)^{1+2} M_b = -c$. With this in mind, find the inverse of the 2×2 matrix

$$\begin{bmatrix} 4 & -7 \\ 3 & -5 \end{bmatrix}$$

Problem 233
Given that

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 0 & -2 \\ 3 & 0 & 6 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & -5 & 0 & 3 \end{bmatrix}, \text{ find } \mathbf{A}^{-1}$$

Problem 234
Find the inverse of the matrix

$$\begin{bmatrix} 8 & 0 & -1 \\ 0 & 2 & 3 \\ -6 & 4 & 3 \end{bmatrix}$$

Problem 235
Suppose one set of variables, x, y, z, is related to another set, r, s, t, by the three simultaneous linear equations

$$\begin{aligned} 3x - 4y + z &= r \\ -2x + y - 5z &= s \\ 4x + 6y - 2z &= t \end{aligned} \qquad \text{(set 1)}$$

It follows that it is also possible to express set 1 in the equivalent form

$$\begin{aligned} ar + bs + ct &= x \\ dr + es + ft &= y \\ gr + hs + it &= z \end{aligned} \qquad \text{(set 2)}$$

provided, of course, that the constant coefficients, a through i, are given the correct values. **Problem**: first express set 1 in matrix form in the manner of eq. (476), and then, by use of the inverse operation, find the values of a through i that will permit the second set of equations to be written in place of the first set.

11.4 Some Properties of the Unit Matrix

The ‘‘unit’’ or ‘‘identity matrix,’’ denoted by **I**, is any square matrix in which all elements of the main diagonal are equal to 1, all other elements being equal to zero. (See Fig. 272 as an example.)

Let us now state that the PRODUCT of any square matrix **A**, of order n, and a unit matrix of the same order n, is equal to the matrix **A**; that is, $\mathbf{AI} = \mathbf{A}$. The truth of this statement will be clear from a study of the following example, in which a square third-order matrix **A** multiplies a third-order unit matrix. Thus, using the procedure of matrix multiplication as defined in section 11.2, you should verify that the following multiplication is correct.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

that is,

$$\mathbf{AI} = \mathbf{A}$$

From the above it’s clear that if **A** is ANY square matrix of order n, and if **I** is a unit matrix of the same order n, then $\mathbf{AI} = \mathbf{A}$. Now reverse the order of multiplication in the above example; doing this, and again keeping the definition of matrix multiplication in mind, you will find it is also true that $\mathbf{IA} = \mathbf{A}$. Thus, if **A** is a square matrix of order n and **I** is a unit matrix of the same order n, then **A** and **I** are commutative in multiplication, and we have the useful relationship

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A} \qquad (486)$$

Thus the unit matrix behaves, in matrix multiplication, much like the number 1 does in ordinary algebraic multiplication. Another useful relationship can be found as follows. Let us begin with the matrix equation

$$\mathbf{AX} = \mathbf{Y}$$

Multiply both sides by $\mathbf{A}^{-1}$:

$$\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{Y} \tag{487}$$

From eq. (480) we know that

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y}$$

From eq. (486), $\mathbf{X} = \mathbf{IX}$, and hence

$$\mathbf{IX} = \mathbf{A}^{-1}\mathbf{Y} \tag{488}$$

Comparison of eqs. (487) and (488) shows that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$; likewise, $\mathbf{AA}^{-1} = \mathbf{I}$, and we thus have the useful fact that the product of a matrix $\mathbf{A}$ and its inverse equals a unit matrix $\mathbf{I}$; that is,

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{489}$$

Problem 236

If we have found the inverse of a given matrix $\mathbf{A}$, then eq. (489) can be used to check the correctness of our work. This is true because, if our work is correct, it has to be true that $\mathbf{AA}^{-1} = \mathbf{I}$, as eq. (489) states. In the example in section 11.3 we found that if

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -4 \\ 1 & 5 & 3 \\ -2 & 0 & 1 \end{bmatrix}$$

then

$$\mathbf{A}^{-1} = -\frac{1}{37}\begin{bmatrix} 5 & -1 & 23 \\ -7 & -6 & -10 \\ 10 & -2 & 9 \end{bmatrix}$$

Your problem here is to double-check the above result by verifying, by actual multiplication, that eq. (489) is satisfied.

11.5 Algebraic Operations. Transpose of a Matrix

We have found that, in general, matrix multiplication is not commutative; that is, in general, $\mathbf{AB}$ does not equal $\mathbf{BA}$ (eqs. (486) and (489) are exceptions to this general rule). Other than the restriction on multiplication, however, most of the other rules of ordinary algebra do also apply to matrix algebra, as follows.

First, as pointed out in section 11.1, matrix *addition* is commutative; that is, $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. Also, of course, $\mathbf{A} - \mathbf{B} = -\mathbf{B} + \mathbf{A}$.

Second, matrix multiplication is *distributive* with respect to addition; that is, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.

Third, matrix multiplication is *associative*; that is, $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

Next, with regard to matrix equations, we may add or subtract the same matrix from both sides of such equations without upsetting the equality of the two sides; for example, if $\mathbf{A} = \mathbf{B}$, then, $\mathbf{A} + \mathbf{C} = \mathbf{B} + \mathbf{C}$. (This assumes, of course, that the matrices all have the same number of m rows and the same number of n columns, as laid down in the requirement for matrix addition in section 11.1.)

We may also multiply both sides of a matrix equation by the same matrix; thus, if $\mathbf{B} = \mathbf{C}$, then $\mathbf{AB} = \mathbf{AC}$, provided, of course, that $\mathbf{A}$ is conformable with $\mathbf{B}$ and $\mathbf{C}$, and that the order of multiplication is the same on both sides of the equation.

Let us next define that the TRANSPOSE of any given matrix $\mathbf{A}$ is another matrix $\mathbf{A}_t$,* *in which the rows of* $\mathbf{A}$ *are written as the columns of* $\mathbf{A}_t$; that is, the *first row* of $\mathbf{A}$ is the *first column* of $\mathbf{A}_t$, the *second row* of $\mathbf{A}$ is the *second column* of $\mathbf{A}_t$, and so on. For example, if

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 0 & 9 \\ 2 & 4 & -7 & 6 \\ 5 & 8 & -3 & 1 \end{bmatrix}$$

then

$$\mathbf{A}_t = \begin{bmatrix} 3 & 2 & 5 \\ -1 & 4 & 8 \\ 0 & -7 & -3 \\ 9 & 6 & 1 \end{bmatrix}$$

Various relationships exist between a matrix and its transpose. The following are the most important ones that you should be aware of.

$$(\mathbf{A}_t)_t = \mathbf{A} \tag{490}$$

$$(\mathbf{A} + \mathbf{B})_t = \mathbf{A}_t + \mathbf{B}_t \tag{491}$$

$$(\mathbf{AB})_t = \mathbf{B}_t\mathbf{A}_t \qquad \text{(``reversal rule'')} \tag{492}$$

$$\det \mathbf{A} = \det \mathbf{A}_t \tag{493}$$

Note that $\mathbf{A}$ can only be a square matrix in eq. (493), because a determinant is defined as a square array of elements.

Problem 237

(a) If

$$\mathbf{A} = \begin{bmatrix} 4 & -4 \\ 3 & 2 \end{bmatrix}, \text{ then } \mathbf{A}_t =$$

(b) If

$$\mathbf{A} = \begin{bmatrix} 6 & 9 & -1 & 4 \\ 0 & 2 & 1 & 3 \\ 4 & 5 & 9 & -2 \\ 0 & 7 & 0 & 8 \end{bmatrix}, \text{ then } \mathbf{A}_t =$$

Problem 238

Verify that eq. (491) is true for the following two given matrices:

$$\mathbf{A} = \begin{bmatrix} 3 & -3 \\ -7 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 5 & 2 \\ 4 & -3 \end{bmatrix}$$

* It should be noted that the alternative notations $\mathbf{A}^T$ (“A, superscript T”) and $\mathbf{A}'$ (“A prime”) are sometimes used to denote the transpose of a matrix $\mathbf{A}$.

Problem 239
Verify that eq. (492) is true for the given two matrices:

$$\mathbf{A} = \begin{bmatrix} 2 & -4 \\ 0 & 6 \\ -7 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 3 & 0 & 3 \\ 2 & -1 & 9 \end{bmatrix}$$

11.6 Matrix Equations for the Two-Port Network

Let us imagine that we have any kind of linear bilateral network, either active or passive (meaning that it may contain generators, as well as passive elements of R, L, and C), enclosed in a box with a pair of INPUT TERMINALS (1, 1) and a pair of OUTPUT TERMINALS (2, 2), brought out as illustrated in Fig. 277.

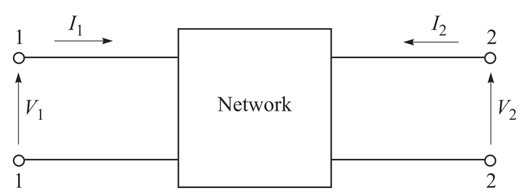


Fig. 277

In the above, each pair of terminals is called a "port"; thus Fig. 277 represents the general form of a TWO-PORT network, with the "input port" on the left and the "output port" on the right. It's also correct to speak of Fig. 277 as a "four-terminal" network.

In all of our work here with such figures, it will be understood that the inputs and outputs are steady-state sinusoidal waves of voltage and current. The letters V, I, and Z that will appear in the equations, will, as usual, denote the complex numbers representing these quantities.*

In working with two-port block diagrams, such as Fig. 277, we make use of *four measurable external quantities*, these being the four quantities V_1, I_1, V_2, and I_2, as shown in the figure.

You'll recall that, in network analysis, the first step is to draw voltage and current arrows to indicate the chosen "positive reference directions" in the network. Once chosen, the directions of the arrows must not, of course, be changed during the investigation of the network. You'll also recall that, in writing network equations, generator voltages are put on one side of the equations with voltage drops on the other side, the signs of the quantities being determined *by the sense in which the arrows are traversed* (eqs. (122) and (123) in section 5.8 illustrate this detail). Let us now emphasize that, in the case of two-port block diagrams, it is standard practice to always draw the voltage and current arrows *in the directions shown in Fig. 277*. The practical application of the notation will be taken up later on, as we progress.

Let us assume that we do not know what is actually inside the box in Fig. 277, except that it is linear and bilateral in nature. Our PROBLEM is to find an EQUIVALENT

* In our work here we'll dispense with the usual "overscore" notation, and simply write V, I, and Z instead of $\bar{V}$, $\bar{I}$, and $\bar{Z}$, since it will be understood that these quantities represent complex numbers.

NETWORK for the unknown contents of the box, GIVEN only the four measurable external quantities shown in Fig. 277.

To do this, we must first write suitable *equations* for the contents of the box, in terms of the four measurable external quantities shown in Fig. 277. This can be done as follows.

In Fig. 277 we can, for example, choose any two of the four external quantities to be independent variables, leaving the other two as dependent variables (two independent variables will require two simultaneous equations). For example, if we choose I_1 and I_2 to be the independent variables, the two equations will be of the form

$$\boxed{\begin{aligned} V_1 &= z_{11}I_1 + z_{12}I_2 \qquad (494) \\ V_2 &= z_{21}I_1 + z_{22}I_2 \qquad (495) \end{aligned}}$$

where the zs are constant coefficients whose values are to be found by actual measurement on any given two-port. Inspection shows that the z coefficients are impedances, and so are measured in ohms. For this reason, eqs. (494) and (495) are said to represent the IMPEDANCE form of an equivalent circuit for the contents of the box in Fig. 277. In matrix form the equations become

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \qquad (496)$$

Note that the double subscripts used with the constant z coefficients, are purposely written, in the two equations (494) and (495), so that they give the proper location, row and column, in the matrix of eq. (496). The above matrix equation is often written in the abbreviated form

$$[\mathbf{V}] = [\mathbf{z}][\mathbf{I}] \qquad (497)$$

or even more simply as $\mathbf{V} = \mathbf{zI}$, if it is understood in the discussions that $\mathbf{V}$ and $\mathbf{I}$ represent the column matrices of eq. (496), and where

$$\mathbf{z} = [\mathbf{z}] = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}$$

this being the "impedance matrix" for the equivalent network.

Now let's pause here, briefly, to discuss how the values of the foregoing z coefficients can be found experimentally in the laboratory. Remember that we do not necessarily know what is actually inside the box in Fig. 277; all we can do, in the laboratory, is measure the EXTERNAL values of V_1, V_2, I_1, and I_2.

Take, for example, the coefficient z_{11}. To find the value of z_{11} for a given two-port, we apply a known voltage V_1 to terminals $(1, 1)$ in Fig. 277, and measure the value of current I_1 with the output terminals $(2, 2)$ *open-circuited*. For this condition $I_2 = 0$, and therefore, setting $I_2 = 0$ in eq. (494), we have

$$z_{11} = \frac{V_1}{I_1}_{(I_2=0)} \qquad (498)$$

which defines the experimental procedure for finding the value of z_{11}.

Next suppose we wish to find, by experimental means, the value of, say, z_{12}. To do this we can apply any convenient voltage to terminals $(2, 2)$ in Fig. 277 and measure the value of the current I_2 with terminals $(1, 1)$ *open-circuited*. This condition makes $I_1 = 0$ and, putting $I_1 = 0$ in eq. (494), we have that

$$z_{12} = \frac{V_1}{I_2}_{(I_1=0)} \tag{499}$$

which thus provides an experimental procedure for finding the value of z_{12}. In the same manner, making use of eq. (495), expressions can be found that will allow us to find the values of z_{21} and z_{22} by experiment in the laboratory.

In regard to the above, we must remember that, except for purely resistive circuits, impedance is a complex number of the form $z = r + jx$, and likewise the resulting currents will be complex numbers of the form $I = I_a + jI_b$. For this reason, experimental determination of the true values of the coefficients requires precision measuring equipment and adequate skill and understanding on the part of the experimenter.

Next recall that *admittance*, which is denoted by y, is the reciprocal of impedance, $y = 1/z$. Thus $I = V/z = yV$, and the equations describing the behavior of the unknown network inside the box can also be written in the ADMITTANCE *form*, thus

$$\boxed{\begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 \qquad (500) \\ I_2 &= y_{21}V_1 + y_{22}V_2 \qquad (501) \end{aligned}}$$

or, in matrix form,

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{502}$$

or, more simply,

$$[\mathbf{I}] = [\mathbf{y}][\mathbf{V}] \tag{503}$$

where

$$[\mathbf{y}] = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$

is the "admittance matrix" for the equivalent network.

Another useful relationship can be found as follows. From eq. (497) we have

$$[\mathbf{z}][\mathbf{I}] = [\mathbf{V}]$$

and hence (by section 11.3)

$$[\mathbf{I}] = [\mathbf{z}]^{-1}[\mathbf{V}]$$

Comparison of the last equation with eq. (503) shows that

$$[\mathbf{y}] = [\mathbf{z}]^{-1} \tag{504}$$

that is, the admittance matrix is the inverse of the impedance matrix (and vice versa, the impedance matrix is the inverse of the admittance matrix).

Equations for the experimental determination of the values of the y coefficients can be found in the same general way as for the z coefficients. For example, if we apply a known voltage V_1 to terminals $(1,1)$ in Fig. 277 and measure I_1 with terminals $(2,2)$ short-circuited, for which $V_2 = 0$, solution of eq. (500) gives

$$y_{11} = \frac{I_1}{V_1}_{(V_2=0)} \tag{505}$$

which thus defines an experimental procedure for finding the value of y_{11}.

Another very useful description of the contents of the box is in terms of the "hybrid" or "h" parameters. In this system I_1 and V_2 are taken as the independent variables, and the simultaneous equations for the equivalent network are written in the form

$$\boxed{\begin{aligned} V_1 &= h_{11}I_1 + h_{12}V_2 \quad (506) \\ I_2 &= h_{21}I_1 + h_{22}V_2 \quad (507) \end{aligned}}$$

or, in matrix form,

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix} \tag{508}$$

Note that eq. (506) says that "V_1 = volts = volts + volts," and thus, by Ohm's law, we see that h_{11} has the dimension of *ohms*, while h_{12} is simply a dimensionless ratio. Likewise, eq. (507) says that "I_2 = amperes = amperes + amperes," and thus, by Ohm's law, we see that h_{21} is a dimensionless ratio while h_{22} has the dimension of "reciprocal ohms"; that is, h_{22} has the dimension of 1/ohms = *mhos*.

In regard to the above, it should be mentioned that the following notation is also often used with the h parameters:

$$\begin{aligned} h_{11} &= h_{\mathrm{i}} = \text{input impedance in ohms,} \\ h_{12} &= h_{\mathrm{r}} = \text{reverse voltage feedback factor,} \\ h_{21} &= h_{\mathrm{f}} = \text{forward current transfer ratio,} \\ h_{22} &= h_{\mathrm{o}} = \text{output admittance in mhos.} \end{aligned}$$

It should also be noted here that the h parameters (and also the y parameters) find especially wide use in practical work involving transistors. One reason for this is that it is relatively easy to find the values of these parameters by direct experiment in the laboratory. From eq. (508) note that the "h-parameter matrix" for the equivalent circuit is

$$[\mathbf{h}] = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

So far we've listed three pairs of simultaneous equations that can be used to describe the network inside the box in Fig. 277. Thus, eqs. (494) and (495) constitute the *impedance* or "z" form of the equations, eqs. (500) and (501) the *admittance* or "y" form, and eqs. (506) and (507) the hybrid or "h" form.

It's also possible to write three more pairs of such equations. Of these three, the two pairs of principal interest are written in terms of what are generally called the "g" and "a" parameters, as follows, beginning with the g parameters.

$$\boxed{\begin{aligned} I_1 &= g_{11}V_1 + g_{12}I_2 \quad (509) \\ V_2 &= g_{21}V_1 + g_{22}I_2 \quad (510) \end{aligned}}$$

and hence, in matrix form,

$$\begin{bmatrix} I_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ I_2 \end{bmatrix} \tag{511}$$

Lastly, the a parameters are defined by the equations

$$\boxed{\begin{aligned} V_1 &= a_{11}V_2 - a_{12}I_2 \\ I_1 &= a_{21}V_2 - a_{22}I_2 \end{aligned}} \qquad \begin{aligned} &(512)^* \\ &(513)^* \end{aligned}$$

or, in matrix form,

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix} \qquad (514)^*$$

The equations show that the g and a coefficients are hybrid-type parameters; some are measured in units of impedance z, some in units of admittance y, and some are ratios of like quantities and thus dimensionless.

Problem 240
Keeping in mind that only like quantities can be added together or set equal to each other, and using the basic relationships $i = v/z = yv$ and $v = iz = i/y$, find the dimensions of each of the g and a coefficients.

Problem 241
Write a relationship that can be used to experimentally find the value of the y_{22} coefficient. Repeat for the g_{21} coefficient.

Problem 242
Solve eq. (514) for

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

Problem 243
If $\mathbf{A}$ is a square matrix of order n, and if c is a constant coefficient, show that

$$[c\mathbf{A}]^{-1} = \frac{1}{c}\mathbf{A}^{-1}$$

(In the above, you may find it convenient to refer to eqs. (482) and (483).)

Problem 244
Express the value of each of the four impedance parameters in terms of the admittance parameters.

Problem 245
For a certain two-port operating at 30 megahertz (30 MHz), it is found that the values of the admittance parameters are, in mhos, equal to

$$y_{11} = 5(4 + j3)10^{-3} \qquad y_{12} = -(2 + j)10^{-3}$$

$$y_{21} = 10(4 - j9)10^{-3} \qquad y_{22} = j(6)10^{-3}$$

Making use of the results of problem 244, find the value of each of the four z parameters at the same frequency.

---

* It is convenient, with a coefficients, to use the minus signs as shown. The reason for doing so is explained later on in the discussion of the "cascade" connection of two-ports. Note that the minus sign appears with I_2 in the matrix equation.

11.7 Continuing Discussion of the Two-Port Network

In the preceding section we were introduced to five different network parameters, denoted by z, y, h, g, and a.

Let us here emphasize that the *values* of these parameters *depend only upon the particular network* INSIDE THE BOX *in Fig. 277*. In other words, the values of the parameters are *independent of any external connections* that might be made to input and output terminals shown in Fig. 277. This is because the parameters are defined in terms of open-circuit and short-circuit values of voltage and current (and thus the parameter values would not be affected by any external connections that might be made to the input and output terminals).

Our object now is threefold, as follows. FIRST, we need to explain why some of the parameters may have ''negative'' values. SECONDLY, we'll derive an ''equivalent circuit'' in terms of the h parameters, since these parameters are much used in practical work. THIRD, we'll find the current that would flow into an external load impedance of Z_L ohms if connected to terminals $(2, 2)$ in Fig. 277.

Let us begin with the first item above. The reason that some of the parameters may be negative is because of the particular directions that have been universally adopted for the voltage and current arrows in Fig. 277.

As you know, ''voltage and current arrows'' indicate what we choose to call the ''positive directions'' of the voltages and currents in any given network.

The showing of such arrows would not generally be important if we always dealt only with simple series circuits. It is, however, definitely necessary to show the arrows when dealing with multiple-loop networks. This is true because in a multi-loop network each voltage and current, in any loop, affects the values of voltage and current in any other loop. Thus the several *simultaneous equations* that describe a given network *cannot be written independently of each other*; that is, the various ''loop equations'' must be written in an orderly manner, each equation taking into account *the mutual effects of all the other loops*. One way to insure that this is done is to first assign, by means of arrows, the chosen ''positive directions'' in the network, and then strictly adhere to these arrows when writing the simultaneous equations for the network.

Now let's get back to the explanation of why, in a given case, some of the two-port parameters may have negative values. In doing this we'll take as examples the h-parameters, as this will have the dual advantage of introducing these important practical parameters.

Therefore, from inspection of the defining eqs. (506) and (507), you can verify that the values of the short-circuit and open-circuit h-parameters are given as follows.

$$h_{11} = \frac{V_1}{I_1}_{(V_2=0)} = \text{short-circuit input impedance (ohms)} \tag{515}$$

$$h_{12} = \frac{V_1}{V_2}_{(I_1=0)} = \text{open-circuit reverse-voltage feedback factor} \tag{516}$$

$$h_{21} = \frac{I_2}{I_1}_{(V_2=0)} = \text{short-circuit forward current gain} \tag{517}$$

$$h_{22} = \frac{I_2}{V_2}_{(I_1=0)} = \text{open-circuit output admittance (mhos)} \tag{518}$$

Note that h_{12} is the ratio of "volts to volts" and h_{21} is the ratio of "amperes to amperes" and hence are equal to pure numbers, and are thus said to be "dimensionless" quantities.

Now, to illustrate why a parameter may carry a negative sign, let us suppose that the network inside the box in Fig. 277 happens to be a T network, as shown in Fig. 278, in which (for simplicity in making our point here) the elements all have equal values of R ohms.

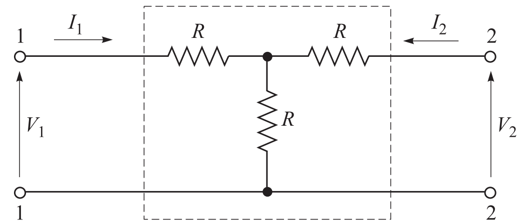


Fig. 278

Now let us note, again, that eqs. (515) through (518) are defined for the *two different conditions* of $V_2 = 0$ and $I_1 = 0$; with this in mind, consider Figs. 279 and 280. In these two figures, note that the voltage and current arrows are all drawn in the "positive sense" *as defined in accordance with Fig. 277*.

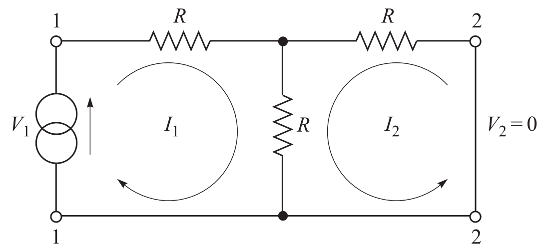


Fig. 279. $V_2 = 0$.

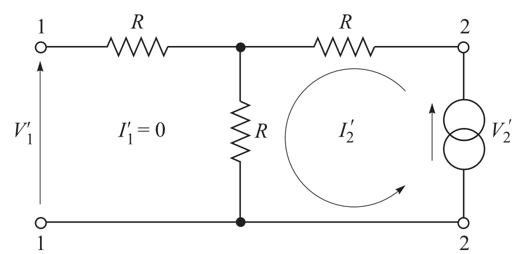


Fig. 280. $I_1' = 0$, hence $V_1' = RI_2'$.

But also note that, although the T network itself is the same in both figures, the values of the voltages and currents will be completely different in the two figures. Therefore, just for this discussion, let us denote the values in Fig. 280 by I_1', I_2', V_1', and V_2', as shown. Then the equations will be, for the given arrow directions,

$$\text{for Fig. 279:} \qquad \left\{ \begin{aligned} 2RI_1 + RI_2 &= V_1 \\ RI_1 + 2RI_2 &= 0 \end{aligned} \right\} \text{(case of } V_2 = 0\text{)}$$

$$\text{for Fig. 280:} \qquad \left\{ \begin{aligned} I_2' &= V_2'/2R \\ V_1' = RI_2' &= V_2'/2 \end{aligned} \right\} \text{(case of } I_1' = 0\text{)}$$

Now, upon making use of the above relationships and eqs. (515) through (518), you should find that

$$h_{11} = \frac{V_1}{I_1} = 3R/2 \qquad h_{21} = \frac{I_2}{I_1} = -1/2$$

$$h_{12} = \frac{V_1'}{V_2'} = 1/2 \qquad h_{22} = \frac{I_2'}{V_2'} = 1/2R$$

Thus, for the network of Fig. 278, h_{21} has a *negative* value. This happens because of the particular definitions of the "positive senses" laid down in Fig. 277.

To show this, suppose, in Fig. 279, that the positive sense of *both* I_1 and I_2 had been given to be in the *clockwise* sense. In that case, note that the equations for Fig. 279 would be

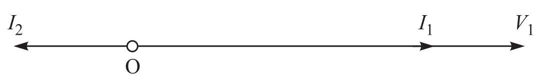
$$2RI_1 - RI_2 = V_1$$
$$-RI_1 + 2RI_2 = 0$$

in which case

$$h_{21} = \frac{I_2}{I_1} = \frac{R\begin{vmatrix} 2 & V_1 \\ -1 & 0 \end{vmatrix}}{R\begin{vmatrix} V_1 & -1 \\ 0 & 2 \end{vmatrix}} = +1/2$$

illustrating how the sign of a parameter can depend upon the positive senses originally chosen in Fig. 277.

Let us be reminded that our procedures are based upon the fact that sinusoidal voltages and currents can be represented as complex numbers and manipulated as vectors on the complex plane. Thus the vector diagram for the purely resistive circuit of Fig. 279, taking V_1 as the reference vector, would have the form



in which the lengths of the vectors represent the magnitudes of the rms values of the voltages and currents. In this particular case the diagram shows that I_2 is 180 degrees out of phase with respect to V_1 (and also I_1, since I_1 is in phase with V_1 in this case). This fact is shown algebraically by eq. (517); thus

$$I_2 = h_{21}I_1 = -I_1/2$$

Lastly, the h-parameter equations for the network inside the box in Fig. 278 are, by eqs. (506) and (507), equal to

$$V_1 = (3R/2)I_1 + (1/2)V_2$$
$$I_2 = -(1/2)I_1 + (1/2R)V_2$$

The following problem is another example of how the contents of the box in Fig. 277 determine the sign of a parameter.

Problem 246

Let the T-network of Fig. 278 be replaced with the transformer-coupled circuits shown in Figs. 281 and 282. In both figures the outputs are short-circuited (making $V_2 = 0$), with reference "grounds" as shown.

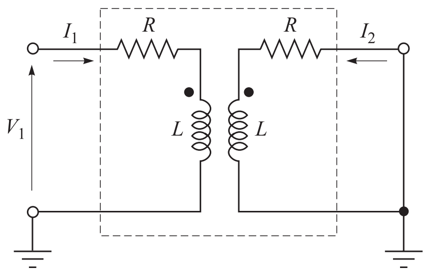


Fig. 281

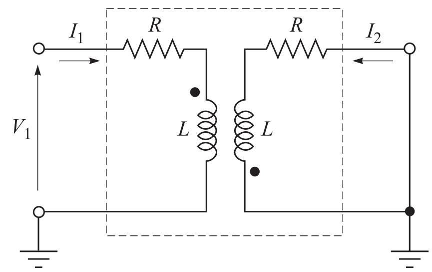


Fig. 282

Note that the ONLY DIFFERENCE between the two figures is in the placement of the transformer ‘‘polarity dots.’’ Also note that, in both figures, the positive sense of I_1 is given to be in the cw sense while the positive sense of I_2 is given to be in the ccw sense.

With these given details in mind, and upon referring back to eq. (385) in section 10.3, find the values of h_{21}, first for Fig. 281 and then for Fig. 282.

Now let’s go on to the second item we set out to explore, which concerns an ‘‘equivalent network’’ for a two-port, given the h-parameters of the two-port.

Let us begin by noting that any network that satisfies the DEFINING EQUATIONS, (506) and (507), can be considered to be an ‘‘equivalent h-parameter network’’ for the general two-port of Fig. 277. (For convenience here we’ve repeated eqs. (506) and (507) below.)

$$V_1 = h_{11}I_1 + h_{12}V_2 \tag{506}$$

$$I_2 = h_{21}I_1 + h_{22}V_2 \tag{507}$$

With the above in mind, we’ll now show that Fig. 283 is a true, valid ‘‘equivalent h-parameter network’’ for a general two-port.

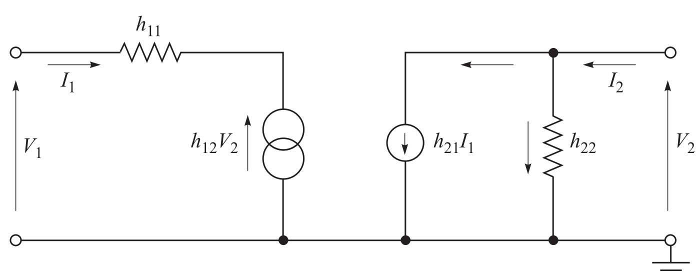


Fig. 283. *h*-parameter equivalent network.

Note that the network has both a dependent voltage generator* and a dependent current generator. Inspection of the network shows that the voltage equation around the left-hand loop is

$$h_{11}I_1 = -h_{12}V_2 + V_1$$

* A ‘‘dependent’’ generator is one whose output is dependent upon the value of a voltage or a current at another location in the network. (For example, the voltage of the left-hand generator in Fig. 283 is dependent upon the value of voltage V_2.)

and the sum of the currents on the right-hand side is

$$I_2 = h_{21}I_1 + h_{22}V_2$$

which are eqs. (506) and (507), thus proving that Fig. 283 is a valid h-parameter equivalent network for the general two-port configuration of Fig. 277.

Lastly, in regard to the third item, the answer simply is that the value of an output current I_2 (that would flow in a passive load impedance of Z_L ohms) would depend upon *the values of the h-parameters* which, in turn, would depend upon the particular network *inside the box* in Fig. 277. Thus we must here postpone any calculation of I_2 until the signs and magnitudes of the h-parameters are known in each individual situation.

11.8 Matrix Conversion Chart for the Two-Port Network

So far we've emphasized the practical importance of the h-parameters. While it is true that the h-parameters are probably the most widely used, it's also true that in some cases it's more convenient to work with one of the other parameters, that is, the z, y, g, or a. Also, manufacturers' data sheets may sometimes be given in terms of a parameter other than the h. Hence, given any one of the five parameters, it's necessary to be able to find the corresponding values of any one of the other four. This can be done as follows.

In section 11.6 we wrote the equations for the equivalent networks in matrix form, which, for convenience here, we've summarized below:

in z parameters:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{519}$$

in y parameters:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{520}$$

in h parameters:

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix} \tag{521}$$

in g parameters:

$$\begin{bmatrix} I_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ I_2 \end{bmatrix} \tag{522}$$

in a parameters:

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix} \tag{523}$$

Since, for any given case, these equations all represent the same network inside the box of Fig. 277, it follows that definite relationships must exist among the z, y, h, g, and a parameters.

Thus, in the "z" matrix of eq. (519) it must be possible to express the z quantities in terms of, say, the g parameters. Or, in the "h" matrix of eq. (521), it must be possible to express the h quantities in terms of, say, the a parameters, and so on.

This can be done, and the relationships can be shown in the form of a *matrix conversion chart*.

Matrix conversion chart

$$[\mathbf{z}] = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} \frac{y_{22}}{dy} & \frac{-y_{12}}{dy} \\ \frac{-y_{21}}{dy} & \frac{y_{11}}{dy} \end{bmatrix} = \begin{bmatrix} \frac{dh}{h_{22}} & \frac{h_{12}}{h_{22}} \\ \frac{-h_{21}}{h_{22}} & \frac{1}{h_{22}} \end{bmatrix} = \begin{bmatrix} \frac{1}{g_{11}} & \frac{-g_{12}}{g_{11}} \\ \frac{g_{21}}{g_{11}} & \frac{dg}{g_{11}} \end{bmatrix} = \begin{bmatrix} \frac{a_{11}}{a_{21}} & \frac{da}{a_{21}} \\ \frac{1}{a_{21}} & \frac{a_{22}}{a_{21}} \end{bmatrix}$$

$$[\mathbf{y}] = \begin{bmatrix} \frac{z_{22}}{dz} & \frac{-z_{12}}{dz} \\ \frac{-z_{21}}{dz} & \frac{z_{11}}{dz} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{h_{11}} & \frac{-h_{12}}{h_{11}} \\ \frac{h_{21}}{h_{11}} & \frac{dh}{h_{11}} \end{bmatrix} = \begin{bmatrix} \frac{dg}{g_{22}} & \frac{g_{12}}{g_{22}} \\ \frac{-g_{21}}{g_{22}} & \frac{1}{g_{22}} \end{bmatrix} = \begin{bmatrix} \frac{a_{22}}{a_{12}} & \frac{-da}{a_{12}} \\ \frac{-1}{a_{12}} & \frac{a_{11}}{a_{12}} \end{bmatrix}$$

$$[\mathbf{h}] = \begin{bmatrix} \frac{dz}{z_{22}} & \frac{z_{12}}{z_{22}} \\ \frac{-z_{21}}{z_{22}} & \frac{1}{z_{22}} \end{bmatrix} = \begin{bmatrix} \frac{1}{y_{11}} & \frac{-y_{12}}{y_{11}} \\ \frac{y_{21}}{y_{11}} & \frac{dy}{y_{11}} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} \frac{g_{22}}{dg} & \frac{-g_{12}}{dg} \\ \frac{-g_{21}}{dg} & \frac{g_{11}}{dg} \end{bmatrix} = \begin{bmatrix} \frac{a_{12}}{a_{22}} & \frac{da}{a_{22}} \\ \frac{-1}{a_{22}} & \frac{a_{21}}{a_{22}} \end{bmatrix}$$

$$[\mathbf{g}] = \begin{bmatrix} \frac{1}{z_{11}} & \frac{-z_{12}}{z_{11}} \\ \frac{z_{21}}{z_{11}} & \frac{dz}{z_{11}} \end{bmatrix} = \begin{bmatrix} \frac{dy}{y_{22}} & \frac{y_{12}}{y_{22}} \\ \frac{-y_{21}}{y_{22}} & \frac{1}{y_{22}} \end{bmatrix} = \begin{bmatrix} \frac{h_{22}}{dh} & \frac{-h_{12}}{dh} \\ \frac{-h_{21}}{dh} & \frac{h_{11}}{dh} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \frac{a_{21}}{a_{11}} & \frac{-da}{a_{11}} \\ \frac{1}{a_{11}} & \frac{a_{12}}{a_{11}} \end{bmatrix}$$

$$[\mathbf{a}] = \begin{bmatrix} \frac{z_{11}}{z_{21}} & \frac{dz}{z_{21}} \\ \frac{1}{z_{21}} & \frac{z_{22}}{z_{21}} \end{bmatrix} = \begin{bmatrix} \frac{-y_{22}}{y_{21}} & \frac{-1}{y_{21}} \\ \frac{-dy}{y_{21}} & \frac{-y_{11}}{y_{21}} \end{bmatrix} = \begin{bmatrix} \frac{-dh}{h_{21}} & \frac{-h_{11}}{h_{21}} \\ \frac{-h_{22}}{h_{21}} & \frac{-1}{h_{21}} \end{bmatrix} = \begin{bmatrix} \frac{1}{g_{21}} & \frac{g_{22}}{g_{21}} \\ \frac{g_{11}}{g_{21}} & \frac{dg}{g_{21}} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

The explanation of the chart is as follows.

First we have the *basic parameter matrices*; thus

$$[\mathbf{z}] = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}, \qquad [\mathbf{y}] = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$

and so on, for eqs. (521) through (523).

Next, "d" indicates the *determinant value* of a basic matrix; thus

$$dz = \det[\mathbf{z}] = \begin{vmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{vmatrix} = z_{11}z_{22} - z_{12}z_{21}$$

$$dy = \det[\mathbf{y}] = \begin{vmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{vmatrix} = y_{11}y_{22} - y_{12}y_{21}$$

and so on, in the same way,

$$dh = \det[\mathbf{h}] = h_{11}h_{22} - h_{12}h_{21}$$

$$dg = \det[\mathbf{g}] = g_{11}g_{22} - g_{12}g_{21}$$

$$da = \det[\mathbf{a}] = a_{11}a_{22} - a_{12}a_{21}$$

The chart is useful because it allows us to find the value of any parameter in terms of any other parameter. For example, suppose, in a certain case, that we wish to work in terms of, say, the z parameters, but are given only the values of, say, the h parameters. From the definition of equal matrices in section 11.1, inspection of the chart then shows that values of the z parameters are calculated from corresponding values of the h parameters by means of the formulas

$$z_{11} = dh/h_{22}, \qquad z_{12} = h_{12}/h_{22}, \qquad z_{21} = -h_{21}/h_{22}, \qquad z_{22} = 1/h_{22}$$

You may be interested in learning how the various entries listed in the chart were arrived at. This was done by making use of the basic equations summarized in eqs. (519) through (523). For example, let's begin with, say, the basic eq. (521); thus

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} V_1 \\ I_2 \end{bmatrix}$$

which becomes, after multiplying both sides by the inverse matrix $[\mathbf{h}]^{-1}$,

$$\begin{bmatrix} I_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}^{-1} \begin{bmatrix} V_1 \\ I_2 \end{bmatrix}$$

and comparison of this last equation with eq. (522) shows that

$$\begin{gathered} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}^{-1} \\ [\mathbf{g}] = [\mathbf{h}]^{-1} \end{gathered} \tag{524}$$

Now find the inverse of the 2×2 h-parameter matrix as indicated. (In connection with finding the inverse of the 2×2 matrix, you may wish to review problem 232 in section 11.3.) Doing this, you should find that

$$[\mathbf{g}] = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{h_{22}}{dh} & \dfrac{-h_{12}}{dh} \\ \dfrac{-h_{21}}{dh} & \dfrac{h_{11}}{dh} \end{bmatrix} \tag{525}$$

Equation (525) verifies that the g-matrix $[\mathbf{g}]$ can be written in terms of the h-parameters in the form shown in the chart. From the condition required for equal matrices given in section 11.1, inspection of eq. (525) shows that h-parameters can be converted into equivalent g-parameters by means of the formulas

$$g_{11} = h_{22}/dh \qquad g_{12} = -h_{12}/dh \qquad g_{21} = -h_{21}/dh \qquad g_{22} = h_{11}/dh$$

Problem 247

Suppose the h-parameter values for an unknown network inside the box in Fig. 277 are found to be

$$h_{11} = 850 \text{ ohms} \qquad h_{12} = (8)10^{-3}$$
$$h_{21} = 26 \qquad h_{22} = (4)10^{-4} \text{ mhos}$$

Using the matrix conversion chart, find the equivalent values of the g parameters.

Problem 248
Making use of the defining eqs. (494) and (495), and (506) and (507), prove that the values of the z-parameters in terms of the h-parameters, as given in the first row of the conversion chart, are correct.

Problem 249
Convert the following g-parameter values into equivalent z-parameter values:

$$g_{11} = 0.068 \text{ mhos} \qquad g_{12} = -0.073$$
$$g_{21} = -228 \qquad g_{22} = 8755 \text{ ohms}$$

Problem 250
Prove that $[\mathbf{h}] = [\mathbf{g}]^{-1}$ by making use of eqs. (508) and (511) and section 11.3.

Problem 251
Convert the h-parameter values in problem 247 into equivalent z-parameter values.

11.9 Matrix Operations for Interconnected Two-Ports

In circuit design work it's often possible to consider a complex system as being composed of an interconnection of separate, individual two-ports. This approach can greatly simplify the work, because it is usually much easier to deal with each such "building block" individually, and then connect them together to form the whole, than it is to deal with the whole complex system as a single unit.

There are five basic ways of interconnecting individual two-ports to form a single equivalent two-port. These five configurations are known as the SERIES, the PARALLEL, the SERIES-PARALLEL, the PARALLEL-SERIES, and the CASCADE connections. We take up each of the five modes in this section, beginning with the series connection. The parameter (z, y, h, g, or a) that will be used in any given case will depend upon the type of connection (series, parallel, and so on).

SERIES CONNECTION OF TWO-PORTS

Two two-port networks, "a" and "b," are said to be connected in series if the two input circuits are in series and the two output circuits are in series. The basic *series connection* of two-ports is thus as shown in Fig. 284.

In Fig. 284 note that V_1 and V_2 are the input and output voltages for the *overall* composite network. It follows that V_1 will divide between the two series-connected inputs and V_2 will divide between the two series-connected outputs, as shown.

Let us now apply the *z-form* of eqs. (494) and (495), to the individual networks a and b in the figure. Let subscript "a" apply to network a and subscript "b" apply to network b. In the figure, note that the same current I_1 flows through both inputs and the same

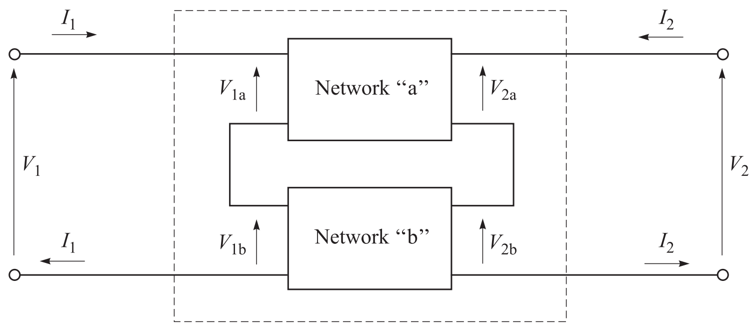


Fig. 284. Series connection of two-ports.

current I_2 flows through both outputs; thus, from eqs. (494) and (495) we have eqs. (526) and (527):

for network "a"
$$V_{1a} = z_{11a}I_1 + z_{12a}I_2 \tag{526}$$
$$V_{2a} = z_{21a}I_1 + z_{22a}I_2 \tag{527}$$
that is,
$$\begin{bmatrix} V_{1a} \\ V_{2a} \end{bmatrix} = \begin{bmatrix} z_{11a} & z_{12a} \\ z_{21a} & z_{22a} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{528}$$

for network "b"
$$V_{1b} = z_{11b}I_1 + z_{12b}I_2 \tag{529}$$
$$V_{2b} = z_{21b}I_1 + z_{22b}I_2 \tag{530}$$
that is,
$$\begin{bmatrix} V_{1b} \\ V_{2b} \end{bmatrix} = \begin{bmatrix} z_{11b} & z_{12b} \\ z_{21b} & z_{22b} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{531}$$

Now write down the sum of eqs. (526) and (529) and also the sum of eqs. (527) and (530). Doing this, and noting, from Fig. 284, that $V_{1a} + V_{1b} = V_1$, and also that $V_{2a} + V_{2b} = V_2$, we find, for the composite network consisting of two series-connected two-ports, that

$$V_1 = (z_{11a} + z_{11b})I_1 + (z_{12a} + z_{12b})I_2 \tag{532}$$
$$V_2 = (z_{21a} + z_{21b})I_1 + (z_{22a} + z_{22b})I_2 \tag{533}$$

or, in matrix form,

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} (z_{11a} + z_{11b}) & (z_{12a} + z_{12b}) \\ (z_{21a} + z_{21b}) & (z_{22a} + z_{22b}) \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{534}$$

Equations (532), (533), and (534) show that *series-connected two-ports* can be replaced with a *single equivalent two-port* whose z parameters are the *sum of the corresponding z-parameters of the individual two-ports*. This fact can be summarized by writing, for series-connected two-ports:

$$[\mathbf{z}] = [\mathbf{z}_a] + [\mathbf{z}_b] \tag{535}$$

where $[\mathbf{z}_a]$ and $[\mathbf{z}_b]$ are the impedance matrices of the individual series two-ports in eqs. (528) and (531), and where $[\mathbf{z}]$ is the impedance matrix of the single equivalent two-port, which appears in eq. (534). Note that (in accordance with the rule for addition of matrices laid down in section 11.1) the sum of the two impedance matrices in eqs. (528) and (531) does produce the impedance matrix of eq. (534), which is what eq. (535) says. It should also be apparent from the preceding work that any number n of series two-ports can be replaced by a single equivalent two-port whose impedance matrix is equal to the *sum of the impedance matrices of the n individual two-ports*.

If the values of the z-parameters are not known in a given case but other parameter values are known, the values of the z-parameters can then be found by use of the conversion chart in section 11.8. For example, suppose the h-parameter values are known but the z-parameter values are unknown, inspection of the conversion chart shows that the required z-parameter values can be found, given the h-values, by means of the formulas

$$z_{11} = dh/h_{22}, \qquad z_{12} = h_{12}/h_{22}, \qquad z_{21} = -h_{21}/h_{22}, \qquad z_{22} = 1/h_{22}$$

PARALLEL CONNECTION OF TWO-PORTS

Two two-port networks, a and b, are *connected in parallel* if the input lines are connected in parallel and the output lines are connected in parallel. The basic *parallel connection* of two-ports is therefore as shown in Fig. 285.

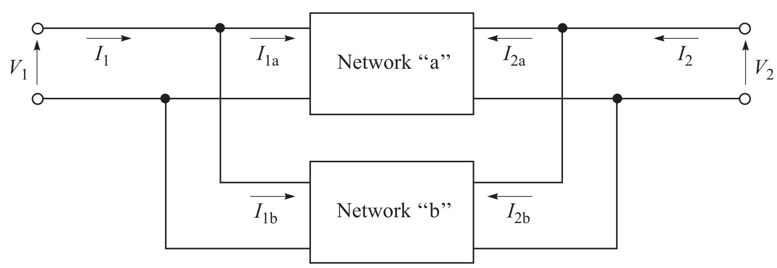


Fig. 285. Parallel connection of two-ports.

First, for the above parallel connection, note that both networks have the same input voltage V_1 and the same output voltage V_2. Next, regarding currents, inspection of Fig. 285 shows that, for the parallel connection,

$$I_1 = I_{1a} + I_{1b} \qquad \text{and} \qquad I_2 = I_{2a} + I_{2b}$$

or, in matrix form,

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} I_{1a} + I_{1b} \\ I_{2a} + I_{2b} \end{bmatrix} = \begin{bmatrix} I_{1a} \\ I_{2a} \end{bmatrix} + \begin{bmatrix} I_{1b} \\ I_{2b} \end{bmatrix}$$

which, upon applying eq. (520) to each of the two right-hand matrices, can be written as

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11a} & y_{12a} \\ y_{21a} & y_{22a} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} y_{11b} & y_{12b} \\ y_{21b} & y_{22b} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

and therefore

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} (y_{11a} + y_{11b}) & (y_{12a} + y_{12b}) \\ (y_{21a} + y_{21b}) & (y_{22a} + y_{22b}) \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{536}$$

Equation (536) shows that *parallel-connected two-ports* can be replaced, for analysis, with a single equivalent two-port whose y-parameters are equal to the *sum of the corre-*

sponding y parameters of the individual two-ports. This fact is summarized by writing that, for parallel-connected two-ports,

$$[\mathbf{y}] = [\mathbf{y}_\mathrm{a}] + [\mathbf{y}_\mathrm{b}] \tag{537}$$

where $[\mathbf{y}_\mathrm{a}]$ and $[\mathbf{y}_\mathrm{b}]$ are the admittance matrices of the individual parallel two-ports in the equation at the top of the page, and where $[\mathbf{y}]$ is the admittance matrix of the single equivalent two-port which appears in eq. (536). Note that, in accordance with the rule for the addition of matrices, the sum of the two admittance matrices in the equation at the top of the page does produce the admittance matrix of eq. (536), which is what eq. (537) says. It's also apparent that any number n of parallel two-ports can, for purposes of analysis, be replaced by a single equivalent two-port whose admittance matrix is equal to the *sum of the admittance matrices of the n individual two-ports*.

If, in a given case, the values of the y-parameters are not known but the values of another set of parameters are known, then the values of the y-parameters can be found by inspection of the conversion chart of section 11.8. For example, if the values of, say, the g parameters are known, then inspection of the chart shows that

$$y_{11} = dg/g_{22}, \qquad y_{12} = g_{12}/g_{22}, \qquad y_{21} = -g_{21}/g_{22}, \qquad y_{22} = 1/g_{22}$$

SERIES-PARALLEL AND PARALLEL-SERIES CONNECTIONS OF TWO-PORTS

In the *series-parallel connection* the two networks are connected in series on the input side and in parallel on the output side. This is illustrated in Fig. 286.

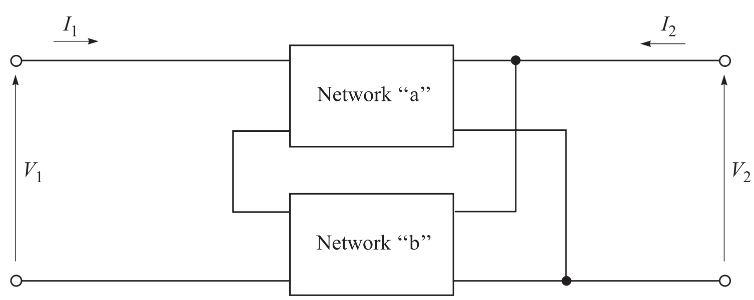


Fig. 286. Series-parallel connection of two-ports.

As will be shown in one of your practice problems, the series-parallel combination of Fig. 286 can be replaced, for purposes of analysis, by a single equivalent two-port network *having h-parameters equal to the sums of the corresponding h-parameters of the individual two-ports*. Thus (where subscript e denotes the h-parameters of the equivalent two-port)

$$h_{11\mathrm{e}} = h_{11\mathrm{a}} + h_{11\mathrm{b}} \qquad h_{21\mathrm{e}} = h_{21\mathrm{a}} + h_{21\mathrm{b}}$$
$$h_{12\mathrm{e}} = h_{12\mathrm{a}} + h_{12\mathrm{b}} \qquad h_{22\mathrm{e}} = h_{22\mathrm{a}} + h_{22\mathrm{b}}$$

or, expressed in matrix notation,

$$[\mathbf{h}_\mathrm{e}] = [\mathbf{h}_\mathrm{a}] + [\mathbf{h}_\mathrm{b}]$$

where $[\mathbf{h}_a]$ and $[\mathbf{h}_b]$ are the h-parameter matrices of the individual series-parallel two-ports, and where $[\mathbf{h}_e]$ is the h-parameter matrix of the equivalent two-port. It's also apparent that the above remarks apply to any number n of series-parallel connected two-ports.

Next, in the *parallel-series connection* of two-port networks, the two networks a and b are connected in parallel on the input side and in series on the output side, as illustrated in Fig. 287.

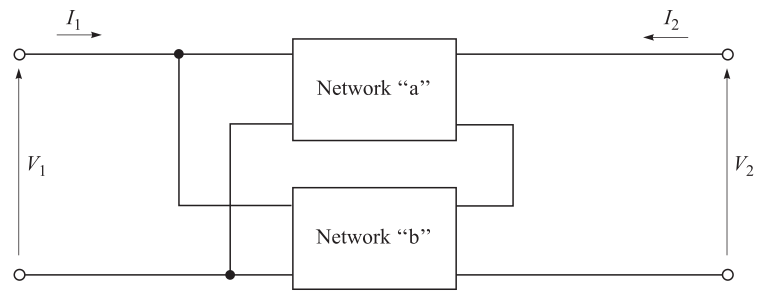


Fig. 287. Parallel-series connection of two-ports.

It can be shown that the parallel-series connection of two-ports can be replaced, for purposes of analysis, by a single equivalent two-port *having g-parameters equal to the sums of the corresponding g-parameters of the individual two-ports*; thus,

$$g_{11e} = g_{11a} + g_{11b} \qquad g_{21e} = g_{21a} + g_{21b}$$

$$g_{12e} = g_{12a} + g_{12b} \qquad g_{22e} = g_{22a} + g_{22b}$$

where "e" denotes the g-parameters of the equivalent two-port. These remarks apply to any number n of parallel-series-connected two-ports.

CASCADE CONNECTION OF TWO-PORTS

By definition, two-port networks are said to be connected in "cascade" if the output of the first is the input to the second, the output of the second is the input to the third, and so on. As an example, a cascade connection of three two-ports is shown in Fig. 288, in which we're using, for each stage, the standard notation of Fig. 277. Note that V_1 and I_1 are the input voltage and current to the cascade, and we're assuming the cascade is terminated in a load impedance of Z_L ohms.

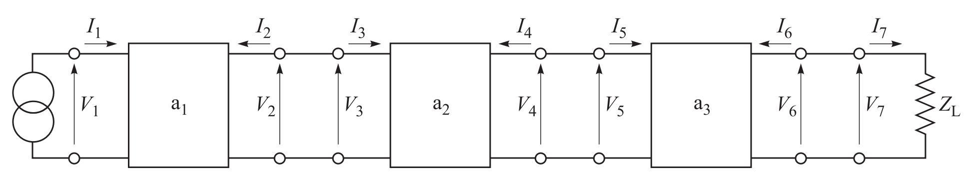


Fig. 288. Cascade connection of three two-ports.

In terms of the standard notation of Fig. 277, note that the negative of the output current of each network equals the input current to the next network; that is, from inspection of Fig. 288,

$$-I_2 = I_3 \qquad -I_4 = I_5 \qquad -I_6 = I_7$$

and also, from inspection of the figure,

$$V_2 = V_3, \qquad V_4 = V_5, \qquad V_6 = V_7$$

and so on, for any number of cascaded networks.

When two-ports are connected in cascade, it's convenient to use the "a" parameters, in which the output current of each network is written with the minus sign, that is, as a negative current, in the form of eq. (514) in section 11.6. Thus, letting V_1 and I_1 denote input voltage and current and V_2 and I_2 denote output voltage and current, the matrix equation for each individual network in a cascade of two-port networks will be, using a parameters from eq. (514),

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{a}] \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix}$$

where

$$[\mathbf{a}] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Referring to Fig. 288 we therefore have, starting at the left (the input end of the cascade), that the matrix equation for the first network is

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{a}_1] \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix} = [\mathbf{a}_1] \begin{bmatrix} V_3 \\ I_3 \end{bmatrix} \tag{538}$$

and then, since

$$\begin{bmatrix} V_3 \\ I_3 \end{bmatrix} = [\mathbf{a}_2] \begin{bmatrix} V_4 \\ -I_4 \end{bmatrix} = [\mathbf{a}_2] \begin{bmatrix} V_5 \\ I_5 \end{bmatrix}$$

eq. (538) becomes

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{a}_1][\mathbf{a}_2] \begin{bmatrix} V_5 \\ I_5 \end{bmatrix} \tag{539}$$

and then, since

$$\begin{bmatrix} V_5 \\ I_5 \end{bmatrix} = [\mathbf{a}_3] \begin{bmatrix} V_6 \\ -I_6 \end{bmatrix} = [\mathbf{a}_3] \begin{bmatrix} V_7 \\ I_7 \end{bmatrix}$$

eq. (539) becomes

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{a}_1][\mathbf{a}_2][\mathbf{a}_3] \begin{bmatrix} V_7 \\ I_7 \end{bmatrix} \tag{540}$$

which is the final matrix equation for the 3-network two-port cascade of Fig. 288.

It's clear that the foregoing procedure can be continued for any number of two-ports in cascade. Thus, if n two-ports are connected in cascade, and if V_1 and I_1 are the input voltage and current to the cascade and V_o and I_o are the final output voltage and current, then eq. (540) extends to the general form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{a}_1][\mathbf{a}_2][\mathbf{a}_3] \cdots [\mathbf{a}_n] \begin{bmatrix} V_\mathrm{o} \\ I_\mathrm{o} \end{bmatrix} \tag{541}$$

that is

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{A}] \begin{bmatrix} V_\mathrm{o} \\ I_\mathrm{o} \end{bmatrix} \tag{542}$$

where [**A**] is the overall network matrix for n two-ports in cascade. We have thus deduced the important fact that

> The *overall network matrix* for two-port networks in cascade is equal to the PRODUCT of the matrices of the individual two-ports if a parameter values are used.

Let us note that the final signs of the currents and voltages in Fig. 288 will depend, in any given case, upon the networks inside the boxes (see ''third item'' discussion in final paragraph of section 11.7).

Problem 252
Prove that the two two-ports in Fig. 286 can be replaced by a single equivalent two-port whose h-parameters are equal to the sums of the corresponding h-parameters of the individual two-ports.

Problem 253
If three individual two-port networks are connected in parallel, express the parameters of the single equivalent two-port in terms of the z-parameters of the individual two-ports.

Problem 254
Two identical two-ports are connected in cascade. Write the matrix expression for the single equivalent two-port in terms of h-parameters.

Problem 255
In problem 254, show that the value of the input current I_1, in terms of the h-parameters, is equal to

$$I_1 = \frac{[1 + h_{11}h_{22} + (1 + dh)h_{22}V_L]V_1}{(1 + dh)h_{11} + [(dh)^2 + h_{11}h_{22}]Z_L}$$

Problem 256
Find the expression for the inverse matrix, $[\mathbf{h}]^{-1}$, if the answer to problem 254 is written in the form

$$\begin{bmatrix} V_o \\ I_o \end{bmatrix} = [\mathbf{h}] \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

11.10 Notes Regarding the Interconnection Formulas

In the preceding work dealing with the interconnection of two-ports, we've assumed that each two-port operates in the normal ''balanced'' mode, meaning that both input terminals carry the same current I_1 and both output terminals carry the same current I_2, as shown in Fig. 289. If we have a situation in which this condition is not true, then we do not

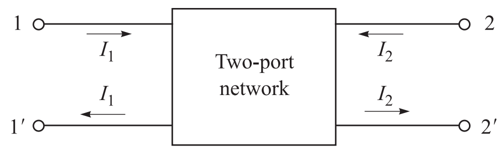


Fig. 289

have the normal two-port network defined by Fig. 277, and the equations we've developed for interconnected two-ports are not valid.

To understand how such current imbalances can occur, consider the series connection of two two-port networks, a and b, shown in Fig. 290, in which we'll concentrate our attention on the input and output currents of network a.

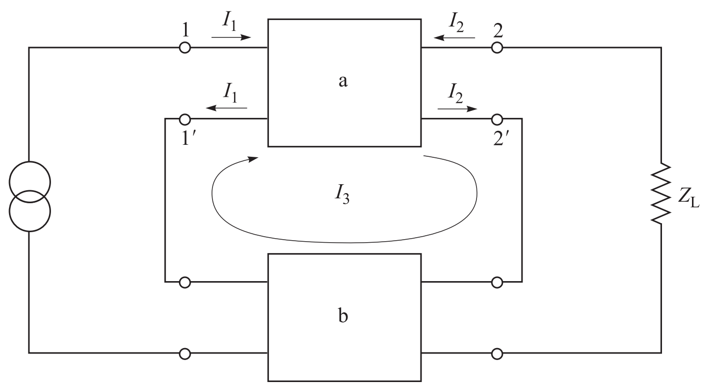


Fig. 290

In Fig. 290, *depending upon the type of networks that blocks a and b represent*, it is possible that an unwanted loop current I_3 can flow in the interconnection, as shown.

If this happens, inspection of the figure shows that, in general, the current at terminal 1 will not be equal to the current at terminal $1'$. In such a case, network a is no longer a normal balanced two-port, and the two-port equations previously derived will not be valid.

Therefore, the condition that the previously derived formulas for series-connected two-ports be valid is that *no circulating current can exist between the networks*, that is, that $I_3 = 0$ in Fig. 290.

Fortunately, a simple test can be applied to determine whether the basic equation (534), found in section 11.9, is valid for a given series connection of two-ports. The test setup is shown in Figs. 291 and 292.

The test is carried out as follows. The first step, shown in Fig. 291, is to connect the inputs of the two networks in series, leaving the outputs open-circuited, as shown. Now imagine a signal voltage V_s to be applied to the series-connected inputs, as shown. Then eq. (534) is valid if, and only if, *the voltage V is zero*. The second step in the test is to apply the test signal V_s to the series-connected outputs with the inputs open-circuited, as shown in Fig. 292. Again, eq. (534) is valid only if $V = 0$.

If the foregoing test shows that the two given two-ports will *not* satisfy the requirement that $V = 0$ for the series connection, they can sometimes be put into a different but equivalent form, for which V will be zero.

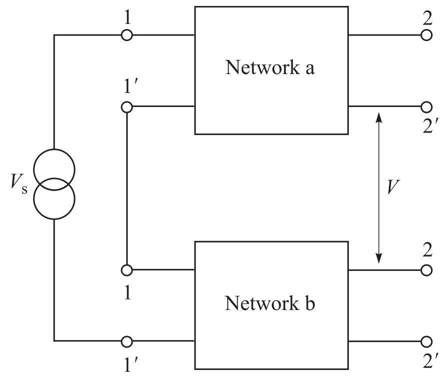


Fig. 291

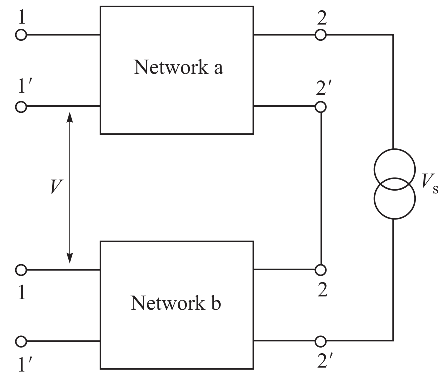


Fig. 292

As an example of this, consider first Fig. 293, which shows the test setup of Fig. 291 for a proposed series connection of two particular two-ports. In Fig. 293, note that V will *not* be equal to zero, because of the voltage drop across R, and hence Eq. (534) will not apply if the two networks are connected in series. If, however, we convert the top network into its "T" equivalent (section 9.2), then Fig. 293 becomes Fig. 294, for which V *does* equal zero. Hence, if Fig. 293 is put into the equivalent form of Fig. 294, then eq. (534) will be valid for the series connection of the two networks.

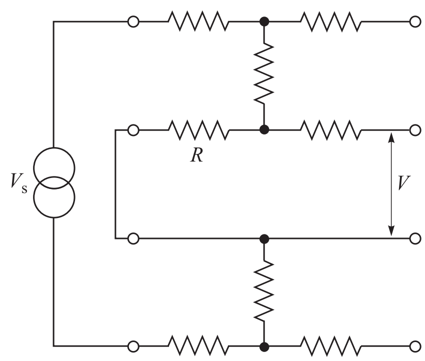


Fig. 293

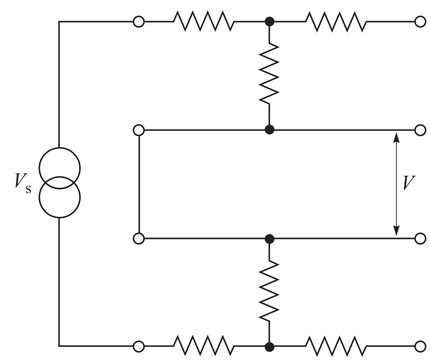


Fig. 294

It's also possible that an unwanted loop current can flow in the *parallel* connection of Fig. 285. The test setup to determine whether eq. (536) is valid for a parallel connection of two two-port networks is shown in Figs. 295 and 296.

To apply the test we begin with Fig. 295, in which a signal voltage V_s is applied to the parallel inputs, with the two outputs, previously connected in parallel, now disconnected and with each short-circuited, as shown. The voltage V is now calculated or measured. The operation is then repeated in the reverse direction, as shown in Fig. 296. Only if $V = 0$ for both test conditions is eq. (536) valid for the parallel connection of the given two-ports.

If V does not equal zero in a given case, it may be possible to transform one, or both, of the two-ports into an equivalent network for which $V = 0$, in a manner such as was done for the series connection of Fig. 293.

In the preceding discussions we've found that if a test setup shows V not equal to zero, we must then try to change one or both of the networks into an equivalent form for which

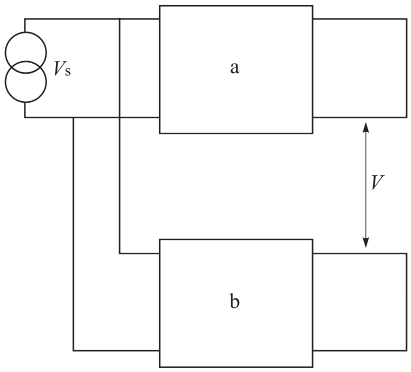


Fig. 295

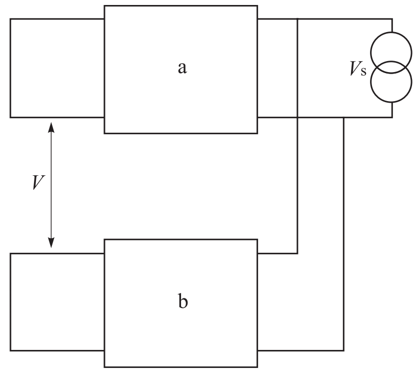


Fig. 296

V will be equal to zero (that is, if we wish to be able to use the formulas found in section 11.9).

In case an equivalent network cannot be found for which $V = 0$, it is always possible, theoretically at least, to use one or more *ideal transformers* to insure that no undesired circulating current can flow in a proposed interconnection. This is illustrated in Fig. 297 for the series connection, where T is assumed to be an *ideal transformer with 1:1 turns ratio*.

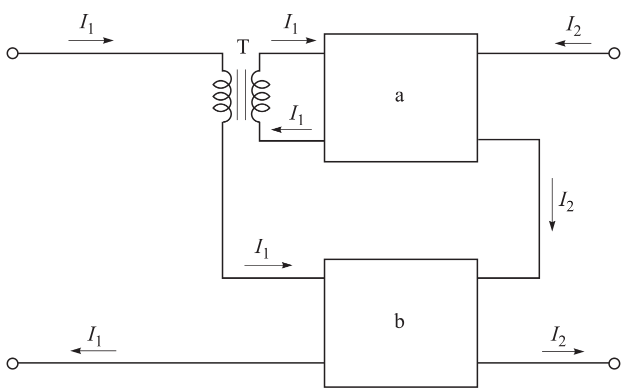


Fig. 297

In regard to the transformer T, the impedance seen between the primary terminals is the same as the impedance seen looking into network a, because T is an ideal "1 to 1" transformer (section 10.6). Therefore, in Fig. 297, the input impedances to networks a and b are connected in series as far as an ac signal is concerned. Also, since T is an ideal 1:1 transformer, the same signal current I_1 flows in both the primary and secondary sides. The important point to see now, in Fig. 297, is that *the transformer forces the currents to be equal in both input terminals to network a*, and hence no undesired loop current I_3 can flow, as might be possible in Fig. 290, depending upon the nature of networks a and b.

Transformer T in Fig. 297 must be as nearly ideal as possible. Whether or not this requirement can be met in a practical case will depend upon such factors as frequency, power level, cost, and limitations as to physical size.

In closing, it should be noted that the problem of possible current imbalance does not arise in the cascade connection of two-ports, Fig. 288. Thus eq. (541) applies without reservation to the cascade connection.

11.11 Some Basic Applications of the Formulas

Let us begin by finding a matrix expression for a single ''shunt-connected'' impedance Z, which we can imagine to be the contents of the box of Fig. 277, as shown in Fig. 298.

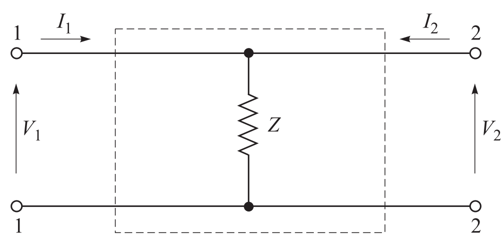


Fig. 298

In the figure we can, in principle, represent the contents of the box in terms of any one of the five parameters z, y, h, g, or a, but here let us suppose we elect to use the z-parameters. To do this we make use of the basic ''z-parameter'' equations, (494) and (495) in section 11.6, as follows, in which, from direct inspection of Fig. 298, we see that

$$\text{by eq. (494) for } I_2 = 0, \quad z_{11} = V_1/I_1 = Z$$

$$\text{by eq. (495) for } I_2 = 0, \quad z_{21} = V_2/I_1 = V_1/I_1 = Z$$

$$\text{by eq. (494) for } I_1 = 0, \quad z_{12} = V_1/I_2 = V_2/I_2 = Z$$

$$\text{by eq. (495) for } I_1 = 0, \quad z_{22} = V_2/I_2 = Z$$

and thus, for the single shunt impedance Z of Fig. 298, eqs. (494) and (495) become

$$V_1 = ZI_1 + ZI_2$$

$$V_2 = ZI_1 + ZI_2$$

or, in matrix notation,

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z & Z \\ Z & Z \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

Thus the matrix representation of a single shunt impedance Z, in the form of Fig. 298, is given by

$$[\mathbf{Z}] = \begin{bmatrix} Z & Z \\ Z & Z \end{bmatrix} \tag{543}$$

Next, suppose the contents of the box in Fig. 277 consisted of a single ''series-connected'' impedance Z, as in Fig. 299.

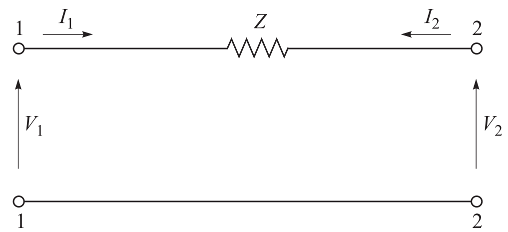


Fig. 299

Now, in conformity with the notation used in the figure, we have, upon applying Ohm's law,

viewed from terminals (1, 1) : $I_1 = (V_1 - V_2)/Z = V_1/Z - V_2/Z$

viewed from terminals (2, 2) : $I_2 = (V_2 - V_1)/Z = V_2/Z - V_1/Z$

Or, if we wish, since $1/Z = Y =$ the admittance of the series element, eqs. (500) and (501) become, for Fig. 299,

$$I_1 = YV_1 - YV_2$$
$$I_2 = -YV_1 + YV_2$$

or, in matrix notation,

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} Y & -Y \\ -Y & Y \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

Thus the matrix representation of a single series impedance Z (or series admittance $Y = 1/Z$) is given by

$$[\mathbf{Y}] = \begin{bmatrix} Y & -Y \\ -Y & Y \end{bmatrix} \tag{544}$$

Next, let's find the matrix representation for the case in which a transistor* is used with an unbypassed emitter impedance of Z ohms, as shown in Fig. 300.

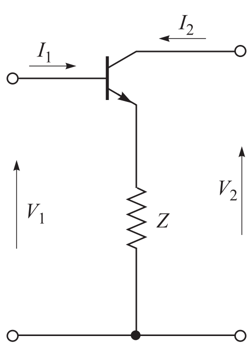


Fig. 300

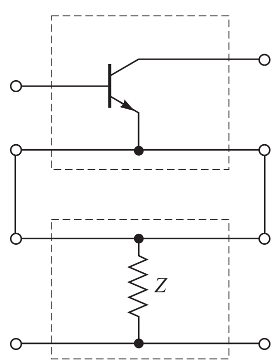


Fig. 301

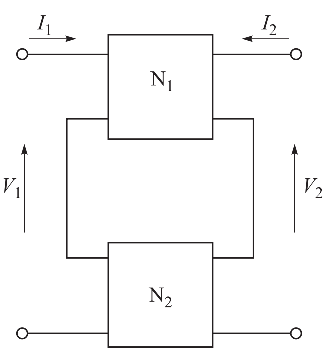


Fig. 302

As Figs. 301 and 302 show, Fig. 300 can be considered to consist of *two two-ports in series*; in Fig. 302 note that N_2 denotes the single shunt impedance Z and N_1 denotes the transistor.

* See note 30 in Appendix.

From our work in section 11.9 concerning series-connected two-ports, we know that the general matrix representation for Fig. 302 will be equal to the SUM of the matrix representations of N_1 and N_2. Thus, since we'll be dealing with the *sum* of two matrices, and since the matrix representation for N_2 is in terms of impedance (eq. (543)), it follows that the matrix representation for N_1 (the transistor) must *also* be expressed in terms of impedance. Therefore, upon making use of eq. (543) and the "impedance matrix" representation of a transistor (note 30 in Appendix), we have that the matrix equation for Fig. 300 is

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} (Z_{11}+Z) & (Z_{12}+Z) \\ (Z_{21}+Z) & (Z_{22}+Z) \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{545}$$

To continue with another example, let us find the matrix representation for a network consisting of a transistor, in the CE mode, using a collector-to-base feedback impedance of Z ohms (or, if we wish, $Y = 1/Z$ mhos), as shown in Fig. 303.

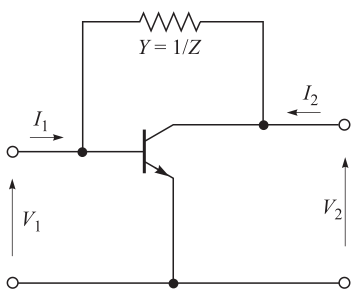


Fig. 303

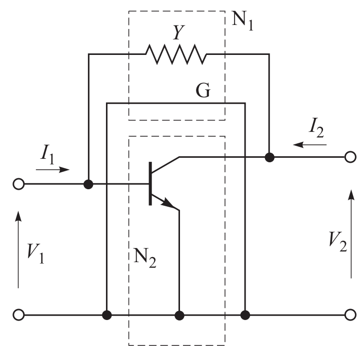


Fig. 304

Now, after some thought, we see that the feedback admittance Y can actually be considered to be a two-port network *in parallel* with the two-port representation of the transistor. This can be seen by redrawing Fig. 303 in the form of Fig. 304, in which we've added a fictitious ground lead G, to show more clearly that Y can be considered to be a two-port network in its own right.

Thus, in Fig. 304, the upper two-port network N_1 is the feedback admittance Y, while the lower two-port N_2 is the transistor itself; as the figure shows, N_1 is in parallel with N_2.

Since the two two-ports are connected in parallel, the admittance matrix of the overall equivalent two-port is the *sum of the admittance matrices* of the individual two-ports. Hence, if Y_{11}, Y_{12}, Y_{21}, and Y_{22} are the admittance parameters of the transistor, the matrix equation for Fig. 303 is (making use of eq. (544))

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} (Y_{11}+Y) & (Y_{12}-Y) \\ (Y_{21}-Y) & (Y_{22}+Y) \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{546}$$

Lastly, as another example, consider two impedances Z and Z', connected in the "L" configuration of Fig. 305.

As Fig. 306 shows, the L-network can be considered to be a *cascade* connection of two two-ports in the manner of Fig. 288. Hence the transmission characteristics for the L-network can be expressed in terms of the product of the a parameter matrices of the individual networks 1 and 2; thus

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = [\mathbf{a}_1][\mathbf{a}_2] \begin{bmatrix} V_o \\ I_o \end{bmatrix} \tag{547}$$

by eq. (541) in section 11.9.

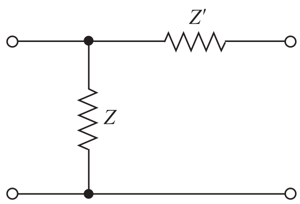


Fig. 305

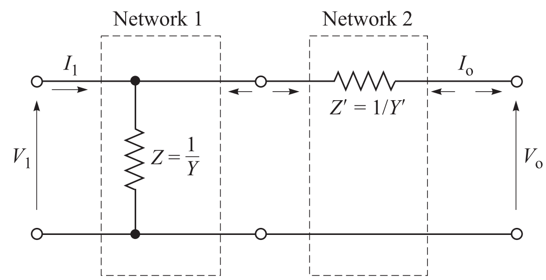


Fig. 306

Now let us make use of the conversion chart in section 11.8 to write the above *a*-parameter matrices in terms of Z and Z' (or Y and Y' if we wish). To do this, note that, for *network 1*, which is the single shunt form of Fig. 298, we've already found that (eq. (543))

$$[\mathbf{Z}] = \begin{bmatrix} Z & Z \\ Z & Z \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$$

and thus, for the particular case of network 1, it is true that

$$Z_{11} = Z_{12} = Z_{21} = Z_{22} = Z, \text{ hence } dZ = Z^2 - Z^2 = 0$$

and upon substituting these values into the fifth row of the conversion chart we have, for network 1, that

$$[\mathbf{a}_1] = \begin{bmatrix} 1 & 0 \\ 1/Z & 1 \end{bmatrix}$$

In a like manner for *network 2* (which is of the single series form of Fig. (299)), we've found that (eq. (544))

$$[\mathbf{Y}'] = \begin{bmatrix} Y' & -Y' \\ -Y' & Y' \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$$

and thus, for the particular case of network 2, it is true that

$$Y_{11} = Y', \quad Y_{12} = -Y', \quad Y_{21} = -Y', \quad Y_{22} = Y', \quad \text{hence } dY = 0$$

and upon substituting these values into the fifth row of the conversion chart we have, for network 2, that

$$[\mathbf{a}_2] = \begin{bmatrix} 1 & 1/Y' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & Z' \\ 0 & 1 \end{bmatrix}$$

Thus, substituting into eq. (547), we have that the transmission characteristics for the L-network of Fig. 305 can be expressed in the matrix form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/Z & 1 \end{bmatrix} \begin{bmatrix} 1 & Z' \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_o \\ I_o \end{bmatrix}$$

that is,

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & Z' \\ 1/Z & (1 + Z'/Z) \end{bmatrix} \begin{bmatrix} V_o \\ I_o \end{bmatrix} \tag{548}$$

Now try the following problems. As you may note, some of the problems would be as easy, or easier, to work without using matrices. But this is, of course, beside the point, as our object here is to provide practice in thinking in terms of matrices.

Matrix algebra is a shorthand method of manipulating systems of simultaneous equations; its real value becomes evident in the analysis of complex networks represented by such equations. It allows us to study systems of interconnected blocks of elements without having to write out the mass of individual equations associated with the system. Digital computer programs for solving matrix equations are available, and are used to provide actual numerical answers if this is required.

Problem 257
For eq. (545) show that

$$I_1 = \frac{(Z_{22} + Z)V_1 - (Z_{12} + Z)V_2}{dz + (Z_{11} + Z_{22} - Z_{12} - Z_{21})Z}$$

where $dz = Z_{11}Z_{22} - Z_{12}Z_{21}$

Problem 258
Write eq. (545) in terms of the h-parameters of the transistor.

Problem 259
Can eq. (504), in section 11.6, be applied directly to eq. (543)?

Problem 260
Solve eq. (548) for the matrix

$$\begin{bmatrix} V_o \\ I_o \end{bmatrix}$$

by taking the inverse of the coefficient matrix.

(**Note:** the above will be easy if you take advantage of the special formula for finding the inverse of a 2×2 matrix given in the solution to problem 256.)

Next, in the basic Fig. 277, suppose a load impedance of Z_L ohms is connected to the output terminals, as in Fig. 307, and that the PROBLEM is to find the value of the output load current I_L.

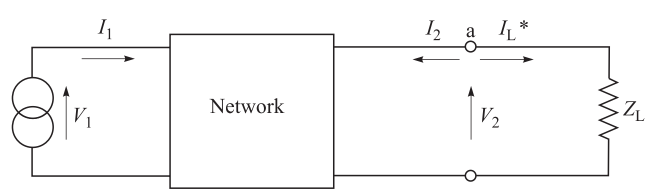


Fig. 307

* At junction point "a" in Fig. 307, by Kirchhoff's current law, $I_2 + I_L = 0$; that is, $I_2 = -I_L$.

Problem 261
For the network of Fig. 307 inside the box, suppose it is found that

$$h_{11} = 400 \text{ ohms} \qquad h_{12} = 0.100$$
$$h_{21} = -20 \qquad h_{22} = 0.002 \text{ mhos}$$

Given that $V_1 = 12$ volts, find the value of load current if $Z_L = R_L = 150$ ohms.
(*Answer*: $I_L = 0.2927$ amps)

Problem 262
Rework problem 261, this time beginning with the matrix equation (514) in section 11.6.

Problem 263
Two identical transistors, operating in common-emitter mode, are connected in cascade as shown in Fig. 308.

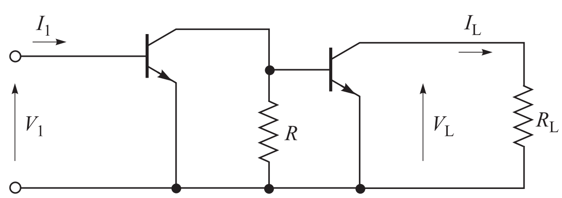


Fig. 308

In the figure, let it be given that the transistor h-parameter values are

$$h_{11} = 1000 \text{ ohms} \qquad h_{12} = 0.004$$
$$h_{21} = 40 \qquad h_{22} = 0.0005 \text{ mhos}$$

If it is given that $R = 500$ ohms and $R_L = 900$ ohms, find the output voltage V_L if the input voltage V_1 is 0.001 volt. (Again, as in problem 262, let us begin with eq. (514) in section 11.6.) (*Answer*: 0.3196 volts)

Problem 264
Write the set of simultaneous eqs. (455), (456), and (457), in Chap. 10, in the form of a single matrix equation.

Problem 265
Note that the answer to problem 264 says that

$$\begin{bmatrix} A_1 \\ A_2 \\ A_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ a & a^2 & 1 \\ a^2 & a & 1 \end{bmatrix}^{-1} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

As an exercise in matrix manipulation, verify that the above expression does produce eqs. (460), (463), and (466) in Chap. 10.

Our final example, which follows, will provide further practice in matrix manipulation and will also bring to light an interesting fact concerning power in unbalanced three-phase systems. In doing this we'll freely make use of our previous work in three-phase theory

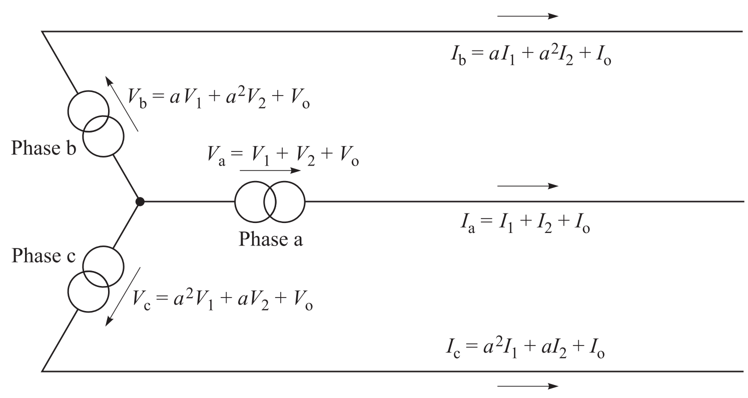


Fig. 309

in sections 10.7 through 10.11. Let us begin with the three-phase generator depicted in Fig. 309.

In the figure, V_a, V_b, and V_c represent the rms values of three unbalanced phase voltages, with I_a, I_b, and I_c representing the rms values of the three corresponding unbalanced phase currents (also the line currents here), as shown. In the work here we wish to concentrate our attention on the POWER produced in the unbalanced condition, ESPECIALLY in regard to expressing the power in terms of the SYMMETRICAL COMPONENTS of the unbalanced system.

To begin, let P_T denote the total "true power" produced by the above unbalanced generator. From inspection of the figure it's clear that P_T is equal to the SUM OF THE POWERS produced by the three individual phases of the generator; thus (see note 29 in Appendix) in terms of the actual phase voltages and currents the value of P_T is equal to the "sum of the real parts" (srp) in the expression

$$P_T = \text{srp}: \; V_a\bar{I}_a + V_b\bar{I}_b + V_c\bar{I}_c \tag{549}$$

in which the "overscore" in $\bar{I}_a$, $\bar{I}_b$, $\bar{I}_c$ denotes the CONJUGATE of the quantity represented by the letter. (It's understood that the Vs and Is are, in general, complex numbers.)

Note that eq. (549) is expressed in terms of the actual phase voltages and currents; but we, however, wish to express the power in terms of the SYMMETRICAL COMPONENTS of the phase voltages and currents. To do this, let us start by writing eq. (549) in matrix notation; thus

$$P_T = \text{srp}: \; [V_a \quad V_b \quad V_c]\begin{bmatrix} \bar{I}_a \\ \bar{I}_b \\ \bar{I}_c \end{bmatrix} \tag{550}$$

Let us now first work on the above *current matrix*, as follows. From inspection of Fig. 309 we have

$$I_a = I_1 + I_2 + I_o$$
$$I_b = aI_1 + a^2I_2 + I_o$$
$$I_c = a^2I_1 + aI_2 + I_o$$

Now take the CONJUGATES of the above equations. Remembering that the conjugate of the *sum* of a number of complex numbers is the *sum of the conjugates* and that the conjugate of the *product* of complex numbers is the *product of the conjugates*, and that the conjugate of a is a^2 and the conjugate of a^2 is a, the above equations become

$$\bar{I}_a = \bar{I}_1 + \bar{I}_2 + \bar{I}_o$$

$$\bar{I}_b = a^2\bar{I}_1 + a\bar{I}_2 + \bar{I}_o$$

$$\bar{I}_c = a\bar{I}_1 + a^2\bar{I}_2 + \bar{I}_o$$

Thus eq. (550) becomes

$$P_T = \text{srp}: [V_a \quad V_b \quad V_c]\begin{bmatrix} 1 & 1 & 1 \\ a^2 & a & 1 \\ a & a^2 & 1 \end{bmatrix}\begin{bmatrix} \bar{I}_1 \\ \bar{I}_2 \\ \bar{I}_o \end{bmatrix} \tag{551}$$

Now, in the above equation, since the "phase voltage matrix" is a "row matrix" it can also be expressed as the TRANSPOSE of the corresponding "column matrix"; doing this, the above *row matrix* can be put into the following form

$$[V_a \quad V_b \quad V_c] = \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix}_t = \begin{bmatrix} (V_1 + V_2 + V_o) \\ (aV_1 + a^2V_2 + V_o) \\ (a^2V_1 + aV_2 + V_o) \end{bmatrix}_t^{*} = \left(\begin{bmatrix} 1 & 1 & 1 \\ a & a^2 & 1 \\ a^2 & a & 1 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \\ V_o \end{bmatrix}\right)_t$$

and thus, upon making use of the "reversal rule" (eq. (492)), we have that

$$[V_a \quad V_b \quad V_c] = [V_1 \quad V_2 \quad V_o]\begin{bmatrix} 1 & a & a^2 \\ 1 & a^2 & a \\ 1 & 1 & 1 \end{bmatrix}$$

Thus eq. (551) becomes

$$P_T = \text{srp}: [V_1 \quad V_2 \quad V_o]\begin{bmatrix} 1 & a & a^2 \\ 1 & a^2 & a \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ a^2 & a & 1 \\ a & a^2 & 1 \end{bmatrix}\begin{bmatrix} \bar{I}_1 \\ \bar{I}_2 \\ \bar{I}_o \end{bmatrix}$$

which, since $a^3 = 1$, $a^4 = a$, and $1 + a + a^2 = 0$, becomes

$$P_T = \text{srp}: [V_1 \quad V_2 \quad V_o]\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} \bar{I}_1 \\ \bar{I}_2 \\ \bar{I}_o \end{bmatrix} = 3[V_1 \quad V_2 \quad V_o]\begin{bmatrix} \bar{I}_1 \\ \bar{I}_2 \\ \bar{I}_o \end{bmatrix}$$

thus, finally,

$$P_T = \text{srp}: 3V_1\bar{I}_1 + 3V_2\bar{I}_2 + 3V_o\bar{I}_o \tag{552}$$

The meaning of eq. (552) is as follows. In section 10.9 we found that the TOTAL POWER produced in a *balanced* three-phase system is *three times the power per phase* (eq. (438)). Equation (552) shows that the total power produced in an *unbalanced* three-phase system is equal to the simple sum of the powers separately produced by the positive-sequence, negative-sequence, and zero-sequence systems; that is, as far as power is concerned, each system acts independently of the other two. Note that this is an unexpected result, because the "principle of superposition" does *not* generally apply to power calculations. (See "note of caution" following problem 73 in section 5.7.)

* See Fig. 309.

CHAPTER 12

Binary Arithmetic. Switching Algebra

12.1 Analog and Digital Signals. Binary Arithmetic

An ANALOG type of signal has in general a *continuous range of amplitude values*, such as is illustrated in Fig. 310.

A DIGITAL signal, on the other hand, is an ordered sequence of *discontinuous* pulse-type signals that can have only a *limited number of different levels of amplitude*. If only *two* different levels are allowed, or can be detected, the digital signal is said to be a BINARY (''BY nary'') type signal, the word ''binary'' meaning ''two-valued.'' The two different levels of a binary signal can be said to represent the ''on'' and ''off'' conditions of the signal, or the ''presence or absence'' of a pulse, and can be denoted by ''1'' and ''0,'' as in Fig. 311.

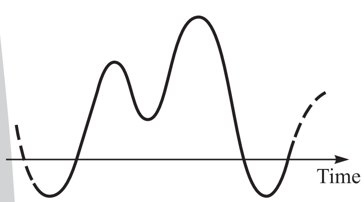


Fig. 310. Analog signal.

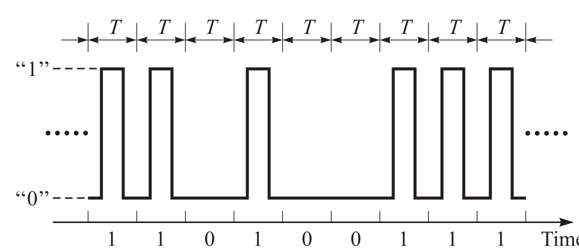


Fig. 311. ''Binary'' digital signal.

In Fig. 311 ''T'' is the measured time allotted to ''one unit of information,'' which is called a ''bit''; thus, ''9 bits of information'' are represented in Fig. 311. The signal is said to be ''binary digital'' because its two different states can be represented by the digits 1 and 0, as shown.

It is true that, in the real world, most signals originate in analog form; for instance, the outputs of microphones, TV cameras, and most other sensing devices, are in analog form. You might therefore very well ask "Of what *use* are digital-type signals; why use a digital system at all?".

In answer to the question, one very important reason is that *the internal operations of* DIGITAL COMPUTERS are handled in the form of binary digital signals. This is because a digital computer uses integrated circuits containing thousands of transistors operating in the binary "on or off" mode.

Another reason is that it is sometimes beneficial to *first transform an analog signal into a coded binary-type signal* before it is fed into a transmission system or channel. This can have a very good effect if the channel is noisy, because at the receiver it is then *only necessary to detect whether a pulse is* PRESENT *or* NOT PRESENT. If this can be done, the original analog signal can then be completely recovered from the binary coded signal, even if the binary signal is mixed with so much noise that it would not be possible to recover the signal if it were in analog form.*

In binary work, especially in regard to digital computers, it is necessary to be fluent in "binary arithmetic," which let us introduce as follows.

A "digit" is a single symbol representing a whole, or integral, quantity. A "number" is a quantity represented by a group of digits. The *number of different digits* a number system uses is called the "base" or the *radix*, R, of the system. Thus the familiar decimal system has the radix "ten," using the ten digits 0 through 9.

In all practical number systems the *value of a digit* in a number depends not only on the digit itself but also upon the *position* of the digit in the number. Consider, as an example, the quantity represented by the decimal system number

2684.735

As you know, the number to the left of the "decimal point" is the whole or integral part of the quantity, while the number to the right of the decimal point is the fractional part of one unit. Note that the digits in the above decimal number have the following values:

$$
\begin{aligned}
\text{the digit ``2'' has the value } 2 \times 10^{3} &= 2000. \\
\text{the digit ``6'' has the value } 6 \times 10^{2} &= 600. \\
\text{the digit ``8'' has the value } 8 \times 10^{1} &= 80. \\
\text{the digit ``4'' has the value } 4 \times 10^{0} &= 4. \\
\text{the digit ``7'' has the value } 7 \times 10^{-1} &= .7 \\
\text{the digit ``3'' has the value } 3 \times 10^{-2} &= .03 \\
\text{the digit ``5'' has the value } 5 \times 10^{-3} &= \underline{.005} \\
\text{Total value of the number} &= 2684.735
\end{aligned}
$$

As the above illustrates, in a decimal system number the value of any digit is equal to *the digit times* 10^n, where the value of n depends upon the *position* of the digit relative to the decimal point.

In the above example, the radix has the value "10," $R = 10$. The same basic principle, however, applies to ANY positional number system of radix R; thus, if any such number

---

* "Noise" consists of both the man-made type and the more fundamental type due to the random motion of electric charges always present in all material substances. Which type predominates, in any given case, depends upon the circumstances in that particular case.

system has R different digits, the value of any number N in that system is equal to a sum of the form

$$\cdots + dR^3 + dR^2 + dR + d + dR^{-1} + dR^{-2} + dR^{-3} + \cdots \tag{553}$$

where "d," in any term, can be any of the R different digits in the system. For instance, the decimal quantity 633.8, expressed in the form of eq. (553), is

$$\begin{aligned} 6 \times 10^2 + 3 \times 10 + 3 + 8 \times 10^{-1} &= 600 + 30 + 3 + 0.8 \\ &= 633.8 \end{aligned}$$

In everyday work, in business, engineering, and so on, amounts are expressed numerically in decimal system numbers, for which $R = 10$, as we know. There's no doubt that the use of $R = 10$ arose from the practice of counting with the aid of the ten fingers of the two hands. But aside from this, there is nothing magical about the use of the decimal system. The use of $R = 10$ does, however, happen to be a good practical compromise between "too large" a value of R and "too small" a value of R *as far as direct use of numbers by human beings is concerned*. However, *for doing arithmetic and storing data through the medium of transistors or magnetic tape* in a digital computer, the use of $R = 10$ is *not at all suitable*; the reason is that the transistors or the tape would have to be able to reliably sense *10 different levels* of voltage or magnetization, a requirement that would call for complicated circuitry. Instead of $R = 10$, it's much more practical, in digital computers, to use a two-state BINARY *system*, because it's easy to reliably drive a transistor into either the "on" or the "off" state, or to "magnetize" or "not magnetize" a spot on magnetic tape. For this reason, the internal mathematical operations in a digital computer are performed using BINARY *arithmetic*, the laws of which let us now consider.

The two-digit binary number system uses JUST THE FIRST TWO DIGITS of the decimal system, "1" and "0." Some examples of the appearance of binary numbers are

$$1101 \qquad\qquad 111001010 \qquad\qquad 1101.101$$

The first two numbers, above, are examples of *whole* binary numbers, while the third consists of the whole part 1101, plus the fractional part 101; the dot separating the whole part and the fractional part is called the "binary point," which corresponds to the "decimal point" in the decimal system.

Now let's consider the *value* represented by each of the above three binary numbers, and let us express these values in terms of equivalent *decimal* numbers. To do this, first note that in the binary system the number *two* is *not represented by a single digit*, just as in the decimal system the value *ten* is not represented by a single digit.

To get the equivalent decimal system values of the above three examples, we will have to use decimal numbers in eq. (553) which, for $R = 2$, becomes

$$N = \cdots d(2)^3 + d(2)^2 + d(2) + d + d(2)^{-1} + d(2)^{-2} + \cdots \tag{554}$$

where the digit d can have only the value 1 or 0. The decimal system equivalent values of the three examples are, therefore,

$$1101 = 1(2)^3 + 1(2)^2 + 0(2) + 1 = 8 + 4 + 0 + 1 = 13, \ \textit{answer}$$

$$\begin{aligned} 111001010 &= 1(2)^8 + 1(2)^7 + 1(2)^6 + 0(2)^5 + 0(2)^4 + 1(2)^3 + 0(2)^2 + 1(2) + 0 \\ &= 256 + 128 + 64 + 0 + 0 + 8 + 4 + 2 + 0 = 462, \ \textit{answer} \end{aligned}$$

$$\begin{aligned} 1101.101 &= 1(2)^3 + 1(2)^2 + 0(2) + 1 + 1(2)^{-1} + 0(2)^{-2} + 1(2)^{-3} \\ &= 8 + 4 + 0 + 1 + 1/2 + 0 + 1/8 = 13.625, \ \textit{answer} \end{aligned}$$

Setting $d = 1$ in eq. (554), it will be seen that the equivalent decimal system values of the *binary 1 digits*, relative to the binary point, are

$$\begin{array}{lccccccccccc} & & & & & & & \downarrow & & & & \\ \text{binary digit:} & \ldots & 1 & 1 & 1 & 1 & 1 & 1 & . & 1 & 1 & 1 & \ldots \\ \text{decimal value:} & & 32 & 16 & 8 & 4 & 2 & 1 & & .5 & .25 & .125 & \end{array}$$

You can now verify, and commit to memory if you wish, that, including zero, the first 16 binary numbers, with their decimal equivalents, are

0000 = 0	0101 = 5	1010 = 10
0001 = 1	0110 = 6	1011 = 11
0010 = 2	0111 = 7	1100 = 12
0011 = 3	1000 = 8	1101 = 13
0100 = 4	1001 = 9	1110 = 14
		1111 = 15

Another important point we should be aware of is that, in any positional number system of radix R, it's very easy to multiply or divide by R raised to any positive or negative integral power. For instance, to multiply or divide a *decimal number* by TEN *raised to any integral power n*, all we need do is *move the decimal point n places to the right or left*, as the case may be. In the same way, to multiply or divide a *binary number* by TWO *raised to any integral power n*, all we need do is *move the binary point n places to the right or left* as the case may be. The following table will make this clear.

Number	Meaning in decimal system	Meaning in binary system
1	one	one
10	ten	two
100	hundred	four
1000	thousand	eight
10000	ten thousand	sixteen
100000	hundred thousand	thirty-two
⋮	⋮	⋮

Thus, while it's easy to multiply or divide by any integral power of *ten* in the decimal system, it's equally easy to multiply or divide by any integral power of *two* in the binary system. For instance, consider a binary number, $N = 101101$; suppose, for example, that we wish to multiply or divide N by say *eight*, which is 1000 in the binary system. Either operation is very simply done, as follows,

$$101101 \times 1000 = 101101000 \quad \text{and} \quad \frac{101101}{1000} = 101.101$$

For the same values, the corresponding operations in the *decimal system* would be "45 times 8" and "45 divided by 8" which, of course, cannot be done by simply moving the "decimal point" to the right or the left.

As you can see from the above, expressing a given value in binary notation requires the writing of *many more digits* than in decimal notation. Thus, while the binary system is admirably suited for internal use in a digital computer, it is *not* well suited for direct use by human beings because it requires the writing down of so many digits to express even small values. This bears out the statement made previously in regard to $R = 2$ being the ideal choice for internal computer operation, and $R = 10$ being a good compromise for use by human beings.

Now let's consider the procedures for converting a decimal number into binary form and vice versa. In stating the rules we'll make use of the terms "most significant digit" (MSD) and "least significant digit" (LSD). The MSD is simply the digit having the most value in a number, while the LSD is the digit having the least value in a number. Thus, in the decimal number 28736, "2" is the MSD, "8" is the next-most significant digit, and so on, until we reach "6," which is the LSD. Or, in the binary number 110101, for example, the 1 at the left-hand end is the MSD, while, going from left to right, we finally arrive at the 1 at the right-hand end, which is the LSD. With this terminology in mind, the rules for conversion from decimal to binary, and from binary to decimal, can be summarized as follows.

DECIMAL-TO-BINARY CONVERSION

First, the following steps can be taken to convert a WHOLE *decimal number* into binary form.

1. Divide the decimal number by 2; this produces a quotient plus a *remainder* of either 1 or 0. The remainder, 1 or 0, is the *LSD* in the equivalent binary number.
2. Divide the quotient found in step (1) by 2; this produces a second quotient plus a remainder of 1 or 0. This remainder, 1 or 0, is the *second least significant* in the binary number.
3. Continue on in this fashion, dividing each quotient by 2, until the quotient is equal to *zero* plus the final remainder of 1 or 0, which is the *MSD* in the binary number.

Example 1

Convert the decimal number 105 to binary form.

Solution

$$105 \text{ divided by } 2 = 52, \text{ plus remainder } 1, \text{ the LSD}$$
$$52 \text{ divided by } 2 = 26, \text{ plus remainder } 0, \text{ second LSD}$$
$$26 \text{ divided by } 2 = 13, \text{ plus remainder } 0, \text{ third LSD}$$
$$13 \text{ divided by } 2 = 6, \text{ plus remainder } 1, \text{ fourth LSD}$$
$$6 \text{ divided by } 2 = 3, \text{ plus remainder } 0, \text{ fifth LSD}$$
$$3 \text{ divided by } 2 = 1, \text{ plus remainder } 1, \text{ sixth LSD}$$
$$1 \text{ divided by } 2 = 0, \text{ plus remainder } 1, \text{ the MSD}$$

Since positional numbers are always written from left to right, with the MSD at the left-hand end, we have that

$$105 \text{ decimal} = 1101001 \text{ binary}, \quad \textit{answer}$$

Next let us consider the conversion of a *decimal fraction* into its equivalent *binary fraction*. Actually the procedure is very simple, but takes a lot of words to describe. Let us therefore first try to describe the procedure in words, and then, by means of an example, show that the procedure is really very easy.

1. Begin by multiplying the given decimal fraction by 2; if the product is *greater than 1* the binary fraction begins as 0.1, but if the product is *less than 1* the binary fraction begins as 0.0.
2. If the product found in step (1) is *greater than 1*, then subtract 1 from the product and then multiply the result by 2; if the result is greater than 1, the binary fraction is now of the form 0.11, but if less than 1 it is of the form 0.10. If, however, the product found in step (1) is *less than 1*, then multiply it by 2; if the result is greater than 1 the binary fraction is now of the form 0.01, but if less than 1 it is of the form 0.00. We continue on in this fashion to any degree of accuracy required.

Example 2

Convert the decimal fraction 0.403 into binary form.

Solution

Here, ‘‘multiply’’ and ‘‘subtract’’ are abbreviated as ‘‘mult’’ and ‘‘sub.’’

The given decimal fraction:	0.403	binary:	
mult by 2:	0.806	0.0	because $0.806 < 1$
mult by 2:	1.612	0.01	because $1.612 > 1$
sub 1, then mult by 2:	1.224	0.011	because $1.224 > 1$
sub 1, then mult by 2:	0.448	0.0110	because $0.448 < 1$
mult by 2:	0.896	0.01100	because $0.896 < 1$
mult by 2:	1.792	0.011001	because $1.792 > 1$
sub 1, then mult by 2:	1.584	0.0110011	because $1.584 > 1$
sub 1, then mult by 2:	1.168	0.01100111	because $1.168 > 1$
sub 1, then mult by 2:	0.336	0.011001110	because $0.336 < 1$

and so on, to whatever accuracy is required. To test the accuracy of the last result, above, let us make use of eq. (554). Since we’re dealing entirely with a binary fraction, we need only use the terms with negative exponents in eq. (554); doing this, we find that

$$0.011001110 = 2^{-2} + 2^{-3} + 2^{-6} + 2^{-7} + 2^{-8} = 0.40234\ldots$$

which may or may not be close enough to 0.403, depending upon accuracy requirements. If you wish to continue with the above example, you can verify, for instance, that

$$0.0110011100101 = 0.40295, \text{ which is of course closer to 0.403 than before.}$$

Although such conversions are very time-consuming when done using pencil and paper, they present no such difficulty when done internally in a digital computer; this is because the computer can execute millions of such routine steps per second.

Also, in regard to decimal-to-binary conversion in general, it should be pointed out that an *integral* decimal number always has an exact equivalent in the binary system, but a decimal *fraction* may or may not have exact representation in the binary system. (However, the binary equivalent of a decimal fraction can always be determined to any desired degree of accuracy.)

BINARY-TO-DECIMAL CONVERSION

To convert an *integral* (whole) binary number to decimal form we move *left to right* through the binary number, from the MSD to the LSD, as follows. Multiply the MSD by 2, then add on the next digit. Multiply the result by 2, then add on the next digit. Multiply this result by 2, then add on the next digit. Continue on until the last digit to the right (the LSD) is included in the conversion.

Example 3
Convert the binary number 1101010 to decimal form.

Solution
Beginning at the left-hand end of the given binary number and using the above procedure, we can chart the results as follows:

binary:	1	1	0	1	0	1	0
decimal:	2	3	6	13	26	53	106

Thus, 1101010 bi = 106 dec, *answer*

Next, to convert a BINARY FRACTION *into a* DECIMAL FRACTION we move through the binary fraction *from right to left* toward the "binary point" in the following manner.

Divide the right-hand digit by 2, then add to this the next digit and divide the result by 2. Now add, to the last result, the next digit and divide the result by 2. Continue on in this way until the binary point is reached.

Example 4
Convert the binary fraction 0.011011 to decimal form.

Solution
Beginning at the right-hand end, and following the above procedure, we can chart the results as follows.

binary:	.0	1	1	0	1	1
						0.5
					0.75	
				0.375		
			0.6875			
		0.84375				

0.421875, thus 0.011011 bi = 0.421875 dec, *answer*

$$CHECK\text{: } 0.011011 = 2^{-2} + 2^{-3} + 2^{-5} + 2^{-6} = \frac{1}{4} + \frac{1}{8} + \frac{1}{32} + \frac{1}{64} = 0.421875$$

Now let us continue on, and study the basic arithmetic ("air ith MET ik") operations of addition, subtraction, multiplication, and division in both the decimal and binary systems. Let us begin with *addition*, first in the familiar decimal system, as follows.

As we know, the decimal system employs the radix "ten," making use of the ten digits $0, 1, 2, \ldots, 9$. In the decimal system the number "ten" is denoted by "10." Now, in regard to the *sum* of any two of the ten digits, two possibilities arise; thus

1. the sum of the two digits is *less than the radix ten*, or
2. the sum of the two digits is *equal to or greater than* the radix ten.

In case (1) no difficulty arises; thus, $2 + 5 = 7$, for example. In case (2), however, a problem does arise in taking a sum such as "$8 + 8$," for example, because *no single digit* exists in the decimal system to denote the quantity "sixteen." Instead, to indicate in number form the quantity "sixteen" we write "16," in which the "1" now has the value "ten." This is the "carry" operation that we are all familiar with in the decimal system; if (in the decimal system) the sum of two digits is equal to or greater than the radix ten, we write down the required digit in the units column and "carry the 1" to the left into the "tens column" where the 1 now has the value ten. If the numbers consist of more than one digit we use the same procedure of "carrying a 1" into the next higher valued column; for example

$$\begin{array}{r} 6\;6\;5\;8 \\ \underline{+9\;6\;2\;6} \\ 1\;6\;2\;8\;4 \end{array}$$

Now consider the addition of two digits in the *binary* system. Here there are only *two* possible two-digit sums, $1 + 0 = 1$, and $1 + 1 =$ two. In the binary system, however, there is no single digit to represent the value "two"; hence we "carry a 1" into the "twos column," which is the next column to the left, and write the result of "$1 + 1$" in the form

$$\begin{array}{r} 1 \\ \underline{+1} \\ 10 \end{array}$$

where the "1" in the "10" now represents the value "two." Now consider, as an example, the addition of the two binary numbers "1011" and "1101," as indicated below to the left

"eights" column
"fours" column
"twos" column
"units" column

$$\begin{array}{rcccc} & 1 & 0 & 1 & 1 \\ + & \underline{1} & \underline{1} & \underline{0} & \underline{1} \end{array} \qquad\qquad \begin{array}{rrrrrl} & 1 & 0 & 1 & 1 & \text{"11" dec} \\ \underline{+} & \underline{1} & \underline{1} & \underline{0} & \underline{1} & \text{"13" dec} \\ 1 & 1 & 0 & 0 & 0 & \text{"24" dec} \end{array}$$

The procedure for performing the addition indicated to the left above is:

1. the sum of the digits in the *units column* is $1 + 1 = 10$; hence we write down 0 and carry 1 into the *twos column*,
2. the sum of the digits now in the *twos column* is $1 + 1 = 10$; hence write down 0 and carry a 1 into the *fours column*,
3. the sum of the digits now in the *fours column* is $1 + 1 = 10$; hence we write 0 and carry a 1 into the *eights column*,

4. the sum of the digits now in the *eights column* is $1 + 1 + 1 = (1 + 1) + 1 = 10 + 1 = 11$, hence we write down 1 and carry a 1 into the *sixteens column*, giving the answer 11000, as shown to the right above with decimal equivalents alongside.

In the above we found the sum of just *two* binary numbers, but suppose the sum of, say, *one thousand* such numbers must be found. In the internal circuitry of a digital computer, this is most conveniently done by adding the binary numbers together two at a time until the final sum is reached. As mentioned before, we must remember that a digital computer can execute millions of such routine operations per second.

Next, let us consider the *subtraction* of one number from another number. Subtraction makes use of the *borrow* operation, illustrated first in the following decimal-system example.

	PROBLEM						SOLUTION				
	9	6	3	7	minuend		8	15	13	7	
−	4	7	7	6	subtrahend	−	4	7	7	6	
					difference		4	8	6	1	*answer*

Discussion. Beginning at the right-hand side in the PROBLEM we first have $7 - 6 = 1$, which presents no difficulty because 7 is larger than 6. Continuing on, from right to left, we next must subtract "7 from 3" (actually, 70 from 30), which does present a difficulty because 3 is smaller than 7. To get around this difficulty we now, in the minuend, "borrow a 1" from the 6, and transfer the borrowed 1 over to the 3; however, since a 1 in the third column has *ten times* the value of a 1 in the second column, this effectively makes the 3 become 13, as shown in the SOLUTION. Since 13 is larger than 7 we now have $13 - 7 = 6$, as shown.

At this point we must not forget that the 6, in the third column, is now changed to 5 (because of the previous borrowing of the 1 from the 6). Therefore, continuing on in the PROBLEM, we must now subtract "7 from 5," which again presents a difficulty because 5 is smaller than 7. To get around this difficulty we now, in the minuend, "borrow a 1" from the 9 and transfer the borrowed 1 to the 5, effectively making the 5 become 15, as shown in the SOLUTION; thus we have $15 - 7 = 8$, as shown. Because of the borrowing of the 1, the 9 becomes 8, as shown in the SOLUTION; hence the last step is to subtract 4 from 8, giving the final answer 4861.

We must remember that the *same basic arithmetic procedures* apply to *all* positional number systems, whatever the particular radix. Thus the "borrow" procedure, illustrated above for the decimal system, is used in the same way to subtract one binary number from another binary number, as the following example illustrates.

	PROBLEM								SOLUTION						
	1	0	1	1	0	1	minuend		0	10	1	0	10	1	
−	0	1	1	0	1	1	subtrahend	−	0	1	1	0	1	1	
							difference		0	1	0	0	1	0	*answer*

Discussion. Beginning at the right-hand side in the PROBLEM we have $1 - 1 = 0$, as shown in the SOLUTION. Continuing on, from right to left in the PROBLEM, the next step is to subtract "1 from 0," which presents a difficulty because 0 is smaller than 1. To surmount this difficulty we now, in the minuend, "borrow a 1" from the third column and transfer it over to the second column in the minuend. In the binary system, however, we must remember that the value of a 1 *doubles* each time we move one position to the left; hence the 1 transferred from the third column effectively becomes *two* in the second column, as shown by the "10" in the SOLUTION. We therefore have $10 - 1 = 1$, as shown in the SOLUTION.

Next, referring to the SOLUTION, note that the transferral of the 1 has left "0" in the third column of the minuend; therefore the next two steps in the subtraction present no difficulty, because $0 - 0 = 0$ and $1 - 1 = 0$, as shown in the SOLUTION. However, in the fifth column of the PROBLEM we again run into the difficulty of subtracting "1 from 0," a difficulty we overcome by borrowing a 1 from the minuend in the sixth column; the borrowed 1 effectively becomes 10 (two) in the minuend in the fifth column, as shown in the SOLUTION. Thus in the fifth column we have $10 - 1 = 1$, giving the *final answer* 10010, as shown (in equivalent decimal notation, $45 - 27 = 18$).

In our work so far we've found that it's generally easier, and less time-consuming, to do *addition* than it is to do subtraction. This is true in both pencil and paper work and in terms of digital computer circuitry requirements. It would therefore be an advantage if subtraction could somehow be performed in a way that used only the *addition* operation. Fortunately it *is* possible to do this by making use of what is called the ONE's COMPLEMENT of a binary number; the basic procedure can most easily be developed by considering the subtraction of one whole binary number from another whole binary number, as follows.

Let N be any whole binary number. Since $R = 2$ for the binary system, and since N is to be a whole number, eq. (553) will contain no negative exponents and will thus be of the form

$$N = d2^n + \cdots + d2^3 + d2^2 + d2^1 + d2^0 \tag{555}$$

in which the digit d is, in any individual term, equal to either 1 or 0. Note that, for convenience, we're expressing the value of the binary number N in terms of the decimal digits 2, 3, and so on. In eq. (555) let n be called the *order* of the binary number N; therefore, in accordance with the terminology of eq. (555), it should be noted that a binary number of *order* n contains $n + 1$ *digits*.

With the foregoing in mind, let us now define that, if N is a binary number of order n, then

$$\text{the } \textit{1's complement} \text{ of } N = (2^{n+1} - N) - 1 \qquad (556)^*$$

The reason the above definition is useful will become clear later on, but first let's investigate the *1's complement* of a binary number as defined above.

To do this, *let N be a binary number of order n*, and note that eq. (556) can also be written as

$$N + (\text{1's comp of } N) = (2^{n+1} - 1) \tag{557}$$

Now remembering that, in binary notation,

$$\begin{array}{ll} 2^0 = 1 & 2^3 = 1000 \\ 2^1 = 10 & 2^4 = 10000 \\ 2^2 = 100 & 2^5 = 100000 \end{array}$$

* It is also possible to work in terms of the "1's complement plus 1," which is called the "2's complement."

and so on, we see that the binary number for 2^n is *1 followed by n zeros*, and hence the binary number for 2^{n+1} is *1 followed by* $n+1$ *zeros*; thus

$$2^{n+1} = 1000\cdots0000, \quad (\text{“}n+1\text{” zeros})$$

and therefore we have that

$$\begin{array}{lrl} & 2^{n+1} \rightarrow & 1000\ \cdots\ 0000 \\ & \text{subtract } 1 \rightarrow & \underline{\qquad\qquad -1} \\ \text{hence,} & (2^{n+1}-1) \rightarrow & 111\ \ \cdots\ \ 1111 \end{array}$$

Thus the quantity $(2^{n+1}-1)$ *is always a string of* "$n+1$" *1's*, and hence eq. (557) can be written in the form

$$N + (\text{1's comp of } N) = 11111\ldots1111$$

which can only be true if N and its 1's complement are *exact opposites* of each other* in regard to the positions of the 1's and 0's. Suppose, for example, that we wish to find the 1's complement of the binary number $N = 101101$, which is of order $n = 5$. By eq. (557), upon transposing N, we have

$$\begin{array}{lrl} & (2^{n+1}-1) = & 1\ 1\ 1\ 1\ 1\ 1 \\ & -N = & -\underline{1\ 0\ 1\ 1\ 0\ 1} \\ \text{thus,} & \text{1's complement of } N = & 0\ 1\ 0\ 0\ 1\ 0, \ \textit{answer} \end{array}$$

Thus it's very easy to find the 1's complement of a binary number N; all we need do is *change the 1's to 0's and the 0's to 1's*. As another example,

$$\begin{array}{rl} \text{if } N = & 1\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 0 \\ \text{the 1's complement of } N = & 0\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1 \end{array}$$

We previously stated that the 1's complement is valuable because it allows *subtraction* to be done through a process of *addition*. To see how this is possible, first solve eq. (557) for N; thus

$$N = 2^{n+1} - 1 - (\text{1's comp of } N)$$

Now let Y be a binary number; then, using the above value of N, the difference, $Y - N$, can be written in the form

$$Y - N = [Y + (\text{1's comp of } N)] - 2^{n+1} + 1 \tag{558}$$

Equation (558) is the basic equation we wish to use in performing the subtraction operation $Y - N$. In regard to eq. (558), it should be pointed out that, as we just learned, it's a simple matter to find the 1's complement of a binary number; it's easy to do this using pencil and paper, and it's also easily done electronically in the internal registers of a digital computer. Also, in regard to the terms "$-2^{n+1}+1$" in eq. (558), we'll find that these two terms can be basically handled together in one simple operation called the "end-around carry." Consider now the following examples.

* For example

$$\begin{array}{r} 1001101011 \\ +\,\underline{0110010100} \\ 1111111111 \end{array}$$

Example 5

Use eq. (558) to perform the binary subtraction

$$\begin{aligned} Y &= 1\ 1\ 0\ 0\ 1 \quad \text{(dec 25)} \\ -N &= -\underline{1\ 0\ 0\ 1\ 1} \quad \text{(dec 19)} \end{aligned}$$

Solution

$$\begin{array}{rl} Y = & 1\ 1\ 0\ 0\ 1 \\ (1\text{'s comp of } N) = & +0\ 1\ 1\ 0\ 0 \\ \hline Y + (1\text{'s comp of } N) = 1 & 0\ 0\ 1\ 0\ 1 = \text{quantity in brackets in eq. (558)} \end{array}$$

In this problem Y and N are of order $n = 4$ (see discussion following eq. (555)). Therefore the value of the "overflow 1" (the "1" circled above) is 2^5, and thus eq. (558) becomes, for this problem

$$Y - N = [2^5 + (0\ \ 0\ \ 1\ \ 0\ \ 1)] - 2^5 + 1$$

Thus the *"overflow 1" cancels out*, and all we need to do is *add 1* to get the value of $Y - N$. This operation is referred to as the "end-around carry" and for this problem can be indicated as follows:

$$\begin{array}{rr} & 1\ 1\ 0\ 0\ 1 \\ + & \underline{0\ 1\ 1\ 0\ 0} \\ 1 & 0\ 0\ 1\ 0\ 1 \\ \longrightarrow & +1 \\ \hline & 0\ 0\ 1\ 1\ 0 = \text{six, } \textit{final answer} \end{array}$$

Thus the "1's complement of N" has allowed us to find the value of $Y - N$ by use of the ADDITION operation only.

Example 6

In example 5, Y and N are both of the same order, $n = 4$, but this is not at all a necessary requirement. For example, suppose the numbers in a certain digital computer are all handled in eight-digit binary form, and suppose the value of $Y - N$ is to be found where, let us say,

$$\begin{aligned} Y &= 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0 \quad (24) \\ N &= 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1 \quad (\ 3) \end{aligned}$$

As we can see, the answer is $Y - N =$ twenty-one; the procedure, using binary numbers and the 1's complement of N, is as follows:

$$\begin{array}{rr} Y = & 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0 \\ 1\text{'s comp of } N = + & 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0 \\ \hline 1 & 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0 \\ \longrightarrow & +1 \\ \hline & 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1, \ \textit{answer} \text{ (dec. 21)} \end{array}$$

Again, we've found the difference, $Y - N$, by use of *addition* only.

Example 7

If Y and N are not integers (whole numbers), exactly the same procedure is used, but we must remember to "line up the binary points" just as we line up the decimal points in decimal addition. For example, suppose we are to find $Y - N$ where, let us say,

$$Y = 1\ 1\ 0\ 1\ .\ 0\ 1\ 1\ 0 \quad \text{(decimal 13.3750)}$$

$$N = 1\ 0\ 0\ 1\ .\ 1\ 0\ 1\ 1 \quad \text{(decimal 9.6875)}$$

Note that the answer in decimal notation is $Y - N = 13.3750 - 9.6875 = 3.6875$. To work out the problem let us use the 1's complement of N in the usual way; thus

$$\begin{array}{rr} Y = & 1\ 1\ 0\ 1\ .\ 0\ 1\ 1\ 0 \\ \text{1's comp of } N = & +0\ 1\ 1\ 0\ .\ 0\ 1\ 0\ 0 \\ \hline 1 & 0\ 0\ 1\ 1\ .\ 1\ 0\ 1\ 0 \\ & \longrightarrow +1 \\ \hline & 0\ 0\ 1\ 1\ .\ 1\ 0\ 1\ 1, \end{array} \quad \textit{answer} \text{ (dec. 3.6875)}$$

Note: The above can also be worked in terms of whole numbers by first shifting the binary point four places to the right and then multiplying by 0.0001; thus

$$\begin{array}{rl} Y = & 11010110 \times 0.0001 \\ \text{1's comp of } N = & 01100100 \times 0.0001 \\ \hline & 00111010 \times 0.0001 \\ & +1 \\ \hline & 00111011 \times 0.0001 = 0011.1011, \text{ as before} \end{array}$$

As you may have noticed, in the foregoing examples the magnitude of Y is greater than the magnitude of N, and hence the values of $Y - N$ are all *positive* numbers. If, however, the magnitude of Y is *less* than the magnitude of N, then $Y - N$ is a *negative* number.

A digital computer must, of course, be able to detect, store, and use both positive and negative numbers. One way a computer can sense whether a difference $Y - N$ is positive or negative is to detect the presence or absence of the "overflow 1" when computing $Y - N$ by use of the 1's complement of N. This is based upon the fact that if Y is greater than N *an overflow 1 will be generated*, but if Y is less than N *no overflow 1 will be generated*. Thus, if $Y - N$ is found by use of the 1's complement of N, then $Y - N$ is a POSITIVE number if an overflow 1 *is* produced, but is a NEGATIVE number if *no* overflow 1 is produced.

To illustrate this, let's return to example 1 above, and this time let $Y = 10011$ and $N = 11001$, so that the problem now becomes

$$Y = 1\ 0\ 0\ 1\ 1 \qquad (Y = 19 \text{ dec})$$

$$-N = -\underline{1\ 1\ 0\ 0\ 1} \qquad (N = 25 \text{ dec})$$

Now using the 1's complement of N in the usual way we find that

$$\begin{array}{rr} Y = & 1\ 0\ 0\ 1\ 1 \\ \text{1's comp of } N = & +\ 0\ 0\ 1\ 1\ 0 \\ \hline Y + \text{1's comp of } N = & 0\ 1\ 1\ 0\ 0\ 1 \end{array}$$

Note that *no overflow 1 is produced*, which tells the computer that *Y is less than N*, and therefore that

(a) the answer to $Y - N$ is a *negative* number whose magnitude is therefore found by

(b) changing the *minuend* number to the 1's complement form and adding; thus

$$
\begin{array}{rl}
\text{1's comp of } Y = & 0\ 1\ 1\ 0\ 0 \\
+N = + & \underline{1\ 1\ 0\ 0\ 1} \\
1 & 0\ 0\ 1\ 0\ 1 \\
\longrightarrow & +1 \\
 & \overline{0\ 0\ 1\ 1\ 0} = \textit{magnitude} \text{ of } Y - N
\end{array}
$$

The computer must now have some way of indicating that the answer, 00110, is a *negative* value, "minus six," and one way of doing this is to use a "sign digit" in the following manner.

In the present example, the numbers in the computer are represented in five-digit binary form, such as 11101, 00010, and so on. In this case an additional binary 1 or 0 could be used in the sixth place as a SIGN DIGIT to indicate "plus" or "minus"; thus, if "0" indicates "plus" and "1" indicates "minus" or negative, the answer to the foregoing problem would be registered in the computer as 100110, indicating "minus six." In the same way, 000110 would indicate "positive six," and so on. The "sign digit," 1 or 0, is generated when the computer processes the difference $Y - N$. Thus (assuming "0" and "1" to denote "plus" and "minus" respectively) if the computer senses, for example, that an overflow 1 is *not* produced, this fact signals the computer to change Y to its 1's complement form, add N to it, and put a "1" in front of the magnitude value of $Y - N$. On the other hand, if an overflow 1 *is* produced, this causes the computer to put the digit "0" in front of the result, to show that $Y - N$ is a positive quantity.

Problem 266
Convert 67 decimal to binary form.

Problem 267
Convert 383 decimal to binary form.

Problem 268
Convert 118.182 decimal to binary form (to 9 binary places).

Problem 269
Convert 1110101 binary to decimal form.

Problem 270
Convert 1001.01101 binary to decimal form.

Problem 271
Using binary addition, write the sums of the following binary numbers, with answers in binary form.

$$
\text{(a)}\ \begin{array}{r} 0\ 1\ 0\ 1\ 1\ 0\ 1 \\ \underline{1\ 0\ 1\ 1\ 0\ 0\ 1} \end{array}
\qquad
\text{(b)}\ \begin{array}{r} 1\ 0\ 1\ 1 \\ 0\ 1\ 0\ 1 \\ \underline{1\ 1\ 0\ 1} \end{array}
\qquad
\text{(c)}\ \begin{array}{r} 1\ 1\ .\ 0\ 1\ 1 \\ \underline{0\ 1\ .\ 1\ 0\ 1} \end{array}
$$

Problem 272
Use "1's complement" to perform the following binary subtractions.

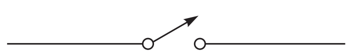
$$\begin{array}{lrl@{\qquad}lrl}
\text{(a)} & 1\ 1\ 0\ 1\ 0\ 1\ 1 & (107 \text{ dec}) & \text{(b)} & 1\ 0\ 0\ 1\ 1\ 1\ 1 & (79 \text{ dec}) \\
 & \underline{-1\ 0\ 0\ 1\ 1\ 1\ 1} & -\ (79 \text{ dec}) & & \underline{-1\ 1\ 0\ 1\ 0\ 1\ 1} & -\ (107 \text{ dec})
\end{array}$$

12.2 Boolean or "Switching" Algebra. Truth Tables

In this section we take up a specialized algebra first introduced in 1847 by the British scientist and philosopher George Boole, and now called "Boolean algebra" in his honor. The algebra was invented by Boole as a method of expressing and reducing statements in logic. The algebra has also, however, been found to be especially useful in the study of electrical switching network problems, and for that reason is often referred to as "switching algebra." The algebra, as applied to the switching operation, can be introduced as follows.

An electrical switch is a "binary" or "two-state" device, because it must always, at any given time, be in one or the other of only *two possible states*; that is, it must be either "open" or "closed" ("off" or "on").

With this in mind, let "1" denote the state of a switch in the *closed* or "on" condition, and "0" denote the state of the switch in the *open* or "off" condition. Then if "A" denotes the state of a switch, the variable A can have *only two different values*, $A = 0$ and $A = 1$, respectively denoting the conditions of "open" and "closed," as depicted in the two figures below.



$A = 0$

$A = 1$

Now consider a *switching network* composed of any arrangement of switches. Let there be an input line and an output line, denoted by "1" and "Z" respectively, as in Fig. 312.

In Fig. 312 the "1' indicates that a signal (say "1 volt") is applied to the input line. Remember that the network inside the box is given to consist only of an arrangement of switches (including, of course, all necessary connecting wires).

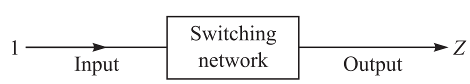


Fig. 312

In the figure Z will be either equal to $Z = 1$ or $Z = 0$ (1 volt or 0 volts), depending upon whether the input signal is able to "get through" or "not get through" the switching network. Whether $Z = 1$ or $Z = 0$ will depend upon the particular configuration of switches inside the box, and upon the particular "on" or "off" state of each switch.

Now, in regard to switching networks, let us begin with the basic *two-switch network*, in which A will denote the state of one switch (A equals 1 or 0) and B will denote the state of

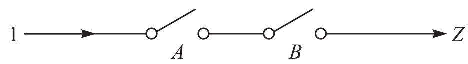


Fig. 313

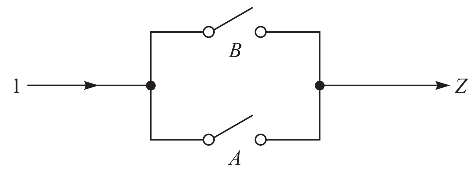


Fig. 314

the other switch (B equals 1 or 0). The two switches can be connected together in TWO *basic ways*, either in *series* or in *parallel*, as indicated in Figs. 313 and 314.

Let's first consider the SERIES *connection* of Fig. 313. Note that the input signal 1 can "get through" the network *only if* BOTH SWITCHES *are in a closed state*; that is, $Z = 1$ only if $A = 1$ *and* $B = 1$. Note that if either switch, or both switches, are open, then $Z = 0$.

Now consider the PARALLEL *connection* of Fig. 314. Note that the input signal 1 can "get through" the network *if* EITHER OR BOTH *switches are closed*; that is, $Z = 1$ if $A = 1$ *or* $B = 1$. Note that $Z = 0$ only if both $A = 0$ and $B = 0$.

The foregoing facts concerning Figs. 313 and 314 can be neatly summarized in the form of what are called *truth tables*. A "truth table" is simply a list of all possible relationships among all the signals involved, presented in a convenient table form. Thus, in Figs. 313 and 314 there are *three variables* involved, A, B, and Z, and upon remembering the facts just stated above about these figures, you can verify that the truth tables are as follows.

Truth table for 313

A	B	Z
1	1	1
1	0	0
0	1	0
0	0	0

Truth table for 314

A	B	Z
1	1	1
1	0	1
0	1	1
0	0	0

The above truth table for Fig. 313 defines what is called the "AND" *operation*, because both A and B must be 1 in order to produce $Z = 1$. The "and" operation is denoted by "$\times$", which must *not* be read as "times" but instead must be read as *and*. Inspection of the above table for Fig. 313 shows that the *four basic "and" relationships* are

$$\boxed{1 \times 1 = 1 \qquad 1 \times 0 = 0 \qquad 0 \times 1 = 0 \qquad 0 \times 0 = 0}$$

We must emphasize again that "1×1" is *not* to be read as "one times one equals one" but as "one *and* one equals zero." Likewise, $1 \times 0 = 0$ is read as "one *and* zero equals zero," and so on. The basic "and" network is thus *two switches in series*, Fig. 313, and this fact is stated in *switching algebra notation* by writing

$$\boxed{A \times B = AB = Z} \tag{559}$$

which is read "A and B is equal to Z." In writing equations we usually omit the "and" symbol "$\times$" and just write AB, as shown in eq. (559), the expression AB again being read "A and B."

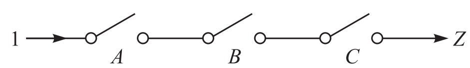
It should be emphasized that eq. (559), $A \times B = AB = Z$, is simply a shorthand way of expressing the information contained in the truth table for Fig. 313. Equation (559) is valid because we have *defined* that

(a) $A, B,$ and Z can have *only the values 1 and 0*, and that

(b) $1 \times 1 = 1, \quad 1 \times 0 = 0, \quad 0 \times 1 = 0, \quad 0 \times 0 = 0,$

and *with these rules agreed upon* it's clear that the "and" operation of eq. (559), $AB = Z$, *does* satisfy the truth table of Fig. 313 and thus represents the basic two-switch series connection of the figure. Also, since it's immaterial in what *order* switches are connected in series, it follows that either of the Boolean products, AB or BA, produces the truth table for Fig. 313. Thus Boolean "multiplication" is *commutative*, meaning that $AB = BA = Z$, just as in ordinary algebra.

It's also apparent that the foregoing extends to *any number of series-connected switches*; for example, if $A, B, C,$ and Z represent the 0 and 1 states of, say, three series-connected switches, thus



it will be noted that the truth table for this series connection is

A	B	C	Z	A	B	C	Z
1	1	1	1	0	1	1	0
1	1	0	0	0	1	0	0
1	0	1	0	0	0	1	0
1	0	0	0	0	0	0	0

which, given that $1 \times 1 \times 1 = 1, 1 \times 1 \times 0 = 0, 1 \times 0 \times 1 = 0, \ldots, 0 \times 0 \times 0 = 0,$ shows that the above truth table *is* represented by the Boolean algebraic equation $A \times B \times C = ABC = Z$, in which the Boolean product ABC is to be read as "A and B and C." Likewise for any Boolean "and" product, $ABCD \cdots = Z$, we have $Z = 1$ only if *all* the variables are equal to 1; otherwise $Z = 0$.

Now let us continue, and examine Fig. 314 and its truth table. The truth table for Fig. 314 defines what is called the "OR *operation*," because $Z = 1$ not only if A and B are both equal to 1, but also if *either A or B* is equal to 1, as inspection of the table shows. The truth table for Fig. 314 is defined to be represented mathematically by the Boolean equation

$$\boxed{A + B = Z} \tag{560}$$

in which the symbol + does *not* mean "plus" in the usual sense but is now used to indicate the "or" relationship between A and B. Thus the quantity "$A + B$" is to be read as *A or B* and *not* as "A plus B."

Remember that Boolean variables, such as A, B, Z, and so on, can have only the two values 1 and 0. With this in mind, and since eq. (560) *is defined to represent the truth table for Fig. 314*, it follows that eq. (560) can have only the four possible states

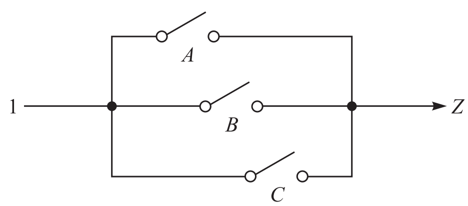
$$1 + 1 = 1 \qquad 0 + 1 = 1$$

$$1 + 0 = 1 \qquad 0 + 0 = 0$$

in which we must remember that a Boolean relationship such as "$1 + 1 = 1$" is to be read as "1 *or* 1 equals 1" and *not* as 1 "plus" 1 equals 1.

We thus have that eq. (560) is the Boolean algebra equation for the basic two-switch parallel network of Fig. 314. Also, the fact that "A and B in parallel" is the same in all respects as "B and A in parallel" is expressed mathematically by writing the Boolean equation $A + B = B + A$, meaning that Boolean "addition" is *commutative* (just as we found that Boolean "multiplication" is commutative, $AB = BA$).

It's also apparent that the foregoing extends to any number of parallel-connected switches; for example, let A, B, C, and Z denote the 0 and 1 states of three parallel-connected switches and their output; thus



Inspection of the figure shows that the truth table for this parallel connection is

A	B	C	Z		A	B	C	Z
1	1	1	1		0	1	1	1
1	1	0	1		0	1	0	1
1	0	1	1		0	0	1	1
1	0	0	1		0	0	0	0

It is, however, defined that the "or" relationship, $A + B + C$, has the value 1 in all cases *except* for the case $0 + 0 + 0$, which has the value 0; thus the above truth table *is* represented by the Boolean equation, $A + B + C = Z$ (which is to be read as "A or B or C equals Z"). Likewise *any number of parallel switches* is represented by the Boolean equation $A + B + C + D + \cdots = Z$, in which $Z = 0$ only if *all* the variables, A, B, C, D, and so on, are equal to zero; otherwise $Z = 1$.

Let's pause here, and summarize our work in the following points:

1. The Boolean variables, such as $A, B, C, \ldots Z$, can have only the two values 1 and 0.
2. The Boolean "and" operation is defined as the expression
$$ABC \cdots = Z \tag{561}$$
read as "A and B and C and ... equals Z," in which $Z = 1$ only if *all* the variables $A, B, C, \ldots$ are equal to 1; otherwise $Z = 0$. The "and" operation, eq. (561), produces the truth table of a *series connection* of switches in which the state of a switch is denoted by 1 or 0, depending upon whether the switch is closed or open.
3. The Boolean "or" operation is defined as the expression
$$A + B + C + \cdots = Z \tag{562}$$
which is read as "A or B or C or ... equals Z," in which $Z = 0$ only if *all* the variables, $A, B, C, \ldots$ are equal to 0; otherwise $Z = 1$. The "or" operation, eq. (562), produces the truth table of a *parallel connection* of switches in which the state of a switch, as in item (2), is denoted by 1 or 0, depending upon whether the switch is closed or open.

Let us now see how Boolean algebra can help in the design and simplification of more complicated switching networks.

First, a few remarks about the interpretation of Boolean variables in regard to switches. In our work we've denoted the states of switches by the Boolean variables A, B, C, and so on. Thus, if A denotes the state of a switch, we define that $A = 1$ if the switch is closed and $A = 0$ if the switch is open.

We must remember, however, that the opening and closing of a switch is controlled by a *signal* of some sort. For instance, if a transistor is being used as a switch, the ''open'' or ''closed'' state of the transistor is controlled by the signal applied to the base of the transistor.

Thus we can regard a Boolean variable, such as A, as denoting the *presence or absence* of a controlling signal; we can say that the *presence* of a signal $(A = 1)$ at a certain switch causes the switch to be in the *closed* state, while the *absence* of the signal $(A = 0)$ causes the switch to be in the *open* state.

To continue, we've already found that Boolean algebra is *commutative*; that is, $AB = BA$ and $A + B = B + A$. Now we wish to show that Boolean algebra is also *distributive*; that is, $A(B + C) = AB + AC$, just as in ordinary algebra. This can be done with the aid of Figs. 315 and 316, in which the object, now, is to show that *both networks yield the same truth table*. If this can be done, it will mean that the two networks are electrically equivalent, and thus that their Boolean equations must be equivalent. Referring now to the figures, the procedure is as follows.

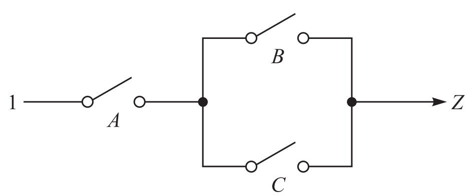


Fig. 315. $\boldsymbol{A(B+C)=Z}$**.**

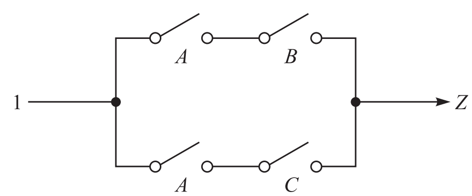


Fig. 316. $\boldsymbol{AB+AC=Z}$**.**

Consider Fig. 315 first. As the figure shows, this is a series-parallel connection of switches in which A is in series with the parallel ''or'' combination ''B or C.'' Thus the network equation is $A(B + C) = Z$ which is read as ''A, *and* the quantity B or C, is equal to Z.''

Next consider Fig. 316. Here we have two series ''and'' connections, AB, AC, connected in a parallel ''or'' configuration; hence the Boolean equation for the network is $AB + AC = Z$, which is read as ''A and B, or A and C, is equal to Z.''

We now wish to prove that both networks obey exactly the same truth table. To do this, we must find, by direct inspection of each network, the value of Z for all possible arrangements of the 1 and 0 values of A, B, and C in each network. Applying this procedure to Figs. 315 and 316 we find that

Truth table for Fig. 315:

A	B	C	Z	A	B	C	Z
1	1	1	1	0	1	1	0
1	1	0	1	0	1	0	0
1	0	1	1	0	0	1	0
1	0	0	0	0	0	0	0

Truth table for Fig. 316:

A	B	C	Z	A	B	C	Z
1	1	1	1	0	1	1	0
1	1	0	1	0	1	0	0
1	0	1	1	0	0	1	0
1	0	0	0	0	0	0	0

Note that the truth tables are identical; hence *both networks perform exactly the same switching function* and *thus their switching equations are equivalent*, proving that

$A(B+C) = AB + AC$. Thus the "distributive" property of ordinary algebra applies also to Boolean algebra.

The foregoing discussion not only shows that $A(B+C) = AB + AC$ but also demonstrates an important application of Boolean algebra, as follows.

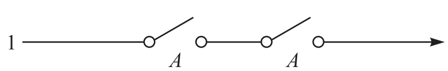
Suppose that, in the process of designing a certain piece of equipment, it is noted that Fig. 316 will correctly perform a required switching operation. Note that Fig. 316 requires *four* switches, two controlled by signal A and one each controlled by signals B and C. But note that the Boolean relationship $AB + AC = A(B+C)$ immediately shows that the *same switching operation* can be performed using just *three* switches (Fig. 315) instead of the four required by Fig. 316. This is, of course, a simple example, but it does illustrate how Boolean algebra can be used to find alternate switching networks that may be more desirable than an originally proposed network.

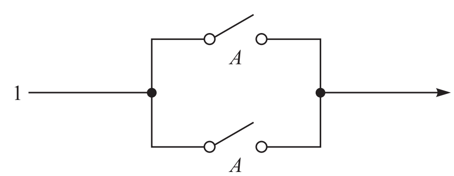
To do this efficiently, however, we must be familiar with some of the more useful theorems of the algebra. A table of such relationships is given below under the title "Theorems of Boolean algebra." Let us now consider the table, item-by-item, as follows.

To begin, we're already familiar with items (1), (2), (3), which simply state that the Boolean "and" and "or" operations are "commutative" and "distributive," just as in ordinary algebra.

4. Next consider item (4). Letting $(A+B) = X$, and applying item (3), the left-hand side of item (4) becomes $X(C+D) = XC + XD$, and upon replacing X with $(A+B)$ and making use of item (3), we have $AC + BC + AD + BD$, as stated.

The next six items, (5) through (10), follow directly from consideration of two switches whose states are *both controlled by the same signal A*; thus





and with reference to these two figures, the truth of items (5) through (10) follows.

5. "A and A equals A," because inspection shows that two such series-connected switches are equivalent to a single switch A.
6. "A or A equals A," because inspection shows that two such parallel-connected switches are equivalent to a single switch A.
7. "A and 1 equals A," because if one of two series switches is *always closed* (always equal to 1) the network effectively consists of only one switch A.
8. "A and 0 equals 0," because if one of two series switches is *always open* (always equal to 0) the network is in the 0 state regardless of the state of the other switch.
9. "A or 1 equals 1," because if one of two parallel switches is *always closed* the network is always closed (always equal to 1).
10. "A or 0 equals A," because if one of two parallel switches is *always open* the state of the network depends only on the state A of the other switch.
11. Next consider item (11), which is read as "A, or A and B, equals A." This can be verified by use of items (3) and (9); thus

$$A + AB = A(1+B) = A$$

Theorems of Boolean algebra

1.	$AB = BA$	11.	$A + AB = A$
2.	$A + B = B + A$	12.	$(A + B)(A + C) = A + BC$
3.	$A(B + C) = AB + AC$	13.	$A\bar{A} = 0$
4.	$(A + B)(C + D) = AC + AD + BC + BD$	14.	$A + \bar{A} = 1$
5.	$AA = A$	15.	$\bar{\bar{A}} = A$
6.	$A + A = A$	16.	$A + \bar{A}B = A + B$
7.	$A \cdot 1 = A$	17.	$\overline{A + B} = \bar{A}\bar{B}$
8.	$A \cdot 0 = 0$	18.	$\bar{A} + \bar{B} = \overline{AB}$
9.	$A + 1 = 1$		
10.	$A + 0 = A$		

12. Next consider item (12), which is read as "A or B, and A or C, equals A or B and C." To confirm this, let us first apply item (4) to the left-hand side of item (12):

$$\begin{aligned}(A + B)(A + C) &= AA + AC + AB + BC \\ &= A + AC + AB + BC, \text{ because } AA = A, \text{ by item (5)} \\ &= A(1 + C + B) + BC \\ &= A + BC, \text{ because } 1 + B + C = 1, \text{ by item (9)}\end{aligned}$$

which establishes the correctness of item (12).

The rest of the items in the table, (13) through (18), involve what is called the *not* operation, which is denoted by a bar placed over the variable, such as $\bar{A}$ (which is read as "not A") or $\overline{AB}$ (which is read as "not the quantity A and B"), and so on.

The "not" symbol (the bar) *reverses the 1 or 0 value of the Boolean expression it is placed above*; that is, $\bar{1} = 0$, and $\bar{0} = 1$. Thus

$$\text{if } A = 1 \quad \text{then} \quad \bar{A} = 0,$$

$$\text{or if } A = 0 \quad \text{then} \quad \bar{A} = 1,$$

or, as another example, if $A + B = 1$, then $\overline{A + B} = 0$, and so on. Electronically, the "not" operation can be performed by a transistor connected in the common-emitter mode, because the collector signal is 180° out of phase with the input base signal in such an amplifier. Thus, if A is the input signal to the base, $\bar{A}$ is the signal at the collector, or if $\bar{A}$ is applied to the base, A appears at the collector. We'll discuss this in more detail later on. Now let's consider items (13) through (18), as follows.

13. The expression $A\bar{A}$ is read as "A and not A." If $A = 1$ we have $A\bar{A} = 1 \times 0 = 0$, or if $A = 0$ we have $A\bar{A} = 0 \times 1 = 0$. Hence, either way, $A\bar{A} = 0$. Electrically, the "and" expression $A\bar{A}$ represents two series switches in which, if the first switch is closed, the second is open, and vice versa; thus the network is *always open*, that is, always is in the "0" state.
14. $A + \bar{A}$ is read as "A or not A." If $A = 1$ we have $1 + 0 = 1$, or if $A = 0$ we have, $0 + 1 = 1$. Note that the expression $A + \bar{A}$ represents two parallel switches in which one or the other will always be closed, so that the network is always in the "1" state.

15. Here, $\bar{\bar{A}}$ is read as "not, not A," in which the second "not" undoes the first "not," thus giving us back A, as would be expected. Thus if $A = 1$ we have $\bar{\bar{1}} = \bar{0} = 1$, or if $A = 0$ we have $\bar{\bar{0}} = \bar{1} = 0$.
16. The relationship is read as "A, or the quantity not A and B," equals "A or B." Let us begin with the right-hand side of item (16); thus

$$\begin{aligned}(A+B) &= (A+B)(A+\bar{A}), \text{ permissible by items (14) and (7),}\\ &= A + AB + \bar{A}\,B, \text{ because } AA = A, \text{ and } A\,\bar{A} = 0,\\ &= A(1+B) + \bar{A}\,B,\\ &= A + \bar{A}\,B, \text{ because } 1 + B = 1 \text{ by item (9). Thus (16) \textit{is} correct.}\end{aligned}$$

17. The relationship is read as "not the quantity A or B" equals "not A and not B." The truth of a Boolean equation can be proved, or disproved, by showing that both sides of the equation are equal *for all possible ways in which 1 and 0 can be assigned to the variables*. This procedure produces a "truth table" for each side of the equation, and the two truth tables must be identical in order for the equation to be valid. In this work we make use of the basic "and" and "or" relationships, which we know to be $1 \times 1 = 1, 1 \times 0 = 0, 0 \times 1 = 0, 0 \times 0 = 0$; also, $1 + 1 = 1, 1 + 0 = 1, 0 + 1 = 1, 0 + 0 = 0$. For item (17) the procedure gives the following results (where "tt" stands for "truth table").

A	B	tt for $\overline{A+B}$	tt for $\bar{A}\,\bar{B}$
1	1	$\overline{1+1} = \bar{1} = 0$	$\bar{1} \times \bar{1} = 0 \times 0 = 0$
1	0	$\overline{1+0} = \bar{1} = 0$	$\bar{1} \times \bar{0} = 0 \times 1 = 0$
0	1	$\overline{0+1} = \bar{1} = 0$	$\bar{0} \times \bar{1} = 1 \times 0 = 0$
0	0	$\overline{0+0} = \bar{0} = 1$	$\bar{0} \times \bar{0} = 1 \times 1 = 1$

The table shows that $\overline{A+B} = \bar{A}\bar{B}$ for all possible combinations of values of A and B, thus proving that item (17) *is correct*.

18. The relationship is read as "not A or not B" is equal to "not the quantity A and B." Let's now use the same "truth table" procedure as for item (17), as follows.

A	B	tt for $\bar{A}+\bar{B}$	tt for $\overline{AB}$
1	1	$\bar{1}+\bar{1} = 0 + 0 = 0$	$\overline{1 \times 1} = \bar{1} = 0$
1	0	$\bar{1}+\bar{0} = 0 + 1 = 1$	$\overline{1 \times 0} = \bar{0} = 1$
0	1	$\bar{0}+\bar{1} = 1 + 0 = 1$	$\overline{0 \times 1} = \bar{0} = 1$
0	0	$\bar{0}+\bar{0} = 1 + 1 = 1$	$\overline{0 \times 0} = \bar{0} = 1$

The table shows that $\bar{A} + \bar{B} = \overline{AB}$ for all possible ways in which 1 and 0 can be assigned to the variables A and B, thus proving that item (18) *is valid*. Consider now the following three examples in the use of the table.

Example 8

Simplify the Boolean expression $\bar{A} + \overline{A+B}$.

Solution
First make use of item (17), then item (9); thus

$$\bar{A} + \bar{A}\,\bar{B} = \bar{A}(1 + \bar{B}) = \bar{A}, \text{ \textit{answer}}$$

Note: In regard to the use of item (9) above, it should be understood that the expression $1 + A = 1$, is true regardless of what Boolean expression *A might actually represent*; for example, $1 + ABCD = 1$, or, $1 + B + C + D = 1$, and so on. We thus regard the expression "$1 + A = 1$" as the generic relationship, valid for whatever the actual form of A might be. The same statement holds for the other relationships in the table; for example, the Boolean relationship

$$\overline{AB} + \overline{C + D}$$

has the basic form of item (18); and thus, by item (18), we have that

$$\overline{AB} + \overline{C + D} = \overline{AB(C + D)}$$

Example 9
Simplify the Boolean expression $(\overline{A} + \overline{B})C + AB$.

Solution
Using item (18), write the given expression in the form

$$\overline{AB}\,C + AB = AB + \overline{AB}\,C \quad \text{(the ``or'' operation is commutative)}$$

to which now apply the generic form of item (16) to get $AB + C$, *answer*.

It should be noted that the new expression, $AB + C$, does exactly the same switching job as the original expression but requires only one "and" and one "or" operation, whereas the original expression requires two "and" operations, two "or" operations, and two "not" (inverter) operations.

Example 10
Given $Z = ABC + D(\overline{A} + \overline{B} + \overline{C})$, find a simpler expression for Z.

Solution
First, by item (18), $\bar{A} + \bar{B} + \bar{C} = \overline{AB} + \bar{C} = \overline{ABC}$, so that the given problem now becomes

$$Z = ABC + \overline{ABC}\,D$$

which is the generic form of item (16); and thus, $Z = ABC + D$, *answer*, which represents a considerably simpler switching network than the original expression.

Problem 273
$A + A + A + A + B + B + B =$

Problem 274
$AAABBCCCC =$

Problem 275
$A + B + C + ABC + 1 =$

Problem 276
Using the table of basic theorems, show that

$$\overline{A + B + C} = \bar{A}\,\bar{B}\,\bar{C}$$

Problem 277
Using the table of basic theorems, show that

$$\bar{A} + \bar{B} + \bar{C} = \overline{ABC}$$

Note: In problems 278 through 284, which follow, simplify the given Boolean equations. Try to minimize both the number of switching operations and the number of "not" operations needed.

Problem 278
$Z = A + AB$

Problem 279
$Z = A(AB + AB\bar{C})$

Problem 280
$Z = A(A + B)(B + C)$

Problem 281
$Z = AC + BC + \bar{B}C + A\bar{C} + \bar{B}\bar{C}$

Problem 282
$Z = B + \bar{C}\bar{D}\overline{(B + D)}$

Problem 283
$Z = \overline{ABC + \bar{A}BC + \bar{A}\bar{B}\bar{C}}$

Problem 284
$Z = \bar{A}\bar{B}\bar{C}\bar{D} + \overline{A + C + D}$

Problem 285
Write $A + \overline{A + \bar{B}}$ in a form without "not" operations.

Problem 286
Write the equation $Z = \overline{A\bar{B}C + \bar{A}\bar{B}C + A\bar{B}\bar{C}}$ in a form that will require only one "not" operation.

12.3 Digital Logic Symbols and Networks

A digital LOGIC NETWORK is an arrangement of "and," "or," and "not" devices that will always produce a required form of output signal for a given set of input binary signals. For example, let A, B, and C denote the 0, 1 values of three binary signals, and suppose it is necessary to generate the output signal, $Z = AB + \bar{C}$. The logic network needed to perform this particular operation is shown in block diagram form in Fig. 317.

We'll not concern ourselves presently with the physical details of what might be actually "inside" each of the above boxes.

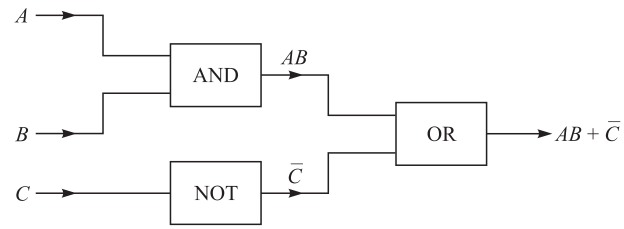
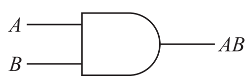
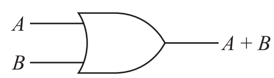


Fig. 317

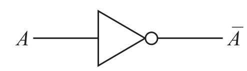
In Fig. 317 we've designated the function of each block with the words ''AND,'' ''OR,'' and ''NOT.'' Instead of words, however, certain symbolic forms have been universally adopted to indicate these operations, as follows.



AND

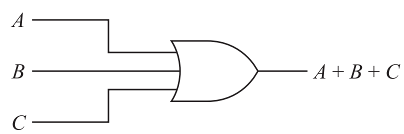


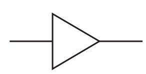
OR



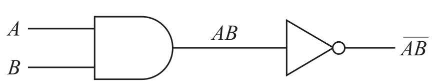
NOT

The above symbols should be committed to memory. In regard to the AND and OR symbols, it should be understood that more than just *two* input lines can be used; for instance, if, say, three binary signals, A, B, and C, are involved, this would be indicated by using three input lines, thus

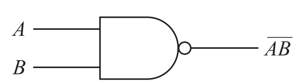


The triangular symbol by itself, , indicates ''amplifier''; it is the addition of the *small circle* at the output side of the triangle that designates that the amplifier is used in the NOT or ''inverting'' mode.

If the output of an AND network is passed through a NOT network, the result is NOT AND (abbreviated NAND), illustrated below to the left.



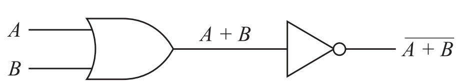
''NOT AND'' or ''NAND'' operation



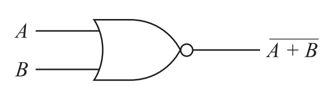
Simplified symbol for the NAND operation

The two-symbol NAND drawing (left above), is often expressed in a shortened form by the simple addition of a small ''not'' circle at the output side of the AND symbol, as shown to the right above.

In the same way, if the output of an OR network is passed through a NOT network, the result is NOT OR (abbreviated NOR), illustrated to the left below, with the simplified version shown in the figure on the right below.



''NOT OR'' or ''NOR'' operation



Simplified symbol for the NOR operation

Now consider the following. Suppose we are given an array of input signals, in the form of on-or-off pulses representing the binary digits 1 and 0, and suppose we must find a switching network that will produce a desired result. In other words, the problem is, *given a truth table*, FIND A DIGITAL SWITCHING NETWORK that will satisfy the given truth table.

One procedure that can be used to find such required circuitry is to begin by writing down the basic or "elemental" Boolean equation for the given truth table. The "elemental equation" for a given truth table is a Boolean AND-OR relationship in which *each* AND *term contains all the variables*. This means that, if, for example, "A" denotes one of the variables, then either "A" or "not A" (A or $\bar{A}$) must appear in each of the "and" terms of the equation.

For instance, if we are dealing with, say, *three* binary input signals, denoted by A, B, and C, then the elemental Boolean equation for a required switching system will be of the AND-OR form.

$$Z = ABC + A\bar{B}C + AB\bar{C} + A\bar{B}\bar{C} + \bar{A}BC + \cdots + \bar{A}\bar{B}\bar{C}$$

and likewise for any number of input variables, $A, B, C, D, \ldots$, *in which only those* AND *terms that will produce an output signal will be used*; that is, *only those* AND *terms for which* $Z = 1$ will be used. Consider the following two examples.

Example 11

Suppose three binary signals, denoted by A, B, and C, are to be switched in such a way as to satisfy the truth table

A	B	C	Z	A	B	C	Z
1	1	1	0	0	1	1	1
1	1	0	0	0	1	0	0
1	0	1	1	0	0	1	1
1	0	0	0	0	0	0	0

Write the elemental equation for the required switching network and simplify as much as possible.

Solution

In accordance with the foregoing rule, inspection of the given truth table shows that the elemental equation is

$$Z = A\bar{B}C + \bar{A}BC + \bar{A}\bar{B}C$$

Now, while the above elemental form *will* do the required switching, it has the disadvantage of requiring two NOT circuits, three AND circuits, and one OR circuit. An equivalent but simpler circuit can, however, be found by applying the Boolean theorems to the above elemental expression; let us begin by factoring out the C signal; thus

$$Z = (A\bar{B} + \bar{A}B + \bar{A}\bar{B})C$$

hence

$$Z = (A\bar{B} + \bar{A})C = (\bar{A} + A\bar{B})C$$

because

$$\bar{A}B + \bar{A}\bar{B} = \bar{A}(B + \bar{B}) = \bar{A} \cdot 1 = \bar{A}, \text{ from the theorems.}$$

The last expression for Z can now be simplified further, as follows. By item (16),

$$A + \bar{A}\,B = A + B$$

therefore

$$\bar{A} + A\,\bar{B} = \bar{A} + \bar{B}$$

because $\bar{A} + A\,\bar{B}$ has the *same basic form* as $A + \bar{A}\,B$, with $\bar{A}$ written in place of A and $\bar{B}$ written in place of B. Therefore the last expression for Z becomes

$$Z = (\bar{A} + \bar{B})C = \overline{AB}\,C, \text{ by item (18), } (\textit{final answer})$$

which, using graphic block diagrams, is drawn as in Fig. 318.

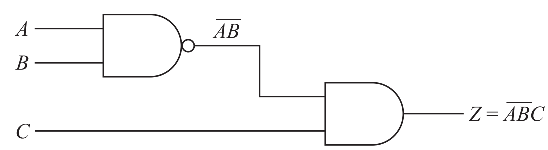


Fig. 318

Commentary. We must remember that the input signals, A, B, and C, are assumed to be in the form of streams of on-or-off pulses, each pulse representing the value "1" or "0," depending upon the presence or absence of the pulse (as illustrated in Fig. 311).

For instance, in Fig. 318 a portion of the streams of simultaneously applied pulses might be as follows,

$$\xrightarrow{\text{Time}}$$

$A \ldots 1\,0\,0\;\; 1\;\; 1\,1\;\; 0\,0\,1\;\; 0\,1\,1\,0\,1\;\; 0\;\; 1\,0 \cdots \rightarrow$

$B \ldots 1\,1\,0\;\; 0\;\; 0\,1\;\; 1\,0\,1\;\; 1\,1\,0\,1\,1\;\; 0\;\; 0\,1 \cdots \rightarrow$

$C \ldots 0\,0\,0\;\; 1\;\; 0\,1\;\; 1\,0\,1\;\; 1\,0\,0\,0\,1\;\; 1\;\; 0\,0 \cdots \rightarrow$

$Z \ldots 0\,0\,0\;\; 1\;\; 0\,0\;\; 1\,0\,0\;\; 1\,0\,0\,0\,0\;\; 1\;\; 0\,0 \cdots \rightarrow$

Hence the network of Fig. 318 delivers an output pulse ($Z = 1$) only for the input combinations of 101, 011, and 001, thus satisfying the requirements of the truth table given with this example. No pulse output appears at the output ($Z = 0$) for any other combination of input 1's and 0's, as required by the truth table.

Example 12

Let A, B, C, and D represent four binary input signals that must be switched so as to always fulfill the following truth table

A	B	C	D	Z	A	B	C	D	Z	A	B	C	D	Z	A	B	C	D	Z
1	1	1	1	0	1	0	1	1	0	0	1	1	1	0	0	0	1	1	0
1	1	1	0	0	1	0	1	0	0	0	1	1	0	0	0	0	1	0	0
1	1	0	1	1	1	0	0	1	1	0	1	0	1	1	0	0	0	1	1
1	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0

Write the elemental equation for the truth table, then find a simplified equivalent.

Solution
Let us apply the foregoing procedure to the given truth table. Doing this gives the elemental equation

$$Z = AB\bar{C}D + AB\bar{C}\bar{D} + A\bar{B}\bar{C}\bar{D} + \bar{A}B\bar{C}D + \bar{A}\bar{B}\bar{C}D$$

The simplification of such AND-OR relationships is largely a trial-and-error procedure, in which the first step is to try various *factoring* arrangements. We then try to apply, to the factored equation, the basic Boolean theorems given earlier.

For the case of the Boolean relationships given above, we could, for example, begin by factoring $AB\bar{C}$ out of terms 1 and 2, and $\bar{B}\bar{C}D$ out of terms 3 and 5, thus giving us

$$Z = AB\bar{C}(D + \bar{D}) + (A + \bar{A})\bar{B}\bar{C}D + \bar{A}B\bar{C}D$$

hence,

$$Z = AB\bar{C} + \bar{B}\bar{C}D + \bar{A}B\bar{C}D, \text{ by items (14) and (7),}$$

then,

$$Z = (AB + \bar{B}D + \bar{A}BD)\bar{C}$$

The next step would be to try to apply the theorems to the quantity inside the parentheses above, and this can be done. Instead of doing this now, however, *let's start again*, this time observing that $\bar{C}D$ factors out of *all terms except the second*, thus putting the original expression for Z in the form

$$Z = [A(B + \bar{B}) + \bar{A}(B + \bar{B})]\bar{C}D + AB\bar{C}\bar{D}$$
$$Z = \bar{C}D + AB\bar{C}\bar{D} = (D + \bar{D}AB)\bar{C}$$

hence, by item (16),

$$Z = (D + AB)\bar{C} = (AB + D)\bar{C} \quad (\textit{final answer})$$

It thus appears that the second approach, in which we factored out $\bar{C}D$, seems preferable to the first approach. The final answer above represents, in Boolean algebra form, the switching network shown in block diagram form in Fig. 319.

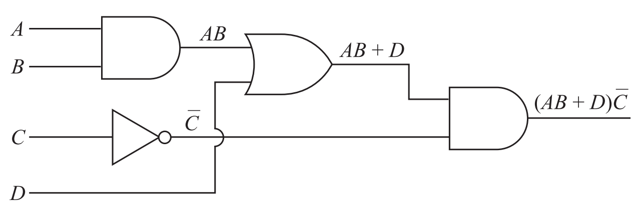


Fig. 319

Now let's turn our attention to an important network called the ''full adder.'' We have already seen that the basic mathematical tool in a digital computer is *addition*; we found, for example, that in binary arithmetic the subtraction operation can be performed by using addition in conjunction with the 1's complement. Multiplication and division are likewise performed by use of addition; thus multiplication is accomplished by repeated additions, while division is accomplished by repeated subtractions. Hence, since addition is such a useful operation in digital computers, it's fitting that we next take up the basic ''full adder'' network.

To begin, consider any two binary numbers that are to be added together, and let A and B be two binary digits of the same order, in the same column, such as is shown below:

$$\begin{array}{r} 1\;1\;A\;0\;1\;1 \\ +\;\underline{1\;0\;B\;1\;1\;0} \end{array}$$

When adding any two digits such as A and B together, we must take into account *not only the sum of the two digits themselves* but also the effects of any CARRY *digits* involved in the operation. For instance, in the particular example above, a "1" would be carried *into* the A, B column from the next lower order column to the right; there then may, or may not, be a "1" carried into the next higher order column to the left, depending upon the values of A and B.

Let us now define that a FULL ADDER is a network capable of producing the *sum of two binary numbers*, including the handling of all "carry" digits that might arise in the operation. This is illustrated in block diagram form in Fig. 320, where "fa" denotes a "full adder" network.

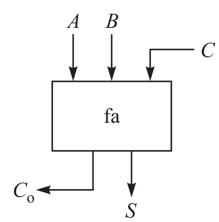


Fig. 320

In Fig. 320, A and B are binary digits, 1 or 0, of the same order; $C =$ binary 1 or 0, carried "in" from the next lower order; $S =$ sum digit, 1 or 0, of same order; and $C_o =$ binary 1 or 0, carried "out" to the next higher order.

Remembering now the binary sums $\begin{array}{r}1\\+\underline{0}\\1\end{array}\quad\begin{array}{r}1\\+\underline{1}\\10\end{array}\quad\begin{array}{r}1\\1\\+\underline{1}\\11\end{array}$, you can verify that the truth table for the above full adder is

A	B	C	S	C_o		A	B	C	S	C_o
1	1	1	1	1		0	1	1	0	1
1	1	0	0	1		0	1	0	1	0
1	0	1	0	1		0	0	1	1	0
1	0	0	1	0		0	0	0	0	0

The elemental equations for the sum digit S and the carry-out digit C_o are, therefore,

$$S = ABC + A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C \tag{563}$$

$$C_o = ABC + AB\bar{C} + A\bar{B}C + \bar{A}BC \tag{564}$$

Now, after considerable trial and error, we find that eq. (563) can be written in the equivalent form

$$S = ABC + (A + B + C)(\bar{A}\bar{B} + \bar{A}\bar{C} + \bar{B}\bar{C})$$

a fact you can verify by multiplying as indicated and remembering that $X\bar{X} = 0$. Now, making use of items (17) and (18), the last expression can be written as

$$S = ABC + (A + B + C)(\overline{A + B} + \overline{A + C} + \overline{B + C})$$

hence,

$$S = ABC + (A + B + C)\overline{(A + B)(A + C)(B + C)}$$

therefore,

$$S = ABC + (A + B + C)\overline{(AB + AC + BC)} \tag{565}$$

The form of eq. (565) is especially useful because, as we'll next show, it will fit in nicely with the generation of the carry-out digit C_{o}. To show this, let us write eq. (564) in the equivalent form

$$C_{\mathrm{o}} = ABC + ABC + ABC + AB\,\bar{C} + A\,\bar{B}\,C + \bar{A}\,BC$$

which is permissible because, in Boolean algebra, $X + X + X + \cdots + X = X$. Now combine together terms 1 and 4, 3 and 5, and 2 and 6; thus

$$C_{\mathrm{o}} = AB(C + \bar{C}) + AC(B + \bar{B}) + BC(A + \bar{A})$$

hence,

$$C_{\mathrm{o}} = (AB + AC + BC)$$

which, as you can see, fits in perfectly with eq. (565); thus the switching equations for the FULL ADDER, Fig. 320, can be written as

$$S = ABC + (A + B + C)\overline{(AB + AC + BC)} \tag{566}$$

$$C_{\mathrm{o}} = (AB + AC + BC) \tag{567}$$

which, as you'll note, requires only one inverse ("not") operation, and no inversions of the individual inputs. Thus the last two equations translate into Fig. 321, which shows in block diagram form the details of the contents of the box in Fig. 320.

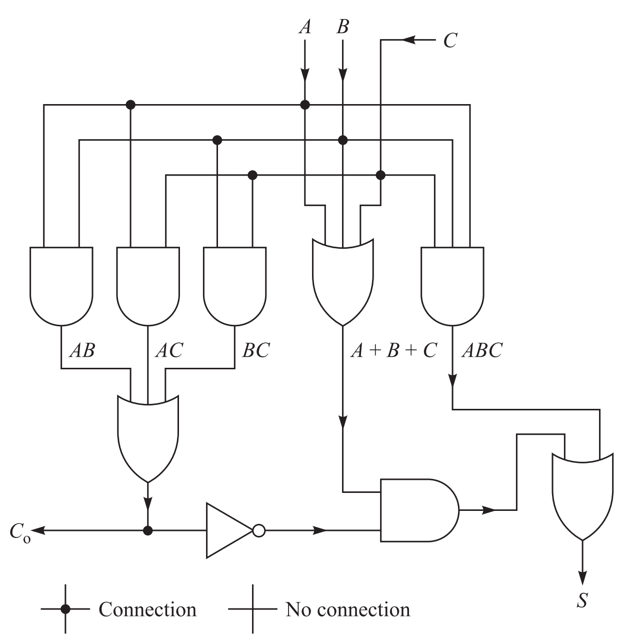


Fig. 321. Full adder.

In the foregoing full adder, A, B, and C represent three binary digits of the same order. In each case the presence or absence of a pulse denotes 1 or 0 respectively. The network delivers a sum digit S of the same order as A, B, and C, and a possible carry digit C_o of the next higher order. Note that, in order to get the sum of two binary numbers, such as, for example,

$$\begin{array}{r} 1\,0\,1\,1\,0\,1\,1\,0 \\ +\underline{1\,1\,0\,1\,1\,1\,0\,1} \end{array}$$

we must use a full adder for each two digits of the same order. This is illustrated in Fig. 322, where "fa" stands for "full adder."

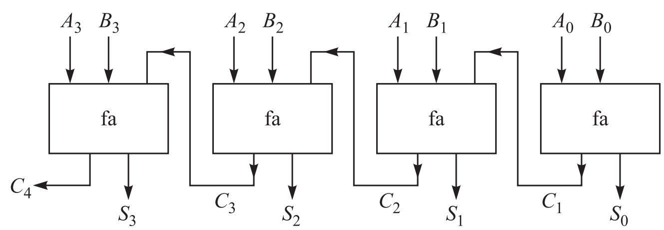


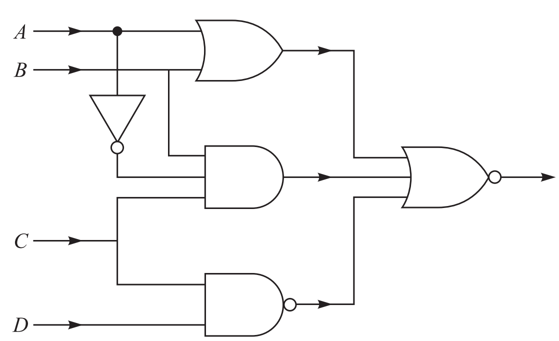
Fig. 322

In the figure, the As denote the digits of one binary number and the Bs denote the corresponding digits of another binary number. The digits of the lowest order (least value) are denoted by A_0 and B_0, with S_0 being the lowest-order sum digit.

In the above circuitry the sum and carry digits appear only during the "on time," that is, during the time during which the A and B pulses are on the input lines. After a given set of A and B pulses is terminated, an "off time" is provided before a next set of A and B pulses is applied to the input lines. During the "off time" all sum and carry digits are zeros (because the As and Bs are all zeros during the off time). Therefore, during the "on time" the sum digits must be fed out of Fig. 322 into what is called an "accumulator" or "register," which is a circuit capable of storing the total sum of all such sums generated.

Problem 287

In the following, A, B, C, and D represent four binary input signals. Write the Boolean expression for the signal appearing on each line in the diagram.



Problem 288
Using Boolean algebra, find a simpler arrangement for the network of problem 287 consisting of just one "nor" and two "and" devices. Resketch the answer to problem 287 showing the new, but equivalent, network.

Problem 289
Two input binary signals A and B are to be switched to satisfy the following truth table:

A	B	Z	A	B	Z
0	0	1	1	0	1
0	1	0	1	1	1

Write the elemental equation, simplify, sketch a circuit to consist of one "or" and one "nor" device.

Problem 290
A network must be devised that will cause two input binary signals A and B to satisfy the following truth table:

A	B	Z	A	B	Z
0	0	1	1	0	0
0	1	0	1	1	1

(a) Write the elemental equation that will satisfy the truth table.
(b) Sketch a network using one "nor" circuit and two "and" circuits that will satisfy the truth table.

Problem 291
A network is required that will cause three input binary signals A, B, C to satisfy the following truth table:

A	B	C	Z	A	B	C	Z
0	0	0	1	1	0	0	0
0	0	1	0	1	0	1	1
0	1	0	0	1	1	0	0
0	1	1	0	1	1	1	1

(a) Write the elemental equation that will satisfy the table.
(b) Sketch a network using one "and" circuit, one "nor" circuit, and one "or" circuit that will satisfy the table.

Problem 292
Let A, B, C, D represent the states of four input binary signals that must be switched so as to fulfill the following truth table:

A	B	C	D	Z	A	B	C	D	Z	A	B	C	D	Z	A	B	C	D	Z
0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	0	0	0
0	0	0	1	0	0	1	0	1	0	1	0	0	1	0	1	1	0	1	0
0	0	1	0	1	0	1	1	0	1	1	0	1	0	0	1	1	1	0	1
0	0	1	1	0	0	1	1	1	0	1	0	1	1	1	1	1	1	1	1

(a) Write the elemental equation that will satisfy the table.
(b) Simplify the answer to part (a) to a form that requires "and" and "or" terms, but just one "not" term.
(c) Using standard symbols, sketch the answer to part (b) in block diagram form.

Problem 293

Let A, B, C, D denote four input binary signals that must be switched to satisfy the following truth table where, as usual, Z is the state of the output binary signal:

A	B	C	D	Z	A	B	C	D	Z	A	B	C	D	Z	A	B	C	D	Z
0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0
0	0	0	1	0	0	1	0	1	0	1	0	0	1	0	1	1	0	1	0
0	0	1	0	1	0	1	1	0	0	1	0	1	0	1	1	1	1	0	0
0	0	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1

(a) Write the elemental equation that will satisfy the table.
(b) Simplify the answer to (a) into a form that requires only one "not" operation.
(c) Using standard symbols, sketch the answer to (b) in block diagram form.

CHAPTER 13

The Digital Processor. Digital Filters

13.1 Bandwidth Requirements for Digital Transmission. Sampling Theorem. PAM and PCM

We have learned that in a digital system information is given, and transmitted, in the form of short "rectangular-type" pulses of voltage or current, the presence or absence of a pulse denoting "1" or "0" (as illustrated in Fig. 311 in Chap. 12).

It should be noted that the transmission of information in pulse form requires that the equipment be able to uniformly amplify and pass a wide range of frequencies; that is, it must possess a relatively WIDE BANDPASS characteristic. This is because a rectangular pulse type of signal is composed of a large number of harmonic frequencies (note 18 in Appendix). Consider now Figs. 323 and 324, in which T is the uniform amount of *time* allotted to the appearance of each pulse, T having the same value in both figures.

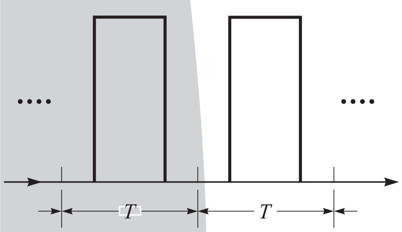


Fig. 323

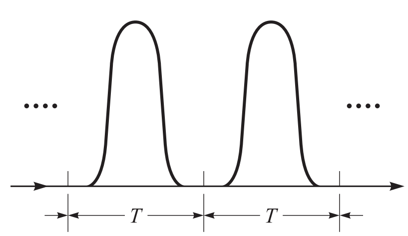


Fig. 324

Now imagine two streams of pulses, one composed of those of Fig. 323, the other of Fig. 324. Then note that the *frequency F* of the *fundamental* component will be the *same* in both cases; thus

$$F = 1/T \qquad \text{(eq. (91), Chap. 5)}$$

because T is given to be the same in both cases (but the fundamental waves would *not*, in general, have equal amplitudes; that is, they would not have equal "peak values").

The point, however, that we wish to emphasize here is that, for practical purposes, it will require *more higher-order harmonics of F* to get a "sufficiently good" representation of Fig. 323 than Fig. 324. This is because the "ideal" rectangular pulse of Fig. 323 has what are called "points of discontinuity," that is, times at which the voltage or current would have to INSTANTLY jump from one value to a different value, which is, of course, impossible in the real world. Another, but less severe, type of discontinuity would occur if a voltage or current were required to instantly alter its "rate of change," such as instantly changing from, say, an increasing value to a decreasing value.

To get an "almost exact" representation of Fig. 323 would require the inclusion of a *large number of higher-order harmonics* of the fundamental frequency; thus, to transmit a nearly exact form of Fig. 323 through a system would require that the system have a relatively WIDE BANDPASS characteristic.

On the other hand, note that the pulses depicted in Fig. 324 are relatively "smooth," having virtually no points of discontinuity. Thus, pulses in the form of Fig. 324 could be transmitted through a system having a considerably NARROWER BANDPASS than that required for Fig. 323.

The point we wish to make is that, if the pulse train of Fig. 324 is "good enough" to do the job (of representing 1's and 0's), then we need not try to make the train more closely resemble the ideal case of Fig. 323.

The amount of circuit "bandwidth" required must especially be considered if the information is to be transmitted by wireless; that is, if the information, in digital form, is used to modulate a high-frequency "carrier wave." This is because, in the case of a modulated wave, the information is not actually contained in the carrier wave itself but, instead, is contained in a cluster of "side-band" waves, with the carrier in the center of the cluster (see Fig. 31-A, note 24 in Appendix).

For this reason the *total bandwidth* required to transmit a modulated wave depends only upon the HIGHEST FREQUENCY COMPONENT present in the information being transmitted. Thus, if f_h is the highest frequency component of importance in the information signal, then, if the carrier is amplitude-modulated (AM), it would require a total bandwidth of $2f_h$ to transmit the information without distortion. (For practical purposes the same bandwidth requirement, $2f_h$, applies to a frequency-modulated carrier.)

In general, to prevent interference between stations transmitting on adjacent frequencies, it's necessary to limit the amount of bandwidth allocated to each station. This means that we must decide what constitutes the "highest frequency of importance" in a given case. This is, of course, an engineering judgement that must be tempered by the restriction on maximum allowable bandwidth.

In the previous chapter we dealt with digital switching systems, and how such systems can be used to make purely arithmetic calculations.

Let's now consider another very important application of digital signals, in which information *in* ANALOG FORM is converted into and transmitted *in* DIGITAL FORM. (Then generally, at the end of the transmission system, the final step is to convert the digital signal back into its original analog form.) Let us, at this point, simply state that the *reason* for using such a system is that it makes it possible to GREATLY REDUCE THE EFFECTS OF ALL TYPES OF NOISE.

The actual conversion of an analog signal into a corresponding digital signal is accomplished by the process of SAMPLING the analog signal. The success of the system depends upon our ability to accurately *reconstruct the original analog signal* from its sampled values at the receiving end of the system. The basic theory rests upon the fundamental and famous SAMPLING THEOREM, which states that, if $v(t)$ denotes an original analog signal, then

> In order to recover an analog signal $v(t)$ from equally spaced samples of $v(t)$, the SAMPLING RATE must be somewhat greater than twice the highest frequency component of significance in $v(t)$.

Note that the sampling rate must be ''somewhat greater'' than the highest significant frequency component of $v(t)$. This is necessary, in practical work, to prevent possible distortion of the recovered signal. For example, if the highest frequency component of $v(t)$ were, say, taken to be 3000 Hz, then $v(t)$ would be sampled at a rate somewhat in excess of 6000 Hz, say 6500 Hz in a practical application.

In order to fully satisfy the sampling theorem, the *height* of each sampling pulse would have to represent the *exact* value of $v(t)$ at the instant of sampling. It is, of course, not possible to do this with 100% accuracy in a practical system; let us discuss this with the aid of Figs. 325 and 326.

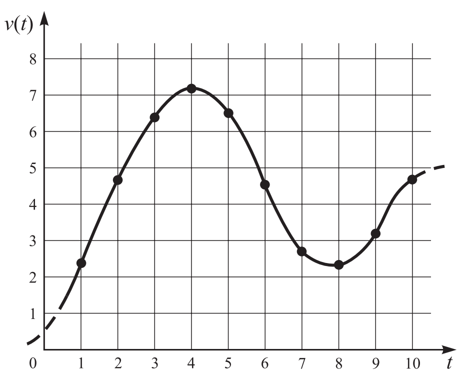


Fig. 325

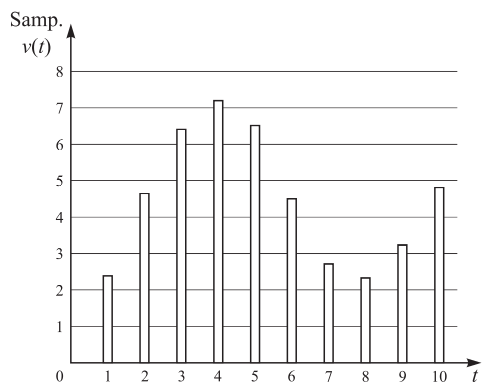


Fig. 326. Sampled form of $v(t)$.

Let Fig. 325 show a portion of an analog signal voltage $v(t)$, where we'll assume $v(t)$ is limited to values between 0 and 8 volts. The regular intervals of time on the horizontal axis can represent seconds, milliseconds, or microseconds, as the case may be.

Now consider Fig. 326, which, let us assume, shows IDEAL SAMPLING of $v(t)$, which means that the *amplitude* or height of each sampling pulse represents the *exact* value of the corresponding amplitude of $v(t)$ at that instant. In such a case it would be theoretically possible to *completely recover the original analog signal $v(t)$ from the sample pulses*; this assumes, of course, that the basic requirement of the sampling theorem, regarding sampling rate, is satisfied.

If the pulse samples in Fig. 326 are used to amplitude-modulate a high-frequency carrier so as to be transmitted by wireless, we have what is called ''pulse-amplitude modulation,'' abbreviated PAM.

The advantage of PAM is that it is ideally suited for use in what is called ''time-division-multiplex,'' a system which allows several different signals to be sent together over the same transmission system, using the same carrier frequency. The *disadvantages* of

PAM are that it requires increased bandwidth with *no basic reduction in the effect of noise* (in comparison with analog transmission). Another disadvantage is that a PAM system must be free of amplitude distortion, a requirement that complicates the design of PAM systems as compared with simple ‘‘present or not-present’’ pulse systems.

The just-mentioned shortcomings of PAM can be overcome by using what is called ‘‘pulse-code modulation,’’ abbreviated PCM. An outline of this remarkably effective system is as follows, in which, as before, $v(t)$ will denote the original ANALOG signal it is desired to transmit in PCM form.

The *first step* in a PCM system is the same as in PAM; that is, the ANALOG signal must be SAMPLED as in Figs. 325 and 326.

The *second step* in forming a PCM signal is to LIMIT THE NUMBER OF DIFFERENT AMPLITUDES the sampling pulses can have. This is called ‘‘quantizing’’ the sampled signal, and is accomplished by using a ‘‘quantizer’’ circuit at the output of the sampling circuit. Thus, at the output of the quantizer the amplitude of each sample pulse represents a RANGE *of possible values of* $v(t)$, instead of an unlimited number of amplitudes as in Fig. 326. Only a quantized signal, which has an *exact* (integral) number of different values, can be converted into a PCM signal.

In Fig. 326 it’s possible for a sample pulse to have ANY amplitude in the range from 0 to 8 volts. Now suppose the pulses of Fig. 326 are fed into a quantizer circuit that is capable of sorting the pulse amplitudes into, let us say, 8 *different ranges*, in accordance with the table to the left of Fig. 327. If we now redraw Fig. 326 in accordance with the table, the result is Fig. 327.

Range of pulse amp.	Quantized value V_q
0 to 1	0
1 to 2	1
2 to 3	2
3 to 4	3
4 to 5	4
5 to 6	5
6 to 7	6
7 to 8	7

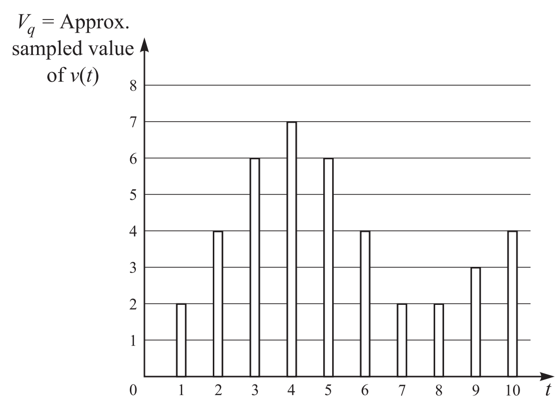


Fig. 327. Quantized form of Fig. 326.

Thus the amplitude of each sample pulse in Fig. 327 represents a RANGE *of possible values of* $v(t)$, in accordance with the given table. Take, for example, time $t = 9$ in Fig. 327. According to Fig. 327, $v(t) = 3$ volts at $t = 9$; actually, however, inspection of the table shows that the value of $v(t)$ could be ANY VALUE in the range from 3 to 4 volts. This is, of course, just an illustration, and we would generally require greater accuracy in our work. We might, for example, require that the pulse amplitudes be broken down into, say, 32 different ranges (‘‘quantization levels’’) for greater accuracy. There is, however, always some amount of ‘‘quantization error’’ produced. The amount of such error allowed depends, of course, upon the particular application.

The *third* and final step is to ‘‘encode’’ the quantized PAM signal of Fig. 327 into a PCM signal, which is a signal in which ONLY TWO DIFFERENT PULSE CONDITIONS, ‘‘on or off’’ (‘‘1’’ or ‘‘0’’), have to be detected, regardless of the different pulse amplitudes actually present in Fig. 327. This is done by representing *each quantized level* by a *specified binary number*, all pulses in the binary numbers having the *same amplitude*. Thus, in PCM it will not be necessary to detect different pulse amplitudes, but only if a pulse is "present" or "not present," "1" or "0." Let us take the case of Fig. 327 as an example, as follows.

As stated above, in PCM each of the various amplitude levels in the quantized sampled form of the analog signal is to be represented by a specific binary number. Since there are 8 different voltage levels possible, including zero, in Fig. 327, this will require the use of 8 binary numbers of 3 binary digits each; thus

Pulse amplitude	Binary form	Pulse amplitude	Binary form
0	000	4	100
1	001	5	101
2	010	6	110
3	011	7	111

Thus the information, contained in Fig. 327, is expressible as follows

Time, t	Amplitude	Coded form	Time, t	Amplitude	Coded form
1	2	010	6	4	100
2	4	100	7	2	010
3	6	110	8	2	010
4	7	111	9	3	011
5	6	110	10	4	100

The above information, in binary-coded form (PCM), can be basically represented (for $t = 1$ through $t = 6$) as shown in Fig. 328.

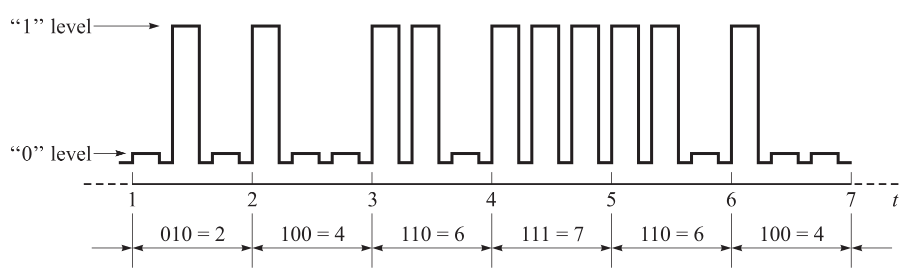


Fig. 328

The advantage of PCM is its great freedom from the effects of noise. This is because, in order to extract the information from such a signal it is only necessary to determine whether a pulse is PRESENT or NOT PRESENT; that is, it is not necessary to know the exact amplitudes or shapes of the pulses. Thus, as long as the pulse amplitudes remain reasonably larger than the noise, the binary numbers, which represent the amplitudes of the sampled values of $v(t)$, can be recovered.

It should be noted that the use of PCM to transmit information at a high rate requires that the equipment have *wide bandwidth*. This is to be expected, because, as we know, a wide bandwidth is required for the fast transmission of pulses through networks.

Now let's consider some of the algebra associated with PCM. To do this, it will be helpful to begin with a block diagram of the circuitry required to generate a PCM signal. Such a diagram is shown in Fig. 329, with explanation as follows.

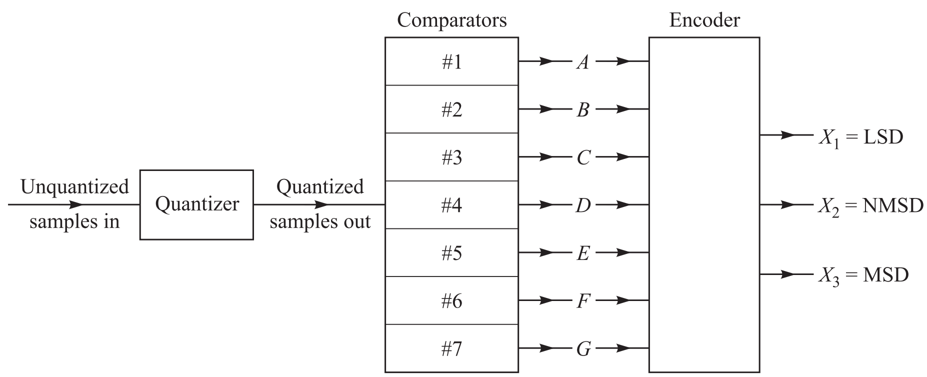


Fig. 329

First, in the figure, the "comparators" are special circuits that compare the amplitude V_q *of each incoming quantized pulse* with a reference voltage of V_{ref} volts, each of the comparators (seven in this case) having a *different value* of V_{ref}. The values of V_{ref} *increase by equal amounts* going from #1 through #7 in the figure.

For each comparator, whenever V_q exceeds the V_{ref} for that comparator, the output voltage for that comparator *becomes and remains* equal to "V volts" until the voltage of the input pulse falls below the value of V_{ref} for that comparator. The comparators are designed so that their output voltages will all have the same value of V volts (whenever V_q exceeds the V_{ref} for each particular comparator). Thus, during the presence of an input quantized pulse, the output of a given comparator is either 0 or V volts (but always 0 volts during the times between input pulses).

Thus suppose (for example) that, in the preceding figure, a particular input pulse to the bank of comparators has, say, an amplitude of 4 volts. In that case (given Fig. 327) comparators #1 through #4 *would all have equal outputs of "V volts,"* while #5 through #7 would have outputs of *zero volts*.

Thus, from the discussion following Fig. 327, *the output of the encoder*, *for this particular pulse*, would be the binary number "100"; that is, every input pulse of amplitude 4 volts would appear, at the output of the encoder, as the binary number 100 (as depicted in Fig. 328).

Again, the ADVANTAGE of this system is that, in this example, instead of having to accurately detect 8 different amplitude levels, all we need now is to detect the simple "present" or "not-present" condition of the pulses at the receiving end of the system.

To summarize, for the case of Fig. 329 the INPUT to the ENCODER consists of groups of equal-amplitude voltage pulses (from 0 to 7 pulses in each group), while the OUTPUT consists of 8 *binary numbers* of 3 digits each (000 to 111). In Fig. 329 the binary output digits are denoted by X_1 (the "least significant digit"), X_2 (the "next more significant digit"), and X_3 (the "most significant digit").

Our PROBLEM now is to find the "inner details" of an *encoder circuit* capable of converting the equal-amplitude input pulses into corresponding binary numbers in the output of Fig. 329.

To do this, let us begin with the following "truth table," in which V_q is the amplitude of the *quantized pulses* being fed into the bank of comparators in Fig. 329. From inspection of Fig. 329 we have that

V_q	A	B	C	D	E	F	G	MSD X_3	NMSD X_2	LSD X_1
0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	1
2	1	1	0	0	0	0	0	0	1	0
3	1	1	1	0	0	0	0	0	1	1
4	1	1	1	1	0	0	0	1	0	0
5	1	1	1	1	1	0	0	1	0	1
6	1	1	1	1	1	1	0	1	1	0
7	1	1	1	1	1	1	1	1	1	1

The job of the *encoder* network is to convert the input signals $A, B, C, \ldots G$ into 3-digit binary numbers, X_3, X_2, X_1, in accordance with the above truth table. In this regard, let us begin by writing an equation for X_1, this being an equation giving all the conditions for which $X_1 = 1$.

One way to do this is to begin with the basic "elemental" Boolean equation for X_1. In this particular case, however, it will be much easier to write the required equation from direct inspection of the truth table, as follows.

From inspection of the table note that it is *not* always true that $X_1 = 1$ when $A = 1$; instead, note that it *is* always true that $X_1 = 1$ whenever

$$A\,\bar{B} = 1 \quad \text{or when} \quad C\,\bar{D} = 1 \quad \text{or when} \quad E\,\bar{F} = 1 \quad \text{or when} \quad G = 1$$

thus the simplest possible equation for X_1 is

$$X_1 = 1 = A\,\bar{B} + C\,\bar{D} + E\,\bar{F} + G \tag{568}$$

Next, close inspection of the table shows that X_2 will be equal to "1" if the relationship

$$X_2 = B\,\bar{D} + F \tag{569}$$

is satisfied. This is true because, as the table shows, X_2 will be "1" when $B\,\bar{D} = 1$, regardless of the value of C. Also from the table, note that $X_2 = 1$ if $F = 1$, regardless of the value of G. Equation (569) shows how important it can be to make a close examination of a truth table.

Lastly, the table reveals that $X_3 = 1$ whenever $D = 1$, independent of the values of E, F, and G; hence our final equation is

$$X_3 = D \tag{570}$$

Thus, in Fig. 329, given the seven "1 or 0" input signals $(A, B, C, \ldots, E, F, G)$ to the encoder, we've found that, to generate the required output PCM signal X_3, X_2, X_1, the *circuitry of the encoder* must be such that

$$X_1 = A\,\bar{B} + C\,\bar{D} + E\,\bar{F} + G$$
$$X_2 = B\,\bar{D} + F$$
$$X_3 = D$$

With these equations as a guide we readily find that Fig. 330 will correctly serve as an encoder circuit to generate the required output signal X_3, X_2, X_1.

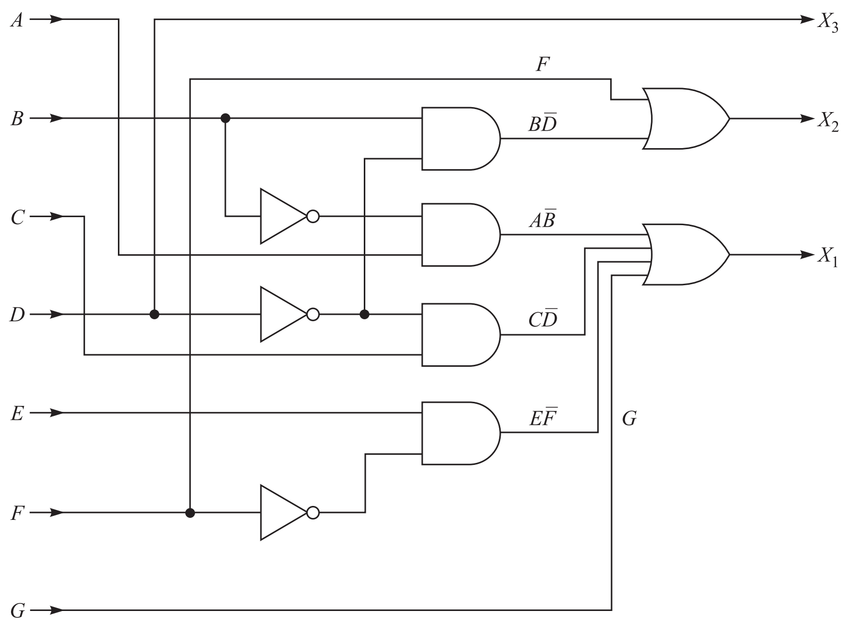


Fig. 330

Problem 294
In the above example, the quantized samples were restricted to 8 different voltage levels ($V_q = 0$ to $V_q = 7$ volts) which the encoder circuit then transformed into 8 different binary numbers (000 to 111).

Now rework the example, this time assuming the quantized pulses were allowed 12 different voltage levels ($V_q = 0$ to $V_q = 11$ volts), which the encoder circuit would then have to transform into twelve binary numbers of the form X_4, X_3, X_2, X_1. Your PROBLEM is to (a) write the truth table for the encoder, then (b) write the four simplest Boolean equations that would generate the required truth table.

13.2 Analog Signal in Sampled Form. Unit Impulse Notation

Let $v(t)$ denote the instantaneous continuous values of an analog signal. Then one way of GRAPHICALLY representing the SAMPLED FORM of $v(t)$ is shown in Fig. 331, where T is the uniform amount of *time between samples*.

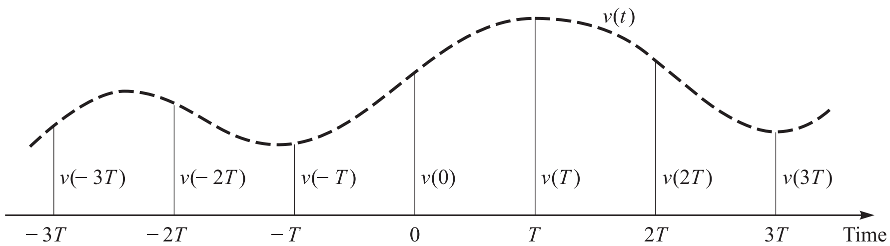


Fig. 331

In the figure, each distance line, drawn from the horizontal axis, represents the value of $v(t)$ at that particular sampling instant. Thus

$$v(-2T) = \text{value of } v(t) \text{ at } t = -2T$$

$$v(-T) = \text{value of } v(t) \text{ at } t = -T$$

$$v(0) = \text{value of } v(t) \text{ at } t = 0$$

$$v(T) = \text{value of } v(t) \text{ at } t = T$$

$$v(2T) = \text{value of } v(t) \text{ at } t = 2T$$

and so on, thus,

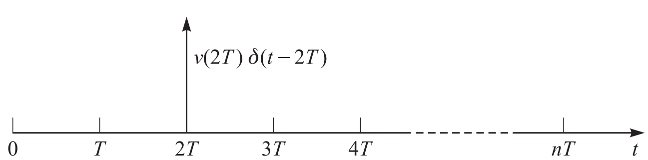
$$v(nT) = \text{value of } v(t) \text{ at any } n\text{th sampling instant}$$

Now let's assume the analog signal starts at some time $t = 0$, that is, $v(t) = 0$ for $t < 0$. For this condition, all the sample values to the left of the origin in the above figure are, of course, equal to zero. Then the equation for $v_s(t)$, *the sampled form of* $v(t)$, can be written in terms of time-delayed unit impulses;* thus

$$v_s(t) = v(0)\delta(t) + v(T)\delta(t - T) + v(2T)\delta(t - 2T) + \cdots + v(nT)\delta(t - nT) \tag{571}$$

where n is a positive integer including zero, $n = 0, 1, 2, 3, \ldots$.

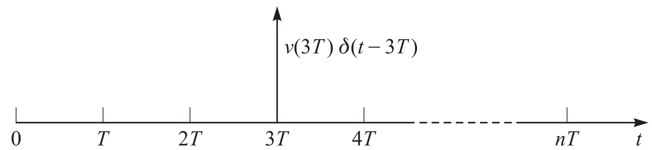
Because of the unit impulse factors, $v_s(t) = 0$ at all times EXCEPT at the instants $t = 0, t = T, t = 2T$, and so on, to $t = nT$. At each such instant *one* of the terms in eq. (571) will *not* be equal to zero; for example, at $t = 2T$, all the terms in the equation are equal to zero *except* the one term $v(2T)\delta(t - 2T)$. The graphical representation of eq. (571) at $t = 2T$ is shown below.



Immediately following the instant at $t = 2T$, all terms in eq. (571) *become and remain equal to zero* until the time $t = 3T$, at which time only the term $v(3T)\delta(t - 3T)$ is not equal to zero; the situation at $t = 3T$ is shown graphically in the following figure.

---

* See note 31, then note 32, in Appendix.



Let us note, now, that eq. (571) can also be written using the convenient "sigma" or "summation" notation, thus

$$v_s(t) = \sum_{n=0}^{n=\infty} v(nT)\delta(t - nT) \tag{572}$$

in which the symbol $\sum$ is the capital Greek letter "sigma." The sigma notation is read as "the SUM of all such terms from $n = 0$ to n equals infinity," where here "infinity" means that n, the number of terms, must be allowed to become infinitely great.

13.3 The *z*-Transform

Equation (572) is called a "time series," because the independent variable is the real quantity time, t.

It has been discovered, however, that the ALGEBRAIC WORK associated with the manipulation of sampled analog signals is much simplified if the situation in Fig. 331 is *mathematically defined* in terms of a COMPLEX VARIABLE z instead of the real variable time. This is done as follows.

Let us begin by arbitrarily writing down the following infinite series, in which $v(0), v(T), v(2T)$ and so on are the actual sampled values of an analog signal, and where z is the complex variable referred to above,

$$F(z) = v(0) + v(T)z^{-1} + v(2T)z^{-2} + \cdots + v(nT)z^{-n}$$

where

$$z = A\epsilon^{j\omega T}$$

in which A is a positive real constant, with the restriction that A be greater than 1 $(A > 1)$.

Note that now the independent variable is sinusoidal FREQUENCY, $\omega = 2\pi f$; thus we are now said to be working in the "frequency domain" instead of the time domain. The quantity $F(z)$ above is called the "z-transform" of the sequence of samples. As always, T is the constant time between samples.

Note now that the above can be neatly summarized by making use of the sigma notation, thus

$$\boxed{F(z) = \sum_{n=0}^{n=\infty} v(nT)z^{-n}} \tag{573}$$

where $z = A\epsilon^{j\omega T}$ and where $A > 1$.*

* Basically the value of A, in the summation of a given $v(nT)$ series, must be large enough so that, as $n \to \infty$, A^{-n} *decreases* faster than the sum of the $v(nT)$ series *increases*; this allows us to get a definite answer for $F(z)$, instead of the indeterminate answer "infinitely great."

As always in our discussions it's understood that $\epsilon = 2.71828\ldots$ (eq. (146) in Chap. 6). If we wish to make use of Euler's formula (also in Chap. 6) we can write that

$$z = A\epsilon^{j\omega T} = A(\cos\omega T + j\sin\omega T)$$

which emphasizes the fact that z is a complex number. Thus the real time function of eq. (572) is now being expressed in terms of a complex number z. The *advantage* is that the algebraic operations, if conducted in the complex plane, are simpler than if we were restricted to the use of real values only.

Next, let's raise the given equation, $z = A\epsilon^{j\omega T}$, to the "$-n$ power"; thus

$$z^{-n} = A^{-n}\epsilon^{-j\omega nT}$$

that is,

$$z^{-n} = \frac{1}{A^n}(\cos\omega nT - j\sin\omega nT)$$

in which n is a positive whole number (the number of terms in the series of eq. (573)).

Note, however, that in accordance with eq. (573) we must allow n to become "infinitely great" (which we indicate by writing $n \to \infty$). But A is a number greater than 1; thus $(1/A^n)$ becomes equal to *zero* when n becomes infinitely great. Hence inspection of the last equation for z^{-n}, above, shows that as n becomes infinitely great z^{-n} becomes equal to *zero*; thus

$$\lim_{n\to\infty} z^{-n} = 0 \tag{574}$$

which can be read as "The limiting value of z^{-n} as n becomes infinitely great is zero," or as "z^{-n} becomes equal to zero if n becomes infinitely great."

A comparison of eqs. (572) and (573) shows that $F(z)$ represents $v_s(t)$ in the complex plane. This means that a given $v_s(t)$ can be manipulated algebraically in terms of z instead of t, which is found to be a great advantage.

In practical applications we work in terms of $v(nT)$, the SEQUENCE OF SAMPLES generated by the sampling of an analog signal.

Thus, suppose we wish to find the result of applying a given $v(nT)$ to the input of a particular digital logic network. To do this, we must first express $v(nT)$ *in terms of* z, which we do by substituting the given $v(nT)$ into eq. (573). The expression for $F(z)$, thus found, is called the *z-transform* of the sequence $v(nT)$.

Let us, therefore, begin by finding the z-transforms of some of the most-used forms of $v(nT)$ encountered in practical work.

The simplest $v(nT)$ signal is called the "unit pulse," which consists of just *one* sample of *unit amplitude* at $t = 0$, as illustrated in conventional form in Fig. 332.

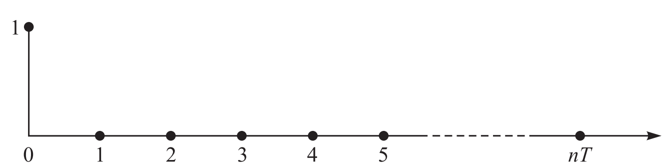


Fig. 332. The unit pulse.

We'll denote the unit pulse by $p(nT)$. Note that $p(nT) = 1$ for $n = 0$, but $p(nT) = 0$ for *all other n*. Thus, in eq. (573), for the case of $v(nT) = p(nT)$, we have $v(nT) = 1$ for $n = 0$ but $v(nT) = 0$ for all other n. Thus, substituting these values into eq. (573), we have that

$$F(z) = 1 \tag{575}$$

that is, *the z-transform* of the unit pulse is "1."

Note that, graphically, the sample values $v(nT)$ are plotted against nT, where n is the number of the sample counted from the $n = 0$ reference.

Next let's consider the very important ''unit-step'' sequence, in which *all* the samples have *unit amplitude*; that is, $v(nT) = 1$ for all values of nT, as shown in Fig. 333.

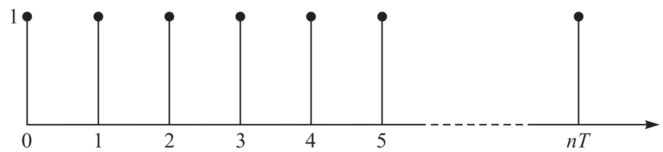


Fig. 333

We'll denote this ''unit-step sequence'' by $U(nT)$. Note that $U(nT) = 1$ for *all* values of n, including $n = 0$. Thus, substituting $v(nT) = U(nT) = 1$ into eq. (573) for all values of nT $(n = 0, 1, 2, 3, \ldots, n)$, we have that

$$F(z) = 1 + z^{-1} + z^{-2} + z^{-3} + \cdots + z^{-n} \quad \text{for } n \to \infty \tag{576}$$

The above is a valid answer, but can be put in a non-series or ''closed'' form as follows.

First, multiply both sides of the equation by $-z^{-1}$, then add the two equations together; doing this will show that (see problem 295 below)

$$F(z) = \sum_{n=0}^{n=\infty} \frac{1 - z^{-n}z^{-1}}{1 - z^{-1}} = \sum_{n=0}^{n=\infty} \frac{z - z^{-n}}{z - 1}$$

(after multiplying numerator and denominator of the first fraction by z).

But note, by eq. (574), that $z^{-n} \to 0$ as n becomes infinitely great. Thus the final fraction above has the limiting value of $z/(z-1)$, and hence *the z-transform of the unit-step sequence* is

$$F(z) = \frac{z}{z - 1} \tag{577}$$

Problem 295

Verify, to your satisfaction, that the suggested operation on eq. (576) does lead to the final result of eq. (577).

Another useful DT sequence* is the *linear rise* of Fig. 334, where T is the constant time between samples, as usual. The z-transform can be found as follows.

Note that Fig. 334 is the sampled form of the CT function $v(t) = t$, so that here

$$v(nT) = nT$$

as you can see from the figure. Thus, upon substituting nT in place of $v(nT)$ in eq. (573) we have that (remembering that T is constant)

$$F(z) = T \sum_{n=0}^{n=\infty} nz^{-n} = T(z^{-1} + 2z^{-2} + 3z^{-3} + 4z^{-4} + 5z^{-5} + \cdots) \tag{578}$$

The above infinite-series form of the answer can be put into closed form by carrying out the following algebraic manipulations.

* Notations such as $v(t), x(t)$, and so on denote continuous-time analog signals, while $v(nT), x(nT)$, and so on denote their sampled (''discrete'') form. For convenience we'll often abbreviate ''continuous time'' as CT and ''discrete time'' as DT.

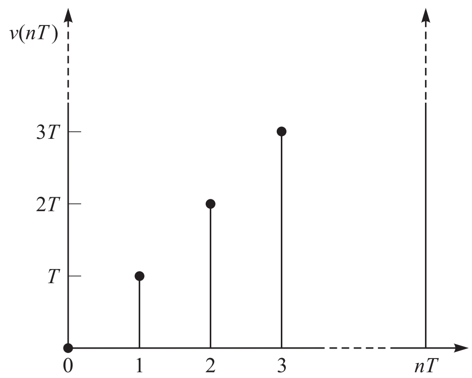


Fig. 334

Let us suppose that, after some experimentation, we decide to try multiplying both sides of eq. (578) by $-z$. Doing this, and noting that $-zz^{-1} = -z^0 = -1$, eq. (578) becomes

$$-zF(z) = T(-1 - 2z^{-1} - 3z^{-2} - 4z^{-3} - 5z^{-4} - \cdots)$$

Now take the algebraic sum of the above equation and eq. (578), thus getting

$$(1 - z)F(z) = T(-1 - z^{-1} - z^{-2} - z^{-3} - z^{-4} - \cdots)$$

which, upon multiplying both sides by -1, becomes

$$(z - 1)F(z) = T(1 + z^{-1} + z^{-2} + z^{-3} + z^{-4} + \cdots)$$

But note that, by eq. (576), the quantity *inside the parentheses* on the right-hand side is equal to the z-transform of the unit-step sequence, which, by eq. (577) equals $z/(z - 1)$. Thus, making this substitution into the last equation and then solving for $F(z)$, you can verify that the *z-transform of the linear-rise sequence* of Fig. 334 is equal to

$$F(z) = \frac{Tz}{(z - 1)^2} \tag{579}$$

Next consider the important continuous-time relationship $v(t) = \epsilon^{-bt}$, this being called the "negative exponential function," where b is a constant.

From Fig. 18-A (note 13 in Appendix), it follows that in *discrete time* the negative exponential function would appear in sampled form as in Fig. 335.

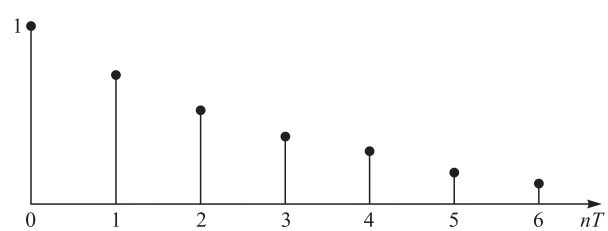


Fig. 335

Thus the CT function $v(t) = \epsilon^{-bt}$ becomes the DT sequence $v(nT) = \epsilon^{-bnT}$, and upon substituting this value into eq. (573) we have that

$$F(z) = \sum_{n=0}^{n=\infty} \epsilon^{-bnT} z^{-n} = \sum_{n=0}^{n=\infty} (\epsilon^{bT} z)^{-n} \tag{580}$$

and thus

$$F(z) = \left[1 + (\epsilon^{bT} z)^{-1} + (\epsilon^{bT} z)^{-2} + \cdots + (\epsilon^{bT} z)^{-n}\right] \quad \text{for } n \to \infty$$

Now note that, since ϵ^{bT} is *constant*, the quantity inside the brackets has the same *form* as eq. (576) which follows Fig. 333, except now we have $(\epsilon^{bT} z)$ in place of z; hence, all we need do is replace "z" with "$\epsilon^{bT} z$" in eq. (577), and we have that the z-transform of the *DT negative exponential sequence* is

$$F(z) = \frac{\epsilon^{bT} z}{\epsilon^{bT} z - 1} = \frac{z}{z - \epsilon^{-bT}} = \frac{z}{z - k} \tag{581}$$

after multiplying numerator and denominator of the first fraction by ϵ^{-bT}, then letting $k = \epsilon^{-bT}$.

All the foregoing results, plus several more, are summarized in Table 1.

Table 1. Some z-Transforms, $n = 0, 1, 2, 3, \ldots$, where $k = \epsilon^{-bT}$, b and T are constants

Number	DT signal $v(nT)$	z-transform of $v(nT) = F(z)$*
1.	unit pulse $= p(nT)$	1
2.	unit step $= U(nT)$	$\dfrac{z}{z-1}$
3.	exponential $= \epsilon^{-bnT}$	$\dfrac{z}{z-k}$
4.	linear rise $= nT$	$\dfrac{Tz}{(z-1)^2}$
5.	product of (3) and (4) $= nT\epsilon^{-bnT}$	$\dfrac{kTz}{(z-k)^2}$
6.	sine $= \sin \omega nT$	$\dfrac{z \sin \omega T}{z^2 - 2z\cos \omega T + 1}$
7.	cosine $= \cos \omega nT$	$\dfrac{z(z - \cos \omega T)z}{z^2 - 2z\cos \omega T + 1}$
8.	damped sine $= \epsilon^{-bnT} \sin \omega nT$	$\dfrac{k(\sin \omega T)z}{z^2 - 2k(\cos \omega T)z + k^2}$
9.	damped cosine $= \epsilon^{-bnT} \cos \omega nT$	$\dfrac{z(z - k\cos \omega T)}{z^2 - 2k(\cos \omega T)z + k^2}$

* We'll use "Z" to indicate that the z-transform of a DT sequence is to be taken; thus, $Zv(nT) = F(z)$, read as "the z-transform of a DT sequence $v(nT)$ is equal to F of Z."

Problem 296

This is an interesting and instructive example of the almost "magical" powers of Euler's formulas (eqs. (153) and (154) in Chap. 6).

Corresponding to the CT sinusoidal function $v(t) = \cos \omega t$, we have the DT sequence $v(nT) = \cos \omega nT$. By making use of Euler's formulas, and also eq. (580), see if you can derive item (7) in the table.

To continue on, in our work it will sometimes be necessary to deal with "time-delayed" DT signals.

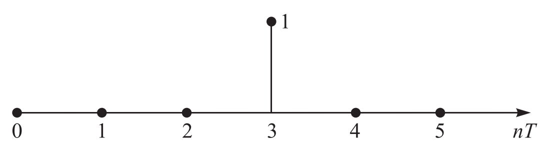
In this regard, recall that if $v(t)$ is a CT signal, then $v(t - T)$ is the same signal *but shifted T seconds to the right of* $v(t)$ *on the time axis* (note 31 in Appendix). Or, if we substitute $t + T$ in place of t we have $v(t + T)$, which is again the exact same form of signal as $v(t)$ but now shifted T seconds to the *left* of $v(t)$ on the time axis.

In the same way, if $v(nT)$ is a DT sequence, then $v(nT - kT)$ denotes the same sequence *but shifted k sample periods to the right of* $v(nT)$ on the nT axis; that is, $v(nT - kT)$ denotes $v(nT)$ *delayed by k sample periods* (delayed by kT seconds).

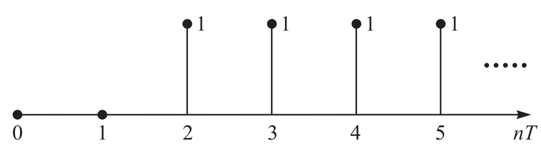
Likewise, $v(nT + kT)$ denotes the same DT sequence $v(nT)$, but now shifted k sample periods to the *left* on the nT axis from $v(nT)$; that is, $v(nT + kT)$ starts kT seconds *before* $v(nT)$ starts.

Consider now a few examples.

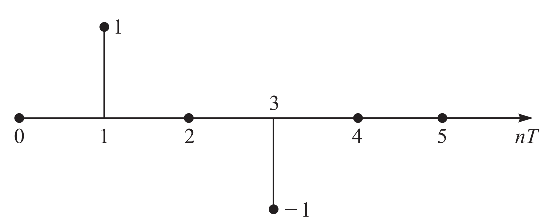
In Fig. 332 $p(nT)$ denotes a "unit pulse" for $nT = 0$; thus, for example, $p(nT - 3T)$ denotes the same pulse but now shifted 3 sample periods to the right of $nT = 0$ to $nT = 3$, as shown below.



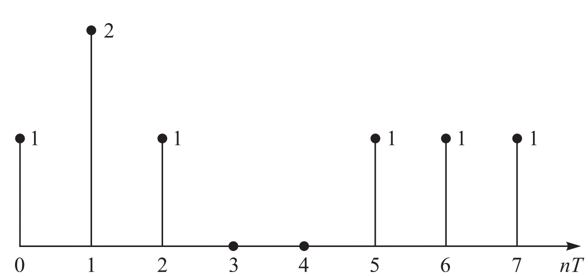
Or, consider the "unit-step sequence" $U(nT)$, illustrated in Fig. 333. The notation $U(nT - 2T)$ would, for example, denote the same sequence as in Fig. 333 but shifted or delayed 2 sample periods to the right; thus



As another example, the expression $v(nT) = p(nT - T) - p(nT - 3T)$ denotes an algebraic *sum* consisting of a positive unit pulse at $nT = 1$ and a negative or "negative-going" pulse at $nT = 3$; thus



As another example, the sum $v(nT) = U(nT) + p(nT - T) - p(nT - 3T) - p(nT - 4T)$ represents the modified unit-step sequence shown below,



In the above example note that $v(nT) = 2$ for $nT = 1$, but $v(nT) = 0$ for $nT = 3$ and $nT = 4$, because at these particular sampling instants the positive and negative sample values cancel each other out.

Problem 297
Show graphically the sequence represented by the DT equation

$$v(nT) = U(nT) + U(nT - T) - 4U(nT - 2T) + U(nT - 3T) + U(nT - 4T)$$

Problem 298
Show graphically the sequence represented by the DT equation

$$v(nT) = nT - (nT - 3T) - 3U(nT - 7T)$$

where nT is the "linear-rise" sequence of Fig. 334.

Now that we've dealt with the graphical representation of $v(nT - kT)$ in the "time domain," let's next consider the corresponding effect in the "z domain."

To do this, let us begin by referring back to the time-domain expression of eq. (572). Now suppose all of the sample values $v(nT)$ remain unchanged but are merely shifted kT *seconds to the right*; to indicate this, the "impulse factor" in eq. (572) would become

$$\delta(t - nT - kT) = \delta[t - (n + k)T]$$

which would indicate to us that, for the time-delayed case, n should be replaced by $n + k$ in the z-domain expression of eq. (573). This is true, and upon substituting $n + k$ in place of n, eq. (573) becomes, for the time-delayed case (where "del" stands for "delayed"),

$$F(z)_{\text{del}} = z^{-k} \sum_{n=0}^{n=\infty} v(nT) z^{-n}$$

because, by the laws of exponents, $z^{-(n+k)} = z^{-n} z^{-k}$, and, because k and z are independent of n, the factor z^{-k} can be put outside, to the left, of the summation sign, as shown.

But note that the quantity to the right of z^{-k} is, by eq. (573), equal to $F(z)$; thus the last equation above shows that

$$F(z)_{\text{del}} = z^{-k} F(z) \tag{582}$$

which says that if $F(z)$ is the z-transform of a given DT sequence, then $z^{-k}F(z)$ is the z-transform of the same sequence DELAYED BY k SAMPLE PERIODS.

In block diagrams of DT networks a delay is represented by a box labeled z^{-k}, which represents some kind of a device capable of producing a delay of k sample periods, as illustrated in Fig. 336.

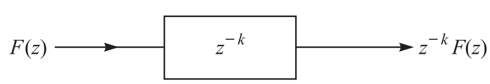


Fig. 336

In the above, the actual contents of the box could, for example, consist of a series connection of flip-flops, or perhaps an array of charge-coupled devices.

It will be a helpful reminder to conclude this section with a summary of certain algebraic operations that apply in the manipulation of z-transform equations. This is done in Table 2, with discussion below.

Table 2. Some Valid Operations in the z-Domain, where $Zv(nT) = F(z)$

1.	$Zav(nT) = aF(z)$, where a is constant
2.	$Z[v_1(nT) + v_2(nT) + \cdots] = F_1(z) + F_2(z) + \cdots$
3.	$Za^n v(nT) = F(z/a)$, where a is constant
4.	$Zv(nT - kT) = Zv(n-k)T = z^{-k}F(z)$, where $v(n-k)T = 0$ for $n < k$. This is the "time-delay" theorem.

First, in the table, item (1) says that if $F(z)$ is the z-transform of sequence $v(nT)$, and if a is constant, then the z-transform of $av(nT)$ is a times the z-transform of $v(nT)$.

Next, item (2) says that the z-transform of the *sum* of two or more sequences equals the sum of the transforms of the individual sequences. This property applies here because we are dealing with linear time-invariant conditions.

Next, item (3) can be established as follows. Since, by eq. (573),

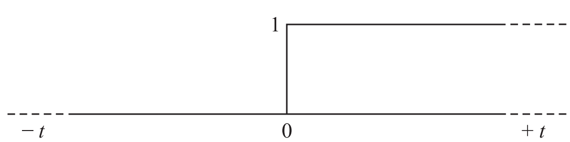
$$Zv(nT) = \sum_{n=0}^{n=\infty} v(nT)z^{-n} = F(z)$$

then,

$$Za^n v(nT) = \sum_{n=0}^{n=\infty} a^n v(nT)z^{-n} = \sum_{n=0}^{n=\infty} v(nT)(z/a)^{-n} = F(z/a)$$

that is, if $F(z)$ is the z-transform of a sequence $v(nT)$, and a is constant, then the z-transform of $a^n v(nT)$ is found by replacing z with z/a in the transform of $v(nT)$.

Problem 299
The notation $U(t)$ denotes a continuous-time (CT) function called the "unit-step function," defined as equal to 1 for all positive time (including $t = 0$) but equal to zero for all negative time, as illustrated below.



Given that a CT function $v(t) = 3U(t) - 20t$ is being sampled 100 times per second, what is the z-transform of $v(nT)$? (*Answer*: $F(z) = z(3z - 3.2)/(z-1)^2$)

Problem 300
Given $v(nT) = U(nT) + U(nT - T) + U(nT - 2T)$, find $F(z)$.

Problem 301
Find $F(z)$ for the $v(nT)$ of problem 298.

13.4 The Inverse z-Transform

We have rightly said that the solution to a DT problem can be simplified if the mathematical work is carried out in the z-domain. After a solution is obtained in the z-domain, the final step is to "inverse transform" the answer back into the time domain. In this section we'll consider how an answer in the z-domain can be inverse-transformed into the time domain. We begin our discussion as follows.

Going back to eq. (573), we see that a function of z, say $Y(z)$, can be expressed in the basic form

$$Y(z) = y(0) + y(T)z^{-1} + y(2T)z^{-2} + y(3T)z^{-3} + \cdots + y(nT)z^{-n} \tag{583}$$

where $y(0) =$ sampled value of $y(t)$ at $t = 0$, $y(T) =$ sampled value of $y(t)$ at $t = T$, $y(2T) =$ sampled value of $y(t)$ at $t = 2T$, and so on.

Then the corresponding answer to the above equation *in the time domain* is given by eq. (571); thus (now writing "y" instead of "v")

$$y_s(t) = y(0)\delta(t) + y(T)\delta(t - T) + y(2T)\delta(t - 2T) + \cdots + y(nT)\delta(t - nT) \tag{584}$$

It thus follows that *if* the answer to a DT problem comes out in the basic form of eq. (583), then there is *no difficulty* in expressing the answer in the time domain, because the sample values of $y(t)$, that is, $y(0), y(T), y(2T)$, and so on, required in eq. (584), appear directly as the *coefficients of the powers of z* in eq. (583).

The practical difficulty, however, is that a solution in the z-domain does not generally come out in the basic form of eq. (583); instead, the solution comes out in the form of the *ratio of two polynomials in z*, in which case the required sample values of $y(t)$ cannot be found by simple inspection of the solution (as would be the case if the solution were in the form of eq. (583). It is possible, however, to put $Y(z)$ directly into the form of eq. (583) by the use of ALGEBRAIC LONG DIVISION,* as the following examples will illustrate.

Example 1

Write the function $Y(z) = z/(z - 0.6)$ in the time domain, that is, in the form of eq. (584).

Solution

The first step is to put the given $Y(z)$ *into the form of eq. (583)*, which, as mentioned above, can be done by using algebraic long division. For the given $Y(z)$ the details are as follows.

$$
\begin{array}{r|l}
 & 1 + 0.6z^{-1} + 0.36z^{-2} + 0.216z^{-3} + 0.1296z^{-4} + \cdots \\
\hline
z - 0.6 & z \\
 & \underline{-z + 0.6} \\
 & \quad 0.6 \\
 & \quad \underline{-0.6 + 0.36z^{-1}} \\
 & \qquad +0.36z^{-1} \\
 & \qquad \underline{-0.36z^{-1} + 0.216z^{-2}} \\
 & \qquad\qquad +0.216z^{-2} \\
 & \qquad\qquad \underline{-0.216z^{-2} + 0.1296z^{-3}} \\
 & \qquad\qquad\qquad +0.1296z^{-3}
\end{array}
$$

* See note 33 in Appendix.

We continue in this manner until the change in the values of the coefficients becomes as small as we wish, at which point we terminate the series. In the above case, we can continue on for several more terms until we have the good approximation that

$$\frac{z}{z-0.6} = 1 + 0.6z^{-1} + 0.36z^{-2} + 0.216z^{-3} + 0.1296z^{-4} + 0.0778z^{-5} + 0.0467z^{-6}$$

which is now in the form of eq. (583), and thus, by direct comparison with (583), we see that the actual sampled values of $y(t)$, at $t = 0, T, 2T, \ldots, 6T$, are

$$\begin{array}{lll} y(0) = 1.0 & y(3T) = 0.216 & y(5T) = 0.0778 \\ y(T) = 0.6 & y(4T) = 0.1296 & y(6T) = 0.0467 \\ y(2T) = 0.36 & & \end{array}$$

Now, putting these values into eq. (584), we have that the equation for the sampled function in the *time domain* is

$$\begin{aligned} y_s(t) = \delta(t) &+ 0.6\delta(t-T) + 0.36\delta(t-2T) + 0.216\delta(t-3T) + 0.1296\delta(t-4T) \\ &+ 0.0778\delta(t-5T) + 0.0467\delta(t-6T) \end{aligned}$$

We can now use the above to pictorially show the form of $y(nT)$ versus nT in the manner previously described. The result is shown in Fig. 337.

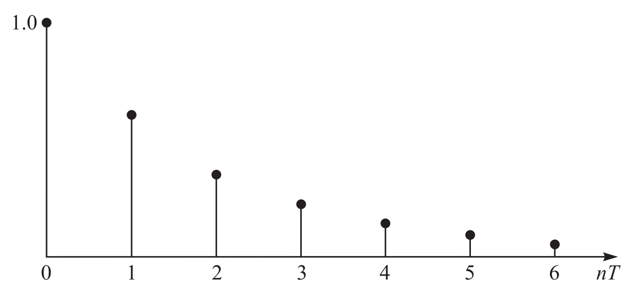


Fig. 337. $y(nT)$ **versus** nT**, for** $\frac{z}{z-0.6}$.

Example 2
Given

$$Y(z) = \frac{z}{z^2 - 1.9z + 1}$$

Find the first 8 sample values of $y(t)$; that is, $y(0), y(T), y(2T), \ldots, y(7T)$.

Solution
We first put $Y(z)$ into the form of eq. (583) by using algebraic long division, as follows, where, to save space, we've rounded off to two decimal places.

$$
\begin{array}{r|l}
 & z^{-1} + 1.90z^{-2} + 2.61z^{-3} + 3.06z^{-4} + 3.20z^{-5} + 3.02z^{-6} + 2.54z^{-7} \\
\hline
z^2 - 1.9z + 1 & z \\
 & \underline{-z + 1.90 - z^{-1}} \\
 & \quad +1.90 - z^{-1} \\
 & \quad \underline{-1.90 + 3.61z^{-1} - 1.90z^{-2}} \\
 & \qquad +2.61z^{-1} - 1.90z^{-2} \\
 & \qquad \underline{-2.61z^{-1} + 4.96z^{-2} - 2.61z^{-3}} \\
 & \qquad\quad +3.06z^{-2} - 2.61z^{-3} \\
 & \qquad\quad \underline{-3.06z^{-2} + 5.81z^{-3} - 3.06z^{-4}} \\
 & \qquad\qquad +3.20z^{-3} - 3.06z^{-4} \\
 & \qquad\qquad \underline{-3.20z^{-3} + 6.08z^{-4} - 3.20z^{-5}} \\
 & \qquad\qquad\quad +3.02z^{-4} - 3.20z^{-5} \\
 & \qquad\qquad\quad \underline{-3.02z^{-4} + 5.74z^{-5} - 3.02z^{-6}} \\
 & \qquad\qquad\qquad +2.54z^{-5} - 3.02z^{-6}
\end{array}
$$

from which we have that the first 7 terms of $Y(z)$ in series form are

$$Y(z) = z^{-1} + 1.90z^{-2} + 2.61z^{-3} + 3.06z^{-4} + 3.20z^{-5} + 3.02z^{-6} + 2.54z^{-7}$$

Now, comparing the above answer with the basic eq. (583), noting that in this case there is no $y(0)$ term, we see that the sample values are

$$
\begin{array}{lll}
y(0) = 0.00 & y(3T) = 2.61 & y(6T) = 3.02 \\
y(T) = 1.00 & y(4T) = 3.06 & y(7T) = 2.54 \\
y(2T) = 1.90 & y(5T) = 3.20 &
\end{array}
$$

Figure 338 is the result of plotting these values versus nT. Fig. 338 thus expresses the z-function $z/(z^2 - 1.9z + 1)$ in terms of time-domain sampled values of $y(t)$. (It so happens that Fig. 338 is a portion of a sine wave in sampled form.)

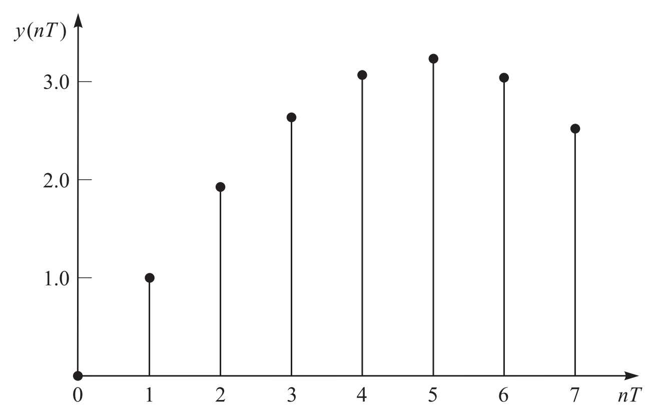


Fig. 338

As you might expect, long division is not the only way to convert a given z-transform into the equivalent time-domain sampled form.

Another method, for example, makes use of ''partial fractions,'' which is a procedure for writing a given fraction as the *sum* of a number of fractions in which the individual fractions are each simpler than the original given fraction. The inverse of each such simpler fraction can then be found by direct inspection of a table such as Table 1. The advantage of the partial fraction method is that the answer comes out in exact or ''closed'' form, while the long-division procedure comes out in series or ''open'' form.

Problem 302
Given that

$$Y(z) = \frac{1}{z - 0.5}$$

find, by means of long division, the time-domain values of $y(0)$, $y(T)$, $y(2T)$, $y(3T)$, $y(4T)$, and $y(5T)$.

Problem 303
Given that

$$Zv_s(t) = F(z) = \frac{z}{z^2 - 0.45}$$

find, by long division, the values of $v(nT)$ for $n = 0$ through $n = 9$.

13.5 The Discrete-Time Processor

The circuitry designed to manipulate DT signals is called a DT (discrete-time) processor, or ''digital processor'' if you wish.

In dealing with such processors it's common practice to associate the letter symbol x with the *input* DT sequence and the letter y with the resulting *output* DT sequence. The symbol (h) will be associated with the processor itself; that is, (h) will describe the digital circuitry needed to convert a given ''x'' input signal into a required ''y'' output signal.

Thus, if $x(nT)$ denotes an input sequence and $y(nT)$ the resulting output sequence, the situation can be represented in block diagram form as shown in Fig. 339, where the box with the (h) notation contains the required digital circuitry.

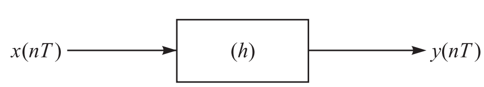


Fig. 339

(As a memory aid, note that ''input to output'' corresponds to the natural order of x to y in the alphabet.)

In the above, the quantity (h) is called the TRANSFER FUNCTION of the system, and is defined as being equal to the RATIO of the OUTPUT sequence $y(nT)$ to the INPUT sequence $x(nT)$; thus

$$\frac{y(nT)}{x(nT)} = (h) \tag{585}$$

that is,

$$y(nT) = (h)x(nT) \tag{586*}$$

Internally, a processor basically consists of the interconnections of THREE DIFFERENT TYPES or ‘‘blocks’’ of circuitry, these being ADDITION, MULTIPLICATION, and TIME-DELAY blocks. (Note that we’ve not mentioned ‘‘subtraction’’ separately because in binary operations subtraction can be performed by addition, as shown in section 12.1.)

Our purpose now is to investigate the manner in which these different ‘‘boxes’’ or ‘‘blocks’’ can be interconnected to form a processor capable of producing a desired output.

In doing this, each of the three basic operations is represented by a different schematic symbol as follows, beginning with the ‘‘addition’’ or ‘‘summer’’ symbol (Fig. 340).

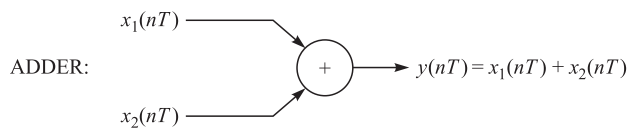


Fig. 340†

The purpose of an ‘‘adder’’ is clear from Fig. 340, in which there could, of course, be more than just two input lines. In the figure, it’s understood that the adder handles x_1 independently of the presence of x_2, and x_2 independently of the presence of x_1; that is, it’s understood that the ‘‘principle of superposition’’ applies to Fig. 340 and hence, as far as doing binary arithmetic is concerned, adders are ‘‘linear’’ devices.

Next, the block diagram symbol for MULTIPLICATION is shown in Fig. 341, in which an input signal $x(nT)$ is multiplied by a CONSTANT FACTOR a to produce an output signal a times the input signal, as shown.

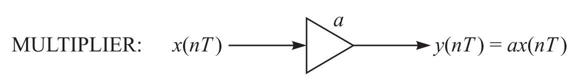


Fig. 341

The third basic requirement is that sample TIME DELAY must be provided for in a DT processor. Time delay is generated in multiples of T, where, as usual, T is the time between samples. The block diagram symbol for time delay has already been given in Fig. 336, but is repeated here in Fig. 342, where $k = 1, 2, 3, \ldots$, depending upon the number of sample periods a signal is to be delayed. (Note that whereas Fig. 336 is in the ‘‘z-domain,’’ Fig. 342 is in the ‘‘nT-domain’’; but it’s usual, in block diagrams, to use the same z^{-k} notation in both cases.)

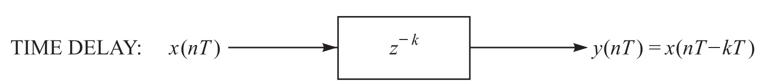


Fig. 342

* We sometimes like to say that (h) ‘‘operates on $x(nT)$’’ to produce $y(nT)$.

† To indicate that x_2 is to be *subtracted* from x_1 a minus sign would be placed alongside the input arrow for x_2 in Fig. 340, in which case $y(nT) = x_1(nT) - x_2(nT)$.

The above delay operation is said to be "time-invariant," because the BASIC INFORMATION carried by the signal is not altered by the time delay. That is, even though the output signal "lags" kT seconds behind the input signal, both signals still carry the same basic information.

The basic reason why time delay is used in digital processors is because the present or "now" output, $y(nT)$, of a processor is generated NOT ONLY by the present or "now" value of the input sequence $x(nT)$ but also by PAST VALUES of $x(nT)$ and, in some cases, by present and past values of the output sequence $y(nT)$.

We must remember that the job of a processor is to *electronically* carry out whatever MATHEMATICAL OPERATION is specified by the transfer function (h) in eq. (586). In general, the required mathematical operation will be too complicated to allow a processor to produce each "now" term of the output sequence, given only a single "now" value of the input sequence. Thus, to do its job, a processor requires *more information* than just each single "now" value. Fortunately, the additional information needed can be obtained by making use of PAST VALUES of the input sequence and, in some cases, also the present and past values of the output sequence. These things will be taken up in more detail in the next section.

13.6 The Form of, and Basic Equations for, a DT Processor

Electronic circuitry, both CT and DT, often makes use of FEEDBACK, which involves a condition in which a portion of the system OUTPUT signal is fed back into the INPUT of the system.

Such feedback, when properly used, can in some cases produce very beneficial results. In regard to DT processors, those that *do* use feedback are said to be *recursive*, while those that do *not* are *non-recursive*. Consider, now, examples of both types, beginning with Fig. 343.

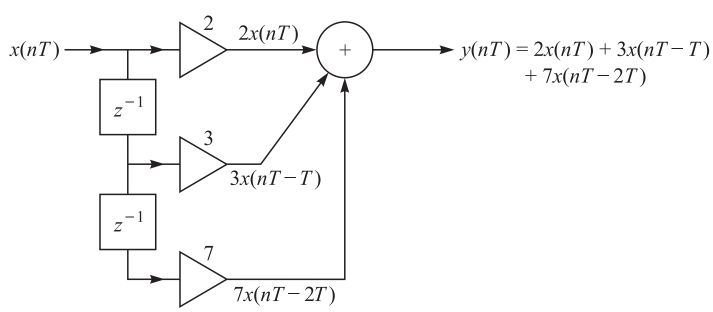


Fig. 343

In the figure note, first of all, that the output signal $y(nT)$ is *not* in any way "fed-back" into the system; thus Fig. 343 is an example of a *non-recursive* system.

Note also that the output sequence $y(nT)$ is the *sum* of the present or "now" value $2x(nT)$ and two PAST values, $3x(nT - T)$ and $7x(nT - 2T)$, which occurred T and $2T$ sample times ago relative to the "now" time of nT seconds. (In the "time-delay boxes" the exponent "-1" means "unit time delay," that is, a time delay of T seconds.)

In regard to interpreting a figure such as Fig. 343, we should note that, while a notation such as $x(nT)$ really denotes an entire SEQUENCE of values, $n = 0, 1, 2, 3, \ldots$, we can, for convenience, think of $v(nT)$ as denoting some particular sample value existing at a time nT.

One more point to note is that Fig. 343 is classified as a "second-order" processor, because it uses two delays.

Next, as a second example, consider Fig. 344. Note that the output sequence $y(nT)$ is fed back to the input adder after being delayed 1 sample period. Thus Fig. 344 is a simple form of *recursive* DT processor (of "first order," because only one delay is used).

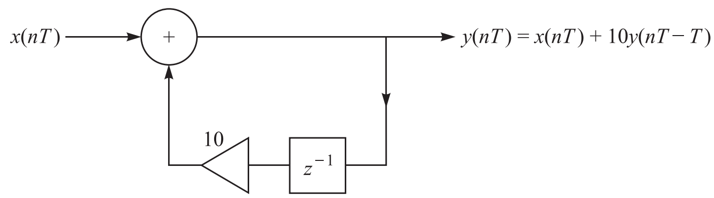


Fig. 344

Figures 343 and 344 are simple examples of DT processors and we need not, at this point in our study, worry about how they are put to work. It is important, however, that you understand the meaning of the notation.

In general, processors will consist of a combination of non-recursive and recursive types; for example, the combination of Figs. 343 and 344 to make a single processor gives Fig. 345, in which

$$y(nT) = y'(nT) + 10y(nT - T)$$

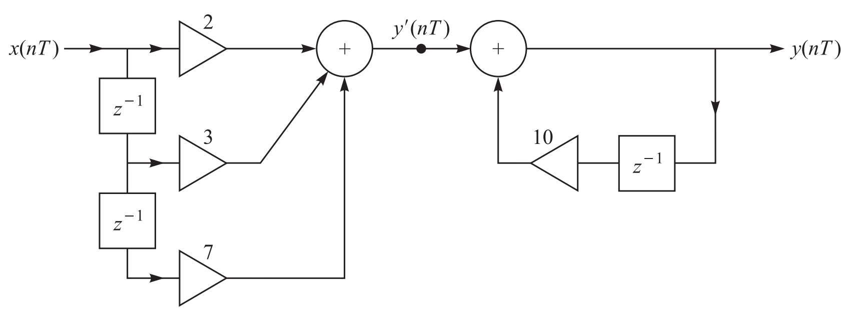


Fig. 345

and thus

$$y(nT) = 2x(nT) + 3x(nT - T) + 7x(nT - 2T) + 10y(nT - T)$$

With the foregoing in mind, let's agree now to adopt the notation illustrated in Fig. 346 for the general form of a DT processor. (There is a non-recursive part to the LEFT of point R and a recursive part to the RIGHT of point R, where R is just a reference point.)

Note, first, that the value of the output at the point R (the output of the first adder) is

$$R(nT) = \sum_{k=0}^{k=p} b_k x(nT - kT)$$

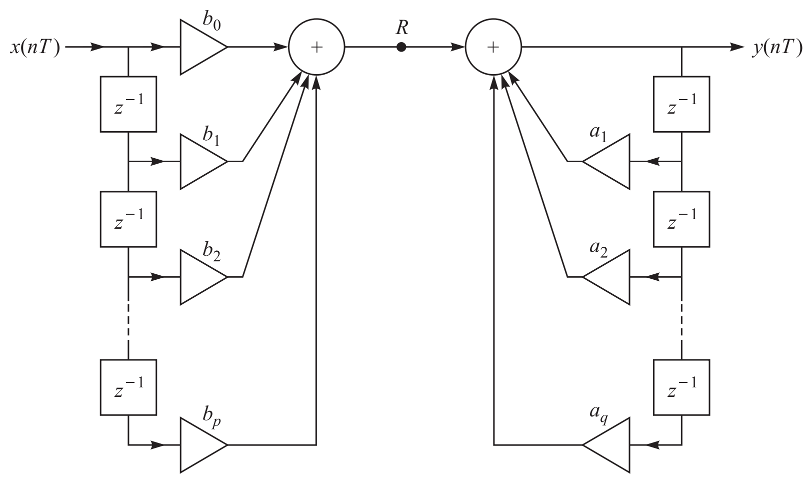


Fig. 346

Thus we have that the final output value of the processor is

$$y(nT) = \sum_{k=0}^{k=p} b_k x(nT - kT) + \sum_{k=1}^{k=q} a_k y(nT - kT) \tag{587}$$

or, written out in expanded form, eq. (587) becomes

$$\begin{aligned} y(nT) &= b_0 x(nT) + b_1 x(nT - T) + b_2 x(nT - 2T) + \cdots + b_p x(nT - pT) \\ &\quad + a_1 y(nT - T) + a_2 y(nT - 2T) + \cdots + a_q y(nT - qT) \end{aligned} \tag{588}$$

Equations (587) and (588) are called "difference equations" because of the presence of the "past history" forms having the difference notation $(nT - kT)$.

Let us now operate on eqs. (587) and (588) in such a way as to express things in terms of the *z-transform*. This can be done as follows.

First note that, in the above equations, $y(nT)$ and $x(nT)$ denote sampled values of $y(t)$ and $x(t)$ *for any particular value of n we might be interested in*. However, to bring the z-transform into the picture we must *summate* the values of $y(nT)$ and $x(nT)$ over the entire range of n that there is, for $n = 0$ to $n \to \infty$ (in accordance with the basic definition of eq. (573) in section 13.3). Therefore (so that we can apply eq. (573)) let us multiply both sides of eq. (588) by z^{-n} and then summate from $n = 0$ to $n \to \infty$; doing this, eq. (588) becomes

$$\begin{aligned} \sum y(nT)z^{-n} &= b_0 \sum x(nT)z^{-n} + b_1 \sum x(nT - T)z^{-n} + \cdots + b_p \sum x(nT - pT)z^{-n} \\ &\quad + a_1 \sum y(nT - T)z^{-n} + a_2 \sum y(nT - 2T)z^{-n} + \cdots + a_q \sum y(nT - qT)z^{-n} \end{aligned}$$

where all summations are understood to be from $n = 0$ to $n = \infty$. Now note that (with same summation from $n = 0$ to $n = \infty$) by eq. (573)

$$\sum y(nT)z^{-n} = Y(z)$$

and by item (4), Table 2

$$\sum y(nT - kT)z^{-n} = Y(z)z^{-k}$$

and the same, of course, for $x(nT)$ in place of $y(nT)$. Now apply the last two relations to the preceding equation; doing this gives the relationship

$$Y(z) = b_0 X(z) + b_1 X(z)z^{-1} + \cdots + b_p X(z)z^{-p} + a_1 Y(z)z^{-1} + a_2 Y(z)z^{-2} + \cdots + a_q Y(z)z^{-q}$$

Now, on the right-hand side of the last equation, factor out $X(z)$ and $Y(z)$ and then solve for the ratio of OUTPUT TO INPUT, that is, $Y(z)/X(z)$. Doing this, and using the notation of eq. (585), you should find that

$$\frac{Y(z)}{X(z)} = H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_p z^{-p}}{1 - (a_1 z^{-1} + a_2 z^{-2} + \cdots + a_q z^{-q})} \tag{589}$$

Equations (588) and (589) are the basic digital processor equations, (588) being in the "time" or "nT" domain and (589) in the z-domain. As we continue, we'll gradually begin to see how they can be applied.

As a final note, remember that the *internal* operations in a digital processor are performed using *binary arithmetic*; that is, internally the information is manipulated in the form of strings of 1's and 0's.

All pulses representing "1" have the same amplitude; a "multiplier" unit does *not* multiply the *actual* amplitudes of the pulses; thus the output of the multiplier unit in Fig. 341 is a binary number "a" times the binary number at the input to the unit. The resulting effect is, of course, the same as if the actual amplitudes of the pulses had been multiplied by a. It might seem as if such a procedure would be too time consuming, but we must remember that a digital processor is capable of performing many many millions of operations per second.

Problem 304
What is the basic equation for the transfer function, in the z-domain, for the processor in Fig. 345?

Problem 305
Write the basic equation, in the z-domain, for the transfer function of a purely non-recursive digital processor.

Problem 306
In the following, $x(nT)$ and $y(nT)$ denote, respectively, DT input and output sequences of a DT processor in the time domain.

(a) $y(nT) = x(nT) + 6x(nT - T)$
(b) $y(nT) = 2x(nT) + 5x(nT - T) + 10x(nT - 2T)$
(c) $y(nT) = 6x(nT) - 8x(nT - T) + 7y(nT - T)$

In each case, state whether the processor is a "non-recursive" or a "recursive" type.

Problem 307
Given that the output $y(nT)$ of a certain processor with $x(nT)$ input is

$$y(nT) = 4x(nT) + 7x(nT - T) - 5x(nT - 2T) + 9x(nT - 3T)$$

(a) Sketch the block diagram of the circuit layout of the processor.
(b) For this processor, $H(z) =$

Let's conclude this section as follows. The "transfer function" of a DT processor has already been defined for the "nT-domain" (eq. (586) in section 13.5). Now, in the z-domain we have, by eq. (589), that

$$Y(z) = H(z)X(z) \tag{590}$$

so that, correspondingly, $H(z)$ is now the "transfer function" expressed in the z-domain (instead of the nT-domain as in eq. (586)). Thus the block diagram form of Fig. 339 in section 13.5 now becomes Fig. 347.

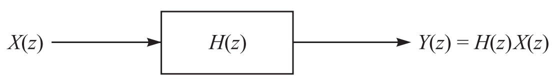


Fig. 347

An interesting fact can be discovered as follows. Suppose the INPUT signal to a processor is the UNIT PULSE $p(nT)$ of Fig. 332. In such a case $X(z) = 1$, because the z-transform of the unit pulse is "1"; and hence, for this particular case, eq. (590) becomes $Y(z) = H(z)$.

> The transfer function $H(z)$ of a linear DT system is equal to the response of the system to UNIT-PULSE INPUT. For this reason the terms "transfer function," "unit-pulse response," and "pulse transfer function" are all used interchangeably.

13.7 Stability and Instability. Poles and Zeros

It is possible for a recursive DT processor to become *unstable* under certain conditions. The *desired* condition of "stability" and the *undesired* condition of "instability" can be defined in general terms as follows.

Let a *momentary signal*, such as the "unit pulse" of Fig. 332, be applied to the input of a recursive DT processor. If the OUTPUT of the processor "dies out" and becomes zero as time increases, the processor is STABLE; if, however, the output does *not* become zero as time increases, the processor is UNSTABLE.

Let us take, as an example to illustrate the basic possibilities, the simple recursive processor shown in Fig. 348.

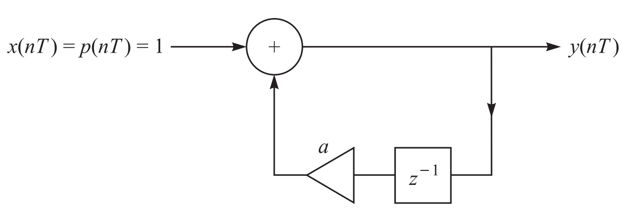


Fig. 348

For brevity here we'll denote the unit pulse by "1," as shown in the figure. Note that the above is the same as Fig. 344, except that now the input is given to be the unit pulse, $x(nT) = p(nT) = 1$, and the multiplier constant is denoted by a. It is the VALUE OF

THE MULTIPLIER a that determines whether the processor of Fig. 348 is stable or unstable. The explanation is as follows.

First remember that, in this case, $x(nT) = 0$ for all time EXCEPT at $t = 0$, when $x(nT) = 1$. Thus the *output* at the instant $t = 0$ is also 1.

Then, following this, *after T seconds has elapsed*, a time-delayed signal, $(1)(a) = a$, arrives at the input to the adder; thus, at $t = T$, the output is a.

Then, *after another T seconds has elapsed*, a time-delayed signal, now equal to $(a)(a) = a^2$, arrives at the input to the adder, so that, at $t = 2T$, the output is a^2.

Then, *after another T seconds has elapsed*, a time-delayed signal, now equal to a^3, is fed back to the input to the adder, so that at $t = 3T$ the output is a^3.

Continuing on in this way we see that, at any integral multiple of time T, $t = nT$, the output is equal to a^n; that is, in Fig. 348, $y(nT) = a^n$. Thus the nature of the output sequence in Fig. 348 *depends upon the value of the multiplier constant a*. Let us discuss this in more detail, as follows.

First, note that our definition of "stability or instability," as given above, could also be stated in the following equivalent way.

Let a single unit pulse $p(nT)$ be applied to the input of a (recursive) processor, and let $y(nT)$ denote the value of the output at any time nT seconds later.

Now let L denote the value that $y(nT)$ would approach if n were allowed to become "infinitely great" (denoted by writing $n \to \infty$, or loosely, for convenience, simply as "$n = \infty$").

We can then say that, in general, a processor is *stable* if $L = 0$, but is *unstable* if L is *not* equal to zero. Let us apply this principle to Fig. 348, where we've already found that

$$y(nT) = a^n \tag{591}$$

Thus, applying the above rule to the particular processor of Fig. 348, we have that

$$\lim_{n=\infty} y(nT) = \lim_{n=\infty} a^n = L \tag{592}$$

It's apparent that, in this case, the value of L will depend upon the *value of the constant multiplier a*. As a matter of fact, after some thought we realize that we must consider three separate possibilities for the value of a, these being the cases for a GREATER than 1, a EQUAL to 1, and a LESS than 1. Let us consider each of the three cases as follows.

Case I. ($a > 1$): Here the values of the output samples, a^n, theoretically become "infinitely great" for $n = \infty$. Thus in this case L is certainly *not* equal to zero, so that Fig. 348 is *unstable* for $a > 1$. This is illustrated in Fig. 349.

Case II. ($a = 1$): Since $1^n = 1$ we have that $L = 1$, and thus Fig. 348 is *unstable* for $a = 1$, as illustrated in Fig. 350.

Case III. ($a < 1$): The integral *power* of a number less than 1 is *less* than the given number; for instance, if $a = 1/3$, then $a^2 = 1/9, a^3 = 1/27$, and so on. Thus for this case L becomes equal to zero ($L = 0$) as n becomes infinitely great; hence Fig. 348 is *stable* for $a < 1$, as illustrated in Fig. 351.

In regard to Fig. 349, the output of an actual processor could not, of course, become "infinitely great"; instead, in such a case the output would cease increasing and stall when the maximum holding capacity of the digital circuits was exceeded.

Now let's return to Fig. 348 and this time apply the *z-transform*. Let us begin by writing down the nT-domain equation for the figure, which, since it's given that $x(nT) = p(nT)$, is

$$y(nT) = p(nT) + ay(nT - T)$$

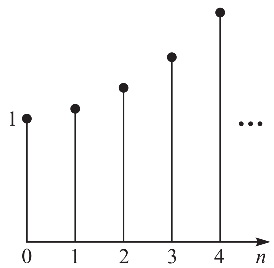


Fig. 349

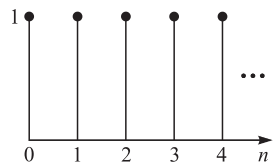


Fig. 350

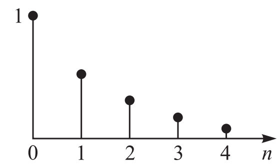


Fig. 351

Now take the z-transform of the above equation. Doing this, and remembering that the z-transform of the unit pulse is "1", we have

$$Y(z) = 1 + aY(z)z^{-1}$$

thus,

$$Y(z) = \frac{1}{1 - az^{-1}}$$

or, upon multiplying numerator and denominator by z, we have

$$Y(z) = \frac{z}{z - a} \tag{593}$$

Now consider the following. We've agreed to define the "stability or instability" of a DT processor in terms of its response $Y(z)$ to *unit pulse input*.

For unit pulse input, however, it is true that $Y(z) = H(z)$, as was pointed out at the end of section 13.6. Hence (always assuming unit pulse input) we can just as well write $H(z)$ instead of $Y(z)$. With this understood, eq. (593) becomes

$$H(z) = \frac{z}{z - a} \tag{594}$$

The behavior of a DT processor is thus determined by the nature of its transfer function $H(z)$. This is done by observing what are called the "zeros" and "poles" of $H(z)$ for a given processor. The zeros and poles of $H(z)$ are defined in accordance with eq. (589) in section 13.6, as follows.

To begin, let $N(z)$ and $D(z)$ denote the numerator and denominator of eq. (589); thus

$$H(z) = \frac{N(z)}{D(z)} \tag{595}$$

We now define that a ZERO of $H(z)$ is any value of z for which $H(z)$ is equal to *zero*, that is, for which $H(z) = 0$.

Or, since $H(z) = 0$ if $N(z) = 0$, this is the same as saying that a "zero" of $H(z)$ is any value of z for which $N(z) = 0$. Thus, by eq. (589), a *zero* of $H(z)$ is any value of z that satisfies the equation

$$b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_p z^{-p} = 0$$

Or, if we wish, upon multiplying through by z^p this becomes

$$b_0 z^p + b_1 z^{p-1} + b_2 z^{p-2} + \cdots + b_p = 0 \tag{596}$$

Equation (596) is the equation that must be solved to find the "zeros" of $H(z)$ for a proposed DT processor. Inspection of Fig. 346 shows that the b coefficients pertain to the *non-recursive* portion of a DT processor, consisting of $1 + p$ multipliers and p delays.

We next define that a POLE of $H(z)$ is any value of z for which $H(z)$ becomes "infinitely great," that is, for $H(z) = \infty$.

Or, since $H(z)$ becomes infinitely great when the *denominator* of eq. (589) becomes equal to *zero*, this is the same as saying that a "pole" of $H(z)$ is any value for which $D(z) = 0$. Thus, by eq. (589), a pole of $H(z)$ is any value of z that satisfies the equation

$$1 - (a_1 z^{-1} + a_2 z^{-2} + \cdots + a_q z^{-q}) = 0$$

Or, if we wish, upon multiplying through by z^q this becomes

$$z^q - a_1 z^{q-1} - a_2 z^{q-2} - \cdots - a_q = 0 \tag{597}$$

Equation 597 is the equation that must be solved to find the "poles" of $H(z)$ for a proposed DT processor. Inspection of Fig. 346 shows that the a coefficients pertain to the *recursive* portion of a DT processor, consisting of q multipliers and delays.

In working with DT processors a knowledge of the locations of both the zeros and poles of $H(z)$ is very important, this being especially true for the locations of the POLES of $H(z)$.

This is because it is the *locations of the poles of H(z)* that determines whether a proposed processor will be *stable or unstable*. This is summarized in the famous dictum that

> A DT processor is STABLE only if ALL THE POLES of the transfer function $H(z)$ lie WITHIN THE UNIT CIRCLE on the complex z plane.

To gain some understanding of why this is true, note first that, as used here, the "unit circle" on the z-plane is defined to be a circle of *unit radius* with *center* at the origin of the complex plane.

That is, a unit circle is defined here to be a circle of radius $r = 1$, center at origin of the complex plane, the unit radius being at any variable angle θ (theta), "positive" θ being measured in the counter-clockwise sense from the positive real axis. This is illustrated in Figs. 352 and 353.

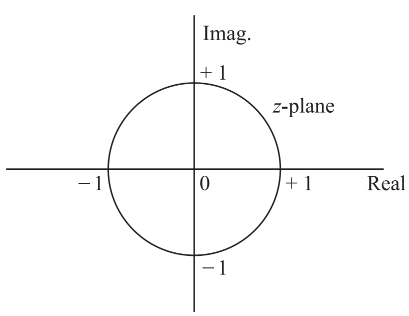


Fig. 352. Unit circle.

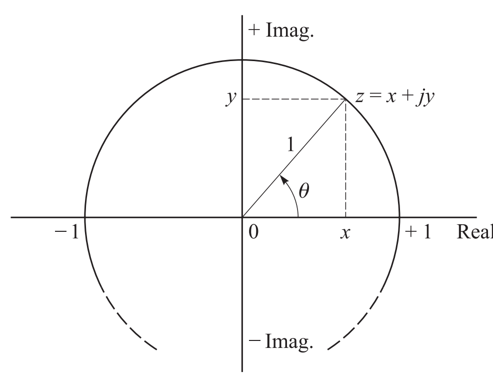


Fig. 353

As can be seen from the enlarged version in Fig. 353, the equation of the unit circle can be written in either the ‘‘rectangular’’ or the ‘‘exponential’’ form; thus

$$z = x + jy = (\cos\theta + j\sin\theta) = \epsilon^{j\theta} \tag{598}$$

Now consider the basic recursive equation, eq. (597). Note, carefully, that the notation used in eq. (597) refers to the notation used in the recursive portion of Fig. 346.

Now consider the fundamental ‘‘first-order’’ processor of Fig. 348. With regard to eq. (597), note that here $q = 1$, with all the a coefficients equal to *zero* EXCEPT for a_1; thus, for Fig. 348, eq. (597) becomes

$$z - a_1 = 0$$

showing that a first-order recursive processor has just ONE POLE located at $z = a_1$. However, as we already know, Fig. 348 is *stable* only if a_1 is *less than 1*. Thus it is true that a first-order recursive processor is stable only if the solution to eq. (597) lies within the unit circle.

What we have just found, for the basic first-order recursive processor, can be extended to ANY ORDER of such processors; that is, any recursive processor is stable only if all the poles of its transfer function $H(z)$ lie within the unit circle on the z-plane.

The poles of a given $H(z)$ may be all real numbers, or all complex numbers, or a combination of real and complex numbers. HOWEVER, it’s an important fact that COMPLEX POLES can occur only in the form of CONJUGATE PAIRS of complex numbers; thus, if $c + jd$ is a pole of $H(z)$, then $c - jd$ is *also* a pole, and vice versa. By way of an explanation, let’s first consider the case of a *second-order* recursive processor, as follows.

For the second-order case, in eq. (597) we would have $q = 2$, with all a coefficients equal to zero EXCEPT for a_1 and a_2; thus, for a second-order recursive processor eq. (597) becomes

$$z^2 - a_1 z - a_2 = 0 \tag{599}$$

which, if we wish, can be put in the *factored* form

$$(z - g)(z - h) = 0 \tag{600}$$

which, in this form, shows that a second-order recursive processor will have TWO POLES, one at $z = g$, the other at $z = h$. If, now, we multiply as indicated in eq. (600), we have that

$$z^2 - (g + h)z + gh = 0$$

and upon comparing this last result with eq. (599) we see that

$$\left.\begin{aligned} a_1 &= (g + h) \\ a_2 &= -gh \end{aligned}\right\} \tag{601}$$

Now, in an actual processor the a_1 and a_2 coefficients will always be *real* numbers. However, even though a_1 and a_2 are themselves real numbers, the two poles, g and h, can be either two *real* numbers or two CONJUGATE complex numbers. To show this, suppose that g and h are two conjugate complex numbers, $g = c + jd$ and $h = c - jd$. Then, by eq. (601), we have

$$a_1 = (c + jd) + (c - jd) = 2c, \text{ a real number}$$

and

$$a_2 = -(c + jd)(c - jd) = -(c^2 + d^2), \text{ a real number.}$$

Thus we have the important fact that a *second-order* recursive processor will always possess TWO POLES, g and h, in which either

(a) g and h are both real numbers, or

(b) g and h are conjugate complex numbers.

However, regardless of whether we have case (a) or case (b) for a given processor, for stability *both* poles must lie within the unit circle on the z-plane.

Now let's consider a *third-order* recursive processor (meaning the use of three delays). For this case $q = 3$ in eq. (597), which thus, for this case, becomes

$$z^3 - a_1 z^2 - a_2 z - a_3 = 0$$

which, theoretically, it's always possible to *factor* into the form

$$(z - f)(z - g)(z - h) = 0$$

showing that a third-order recursive processor possesses THREE POLES, denoted here by f, g, and h. Since complex poles can exist only in conjugate form it follows that the possibilities for a third-order (recursive) processor are

(a) three real poles, or

(b) one real and one pair of conjugate poles.

Again, for stability it's necessary that all three poles lie within the unit circle.

To continue on, it's a fundamental fact that any algebraic equation of the form of eq. (597), in which the highest power of the unknown, z, is an integer, q, can always be factored into the form

$$(z - h_1)(z - h_2)(z - h_3) \cdots (z - h_q) = 0 \qquad (602)$$

which clearly shows that any such equation has "q solutions"* which we're denoting here by $h_1, h_2, h_3, \ldots h_q$.

You may have noticed that, so far, we've not said much about eq. (596). As we already know, solutions to eq. (596) are called "zeros" because these are the values of z for which $H(z) = 0$. For now, however, let us just note that certain procedures do exist that require the use of both the zeros and the poles of a processor.

Problem 308
Write the equation for finding the poles of a fourth-order recursive processor and list the possible combinations of real and complex poles that might exist.

Problem 309
Repeat problem 308 for a fifth-order (recursive) processor.

Note: The following problems will call for a certain amount of factoring. In some cases this can be done by direct inspection, while in other cases you may wish to review the "standard quadratic formula" found in note 1 in the Appendix.

* In the general language of algebra the "solutions" are said to be the "roots" of the equation, but in the specific application here it's customary to call the roots "poles."

Problem 310
Find the zeros and poles, given the transfer function

$$H(z) = \frac{4z + 9}{z(z - 9)(z^2 + 5z + 7)}$$

Problem 311
The following are transfer functions for certain DT processors. Determine, in each case, whether the processor is stable or unstable.

(a) $H(z) = \dfrac{z}{(z - 0.46)(z - 0.22)}$

(b) $H(z) = \dfrac{z^2}{(z - 0.61)(z^2 - 1.6z + 0.48)}$

(c) $H(z) = \dfrac{z^{-1} - 1.2z^{-2}}{1 - 1.37z^{-1} + 0.305z^{-2}}$

13.8 Structure of DT Processors

Let us, for ready reference here, begin by redrawing Fig. 346 as Fig. 354, where, as before, $H_1(z)$ = the non-recursive part of the network and $H_2(z)$ = the recursive part of the network.

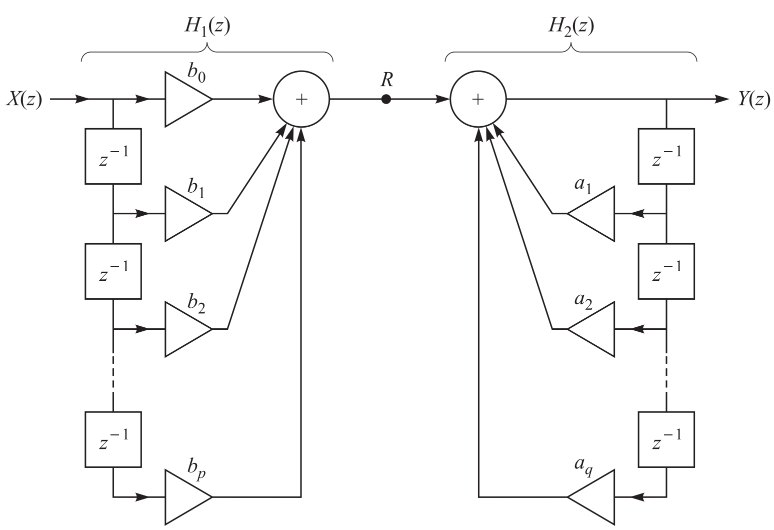


Fig. 354

Figure 354 is generally referred to as the *Direct Form I* DT structure. Now, while Fig. 354 is the basic form of DT processor, it can be considerably simplified into what is called the *Direct Form II* structure. To do this, note that in Fig. 354 the two divisions of the network are really connected in series ("cascade"), and thus the basic network equation, eq. (590) in section 13.6, can be written in either of the forms

$$Y(z) = [H_1(z)H_2(z)]X(z) \tag{603}$$

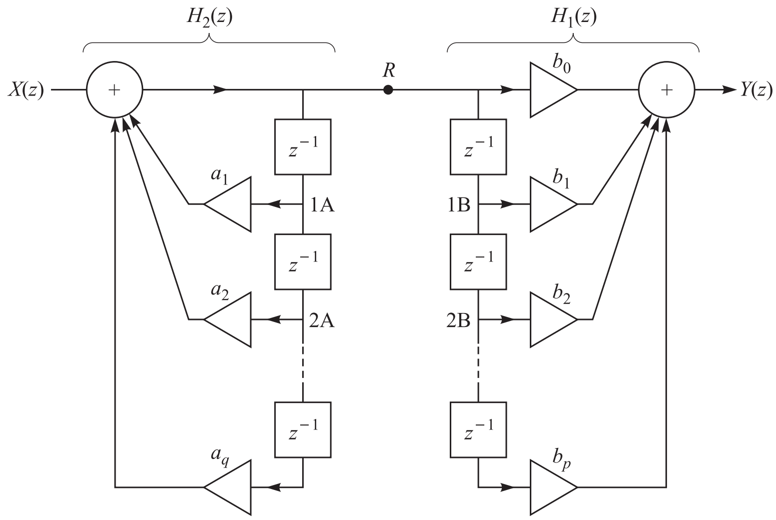


Fig. 355

or

$$Y(z) = [H_2(z)H_1(z)]X(z) \tag{604}$$

Thus it basically makes no difference whether we feed the input signal $X(z)$ to the $H_1(z)$ section first, as in Fig. 354, or to the $H_2(z)$ section first, as in Fig. 355.

Now note, in Fig. 355, that the signals at points 1A and 1B are equal; likewise, the signals at points 2A and 2B are equal, and so on down the ladder of delays.

Thus Fig. 355 simplifies into Fig. 356, which is called the *Direct Form II* DT structure. (Figure 356 happens to be drawn for the case where q is greater than p.)

Now let us verify that the transfer function for Fig. 356 is equal to the transfer function for Fig. 354. That is, let us verify that $H(z)$ for Fig. 356 is the same as eq. (589) in section

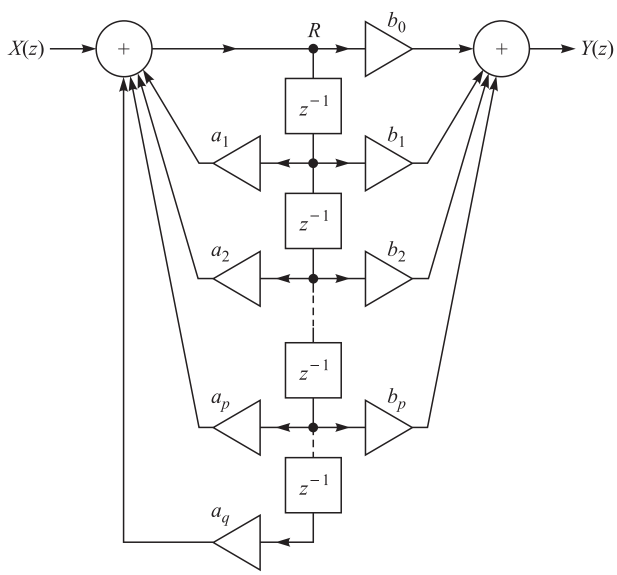


Fig. 356. Direct Form II.

13.6. To do this, let $R(z)$ denote the signal at point R in Fig. 356; then inspection of the figure shows that

$$R(z) = X(z) + a_1 R(z)z^{-1} + a_2 R(z)z^{-2} + \cdots + a_q R(z)z^{-q}$$

that is,

$$R(z) = X(z) + (a_1 z^{-1} + a_2 z^{-2} + \cdots + a_q z^{-q})R(z)$$

which, upon solving for $X(z)$, gives

$$R(z)[1 - (a_1 z^{-1} + a_2 z^{-2} + \cdots + a_q z^{-q})] = X(z) \tag{605}$$

Next, from inspection of Fig. 356 we see that the output signal $Y(z)$ is equal to

$$Y(z) = b_0 R(z) + b_1 R(z)z^{-1} + b_2 R(z)z^{-2} + \cdots + b_p R(z)z^{-p}$$

that is,

$$Y(z) = (b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_p z^{-p})R(z)$$

Now solve the last equation for $R(z)$, and substitute the result in place of $R(z)$ in eq. (605). Upon doing this, and then solving for the ratio $Y(z)/X(z)$, you should find that, for Fig. 356,

$$\frac{Y(z)}{X(z)} = H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_p z^{-p}}{1 - (a_1 z^{-1} + a_2 z^{-2} + \cdots + a_q z^{-q})} \tag{606}$$

Note that this is exactly the same as eq. (589), in which the b multipliers are in the non-recursive part of the network while the a multipliers are in the recursive or feedback part of the network. It should be noted that the ‘‘Direct Form II’’ configuration requires the *least number of delays* needed to perform a given processing task.

The Direct Forms I and II are basic to the construction of DT processors. It is found, however, that practical difficulties sometimes arise if we attempt to do too much with just one, single, high-order processor. (The ‘‘order’’ of a processor is the number of delays required to make the processor accomplish its task.) Thus it’s often better to use a number of lower-order processors, *connected either in cascade or parallel*, instead of a single high-order processor. The procedure is based upon the assumed linearity of binary-type circuits;* that is, on the assumption that if $H_1(z), H_2(z), \ldots, H_n(z)$ denote the n transfer functions of n individual processors, then

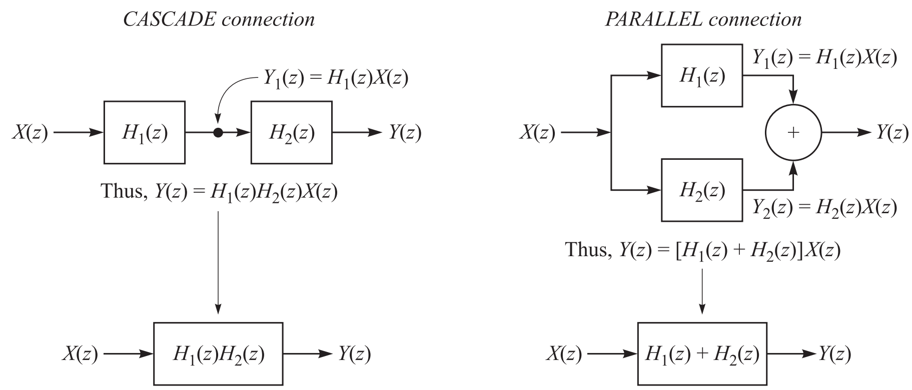
$$H(z) = H_1(z)H_2(z)\ldots H_n(z) = \text{transfer function of the } \textit{cascade} \text{ (series) connection of the } n \text{ processors}$$

and

$$H(z) = H_1(z) + H_2(z) + \cdots + H_n(z) = \text{transfer function of the } \textit{parallel} \text{ connection of the } n \text{ processors}$$

This is illustrated below in block diagram form for the case of $n = 2$, that is, for the case of two separate processors. We’re using the basic notation of eq. (590) in section 13.6.

* If all we need to do is distinguish between 0’s and 1’s the circuits need not be linear. This is another advantage of encoding a CT signal in DT form.



The same block diagram notation can, of course, be extended to any number of cascaded or paralleled stages.

Problem 312
Write the equation

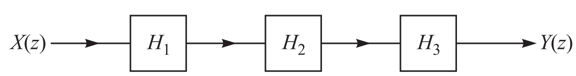
$$H(z) = \frac{2z^3 + 4z^2 + z}{z^3 - 5z^2 + 6z + 9}$$

in the form of eq. (606).

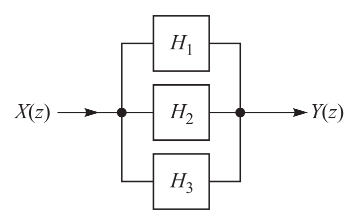
Problem 313
The individual unit pulse responses of three DT processors are as follows:

$$H_1(z) = \frac{z}{z - 0.2} \qquad H_2(z) = \frac{z}{z - 0.4} \qquad H_3(z) = \frac{z}{z^2 - 0.8z + 0.15}$$

If the three processors are connected as in the block diagram below, find the unit pulse response $H(z)$ for the entire connection. Write final answer in form of eq. (606).



Problem 314
Repeat problem 313 if the same three processors are connected in the configuration shown to the right.



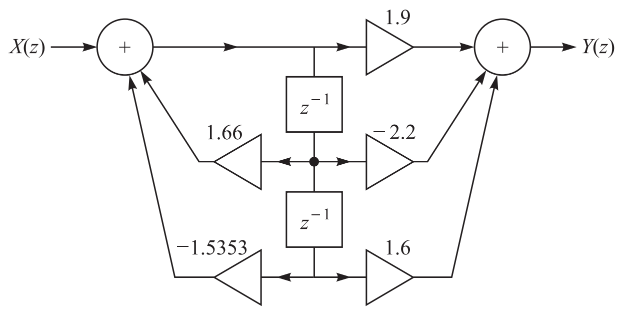
Problem 315
Suppose a single unit pulse of voltage $p(nT)$ is applied to the input of the processor of problem 313. Find the output of the network $3T$ seconds later.

(*Answer*: 1.25 volts)

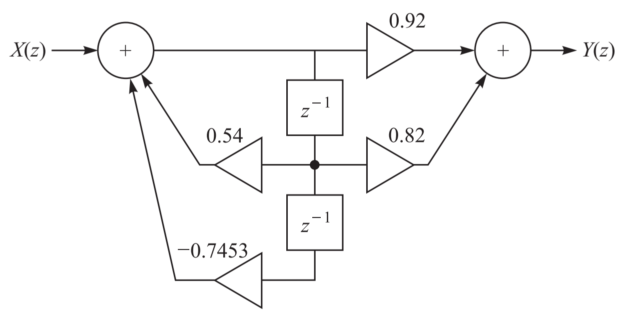
Problem 316
Sketch the block diagram of the Direct Form II processor (Fig. 356) having the transfer function

$$H(z) = \frac{2z^3 + 1.3z^2 + 0.9z}{z^3 - 2.2z^2 - 1.5z + 0.75}$$

Problem 317
Determine whether the following Direct Form II processor is stable or unstable.



Problem 318
Determine whether the Direct Form II processor to the right is stable or unstable.



13.9 Digital Filters; The Basic Algebra

Signals can be, and are, studied in both the time domain and the frequency domain.

In the time domain we principally study the manner in which the amplitude and time delay of a signal change with time.

In the frequency domain we study the amplitudes and phase shifts of the different sinusoidal frequency components present in a signal (the "fundamental" and "harmonics," as outlined in note 18 in the Appendix). The result of such a study is summarized in terms of the "frequency response characteristic" of a system.

An "electric FILTER" is a network designed to have a SPECIFIC FORM of frequency response characteristic. Thus we have "low-pass" filters, "high-pass" filters, and so on. You'll recall that it's convenient to display such results graphically, in the form of frequency response curves.

By "frequency" it's always understood that we mean the frequencies of the *sinusoidal* component waves of a signal. As always, frequency in radians per second is denoted by omega, ω, while frequency in cycles per second (hertz) is denoted by f, in which, as you know, $\omega = 2\pi f$.

In practical work it's convenient to express results in terms of FREQUENCY of SINUSOIDAL waves of voltage and current. Thus in the time domain we work with the basic equations $v = \sin \omega t$ and $v = \cos \omega t$.

We have found, however, that the ALGEBRAIC work can be greatly simplified if we are not restricted to the use of real numbers only but are allowed to work in the total "complex plane" of all numbers. This is because the algebraic operations of multiplication, division, roots, and powers are easier to express and carry out in the complex plane than in the ordinary x, y plane of real number pairs. (This is fundamentally true because complex numbers can be expressed in EXPONENTIAL FORM to which the "laws of exponents" can be applied to simplify the foregoing mentioned algebraic operations.)

But now let us get on with the subject of "digital filters." In the study of such filters much use is made of the "unit circle in the complex plane." Let us therefore begin by returning briefly to Figs. 352 and 353 and eq. (598) in section 13.7. In that discussion we show that the EQUATION of the unit circle IN THE COMPLEX PLANE is given by

$$z = \epsilon^{j\theta} = (\cos\theta + j\sin\theta)$$

which we can also regard as being the basic EQUATION of a SINUSOIDAL WAVE of peak value 1 when expressed in complex numbers. Note that the real and imaginary parts of the equation each separately represent sinusoidal waves if sketched on the ordinary x, y plane of real numbers.

Now let us note that the meaning of "frequency response," as applied to digital networks, is basically the same as that defined for analog networks,* except that in the digital case the input test signal will be a SAMPLED sinusoidal wave instead of a continuous-time sinusoidal wave as in the analog case.

In this regard consider Fig. 357, which shows, in block diagram form, the test setup required to experimentally determine the frequency response of a digital filter (abbreviated DF). Note that the setup uses both analog-to-digital (A/D) an digital-to-analog (D/A) circuits.

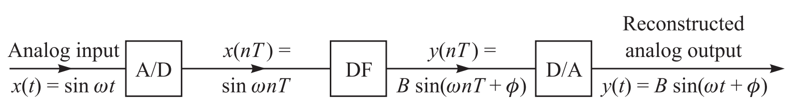


Fig. 357

Note that the actual input to the DF is a *sampled* sinusoidal wave of unit amplitude, $x(nT) = \sin \omega nT$. Then note that, corresponding to $x(nT)$, the *output* of the DF is $y(nT) = B \sin(\omega nT + \phi)$, where B and ϕ are *both functions of frequency* ω. Thus the frequency response of a DF is expressed in terms of the manner in which B and ϕ vary with the frequency ω. (If we wish, B and ϕ can be measured at the output of the D/A circuit, as shown.) Consider now the following.

In eq. (573), section 13.3, the variable z is defined to be a complex quantity $z = A\epsilon^{j\omega T}$, where A is a real variable assumed to be always large enough to assure that the value of $F(z)$ does not become "infinitely great" as $n \to \infty$.

Now consider the SPECIAL CASE that arises for the specific value $A = 1$. For this particular condition we have

$$z = \epsilon^{j\omega T} = \cos\omega T + j\sin\omega T$$

which is the equation, in the complex plane, for a purely SINUSOIDAL WAVE of UNIT AMPLITUDE. Therefore the equation for a SAMPLED unit-amplitude sinusoidal wave at the INPUT *to the* DF in Fig. 357 is

$$X(z) = (\cos\omega nT + j\sin\omega nT) = \epsilon^{j\omega nT} = z^n$$

* See section 9.5, Fig. 186, and discussion following the figure.

which produces $Y(z)$ at the OUTPUT *of the* DF, a corresponding sampled sinusoidal wave of amplitude B and phase angle ϕ; thus

$$Y(z) = B\epsilon^{j(\omega nT+\phi)} = (B\epsilon^{j\phi})(\epsilon^{j\omega nT}) = (B\epsilon^{j\phi})z^n$$

in which we made use of the basic relationship $\epsilon^{a+b} = \epsilon^a\epsilon^b$. Thus, letting $H(z)$ denote the TRANSFER FUNCTION of the filter, we have that

$$H(z) = Y(z)/X(z) = B\epsilon^{j\phi} \tag{607}$$

from which we may infer, correctly, the IMPORTANT PRACTICAL RULE that

> If $H(z)$ is the transfer function of a DT processor, then, to find the steady-state sinusoidal frequency response of the processor, simply set $z = \epsilon^{j\omega T}$ in $H(z)$.

Equation (607) thus gives the AMPLITUDE and PHASE ANGLE, B and ϕ, of the *output sinusoidal wave* produced by the input unit reference signal sin ωt in the test setup of Fig. 357.

In connection with the above procedures it's convenient to use the following notation. Let f = input analog frequency (Hz), or $2\pi f = \omega$ = input analog frequency (rad/sec); now let f_s = constant "sampling frequency," where (see eq. (91) in Chap. 5) $f_s T = 1$. Thus,

$$T = \frac{1}{f_s} = \frac{2\pi}{2\pi f_s} = \frac{2\pi}{\omega_s}$$

hence,

$$\omega T = \omega(2\pi/\omega_s) = 2\pi r$$

where $r = (\omega/\omega_s)$; that is, r = *ratio* of analog frequency to fixed sampling frequency.

Thus we can, if we wish, state the foregoing rule as: "To find the sinusoidal frequency response of a DT processor, set $z = \epsilon^{j2\pi r}$ in $H(z)$, where $r = (\omega/\omega_s)$."

In our work we'll basically make use of eq. (606) in section 13.8. With this in mind, note that if a processor is purely "non-recursive" (does not use feedback)* then all of the a coefficients in eq. (606) are equal to *zero*, and thus, for a non-recursive case, eq. (606) becomes

$$H(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_p z^{-p} \tag{608}$$

As the equation shows, an actual calculation of $H(z)$ will require finding the *sum* of a number of complex numbers. We recall, however, that this will require that the numbers first be put into the "rectangular" $(a + jb)$ form; thus we'll need to make use of the famous Euler relationship

$$\epsilon^{\pm j2\pi r} = (\cos 2\pi r \pm j \sin 2\pi r) \tag{609}$$

Another important point to note is as follows. In Fig. 357 let ω_h denote the HIGHEST FREQUENCY COMPONENT of importance present in the input analog signal. Then, in order to satisfy the basic requirement of the "sampling theorem," it has to be true that $\omega_s = 2\omega_h$; thus, for $\omega = \omega_h$ and $\omega_s = 2\omega_h$, we have that

$$r = (\omega_h/2\omega_h) = 0.5$$

* "Non-recursive" filters are called "finite impulse response" filters, abbreviated FIR.

which means that, FOR A GIVEN SAMPLING FREQUENCY ω_s and variable analog frequency ω, the ratio $r = (\omega/\omega_s)$ should not exceed the value of $r = 0.5 = 1/2$. Thus the FREQUENCY RESPONSE of a digital filter generally needs to be calculated only over the range of $r = 0$ to $r = 0.5$.

Example
Show that the DT processor in Fig. 358 will serve as a "low-pass" filter.

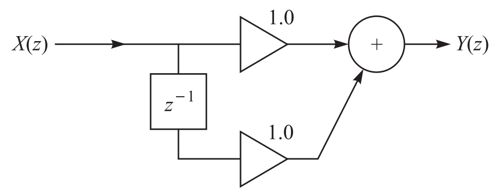


Fig. 358

Solution
Note that this is a non-recursive filter in which $b_0 = 1.0$ and $b_1 = 1.0$ (see left-hand side of Fig. 354 in section 13.8). Thus, upon substituting these values into eq. (608) we FIRST have that

$$H(z) = 1.0 + z^{-1}$$

Next, since we wish to find the steady-state "sinusoidal frequency response" of the given processor, we now make the substitution $z = \epsilon^{j2\pi r}$ into the above $H(z)$. If we do this, changing the notation $H(z)$ to $H(r)$ and making use of Euler's formula, the above equation for $H(z)$ becomes

$$H(r) = 1.0 + \epsilon^{-j2\pi r} = 1.0 + \cos 2\pi r - j \sin 2\pi r \tag{610}$$

Thus $H(r)$ is now in the rectangular form $H(r) = a + jb = |H(r)|\underline{/\phi}$, where

$$|H(r)| = \sqrt{a^2 + b^2} \tag{611}$$

and

$$\phi = \arctan(b/a) \tag{612}$$

where, in this particular case, we have $a = (1.0 + \cos 2\pi r)$ and $b = -\sin 2\pi r$. Hence we have, here, by eq. (611),

$$|H(r)| = \sqrt{(1 + \cos 2\pi r)^2 + \sin^2 2\pi r} = \sqrt{2(1 + \cos 2\pi r)}$$

in which we made use of the identity $\sin^2 x + \cos^2 x = 1$ (from problem 64), and also, by eq. (612),

$$\phi = -\arctan[\sin 2\pi r/(1 + \cos 2\pi r)]$$

because $\arctan(-x) = -\arctan x$.

Note: Euler's formula, in the form $\epsilon^{\pm jx} = \cos x \pm j \sin x$, is valid for x in *radians*, where degrees = (radians)$(180/\pi)$. Thus, if we wish to work in degrees, we would write $\sin 360r$ and $\cos 360r$.

It will be informative, now, to show graphically how $|H(r)|$ and ϕ change with changing values of r, for the case of Fig. 358.

To do this we begin with the following ‘‘table of values’’ which, as you can verify, was found by making use of the above formulas for $|H(r)|$ and ϕ.

r	$\lvert H(r)\rvert$	ϕ°
0.0	2.00	0.00
0.1	1.90	−18.00
0.2	1.62	−36.00
0.3	1.18	−54.00
0.4	0.62	−72.00
0.5	0.00	−90.00*

The above results are shown graphically in Figs. 359 and 360.

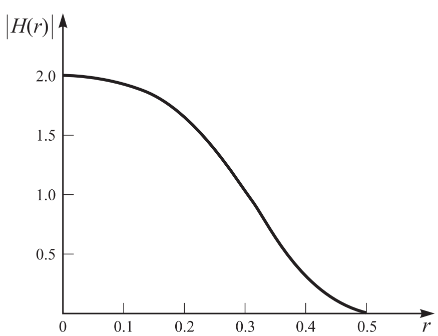


Fig. 359

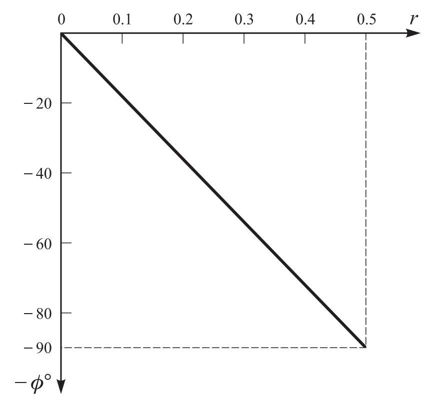


Fig. 360

First, Fig. 359 shows that, while the circuit of Fig. 358 *is* a basic form of low-pass filter, it is very broad in its action, not possessing the ‘‘sharp cutoff’’ characteristic we generally would want such a filter to have. This is understandable, because Fig. 358 is the most basic type of low-pass digital filter.

Next, Fig. 360 shows that the *ratio* of ‘‘phase shift to frequency’’ is constant, which means that the non-recursive (FIR) filter of Fig. 358 has constant time delay and thus produces no time-delay distortion (note 20 in Appendix). The fact that FIR filters have constant time delay is an advantage in certain applications. Another advantage of FIR filters is that they are always stable (because of the absence of feedback).

Thus our ALGEBRA has let us to the fact that the OUTPUT of Fig. 358 depends upon the FREQUENCY ω of the ORIGINAL ANALOG SIGNAL; the details of WHY this happens can be explained as follows. Let us begin by recalling, with the aid of Fig. 361, the action of the UNIT DELAY circuit in Fig. 358.

In Fig. 361, A represents the sampled value of an analog signal at a time t, while B is the *same value as A* but *delayed T seconds from A*, as shown.

Thus timewise, *at* $t+T$, the output of the time delay unit is the ‘‘past value’’ of the analog signal at time t.

---

* Direct substitution of $r = 1/2$ into the equation for ϕ gives $\phi = -\arctan(0/0)$, where 0/0 has, itself, an indeterminate value. If, however, using your calculator, you successively find the values of ϕ for, say, $r = 0.496, 0.497, 0.498$, and so on, it will be apparent that ϕ approaches the limiting value of -90° for $r = 1/2$.

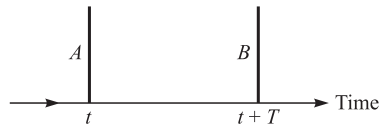


Fig. 361

Thus, from inspection of Fig. 358, we see that *the value of each* OUTPUT SAMPLE is the algebraic SUM of the ‘‘present’’ value of the input sample and the ‘‘past value’’ of the previous sample.

With this in mind suppose, first, that the INPUT to Fig. 358 is the sampled form of a relatively LOW FREQUENCY analog signal, such as is illustrated in Fig. 362 which shows a portion of one cycle.

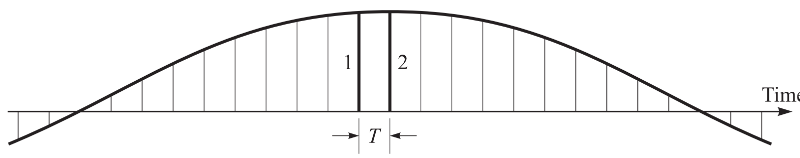


Fig. 362

In Fig. 362 we label, for purposes of explanation, just TWO of the many samples that would be fed into the circuit of Fig. 358. We’ve labeled the chosen samples ‘‘1’’ and ‘‘2,’’ as shown. The two successive samples are the usual T seconds apart. (In the above, for the indicated T, about 20 samples would be taken in each half cycle.) The point we wish to make in connection with Fig. 362 is as follows.

Note that (in accordance with the discussion given with Fig. 361) when *sample* 2 appears at the INPUT to the circuit of Fig. 358 the algebraic SUM of samples 1 and 2 appears at the OUTPUT of the circuit. The important point to note is that the *sum of samples 1 and 2* is very closely *twice the value of sample 2 alone**; this is because the amplitude of a low-frequency signal changes slowly with time. This is the basic reason why the output magnitude, in Fig. 359, is highest for low values of ω ($r = \omega/\omega_s$).

Now suppose the input to Fig. 358 is the sampled form of a relatively HIGH FREQUENCY analog signal, such as is illustrated in Fig. 363 (same peak value, same T, as in Fig. 362).

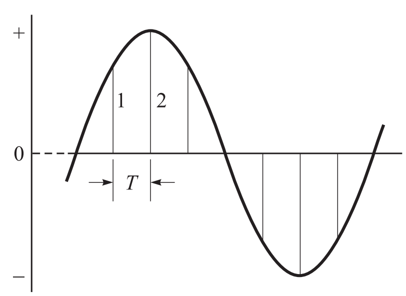


Fig. 363

In Fig. 363 note that we cannot say that the algebraic sum of samples 1 and 2 is ‘‘very closely’’ equal to twice the value of 2 alone (as we could for the low-frequency case of

* As you can from Fig. 362, this is closely true for samples taken near the maximum value of the analog wave, while increasingly less so for samples taken farther from the maximum.

Fig. 362), and this effect is increasingly great between samples taken farther from the maximum value. This is why the magnitude of the output decreases with increasing analog frequency (Fig. 359), and it is due to the fact that the amplitude of the high-frequency analog wave changes rapidly with time.

Now consider Fig. 364. The figure is based upon the requirements of the basic ‘‘sampling theorem’’ (section 13.1), and illustrates the condition in which r has the maximum allowable value of $r = 0.5$. In this condition the HIGHEST ANALOG FREQUENCY that can be sampled and allowed to enter a digital filter system is equal to ‘‘one-half the sampling frequency’’ the system uses. Thus *two samples per cycle* must be taken of the highest permitted frequency component of the analog signal; this is the condition shown in Fig. 364. From the figure, note that the algebraic sum of samples 1 and 2 is always *zero*; that is, the output of the filter of Fig. 358 is zero for the highest allowable frequency analog signal (as we see in Fig. 359).

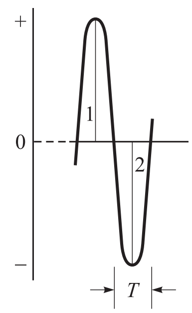


Fig. 364

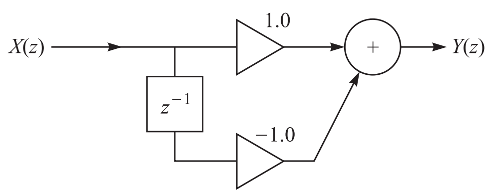
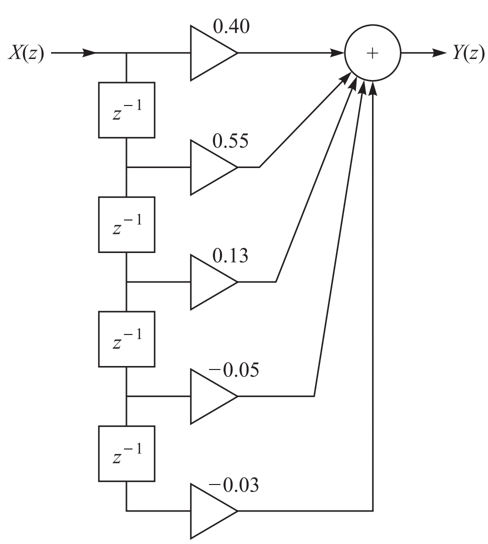


Fig. 365

Problem 319
Rework the foregoing example for $b_0 = 1.0$ and $b_1 = -1.0$, as shown in Fig. 365.

Problem 320
Here we wish to sketch the magnitude of the frequency response curve of the following low-pass digital filter.



For the above, find the values of $|H(r)|$ for $r = 0, 0.1, 0.2, 0.3, 0.4, 0.5$, then sketch the curve of $|H(r)|$ versus r, from $r = 0$ to $r = 0.5$.

Problem 321

Digital filters can be of the band-pass and band-elimination types as well as the low-pass and high-pass types. To illustrate this, consider Fig. 366, which is drawn in the "Direct Form II" configuration of Fig. 356 in section 13.8 (where, in this particular case, a_1 and b_1 are both equal to zero).

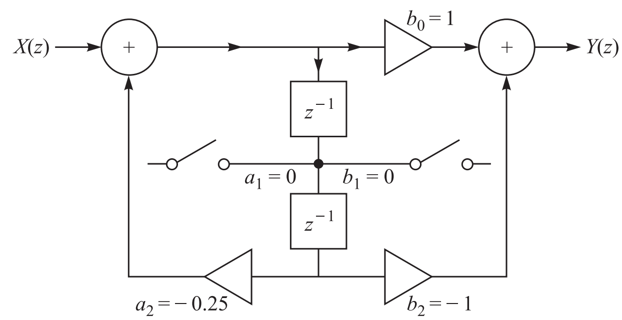


Fig. 366

(a) Show that the given filter is stable.

(b) Fill in the following table of values for the given values of r, then sketch the curve of $|H(r)|$ versus r.

r	$H(r)$	$\|H(r)\|$
0.00		
0.05		
0.10		
0.20		
0.25		
0.30		
0.40		
0.45		
0.50		

Appendix

Note 1. Some Basic Algebra

This item constitutes a very brief review of some basic rules and operations of algebra. We begin with the notation used to denote multiplication, as follows. Let A and B represent two numbers; then "A times B," called the "product" of A and B, is, in algebra, denoted in any of the following ways

$$A \times B = A \cdot B = AB$$

In the same way, if A, B, and C represent three numbers, then the product of the three can be denoted in any of the following ways

$$A \times B \times C = A \cdot B \cdot C = ABC$$

In the above, A, B, and C are referred to as the "factors" of the product ABC. It should be noted that multiplication is a "commutative" operation, which simply means that it makes no difference in what *order* the factors of a product are written; that is

$$AB = BA$$

Multiplication is also "associative," which means that the product of three or more numbers is the same in whatever way they may be grouped together; that is

$$ABC = A(BC) = (AB)C$$

Lastly, multiplication is "distributive" with respect to addition, which is summarized in the statement that

$$A(B + C) = AB + AC$$

which is read as "A, times the quantity B plus C, is equal to A times B, plus A times C."

In regard to positive and negative numbers, the rules concerning MULTIPLICATION are

the PRODUCT of two numbers having LIKE SIGNS is POSITIVE,

the PRODUCT of two numbers having UNLIKE SIGNS is NEGATIVE.

The "absolute" or "numerical" value of a number is its value *without regard to sign*. The absolute value of $-A$ is A, which is shown symbolically by writing

$$|-A| = A$$

Thus $|-2| = 2$, which is read "the absolute value of minus 2 is 2."

If no sign is shown with a number, the number is understood to be positive; thus, $2 = +2$, and so on.

In the *addition* of two numbers the following rules apply.

(a) To add two numbers having LIKE SIGNS, add their *absolute values* and prefix the common sign. Thus, $2 + 5 = 7$, $-2 - 5 = -7$, and so on.

(b) To add two numbers having UNLIKE SIGNS, take the *difference of their absolute values* and prefix to it *the sign of the number having the larger absolute value*. For example, $-2 + 5 = 3$, and $2 - 5 = -3$.

To SUBTRACT one number from another, change the sign of the number to be taken away and proceed as in addition. Thus, to subtract 5 (meaning $+5$) from 2, we have $2 - 5 = -3$. Or, to subtract -5 from 2 we have $2 + 5 = 7$, that is, $2 - (-5) = 7$.

Next, if one number A is to be DIVIDED BY another number B, this is indicated algebraically by the fractional form,

$$\frac{A}{B} = A/B$$

which is read as "A over B," meaning "A divided by B." If we write $\frac{A}{B} = C$, this says that "A divided by B is equal to C," in which the SIGN of the "quotient" C is POSITIVE if A and B have LIKE SIGNS but NEGATIVE if A and B have UNLIKE SIGNS. Thus, $6/3 = -6/-3 = 2$, but, $-6/3 = 6/-3 = -2$. In the expression A/B, A is called the "numerator" of the fraction and B is called the "denominator" of the fraction. The value of a fraction is not changed if the numerator and denominator are *both* multiplied or divided by the same quantity. In regard to the multiplication of fractions, the PRODUCT of two fractions is equal to "the product of the two numerators over the product of the two denominators"; that is

$$\frac{A}{B} \cdot \frac{C}{D} = \frac{AC}{BD}$$

In regard to an EQUATION, the equality of the two sides is preserved if the *same operation* is applied to *both sides* of the equation. For instance, multiplying both sides of the equation $A/B = C$ by B shows that $A = BC$; thus, $A/B = C$ and $A = BC$ denote the same relationship among the quantities A, B, and C.

Next we have the algebraic form B^a, in which "B" is called the *base number* and in which the *exponent "a"* is the "power" to which B is to be raised. The exponent a can be any positive or negative integer or fraction. If a is a positive integer (positive whole number), then B^a is simply a shorthand notation for the number of times B is to be multiplied by itself; thus, $B^2 = BB$, $B^3 = BBB$, and so on. Or, if a is an integer, then B^{-a} denotes the *reciprocal of* B^a; that is, $B^{-a} = 1/B^a$. Thus, $B^{-1} = 1/B$, $B^{-2} = 1/BB$, $B^{-3} = 1/BBB$, and so on.

If the exponent is a *fraction* $1/a$, then $B^{1/a}$ is the ath ("aye th"), *root* of B. For instance, if $a = 2$, then $B^{1/a} = B^{1/2}$, which is called the "square root" of B, which is also written using the "radical sign", thus

$$B^{1/2} = \sqrt{B} = C$$

which is defined to mean that

$$B = C^2$$

Likewise, $B^{1/3}$ is the "cube root" of B, which can be written in the form

$$B^{1/3} = \sqrt[3]{B} = C$$

meaning that

$$B = C^3$$

and so on. Thus $(16)^{1/2} = \sqrt{16} = \pm 4$, because $(\pm 4)^2 = 16$, and $(8)^{1/3} = \sqrt[3]{8} = 2$, since $8 = 2^3$.

All of the foregoing operations with exponents can be summarized in the following "laws of exponents," which are valid for all positive and negative integral and fractional values of the exponents.

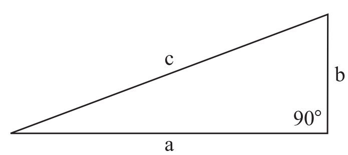
1. $B^a \cdot B^b = B^{a+b}$ (exponents *add* in multiplication)
2. $B^a / B^b = B^{a-b}$ (numerator exponent *minus* denominator exponent in division)
3. $(B^a)^b = B^{ab}$ (B to power a, raised to power b)
4. $(B/C)^a = B^a/C^a$ (fraction B/C raised to power a)

In the above, note that (1) and (2) apply for *like base numbers* only. To close our discussion of exponents, suppose $a = 0$ in law (1) above; for this case, using law (1), we have that $B^0 B^b = B^{0+b} B^b$, which can be true *only if* $B^0 = 1$. Thus it is *defined* that "B to the zero power is *one*," that is, $B^0 = 1$, where B has any finite value except zero (because *no value* is assigned to the expression 0^0).

The "right triangle" is important in algebraic applications. Let us denote the sides of a right triangle by a, b, and c, where side c, opposite the 90° angle, is the "hypotenuse," as shown below.

In any right triangle it is true that

$$c^2 = a^2 + b^2 \quad \text{the "Pythagorean theorem"}$$



thus,

$$c = \sqrt{a^2 + b^2}$$

The "degree" of an algebraic equation is equal to the *highest power of the unknown* in the equation. Thus, if a and b are known constant values, and x denotes the value of an unknown quantity, the basic FIRST DEGREE or "linear" equation is of the form

$$ax + b = 0$$

the solution of which is

$$x = -b/a$$

Or, if a, b, and c are known constant values, with x denoting the value of an unknown quantity, the basic SECOND DEGREE or "quadratic" equation is of the general form

$$ax^2 + bx + c = 0$$

the solutions of which are

$$x = \frac{1}{2a}\left(-b \pm \sqrt{b^2 - 4ac}\right)$$

In closing this item, let us be reminded that DIVISION BY ZERO IS NOT PERMITTED in algebra. The reason for this restriction can be shown as follows. Let A, B, and C represent three numbers, thus

$$\frac{A}{B} = C, \text{ meaning that } A = BC$$

If, however, $B = 0$, then $A = 0 \cdot C$, which means that $A = 0$ regardless of the value of C; thus, setting $B = 0$ leads to the expression $\frac{0}{0} = C$, which is meaningless, since C has no definite value.

Note 2. Fundamental Units

The four fundamental quantities in the mks system are

$$\begin{aligned} \text{LENGTH, } l, &\text{ measured in ``meters'' (m)} \\ \text{MASS, } m, &\text{ measured in ``kilograms'' (kg)} \\ \text{TIME, } t, &\text{ measured in ``seconds'' (s)} \\ \text{ELECTRIC CHARGE, } q, &\text{ measured in ``coulombs'' (C)} \end{aligned}$$

$$\begin{aligned} 1 \text{ meter} &= 39.37 \text{ inches} \\ &= 3.281 \text{ feet} \\ 1 \text{ kilometer} &= 1000 \text{ meters} \\ &= 0.6214 \text{ miles} \\ 1 \text{ kilogram} &= 1000 \text{ grams} \\ &= 2.205 \text{ pounds} \end{aligned}$$

In the mks system *force* is measured in NEWTONS, where

$$\begin{aligned} 1 \text{ newton} &= 0.2248 \text{ pounds of force} \\ 1 \text{ pound of force} &= 4.448 \text{ newtons} \end{aligned}$$

In physics, ENERGY is measured in terms of the ability to do *work*; in the mks system the basic unit of "energy" is the JOULE, where

$$\begin{aligned} 1 \text{ joule} &= 0.7376 \text{ foot-pounds of work} \\ 1 \text{ foot-pound of work} &= 1.356 \text{ joules} \end{aligned}$$

We note that *heat* is a form of energy; thus, mechanical and electrical energy can be transformed into heat, and heat can be transformed into mechanical and electrical energy. Heat is measured in "calories", where

$$\begin{aligned} 1 \text{ calorie} &= 4.186 \text{ joules of work} \\ &= 3.0876 \text{ foot-pounds of work} \end{aligned}$$

To conclude, POWER is the time *rate* of doing work; the UNIT OF POWER in the mks system is the *watt*, where

$$
\begin{aligned}
1\text{ watt} &= 1\text{ joule of work per second} \\
&= 0.7376\text{ foot-pounds per second} \\
1\text{ horsepower} &= 550\text{ foot-pounds per second} \\
&= 745.7\text{ watts} \\
1\text{ kilowatt} &= 1000\text{ watts} \\
&= 1.341\text{ horsepower}
\end{aligned}
$$

Note 3. Prefix Nomenclature

The following prefixes are generally accepted as denoting powers of 10.

p = pico = 10^{-12} ("trillionth," as in "picofarad")

n = nano = 10^{-9} ("billionth," as in "nanosecond")

μ = micro = 10^{-6} ("millionth," as in "microcoulomb")

m = milli = 10^{-3} ("thousandth," as in "milliampere")

k = kilo = 10^{3} ("thousand," as in "kilowatt")

M = mega = 10^{6} ("million," as in "megavolts")

G = giga = 10^{9} ("billion," as in "gigahertz")

Note 4. Vectors

A VECTOR quantity, as distinguished from ordinary "scalar" quantities, is a quantity having both MAGNITUDE and *a* SENSE OF DIRECTION and which obeys the PARALLELOGRAM LAW OF ADDITION (as will be discussed very shortly).

To show that a letter represents a vector quantity, we'll generally write the letter with an "overscore"; thus, as examples,

$$\bar{F}, \quad \bar{v}, \quad \bar{E}$$

A vector quantity is represented geometrically by a "directed line segment," which is a straight line with an arrowhead at one end. The LENGTH of the line is made proportional to the MAGNITUDE of the vector quantity, while the arrowhead shows the DIRECTION or sense of the vector.

Thus, if $\bar{A}$ is a vector quantity, $\bar{A}$ is represented geometrically by drawing a straight line (using any convenient scale such as, for example, 1 inch = 10 pounds of force) and then affixing an arrowhead to show the sense of direction, as shown in Fig. 1-A.

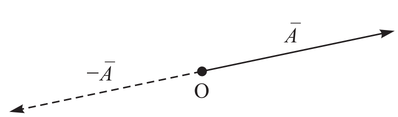


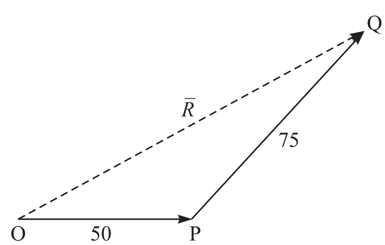
Fig. 1-A

Thus the *magnitude* of $\bar{A}$ is represented by the *length* of the straight line, while the *direction* or sense is indicated by the arrowhead. The *negative* of vector $\bar{A}$, which is denoted algebraically by writing $-\bar{A}$, is represented geometrically by the dashed line in Fig. 1-A. Notice that $-\bar{A}$ has the *same magnitude as* $\bar{A}$ but is drawn in exactly the opposite direction from $\bar{A}$. In Fig. 1-A, the point O and the arrowhead are often called the "tail" and "head" of the vector.

Probably the most basic or "prototype" vector quantity is "displacement," which involves both distance or length, and direction.

To illustrate the vector nature of displacement, suppose, for example, that a person starts at a point O, and walks 50 feet in a straight line to a point P. Suppose this person then turns sharply and then walks, say, 75 feet to point Q. We now have *two displacement vectors*, one scaled in length to represent a distance of 50 feet, the other 1.5 times as long to represent a distance of 75 feet.

We now ask the question, "At the end of the above action, where is the person relative to the starting position at O?". The answer, of course, depends not only on the *magnitudes* of the two displacements, 50 and 75 feet, but also upon the *directions* associated with these magnitudes. Since the directions are not given, we cannot, in this case, give a definite answer to the question. However, one POSSIBLE location of the point Q relative to the point O is illustrated in the figure below.



In this figure the vector $\bar{R}$ is called the *vector sum* of the two vectors of magnitudes 50 and 75; that is, the single vector $\bar{R}$ is mathematically and geometrically equivalent to the combined effects of the two component vectors, so that $\bar{R}$ is called the "vector sum" of the two component vectors.

From the figure, it's apparent that "vector addition" is quite different from the ordinary algebraic addition of scalar quantities. This can be investigated in more detail as follows.

Using the above figure, let $\bar{A}$ denote the displacement vector of magnitude 50, that is, $|\bar{A}| = 50$, and let $\bar{B}$ denote the displacement vector of magnitude 75, $|\bar{B}| = 75$.

If we now let $\bar{R}$ be the combined effect (the "vector sum") of $\bar{A}$ and $\bar{B}$, we then find *by direct experiment with displacements* that the value of $\bar{R}$ can be found graphically by any of the three equivalent procedures shown in Figs. 2-A, 3-A, and 4-A. Note that $\bar{R}$ has the *same magnitude and direction* in all three figures.

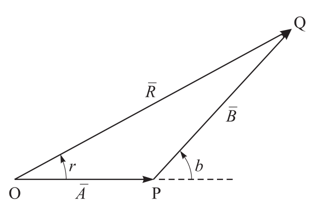


Fig. 2-A

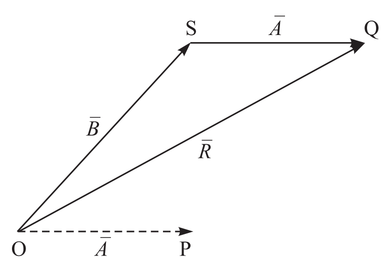


Fig. 3-A

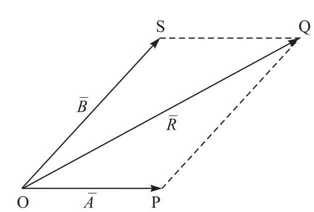


Fig. 4-A

In the figures, the "vector quantities" all represent *distances* measured off at different angles relative to some specific "reference line."

In the figures, let us suppose that the distance represented by $\bar{A}$ happens to lie exactly along the reference line; that is, suppose that $\bar{A}$ is at an angle of *zero degrees* with respect to the reference line; symbolically this can be shown by writing that

$$\bar{A} = |\bar{A}|\underline{/0^\circ}$$

Or, if $\bar{A}$ were measured (for example) at an angle of, say, 20° with respect to the reference line, this would be shown by writing

$$\bar{A} = |\bar{A}|\underline{/20^\circ}$$

In Fig. 2-A, $\bar{B}$ is a distance, of magnitude $|\bar{B}|$, measured from point P at an angle let us denote by "b" degrees relative to the vector $\bar{A}$, as shown in the figure. (We're assuming that $\bar{A}$ is at 0° relative to the reference line.) The *resultant* of these two measurements is a distance of magnitude $|\bar{R}|$ at, let us say, an angle of "r" degrees relative to the reference line. Algebraically, the whole operation can be expressed by writing that

$$|\bar{A}|\underline{/0^\circ} + |\bar{B}|\underline{/b^\circ} = |\bar{R}|\underline{/r^\circ}$$

or, in a more abbreviated form, $\bar{A} + \bar{B} = \bar{R}$, which says that vector $\bar{R}$ is equal to the *sum* of vectors $\bar{A}$ and $\bar{B}$.

Now, in regard to actually *finding* the value of $\bar{R}$, where $\bar{R} = \bar{A} + \bar{B}$, two procedures, one graphical and the other mathematical, are available. Here we'll mainly emphasize the graphical procedure, as follows.

At the beginning of this discussion we defined that quantities are truly *vector* quantities only if they obey the PARALLELOGRAM LAW OF ADDITION. This simply means that, vectorially speaking, in order for $\bar{R}$ to be the true vector sum of $\bar{A}$ and $\bar{B}$, $\bar{R}$ must be equal to the DIAGONAL OF THE PARALLELOGRAM having $\bar{A}$ and $\bar{B}$ as opposite sides of the parallelogram.

With this in mind, consider Fig. 2-A; we know, *from actual experience*, that displacement $\bar{A}$, plus displacement $\bar{B}$, produces exactly the same final result as would the single vector $\bar{R}$. But study of Figs. 3-A and 4-A shows that $\bar{R}$ *is* equal to the diagonal of the parallelogram having $\bar{A}$ and $\bar{B}$ as opposite sides; thus displacement is a true "vector" quantity.

To illustrate the above graphical procedure, consider Fig. 5-A, in which $\bar{A}$ and $\bar{B}$ are given to be two vector quantities at an angle a, as shown. Let the problem be to find, graphically, the "resultant" vector $\bar{R}$, where $\bar{R} = \bar{A} + \bar{B}$.

Since $\bar{A}$ and $\bar{B}$ are given to be vector quantities, the first step is to construct a parallelogram with $\bar{A}$ and $\bar{B}$ as opposite sides, as in Fig. 6-A. The *diagonal* of the parallelogram, drawn from the junction of $\bar{A}$ and $\bar{B}$, is the resultant of the two vectors $\bar{A}$ and $\bar{B}$; that is, it is the graphical solution of the equation $\bar{R} = \bar{A} + \bar{B}$.

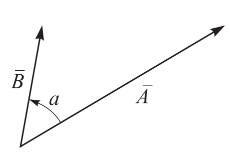


Fig. 5-A

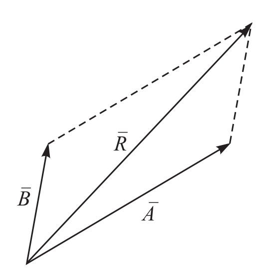


Fig. 6-A

In the foregoing discussion we've referred to displacement as the basic vector quantity. A number of other quantities, such as force and velocity, are also vector quantities. Of special interest to us, however, is the fact that "rms values" of alternating currents can be treated as vector quantities.

If we have THREE vectors, we first find the resultant of any two, then combine that resultant with the third vector to get the final resultant. We proceed in the same way to find the sum of any number of vectors. In this way we can state that the rule for finding the *sum of n vectors by geometric means* is as follows.

> Keeping the *directions* of the vectors unchanged, move them by "translation" (that is, without rotation) until the tail of the second touches the head of the first, the tail of the third touches the head of the second, and so on, for all the *n* vectors.
>
> The *sum* or "resultant" of the *n* vectors is the vector $\bar{R}$, which is drawn from the tail of the first vector to the head of the last *n*th vector.

The above rule is illustrated for the addition of four vector quantities, $\bar{A}, \bar{B}, \bar{C}, \bar{D}$, in Fig. 7-A. This figure is thus the geometric solution to the vector equation $\bar{A} + \bar{B} + \bar{C} + \bar{D} = \bar{R}$, where the vectors are given to have the magnitudes and directions as shown.

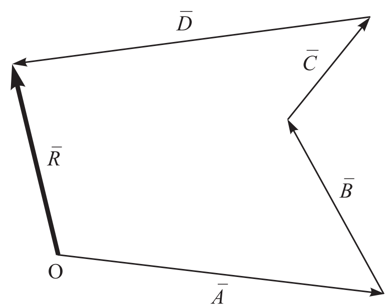


Fig. 7-A

To geometrically find the *difference* of two vectors, $\bar{A} - \bar{B}$, we draw the vector $-\bar{B}$ in accordance with Fig. 1-A, then combine it with vector $\bar{A}$ by means of the parallelogram law in the usual way. This is illustrated in Fig. 8-A, which shows the geometric solution to the vector equation $\bar{R} = \bar{A} - \bar{B}$.

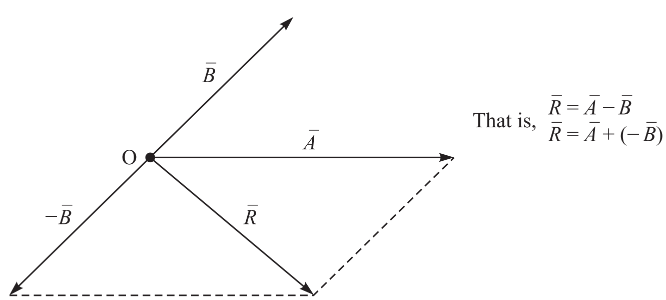


Fig. 8-A

It is sometimes convenient to represent a given vector as being the sum of two or more "component" vectors. This is illustrated in Fig. 9-A, in which a given vector $\bar{A}$ is represented as being the vector sum of the two components $\bar{A}'$ and $\bar{A}''$ (A prime and A double prime).

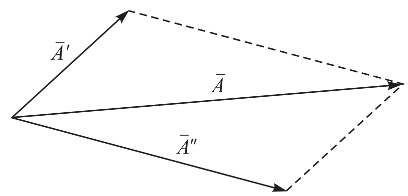


Fig. 9-A

It's especially useful, when finding the sum of a number of vectors, to express each vector in terms of its HORIZONTAL AND VERTICAL COMPONENTS. To do this, we place the "tails" of all the vectors at the origin O of the x, y plane, then resolve each vector into *horizontal* components, all lying on the x-axis, and *vertical* components, all lying on the y-axis. This is illustrated in Fig. 10-A for two given vectors $\bar{A}$ and $\bar{B}$, with their horizontal and vertical components denoted by the subscripts "h" and "v" respectively.

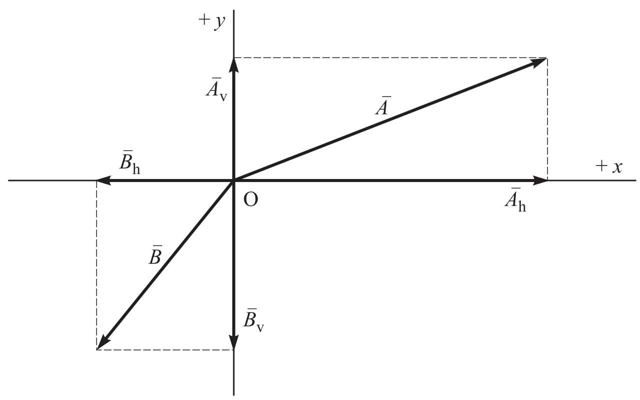


Fig. 10-A

The above procedure applies to the problem of finding the resultant sum, $\bar{R}$, of any number of vectors. The *advantage* of the procedure is that the *total* HORIZONTAL COMPONENT of $\bar{R}$ is the simple *algebraic sum of all the horizontal components*, and the *total* VERTICAL COMPONENT of $\bar{R}$ is the simple *algebraic sum of all the vertical components*.

Note 5. Increment (Delta) Notation

The symbol Δ is the Greek letter "delta." A term such as "Δq" denotes an optionally SMALL CHANGE in the value of a variable q, and is read as "delta q." Note that Δq does *not* mean "delta times q."

Terms such as Δq and Δt are called "increments" of the variables q and t, and denote small changes in the values of q and t. As used here, q denotes a total amount of electric charge in coulombs, and t denotes a total amount of time in seconds, counted from some optionally chosen time at which $t = 0$.

Hence the *ratio* of the increments, $\Delta q/\Delta t$, is the AVERAGE "time rate of change" of q in coulombs per second, which is given the name "amperes."

At a time $t + \Delta t$ the total charge is $q + \Delta q$; thus, as Δt becomes smaller and smaller, Δq *also* becomes smaller and smaller, and the *ratio* $\Delta q/\Delta t$ comes closer and closer to being the EXACT value of current, i, flowing at the *beginning* of the interval of time Δt, at time t. This idea is expressed mathematically by writing that

$$\lim_{\Delta t \to 0} \frac{\Delta q}{\Delta t} = i = \text{exact current at a time } t$$

which says that "the limiting value of the ratio $\Delta q/\Delta t$, as Δt is allowed to approach zero as a limit, is the exact value of current at a time t." If the mathematical relationship between q and time t is known, then, using the formulas of differential calculus, the value of current i at any time t can be calculated. Mathematically, the above limit is denoted by the symbol "$\mathrm{d}q/\mathrm{d}t$," which is read as "dee q, dee t." Thus

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \text{coulombs per second} = \text{amperes}$$

Note 6. Similar Triangles. Proof of Eq. (98)

Two triangles are called "similar" if their ANGLES *are all equal*. Thus the two right triangles in Fig. 11-A are similar triangles.

By definition, if two triangles are similar *their "corresponding angles" are equal* (in Fig. 11-A the corresponding angles are 30 and 30, 60 and 60, and 90 and 90 degrees). The "corresponding sides" are A and A', B and B', and C and C', as shown. Note that the *ratio* A/B is equal to the *ratio* A'/B', the *ratio* A/C is equal to the *ratio* A'/C', and so on.

In Fig. 11-A the two similar triangles are in a position such that their corresponding sides are PARALLEL; thus, if the corresponding sides of two triangles can be shown to be mutually PARALLEL, this establishes that the two are *similar* triangles.

Also, *if* two triangles are *similar* they can always be moved and rotated into a position such that *their corresponding sides are* PERPENDICULAR, as illustrated in Fig. 12-A. Thus, if the corresponding sides of two triangles can be shown to be mutually PERPENDICULAR, this is sufficient to establish that the two are *similar* triangles.

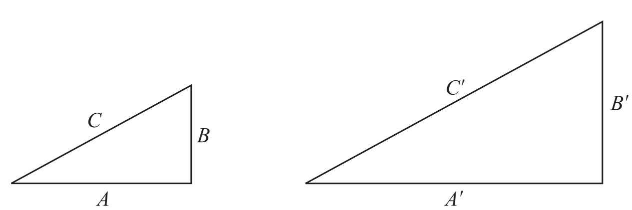


Fig. 11-A

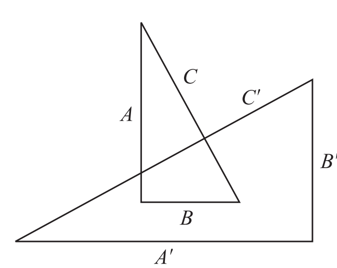


Fig. 12-A

Now, with the above facts in mind, let x and y be two adjacent angles, as in Fig. 13-A.

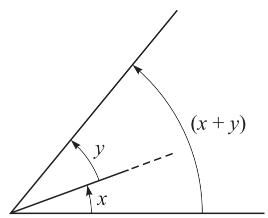


Fig. 13-A

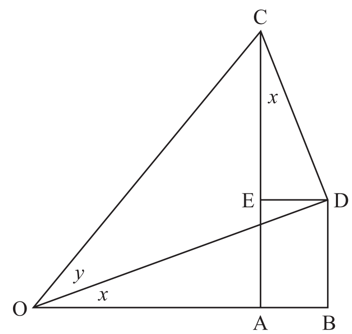


Fig. 14-A

Now consider Fig. 14-A, in which it is given that

lines CA and DB are perpendicular to line OB,

line ED is perpendicular to line CA,

line CD is perpendicular to line OD.

In Fig. 14-A, consider the two triangles OBD and CED. Close inspection will show that the corresponding sides of these two triangles are *perpendicular* and thus that they are *similar* triangles, with angle x as shown. Hence, from direct inspection of Fig. 14-A we have that

$$\cos(x+y) = \frac{\text{OA}}{\text{OC}} = \frac{\text{OB} - \text{AB}}{\text{OC}} = \frac{\text{OB}}{\text{OC}} - \frac{\text{ED}}{\text{OC}} \text{ (since AB = ED)}$$

Then, since OB = OD cos x and ED = CD sin x, we have

$$\cos(x+y) = \frac{\text{OD}}{\text{OC}} \cos x - \frac{\text{CD}}{\text{OC}} \sin x$$

But note that OD/OC = cos y and CD/OC = sin y; thus the preceding equation can be written in the standard form

$$\cos(x+y) = \cos x \cos y - \sin x \sin y$$

which, for the special case of $y = x$, becomes

$$\cos 2x = \cos^2 x - \sin^2 x$$

Now make use of the identity $\sin^2 x + \cos^2 x = 1$ (from problem 64). Doing this, then writing "θ" in place of "x," gives the required eq. (98).

Note 7. Identity for $\sin(x+y)$

From Fig. 14-A:

$$\sin(x+y) = \frac{\text{AC}}{\text{OC}} = \frac{\text{CE} + \text{AE}}{\text{OC}} = \frac{\text{CE}}{\text{OC}} + \frac{\text{BD}}{\text{OC}} \text{ (since BD = AE)}$$

Then, since, CE = CD cos x and BD = OD sin x, we have

$$\sin(x+y) = \frac{\mathrm{CD}}{\mathrm{OC}}\cos x + \frac{\mathrm{OD}}{\mathrm{OC}}\sin x$$

But note that CD/OC = sin y and OD/OC = cos y; thus the preceding equation can be written in the standard form

$$\sin(x+y) = \sin x \ \cos y + \cos x \sin y$$

Note 8. Often-Used Greek Letters

The following are Greek letters most often used in engineering work.

α = alpha	Ω = omega (capital)
β = beta ("BAY tah")	θ = theta ("THAY tah")
ϵ = epsilon	μ = mu
δ = delta (small)	π = pi
Δ = delta (capital)	ϕ = phi ("fee")
ω = omega (small)	

Note 9. Sinusoidal Waves of the Same Frequency

Applying the identity found in note 7 to the *left-hand side* of eq. (124) gives the equality

$$A\sin\omega t + B\sin(\omega t + a) = (A + B\ \cos a)\sin\omega t + B\sin a\ \cos\omega t$$

Or, letting E and F denote the CONSTANT values $A + B\ \cos a$ and $B\sin a$, the above becomes

$$A\sin\omega t + B\sin(\omega t + a) = E\sin\omega t + F\ \cos\omega t$$

which can also be written in the form

$$A\sin\omega t + B\sin(\omega t + a) = \sqrt{E^2+F^2}\left(\frac{E}{\sqrt{E^2+F^2}}\sin\omega t + \frac{F}{\sqrt{E^2+F^2}}\ \cos\omega t\right) \qquad \text{(1-A)}$$

Now, letting ϕ be a constant angle, construct the right triangle shown in Fig. 15-A.

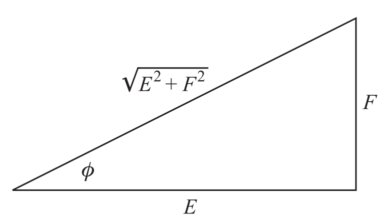


Fig. 15-A

Note that $\dfrac{E}{\sqrt{E^2+F^2}} = \cos\phi$ and $\dfrac{F}{\sqrt{E^2+F^2}} = \sin\phi$, and thus eq. (1-A) becomes

$$A\sin\omega t + B\sin(\omega t + a) = \sqrt{E^2+F^2}\,(\sin\omega t\ \cos\phi + \cos\omega t\sin\phi)$$

Now apply the identity found in note 7 to the right-hand side above (setting $x = \omega t$ and $y = \phi$) to get the final result that

$$A\sin\omega t + B\ \sin(\omega t + a) = \sqrt{E^2+F^2}\sin(\omega t + \phi) \tag{2-A}$$

thus proving that the *sum* of two sinusoidal waves of the same frequency is equivalent to a single sinusoidal wave of the same frequency.

Note 10. Sinusoidal Waves as Vectors

Let A and B denote the PEAK VALUES of two sine waves of the same frequency, with the first wave "lagging" the second wave by a degrees. In deriving eq. (2-A) above for the same two waves we algebraically showed that

$$\begin{aligned}\text{peak value of the SUM of the two waves} &= \sqrt{E^2+F^2}\\ &= \sqrt{A^2+B^2+2AB\ \cos a}\end{aligned}$$

(making use of the identity $\sin^2 a + \cos^2 a = 1$), and also, from inspection of Fig. 15-A, that

$$\text{PHASE ANGLE of resultant wave} = \phi = \arctan\frac{F}{E} = \arctan\frac{B\sin a}{A + B\ \cos a}$$

We now wish to show that the *same results* can be obtained *graphically* by using the "phasor" representation of sine waves, and by assuming that phasors can be treated as *vector* quantities; that is, that phasors ADD together in accordance with the PARALLELOGRAM LAW of addition of vectors. To do this, let us begin with a phasor diagram of the two waves, such as shown in Fig. 16-A.

Now let R be the VECTOR sum of A and B, as shown in Fig. 17-A. Then, upon applying the Pythagorean theorem to the large right triangle, we have

$$R^2 = (A + B\ \cos a)^2 + B^2\sin^2 a$$

hence, after applying the identity $\sin^2 a + \cos^2 a = 1$, we have that the MAGNITUDE of the resultant vector R is

$$R = \sqrt{A^2+B^2+2AB\cos a}$$

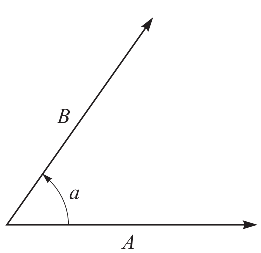


Fig. 16-A

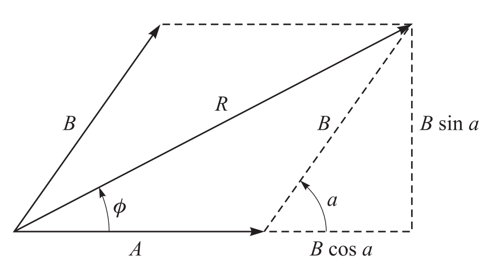


Fig. 17-A

Next, inspection of Fig. 17-A shows that the PHASE ANGLE of the resultant vector R is

$$\phi = \arctan \frac{B \sin a}{A + B \ \cos a}$$

Note that the above values of R and ϕ are the same as the true values of peak voltage and phase angle found by mathematical means in note 9. Thus it *is correct* to represent peak values of sinusoidal waves of the same frequency as *vector* quantities, provided that we are interested in knowing only the "peak values" of the waves.

Now suppose that the lengths of all the lines in Fig. 17-A were multiplied by 0.7071; doing this would not change the SHAPE of the figure in any way, but now the lengths would represent "rms" values instead of peak values. Thus the foregoing statement, concerning the vector representation of peak values of sinusoidal waves, holds also for the *rms values* of such waves.

Note 11. Rational and Irrational Numbers

Real numbers are classified as being either *rational* or *irrational*. Let us first consider the "rational" type, as follows.

A *rational* number is defined to be the *ratio of two integers*, that is, a rational number is of the form a/b, where a and b are integers (whole numbers). Thus, 1/3, 8/9, 29/50 are examples of rational numbers.

When expressed in DECIMAL *form*, rational numbers are always of the REPEATING type; that is, the decimal form of a rational number always consists of a single number, or a block of numbers, that is REPEATED OVER AND OVER, endlessly. As examples,

$$\frac{1}{3} = 0.33333\cdots \qquad \frac{4}{11} = 0.36363636\cdots \qquad \frac{1}{5} = 0.200000\cdots$$

(in the last example, the answer is 0.2 followed by *zero* repeated over and over).

As an aid to understanding WHY the decimal form of a rational number is always of the "repeating" type, let us convert the rational number 23/37 into decimal form. To do this, let us use ordinary "long division" to divide 23 by 37, as follows.

```
        0.6216
   37 ) 23.000000 ...
       -22 2
          80 (remainder of 8)
         -74
           60 (remainder of 6)
          -37
           230 (remainder of 23)
          -222
             8
```

which is the *second time* that the remainder 8 has appeared; hence the preceding three results will be repeated, resulting in a remainder of 8 for the *third time*, and so on in this manner, endlessly. Thus we have that

$$\frac{23}{37} = 0.621\,621\,621\,621\cdots$$

In any such division (a/b), there is only a certain number of possible remainders. As the division operation proceeds, sooner or later there has to occur a *repetition* of a previous remainder. At this point the process "starts over," exactly duplicating the preceding cycle of remainders, until it again returns to the repeated remainder, at which point the process is again repeated, and so on, endlessly. Thus a "rational" number can be said to be an "orderly" type of number.

Now let us consider *irrational* numbers; an "irrational" number is defined as one which *cannot be expressed exactly as the quotient of two integers*. An irrational number has no such order as a "rational" number has. The difference between rational and irrational numbers is simply that an *irrational* number involves an infinite number of decimal places which are unordered, that is, not in the form of an endless repetition of a block of numbers. In any case, an engineer, when manipulating numbers, need not concern himself with the particular arrangement of the digits in the decimal fractions.

Note 12. The Concept of Power Series

As an example of a "power series in x," consider the following. Let

$$S_n = 1 + x + x^2 + x^3 + \cdots + x^{n-1} + x^n \tag{3-A}$$

where S_n denotes *the* SUM *of the first $n+1$ terms of the series*. Thus

$$\text{for } n = 1, \quad S_1 = 1 + x,$$
$$\text{for } n = 2, \quad S_2 = 1 + x + x^2,$$
$$\text{for } n = 3, \quad S_3 = 1 + x + x^2 + x^3, \text{ and so on.}$$

Now suppose that the number of terms in eq. (3-A) is allowed to "increase without bound"; that is, suppose that n is allowed to become INFINITELY GREAT ($n \to \infty$). For this case it's clearly true that S_n, the *sum* of the $n+1$ terms, will *also* become infinitely great *if x is* EQUAL TO OR GREATER THAN *1*.

But suppose that the value of x is LESS THAN *1*, that is, $x < 1$. In this case the result is open to question because *as n increases, the value of x^n decreases* (because $x < 1$). A definite answer can be found, however, by taking the following steps. First, multiply both sides of eq. (3-A) by x to get

$$xS_n = x + x^2 + x^3 + \cdots + x^n + x^{n+1}$$

Now *subtract* the last equation from eq. (3-A); doing this, you should find that

$$S_n = \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x} - \frac{x^{n+1}}{1 - x} \tag{4-A}$$

Now let the number of terms, n, increase without bound; that is, let $n \to \infty$. Doing this, and noting that *if x is* LESS THAN *1* then

$$\lim_{n\to\infty} x^{n+1} = 0 \text{ (for } x < 1)$$

we find that eq. (4-A) becomes

$$\lim_{n\to\infty} S_n = \frac{1}{1 - x} \text{ (for } x < 1)$$

Hence, if the number of terms n is allowed to *increase without bound*, S_n becomes equal to $\dfrac{1}{1-x}$ and thus eq. (3-A) can be written as

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n \quad (\text{for } x < 1) \tag{5-A}$$

which, it should be understood, is EXACTLY TRUE only in the limit as n becomes infinitely great.

In the same way, the functions ϵ^x, $\sin x$, and $\cos x$ can be represented by power series in x. In these cases, however, the nature of the series is such that the series representation is valid for ALL positive and negative values of the variable x.

Note 13. Series *RL* Circuit. *L/R* Time Constant

We must first note the nature of the "negative exponential function," $\epsilon^{-x} = 1/\epsilon^x$, where ϵ (epsilon) denotes the irrational number $\epsilon = 2.71828\ldots$, defined by eq. 146 in section 6.5.

In the discussion here, we'll be interested only in the case where x is a positive real number. Using your calculator, you can verify the values listed in the following "table of values" (values of ϵ^{-x} rounded off to two decimal places). These values are plotted against x in Fig. 18-A.

x	ϵ^{-x}	x	ϵ^{-x}
0.00	1.00	0.80	0.45
0.10	0.91	0.90	0.41
0.20	0.82	1.00	0.37
0.30	0.74	1.50	0.22
0.40	0.67	2.00	0.14
0.50	0.61	3.00	0.05
0.70	0.50	5.00	0.01

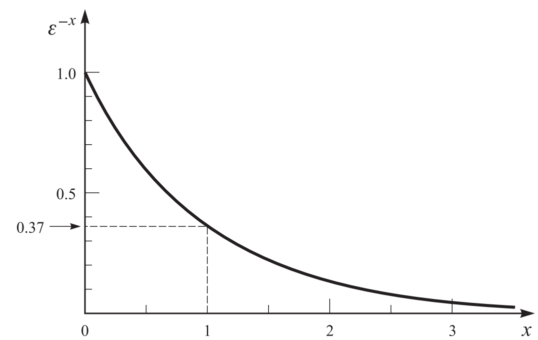


Fig. 18-A

Now consider the basic *series "RL" circuit*, to which a constant voltage of V volts is applied at the closing of a switch, as shown in Fig. 19-A.

In Fig. 19-A, L is inductance in henrys, R is resistance in ohms, and i is current in amperes flowing any time t seconds after the switch is closed at $t = 0$.

At $t = 0$ *the current i is zero*, at which time the entire applied voltage V appears across the coil L; then, as time increases, the current increases slowly toward the limiting value of $I = V/R$, as shown in Fig. 20-A. As this occurs, the voltage drop across L decreases, while the voltage drop across R rises toward the limiting value of $IR = V$ volts.

The exact relationship between the current i and time t is given by the equation

$$i = \frac{V}{R}\left(1 - \epsilon^{-Rt/L}\right) \tag{6-A}$$

Note that when $t = 0$, then $i = 0$, as already mentioned, and as shown in Fig. 20-A. Then, as time increases, the term $\epsilon^{-Rt/L}$ *decreases* exponentially toward the value zero (as

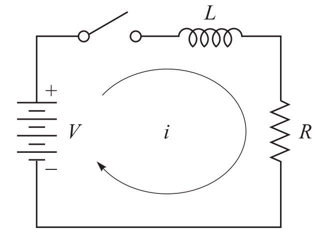


Fig. 19-A

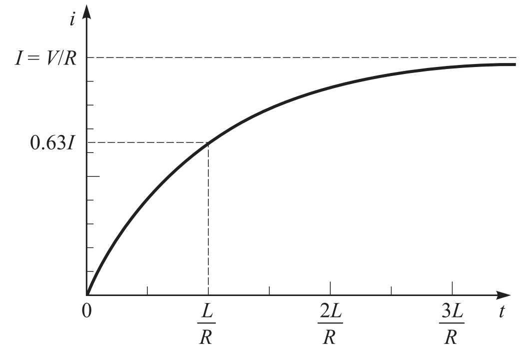


Fig. 20-A

in Fig. 18-A). Thus, as time t increases, the current i increases toward the *limiting value* of $i = I = V/R$ amperes, as shown in Fig. 20-A.

In Fig. 20-A, note that "time" is expressed in multiples of L/R. This can be done because the *ratio* of henrys to ohms is *time* in seconds, as the following shows.

By eq. (181),

$$L = \frac{v}{\mathrm{d}i/\mathrm{d}t} = \frac{\text{volts}}{\text{amp/sec}} = \frac{\text{volts} \times \text{sec}}{\text{amp}}$$

and, by Ohm's law,

$$\frac{1}{R} = \frac{\text{amp}}{\text{volts}}$$

hence,

$$\frac{L}{R} = \frac{\text{volts} \times \text{sec}}{\text{amp}} \times \frac{\text{amp}}{\text{volts}} = \text{seconds}$$

The ratio of henrys to ohms, L/R, is called the "time constant" of the basic series circuit of Fig. 19-A.

As Fig. 20-A shows, at the end of one time constant (L/R seconds) the current in Fig. 19-A will have risen to approximately 63% of its final value of $I = V/R$ amperes. (To show this, set $t = L/R$ in eq. (6-A).)

Note 14. Series *RC* Circuit. *RC* Time Constant

Here we wish to emphasize that *time* is required to change the amount of energy stored in the electric field of a capacitor. To illustrate this, consider the basic *series "RC" circuit*, to which a constant voltage of V volts is applied at the closing of a switch, as shown in Fig. 21-A. We wish to examine the manner in which the VOLTAGE ACROSS THE CAPACITOR increases after the switch is closed.

In Fig. 21-A, R is resistance in ohms, C is capacitance in farads, and i is current in amperes flowing at any time t seconds after the switch is closed. We'll assume that initially (at $t = 0$) there is zero voltage across the capacitor.

We first note that, at the instant the switch is closed at $t = 0$, the capacitor momentarily behaves like a "short circuit"; thus, *at* $t = 0$ the current is equal to V/R amperes. Then, as time increases and the capacitor begins to charge, the current i *decreases* exponentially, in the manner of Fig. 18-A, approaching the limiting value of *zero*. Thus, *as time increases*, the voltage across the capacitor rises, in an exponential-type curve, toward the final limiting value of V volts. Letting "v_c" denote the *voltage across the capacitor*, the relationship between v_c and time t is shown graphically in Fig. 22-A, the equation of the curve being

$$v_c = V(1 - \epsilon^{-t/RC}) \tag{7-A}$$

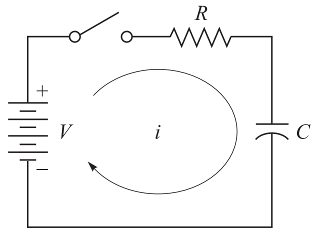


Fig. 21-A

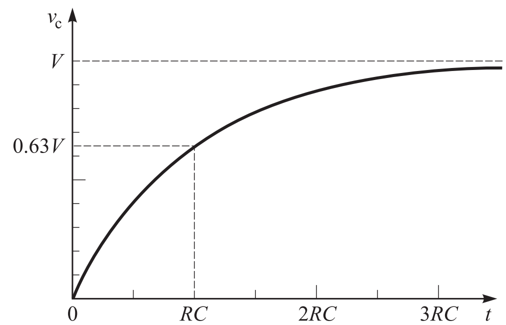


Fig. 22-A

Note that when $t = 0$, then $v_c = 0$, as already mentioned and as shown in Fig. 22-A. Then, as time increases, the term $\epsilon^{-t/RC}$ *decreases* exponentially toward the value *zero* (as in Fig. 18-A). Thus, as t increases, the voltage v_c increases toward the limiting value of V volts, as shown in Fig. 22-A.

In Fig. 22-A, note that time is expressed in multiples of RC. This can be done because the product "ohms times farads" is *time* in seconds, as the following shows.

First, by Ohm's law,

$$R \text{ in ohms} = \frac{\text{volts}}{\text{amperes}} = \frac{\text{volts}}{\text{coulombs/sec}} = \frac{\text{volts} \times \text{seconds}}{\text{coulombs}}$$

and then by eq. (184),

$$C \text{ in farads} = \frac{\text{coulombs}}{\text{volts}}$$

hence,

$$RC = \frac{\text{volts} \times \text{seconds}}{\text{coulombs}} \times \frac{\text{coulombs}}{\text{volts}} = \textit{seconds}$$

The product, ohms times farads, is called the "time constant" of the basic series circuit of Fig. 21-A. As Fig. 22-A shows, at the end of one time constant (RC seconds) the voltage across the capacitor has risen to 63% of its final value of V volts. (To show this, set $t = RC$ in eq. (7-A).)

Note 15. ωL is in Ohms

First, $\omega = 2\pi f = 2\pi/T$, where T is time of one cycle (eq. (91), Chap. 5). Thus, since π is simply the ratio of two lengths, we see that ω is basically measured in terms of $1/T$, that is, in "reciprocal seconds."

Next, by eq. (181) of Chap. 7,

$$L = \frac{\text{volts}}{\text{amperes/seconds}} = \frac{\text{volts} \times \text{seconds}}{\text{amperes}}$$

hence,

$$\omega L = \frac{1}{\text{seconds}} \, \frac{\text{volts} \times \text{seconds}}{\text{amperes}} = \frac{\text{volts}}{\text{amperes}} = \text{ohms}$$

Note 16. $j\bar{Z} = \bar{Z}$ Rotated through 90 Degrees

Let us make use of the exponential form of a complex number (section 6.5) as follows.

Let $\bar{Z}$ be a complex number of magnitude A and angle θ; thus

$$\bar{Z} = A\epsilon^{j\theta}$$

then

$$j\bar{Z} = Aj\epsilon^{j\theta}$$

But note that $\epsilon^{j90^\circ} = \cos 90^\circ + j \sin 90^\circ = j$; thus the preceding expression becomes

$$j\bar{Z} = A\epsilon^{j90^\circ}\epsilon^{j\theta} = A\epsilon^{j(\theta+90^\circ)}$$

showing that $j\bar{Z}$ is equal to $\bar{Z}$ rotated through 90°.

Note 17. $1/\omega C$ is in Ohms

From note 15, ω is measured in reciprocal time, $1/T$, while capacitance C is measured in "coulombs per volt," q/v (eq. (184) in Chap. 7). Thus $1/\omega C$ is basically measured in units of

$$\frac{1}{\dfrac{1}{\text{second}} \, \dfrac{\text{coulombs}}{\text{volts}}} = \frac{\text{volts}}{\text{coulombs/second*}} = \frac{\text{volts}}{\text{amperes}} = \text{ohms}$$

Note 18. Harmonic Frequencies. Fourier Series

If any particular frequency, f, is taken to be a "fundamental" frequency, then any INTEGRAL MULTIPLE of f is said to be a *harmonic* of f. Thus, $2f$ is the *second harmonic* of f, $3f$ is the *third harmonic* of f, and so on, so that nf is any nth *harmonic* of f, where $n = 1, 2, 3, \ldots$ (for $n = 1$ we have "fundamental" instead of "first harmonic").

Now suppose we have some kind of *non-sinusoidal* function which occupies the interval from $x = 0$ to $x = 2\pi$, and which is exactly *repeated*, over and over, endlessly, for all positive and negative intervals of 2π, as in Fig. 23-A.

* "Coulombs per second" is amperes.

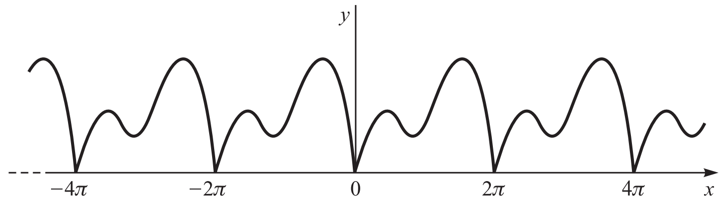


Fig. 23-A

It is a fact that ANY such repeating function, as met in engineering, can, for purposes of analysis, be considered to be composed of a FUNDAMENTAL *sinusoidal wave* plus, in general, an infinite number of sinusoidal *harmonics* of the fundamental wave.*

This is a fact of great usefulness because it allows us, by the principle of superposition, to apply the ordinary algebra of complex numbers to the analysis of networks to which *non-sinusoidal* waves are applied.

In such a representation, each complete *fundamental* wave (which is the *lowest* frequency component) covers a distance of 2π radians on the x-axis. Hence, in a distance of 2π radians there will be *two* complete *second* harmonic waves, *three* complete *third* harmonic waves, and so on.

Let us discuss, as an interesting example, the symmetrical "square wave" of Fig. 24-A.

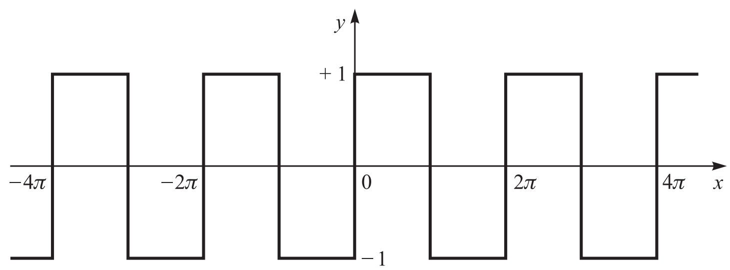


Fig. 24-A

Using a procedure called Fourier ("foo ree AYE") analysis,† it is found that the above square wave can be represented by the following infinite series:

$$y = \frac{4}{\pi}\left[\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \frac{1}{7}\sin 7x + \cdots\right] \tag{8-A}$$

where $\sin x$ is the *fundamental* sinusoidal component, the angle x being in radians. The equation shows that the square wave is composed of *odd* harmonics only ($3x, 5x, 7x$, and so on). Note that the higher the *order* of the harmonic, the *lower* is its amplitude. Also, because of its symmetry and its position relative to the x-axis, the wave has no constant term (no dc component).

If, now, you were to take the time to actually calculate a number of values of y, using eq. (8-A),‡ then plot the values of y versus x, you would get the result shown in Fig. 25-A, for $x = 0^\circ$ to $x = 360^\circ$, where degrees = (radians)($180/\pi$) (section 5.4).

* The function can include a possible CONSTANT term (often referred to as the "dc" component).

† Named for Joseph Fourier (1768–1830), French mathematician.

‡ Using the first four terms of eq. (8-A).

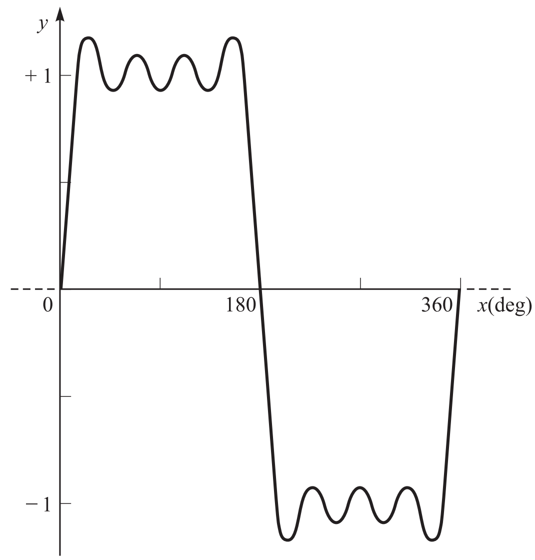


Fig. 25-A

It's obvious that the sum of just the first *four* terms of the series doesn't produce a very good square wave. If, for example, we had taken the sum of the first *eight* terms of the series (up to the 15th harmonic) the result would have been considerably improved.

Note 19. Logarithms. Decibels

LOGARITHMS are EXPONENTS. By definition, the LOGARITHM *of a number* is the POWER to which a fixed number, called the *base number*, must be raised to, to equal the number.

In theoretical work, the base number is taken to be the irrational number denoted by ϵ (section 6.5). In certain practical work, however, it's more meaningful to use the number *ten* as the base number.*

The OBJECT of this note is to show, in just a general way, why eq. (315) does make sense in a practical way. To do this, let us begin with the previous definition that

> If x is any positive number, then $\log x$ is the POWER that 10 has to be raised to, to equal x; that is, by definition,
>
> $$x = 10^{\log x}$$

Note that x will always be greater than the exponent $\log x$. To see this relationship more clearly, consider first the table of values to the left of Fig. 26-A, in which the calculator values of $\log x$ have been rounded off to two decimal places.

* The notation "ln x" denotes the logarithm of x to the base ϵ, while "log x" denotes the logarithm of x to the base 10.

x	$\log x$
0.1	−1.00
0.5	−0.30
1.0	0.00
2.0	0.30
3.0	0.48
4.0	0.60
5.0	0.70
6.0	0.78
7.0	0.85
8.0	0.90
9.0	0.95
10.0	1.00

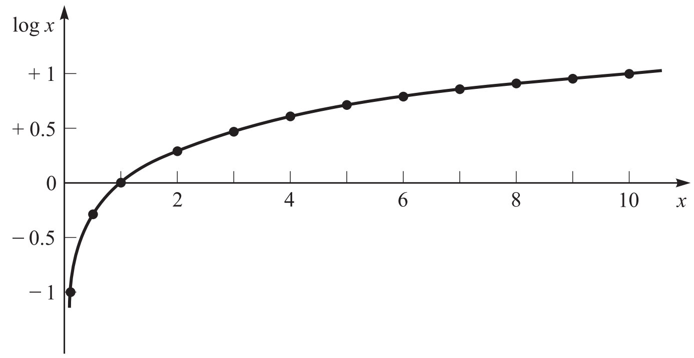


Fig. 26-A

We see that, as x increases in value, $\log x$ also increases in value. Note, however, that, as x increases in value, the *amount of* CHANGE *produced in* $\log x$ depends not only upon the amount of change in x but also upon the particular *value* of x; for example, from the table we see, for equal unit increases in the value of x, that

if x increases from	$\log x$ increases by the amount of
$x = 1$ to $x = 2$	0.30
$x = 2$ to $x = 3$	0.18
$x = 3$ to $x = 4$	0.12
$\vdots$	$\vdots$
$x = 9$ to $x = 10$	0.05

and so on, showing that the greater the *value* of x, the slower is the *rate of increase* in the value of $\log x$ with respect to x. This is evident from inspection of Fig. 26-A.

Let us now turn to the relationship between ACOUSTIC POWER ("sound power") and the sensation of LOUDNESS, as registered by the human ear.

It should be noted that the ear is capable of responding to an ENORMOUS RANGE of acoustical power; for example, in the case of a full symphony orchestra, the sound power produced during the loudest passages can be *10 million times* the sound power produced during the softest passages. The ear can handle such a tremendous range of sound power because, as the power *increases*, the sensitivity of the ear *automatically decreases*, so that the RATE OF INCREASE in the sensation of loudness decreases as the power increases, in the same manner that the *rate of increase* in the value of $\log x$ decreases as x increases, as shown in Fig. 26-A. Hence the effect that sound power has on the ear for two different power levels is *not* proportional to the power ratio itself but, instead, is approximately proportional to the LOGARITHM of the power ratio. It is this fact that led to the definition of eq. (315), and is responsible for the statement that "the ear hears logarithmically."

In closing this discussion, we should note that, for prescribed conditions, eq. (315) can be written in terms of a VOLTAGE RATIO. To show this, we first need to prove that if x is any positive number raised to a power n, then

$$\log x^n = n \log x \tag{9-A}$$

The truth of eq. 9-A can be established as follows.

$$\begin{array}{rl} \text{let} & \log x = y \\ \text{meaning that} & x = 10^y \\ \text{raise both sides to power } n & x^n = 10^{yn} \\ \text{by definition} & \log x^n = yn \\ \text{hence, since } y = \log x & \log x^n = n \ \log x \end{array}$$

thus proving eq. (9-A).

Let us now make use of eq. (9-A) as follows. Recall that, in a purely resistive circuit, POWER can be calculated by the equation V^2/R. Thus, for two voltages, V_1 and V_2, applied to a resistance of R ohms, the POWER RATIO is

$$P = \frac{V_1^2/R}{V_2^2/R} = (V_1/V_2)^2$$

and thus, substituting this value of P in eq. (315) then making use of eq. (9-A), eq. (315) becomes

$$\mathrm{dB} = 20 \ \log(V_1/V_2) \tag{10-A}$$

which, it should be noted, is true only if V_1 and V_2 are both applied to the *same value of resistance* of R ohms. Actually, however, in practice eq. (10-A) is often applied in cases where the two resistances are *not* equal; in such a case the results are not really in decibels but in what we could call "logarithmic units."

Note 20. Phase (Time-Delay) Distortion

It's been pointed out that frequency discrimination (frequency distortion) is produced by the presence of AMPLITUDE DISTORTION and TIME-DELAY (PHASE) DISTORTION.

First, in regard to "amplitude" distortion, it's clear that *no amplitude distortion* can occur if, in passing through a network, the *amplitudes* of all frequency components are multiplied by a fixed constant value k; that is, if all frequency components are treated the same, as far as amplitudes are concerned.

Let us, therefore, turn our attention to "time-delay distortion," as follows. Since there is *energy storage* associated with inductance and capacitance, it's understandable that *time* is required to change the state of energy level in these parameters. Because of this, a *time delay* exists between the input and output waves of voltage and current in a network. If such time delay is the *same for all frequency components*, then there is *no distortion* due to time delay. This is illustrated in Figs. 27-A and 28-A, in which the INPUT signal, Fig. 27-A, consists of fundamental, second-harmonic, and fourth-harmonic waves, having amplitudes and positions as shown, the independent variable being time, t.

Now let T denote the "time delay" between input and output waves, and suppose we have the *desired condition* in which *T is the same for all frequency components*, which is the condition illustrated in Fig. 28-A.

Comparison of the two figures makes clear that if there is no amplitude distortion, and if the TIME DELAY is the SAME FOR ALL FREQUENCIES, then the output wave will be delayed, relative to the input wave, by T seconds, but the basic WAVESHAPE of the output wave will be the SAME as that of the input wave.

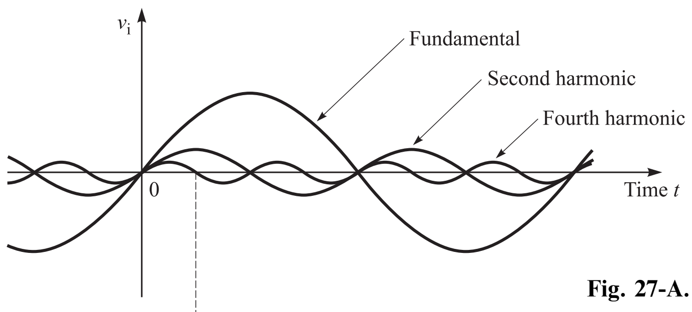


Fig. 27-A. INPUT wave.

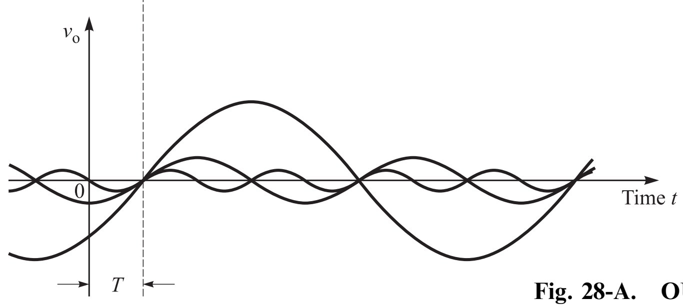


Fig. 28-A. OUTPUT wave.

As already noted, time-delay distortion is also called "phase" distortion. If we choose to talk in terms of phase distortion, the two conditions for distortionless transmission through a network are that (1) there be no amplitude distortion, and (2) a *linear* (first degree) relationship exist between phase shift and order of harmonic in the output wave. The meaning of this statement can be explained as follows.

We recall, from note 18, that any non-sinusoidal repetitive waveform can be expressed as the sum of a fundamental sinusoidal wave and its harmonics, in which the amplitudes of the harmonics decrease in a general way as the order of the harmonic increases. With this in mind, let v_i denote the instantaneous value of such a voltage waveform applied to the INPUT of a network, and let the Fourier series for v_i be of the form

$$v_i = a_1 \sin \omega t + a_2 \sin 2\omega t + \cdots + a_n \sin n\omega t + \cdots \tag{11-A}$$

where $a_1 \sin \omega t$ is the fundamental (lowest frequency) component, $a_2 \sin 2\omega t$ is the second-harmonic component, and so on, to any nth *harmonic component* of amplitude a_n and frequency $n\omega$.

Now suppose, in passing through the network, that all the *amplitudes* are multiplied by the same constant value k (thus no "amplitude" distortion), and that all the components are *delayed* by the *same amount* of T seconds. In such a case, upon setting $ka_1 = b_1, ka_2 = b_2$, and so on, and upon replacing t with $(t - T)$,* eq. (11-A) becomes the OUTPUT VOLTAGE of the network

$$v_o = b_1 \sin(\omega t - \omega T) + b_2 \sin(2\omega t - 2\omega T) + \cdots + b_n \sin(n\omega t - n\omega T) + \cdots \tag{12-A}$$

Equation (12-A) is true for the *ideal condition* in which all frequency components are delayed the *same amount of time*, T seconds, in passing through a network. Thus $a_n \sin n\omega t$, applied at the INPUT, appears at the OUTPUT as

$$b_n \sin(n\omega t - \phi_n) = b_n \sin(n\omega t - n\omega T)$$

* v_o thus LAGS v_i by T seconds, as in Fig. 28-A.

where $\phi_n = (n\omega)T$ = PHASE SHIFT in radians, where $(n\omega)$ = frequency, $(n = 1, 2, 3, \ldots)$. Thus we have the ratio

$$\frac{\phi_n}{(n\omega)} = T \tag{13-A}$$

which, since T is *constant* for all frequencies in the ideal case, shows that, in order to have *zero phase-shift distortion* (constant time delay), the RATIO of PHASE SHIFT TO FREQUENCY must have the same constant value T at all frequencies.

Note 21. Logarithmic Graph Paper

Ordinary "linear" graph paper is not generally suitable for plotting frequency response curves. One reason is that detection by the ear of a CHANGE IN FREQUENCY, $f_2 - f_1$, depends not only upon the value of $f_2 - f_1$ but also upon the values of f_2 and f_1 themselves.

Thus, for example, an increase in frequency from 50 Hz to 70 Hz produces a CHANGE of 20 Hz which, as experience shows, would *most definitely* be detected by the ear. On the other hand, a frequency change from 7000 Hz to 7020 Hz would *also* produce a change of 20 Hz which, however, as experience shows, the ear would scarcely detect, if at all. (The ear would, however, readily detect, for example, a change from 7000 Hz to 8000 Hz.) Hence, what is needed is a type of horizontal frequency axis in which EQUAL DISTANCES represent *larger and larger values* of $f_2 - f_1$. To show that a "logarithmic" scale meets this requirement, let us examine the horizontal axis in Fig. 189, as follows.

In Fig. 189 notice, for example, that a frequency change from *50 to 70 Hz* occupies a *distance* of 0.2 inch, a frequency change from *500 to 700 Hz* occupies the *same distance* of 0.2 inch, a frequency change from *5000 to 7000 Hz* occupies the *same distance* of 0.2 inch, and so on. Thus *equal segments of distance* on the axis have approximately EQUAL EFFECTS as far as the ear is concerned.

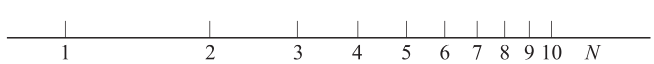
It's of interest to note that a "log" scale can be constructed using the formula

$$L = K \log N$$

where N = any *number* from 1 to 10, L = *distance* that the point representing N is to be located from the point where $N = 1$, and K = *scale factor*, which determines the physical size of the scale.

To illustrate, a log scale for $N = 1$ to $N = 10$, for a scale factor of $K = 4$ inches, would be constructed as follows (rounding calculator values to three decimal places):

for $N = 1$,	$L = 4 \log 1 = 0.000$	for $N = 2$,	$L = 4 \log 2 = 1.204$
for $N = 3$,	$L = 4 \log 3 = 1.905$	for $N = 4$,	$L = 4 \log 4 = 2.408$
for $N = 5$,	$L = 4 \log 5 = 2.796$	for $N = 6$,	$L = 4 \log 6 = 3.113$
for $N = 7$,	$L = 4 \log 7 = 3.380$	for $N = 8$,	$L = 4 \log 8 = 3.612$
for $N = 9$,	$L = 4 \log 9 = 3.817$	for $N = 10$,	$L = 4 \log 10 = 4.000$



The scale can be repeated as many times as needed, from $N = 10$ to 100, $N = 100$ to 1000, and so on, to as high a value of N as is needed.

Note 22. $\operatorname{Log} XY = \operatorname{Log} X + \operatorname{Log} Y$

Let

$$\log X = a \quad \textit{and} \quad \log Y = b$$

thus

$$X = 10^a \quad \textit{and} \quad Y = 10^b$$

so that

$$XY = 10^a 10^b = 10^{(a+b)}$$

By definition,

$$\log XY = (a+b) = \log X + \log Y, \text{ as stated.}$$

Note 23. Discussion of Eq. (344)

In terms of loop-current notation, the basic Fig. 202 becomes Fig. 29-A.

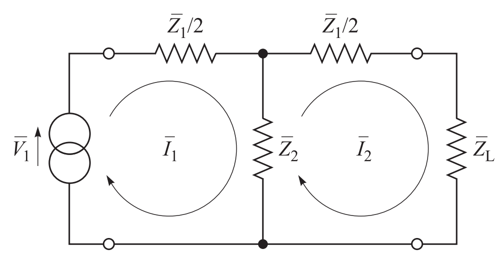


Fig. 29-A

By inspection,

$$(\bar{Z}_1/2 + \bar{Z}_2)\bar{I}_1 - \bar{Z}_2\bar{I}_2 = \bar{V}_1$$
$$-\bar{Z}_2\bar{I}_1 + (\bar{Z}_2 + \bar{Z}_1/2 + \bar{Z}_L)\bar{I}_2 = 0$$

$$\bar{I}_1 = \frac{\begin{vmatrix} \bar{V}_1 & -\bar{Z}_2 \\ 0 & (\bar{Z}_2 + \bar{Z}_1/2 + \bar{Z}_L) \end{vmatrix}}{\Delta} = (\bar{Z}_2 + \bar{Z}_1/2 + \bar{Z}_L)\bar{V}_1/\Delta$$

$$\bar{I}_2 = \frac{\begin{vmatrix} (\bar{Z}_1/2 + \bar{Z}_2) & \bar{V}_1 \\ -\bar{Z}_2 & 0 \end{vmatrix}}{\Delta} = \bar{Z}_2\bar{V}_1/\Delta$$

thus

$$\frac{1}{\bar{I}_2} = \frac{\Delta}{\bar{Z}_2\bar{V}_1}$$

hence

$$\frac{\bar{I}_1}{\bar{I}_2} = \frac{\bar{Z}_2 + \bar{Z}_1/2 + \bar{Z}_L}{\bar{Z}_2} = 1 + \frac{\bar{Z}_1}{2\bar{Z}_2} + \frac{\bar{Z}_L}{\bar{Z}_2}$$

which is eq. (344).

Note 24. Amplitude Modulation. Sidebands

Let us take "amplitude modulation" as an example. In amplitude modulation, the amplitude of the high-frequency carrier wave is varied or "modulated" in accordance with audio, video, or other type of signal information to be transmitted. The general principle of amplitude modulation, AM, will be clear from a study of Fig. 30-A, which let us now concentrate on, as follows.

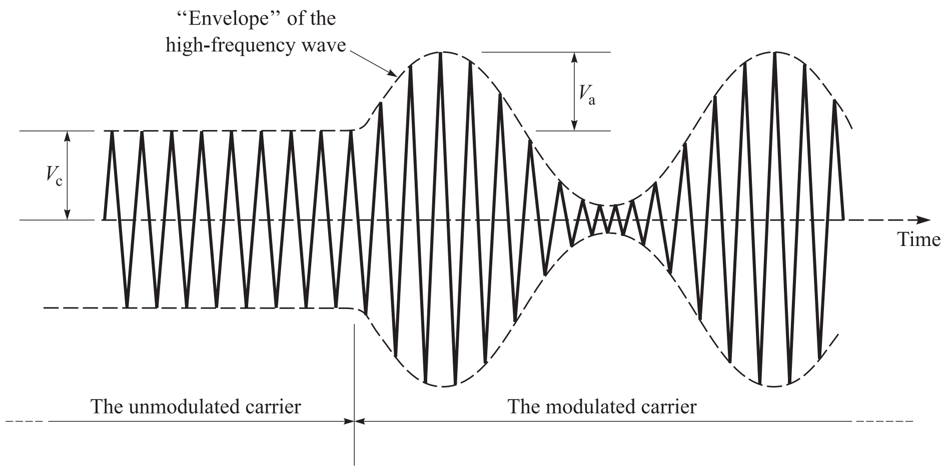


Fig. 30-A

To the left, in the figure, is shown a portion of an *unmodulated* high-frequency carrier wave. (For convenience in drawing, the carrier is shown as a triangular wave, but it will actually be a sinusoidal wave.) Note that the peak value of the unmodulated carrier wave is denoted by V_c.

Now suppose the carrier is AMPLITUDE MODULATED by, say, a sinusoidal wave of audio frequency voltage. The appearance of the resulting wave would then be such as shown to the right in the figure. The dashed line connecting the peaks of the carrier wave is called the "envelope" of the wave. For distortionless modulation the envelope must have the *same waveshape as the modulating voltage*.

Now let V_a be the peak value of the modulated component of the carrier, as indicated in the figure. Then the ratio of V_a to V_c is called the "modulation factor" and is denoted by m; thus

$$\frac{V_a}{V_c} = m \tag{14-A}$$

In the discussion here we'll assume a sinusoidal carrier wave and sinusoidal modulating signal, and use the notation

$$\omega_c = \text{carrier frequency in radians/second, and}$$
$$\omega_a = \text{frequency of modulating signal, rad/sec.}$$

hence

$$v_c = V_c \sin \omega_c t = \text{instantaneous value of unmodulated carrier, and}$$
$$v_a = V_a \sin \omega_a t = \text{instantaneous value of modulation envelope}$$

or, by eq. (14-A)

$$v_a = mV_c \sin \omega_a t$$

After a detailed consideration of the above we come to the conclusion that the instantaneous value v of the resulting *amplitude-modulated wave* must be equal to

$$v = V_c(1 + m \sin \omega_a t) \sin \omega_c t$$

Note that this equation satisfies the requirement that if $m = 0$ (the condition of no modulation), then all that is left is the unmodulated carrier, $V_c \sin \omega_c t$. Next, upon multiplying as indicated, the equation becomes

$$v = V_c \sin \omega_c t + mV_c \sin \omega_c t \sin \omega_a t$$

Now, in the trigonometrical identity for $\sin x \sin y$ (see note 25), set $x = \omega_c t$ and $y = \omega_a t$. Upon doing this, the last equation becomes the very important result that

$$v = V_c \sin \omega_c t + \frac{mV_c}{2} \cos(\omega_c - \omega_a)t - \frac{mV_c}{2} \cos(\omega_c + \omega_a)t \tag{15-A}$$

The equation brings out the important fact that, when a high-frequency sinusoidal carrier of frequency $\omega_c = 2\pi f_c$ is amplitude-modulated by a single sinusoidal signal of frequency $\omega_a = 2\pi f_a$ the resulting amplitude-modulated wave is composed of *three component sinusoidal waves*, thus

1. the CARRIER wave, of frequency $\omega_c = 2\pi f_c$,
2. the LOWER SIDEBAND wave, of frequency $(\omega_c - \omega_a) = 2\pi(f_c - f_a)$,
3. the UPPER SIDEBAND wave, of frequency $(\omega_c + \omega_a) = 2\pi(f_c + f_a)$.

For several practical reasons the carrier must be a high-frequency (''radio-frequency'') wave, much higher in frequency than the frequency of the modulating signal; that is, f_c must be much higher than f_a. Thus the sidebands $(f_c \pm f_a)$ are *also* high-frequency waves, centered around the carrier wave. If the carrier is being modulated by a *non-sinusoidal wave*, as would normally be the case, then *a pair of sidebands exists for each harmonic of the modulating wave*. Thus the carrier wave is at the center of a cluster of high-frequency sideband waves, as indicated in Fig. 31-A, in which V_c is the rms value of the carrier wave. A brief but fundamental discussion follows the figure.

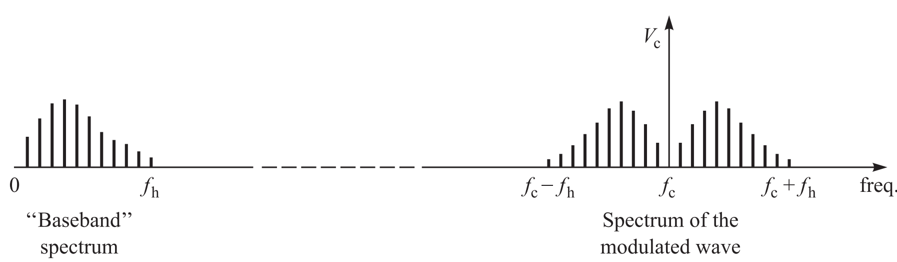


Fig. 31-A

We should first explain that the term ''baseband'' spectrum refers to the original location of the spectrum of a signal.

The baseband spectrum will normally extend from near $f = 0$ to $f = f_h$, where f_h is the highest frequency component in the signal. The shape of a possible baseband spectrum is illustrated in Fig. 31-A. The baseband of an audio signal, as produced by a microphone, is generally taken to extend from approximately 16 Hz to approximately 16,000 Hz.

In a simple telephone system, the voice signal is transmitted in its original baseband form, over copper wires, from transmitter to receiver. Such a case does not involve modulation.

If, however, the same voice signal is to be transmitted by wireless, then some kind of modulation of a high-frequency carrier wave is required.

Regardless of the type of modulation used, the result is always the production of a band of side-band frequencies, clustered symmetrically about the carrier. Since the original signal information is contained in this band of frequencies, it's desirable that the entire band be passed through circuits having a reasonably good band-pass characteristic, such as is illustrated in Fig. 240.

Note 25. Trigonometric Identity for ($\sin x \ \sin y$)

First, by *note 6*:

$$\cos(x+y) = \cos x \cos y - \sin x \sin y \tag{I}$$

In the above equation, replace y with $-y$; then, since (section 5.3) $\cos(-y) = \cos y$ and $\sin(-y) = -\sin y$, eq. (I) becomes

$$\cos(x-y) = \cos x \cos y + \sin x \sin y$$

hence

$$-\cos(x-y) = -\cos x \cos y - \sin x \sin y \tag{II}$$

Lastly, addition of eqs. (I) and (II) gives the identity we are after, thus

$$\sin x \sin y = \tfrac{1}{2}[\cos(x-y) - \cos(x+y)]$$

Note 26. L Proportional to N^2

Consider a coil of TWO TURNS in which each turn, considered by itself, has L henrys of inductance. If there were NO COUPLING WHATEVER between the two turns the total inductance L_t would simply be the *sum* of the inductances of the individual turns; thus

$$L_t = L + L = 2L \tag{16-A}$$

If, however, some amount of coupling does exist between the two turns, then, as we found in problem 198, the total inductance would be equal to

$$L_t = L + L + 2M$$

or, by eq. (371)

$$L_t = L + L + 2kL$$

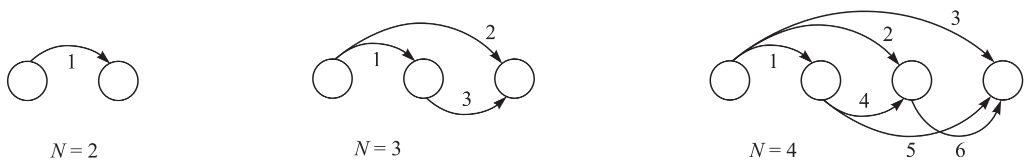
Hence, for the *ideal case* in which there is COMPLETE COUPLING between the two turns ($k = 1$), eq. (16-A) becomes

$$L_t = L + L + 2L \tag{17-A}$$

Now consider a coil of N turns in which each turn, considered by itself, has the same inductance of L henrys. If there were NO COUPLING WHATEVER between any turn and any other turn, the *total inductance* L_t would simply be the *sum* of the individual inductances of the individual turns; thus

$$L_t = L + L + L + L + \cdots + L = NL \tag{18-A}$$

Now imagine an inductor coil of N turns, of L henrys each, in which 100% COUPLING EXISTS BETWEEN EACH TURN AND ALL THE OTHER TURNS; this would constitute a true "ideal" inductor and, in such a case, the amount of "$2L$" would have to be added to eq. (18-A) FOR EVERY POSSIBLE COMBINATION OF TWO TURNS in the coil. Thus, for example, $2L$ would be added *one time* for $N = 2$, *three times* for $N = 3$, *six times* for $N = 4$, *ten times* for $N = 5$, and so on; this is illustrated in the figures below for $N = 2, N = 3$, and $N = 4$.



Thus eq. (18-A) would become

$$\begin{aligned}
\text{for } N = 2, \quad L_t &= L + L + L + 2L = 4L = (2)^2 L \\
\text{for } N = 3, \quad L_t &= L + L + L + 2L + 2L + 2L = 9L = (3)^2 L \\
\text{for } N = 4, \quad L_t &= L + L + L + L + 2L + 2L + 2L + 2L + 2L + 2L \\
&= 16L = (4)^2 L \\
\text{for } N = 5, \quad L_t &= 5L + 10(2L) = 25L = (5)^2 L
\end{aligned}$$

and upon continuing on in this manner it soon becomes evident that an ideal inductor of N turns has a total inductance of

$$L_t = N^2 L = kN^2, \text{ as stated,}$$

where $L =$ inductance/turn. (It should be mentioned that the above result can also be derived by direct application of the formula for the combination of N things taken two at a time.)

Note 27. Arrow and Double-Subscript Notation

Network equations are written in accordance with Kirchhoff's voltage and current laws. The voltage law, for example, says that around any closed loop in a network the "algebraic sum" of the voltage drops and generator voltages must be equal to *zero*; thus

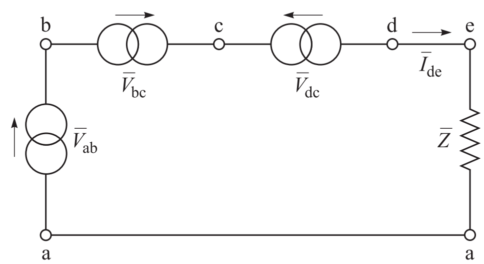
$$(\text{algebraic sum of voltage drops}) + (\text{algebraic sum of generator voltages}) = 0$$

so that a plus or a minus sign must be attached to each quantity in the equation for each loop of a network. Actually, however, in writing network equations we'll adhere to the THREE RULES stated in section 4.4, and write that

$$(\text{algebraic sum of voltage drops}) = (\text{algebraic sum of generator voltages})$$

in which we must remember that the signs are written in accordance with the rules given in that section (except that now, in using vector algebra, we'll write $\bar{Z}$, $\bar{V}$, and $\bar{I}$ instead of R, V, and I—see discussion given with Fig. 135 in Chap. 8).

In regard to the above-mentioned signs, it should be noted that if a network contains just *one* generator, then no serious problem exists. If, however, a network contains TWO OR MORE GENERATORS then it is absolutely necessary that some method be used to indicate the DIRECTIONS of the voltage vectors. One such method uses the "double-subscript" notation, in which the positive sense of the vector quantity is denoted by the *order* in which the subscripts are written. This is illustrated in the series circuit below (in which, for comparison, the familiar "arrow" notation is also shown).



Thus the notation $\bar{V}_{ab}$ says there is an *increase* in vector voltage in going through that particular generator in the direction *from a to b*, while $\bar{V}_{dc}$ says there is an increase or rise in voltage in going through that generator *from d to c*. (Note that the voltage "arrows" can be used to give us the same information.) In regard to the current notation, $\bar{I}_{de}$ says that the vector current is positive in the direction of *d to e*.

Hence, using double-subscript notation, and with the previously-mentioned rules in mind, the equation for the above figure can be written as (going around in the cw sense)

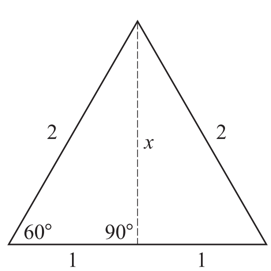
$$\bar{I}_{de}\bar{Z} = \bar{V}_{ab} + \bar{V}_{bc} - \bar{V}_{dc}$$

Note 28. Square Root of 3 in Three-Phase Work

The factor $\sqrt{3}$ is prominent in three-phase work; our object here is to show how this comes about. To do this, let us note that the factor first appears in eq. (424), because

$$\sin 120^\circ = \sin 60^\circ = \sqrt{3}/2 = 0.8660$$

a fact we can establish with the aid of the equilateral triangle shown below.



For convenience, we've let the sides of the triangle be 2 units in length, as shown. Hence, by the Pythagorean theorem,

$$x^2 + 1 = 4$$

so that

$$x = \sqrt{3}$$

and therefore, by definition, we have

$$\sin 120^\circ = \sin 60^\circ = x/2 = \sqrt{3}/2 = 0.8660$$

thus the factor $\sqrt{3}$ fundamentally appears in three-phase work.

Note 29. Proof of Eq. (467) (True Power)

Let

$$\bar{V} = V\underline{/a} = V\epsilon^{ja} = \text{generator voltage}$$

and let

$$\bar{I} = I\underline{/b} = I\epsilon^{jb} = \text{generator current}$$

hence $(a - b)$ = angle between $\bar{V}$ and $\bar{I}$, and hence, by eq. (117) in Chap. 5,

$$P = VI\ \cos(a - b) = \text{true power produced by the generator}$$

Now note that

$$\bar{\bar{I}} = I\underline{/-b} = I\epsilon^{-jb}$$

and therefore

$$\bar{V}\bar{\bar{I}} = VI\epsilon^{ja}\epsilon^{-jb} = VI\epsilon^{j(a-b)}$$

hence

$$\bar{V}\bar{\bar{I}} = VI\ [\cos(a - b) + j\ \sin(a - b)]$$

showing that the REAL PART of the product $\bar{V}\bar{\bar{I}}$ is equal to the TRUE POWER produced by the generator, as eq. (467) states.

Note 30. The Transistor as Amplifier

In general, in electronics, an "amplifier" is a circuit having an input signal and an output signal, in which the output signal is a reasonably good replica of the input signal and in which the POWER of the output signal is greater than the power of the input signal.

An amplifier basically consists of a SOURCE OF DC POWER, an OUTPUT LOAD IMPEDANCE Z_L, and a CONTROL DEVICE capable of controlling the instantaneous output current of the dc power source. This is illustrated in block-diagram form in Fig. 32-A.

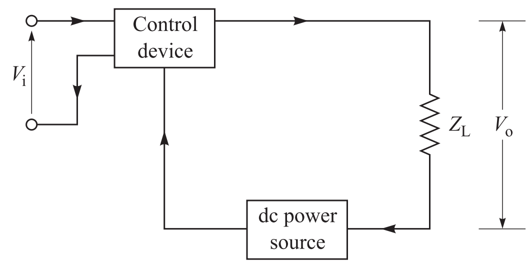


Fig. 32-A

In the above, the input ac signal component V_i, on the left, controls, by means of the "control device," the flow of power from the dc power source; thus the POWER of the output ac signal component V_o across the output impedance Z_L, can be greater than the power of the input signal.

It should be noted that the total voltage across Z_L actually consists of a dc component plus an ac component; it is, however, the *useful, information-carrying ac signal component* that we are interested in. Since, in our work, we deal with steady-state sinusoidal conditions, it follows that, in the above, V_i and V_o denote rms values of sinusoidal voltages.

Let us here assume the control device to be a *transistor*. A transistor is a solid-state control device having three terminals, called the *base* (B), the *collector* (C), and the *emitter* (E), as shown in schematic form in Fig. 33-A.

In Fig. 34-A, a transistor is shown connected in what is called the *common-emitter* (CE) mode. The designation "common-emitter" mode is appropriate because, as you can see, the emitter is common to both the input and output circuits. It should be noted that in Fig. 34-A we've omitted all dc voltages and currents and show only the ac signal components of voltage and current.

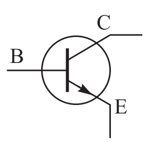


Fig. 33-A

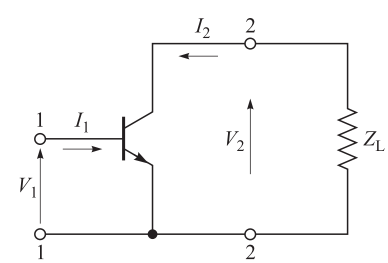


Fig. 34-A

Inspection of Fig. 34-A shows that, as far as ac signal components are concerned, a *transistor* can be considered to be a two-port network in the standard form of Fig. 277. Hence, as we showed in section 11.6, the "impedance matrix" representation of a transistor is of the form

$$[\mathbf{Z}] = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$$

Let us now examine, in more detail, the operation of the basic CE circuit. To do this we'll make use of Figs. 35-A and 36-A, in which we've taken the load impedance to be a pure resistance of R_L ohms, as shown. Our OBJECT now is to show that AC OUTPUT VOLTAGE in a CE circuit with resistive load is *180 degrees out of phase with ac input voltage*.

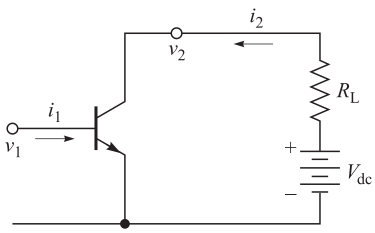


Fig. 35-A

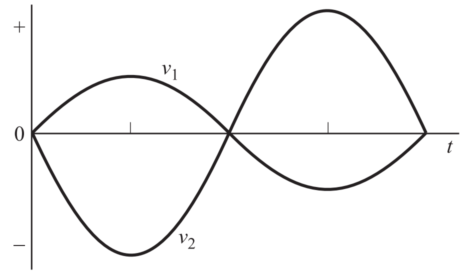


Fig. 36-A

First let us agree, in Fig. 35-A, to take the *emitter* E as being the zero-voltage "reference point" in the circuit.

Now consider some instant of time at which the input voltage v_1 becomes MORE POSITIVE with respect to E; this causes increased current i_1 to flow into the base.

This, however (as a result of the fundamental physics of transistor action), causes an *increase* of output current i_2 to flow into the collector, as shown in the figure; this increase of current through R_L causes an *increase* in the voltage drop across R_L, thus making the collector become LESS POSITIVE with respect to the emiter. Thus, as v_1 becomes MORE POSITIVE, v_2 becomes LESS POSITIVE.

On the other hand, as v_1 becomes LESS POSITIVE (with respect to E), then i_1 and i_2 both decrease, thus causing the voltage drop across R_L to decrease, thus causing the collector voltage to become MORE POSITIVE with respect to E. Hence, as v_1 become LESS POSITIVE v_2 becomes MORE POSITIVE.

The foregoing *variations* in voltages constitute *the useful ac signal*; thus, in terms of sinusoidal waves, we have that the OUTPUT SIGNAL, produced by the transistor operating in the CE mode into a resistive load, is 180° OUT OF PHASE WITH THE INPUT SIGNAL.

As a last comment, note that, in Fig. 35-A, the required source of dc power is supplied by a battery of V_{dc} volts. Thus there is present, at the collector of the transistor, a *dc* component of voltage as well as the useful ac signal component. Hence the output of the amplifier, at the collector, is usually then passed through a simple *RC* high-pass filter, passing the ac component but not the zero-frequency dc component. The final result is illustrated in Fig. 36-A, where v_2 is the ac output signal (shown to be 180° out of phase with the input signal v_1).

Note 31. Shifting Theorem

Here we wish to establish a relationship called the "shifting theorem," which can be done with the aid of Fig. 37-A.

In the figure, let t be time *measured from the origin*, as shown. Let curve A be a portion of the curve of some function $y = f(t)$. Now let it be given that curves A and B are identical *except* that curve B is shifted horizontally T units to the right of curve A, as shown in the figure. That is, curve B *lags T seconds behind curve A*; notice that curve A has the value y_0 at $t = 0$, but curve B does not reach that value until T seconds later.

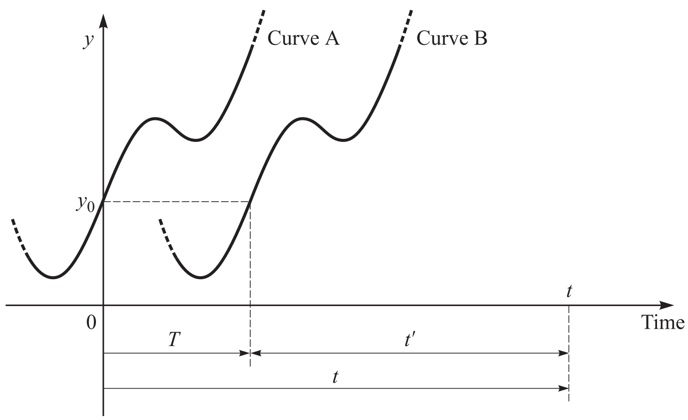


Fig. 37-A

Now let t' ("t prime") be time *measured from the instant* $t = T$, as shown in the figure. Thus $t' = 0$ when $t = T$. If we use t for time in curve A, and t' for time in curve B, then the *equations* for curves A and B will be *identical*; thus

$$y = f(t) \quad \text{for curve A}$$
$$y = f(t') \quad \text{for curve B}$$

From the figure note that $t = T + t'$, so that $t' = t - T$. Using this relation, we can write the equations of both curves *in terms of t*; thus

$$y = f(t) \qquad \text{equation of curve A}$$
$$y = f(t - T) \quad \text{equation of curve B}$$

What we've tried to demonstrate here is called the "shifting theorem," which can be summarized as follows.

> For any function $f(t)$, the substitution of $t - T$ *in place of t* has the effect of shifting the curve of $f(t)$ horizontally T units to the right.

That is, the curve of $f(t - T)$ is exactly the same as the curve of $f(t)$ except that it is *shifted T units to the right*, as illustrated in Fig. 37-A.

Note: If, in Fig. 37-A, curve B had been drawn to the left of curve A, then, $t' = t + T$, showing that substitution of "$t + T$" in place of "t" in $f(t)$ has the effect of shifting the curve of $f(t)T$ units to the *left* of its original position.

Note 32. Unit Impulse

Here we wish to introduce the very useful concept of "unit impulse." Let us begin with drawings A, B, and C in Fig. 38-A.

As illustrated, the particular form of pulse we'll be interested in here will always have a *time duration* of a seconds and a *value* equal to $1/a$; thus, since, $a(1/a) = 1$, such a pulse will always enclose *unit area*, as shown in the figure.

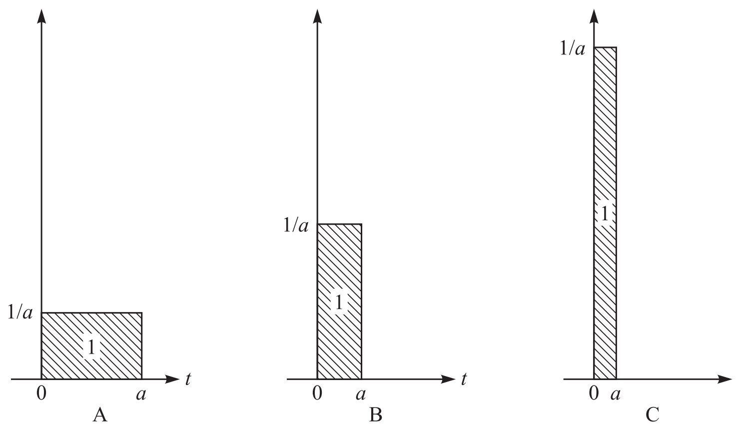


Fig. 38-A

Now, in C, let us allow a to approach *zero* as a limiting value ($a \to 0$) which, in turn, will make the *value* of the pulse, $1/a$, become "infinitely great" ($1/a \to \infty$), with the *area* enclosed by the pulse always remaining equal to 1.

In words, we are hypothesizing, at $t = 0$, the existence of a pulse of infinitely great amplitude but vanishingly short time duration, the pulse always enclosing unit area.

The hypothetical pulse so described is called a UNIT IMPULSE, and is denoted by the symbol $\delta(t)$, which can be read as "delta of t" ("δ" is the small Greek letter "delta").

Such a pulse cannot, of course, exist in the real physical world. It is, nevertheless, a very useful mathematical device, for the following reasons.

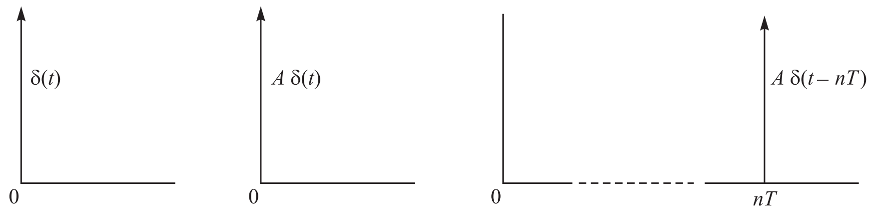
First of all, very short, high-valued pulses of voltage and current *do* exist in the real world, and such actual pulses, when applied to a network, have the same general effect as would the application of a theoretical impulse to the network. In other words, the theoretical analysis of a network to applied $\delta(t)$ yields results that will closely approximate the actual results produced by the application of a very high, "sharp," pulse to the network.

A second reason lies in the fact that a function of time t can be expressed in terms of a particular secondary variable "s," where s is a complex number of the form $s = a + j\omega$. This is important, because the work required in circuit analysis can often be greatly reduced when carried out in terms of s instead of t. This is especially true if we're investigating the effect of applying an impulse type of signal to a network. This is because it turns out that $\delta(t)$ is simply replaced by "1" when working in terms of s, a fact that can considerably reduce the algebraic complications in impulse-type problems.

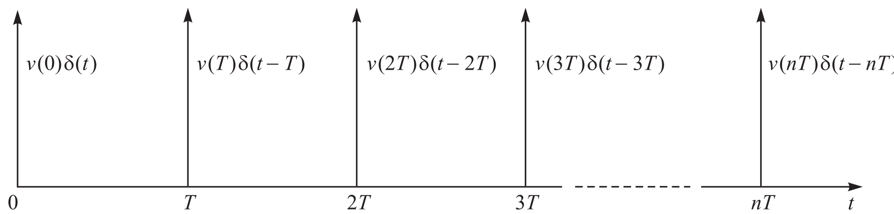
Lastly, $\delta(t)$ is a convenient symbol to use in writing "sampling equations," which let us discuss in more detail, as follows.

First, it will be convenient to refer to $\delta(t)$ as an impulse of "unit strength." Then the notation "$A\delta(t)$" will naturally be called an impulse of "strength A." Hence, by note 31, the notation $A\delta(t - nT)$ will denote an impulse of strength A "shifted nT units to the right."

GRAPHICALLY, an impulse is represented by a vertical line with arrowhead, with notation given alongside the line. Thus, graphically, the three impulse cases just mentioned above would appear as illustrated in the following.



Thus Fig. 331, in the main text, can be described in algebraic form by eq. (571) or in graphical form as below, starting at $t = 0$.



Note: Impulse "strength" is also called impulse "weight." Thus the above sequence is also referred to as a sequence of "weighted impulses."

In conclusion, let us note that it is *not* correct to write that $\delta(t) = 1$ for $t = 0$, or that $\delta(t - nT) = 1$ for $t = nT$. All we say is that $\delta(t)$ "exists" only for $t = 0$, and that $\delta(t) = 0$ for all other values of t. Thus "impulse notation," as we're using it here, is useful in the mathematical description of impulse-type sampled signals.

The term "unit impulse" refers to the fact that $\delta(t)$ is defined to enclose "unit area" and it is this conception that leads to very useful results when, in the calculus, the process of integration is applied to the study of sampled signals. Right now, however, we'll just view $\delta(t)$ and $\delta(t - nT)$ as useful shorthand notations.

Note 33. Algebraic Long Division

Basic terminology: $\dfrac{\text{dividend}}{\text{divisor}} = \dfrac{\text{numerator}}{\text{denominator}} = \text{quotient}.$

If the numerator and denominator are both algebraic polynomials in x, a useful procedure to find the quotient can be summarized as follows.

1. Arrange both numerator and denominator in descending powers of x.*
2. DIVIDE the FIRST TERM OF THE NUMERATOR by the FIRST TERM OF THE DENOMINATOR.
3. Now MULTIPLY the ENTIRE DENOMINATOR by the result of step (2), then SUBTRACT the result from the numerator.
4. Now consider the result of step (3) as being a "new numerator," and repeat steps (2) and (3).

* To keep track of the work it's helpful to write in any missing powers of x with "zero coefficients." For example, $x^3 - x + 1$ would be written as $x^3 + 0x^2 - x + 1$.

The following examples will help you to check your understanding of the above four steps. It should be noted that it's generally not necessary, or even desirable, to apply long division to a given algebraic fraction; it depends upon the particular situation, such as, here, finding an inverse z-transform.

It should also be noted that algebraic fractions are classified as being "proper" or "improper" as follows. If the *highest power of x* is located in the *denominator* the fraction is said to be "proper," but if this is not true the fraction is said to be "improper." Both types will appear in the following examples, with suitable comments.

Example 1

Write the improper algebraic fraction

$$\frac{5x^3 + 13x^2 + 2x + 2}{x + 2}$$

in a form that contains only a proper fraction.

Solution

This can be done by using algebraic long division, as follows. First, for this operation, let us begin by writing the indicated division in the more convenient form

$$x + 2 \,\overline{\big)\, 5x^3 + 13x^2 + 2x + 2}$$

where $x + 2$ is the divisor. Now carefully follow the prescribed procedure until the "new numerator" *becomes free of the variable x*. The detailed results are as follows.

$$\begin{array}{r|l}
 & 5x^2 + 3x - 4 \\
\hline
x + 2 & 5x^3 + 13x^2 + 2x + 2 \\
 & \underline{-5x^3 - 10x^2} \\
 & \qquad 3x^2 + 2x \\
 & \qquad \underline{-3x^2 - 6x} \\
 & \qquad\qquad -4x + 2 \\
 & \qquad\qquad \underline{+4x + 8} \\
 & \qquad\qquad\qquad +10
\end{array}$$

In the above result "10" is called the "remainder," and all we can do is write $10/(x + 2)$ to indicate that the remainder, 10, is to be divided by the divisor, $x + 2$. Thus the final answer is

$$\frac{5x^3 + 13x^2 + 2x + 2}{x + 2} = 5x^2 + 3x - 4 + \frac{10}{x + 2}$$

Since the only fraction, $\frac{10}{x + 2}$, is a proper fraction, the problem requirement is met.

Example 2

Apply algebraic long division to the proper algebraic fraction $\frac{1}{x^2 - 2}$.

Solution
We can begin by writing the indicated division in the convenient form $x^2 - 2 \overline{)\,1}$, then applying the "four-step" procedure, as follows.

First, by step (2),

$$\frac{1}{x^2} = x^{-2}$$

then, applying step (3), the work appears as

$$\begin{array}{r|l} & \;\; x^{-2} \\ \cline{2-2} x^2 - 2 & \;\; 1 \\ & \underline{-1 + 2x^{-2}} \\ & \qquad 2x^{-2} \end{array}$$

The "new numerator" is thus $2x^{-2}$, to which, by step (4), we must again apply steps (2) and (3). Doing this, the work now appears as

$$\begin{array}{r|l} & \;\; x^{-2} + 2x^{-4} \\ \cline{2-2} x^2 - 2 & \;\; 1 \\ & \underline{-1 + 2x^{-2}} \\ & \qquad 2x^{-2} \\ & \qquad \underline{-2x^{-2} + 4x^{-4}} \\ & \qquad\qquad\quad 4x^{-4} \end{array}$$

The "new numerator" is now $4x^{-4}$, to which, by step (4), we would again apply steps (2) and (3). How long we continue on in this way depends upon the particular problem and the range of values of x to be encountered in the problem. For instance, if we continue on and take two more steps in the above work we have the result that

$$\frac{1}{x^2 - 2} = x^{-2} + 2x^{-4} + 4x^{-6} + 8x^{-8} + \left(\frac{16}{x^8(x^2 - 2)}\right)$$

where the last term to the right, in the large parentheses, is the "remainder term," corresponding to the term $10/(x + 2)$ in example 1. *If*, in applying the above expression, the *range of values of* x is such that the value of the remainder term is very small (compared with the sum of the first four terms), then we can drop the remainder term and write that, for practical purposes,

$$\frac{1}{x^2 - 1} = x^{-2} + 2x^{-4} + 4x^{-6} + 8x^{-8}$$

The above considerations are understood to apply in the application of "algebraic long division" in section 13.4.

Solutions to Problems

1. Setting $q_1 = q_2 = 6 \times 10^{-6}$ and $r = 0.25$ in eq. (3) we have

$$F = 9 \times 10^9 \times 36 \times 10^{-12}/0.0625 = 5.184 \text{ newtons}$$
$$= (5.184)(0.2248) = 1.165 \text{ pounds, } \textit{answer}.$$

2. Let us use the method of superposition, in connection with Fig. 367.

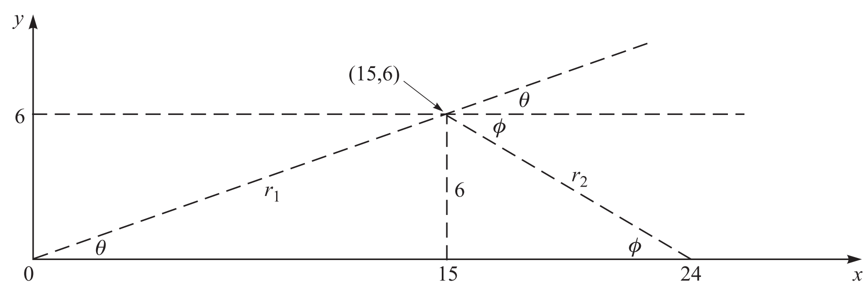


Fig. 367

From the figure we have

$$r_1 = \sqrt{261} \text{ and } \theta = \arctan(6/15) = 21.8014^\circ$$

and

$$r_2 = \sqrt{117} \text{ and } \phi = \arctan(6/9) = 33.6901^\circ$$

Let $q_1 = (3)10^{-6}$ coulombs at the origin, and let $\bar{E}_1$ be the field strength at the point (15, 6) due to q_1. Now setting $q = q_1 = (3)10^{-6}$ and $r^2 = 261$ in eq. (6) gives

$$\bar{E}_1 = 103.448\underline{/21.8014^\circ}$$

where, since the force between q_1 and q_0 is a force of repulsion, the angle $\theta = 21.8014^\circ$ is relative to the positive direction of the x-axis, as inspection of Fig. 367 shows.

We then have that

$$\text{horizontal component of } \bar{E}_1 = 103.448 \cos 21.8014^\circ = 96.0491$$

$$\text{vertical component of } \bar{E}_1 = 103.448 \sin 21.8014^\circ = 38.4196$$

Next, let $q_2 = (-2)10^{-6}$ coulombs at the point $x = 24$, and let $\bar{E}_2$ be the field strength at the point $(15, 6)$ due to q_2. Setting $q = q_2 = (-2)10^{-6}$ and $r^2 = 117$ in eq. (6) gives

$$\bar{E}_2 = 153.846\underline{/-33.6901^\circ}$$

in which there is a force of attraction between q_0 and q_2 and thus the force vector acting on q_0 points toward q_2, resulting in a negative value of ϕ relative to the x-axis, as can be seen from inspection of the figure. We then have that

$$\text{horizontal component of } \bar{E}_2 = 153.846 \cos(-33.6901^\circ) = 128.0076$$

$$\text{vertical component of } \bar{E}_2 = 153.846 \sin(-33.6901^\circ) = -85.3385$$

Now, letting $\bar{E}$ be the resultant field strength at the point $(15, 6)$, we have

$$\text{horizontal component of } \bar{E} = 96.0491 + 128.0076 = 224.0567$$

$$\text{vertical component of } \bar{E} = 38.4196 - 85.3385 = -46.9189$$

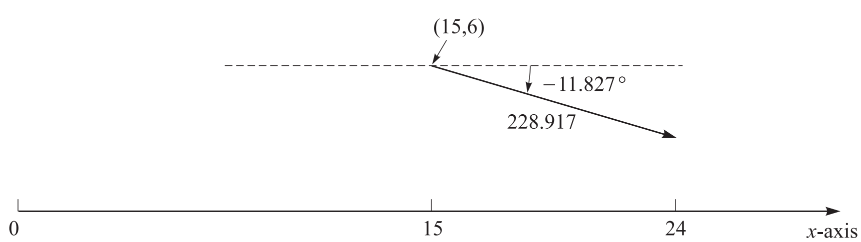
Therefore,

$$\bar{E} = \sqrt{52402.79}\,\underline{/-\arctan(46.9189/224.0567)}$$

hence,

$$\bar{E} = 228.917\underline{/-11.827^\circ} \text{ newtons/coulomb, } \textit{answer}$$

shown graphically in the following diagram.



3. From eq. (8), $qV = W =$ joules, *answer*.

4. By eq. (8),

$$V = \frac{W}{q} = \frac{2.65}{0.0078} = 339.74 \text{ volts, } \textit{answer}.$$

5. The answer is ''no,'' because, in order to calculate the potential difference between a and b, the field strength must be known at ALL POINTS along any path connecting a and b and NOT at just the end points of the path. (Such calculations, however, need not concern us here, because potential difference is normally a known quantity in practical work.)

More basically, the equation for $\bar{E}$ must be given before the potential difference between two points can be calculated. This is of little importance, however, because potential difference is generally the known quantity in practical work.

6. If n = number of electrons, then $n(1.602)10^{-19} = 1$ (1 coulomb), and thus

$$n = \frac{1}{1.602 \times 10^{-19}} = \frac{10^{19}}{1.602} = 6.24 \times 10^{18} \text{ electrons, } \textit{answer}.$$

7. In eq. (9), for this problem, $\Delta q = 9 \times 10^6 \times 1.602 \times 10^{-19} = 1.44 \times 10^{-12}$ coulombs, and $\Delta t = 10^{-7}$ seconds; thus

$$\frac{\Delta q}{\Delta t} = \frac{1.44 \times 10^{-12}}{10^{-7}} = 1.44 \times 10^{-5} \text{ amperes, } \textit{answer}, \text{ or}$$

$$= 14.4 \text{ microamperes, } \textit{answer}.$$

8. By eq. (11), $I = 48/6 = 8$ amperes, *answer*.

9. By eq. (15), $(48)(8) = 384$ watts, *answer*.
By eq. (16), $(48)^2(6) = 384$ watts, *answer*.
By eq. (17), $(8)^2(6) = 384$ watts, *answer*.

10. By eq. (17), $I = \sqrt{P/R} = \sqrt{18/75} = 0.4899$ amperes, *answer*.

11. By eq. (17), $P = (1.86)^2(25) = 86.49$ watts = 86.49 joules per second, and therefore $86.49/4.186 = 20.6617$ calories/second, *answer*.

12. The problem is to calculate the values to put into eq. (18), beginning with the length L, as follows. Manipulating "units" like algebraic quantities (as discussed just prior to problem 3 in Chap. 1), and noting that 1 foot = 12 inches and 1 meter = 39.370 inches, we have, for the value of L

$$L = 450 \text{ feet} = \frac{(450 \text{ feet})(12 \text{ inches})(1 \text{ meter})}{(1 \text{ foot})(39.370 \text{ inches})} = 137.160 \text{ meters}$$

Next, the *radius* is

$$0.25 \text{ inch} = \frac{(0.25 \text{ inch}(1 \text{ meter})}{(39.370 \text{ inches})} = (6.35)10^{-3} \text{ meters;}$$

hence,

$$A = \pi(6.35)^2 10^{-6} = (1.2668)10^{-4} \text{ square meters}$$

Next, noting that for this problem $T = 86°\text{F} = 30°\text{C}$, we have, using eq. (19),

$$\rho = (2.83)(10^{-8})[1 + (4.03)(10^{-3})(10)] = (2.94405)10^{-8}$$

Upon substituting the above values of L, A, and ρ into eq. (18) you should find that $R = 0.03188$ ohms, *answer*.

13. Note that only the temperature T is to change in this problem. Let R_1 and ρ_1 be the values at T_1, and R_2 and ρ_2 be the values at T_2. Then, by eq. (18),

$$R_1 = \rho_1(L/A) \quad \text{and} \quad R_2 = \rho_2(L/A)$$

from which, after dividing the second equation by the first, we have

$$R_2 = (\rho_2/\rho_1)R_1$$

Hence, using eq. (19),

$$R_2 = \left[\frac{1 + \alpha_0(T_2 - 20)}{1 + \alpha_0(T_1 - 20)}\right] R_1$$

from which you should find that

$$R_2 = \frac{1.076}{1.038}(2.625) = 2.721 \text{ ohms}, \textit{ answer}.$$

14. By eq. (19), $\rho = (49)(10^{-8})(1.00545) = (49.2671)10^{-8}$. Then, from eq. 18,

$$L = \frac{RA}{\rho} = \frac{(35)(\pi)(25)10^{-8}}{(49.2671)10^{-8}} = 55.7957 \text{ meters} = 183.056 \text{ feet}, \textit{ answer}.$$

15. (a) $I = V/R_T = 48/25 = 1.92$ amperes, *answer*.
(b) M2 reads reads the voltage across the 5-ohm resistor; hence, $R_x = 5$ ohms; thus,

$$V_x = IR_x = (1.92)(5) = 9.6 \text{ volts}, \textit{ answer}.$$

(c) $P = VI = 48 \times 1.92 = 92.16$ watts, *answer*.

16. The power to each resistor is I^2R, by eq. (17). Hence, for the

3-ohm resistor $= (1.92)^2(3) = 11.0592$ watts
7-ohm resistor $= (1.92)^2(7) = 25.8048$ watts
5-ohm resistor $= (1.92)^2(5) = 18.4320$ watts
4-ohm resistor $= (1.92)^2(4) = 14.7456$ watts
6-ohm resistor $= (1.92)^2(6) = 22.1184$ watts

The sum of all the individual powers is 92.16 watts, which agrees with the answer found in part (c) of problem 15.

17. (a) By eq. (32), $1/R_T = 1/9 + 1/15 + 1/22 + 1/17 + 1/12 = 0.36539$; hence,

$$R_T = 1/0.36539 = 2.736808 \text{ ohms}, \textit{ answer}.$$

(b) $I = V/R_T = 18/2.736808 = 6.577005$ amperes, *answer*.
(c) $P = VI = (18)(6.577005) = 118.3861$ watts, *answer*.
(d) current in 9-ohm $= 18/9 = 2.00000$ amp.
current in 15-ohm $= 18/15 = 1.20000$ amp.
current in 22-ohm $= 18/22 = 0.81818$ amp.
current in 17-ohm $= 18/17 = 1.05882$ amp.
current in 12-ohm $= 18/12 = 1.50000$ amp.
The sum of the above currents = total current found in part (b); so eq. (28) *is* satisfied.
(e) Using $V^2 = (18)^2 = 324$, we have, by eq. (33),
power to 9-ohm $= 324/9 = 36.0000$ watts
power to 15-ohm $= 324/15 = 21.6000$ watts

power to 22-ohm = 324/22 = 14.7273 watts
power to 17-ohm = 324/17 = 19.0588 watts
power to 12-ohm = 324/12 = 27.0000 watts

Sum of above powers = output of battery in part (c), so eq. (26) *is* satisfied.

18. By eq. (34) the battery sees $R_T = (25)(38)/(25 + 38) = 15.0794$ ohms, and hence

$$\text{battery current} = I = V/R_T = 15/15.0794 = 0.9947 \text{ amperes, } \textit{answer}.$$

$$\text{battery power} = VI = 14.9205 \text{ watts, } \textit{answer}.$$

19. By eq. (35), $R_T = 25/16 = 1.5625$ ohms; therefore

(a) battery current $= I = V/R_T = 12/1.5625 = 7.68$ amperes, *answer*.

(b) $I_x = V/R_x = 12/25 = 0.48$ amperes, *answer*.

(c) $P = VI = 12(7.68) = 92.16$ watts, *answer*.

(d) $P_x = VI_x = 12(0.48) = 5.76$ watts, *answer*.

Note of interest: Let us pause and consider the question "At what speed do the charge carriers actually move in a solid conductor such as a copper wire?"

As we know, the speed of propagation of electrical energy along wires is very great, being only slightly less than the speed of light in free space, which is approximately 300 million meters *per second* (186,300 miles/sec). It should be emphasized, however, that this is the speed at which the electric and magnetic fields are propagated along the line, and is *not* the speed at which the charge carriers actually move. In a solid conductor, such as a copper wire, the charge carriers move at considerably less than 2.5 centimeters (one inch) per second.

To put it somewhat loosely, the "wave of electromotive force" is propagated along wires at very great speed, but the actual charge carriers, that constitute the current, move quite slowly, with an average speed of less than 1 inch per second. The amount of current flowing in a wire depends upon the number of charge carriers in motion, not on the speed of the individual charge carriers.

20. (a) By eq. (34), the parallel combination of the 12- and 6-ohm resistance is equal to a single resistance of (12)(6)/(12 + 6) = 4 ohms. Hence Fig. 34 can be redrawn as Fig. 368.

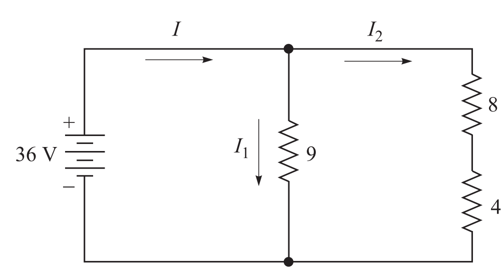


Fig. 368

By Ohm's law, $I_1 = 36/9 = 4$ amp and $I_2 = 36/12 = 3$ amp. Therefore,

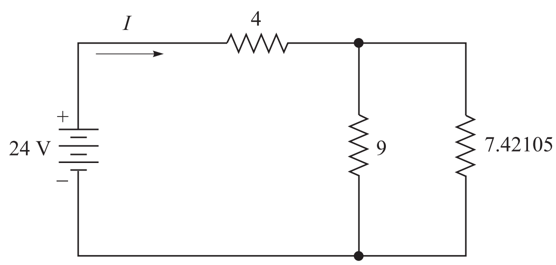
$$\text{battery current} = I = I_1 + I_2 = 4 + 3 = 7 \text{ amperes, } \textit{answer}.$$

(b) $P = VI = (36)(7) = 252$ watts, *answer*.

(c) Referring to Fig. 368, we have already found, in part (a), that $I_2 = 3$ amperes. Hence the voltage drop across the 4-ohm resistor in Fig. 368 is, by eq. (13), equal to $(3)(4) = 12$ volts, *which is the voltage across both the 12-ohm and the 6-ohm resistors* in Fig. 34. Therefore the current in the *6-ohm resistor* is equal to

$$I = 12/6 = 2 \text{ amperes, } \textit{answer}.$$

21. First, by eq. (34), the parallel 7-ohm and 12-ohm resistors are equivalent to a single resistance of $(7)(12)/19 = 4.42105$ ohms, approx. Since this is in series with the 3-ohm resistor, the figure reduces to the following:



Note that we now have 4 ohms in series with the parallel combination of 9 ohms and 7.42105 ohms. Hence $R_T = 4 + 9(7.42105)/16.42105 = 8.06731$ ohms, and therefore, by Ohm's law, eq. (11), we have

$$I = \text{battery current} = 24/8.06731 = 2.97497 \text{ amperes, } \textit{answer}.$$

22. First let us find the battery current I, as follows, Let R be the equivalent resistance of 14, 16, 18, and 20 ohms in parallel. To do this we use eq. (32), thus $1/R = 1/14 + 1/16 + 1/18 + 1/20 = 0.239484$, hence $R = 4.17564$ ohms. The battery thus sees a resistance $R_T = 3 + 4.17564 = 7.17564$ ohms, and thus the battery current I is equal to

$$I = V/R_T = 18/7.17564 = 2.50849 \text{ amperes}$$

(a) The resistance from point x to ground is $R_x = 4.17564$ ohms, and since the battery current I flows through this resistance we have

$$V_x = IR_x = 10.47455 \text{ volts, } \textit{answer}.$$

(b) The *current in the 20-ohm branch* $(7 + 13 = 20)$ is the voltage at point x divided by 20 ohms, that is, $10.47455/20 = 0.523728$ amperes, and this current *flowing through the 13-ohm resistor* gives

$$V_y = (0.523728)(13) = 6.80846 \text{ volts, } \textit{answer}.$$

23. Let us use eq. (34) as follows. Solving for R_2, we have $R_2 = \dfrac{R_1 R_T}{R_1 - R_T}$, and setting $R_1 = 36$ and $R_T = 20$, we have

$$R_2 = (36)(20)/16 = 45 \text{ ohms, } \textit{answer}.$$

24. (a) $7 \times 7 = 49$ elements.

(b) Written as a_{53} or, if you wish, $a_{5,3}$.

(c) $a_{1,11}$ denotes the element at the intersection of the first row and the eleventh column, whereas $a_{11,1}$ denotes the element at the intersection of the eleventh row and the first column.

Note: The answers to problems 25 through 29 all follow from eq. (36).

25. (a) $(12 - 8) = 4$, *answer*.
(b) $3(2 + 8) = 30$, *answer*.

26. $(24 - 6) + (-3 + 2) + (4 + 6) = 27$, *answer*.

27. $\dfrac{(20 - 60)}{(20 - 28)} = -40/-8 = 5$, *answer*.

28. $10x^2y + y^2 = y(10x^2 + y)$, *answer*.

29. $(x + 2)(2 - x) - 4(x - 5) = -x^2 - 4x + 24$, *answer*.

Note: The solutions to problems 30 through 34 follow the 3-step procedure of section 3.3.

30. (1) Here, $a_{11} = 3$, $a_{21} = -5$, $a_{31} = 1$.

(2) $$a_{11}A_{11} = a_{11}(-1)^2M_{11} = a_{11}M_{11} = 3\begin{vmatrix} 7 & -4 \\ -2 & 3 \end{vmatrix} = 3(21 - 8) = 39$$

$$a_{21}A_{21} = a_{21}(-1)^3M_{21} = -a_{21}M_{21} = 5\begin{vmatrix} 6 & 1 \\ -2 & 3 \end{vmatrix} = 5(18 + 2) = 100$$

$$a_{31}A_{31} = a_{31}(-1)^4M_{31} = a_{31}M_{31} = \begin{vmatrix} 6 & 1 \\ 7 & -4 \end{vmatrix} = (-24 - 7) = -31$$

(3) $D = 39 + 100 - 31 = 108$, *answer*.

31. (1) Here $a_{31} = 1$, $a_{32} = -2$, $a_{33} = 3$

(2) $$a_{31}A_{31} = a_{31}(-1)^4M_{31} = a_{31}M_{31} = \begin{vmatrix} 6 & 1 \\ 7 & -4 \end{vmatrix} = (-24 - 7) = -31$$

$$a_{32}A_{32} = a_{32}(-1)^5M_{32} = -a_{32}M_{32} = 2\begin{vmatrix} 3 & 1 \\ -5 & -4 \end{vmatrix} = 2(-12 + 5) = -14$$

$$a_{33}A_{33} = a_{33}(-1)^6M_{33} = a_{33}M_{33} = 3\begin{vmatrix} 3 & 6 \\ -5 & 7 \end{vmatrix} = 3(21 + 30) = 153$$

(3) $D = -31 - 14 + 153 = 108$, *answer*.

32. The easiest way is to expand the determinant in terms of the elements of the *third column*, because two of the three elements there are zeros. Thus, since $a_{13} = 0$,

$a_{23} = 0$, and $a_{33} = -4$, we have

$$D = 0 + 0 + a_{33}(-1)^6 M_{33} = -4\begin{vmatrix} 6 & 1 \\ 2 & 4 \end{vmatrix} = -88, \ \textit{answer}.$$

33. (1) Let us use the *second column* (because it contains two zeros). Thus we have $a_{12} = 2$, $a_{22} = 0$, $a_{32} = -1$, $a_{42} = 0$.

(2)* Since $a_{22} = 0$ and also $a_{42} = 0$, we have

$$a_{12}M_{12} = -2\begin{vmatrix} 1 & 2 & 2 \\ -3 & 3 & 0 \\ -1 & 6 & 3 \end{vmatrix} \quad \text{and} \quad a_{32}M_{32} = -(-1)\begin{vmatrix} 2 & 5 & -4 \\ 1 & 2 & 2 \\ -1 & 6 & 3 \end{vmatrix}$$

Let D_1 and D_2 be the values of the two determinants above, and let us expand both in terms of the elements of the third column. Doing this gives us

$$D_1 = -2\left(2\begin{vmatrix} -3 & 3 \\ -1 & 6 \end{vmatrix} + 3\begin{vmatrix} 1 & 2 \\ -3 & 3 \end{vmatrix}\right) = -2(-30 + 27) = 6$$

$$D_2 = -4\begin{vmatrix} 1 & 2 \\ -1 & 6 \end{vmatrix} - 2\begin{vmatrix} 2 & 5 \\ -1 & 6 \end{vmatrix} + 3\begin{vmatrix} 2 & 5 \\ 1 & 2 \end{vmatrix} = -4(8) - 2(17) + 3(-1) = -69$$

(3) $D = D_1 + D_2 = 6 - 69 = -63$, *answer*.

34. Let us begin by expanding the given determinant in terms of the elements of the fourth row. Since all the elements of the fourth row except the -5 are zeros, the value of the determinant reduces to the value of the single third-order determinant

$$D = -5\begin{vmatrix} 1 & 3 & 2 \\ 6 & 1 & -3 \\ 4 & -2 & 0 \end{vmatrix}$$

Now expand the above result in terms of, let us say, the third column; thus

$$D = -5\left(2\begin{vmatrix} 6 & 1 \\ 4 & -2 \end{vmatrix} - (-3)\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix}\right) = -5(-32 - 42) = 370, \ \textit{answer}.$$

35. First factor 6 from column 1 and 5 from column 3. Then factor 4 from the second row. We can then, if we wish, expand the determinant in terms of the minors of the first column; thus

$$(6)(5)(4)\begin{vmatrix} 1 & -3 & 1 \\ 1 & -4 & 6 \\ -2 & 1 & 5 \end{vmatrix} = 120\left(\begin{vmatrix} -4 & 6 \\ 1 & 5 \end{vmatrix} - \begin{vmatrix} -3 & 1 \\ 1 & 5 \end{vmatrix} - 2\begin{vmatrix} -3 & 1 \\ -4 & 6 \end{vmatrix}\right)$$

$$= 120(-26 + 16 + 28) = 2160, \ \textit{answer}.$$

* Step (2) can also be stated as follows. Let a denote any element in a determinant, and let aM denote the product of a and its minor determinant M.

Let s denote the *sum* of the number of the row and the number of the column in which a is located; if s is an *even* number write aM, but if s is an *odd* number write $-aM$.

36. Factoring in accordance with property 5, then expanding in terms of the minors of column 2, we have

$$(7)(5)(2)(6)\begin{vmatrix} 3 & 1 & -2 & 1 \\ 1 & 0 & 1 & -2 \\ -4 & 0 & 2 & 1 \\ 2 & 0 & -3 & 2 \end{vmatrix} = -420\begin{vmatrix} 1 & 1 & -2 \\ -4 & 2 & 1 \\ 2 & -3 & 2 \end{vmatrix} = -420, \ \textit{answer}.$$

37. *No*, because this is simply a special case of multiplying by $+1$ or -1 in property 7.

38. Factoring "2" from column 4, we have

$$2\begin{vmatrix} 3 & -1 & 2 & 3 \\ 4 & 3 & 0 & 4 \\ 1 & -4 & 4 & 1 \\ 6 & 7 & 5 & 6 \end{vmatrix} = (2)(0) = 0, \text{by property 4.}$$

39. Note, first, that no factoring can be done, and that no two rows or no two columns are identical. Therefore, using property 7, let's try to write the determinant in a more convenient form, so that most of the elements in one particular row or column are zeros. One way is as follows.

First, to each element of column 4 add the corresponding element of column 2 multiplied by 2.

Next, to each element of column 5 add the corresponding element of column 2 multiplied by 4. Taking these steps, the original determinant becomes

$$\begin{vmatrix} 2 & 0 & -1 & (5+0) & (0+0) \\ 0 & -1 & 0 & (2-2) & (4-4) \\ -1 & 3 & 2 & (0+6) & (-1+12) \\ 3 & -2 & 0 & (-1-4) & (-1-8) \\ -2 & 4 & -3 & (2+8) & (0+16) \end{vmatrix} = \begin{vmatrix} 2 & 0 & -1 & 5 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ -1 & 3 & 2 & 6 & 11 \\ 3 & -2 & 0 & -5 & -9 \\ -2 & 4 & -3 & 10 & 16 \end{vmatrix}$$

In the new, but equivalent, determinant to the right, notice that all the elements in row 2 are *zeros* except the one element "-1." This is good, because it is now easy to expand the determinant in terms of the elements of row 2; upon doing this, we are able, very easily, to reduce the given *fifth-order* determinant to a single *fourth-order* determinant; thus

$$(-1)\begin{vmatrix} 2 & -1 & 5 & 0 \\ -1 & 2 & 6 & 11 \\ 3 & 0 & -5 & -9 \\ -2 & -3 & 10 & 16 \end{vmatrix}$$

Let us now repeat the procedure to reduce the fourth-order to a third-order determinant; one way is as follows. First, multiply all the terms of column 3 by "3"; this is permissible *provided* that we also multiply the determinant by "1/3";

thus (basically making use of property 5)

$$-\tfrac{1}{3}\begin{vmatrix} 2 & -1 & 15 & 0 \\ -1 & 2 & 18 & 11 \\ 3 & 0 & -15 & -9 \\ -2 & -3 & 30 & 16 \end{vmatrix}$$

Now take the following steps: first, to each element of column 3 add the corresponding element of column 1 multiplied by "5"; then, to each element of column 4 add the corresponding element of column 1 multiplied by "3"; thus

$$-\tfrac{1}{3}\begin{vmatrix} 2 & -1 & (15+10) & (0+6) \\ -1 & 2 & (18-5) & (11-3) \\ 3 & 0 & (-15+15) & (-9+9) \\ -2 & -3 & (30-10) & (16-6) \end{vmatrix} = -\tfrac{1}{3}\begin{vmatrix} 2 & -1 & 25 & 6 \\ -1 & 2 & 13 & 8 \\ 3 & 0 & 0 & 0 \\ -2 & -3 & 20 & 10 \end{vmatrix}$$

In the determinant to the right notice that all the elements of row 3 are *zeros* except the one element 3. Hence it's now easy to expand the determinant in terms of the minors of the elements of row 3, and doing this reduces the *fourth-order* determinant to a single equivalent *third-order* determinant; thus,

$$-\begin{vmatrix} -1 & 25 & 6 \\ 2 & 13 & 8 \\ -3 & 20 & 10 \end{vmatrix} = -2\begin{vmatrix} -1 & 25 & 3 \\ 2 & 13 & 4 \\ -3 & 20 & 5 \end{vmatrix}$$

The final step is to expand the third-order determinant, on the right-hand side, into a sum of basic second-order determinants. Fundamentally, this is done by expanding the third-order determinant into a sum of three second-order determinants in the basic way. Or, using property 7, we can reduce the number of second-order determinants to just one, as follows.

First, to each element of column 2 add the corresponding element of column 1 multiplied by 25. Next, to each element of column 3 add the corresponding term of column 1 multiplied by 3. Doing this reduces the third-order determinant to a single equivalent *second-order* determinant; thus

$$-2\begin{vmatrix} -1 & 0 & 0 \\ 2 & 63 & 10 \\ -3 & -55 & -4 \end{vmatrix}_{*} = (-2)(-1)\begin{vmatrix} 63 & 10 \\ -55 & -4 \end{vmatrix} = 2(-252+550) = 596, \ \textit{answer}.$$

The foregoing solution shows how the use of property 7 can greatly reduce the amount of work needed to find the value of a determinant. Thus, in *one step* we reduced the original 5th-order determinant to a 4th-order determinant. Then, in a second step we reduced the 4th-order determinant to a 3rd-order, and in a third step we reduced the 3rd-order to the basic 2nd-order form. Thus in three steps the given 5th-order determinant was reduced to one basic 2nd-order determinant.

40. Step 1 is already satisfied. Next (step 2) we have

$$\Delta = \begin{vmatrix} 5 & 2 \\ -3 & 4 \end{vmatrix} = (20+6) = 26$$

* At this point we could have factored a "−1" from the third row, making the elements of the third row 3, 55, 4 and thus the multiplier of the determinant 2 instead of −2.

Step 3:
Let us solve for x first:

$$\Delta' = \begin{vmatrix} -7 & 2 \\ 25 & 4 \end{vmatrix} = (-28 - 50) = -78$$

Step 4:

$$x = \Delta'/\Delta = -78/26 = -3, \text{ *answer*, the value of } x.$$

To find the value of y we must now rework step 3 to find the new value of Δ'; thus

$$\Delta' = \begin{vmatrix} 5 & -7 \\ -3 & 25 \end{vmatrix} = (125 - 21) = 104$$

thus, repeating step 4, $y = \Delta'/\Delta = 104/26 = 4$, *answer*, the value of y.

To check the correctness of the answers, go back to the *original two equations* and set $x = -3$ and $y = 4$. Doing this we find that

$$5(-3) + 2(4) = -7 \text{ (which is true)},$$
$$-3(-3) + 4(4) = 25 \text{ (which is true)},$$

showing that the answers *are* correct. It should be noted that $x = -3$ and $y = 4$ are the *only* values of x and y that simultaneously satisfy both of the given equations.

41. Step 1 is already satisfied. Next, for step 2 we have (note that $x = 1$ times x)

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & -1 \end{vmatrix}$$

Now, using property 7, to each element of column 3 add the corresponding element of column 1 multiplied by "1," thus getting

$$\Delta = \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & -2 & 0 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = -6$$

thus, $\Delta = -6$.

Step 3: Let us solve for the values of x, y, and z, in that order, as follows.

$$x = \frac{\begin{vmatrix} 6 & 1 & 1 \\ 0 & 1 & -1 \\ 3 & -2 & -1 \end{vmatrix}}{-6} = \frac{\begin{vmatrix} 0 & 5 & 3 \\ 0 & 1 & -1 \\ 3 & -2 & -1 \end{vmatrix}}{-6} = \frac{3\begin{vmatrix} 5 & 3 \\ 1 & -1 \end{vmatrix}}{-6} = 4, \text{ *answer*, value of } x.$$

(**Note:** to get the second determinant above from the first, to each element of row 1 add the corresponding element of row 3 multiplied by -2.)

Next,

$$y = \frac{\begin{vmatrix} 1 & 6 & 1 \\ 1 & 0 & -1 \\ 1 & 3 & -1 \end{vmatrix}}{-6} = \frac{\begin{vmatrix} -1 & 0 & 3 \\ 1 & 0 & -1 \\ 1 & 3 & -1 \end{vmatrix}}{-6} = \frac{-3\begin{vmatrix} -1 & 3 \\ 1 & -1 \end{vmatrix}}{-6} = -1, \textit{ answer}, \text{ value of } y.$$

$$z = \frac{\begin{vmatrix} 1 & 1 & 6 \\ 1 & 1 & 0 \\ 1 & -2 & 3 \end{vmatrix}}{-6} = \frac{\begin{vmatrix} -1 & 5 & 0 \\ 1 & 1 & 0 \\ 1 & -2 & 3 \end{vmatrix}}{-6} = \frac{3\begin{vmatrix} -1 & 5 \\ 1 & 1 \end{vmatrix}}{-6} = 3, \textit{ answer}, \text{ value of } z.$$

Hence the *answers* are $x = 4$, $y = -1$, and $z = 3$.

42. Step 1 is satisfied. Next, for step 2 we have

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 3 & 4 & -2 \\ 0 & -4 & -5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & -5 \\ 0 & -4 & -5 \end{vmatrix} = \begin{vmatrix} 1 & -5 \\ -4 & -5 \end{vmatrix} = -5 - 20 = -25$$

thus, $\Delta = -25$.

Step 3: Let us solve for the values of x, y, and z, in that order, as follows. First, for the value of x, the value of Δ' is

$$\Delta' = \begin{vmatrix} 4 & 1 & 1 \\ -2 & 4 & -2 \\ 1 & -4 & -5 \end{vmatrix}$$

Now take the following steps: to each element of row 3 add the corresponding element of row 2 multiplied by 1; then, to each element of row 2 add the corresponding element of row 1 multiplied by -4. Doing this, the above expression for Δ' becomes

$$\Delta' = \begin{vmatrix} 4 & 1 & 1 \\ -2 & 4 & -2 \\ -1 & 0 & -7 \end{vmatrix} = \begin{vmatrix} 4 & 1 & 1 \\ -18 & 0 & -6 \\ -1 & 0 & -7 \end{vmatrix} = -\begin{vmatrix} -18 & -6 \\ -1 & -7 \end{vmatrix} = -120$$

Hence, $x = \Delta'/\Delta = -120/-25 = 4.8$, *answer*, value of x.

Next, for the value of y,

$$\Delta' = \begin{vmatrix} 1 & 4 & 1 \\ 3 & -2 & -2 \\ 0 & 1 & -5 \end{vmatrix} = 75$$

thus, $y = \Delta'/\Delta = \dfrac{75}{-25} = -3$, *answer*, value of y.

Next for the value of z,

$$\Delta' = \begin{vmatrix} 1 & 1 & 4 \\ 3 & 4 & -2 \\ 0 & -4 & 1 \end{vmatrix} = -55$$

thus, $z = \Delta'/\Delta = \dfrac{-55}{-25} = 2.2$, *answer*, value of z.

Thus the correct *answers* are $x = 4.8$, $y = -3.0$, and $z = 2.2$, which you can verify by replacing x, y, and z with these values in the original three equations.

43. Step 1 is satisfied. Next, for step 2, we have

$$\Delta = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 3 & -2 & 4 & 4 \\ -2 & 5 & 7 & 0 \\ 0 & 3 & 2 & -3 \end{vmatrix}$$

The above 4th-order determinant can be reduced to a single 3rd-order by applying property 7; one procedure is as follows.

First, to each element of row 2 add the corresponding element of row 1 multiplied by -4. Then, to each element of row 4 add the corresponding element of row 1 multiplied by 3, thus getting

$$\Delta = \begin{vmatrix} 1 & 1 & 1 & 1 \\ -1 & -6 & 0 & 0 \\ -2 & 5 & 7 & 0 \\ 3 & 6 & 5 & 0 \end{vmatrix} = -\begin{vmatrix} -1 & -6 & 0 \\ -2 & 5 & 7 \\ 3 & 6 & 5 \end{vmatrix}$$

Now let us reduce the 3rd-order determinant to a single 2nd-order; one way is as follows. To each element of column 2 add the corresponding element of column 1 multiplied by -6, thus giving us

$$\Delta = -\begin{vmatrix} -1 & 0 & 0 \\ -2 & 17 & 7 \\ 3 & -12 & 5 \end{vmatrix} = \begin{vmatrix} 17 & 7 \\ -12 & 5 \end{vmatrix} = 169$$

Thus, $\Delta = 169$ for this problem. We now go to step 3 to find the values of w, x, y, and z, in that order, as follows.

For w:

$$\Delta' = \begin{vmatrix} -4 & 1 & 1 & 1 \\ 0 & -2 & 4 & 4 \\ -12 & 5 & 7 & 0 \\ 5 & 3 & 2 & -3 \end{vmatrix} = 338$$

thus, $w = \Delta'/\Delta = \dfrac{338}{169} = 2.$

For x:

$$\Delta' = \begin{vmatrix} 1 & -4 & 1 & 1 \\ 3 & 0 & 4 & 4 \\ -2 & -12 & 7 & 0 \\ 0 & 5 & 2 & -3 \end{vmatrix} = -507$$

thus, $x = \Delta'/\Delta = -\dfrac{507}{169} = -3$

For y:

$$\Delta' = \begin{vmatrix} 1 & 1 & -4 & 1 \\ 3 & -2 & 0 & 4 \\ -2 & 5 & -12 & 0 \\ 0 & 3 & 5 & -3 \end{vmatrix} = 169$$

thus, $y = \Delta'/\Delta = \dfrac{169}{169} = 1.$

For z:

$$\Delta' = \begin{vmatrix} 1 & 1 & 1 & -4 \\ 3 & -2 & 4 & 0 \\ -2 & 5 & 7 & -12 \\ 0 & 3 & 2 & 5 \end{vmatrix} = -676$$

thus, $z = \Delta'/\Delta = \dfrac{-676}{169} = -4.$

Thus the correct *answers* are $w = 2$, $x = -3$, $y = 1$, and $z = -4$.

44. First you should verify that

$$\Delta = \begin{vmatrix} 3 & -2 & -5 \\ 1 & -1 & -1 \\ 2 & -1 & -4 \end{vmatrix} = 0$$

which assures us that the given system does have non-trivial solutions. Then, from inspection of the equations

$$\begin{array}{lll} a = 3, & b = -2, & c = -5 \\ d = 1, & e = -1, & f = -1 \\ g = 2, & h = -1, & i = -4 \end{array}$$

which, upon substitution into eq. (55), gives the *answers* $x = 3$, $y = 2$, and $z = 1$.

Note: An important point concerning systems of homogeneous linear equations can be shown as follows. Let k be any constant, and let us multiply through each of the three equations of eq. (53) by k, thus getting

$$a(kx) + b(ky) + c(kz) = 0$$
$$d(kx) + e(ky) + f(kz) = 0$$
$$g(kx) + h(ky) + i(kz) = 0$$

Note that the last three equations have *exactly the same form as eq. (53)*; thus, if x, y, and z represent values that satisfy eq. (53), the values of kx, ky, and kz *also* satisfy eq. (53).

It's apparent that the foregoing is true for any system of n homogeneous linear equations; thus, if $x_1, x_2, \ldots, x_n$ are found to be a solution set of such a system, then $kx_1, k_{x2}, \ldots, kx_n$ is also a solution set to the system, where k is any constant.

45. Let us expand the determinant of the coefficients in terms of the elements of the first row; doing this, you should find that

$$\Delta = \begin{vmatrix} 4 & -18 & -7 \\ 2 & -4p & p \\ p & 3 & 5 \end{vmatrix} = -46p^2 - 92p + 138$$

The system will have non-trivial solutions only if $\Delta = 0$. Thus, setting $\Delta = 0$, we have the requirement that $-46p^2 - 92p + 138 = 0$, the solutions of which are (using the quadratic formula),* $p = 1$ or $p = -3$, *answers* to first part. Next, setting first $p = 1$ and then $p = -3$ in the original given system gives the two possible systems

$$\begin{aligned} 4x - 18y - 7z &= 0 \qquad & 4x - 18y - 7z &= 0 \\ 2x - 4y + z &= 0 \qquad & 2x + 12y - 3z &= 0 \\ x + 3y + 5z &= 0 \qquad & -3x + 3y + 5z &= 0 \end{aligned}$$

Now write down the values of "*a* through *i*" for each of the above two equations, then substitute into eq. (55). Doing this, you should find that the *answers* are

$$\begin{array}{llll} \text{for } p = 1: & & \text{for } p = -3: & \\ x = -23, & \text{thus } x = -23k & x = 69, & \text{thus } x = 69k \\ y = -9, & \text{thus } y = -9k & y = -1, & \text{thus } y = -k \\ z = 10, & \text{thus } z = 10k & z = 42, & \text{thus } z = 42k \end{array}$$

46. In Fig. 48 the current arrows show the direction of flow of what we will call "positive" current. Let us therefore write the voltage equation generated by tracing around the circuit in the cw sense, putting voltage "drops" on the left-hand side and voltage "rises" on the right-hand side, as agreed upon in the discussion of Fig. 47. Doing this, starting at point A (keeping Fig. 46 in mind), we have that the voltage equation for Fig. 48 is

$$12 + 4I + 2I + I + 26 + 5I = 17$$

thus,

$$I = -21/12 = -1.75 \text{ amperes, } \textit{answer}.$$

The answer, "minus" 1.75, means that, for the given battery voltages and polarities, the current I would actually flow in the ccw sense.

47.† Let us follow the suggested three steps (as illustrated in the example problem) thus:

Step I. Note that *two* loop currents will satisfy the requirement that all the circuit elements be traversed at least once by a current. Call the current in the left-hand loop I_1 and the current in the right-hand loop I_2, and let us elect to draw the arrow-heads to indicate that I_1 and I_2 both flow in the *clockwise* sense.

* From note 1 in the Appendix.

† Unless specifically stated otherwise, we'll assume that all generators (batteries in these problems) have negligible internal resistance. (See Figs. 25 and 26 in section 2.5 for a discussion of "internal resistance.")

Step II.

voltage equation around left-hand loop: $15I_1 - 7I_2 = 6$

voltage equation around right-hand loop: $-7I_1 + 11I_2 = 21$

Step III. The current in the 7-ohm resistance is the algebraic sum of I_1 and I_2, and therefore we must find the values of *both* I_1 and I_2, as follows.

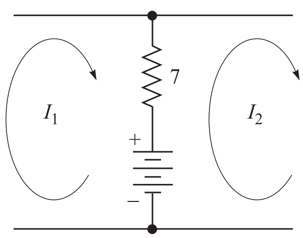
First,

$$\Delta = \begin{vmatrix} 15 & -7 \\ -7 & 11 \end{vmatrix} = 116$$

Then,

$$I_1 = \frac{\begin{vmatrix} 6 & -7 \\ 21 & 11 \end{vmatrix}}{116} = 1.83621 \text{ amperes and } I_2 = \frac{\begin{vmatrix} 15 & 6 \\ -7 & 21 \end{vmatrix}}{116} = 3.07759 \text{ amperes.}$$

For the directions we have elected to take for I_1 and I_2, we have the condition shown in the figure to the right (showing just that part of the circuit we're interested in now).



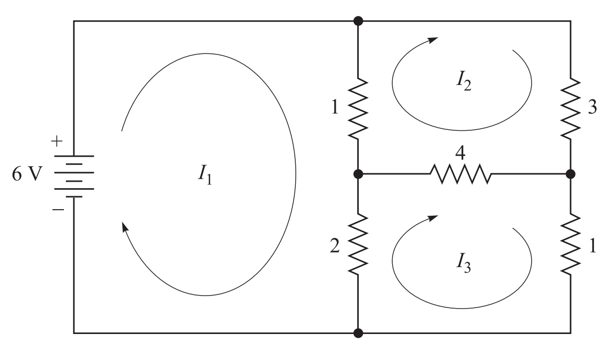
Since the two currents in the 7-ohm resistance flow in *opposite directions* through the resistance, we must take the *difference* of the two currents; thus the current in the 7-ohm resistance is

$$I_1 - I_2 = -1.24138 \text{ amperes, } \textit{answer}$$

$$I_2 - I_1 = 1.24138 \text{ amperes, } \textit{answer}.$$

Both answers mean the same thing; the first answer means the resultant current (in the 7-ohm resistor) is flowing opposite to I_1, that is, in the direction of I_2, while the second answer directly states that the resultant current is flowing in the direction of I_2. An ammeter placed in series with the 7-ohm resistor would read 1.24138 amperes.

48. **Step I.** Three loop currents must be used, to satisfy the requirement that all the circuit elements must be included in the analysis. Let us suppose we elect to have all the currents flow in the clockwise sense, as in the figure below.



Step II.

$$\begin{aligned}
&\text{voltage equation for loop 1:} && 3I_1 - I_2 - 2I_3 = 6\\
&\text{voltage equation for loop 2:} && -I_1 + 8I_2 - 4I_3 = 0\\
&\text{voltage equation for loop 3:} && -2I_1 - 4I_2 + 7I_3 = 0
\end{aligned}$$

Note that since there is zero battery voltage in loops 2 and 3, the right-hand sides of the voltage equations for those loops is zero.

Step III. In this problem we're asked to find the voltage drop across the 4-ohm resistance. From the above figure we see that, since I_2 and I_3 *both* flow in that resistor, we must find the values of both I_2 and I_3. The first step in doing this is, as usual, to find the value of "delta," which is found from step II to be

$$\Delta = \begin{vmatrix} 3 & -1 & -2 \\ -1 & 8 & -4 \\ -2 & -4 & 7 \end{vmatrix} = 65$$

Next you can verify that

$$I_2 = \frac{6\begin{vmatrix} 3 & 1 & -2 \\ -1 & 0 & -4 \\ -2 & 0 & 7 \end{vmatrix}}{65} = \frac{90}{65} = 1.384615 \text{ amperes,}$$

$$I_3 = \frac{6\begin{vmatrix} 3 & -1 & 1 \\ -1 & 8 & 0 \\ -2 & -4 & 0 \end{vmatrix}}{65} = \frac{120}{65} = 1.846154 \text{ amperes.}$$

From inspection of the figure, the net resultant current flowing from left to right in the 4-ohm resistor is $I_3 - I_2 = 0.461539$ amperes, and therefore the voltage drop across the 4-ohm resistor is

$$IR = (0.461539)4 = 1.8462 \text{ volts approx., } \textit{answer}.$$

49. Going from left to right in Fig. 54, draw three loop currents, I_1, I_2, and I_3, which we'll assume to all be in the clockwise sense. Then the three voltage equations are, going from left to right in Fig. 54,

$$\begin{aligned}
15I_1 - 10I_2 + 0I_3 &= 32\\
-10I_1 + 17I_2 - 4I_3 &= -21\\
0I_1 - 4I_2 + 11I_3 &= 12
\end{aligned}$$

In this problem we must find current I_3, because this is the current through the 7-ohm resistor. To do this we now form a determinant from the coefficients of the unknown currents; thus

$$\Delta = \begin{vmatrix} 15 & -10 & 0 \\ -10 & 17 & -4 \\ 0 & -4 & 11 \end{vmatrix} = 1465$$

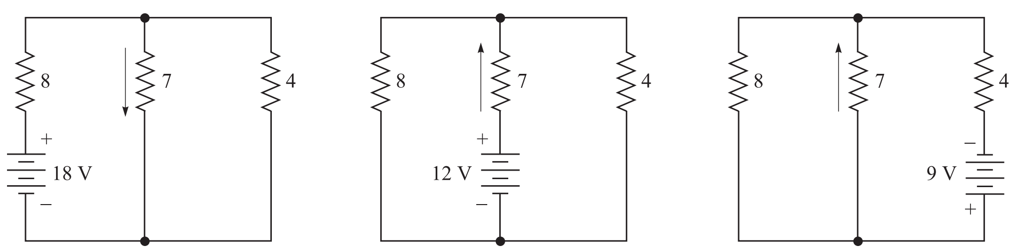
thus,

$$I_3 = \frac{\begin{vmatrix} 15 & -10 & 32 \\ -10 & 17 & -21 \\ 0 & -4 & 12 \end{vmatrix}}{1465} = \frac{1880}{1465} = 1.283277 \text{ amperes}$$

hence,

$$V_{\mathrm{a}} = 7I_3 = 8.98294 \text{ volts, } \textit{answer}.$$

50. The procedure is to find the current in the 7-ohm resistance in Fig. 52 due to *each battery considered separately*, the other batteries being replaced by their internal resistances (considered to be zero in this case). We must therefore, in this case, find the current in the 7-ohm resistance in each of the following three series-parallel networks. (Let us call "down" currents "negative" and "up" currents "positive.")



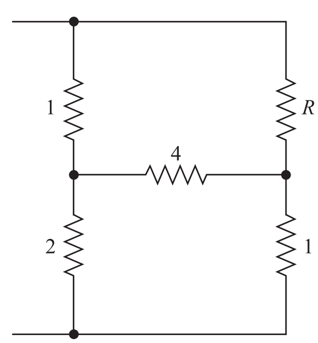
$I_7 = -0.620\,6896$ $I_7 = 1.241\,3793$ $I_7 = 0.620\,6896$

The three separate current values below the figures were calculated using the procedures of section 2.7. The net current in the 7-ohm resistance is then the algebraic sum of the three currents, which is, in this case, to five decimal places, 1.24138 amperes, the *same value* as found in problem 47. We might mention that, as was pointed out in section 2.5, current is assumed to flow out of a battery at the positive terminal and re-enter at the negative terminal.

51. A change in current causes the temperature of a physical resistor to change, which in turn causes the resistance to change (section 2.4). In our problems here we're assuming such changes in resistance are small enough to be disregarded.

52. This problem can be worked in two different ways, as follows.

FIRST WAY: Note the figure to the right. If the *ratio* of 1 to 2 is equal to the *ratio* of R to 1, then *zero potential difference* will exist between the terminals of the 4-ohm resistor, and thus zero current will flow in that resistor. That is, we must have



$$\frac{1}{2} = \frac{R}{1}, \text{ hence } R = 1/2 \text{ ohm, } \textit{answer}.$$

SECOND WAY: We can make use of the solution given with problem 48, as follows. Note that requiring zero current in the 4-ohm resistance is the same as requiring that currents I_2 and I_3 *cancel each other out in the 4-ohm resistance*, which, since I_2 and I_3 flow through the 4-ohm resistance in opposite directions, will be done only if I_2 and I_3 have *equal magnitudes*, that is, if $I_3 = I_2$. Therefore, in the network diagram given with the solution to problem 48, change "3" to "R" and I_3 to I_2. Doing this, the three loop equations become

$$I_1 - I_2 = 2 \qquad \text{(A)}$$
$$-I_1 + (1 + R)I_2 = 0 \qquad \text{(B)}$$
$$-2I_1 + 3I_2 = 0 \qquad \text{(C)}$$

Now add (A) and (B) together to get $I_2 = 2/R$. Now replace I_2 with $2/R$ in (B) and (C) to get

$$-I_1 + (1 + R)\frac{2}{R} = 0$$
$$-I_1 + \frac{3}{R} = 0$$

In these two equations, multiply through the first equation by R and the second by $-R$, then add the two together to get $R = 1/2$ ohm, *answer*.

53. Figure 56 is of the form of Fig. 55, and thus eq. (63) applies. Let us agree to number the branches from 1 through 5, from left to right in Fig. 56.

The first step is to convert "ohms" to "mhos," using the definition of eq. (58). Doing this, and noting that, in Fig. 56, $V_2 = V_5 = 0$, eq. (63) becomes

$$V_o = \frac{0.08333 \times 15 + 0.06666 \times 22 + 0.11111 \times 12}{0.08333 + 0.12500 + 0.06666 + 0.11111 + 0.10000}$$
$$= \frac{4.04979}{0.48611} = 8.33102 \text{ volts, } \textit{answer}.$$

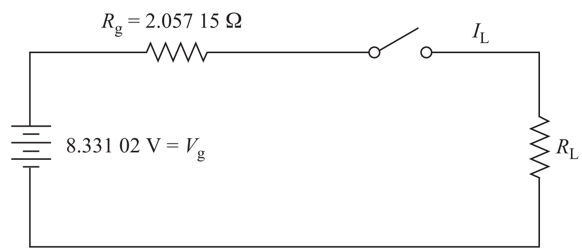
54. We would not have $V_4 = -12$ volts, and upon making this change in the above solution to problem 53 we find that

$$V_o = \frac{1.38315}{0.48611} = 2.84534 \text{ volts, } \textit{answer}.$$

55. The first step is to draw the Thevenin equivalent generator for Fig. 56. To do this, imagine that we are at the right-hand side of Fig. 56, looking to the left at the 10-ohm resistance. Then, from the solution to problem 53, we have

$$V_o = V_g = 8.33102 \text{ volts, and}$$
$$R_g = 1/0.48611 = 2.05715 \text{ ohms}$$

and thus, by Thevenin's theorem, we now have the simple series circuit condition shown as follows.



Thus, when the switch is closed:

for $R_L = 2$ ohms, $I_L = 8.33102/4.05715 = 2.05342$ amperes, *answer*,

for $R_L = 3$ ohms, $I_L = 8.33102/5.05715 = 1.64738$ amperes, *answer*,

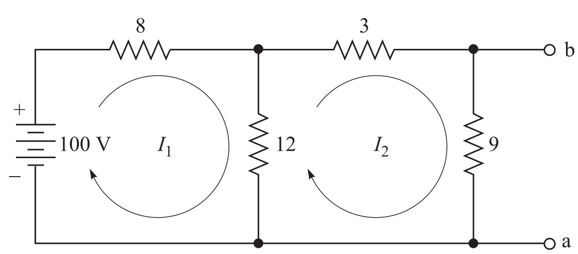
for $R_L = 4$ ohms, $I_L = 8.33102/6.05715 = 1.37540$ amperes, *answer*.

Without the use of Thevenin's theorem it would be necessary to rework problem 53 for each new value of R_L.

56. The first step is to remove R_L and find the open-circuit voltage between terminals a and b. One way to do this is to write the two loop-voltage equations (see figure below), thus

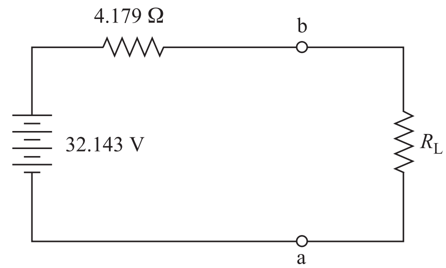
$$20I_1 - 12I_2 = 100$$
$$-12I_1 + 24I_2 = 0$$

from which $I_2 = 25/7$ amperes, and hence *the open-circuit voltage at a, b* is equal to $9(25/7) = 32.143$ volts.



The next step is to find the resistance looking back into terminals a, b (keeping R_L disconnected, as before). The 8-ohm and 12-ohm resistances are in parallel when looking back from terminals a, b and hence combine together (product over sum, eq. (34), Chap. 2) to give $(12)(8)/20 = 4.8$ ohms. The 4.8 ohms is now in series with the 3-ohm resistance, giving us a total of 7.8 ohms in parallel with the 9-ohm resistance. Therefore, looking into a, b we see 7.8 ohms in parallel with 9 ohms, which, using the "product of the two, over the sum," is equal to $(7.8)(9)/16.8 = 4.179$ ohms, approximately.

Thus the equivalent Thevenin generator for Fig. 59 is therefore as shown in the following figure, *answer*.



57. First, I_{sc} flows into a conductance equal to, by eq. (61), $G_T = G_g + G_L$, and thus, by eq. (59),

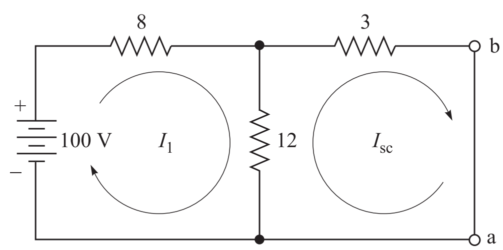
$$I_{sc} = (G_g + G_L)V_L$$

also,

$$I_L = G_L V_L$$

thus $V_L = I_L/G_L$, and putting this value of V_L into the first equation and solving for I_L gives the *answer*, eq. (67).

58. (a) The first step is to find the *short-circuit current* I_{sc}, which is the current that flows into the short-circuited terminals a, b as shown in the figure to the left below, with the two loop-voltage equations to the right.



$$5I_1 - 3I_{sc} = 25$$
$$-4I_1 + 5I_{sc} = 0$$

thus,

$$I_{sc} = \frac{\begin{vmatrix} 5 & 25 \\ -4 & 0 \end{vmatrix}}{\begin{vmatrix} 5 & -3 \\ -4 & 5 \end{vmatrix}} = \frac{100}{13} = 7.692308 \text{ amperes, approx.}$$

We must next find the value of the *conductance* component G_g of the Norton equivalent generator.

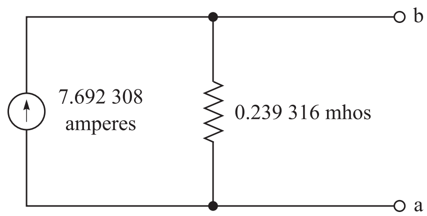
To do this, we now *remove the short-circuit* from the terminals a, b in the pre ceding figure; G_g now equals the conductance seen looking to the left into the now open-circuited terminals a, b. From inspection of Fig. 59 (disregard the 9-ohm branch for the moment), note that we look into one branch consisting of 3 ohms in series with the parallel combination of 8 ohms and 12 ohms, that is, into $3 + (8)(12)/20 = 7.8$ ohms. Since this resistance is in parallel with the 9-ohm resistance, the total resistance seen looking into the open-circuited terminals a, b is

$$R_g = (9)(7.8)/16.8 = 4.178571 \text{ ohms}$$

thus,

$$G_g = 1/R_g = 0.239316 \text{ mhos, approx.}$$

and hence the *answer* is that the Norton equivalent generator for Fig. 59 is



(b) The Thevenin equivalent of Fig. 59 appears in the solution to problem 56, and the Norton equivalent appears above. Setting $R_L = 10$ ohms in the Thevenin equivalent generator and applying Ohm's law, and setting $G_L = 1/10 = 0.1$ mho in eq. (67), we have that

$$\text{Thevenin case: } I_L = 32.143/14.179 = 2.267 \text{ amperes, approx.}$$

$$\text{Norton case: } I_L = 0.7692308/0.339316 = 2.267 \text{ amperes, approx.}$$

there being some difference beyond the third decimal place because the same degree of accuracy was not used in calculating in problem 56 as in problem 58.

59. Let a, b be the output terminals of any given network; to convert the network into the Thevenin equivalent generator we must take the following two steps.

(a) Find the open-circuit voltage at the terminals a, b; this is the generated *voltage*, V_g, of the equivalent generator (Fig. 57).

(b) Find the resistance looking into the open-circuited terminals a, b; this is the internal *resistance*, R_g, of the equivalent generator (Fig. 57).

Now apply the two steps to Fig. 60 (keeping the switch open), as follows.

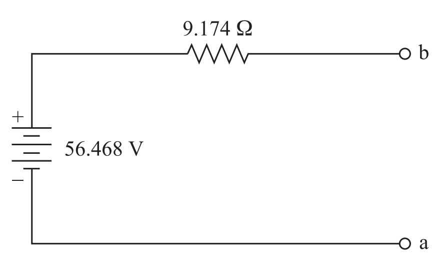
(a) Here we use the basic Ohm's law formula $V = RI = (1/G)I$, thus

$$V = \frac{I_{sc}}{G_g} = \frac{6.155}{0.109} = 56.468 \text{ volts} = V_g$$

(b) We must now replace all generators with their internal resistances. In this case we must remember that a constant-current generator has INFINITELY GREAT internal resistance. Thus, as far as resistance is concerned, the constant-current generator has no "shunting effect" on G_g, and thus

$$1/G_g = 1/0.109 = 9.174 \text{ ohms} = R_g$$

hence the Thevenin equivalent generator is



60. First note that Fig. 63 contains 7 node points (including the zero-volt reference node at "ground" potential), but note that only *two* of the node voltages, at nodes 1 and 2, are *unknown*. Let us therefore begin by writing the Kirchhoff current law equations for nodes 1 and 2; if we take the current direction as they happen to be drawn in Fig. 63 we have

$$\text{at node 1:} \quad I_1 - I_2 + I_3 - I_4 = 0 \quad \text{(A)}$$

$$\text{at node 2:} \quad I_4 - I_5 + I_6 = 0 \quad \text{(B)}$$

Now apply eq. (68) to each current at nodes 1 and 2, then substitute into (A) and (B). Doing this (paying careful attention to battery polarities), eqs. (A) and (B) become

$$\frac{8 - V_1}{6} - \frac{V_1}{15} + \frac{6 - V_1}{5} - \frac{V_1 - V_2}{8} = 0$$

and

$$\frac{V_1 - V_2}{8} - \frac{V_2 + 12}{9} + \frac{10 - V_2}{7} = 0$$

Now multiply the first equation by 240 and the second by 504; this should give you the two simultaneous equations

$$67V_1 - 15V_2 = 304$$

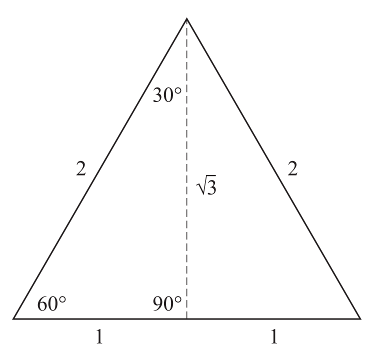
$$-63V_1 + 191V_2 = 48$$

the solutions of which are (using determinants is probably easiest)

$$V_1 = 4.9598 \text{ volts, } \textit{answer}\text{, and } V_2 = 1.8873 \text{ volts, } \textit{answer}\text{,}$$

both voltages with respect to the zero-volt reference node.

61. Since the trigonometric functions are defined in terms of the *ratios* of the lengths of the sides of a right triangle to one another, let us, for simplicity, use the triangle below.



(By the Pythagorean theorem, the length of the vertical dashed line is equal to $\sqrt{3}$, as shown.) The *answers* are now found as follows. First, if we take the 60° angle as the "reference angle," we have (see the definitions just prior to eq. 69)

$$\sin 60^\circ = \sqrt{3}/2 = 0.866025 \text{ approx.}$$

and

$$\cos 60^\circ = 1/2 = 0.500000$$

Next, taking the 30° angle as the reference angle, we have

$$\sin 30^\circ = 1/2 = 0.500000$$

and

$$\cos 30^\circ = \sqrt{3}/2 = 0.866025, \text{ approx.}$$

62. First, by definition, $\sin\theta = b/h$ (using θ as reference angle); also, by definition, $\cos\phi = b/h$ (using ϕ as reference angle).

Thus, $\sin\theta = \cos\phi$, or, since $\phi = 90 - \theta$, we have

$$\sin\theta = \cos(90^\circ - \theta), \textit{ answer}.$$

63. From the answer to problem 62 we have that $\sin 62^\circ 38' = \cos(90^\circ - 62^\circ 38')$. Thus the required angle is $90^\circ - 62^\circ 38'$, and hence (note that $90^\circ = 89^\circ 60'$) we have

$$\begin{array}{r} 89^\circ 60' \\ -\ \underline{62^\circ 38'} \\ 27^\circ 22' \end{array}, \textit{ answer}.$$

64. (a) In the right triangle of Fig. 65, by the theorem of Pythagoras, $a^2 + b^2 = h^2$. However, by eqs. (69) and (70), $a = h\cos\theta$ and $b = h\sin\theta$. Making these substitutions gives the required identity.

(b) Using eq. (71) and then eqs. (69) and (70), we have

$$\frac{b}{a} = \tan\theta = \frac{h\sin\theta}{h\cos\theta} = \frac{\sin\theta}{\cos\theta}, \text{ as required.}$$

65. (a) 115 degrees is a second quadrant angle, hence (Fig. 76) $\phi = 180 - 115 = 65$; thus

$$\cos 115 = -\cos 65 = -0.4226, \textit{ answer}.$$

(b) By eq. (76),

$$\sin(-35) = -\sin 35 = -0.5736, \textit{ answer}.$$

(c) From Fig. 76,

$$\tan 155 = -\tan 25 = -0.4663, \textit{ answer}.$$

(d) 255 degrees is a third quadrant angle, hence (Fig. 77) $\phi = 255 - 180 = 75$; thus

$$\sin 255 = -\sin 75 = -0.9659, \textit{ answer}.$$

(e) From Fig. 76,

$$\cos 95 = -\cos 85 = -0.0872, \textit{ answer}.$$

(f) By eq. (78), $\tan(-285) = -\tan 285$. Since 285 degrees is a fourth quadrant angle, we have (Fig. 78),

$$-\tan 285 = -(-\tan 75) = 3.7321, \textit{ answer}.$$

(g) $\sin 285 = -\sin 75 = -0.9659$, *answer*.

(h) By eq. (76), $\sin(-188) = -\sin 188$. Since 188 is a third quadrant angle, we have (Fig. 77)

$$-\sin 188 = -(-\sin 8) = 0.1392, \textit{ answer}.$$

66. Let us make use of the right triangle of Fig. 64 and eqs. (69) and (70) in section 5.2, as follows.

First,

$$b = h \sin\theta \text{ (with } \theta \text{ as reference angle)},$$
$$b = h \cos\phi \text{ (with } \phi \text{ as reference angle)};$$

thus

$$\sin\theta = \cos\phi, \text{ or, since } \phi = 90 - \theta,$$
$$\sin\theta = \cos(90 - \theta), \text{ as required.}$$

Next,

$$a = h \cos\ \theta \text{ (with } \theta \text{ as reference angle)},$$
$$a = h \sin\phi \text{ (with } \phi \text{ as reference angle)};$$

thus,

$$\cos\ \theta = \sin\phi, \text{ or, since } \phi = 90 - \theta,$$
$$\cos\theta = \sin(90 - \theta), \text{ as required.}$$

67. (a) From eq. (81), $360f = 180{,}000$; thus

$$f = 500\,\text{Hz}, \textit{ answer}.$$

(b) By eq. (81), $v = 100\ \sin\ 27{,}000^\circ$. Since $27{,}000/360 = 75$ complete cycles, with "zero degrees left over," we have

$$v = 100\ \sin 0^\circ = 0, \textit{ answer}.$$

68. Here, $\omega = 2\pi f = 533{,}850$, which gives (using ordinary calculator values for π)

$$f = 84{,}964.866\,\text{Hz}, \text{ approx.}$$

Hence, by eq. (91),

$$T = 1/84{,}964.866 = 1.17695 \times 10^{-5} \text{ seconds}$$
$$= 11.7695 \text{ microseconds}, \textit{ answer}.$$

69. From eq. (91), $f = 1/T$. By inspection, T is smaller for B than for A; thus B has the higher frequency, *answer*.

70. The answer is "yes," because there is always an average power P associated with any such waveform. The problem, however, would be to *find* the rms value in such a case. We should emphasize that the simple relationships of eqs. (102) and (103) are true only for sinusoidal voltages and currents. This is one advantage, among others, of using sinusoidal voltages and currents.

71. (a) By eq. (104),

$$P = (120)(8.5) = 1020 \text{ watts}, \textit{ answer}$$

(b) By eq. (102), $V_\text{p} = \sqrt{2}(120)$, and by eq. (103), $I_\text{p} = \sqrt{2}(8.5)$. Hence,

$$\text{peak power} = V_\text{p} I_\text{p} = 2(120)(8.5) = 2040 \text{ watts}, \textit{ answer}.$$

72. (a) First, the horizontal and vertical components of each generator voltage are

for V_1: $V_\text{h} = 65$ volts, $V_\text{v} = 0$ volts;

for V_2: $V_\text{h} = 90 \cos 60^\circ = 45.000$ volts, $V_\text{v} = 90 \sin 60^\circ = 77.942$ volts;

for V_3: $V_\text{h} = 75 \cos 150^\circ = -64.952$ volts, $V_\text{v} = 75 \sin 150^\circ = 37.500$ volts.

Hence the *horizontal* component of the output voltage $\bar{V}$ is $65 + 45 - 64.952 = 45.048$ volts, and the *vertical* component of $\bar{V}$ is $0 + 77.942 + 37.500 = 115.442$ volts. The polar notation for $\bar{V}$ is of the form $\bar{V} = |\bar{V}|\underline{/\theta}$, and thus, using eqs. (109) and (110), we have

$$|\bar{V}| = \sqrt{15{,}356.178} = 123.920 \text{ volts}$$

and

$$\theta = \arctan(115.442/45.048) = 68.683 \text{ degrees}$$

hence the *answer* is

$$\bar{V} = 123.920\underline{/68.683^\circ}$$

(b) 123.920 volts, *answer*, because ac voltmeters (and ammeters) read magnitude of rms values.

(c) Upon substituting the answer to part (a) into eq. (111) we have

$$\bar{I} = \frac{123.920}{25}\underline{/68.683^\circ} = 4.957\underline{/68.683} \text{ amperes, } \textit{answer}.$$

Since ac ammeters read magnitude of rms current, the *second answer* is 4.957 amperes.

(d) The *answer* is "yes," because, from part (c), the phase angle of $\bar{I}$ is 68.683°, the same as for $\bar{V}$ in part (a). This is true because $\bar{V}$ is the resultant voltage *across the resistance R,* and there is always zero phase shift between voltage across a resistance and the current through the resistance (as emphasized in the discussion following Fig. 98).

(e) The *answer* is "no," because the phase angles of the generator voltages are given to be 0°, 60°, and 150°, while, from part (c), the phase angle of $\bar{I}$ is 68.683°. (This is a similar situation to Fig. 101, in which $\bar{I}$ is not in phase with either generator voltage $\bar{V}_1$ or $\bar{V}_2$.)

73. (a) From problem 72, part (a), the magnitude of the output voltage across the 25-ohm resistive load is $V = 123.920$ volts rms, and from part (c) the magnitude of the current in the 25-ohm load is $I = 4.957$ amperes rms. Since the load is a pure *resistance*, $\bar{V}$ and $\bar{I}$ are IN PHASE with each other, making $\cos\theta = \cos 0 = 1.000$, and thus, by eq. (117), we have

$$P = VI = (123.920)(4.957) = 614.27 \text{ watts approx., } \textit{answer}.$$

(b) Let us apply eq. (117) to each generator individually, as follows. First note that, since Fig. 102 is a series circuit, the *same current*, $\bar{I} = 4.957\underline{/68.683^\circ}$, flows through each generator. Since the phase angles of the generators are given to be $0^\circ, 60^\circ$, and 150°, the phase angles between the generator voltages and current are, in each case, equal to

$$\text{for } V_1\text{: } \theta = 68.683^\circ$$

$$\text{for } V_2\text{: } \theta = 68.683 - 60 = 8.683^\circ$$

$$\text{for } V_3\text{: } \theta = 150 - 68.683 = 81.317^\circ$$

Thus, applying eq. (117), the *answers* are

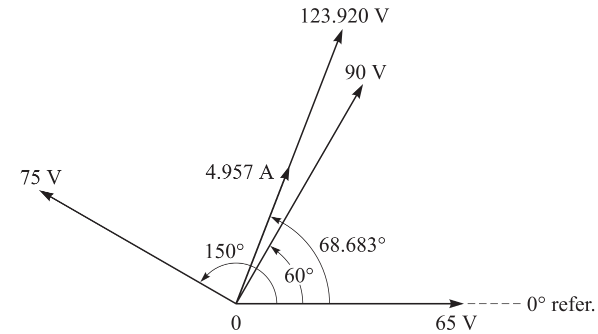
$$\text{power produced by } V_1 = (65)(4.957)\cos 68.683^\circ = 117.130 \text{ watts},$$

$$\text{power produced by } V_2 = (90)(4.957)\cos 8.683^\circ = 441.017 \text{ watts},$$

$$\text{power produced by } V_3 = (75)(4.957)\cos 81.317^\circ = 56.126 \text{ watts}.$$

(c) The sum in part (b), to two decimal places, is 614.27 watts, which checks with the answer found in part (a).

74. The only change in eqs. (129) and (130) would be that the $-\bar{V}_3$ term would become
(d)



$+\bar{V}_3$. This would cause the term $-25.456\bar{V}_3$ in eq. (132) to become $+25.456\bar{V}_3$. Then, upon making use of eq. (128) and applying the same procedure as in the original solution, you should find that

$$\bar{I}_2 = \frac{1874.492}{591.991} \underline{/62.864^\circ} = 3.166\underline{/62.864^\circ}$$

thus

$$|\bar{V}_o| = (12)(3.166) = 37.992 \text{ volts approx., } \textit{answer}.$$

75. (a) $\pm j12$, *answer* (procedure same as in example 1, part (a)).

(b) Using fractional exponents and the same procedure as in example 1 (b), we have

$$\pm j10x^2y^{-5} = \pm j10x^2/y^5, \ \textit{answer}.$$

(c) $[(-1)(4)^{-1}x^2y^{-2}z^{-4}]^{1/2} = \pm jx/2yz^2$, *answer*; or, since

$$\left(\frac{A}{B}\right)^{1/2} = \frac{A^{1/2}}{B^{1/2}}$$

we have $\pm \dfrac{jx}{2yz^2}$, *same answer*.

76. (a) Since $j^2 = -1$, we have, $-j(-1) = j$, *answer*.

(b) Upon adding exponents (see note following eq. (136)), we have j^{13}. Since 4 goes into 13 "3 times with 1 left over," we have $j^{13} = j$, *answer*.

(c) Since 4 goes into 31 "7 times with 3 left over," $j^{31} = j^3 = -j$, *answer*.

(d) Since 4 goes into 342 "85 times with 2 left over," $-j^{342} = -j^2 = +1$, *answer*.

(e) $\dfrac{1}{j^3} = \dfrac{j}{j^4} = j$, *answer*.

(f) $\dfrac{1}{j^{34}} = \dfrac{1}{j^2} = \dfrac{1}{-1} = -1$, *answer*.

(g) $\dfrac{1}{j^{17}} = \dfrac{1}{j} = -j$, *answer*, by eq. (139).

77. (a) Transposing, then taking the square root,

$$x = \sqrt{-16} = \pm j4, \ \textit{answers}.$$

(b) Transposing, then taking the square root,

$$x = \sqrt{900} = \pm 30, \ \textit{answers}.$$

Thus the answers here are *real*, $x = 30$ and $x = -30$.

(c) Here, $x = \sqrt{-94.8/6} = \sqrt{-15.8} = \pm j3.9749$, *answers*.

78. $(6 - 8 + 4) + j(5 - 4 - 3) = 2 - j2 = 2(1 - j)$, *answer*.

79. Since $j^3 = -j, j^2 = -1$, and $j^{100} = 1$, the problem becomes

$$\begin{aligned} j5 - 7 - j + 1 - j10 + 4 + 3 + 10 &= (-7 + 1 + 4 + 3 + 10) + j(5 - 1 - 10) \\ &= 11 - j6, \ \textit{answer}. \end{aligned}$$

80. $12 - j2 - j18 - 3 = (9 - j20)$, *answer*.

81. The product of the first two factors is $(-1 + j3)$, and multiplying this result by the third factor gives

$$-12 - j14 = -2(6 + j7), \ \textit{answer}.$$

82. $ac + jad + jbc - bd = (ac - bd) + j(ad + bc)$, *answer*.

83. The easiest way is to first factor 6 from both parts of the complex number and then square; thus

$$\begin{aligned} [6(1 + j2)]^2 = (6)^2(1 + j2)^2 &= 36(1 + j2)(1 + j2) \\ &= 36(-3 + j4), \ \textit{answer}. \end{aligned}$$

84. First note that $(1 + j)^2 = (1 + j)(1 + j) = j2$. Hence the given problem can be written in the form

$$(1 + j)^5 = (j2)(j2)(1 + j) = -4(1 + j), \ \textit{answer}.$$

85. (a) $\dfrac{(3 + j4)(1 - j)}{(1 + j)(1 - j)} = \dfrac{7 + j}{2} = 0.5(7 + j)$, *answer*.

(b) The conjugate of $0 + j5$ is $0 - j5 = -j5$; hence

$$\frac{(14 - j25)(-j5)}{(j5)(-j5)} = \frac{-125 - j70}{25} = -(5 + j2.8), \ \textit{answer}.$$

86. Upon multiplying in the numerator and denominator as indicated, then collecting like terms, we have

$$\frac{9-j2}{5+j} = \frac{(9-j2)(5-j)}{(5+j)(5-j)} = \frac{43-j19}{26} = 1.6539 - j0.7308, \ \textit{answer}.$$

87. The easiest way is to note that $(1+j)^2 = j2$; the problem then becomes

$$\frac{j12}{(j2)(j2)} = \frac{j12}{-4} = -j3, \ \textit{answer}.$$

88. Yes, because two EQUAL complex numbers are represented by the SAME POINT on the complex plane, so that $x = a$ for both numbers and $y = b$ for both numbers.

89. (a) Applying the SUM rule of Article 2, the problem becomes $8.1 + j10.4$. Hence, by eqs. (143) and (144), $A = 13.182$ and $\theta = \arctan(10.4/8.1) = 52.087°$; thus we have

$$13.182(\cos 52.087° + j \sin 52.087°), \ \textit{answer}.$$

(b) First, by section 6.3,

$$\frac{(7-j2)(4-j9)}{(4+j9)(4-j9)} = \frac{10-j71}{97} = 0.103 - j0.732$$

Hence, by eqs. (143) and (144)

$$A = 0.739 \quad \text{and} \quad \theta = \arctan(-7.107) = -81.99°$$

and thus (making use of eqs. (77) and (76)) we have

$$0.739(\cos 81.99° - j \sin 81.99°), \ \textit{answer}.$$

90. (a) $90(\cos 166° + j \sin 166°) = -87.327 + j21.773$, *answer*.

(b) Since $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$, let us write

$$400(\cos 126° - j \sin 126°) = -(235.114 + j323.607), \ \textit{answer}$$

(c) $17\underline{/45°} = 12.021 + j12.021$ and $-22\underline{/265°} = 1.917 + j21.916$. Hence, by section 6.2, we have

$$13.938 + j33.937, \ \textit{answer}.$$

91.

$$\text{Total REAL COMPONENT} = 16\cos 36° - 22\cos 315° + 9.15 = 6.5379$$

$$\text{total IMAGINARY COMPONENT} = 16\sin 36° - 22\sin 315° - 6.88 = 18.0809$$

thus we have

$$(6.5379 + j18.0809), \ \textit{answer}.$$

92. In all four cases $|b| = 4$ and $|a| = 3$; thus, in all four cases, $A = 5$ and $h = \arctan 4/3 = 53.13°$. Hence the four *answers* are (noting the quadrant in each case)

(a) $5\epsilon^{j53.13°}$ (b) $5\epsilon^{j126.87°}$ (c) $5\epsilon^{j233.13°}$ (d) $5\epsilon^{j306.87°}$

Verification of the above answers, making use of eq. (153):

(a) $5(\cos 53.13^\circ + j\sin 53.13^\circ) = 3 + j4$, which checks,

(b) $5(\cos 126.87^\circ + j\sin 126.87^\circ) = -3 + j4$, which checks,

(c) $5(\cos 233.13^\circ + j\sin 233.13^\circ) = -3 - j4$, which checks,

(d) $5(\cos 306.87^\circ + j\sin 306.87^\circ) = 3 - j4$, which checks.

93. To find the SUM of a number of complex numbers we must first express each one in the form of eq. (156) or (158). We can do this by applying eq. (153) to each term in the given problem; thus

$$14(\cos 112^\circ + j\sin 112^\circ) = -5.2445 + j12.9806$$

$$8(\cos 28^\circ + j\sin 28^\circ) = 7.0636 + j3.7558$$

$$19(\cos 155^\circ - j\sin 155^\circ) = -17.2199 - j8.0298$$

in which we made use of the identities $\cos(-x) = \cos x$ and $\sin(-x) = -\sin x$. Thus the answer in the form of eq. (156) is $-15.4008 + j8.7066$, which lies in the SECOND QUADRANT of the complex plane, where $a = -15.4008$ and $b = 8.7066$ (see Fig. 110).

First, therefore, by eq. (143), $A = 17.6915$. Next we have that $h = \arctan(8.7066/15.4008) = 29.4810^\circ$, hence $\theta = 180 - h = 150.52^\circ$, and thus the required *answer* is (rounded to two decimal places)

$$17.69\epsilon^{j150.52^\circ}$$

94. *First* apply eq. (161) (extended to cover three factors), *then* apply eq. (158), thus getting

$$\begin{aligned} 84\epsilon^{j215^\circ} &= 84(\cos 215^\circ + j\sin 215^\circ) \\ &= -(68.809 + j48.180), \ \textit{answer}. \end{aligned}$$

95. By eq. (164), and also eqs. (157) and (158), we have

$$12\underline{/58^\circ} = 12(\cos 58^\circ + j\sin 58^\circ) = 6.359 + j10.177, \ \textit{answer}.$$

96. By eqs. (143) and (144),

$$2 + j3 = \sqrt{13}\epsilon^{j56.3099^\circ}$$

Hence, making use of the basic relationship $(A\epsilon^{j\theta})^n = A^n\epsilon^{jn\theta}$, we have that

$$\begin{aligned} (2 + j3)^6 = (\sqrt{13})^6\epsilon^{j337.86^\circ} &= 2197(\cos 337.86^\circ + j\sin 337.86^\circ) \\ &= 2035.01 - j827.99, \ \textit{answer}. \end{aligned}$$

97. By eq. (163), noting that $\cos(-x) = \cos x$ and $\sin(-x) = -\sin x$,

$$\begin{aligned} \left(\frac{15}{36}\right)\epsilon^{-j23^\circ} &= \frac{15}{36}(\cos 23^\circ - j\sin 23^\circ) \\ &= 0.3835 - j0.1628, \ \textit{answer}. \end{aligned}$$

98. One way is to write the problem as the difference of two fractions, then apply eq. (165), thus

$$\frac{16\angle 102^\circ}{7\angle 75^\circ} - \frac{9\angle 390^\circ}{7\angle 75^\circ} = \frac{16}{7}\angle 27^\circ - \frac{9}{7}\angle 315^\circ$$

each term now being in the form of eq. (157). Now applying eq. (158) we have

$$\frac{16}{7}(\cos 27^\circ + j\,\sin 27^\circ) - \frac{9}{7}(\cos 315^\circ + j\,\sin 315^\circ) = 1.1275 + j1.9468,\ \textit{answer}.$$

99. Setting $n = 2$ in eq. (166) (and writing "x" in place of "θ") you should find that

$$(\cos^2 x - \sin^2 x) + j(2\,\sin x\,\cos x) = \cos 2x + j\,\sin 2x$$

hence we have

(a) $\cos 2x = \cos^2 x - \sin^2 x$, *answer*.

(b) $\sin 2x = 2\sin x\,\cos x$, *answer*.

The above illustrates the fact that the algebra of complex numbers often leads to important and entirely REAL results.

100. Noting that $2^5 = 32$, and upon setting $\theta = 48^\circ$ and $n = 5$ in eq. (166), we have

$$32(\cos 240^\circ + j\,\sin 240^\circ) = -(16.000 + j27.713),\ \textit{answer}.$$

101. Writing the problem as $0.25(\cos 17^\circ + j\,\sin 17^\circ)^{-3}$, then applying eq. (166) for $n = -3$, and remembering that $\cos(-x) = \cos x$ and $\sin(-x) = -\sin x$, we have

$$0.25(\cos 51^\circ - j\,\sin 51^\circ) = 0.157 - j0.194,\ \textit{answer}.$$

102. First, by eq. (153),

$$\begin{aligned}\epsilon^{jx}\epsilon^{jy} &= (\cos x + j\,\sin x)(\cos y + j\,\sin y)\\ &= (\cos x\,\cos y - \sin x\,\sin y) + j(\sin x\,\cos y + \cos x\,\sin y)\end{aligned}$$

By eqs. (161) and (153), or by putting $A = B = 1$ in eq. (161), we find it's also true that

$$\epsilon^{jx}\epsilon^{jy} = \epsilon^{j(x+y)} = \cos(x+y) + j\,\sin(x+y)$$

Thus the right-hand sides of the last two equations are equal, and hence, invoking the principle of problem 99, we have

(a) $\cos(x+y) = \cos x\,\cos y - \sin x\,\sin y$, *answer*.

(b) $\sin(x+y) = \sin x\,\cos y + \cos x\,\sin y$, *answer*.

103. Since $3 - j2$ lies in the FOURTH QUADRANT of the complex plane, we have (see Fig. 110)

$$h = \arctan(2/3) = 33.6901^\circ$$

thus,

$$\theta = 360 - h = 326.3099^\circ$$

Hence, by eqs. (168), (143), and (144), we have

$$\begin{aligned}(3-j2)^7 &= (\sqrt{13})^7(\cos 2284.17^\circ + j\ \sin 2284.17^\circ)\\ &= 7921.40(\cos 124.17^\circ + j\ \sin 124.17^\circ)\\ &= (-4449.06 + j6553.97),\ \textit{answer}.\end{aligned}$$

104. We begin by writing the problem in the form $6000(3+j4)^{-5}$. Thus $n=-5$ in eq. (168). Next, since $3+j4$ lies in the first quadrant, we have (Fig. 110) $\theta = h = \arctan(4/3) = 53.1301^\circ$. Also, by eq. (143), $A=\sqrt{25}=5$. Thus, using these values and setting $n=-5$ in the right-hand side of eq. (168), and noting that $\cos(-x)=\cos x$ and $\sin(-x) = -\sin x$, we have

$$6000(5)^{-5}(\cos 265.65^\circ - j\ \sin 265.65^\circ) = -0.1457 + j1.9144 \text{ approx.},\ \textit{answer}.$$

105. The four "roots" are the values of $(3+j7)^{1/4}$. Thus, for use in eq. (176), we have, for this problem, $a=3, b=7$, and $n=4$. We can proceed as follows.

First, the magnitude A of the complex number $(3+j7)$ is, by eq. (143), equal to

$$A=\sqrt{58}=(58)^{1/2}$$

Thus, for this problem ($n=4$), we have that

$$A^{1/4} = (58)^{1/8} = 1.6612 \text{ approx., by calculator.}$$

Next, noting that $(3+j7)$ lies in the first quadrant of the complex plane, the value of θ is, by eq. (144), equal to

$$\theta = \arctan(7/3) = 66.8014^\circ \text{ approx.,}$$

thus

$$\frac{\theta}{n} = \frac{66.8014^\circ}{4} = 16.7004^\circ = 16.7^\circ \text{ approx.}$$

Let us now denote the four roots by r_1, r_2, r_3, and r_4. Since, in this problem, $n=4$, the four values of k to be substituted into eq. (176) are $k=0,1,2$, and 3, and upon doing this we find that

for $k=0$: $r_1 = 1.6612(\cos 16.7^\circ + j\ \sin 16.7^\circ) = 1.591 + j0.477$, *answer*.

for $k=1$: $r_2 = 1.6612(\cos 106.7^\circ + j\ \sin 106.7^\circ) = -0.477 + j1.591$, *answer*.

for $k=2$: $r_3 = 1.6612(\cos 196.7^\circ + j\ \sin 196.7^\circ) = -1.591 - j0.477$, *answer*.

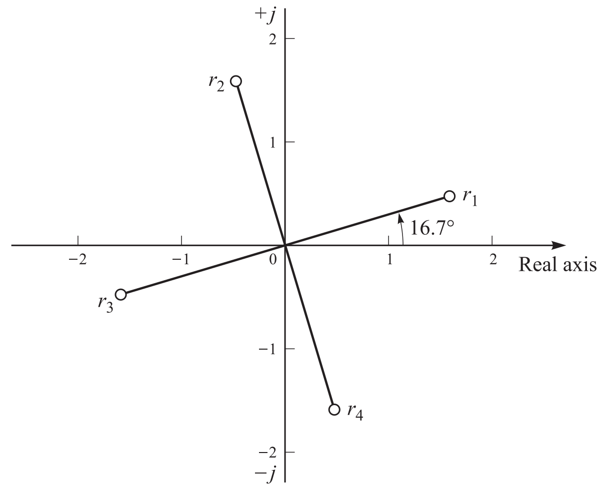
for $k=3$: $r_4 = 1.6612(\cos 286.7^\circ + j\ \sin 286.7^\circ) = 0.477 - j1.591$, *answer*.

The locations of the four roots on the complex plane are shown in the figure below.

In the figure, the LENGTH of a line drawn from the origin to any point represents the MAGNITUDE of the complex number associated with that point. Reference to eq. (176) shows that *all roots will have the* SAME MAGNITUDE ($A^{1/n}$), and this fact is evident, geometrically, from inspection of the following figure.

Next, the ANGULAR factors associated with each root can be clearly seen by writing eq. (176) in the exponential form, thus

$$(a+jb)^{1/n} = (A^{1/n})\epsilon^{j(\theta/n+360k/n)^\circ}$$



Note that the angle for the FIRST ROOT (the root for $k = 0$) is equal to θ/n, where θ (eq. (144)) is relative to the real axis; here, in this problem (where $n = 4$), we have $\theta/n = \theta/4 = 16.7°$, as shown in the figure.

Next notice that $360(k/n) = (360/n)k$ is the ANGULAR SEPARATION, in degrees, between the lines drawn from the origin to the points at the root locations. Since n is a given constant in any given problem, the angular separation is *the* SAME *for all roots* in any given problem. Thus, in the present problem, where $n = 4$, the angular separation is $(90k)°$, that is, *90 degrees*, as inspection of the figure shows.

106. Here $a = 19, b = -33, n = 5$. Thus, first, we have

$$A^{1/n} = (1450)^{1/10} = 2.07 \text{ approx.}$$

Next, since the point $(19, -33)$ lies in the fourth quadrant (Fig. 110), we have

$$h = \arctan(33/19) = 60.07°$$

thus,

$$\theta = 360 - h = 299.93° \text{ approx.}$$

therefore, $\theta/n = 59.99°$, and hence eq. (176) becomes

$$\begin{aligned}(19 - j33)^{1/5} &= 2.07[\cos(59.99 + 72k)° + j\sin(59.99 + 72k)°] \\ &= 2.07\underline{/(59.99 + 72k)°} \text{ (see eqs. (157) and (158), section 6.5)}\end{aligned}$$

(see eqs. (157) and (158), section 6.5)

Now, setting, successively, $k = 0, 1, 2, 3$, and 4 into the last expressions gives

for $k = 0$: $2.07\underline{/59.99°} = 1.04 + j1.79$, *answer*.

for $k = 1$: $2.07\underline{/131.99°} = -1.39 + j1.54$, *answer*.

for $k = 2$: $2.07\underline{/203.99°} = -1.89 - j0.84$, *answer*.

for $k = 3$: $2.07\underline{/275.99°} = 0.22 - j2.06$, *answer*.

for $k = 4$: $2.07\underline{/347.99°} = 2.03 - j0.43$, *answer*.

Thus all five roots have the same magnitude, 2.07. To show the above results graphically, draw a circle with center at the origin of the complex plane, and radius of any convenient length to represent 2.07. The five roots are distributed around the circumference of the circle, 72° apart, in accordance with the above angles.

107. We ordinarily say that "1 raised to a power is 1"; thus, it might seem, offhand, that the ONLY possible answer is that $(1)^{1/3} = 1$. This, however, assumes that the answer has to be a *real* number, which is *not true*, because in mathematical applications "one number is as valid as any other number" on the total number plane of Fig. 108. Thus, as we'll now find, the COMPLETE cube root of 1 is equal to the real number 1, plus *two other roots*, both complex numbers.

To see why this is true, let us begin by noting that $1 = 1 + j0$; thus, "1" can be regarded as being a complex number, $a + jb$, having $a = 1$ and $b = 0$. Then

$$A = (1)^{1/2} = 1 \quad \text{(by eq. (177))}$$

and

$$\theta = \arctan 0 = 0^\circ \quad \text{(by eq. (178))}$$

thus $(n = 3)$

$$A^{1/n} = (1)^{1/6} = 1 \quad \text{(by eq. (179))}$$

Then, since $\theta/n = 0/n = 0$, eq. (176) becomes

$$(1)^{1/3} = (1)[\cos(120k)^\circ + j\sin(120k)^\circ] = 1\underline{/(120k)^\circ}$$

which, upon successively setting $k = 0, 1$, and 2, gives the "three cube roots of unity," thus

for $k = 0$: $r_1 = \cos 0 + j\sin 0 = 1.000$, *answer*.

for $k = 1$: $r_2 = \cos 120^\circ + j\sin 120^\circ = -0.500 + j0.866$, *answer*.

for $k = 2$: $r_3 = \cos 240^\circ + j\sin 240^\circ = -0.500 - j0.866$, *answer*.

108. The purpose here is to emphasize that induced voltage depends *not* upon the *amount* of current, but only upon the *rate of change* of current. Here, in both (a) and (b), the current is changing at the same constant rate of 2 *amperes per second*, $di/dt = 2$ amp/sec. Hence, by eq. (181), in both cases, $v = (0.62)(2) = 1.24$ volts, *answer*.

109. By eq. (181),

$$di/dt = v/L = 5.52/0.62 = 8.903 \text{ amp/sec, } \textit{answer}.$$

110. By eq. (181),

$$L = v/(di/dt) = 0.048/76 = 6.316 \times 10^{-4} \text{ henrys}$$
$$= 631.6 \text{ microhenrys, } \textit{answer}.$$

111. Since 0.00065 meters = 0.65 mm, and since $E = 3000$ volts per mm, we have that

$$v = (3000)(0.65) = 1950 \text{ volts, } \textit{answer}.$$

112. $0.015\,\mu F = 1.5\times 10^{-2}\times 10^{-6}\,F = 1.5\times 10^{-8}\,F$; thus, by eq. (184),

$$q = Cv = 1.5\times 10^{-8}\times 2.9\times 10^{2} = 4.35\times 10^{-6} \text{ coulombs, } \textit{answer}.$$

113. (a) Using either eq. (188) or (189), you should find that

$$C_T = 0.0368\,\mu F \text{ approx., } \textit{answer}.$$

(b) By eq. (191),

$$C_T = 0.93\,\mu F, \ \textit{answer}.$$

114. We must use eq. (190), but to do this we *first* need to know the value of C_T. Upon using either eq. (188) or (189), you should find that $C_T = 0.03934\,\mu F$ approx. Hence, for $V = 450$ volts and $C_T = 0.03934\,\mu F$, eq. (190) becomes

$$V_x = 17.703/C_x \text{ volts}$$

where, since C_T is in mfd, C_x is also in mfd.

Thus we have that

$$V_1 = 17.703/0.15 = 118.02 \text{ volts},$$
$$V_2 = 17.703/0.06 = 295.05 \text{ volts},$$
$$V_3 = 17.703/0.48 = 36.88 \text{ volts}.$$

Since none of the capacitor voltages will exceed 300 volts, it *is* theoretically proper to specify a capacitor voltage rating of 300 volts.

115. The three capacitors are, together, equivalent to a *single capacitor* of capacitance $C_T = 0.03934\times 10^{-6}$ farads. Hence, by eq. (185),

$$W = \tfrac{1}{2}(0.03934)(10^{-6})(450)^2 = 3.983\times 10^{-3} \text{ joules, } \textit{answer}.$$

116. (a) Here, $\bar{V} = V = 115$ volts, $R = 28$ ohms, $\omega L = 2\pi(60)(0.12) = 45.2389$ ohms. Now substituting these values into eq. (197) we have

$$\bar{I} = \frac{115}{28 + j45.2389}$$

hence

$$|\bar{I}| = \frac{115}{\sqrt{(28)^2 + (45.2389)^2}} = 2.1615 \text{ amperes, } \textit{answer}.$$

(b) By eq. (203), $\phi = \arctan(45.2389)/28 = 58.245^\circ$, *answer*.

(c) Voltmeters are calibrated to read rms volts; thus (see Fig. 129)

$$V_L = \omega L I = (45.2389)(2.1615) = 97.784 \text{ volts rms, } \textit{answer}.$$

117. Here,

$$\bar{V} = V = 95 \text{ volts}, \ R_T = 30 \text{ ohms}, \ \omega L = 10^3(37)10^{-3} = 37 \text{ ohms } (1\,\text{mH} = 10^{-3}\,\text{H}).$$

Hence

(a) $\bar{I} = \dfrac{95}{30 + j37}$, thus

$$|\bar{I}| = \frac{95}{\sqrt{(30)^2 + (37)^2}} = 1.994 \text{ amperes, } \textit{answer}.$$

(b) By eq. (203),

$$\phi = \arctan(37/30) = 50.965^\circ \text{ approx., } \textit{answer}.$$

(c) Let L' ("L prime") denote the 25 mH inductor; then

$$V = \omega L' I = (10^3)(25)(10^{-3})(1.994) = 49.85 \text{ volts rms, } \textit{answer}.$$

118. (a) Setting $\bar{Z}_1 = 8$ ohms, $\bar{Z}_2 = j12$ ohms, $\bar{Z}_3 = (4 + j6)$ ohms in eq. (207), you should find that

$$\frac{1}{\bar{Z}_\mathrm{T}} = \frac{5 - j24}{24(3 - j2)}$$

thus

$$\bar{Z}_\mathrm{T} = \frac{24(3 - j2)}{5 - j24}$$

which, upon rationalizing (section 6.3), becomes

$$\bar{Z}_\mathrm{T} = (2.516 + j2.476) \text{ ohms approx., } \textit{answer}.$$

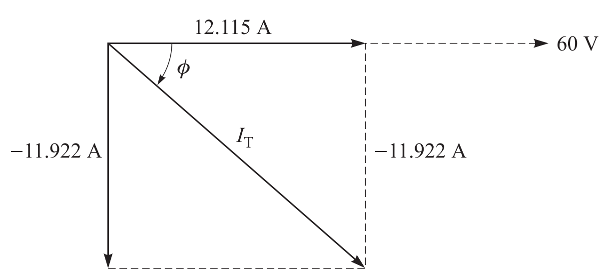
(b) By eq. (206),

$$\bar{I}_\mathrm{T} = \frac{60}{2.516 + j2.476}$$

which, after rationalizing, becomes

$$\bar{I}_\mathrm{T} = \frac{60(2.516 - j2.476)}{12.461} = (12.115 - j11.922) \text{ amperes, } \textit{answer}.$$

(c) From part (b) we have the vector diagram:



thus

$$\phi = \arctan\frac{-11.922}{12.115} = -44.54^\circ \text{ approx.}$$

meaning that $\bar{I}_\mathrm{T}$ "lags" $\bar{V}$ by approximately 44.54°, *answer*. (In polar form, $\bar{I}_\mathrm{T} = 16.997\underline{/-44.54^\circ}$.)

(d) $\bar{I}_1 = \dfrac{60}{8} = 7.50$ amperes, *answer*. (In polar form, $7.50\underline{/0^\circ}$)

(e) $\bar{I}_2 = \dfrac{60}{j12} = -j5$ amperes, *answer*. (In polar form, $5\underline{/-90^\circ}$)

(f) $\bar{I}_3 = \dfrac{60}{4+j6} = \dfrac{30}{2+j3}$ which, upon rationalizing, becomes

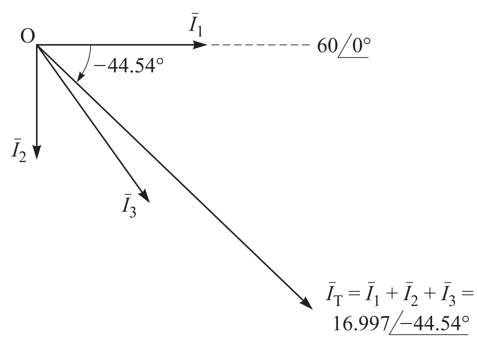
$$\bar{I}_3 = \frac{30(2-j3)}{13} = (4.615 - j6.923) \text{ amperes, } \textit{answer}.$$

(In polar form, $8.320\underline{/-56.31^\circ}$)

(g) By eq. (206),

$$\begin{aligned}\bar{I}_T &= 7.50 - j5.00 + 4.615 - j6.923 \\ &= (12.115 - j11.923) \text{ amperes, which } \textit{does check}.\end{aligned}$$

(h) Relative to the reference voltage $60\underline{/0^\circ}$, the vector diagram can be drawn as follows:



The basic graphical procedures is as follows.

First find the vector sum of $\bar{I}_1$ and $\bar{I}_2$, then combine this vector with the $\bar{I}_3$ vector to get the final resultant vector $\bar{I}_T$.

119. For this particular case ($\bar{Z}_1 = R, \bar{Z}_2 = j\omega L$), eq. (209) becomes ($\bar{V} = V\underline{/0^\circ} =$ real number V)

$$\bar{Z}_T = \frac{jR\omega L}{R + j\omega L}$$

thus

$$\bar{I}_T = \frac{V}{\bar{Z}_T} = \frac{V(R + j\omega L)}{jR\omega L}$$

Now, to express the current in the rectangular form $I' + jI''$, all we need do is multiply the numerator and denominator by $-j$, thus getting

$$\bar{I}_T = \frac{V(-jR + \omega L)}{R\omega L} = V\left(\frac{1}{R} - j\frac{1}{\omega L}\right), \ \textit{answer}.$$

120. Drawing the loop currents in the manner of Fig. 135 (denoting $\bar{I}_1$ by $\bar{I}_T$), the three simultaneous equations are

$$\begin{aligned} 8\bar{I}_T \quad & - 8\bar{I}_2 \quad & + 0\bar{I}_3 = 60 \\ -8\bar{I}_T + & (8 + j12)\bar{I}_2 \quad & - j12\bar{I}_3 = 0 \\ 0\bar{I}_T \quad & - j12\bar{I}_2 + & (4 + j18)\bar{I}_3 = 0 \end{aligned}$$

We next must find the value of the "denominator determinant" Δ, thus

$$\Delta = \begin{vmatrix} 8 & -8 & 0 \\ -8 & (8+j12) & -j12 \\ 0 & -j12 & (4+j18) \end{vmatrix} = (8)(4)(2)\begin{vmatrix} 1 & -1 & 0 \\ -2 & (2+j3) & -j3 \\ 0 & -j6 & (2+j9) \end{vmatrix}$$

The easiest way, now, is to multiply each element of the first row by 2 and add the result to the corresponding element of the second row, thus getting

$$\Delta = (64)\begin{vmatrix} 1 & -1 & 0 \\ 0 & j3 & -j3 \\ 0 & -j6 & (2+j9) \end{vmatrix} = j192\begin{vmatrix} 1 & -1 \\ -j6 & (2+j9) \end{vmatrix} = -192(3 - j2)$$

hence

$$\bar{I}_T = \frac{(60)(4)(2)\begin{vmatrix} 1 & -8 & 0 \\ 0 & (2+j3) & -j3 \\ 0 & -j6 & (2+j9) \end{vmatrix}}{-192(3 - j2)} = \frac{480}{192}\,\frac{5 - j24}{3 - j2}$$

thus, rationalizing,

$$\bar{I}_T = (12.115 - j11.923) \text{ amperes (A)}, \ \textit{answer}, \text{ as in problem 118(b).}$$

121. (a) Here, $\omega = 2\pi f = (2.5133)10^4$ rad/sec, $C = (5)10^{-7}$ farads; thus,

$$\frac{1}{\omega C} = \frac{1}{(2.5133)10^4(5)10^{-7}} = 79.577 \text{ ohms}$$

thus, by eq. (219),

$$|\bar{Z}| = \sqrt{16332.50} = 127.799 \text{ ohms, approx.}$$

hence, by eq. (220),

$$|\bar{I}| = \frac{95.00}{127.799} = 0.743 \text{ A}, \ \textit{answer}.$$

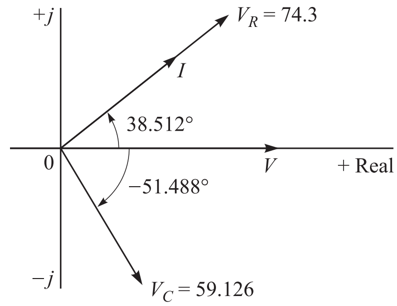
(b) By eq. (221),

$$\phi = \arctan 0.7958 = 38.512^\circ, \ \textit{answer}.$$

(c) Ac meters read magnitude of rms values; hence, voltmeter reading across $R = (0.743)(100) = 74.3$ volts, *answer*.

(d) Voltmeter reading across $C = (1/\omega C)|\bar{I}| = (79.577)(0.743) = 59.126$ volts, *answer*.

(e) The details of the arrival of the full answer are as follows. Using rms values (instead of peak values), Fig. 138 becomes, on the complex plane, for the above values,



Hence

$$\bar{V}_R = 74.300(\cos 38.512^\circ + j\ \sin 38.512^\circ) = (58.138 + j46.265)\ \text{volts}$$

and

$$\bar{V}_C = 59.126(\cos 51.488^\circ - j\ \sin 51.488^\circ) = (36.817 - j46.265)\ \text{volts}$$

thus

$$\bar{V}_R + \bar{V}_C = 94.955 + j0 = 95\underline{/0^\circ}\ \text{approx., the applied reference voltage, } \textit{answer}.$$

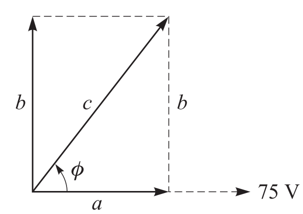
122. (a) Since the value of each capacitance is 0.12×10^{-6} farads, eq. (222) becomes

$$\bar{I} = \frac{75}{36 - \dfrac{j3}{(5)(10^5)(0.12)10^{-6}}} = \frac{37.5}{18 - j25}$$

which, upon rationalizing (multiplying numerator and denominator by conjugate of denominator), becomes

$\bar{I} = 0.7113 + j0.9879$, thus $|\bar{I}| = 1.2173$ A, *answer*.

(b) Here, $\bar{I} = a + jb$ amperes, where a and b have the values found in part (a); thus, from the vector diagram to the right,



$$\phi = \arctan(b/a) = \arctan\frac{0.9879}{0.7113} = 54.246^\circ,\ \textit{answer}$$

(c) $|\bar{V}_C| = |\bar{I}|(1/\omega C) = (1.2173)(1/0.06) = 20.288$ V, *answer*.

123. (a) First, in complex notation,

$$\text{for } 0.25\,\mu\text{F capacitor, } -j(1/\omega C) = -j40 \text{ ohms,}$$

$$\text{for } 0.32\,\mu\text{F capacitor, } -j(1/\omega C) = -j31.25 \text{ ohms.}$$

Here we're asked to find the total (generator) current $\bar{I}_T$; let us do this in two different ways, as follows.

FIRST WAY: Since $\bar{I}_T = \dfrac{\bar{V}}{\bar{Z}_T} = \dfrac{60}{\bar{Z}_T}$, where $\bar{Z}_T$ is the total impedance seen by the generator, let us begin by making use of eq. (207); thus, noting that $1/(-j40) = j/40$,

$$\frac{1}{\bar{Z}_T} = \frac{j}{40} + \frac{1}{32} + \frac{1}{16 - j31.25} = \frac{40 + j32}{1280} + \frac{1}{16 - j31.25}$$

hence

$$\bar{I}_{\mathrm{T}} = \frac{60}{\bar{Z}_{\mathrm{T}}} = \frac{3}{8}(5 + j4) + \frac{60}{16 - j31.25}$$

which, after rationalizing the fraction, becomes

$$\bar{I}_{\mathrm{T}} = \frac{3}{8}(5 + j4) + \frac{60(16 + j31.25)}{1232.5625} = 2.6539 + j3.0212 \text{ A approx.}$$

which is a leading current of magnitude $|\bar{I}_{\mathrm{T}}| = 4.021\,\mathrm{A}$, *answer*.

SECOND WAY: Let us now apply the "loop current" procedure, explained in connection with Fig. 135 and problem 120. Using the same current notation and reference directions as in Fig. 135 (denoting $\bar{I}_1$ by $\bar{I}_{\mathrm{T}}$), the three simultaneous equations for Fig. 140 are

$$\begin{aligned}
-j40\bar{I}_{\mathrm{T}} + \quad j40\bar{I}_2 \quad + 0\bar{I}_3 &= 60\\
j40\bar{I}_{\mathrm{T}} + (32 - j40)\bar{I}_2 + \quad -32\bar{I}_3 &= 0\\
0\bar{I}_{\mathrm{T}} \quad - 32\bar{I}_2 + (48 - j31.25)\bar{I}_3 &= 0
\end{aligned}$$

$$\begin{aligned}
-j2\bar{I}_{\mathrm{T}} \quad + j2\bar{I}_2 \quad + 0\bar{I}_3 &= 3\\
= \quad j5\bar{I}_{\mathrm{T}} + (4 - j5)\bar{I}_2 \quad - 4\bar{I}_3 &= 0\\
0\bar{I}_{\mathrm{T}} \quad - 32\bar{I}_2 + (48 - j31.25)\bar{I}_3 &= 0
\end{aligned}$$

The value of the denominator determinant Δ is therefore equal to

$$\Delta = j2\begin{vmatrix} -1 & 1 & 0 \\ j5 & (4 - j5) & -4 \\ 0 & -32 & (48 - j31.25) \end{vmatrix} = j2\begin{vmatrix} -1 & 0 & 0 \\ j5 & 4 & -4 \\ 0 & -32 & (48 - j31.25) \end{vmatrix}$$

$$= -8(31.25 + j16)$$

$$\bar{I}_{\mathrm{T}} = \frac{\begin{vmatrix} 3 & j2 & 0 \\ 0 & (4 - j5) & -4 \\ 0 & -32 & (48 - j31.25) \end{vmatrix}}{-8(31.25 + j16)} = \frac{3[(4 - j5)(48 - j31.25) - 128]}{-8(31.25 + j16)}$$

$$= \frac{3(92.25 + j365)}{8(31.25 + j16)}$$

Hence

$$|\bar{I}_{\mathrm{T}}| = \frac{3|92.25 + j365|}{8|31.25 + j16|} = 4.021 \text{ A, } \textit{answer}.$$

(b) $\phi = \arctan\dfrac{3.0212}{2.6539} = 48.703^\circ$ approx., *answer*.

124. Note: To save space, we've generally rounded calculator values off to five decimal places.

First, $\omega = 2\pi f = 2\pi(28)10^3 = (1.75929)10^5$ rad/sec. Hence the reactances are

for 35 microhenry coil, $j\omega(35)10^{-6} = j6.15752$ ohms,

for 65 microhenry coil, $j\omega(65)10^{-6} = j11.43539$ ohms,

for 0.22 microfarad cap., $-j/\omega(22)10^{-8} = -j25.83687$ ohms,

total resistance $= 2 + 4 = 6$ ohms, a real number.

Let $\bar{Z}_T$ be the total impedance seen by the generator. Since Fig. 141 is a *series* circuit, the REAL PART of $\bar{Z}_T$ is the sum of all the real components and the IMAGINARY PART is the sum of all the imaginary components; hence, for the case of Fig. 141 we have

$$\bar{Z}_T = 6 - j8.24396$$

thus, by Ohm's law,

$$\bar{I} = \frac{20}{6 - j8.24396} = \frac{20(6 + j8.24396)}{103.96288} = (1.15426 + j1.58594) \text{ A, leading.}$$

Therefore, since the reactance of the 65 microhenry coil is $j11.43539$ ohms, and since the *voltage drop across the coil* is equal to the *coil reactance times the current* $\bar{I}$, we have that

$$\bar{V}_a = (1.15426 + j1.58594)(j11.43539)$$

thus

$$\bar{V}_a = (-18.13584 + j13.19941) \text{ volts approx., } \textit{answer}.$$

The above answer is in terms of "rectangular coordinates," expressed by a complex number lying in the second quadrant. To put the same answer in terms of polar coordinates, note that

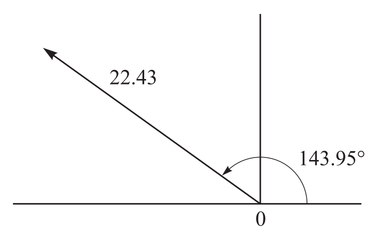
$$|\bar{V}_a| = 22.43063 \quad \text{and} \quad \phi = \arctan\left|\frac{13.19941}{18.13584}\right| = 36.05^\circ, \text{ approx.}$$

Since ϕ actually ends in the second quadrant, we have that

$$\phi = 180 - 36.05 = 143.95^\circ$$

Thus, in polar coordinates the *answer* is

$$\bar{V}_a = 22.43\underline{/143.95^\circ}$$



as shown to the right.

Thus an ac voltmeter* connected from point "a" to ground would read 22.43 volts, which, you'll note, is greater than the generator voltage of 20 volts. This is the result of the phenomenon of "series resonance," which we take up in section 8.6.

125. First,

$$\text{reactance of inductor} = j\omega L = j(10^6)(25)10^{-6} = j25 \text{ ohms}$$

and

$$\text{reactance of capacitor} = -j/\omega C = -j/(10^6)(5)10^{-8} = -j20 \text{ ohms}$$

* Be reminded that ac meters are normally calibrated to read magnitudes of rms voltages and currents.

Let us now use the method of loop currents, and let $\bar{I}_1$ and $\bar{I}_2$ denote the two loop currents required for Fig. 142. In doing this, let us designate that positive current flows in the clockwise sense around each loop (as in Fig. 135). Thus (see discussion with Fig. 135) the simultaneous vector equations for Fig. 142 are

$$\begin{aligned}(15+j25)\bar{I}_1 \quad -15\bar{I}_2 &= 30\\ -15\bar{I}_1 + (25-j20)\bar{I}_2 &= 0\end{aligned} \quad = \quad \begin{aligned}(3+j5)\bar{I}_1 \quad -3\bar{I}_2 &= 6\\ -3\bar{I}_1 + (5-j4)\bar{I}_2 &= 0\end{aligned}$$

hence

$$\Delta = \begin{vmatrix} (3+j5) & -3 \\ -3 & (5-j4) \end{vmatrix} = 13(2+j)$$

and hence

$$\bar{I}_2 = \frac{\begin{vmatrix} (3+j5) & 6 \\ -3 & 0 \end{vmatrix}}{13(2+j)} = \frac{18}{13(2+j)} = \frac{18(2-j)}{65} = 0.55385 - j0.27692 \text{ A}$$

thus

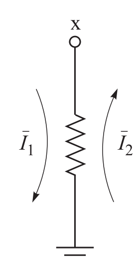
$$\bar{V}_y = 10\bar{I}_2 = (5.5385 - j2.7692) \text{ volts (rect. form), } \textit{answer}, \text{ or}$$

$$\bar{V}_y = 6.19221\underline{/-26.566^\circ} \text{ volts (polar form), } \textit{answer}.$$

126. We must find the TOTAL CURRENT flowing through the 15-ohm resistance. Since we already know the value of $\bar{I}_2$, this means we must now find the value of $\bar{I}_1$. Since $\Delta = 13(2+j)$ as before, we have that

$$\bar{I}_1 = \frac{\begin{vmatrix} 6 & -3 \\ 0 & (5-j4) \end{vmatrix}}{13(2+j)} = \frac{6(5-j4)}{13(2+j)} = \frac{6(5-j4)(2-j)}{65} = 0.55385 - j1.20000 \text{ amperes}$$

Since our equations have been written for the case where "positive" current flows in the "clockwise" sense, the situation for the 15-ohm resistance is as shown to the right. From inspection we see that the total resultant current $\bar{I}_T$ in the 15-ohm resistance is equal to



$$\bar{I}_T = \bar{I}_1 - \bar{I}_2 = -j0.92308 \text{ A approx.}$$

and thus, by Ohm's law,

$$\bar{V}_x = (15)\bar{I}_T = -j13.8462 \text{ volts, } \textit{answer},$$

or, in polar form

$$\bar{V}_x = 13.8462\underline{/-90^\circ}, \ \textit{answer}.$$

Note: In problems 125 and 126 we chose the "clockwise" direction around the loop to be the "positive" direction. It should be noted, however, that the same *rms values* of currents and voltages will be obtained regardless of which direction, cw or ccw, is chosen to be the positive direction. Of course, once the positive direction is designated for each loop, in a given problem, that designation must not be changed during the writing of the network equations.

127. The first step is to find the voltage *between terminals a and b* in Fig. 143; this is the voltage $\bar{V}'$ of the equivalent generator, and is found by applying Ohm's law to Fig. 143 as follows.

Note that the voltage between terminals a and b is the voltage drop *across capacitor C*, which is equal to the *current* $\bar{I}$ times the *reactance* $-jX_C$ of capacitor C; that is, we have $\bar{V}' = (\bar{I})(-jX_C)$. However, from inspection of Fig. 143, $\bar{I} = V/(R - jX_C)$, and thus we have

$$\bar{V}' = \frac{-jVX_C}{R - jX_C}$$

or, rationalizing,

$$\bar{V}' = \frac{-jVX_C(R + jX_C)}{R^2 + X_C^2} = \frac{VX_C(X_C - jR)}{R^2 + X_C^2}, \ \textit{answer}, \text{ in terms of } X_C.$$

Next, the impedance $\bar{Z}'$ of the equivalent generator is the impedance seen looking into terminals a, b in Fig. 143; thus, since R and $-jX_C$ are in parallel, we have, using eq. (209),

$$\bar{Z}' = \frac{-jRX_C}{R - jX_C} = \frac{-jRX_C(R + jX_C)}{R^2 + X_C^2} = \frac{RX_C(X_C - jR)}{R^2 + X_C^2}, \ \textit{answer}, \text{ in terms of } X_C.$$

Now setting $X_C = 1/\omega C$ in the above answers, you can verify that

$$\bar{V}' = \frac{V(1 - jR\omega C)}{1 + (R\omega C)^2} \quad \text{and} \quad \bar{Z}' = \frac{R(1 - jR\omega C)}{1 + (R\omega C)^2}, \ \textit{answers}.$$

Thevenin's theorem thus allows the replacement of a somewhat complicated network with a simple *series circuit* in the form of Fig. 144. This is useful when, given a complicated network, we wish to find the current that would flow in a number of different loads when connected to the original complicated network.

128. (a) First, $X_C = 2$ ohms and $X_L = 3$ ohms. Now, from left to right in Fig. 145, let $\bar{I}_1$ and $\bar{I}_2$ denote the two loop currents, with the positive sense to be the clockwise direction. Then, noting that $5\underline{/90^\circ} = j5$ (note 16 in Appendix), the two simultaneous equations for Fig. 145 are

$$\begin{aligned} 2(1 - j)\bar{I}_1 \qquad\quad - 2\bar{I}_2 &= 5(2 + j) \\ -2\bar{I}_1 + (4 + j3)\bar{I}_2 &= -j5 \end{aligned}$$

Since, from inspection of Fig. 145, $\bar{V}_x = 2\bar{I}_2$, let us solve for $\bar{I}_2$, as follows. First,

$$\Delta = 2\begin{vmatrix} (1 - j) & -2 \\ -1 & (4 + j3) \end{vmatrix} = 2(5 - j)$$

and hence

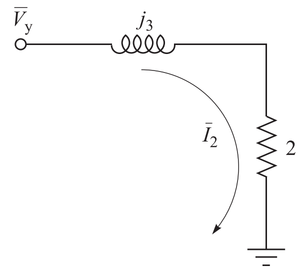
$$\bar{I}_2 = \frac{(2)(5)\begin{vmatrix} (1 - j) & (2 + j) \\ -1 & -j \end{vmatrix}}{2(5 - j)} = \frac{5}{5 - j}$$

thus

$$\bar{V}_x = 2\bar{I}_2 = \frac{10}{5 - j}\,\frac{10(5 + j)}{26} = (1.9231 + j0.3846) \text{ volts}, \ \textit{answer}.$$

(b) $|\bar{V}_x| = 1.9612$ volts approx., *answer*.

129. The value of $\bar{I}_2$ is known from problem 128; hence (see figure to the right) the value of $\bar{V}_y$ is equal to



$$\bar{V}_y = (2 + j3)\bar{I}_2 = \frac{5(2 + j3)}{5 - j}$$

thus

$$|\bar{V}_y| = \frac{5|2 + j3|}{|5 - j|} = 5\sqrt{\frac{13}{26}} = 3.5355 \text{ V}, \ \textit{answer}.$$

130. Note that we have a total of five node points (not counting the reference ground node) labeled a, b, c, x, and y, in Fig. 147. From inspection of the figure we have *three known node voltages*, thus

$$\bar{V}_a = 10\underline{/0^\circ} = 10 \text{ volts}, \ \bar{V}_b = 15\underline{/90^\circ} = j15 \text{ volts}, \ \bar{V}_c = 20\underline{/0^\circ} = 20 \text{ volts}$$

Next note that $X_L = \omega L = 4$ ohms and $X_C = 1/\omega C = 2$ ohms. Then, from inspection of the figure and in accordance with eq. (223), we have that

$$\bar{I}_1 = \frac{\bar{V}_a - \bar{V}_x}{j4}, \quad \bar{I}_2 = \frac{\bar{V}_b - \bar{V}_x}{2}, \quad \bar{I}_3 = \frac{\bar{V}_x - \bar{V}_y}{5}, \quad \bar{I}_4 = \frac{\bar{V}_y}{3}, \quad \bar{I}_5 = \frac{\bar{V}_c - \bar{V}_y}{-j2}$$

Next, at *node x* we have $\bar{I}_1 + \bar{I}_2 - \bar{I}_3 = 0$, and at *node y* $\bar{I}_3 - \bar{I}_4 + \bar{I}_5 = 0$, which, after substituting in the above values, gives the two simultaneous equations

$$\begin{aligned} -(5 + j14)\bar{V}_x \qquad\quad + j4\bar{V}_y &= 100 \\ j6\bar{V}_x + (15 - j16)\bar{V}_y &= 300 \end{aligned}$$

the solution of which, by determinants, gives the required *answers*

$$\bar{V}_x = (-0.5242 + j10.4296) \text{ volts}, \text{ and } \bar{V}_y = (11.2024 + j12.1589) \text{ volts}.$$

131. Ac voltmeters read magnitude of resultant rms voltage; thus

$$|\bar{V}_x - \bar{V}_y| = |-11.7266 - j1.7293| = 11.8534 \text{ V}, \ \textit{answer}.$$

132. First

reactance of 2 microfarad cap. $= -j5$ ohms,

reactance of 60 microhenry ind. $= j6$ ohms,

reactance of 160 microhenry ind. $= j16$ ohms,

reactance of 0.5 microfarad cap. $= -j20$ ohms,

$10\underline{/90^\circ} = j10$ volts, $\qquad 20\underline{/270^\circ} = -j20$ volts.

Then,

BRANCH 1: $\bar{V}_1 = 12\,\text{V}$, $\bar{Y}_1 = 1/8 = 0.1250$ mhos;

BRANCH 2: $\bar{V}_2 = 15\,\text{V}$, $\bar{Y}_2 = 1/-j5 = j0.2000$ mhos;

BRANCH 3: $\bar{V}_3 = j10\,\text{V}$, $\bar{Y}_3 = 1/j6 = -j0.1667$ mhos;

BRANCH 4: $\bar{V}_4 = 0\,\text{V}$, $\bar{Y}_4 = \dfrac{1}{3-j4} = \dfrac{3+j4}{25} = 0.1200 + j0.1600$ mhos;

BRANCH 5: $\bar{V}_5 = -j20\,\text{V}$, $\bar{Y}_5 = 1/5 = 0.2000$ mhos.

Substitution of the above values into eq. (225) gives

$$|\bar{V}_{\text{a}}| = |\bar{V}_{\text{o}}| = \left|\frac{3.1667 - j1.0000}{0.4450 + j0.1933}\right| = \frac{3.3208}{0.4852} = 6.844\,\text{V approx., } \textit{answer}.$$

133. Here, $V = 115$ volts, $R = 28$ ohms, $X = 45.2389$ ohms, thus $Z = \sqrt{R^2 + X^2} = 53.2030$ ohms approximately. Hence (using the value of I found in the solution to problem 116) we have

(a) $P_{\text{a}} = VI = (115)(2.1615) = 248.5725$ watts, *answer*.

(b) By eq. (232), $P_{\text{t}} = VI(\text{R}/\text{Z}) = 130.82$ watts, *answer*.

(c) By eq. (233), $P_{\text{x}} = \sqrt{P_{\text{a}}^2 - P_{\text{t}}^2} = 211.363$ watts, *answer*.

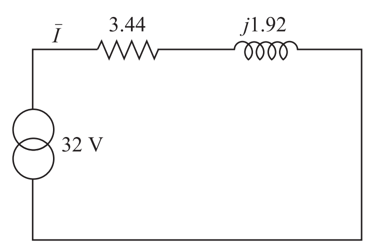
134. The EASY way is to note that, in Fig. 134, since R and L are in parallel (and since only R actually *consumes* energy) the true power output of the generator is equal to the power delivered to the 12-*ohm resistance*, which is (since $P = VI = V^2/R$)

$$P_{\text{t}} = (28)^2/12 = 65.333\ \text{W (watts) approx., } \textit{answer}.$$

135. Note, first, that the reactance of the inductor coil is $\omega L = X = 3$ ohms. Then, since the 2-ohm resistance is in series with the parallel combination of the 4-ohm resistance and the 3-ohm inductive reactance, the generator sees an impedance equal to

$$\bar{Z} = 2 + \frac{j12}{4+j3} = 2 + \frac{j12(4-j3)}{(4+j3)(4-j3)} = 2 + 1.44 + j1.92 = 3.44 + j1.92 \text{ ohms}$$

Thus the generator sees an impedance $\bar{Z} = R + jX = 3.44 + j1.92$ ohms, as shown in the following figure.



Hence (see paragraph following eq. (234)) we now have

$$|\bar{Z}| = Z = \sqrt{(3.44)^2 + (1.92)^2} = 3.9395 \text{ ohms}$$

thus

$$I = 32/3.9395 = 8.1229 \text{ A}$$

Hence, in this problem, $V = 32, I = 8.1229, R = 3.44$, and $Z = 3.9395$; thus

$$P_t = VI(R/Z) = 226.98 \text{ watts, } \textit{answer}.$$

136. Let us denote the two loop currents (from left to right in Fig. 158) by $\bar{I}_1$ and $\bar{I}_2$, and let us draw both current arrows in the clockwise sense. Then the two network equations are

$$6\bar{I}_1 \qquad - 4\bar{I}_2 = 32$$

$$-4\bar{I}_1 + (4 + j3)\bar{I}_2 = 0$$

We must now solve the above two simultaneous equations for the values of $\bar{I}_1$ and $\bar{I}_2$. Using the method of determinants, we first have that

$$\Delta = \begin{vmatrix} 6 & -4 \\ -4 & (4 + j3) \end{vmatrix} = 2(4 + j9)$$

and next, upon applying the standard procedure of solution, you should find that

$$\bar{I}_1 = \frac{32(4 + j3)}{\Delta} = (7.092784 - j3.958763) \text{ A approx., and}$$

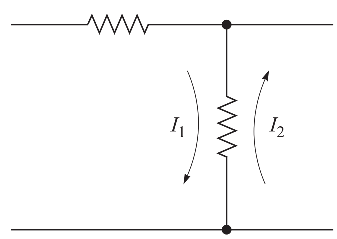
$$\bar{I}_2 = \frac{128}{\Delta} = (2.639175 - j5.938144) \text{ A approx.}$$

The solutions to parts (a), (b), and (c) are now as follows.

(a) From the above, $|\bar{I}_1| = I_1 = 8.122770$ A approx., and thus the power, P_2, to the 2-*ohm resistance* is

$$P_2 = I_1^2 R = 131.959 \text{ W, } \textit{answer}.$$

(b) Now, in regard to the 4-*ohm resistance*, since both currents were drawn in the clockwise sense, we see, from the diagram to the right, that $\bar{I}_4 = \bar{I}_1 - \bar{I}_2 = 4.453609 + j1.979381$, and thus $|\bar{I}_4| = |\bar{I}_1 - \bar{I}_2| = I = 4.873662$ A approx.; thus the power to the 4-*ohm resistance* is



$$P_4 = I^2 R = 95.010 \text{ W approx., } \textit{answer}.$$

(c) $P_T = P_2 + P_4 = 131.959 + 95.010 = 226.97$ watts approx., which *does check* with the answer obtained in problem 135.

137. From inspection of Fig. 141, and also from the solution to problem 124, we have $V = 20$ volts, $R = 6$ ohms, $|\bar{Z}| = 10.19622$ ohms, and $|\bar{I}| = 1.961510$ A. Hence

(a) $P_a = VI = 39.230$ W, *answer*.

(b) By eq. (232), $P_t = VI(R/Z) = 23.085$ W, *answer*.

138. From the solution to problem 126, $\bar{I}_1 = 0.55385 - j1.20000$ A. Hence the component of generator current IN PHASE with the generator voltage of 30 volts is 0.55385 amperes, and thus the true power output of the generator is (see discussion with Fig. 150)

$$P_t = (30)(0.55385) = 16.6155 \text{ W approx., } \textit{answer}.$$

139. (a) From the solution to problem 126, the magnitude of current in the 15-ohm resistance is equal to 0.92308 A, and thus the power to the 15-ohm resistance is

$$I^2R = 12.7812 \text{ W, } \textit{answer}.$$

(b) From the solution to problem 125, the magnitude of current in the 10-ohm resistance equals 0.61922 A, thus the power to the 10-ohm resistance is

$$I^2R = 3.8344 \text{ W, } \textit{answer}.$$

Since power is a scalar quantity, not a vector quantity, the *total power* is the *sum* of the above two answers, which *does check* with the answer found in problem 138.

140. Disregard the value of R because, in a series circuit, the resonant frequency is independent of the value of R. Then, from sections 7.5 and 7.6, we have that $L = 4 \times 10^{-6}$ henrys and $C = 25 \times 10^{-10}$ farads; thus, by eqs. (238) and (239), we have

$$\omega_0 = \frac{1}{\sqrt{10^{-14}}} = \frac{1}{10^{-7}} = 10^7 \text{ rad/sec, } \textit{answer}, \text{ or}$$

$$f_0 = \frac{10^7}{2\pi} = 1.5916 \times 10^6 \text{ Hz, } \textit{answer}, \text{ or}$$

$$= 1.5916 \text{ MHz (megahertz).}$$

141. From eq. (239), $C = \dfrac{1}{4\pi^2 f_0^2 L}$, upon which, after substituting in $f_0 = 5 \times 10^5$ Hz and $L = 4 \times 10^{-4}$ henry, you should find that

$$C = 10^{-8}/4\pi^2 = 0.02533 \times 10^{-8} \text{ F} = 253.3 \text{ pF, } \textit{answer}.$$

142. (a) At resonance, $I_0 = V/R = 20/5 = 4$ amperes; hence

$$P = I_0^2 R = (16)(5) = 80 \text{ watts, } \textit{answer}.$$

(b) By eq. (238), $\omega_0 = (2)10^7$ rad/sec, at which frequency $X_C = 1/\omega_0 C = 20$ ohms; thus, in magnitude,

$$V_C = IX_C = 4(20) = 80 \text{ volts, } \textit{answer}.$$

(c) We now have $X_C = 1/\omega C = 40$ ohms, $X_L = \omega L = 10$ ohms, $R = 5$ ohms. Thus, by eq. (236), $|\bar{I}| = I = 20/\sqrt{925}$; hence

$$V_C = IX_C = 800/\sqrt{925} = 26.3038 \text{ volts, } \textit{answer}.$$

(d) By eq. (237),

$$\phi = \arctan(30/5) = 80.5377^\circ, \text{ leading, } \textit{answer}.$$

(e) By eq. (228), and also using the value of I from (c) above, we have

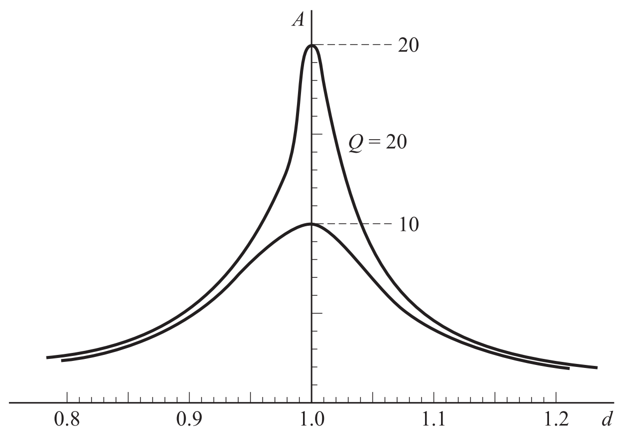
$$P = (20)(20/\sqrt{925}) \cos 80.5377^\circ = 2.1620 \text{ watts approx., } \textit{answer}.$$

OR, using the equation,

$$P = I^2R = (400/925)(5) = 2.1622 \text{ watts approx., } \textit{answer}.$$

143. The values in the following table were calculated using eq. (248), with freehand plots of the data shown to the right. As the curves show, the higher Q circuit is clearly superior to the lower Q circuit as far as MAXIMUM OUTPUT VOLTAGE and SELECTIVITY are concerned. Note, however, that the higher Q circuit has a relatively NARROW BANDWIDTH as compared with the low Q circuit, and this may or may not be an advantage, depending upon the *rate of transmission of information* through the circuit. (It is a fundamental fact of nature that, the more information that is to be transmitted through a system per unit time, the greater must be the bandwidth of the system.)

d	A, for $Q = 10$	A, for $Q = 20$
0.80	2.71	2.76
0.85	3.45	3.56
0.90	4.76	5.12
0.93	6.10	7.00
0.95	7.35	9.22
0.97	8.80	13.08
0.99	9.90	18.74
1.00	10.00	20.00
1.01	9.71	18.40
1.03	8.36	12.54
1.05	6.82	8.68
1.07	5.55	6.47
1.10	4.22	4.61
1.15	2.92	3.05
1.20	2.19	2.25



144. Since the generator is operating at the resonant frequency defined by eq. (252), it would see a pure resistance of R_0 ohms, given by eq. (254). Thus, upon substituting the given values into eq. (254) (see conversion formulas following eq. (184) in section 7.6), we find that

$$R_0 = \frac{10^{-4}}{(50)(10^{-10})} = 20{,}000 \text{ ohms}, \ \textit{answer}.$$

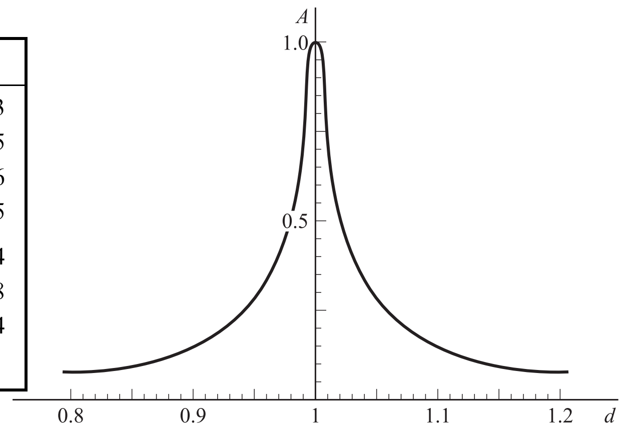
Thus the LOW RESISTANCE load of 50 ohms appears to the generator as a HIGH RESISTANCE load of 20,000 ohms. This can be of great practical advantage if, for instance, the generator is a device having a relatively high internal resistance.

145. For $Q = 20$, eq. (269) becomes

$$A = \frac{\sqrt{1 + 400d^2(d^2 - 0.9975)^2}}{d^2 + 400(d^2 - 1)^2}$$

thus

d	A	d	A
0.80	0.11	1.01	0.93
0.85	0.15	1.03	0.65
0.90	0.23	1.05	0.46
0.93	0.33	1.07	0.35
0.95	0.44	1.10	0.24
0.97	0.64	1.15	0.18
0.99	0.93	1.20	0.14
1.00	1.00		



The parallel circuit of Fig. 167 is widely used because it presents a very HIGH IMPEDANCE at, and near, its resonant frequency (as the figure illustrates). Thus, when used as a "tuned load" in an amplifier stage, the gain of the stage is great at, and in the immediate neighborhood of, the resonant frequency, but low at undesired frequencies away from the resonant frequency.

146. (a) Putting $L = (250)10^{-6}$ and $C = (4)10^{-9}$ into eq. (259), the approximate value of ω_0 is

$$\omega_0 = 10^6 \text{ radians/second}, \textit{ answer}, \text{ or}$$

$$f_0 = \omega_0/2\pi = 159.155 \text{ kHz (kilohertz)}.$$

(b) Putting the given values of L, C, and $>R$ into eq. (252) gives $\omega_0 = 10^6\sqrt{1-0.0064}$, which for almost all practical purposes can be taken to be $\omega_0 = 10^6$ rad/sec, the same value as found in part (a). Thus the *answer* to the question is "yes."

147. (a) The generator sees a pure resistance at resonance, given by eq. (254); thus

$$R_0 = \frac{(250)10^{-6}}{20(4)10^{-9}} = 3125 \text{ ohms}, \textit{ answer}.$$

(b) By Ohms's law, at resonance

$$\bar{I}_g = I_g\underline{/0^\circ} = 90/3125 = 0.0288 \text{ amperes}, \textit{ answer}.$$

(c) $P = V^2/R_0 = 8100/3125 = 2.592$ watts, *answer*; or, if you wish,

$$P = I_g^2 R_0 = (0.0288)^2(3125) = 2.592 \text{ watts, same } \textit{answer}.$$

(d) $X_C = 1/\omega_0 C = 250$ ohms. Hence, by Ohm's law,

$$\bar{I}_C = \frac{90}{-j250} = j0.36 \text{ amperes, leading } V \text{ by } 90^\circ, \textit{ answer}.$$

(e) $X_L = \omega_0 L = 250$ ohms. Hence, by Ohm's law,

$$\bar{I}_L = \frac{90}{20 + j250}$$

and thus, upon rationalizing,

$$\bar{I}_L = (0.02862 - j0.35771) \text{ amperes, lagging, } \textit{answer}.$$

(f) By eq. (256),

$$Q = 250/20 = 12.5, \textit{ answer}.$$

(g) $P = |\bar{I}_L|^2 R = (0.12878)(20) = 2.576$ watts approx., *answer*.

The answer here differs, by a very small amount, from the answer found in part (c), because the true value of ω_0 is SLIGHTLY LESS than the value of ω_0 found by using eq. (259) (as was brought out in the solution to problem 146). Hence the value of $\omega_0 L$ used in calculating $\bar{I}_L$ in part (e) is SLIGHTLY MORE than the true value of $\omega_0 L$.

148. (a) Since the values of C, L, and R have not been changed, it follows that the values of both ω_0 and Q, defined by eqs. (259) and (256), will have the same values as found in problems 146 and 147. Likewise, R_0 will still have the same value, $R_0 = L/RC = 3125$ ohms. Also note that, for this problem, by eq. (262), $d = 1.05$. Thus, upon substituting $R_0 = 3125, Q = 12.5$, and $d = 1.05$ into eq. (268), we have that

$$\bar{Z}_\text{p} = 3125\left[\frac{1 - j1.429313}{2.744102}\right] = 3125(0.364418 - j0.520867) \text{ ohms}$$

hence, upon applying Ohm's law, then rationalizing, we have

$$\bar{I}_\text{g} = \frac{V}{\bar{Z}_\text{p}} = \frac{90}{3125(0.364418 - j0.520867)} = (0.025972 + j0.037122) \text{ amperes, } \textit{answer}.$$

(b) $\phi = \arctan\dfrac{0.037122}{0.025972} = 55.022^\circ$, *answer*.

(c) The *first* (and easiest) way is to multiply the generator voltage by the "in phase" component of the generator current (section 5.7); thus

$$P = (90)(0.025972) = 2.338 \text{ watts approx., } \textit{answer}.$$

The *second* way is to make use of eq. (117) in Chap. 5; that is, $P = VI \cos\phi$, where V and I_g are rms magnitudes of voltage and current. From part (a) you can verify that $I_\text{g} = 0.045305$ amperes; thus

$$P = (90)(0.045\,305) \cos 55.022^\circ = 2.338 \text{ watts, same } \textit{answer}.$$

149. We must first determine whether to use eq. (270) or eq. (274). To do this, make use of eq. (271) or (275)); thus

$$R_\text{in} = L/RC = (28.5)10^{-6}/(135)(36)10^{-10} = 58.848 \text{ ohms}$$

This shows that R_in is *less than* R, and therefore we must use eq. (274), which gives the value

$$\omega_0 = (2.348)10^6 \text{ Hz} = 2.348 \text{ megahertz, } \textit{answer}.$$

150. Here $R = 16$ ohms and $R_{\text{in}} = 75$ ohms, so that we must use the L-section of Fig. 170. Hence, substituting the given values into eqs. (272) and (273), we find that

$$L = (16)(10^{-6})\sqrt{3.6875} = (30.725)10^{-6}\,\text{H} = 30.725\,\mu\text{H}, \ \textit{answer}.$$

$$C = \frac{10^{-6}}{75}\sqrt{3.6875} = (0.0256)10^{-6}\,\text{F} = 0.0256\,\mu\text{F}, \ \textit{answer}.$$

151. Here $R = 125$ ohms and $R_{\text{in}} = 85$ ohms, so that we must use the reverse L section of Fig. 171. Also note that, here, $\omega_0 = 2\pi f_0 = (2.26195)10^6$ rad/sec, and upon substituting all these values into eqs. (276) and (277) you should find that

$$L = (25.778)10^{-6}\,\text{H} = 25.778\,\mu\text{H}, \ \textit{answer}.$$

$$C = (2.426)10^{-9}\,\text{F} = 0.002426\,\mu\text{F}, \ \textit{answer}, \text{ or}$$

$$= 2426\,\text{pF (picofarads)}.$$

152. Let $\bar{Z}_{\text{in}}$ denote the impedance looking to the right, into terminals $(1, 1)$, in Fig. 171. Since the inductor L is in series with the parallel combination of C and R, we have,

$$\bar{Z}_{\text{in}} = jX_L + \frac{-jRX_C}{R - jX_C}$$

which, after rationalizing the fraction, then separating real and imaginary parts, becomes

$$\bar{Z}_{\text{in}} = \frac{RX_C^2}{R^2 + X_C^2} + j\left(X_L - \frac{R^2 X_C}{R^2 + X_C^2}\right)$$

Looking to the right, into terminals $(1, 1)$, we are to see, at the resonant frequency ω_0, a pure resistance; this means that, at the resonant frequency, the IMAGINARY PART of the above equation must have the value ZERO. Thus, equating the above imaginary part equal to zero gives the equation

$$R^2 X_L + X_L X_C^2 - R^2 X_C = 0$$

which is true *at the resonant frequency* ω_0, at which $X_L = \omega_0 L$, and $X_C = 1/\omega_0 C$. Now, making these substitutions into the last equation, then multiplying by ω_0 and solving for ω_0, should give eq. (274), *answer*.

153. Set $X_L = \omega_0 L$ and $X_C = 1/\omega_0 C$ in the equation for $\bar{Z}_{\text{in}}$ found in the solution to problem (152). Doing this, the imaginary part of the equation vanishes, leaving $\bar{Z}_{\text{in}} = R_{\text{in}}$; thus

$$R_{\text{in}} = \frac{R/\omega_0^2 C^2}{R^2 + 1/\omega_0^2 C^2} = \frac{R}{1 + \omega_0^2 R^2 C^2}$$

Now substitute into the last equation the value of ω_0 given by eq. (274); doing this gives the value L/RC, which is eq. (275), *answer*.

154. By eq. (275), $C = L/R_{\text{in}}R$, and substituting this value of C into eq. (274) gives, after a bit of algebra, eq. (276). Then, also by eq. (275), $L = R_{\text{in}}RC$, and substituting this value of L into eq. (274) gives eq. (277), *answers*.

155. By inspection,

$$\bar{Z}_{1O} = 3 + (9)(9)/18 = 7.500 \text{ ohms},$$
$$\bar{Z}_{1S} = 3 + (9)(4)/13 = 5.769 \text{ ohms},$$
$$\bar{Z}_{2O} = (5)(13)/18 = 3.611 \text{ ohms};$$

thus

by eq. (282), $\bar{Z}_3 = \sqrt{6.251} = 2.500$ ohms, *answer*,

by eq. (283), $\bar{Z}_2 = 3.611 - 2.500 = 1.111$ ohms, *answer*,

by eq. (284), $\bar{Z}_1 = 7.500 - 2.500 = 5.000$ ohms, *answer*.

In regard to the question asked, the *answer* is ''yes,'' because the value of a pure resistance is theoretically the same at all frequencies.

156. The reactances are $-j/\omega C = -j5$ ohms and, $j\omega L = j15$ ohms. Then, by inspection of Fig. 178, we have

$$\bar{Z}_{1O} = \frac{10(5-j5)}{15-j5} = \frac{10(1-j)}{3-j} = \frac{10(1-j)(3+j)}{(3-j)(3+j)} = (4-j2) \text{ ohms}$$

Next, to find the value of $\bar{Z}_{1S}$, first note that ''5 ohms in parallel with $j15$ ohms'' is

$$\frac{5(j15)}{5+j15} = \frac{j15}{1+j3} = (4.5+j1.5) \text{ ohms},$$

and with this in mind, inspection of Fig. 178 then shows that

$$\bar{Z}_{1S} = \frac{10(4.5-j3.5)}{14.5-j3.5} = \frac{775.0-j350.0}{222.50} = (3.483-j1.573) \text{ ohms}$$

Next, looking into terminals (2, 2) with (1, 1) open-circuited, we see that

$$\bar{Z}_{2O} = j15 + \frac{5(10-j5)}{15-j5} = (3.5+j14.5) \text{ ohms}$$

The final step is to substitute the values of $\bar{Z}_{1O}, \bar{Z}_{1S}, \bar{Z}_{2O}$, just found, into eqs. (282) through (284). Let us begin with eq. (282); thus

$$\bar{Z}_3 = \sqrt{(3.5+j14.5)(0.517-j0.427)} = \sqrt{8.00+j6.00}$$

To find the indicated square root, let us write the complex number $8.00 + j6.00$ in the exponential form $A\epsilon^{j\theta}$, as follows.

First, $(8.00 + j6.00) = 10\epsilon^{j\theta}$. Then, since $(8 + j6)$ lies in the first quadrant of the complex plane, we have $\theta = \arctan(6/8) = 36.869\,90^\circ$, and hence

$$(8.00+j6.00) = 10\epsilon^{j36.869\,90^\circ} *$$

Now, substituting this value into the above value of $\bar{Z}_3$, and remembering that

$$\sqrt{A\epsilon^{j\theta}} = \sqrt{A}\epsilon^{j\theta/2} = \sqrt{A}\left(\cos\frac{\theta}{2} + j\sin\frac{\theta}{2}\right)$$

* In regard to the use of degrees here, see last footnote in section 6.5.

(Euler's formula), gives

$$\bar{Z}_3 = \sqrt{10}\,(\cos 18.4350^\circ + j\,\sin 18.4350^\circ) = 3.000 + j1.000 \text{ ohms},$$

then, by eqs. (283) and (284),

$$\bar{Z}_2 = 0.5 + j13.5 \text{ ohms, and } \bar{Z}_1 = 1.0 - j3.0 \text{ ohms.}$$

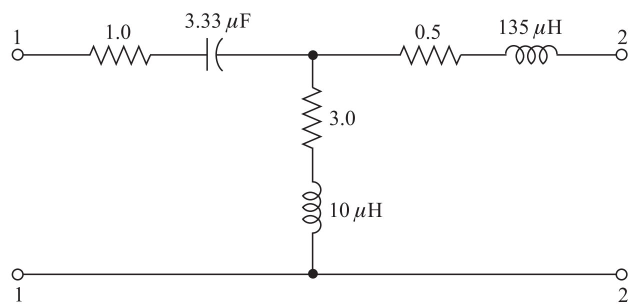
Next, to find the actual values of inductance and capacitance required, we make use of the reactance formulas $X_L = \omega L$ and $X_C = 1/\omega C$, that is, $L = X_L/\omega$ and $C = 1/\omega X_C$. Thus, using $\omega = 10^5$ and the reactance values found above, we have that

$$\text{for } \bar{Z}_3, \quad X_L = 1.0 \text{ ohm, thus } L = 1/10^5 = 10^{-5}\text{ H (henrys)} = 10\,\mu\text{H},$$

$$\text{for } \bar{Z}_2, \quad X_L = 13.5 \text{ ohms, thus } L = 13.5(10^{-5})\text{ H} = 135\,\mu\text{H},$$

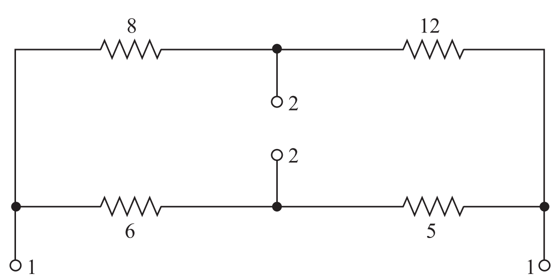
$$\text{for } \bar{Z}_1, \quad X_C = 3.0 \text{ ohms, thus } C = 1/(10^5)(3)\text{ F (farads)} = 3.33\,\mu\text{F}.$$

Thus the complete equivalent T for Fig. 178 is as shown below, with resistance values in ohms.



Answer **to problem 156.**

157. It may be less confusing if we redraw the network in the form shown in the figure below.



From the figure we see that

$$Z_{1\mathrm{O}}{}^{*} = 11 \text{ ohms in parallel with 20 ohms} = 220/31 = 7.097 \text{ ohms},$$

$$Z_{1\mathrm{S}} = 48/14 + 60/17 = 6.958 \text{ ohms},$$

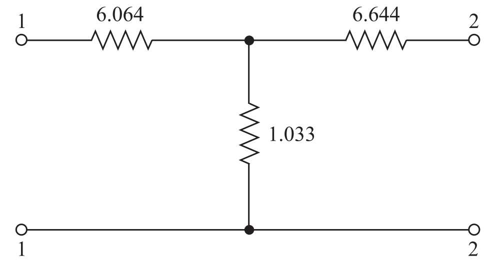
$$Z_{2\mathrm{O}} = 14 \text{ ohms in parallel with 17 ohms} = (14)(17)/31 = 7.677 \text{ ohms.}$$

* The "overscore" notation, $\bar{Z}$, indicates that Z is, or may be, a complex number. Since, in this problem, the Zs can only represent real numbers, we can, if we wish, write plain Z instead of $\bar{Z}$.

We now put the above values into eqs. (282), (283), and (284) to get the required answers; thus

$$Z_3 = \sqrt{7.677(7.097 - 6.958)}$$
$$= 1.033 \text{ ohms,}$$
$$Z_2 = 7.677 - 1.033 = 6.644 \text{ ohms,}$$
$$Z_1 = 7.097 - 1.033 = 6.064 \text{ ohms.}$$

The equivalent network is shown to the right.



158. Here we must use eqs. (285) through (289), and in using these equations the required *divisions* will generally be easier to do if the complex numbers are expressed in either the exponential or the polar form. Thus, if $A\underline{/p}$ and $B\underline{/q}$ are two complex numbers, then (eq. (165) in Chap. 6)

$$\frac{A\underline{/p}}{B\underline{/q}} = \frac{A}{B}\underline{/p-q}$$

and for this reason we've elected to write some of the following answers in both the polar and rectangular forms, with final answers in rectangular form.

First, by eq. (288),

$$\bar{Z}' = (20 - j12)(8 + j18) = 8(5 - j3)(4 + j9)$$

thus,

$$\bar{Z}' = 8(47 + j33) = 459.426\underline{/35.074^\circ} \text{ ohms}$$

Next, by eq. (289),

$$\bar{Z}'' = \sqrt{8(5 - j3)(5 - j3)} = \sqrt{8}(5 - j3) = 16.492\underline{/-30.964^\circ} \text{ ohms}$$

Then,

$$\bar{Z}_{2\text{O}} - \bar{Z}'' = 5.858 - j3.515 = 6.832\underline{/-30.965^\circ} \text{ ohms}$$

and

$$\bar{Z}_{1\text{O}} - \bar{Z}'' = 3.858 + j20.485 = 20.845\underline{/79.334^\circ} \text{ ohms}$$

Now substitute the above values, in polar form, into eqs. (285) through (287); then, by use of the polar formula $C\underline{/\phi} = C(\cos\phi + j\sin\phi)$, express the final answers in rectangular form; thus

by eq. (285), $\bar{Z}_\text{A} = 67.246\underline{/66.039^\circ} = 27.310 + j61.451$ ohms, *answer*.

by eq. (286), $\bar{Z}_\text{B} = 27.858\underline{/66.038^\circ} = 11.314 + j25.457$ ohms, *answer*.

by eq. (287), $\bar{Z}_\text{C} = 22.040\underline{/-44.260^\circ} = 15.785 - j15.382$ ohms, *answer*.

159. In eq. (298), combine the two quantities on the left-hand side over the common denominator $\bar{Z}_2 + \bar{Z}_3$, then multiply both sides by $\bar{Z}_2 + \bar{Z}_3$, then replace $\bar{Z}_2 + \bar{Z}_3$ with the right-hand side of eq. (299) to get

$$\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3 = \frac{\bar{Z}_\text{A}\bar{Z}_\text{B}\bar{Z}_\text{C}}{\bar{Z}_\text{A} + \bar{Z}_\text{B} + \bar{Z}_\text{C}} \tag{A}$$

or, if we wish,

$$\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3 = \bar{Z}_\mathrm{A}\left(\frac{\bar{Z}_\mathrm{B}\bar{Z}_\mathrm{C}}{\bar{Z}_\mathrm{A} + \bar{Z}_\mathrm{B} + \bar{Z}_\mathrm{C}}\right)$$

hence, by eq. (303),

$$\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3 = \bar{Z}_\mathrm{A}\bar{Z}_2$$

which, upon solving for $\bar{Z}_\mathrm{A}$, proves that eq. (304) *is correct*.

Next, upon making use of eq. (A) above and eq. (301), we have that

$$\bar{Z}_1\bar{Z}_2 + \bar{Z}_1\bar{Z}_3 + \bar{Z}_2\bar{Z}_3 = \bar{Z}_\mathrm{B}\left(\frac{\bar{Z}_\mathrm{A}\bar{Z}_\mathrm{C}}{\bar{Z}_\mathrm{A} + \bar{Z}_\mathrm{B} + \bar{Z}_\mathrm{C}}\right) = \bar{Z}_\mathrm{B}\bar{Z}_3$$

which, upon solving for $\bar{Z}_\mathrm{B}$, proves that eq. (305) *is correct*.

Next, again making use of eq. (A) and also eq. (302), you can verify that eq. (306) is *also correct*.

160. First, $\omega = 2\pi f = (3.1416)10^6$ rad/sec. Next, from Fig. 180 we have $\bar{Z}_\mathrm{A} = 30$ ohms, $\bar{Z}_\mathrm{B} = j\omega L = j62.832$ ohms, $\bar{Z}_\mathrm{C} = 20$ ohms. Putting these values in equations (302), (303), and (301), we have

$$\bar{Z}_1 = \frac{j1884.96}{50 + j62.832} = (18.368 + j14.617) \text{ ohms}$$

$$\bar{Z}_2 = \frac{j1256.64}{50 + j62.832} = (12.246 + j9.745) \text{ ohms}$$

$$\bar{Z}_3 = \frac{600}{50 + j62.832} = (4.653 - j5.847) \text{ ohms}$$

where the REAL PART of each answer represents *resistance* in ohms and the IMAGINARY PART represents *reactance* in ohms. Thus, since $X_L = \omega L$ and $X_C = 1/\omega C$, we have that $L = X_L/\omega$ and $C = 1/\omega X_C$, and therefore, using the known value of ω and the above values of $\bar{Z}_\mathrm{A}, \bar{Z}_\mathrm{B}$, and $\bar{Z}_\mathrm{C}$, we find that

$\bar{Z}_1$ consists of a resistance of 18.368 ohms in series with an inductor coil having 4.653 μH of inductance, *answer*.

$\bar{Z}_2$ consists of a resistance of 12.246 ohms in series with an inductor coil having 3.102 μH of inductance, *answer*.

$\bar{Z}_3$ consists of a resistance of 4.653 ohms in series with a capacitor having $(5.444)10^{-8}$ F = 0.05444 μF of capacitance, *answer*.

161. In Fig. 181, $\bar{Z}_1 = -j12, \bar{Z}_2 = -j9$, and $\bar{Z}_3 = j6$. Putting these values into eqs. (304), (305), and (306), you should find that

$$\bar{Z}_\mathrm{A} = \frac{18}{-j9} = j2 \text{ ohms, } \textit{answer},$$

$$\bar{Z}_\mathrm{B} = \frac{18}{j6} = -j3 \text{ ohms, } \textit{answer},$$

$$\bar{Z}_\mathrm{C} = \frac{18}{-j12} = j1.5 \text{ ohms, } \textit{answer}.$$

162. To produce maximum power in the load, the generator must see a resistance of 36 ohms looking into the input terminals of the T network. Since we're to use a balanced T network, the magnitudes of the reactances will all have the same value. By eq. (312),

$$X = \sqrt{(36)(115)} = 64.343 \text{ ohms.}$$

(a) Using the formulas $X_L = 2\pi fL$ and $X_C = 1/2\pi fC$, we find that

$$L = \frac{64.343}{(6.2832)(1.75)10^5} = (5.852)10^{-5} \text{ henry} = 58.52 \text{ microhenrys, } \textit{answer}.$$

$$C = \frac{1}{(6.2832)(1.75)10^5(64.343)} = (1.413)10^{-8} \text{ farads}$$

$$= 0.01413 \text{ microfarads, } \textit{answer}.$$

(b) $I_{gen} = 90/(R_g + R_L) = 90/72 = 1.25$ amperes, *answer*.

(c) $V_{in} = IR_{in} = (1.25)(36) = 45$ volts, *answer*.

(d) The power input to the T network is $P_{in} = V_{in}I = (45)(1.25) = 56.25$ watts. Since there is assumed to be no energy loss in the T network itself, the output power to the 115-ohm load is also 56.25 watts. Hence we have, where I_L =load current,

$$P_{out} = I_L^2 R$$

thus

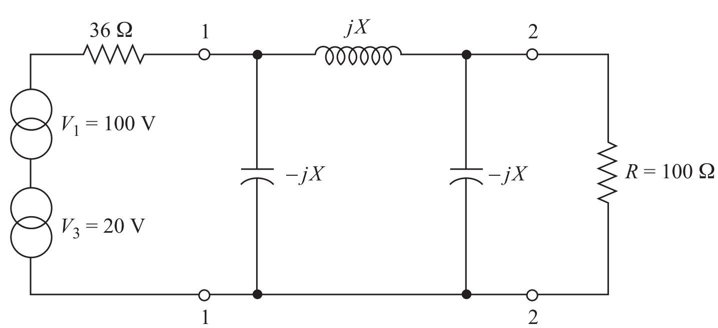
$$I_L = \sqrt{56.25/115} = 0.699 \text{ amperes, } \textit{answer}.$$

Then, since $P_{out} = V_L I_L$, we have

$$V_L = 56.25/0.699 = 80.47 \text{ volts, } \textit{answer}.$$

163. Here we're assuming a linear system, so that the principle of superposition applies; that is, we can *separately* calculate the output of the 100-volt fundamental generator and the output of the 20-volt third-harmonic generator, just as if each acted separately.

This is illustrated in the figure below, where V_1 and V_3 are the separate output voltages of the fundamental and third-harmonic generators.



(a) Let us first deal with the fundamental frequency, as follows. In order for V_1 to produce maximum power in the 100-ohm load, the V_1 generator must see a resistance of 36 ohms looking to the right into terminals (1, 1). Hence, by eq.

(313), each element in the balanced pi network must have a magnitude of reactance equal to

$$X = \sqrt{3600} = 60 \text{ ohms}$$

Now, using the formulas $X_L = \omega L$ and $X_C = 1/\omega C$, we calculate the necessary values of L and C for a frequency of 300 kiloradians/sec; thus

$$\begin{aligned} L &= \frac{60}{(3)10^5} = (20)10^{-5} \text{ henry} \\ &= 200\,\mu\text{H}, \ \textit{answer}. \\ C &= \frac{1}{(60)(3)10^5} = (5.555)10^{-8}\,\text{F} \\ &= 0.0555\,\mu\text{F}, \ \textit{answer}. \end{aligned}$$

(b) At 300 kilorad/sec we see a pure resistance of 36 ohms looking to the right into terminals (1, 1) in the figure. The 100-volt, 300 kilorad/sec generator thus sees a total resistance of $36 + 36 = 72$ ohms, and therefore delivers a current of $I_1 = 100/72 = 1.3889$ amp, thus generating a total power of $V_1 I_1 = 138.89$ watts. Since we have matched conditions for V_1, *half* this power of 138.89 watts goes to the useful 100-ohm load. Thus, if I_f denotes the load current at the fundamental frequency, then, since $P_f = I_f^2 R$, we have

$$I_f = \sqrt{\frac{P_f}{R}} = \sqrt{\frac{69.45}{100}} = 0.8334 \text{ amperes}$$

and thus the voltage across the 100-ohm load *at the fundamental frequency* is

$$V_f = I_f R = (0.8334)100 = 83.34 \text{ volts}, \ \textit{answer}.$$

(c) At *900 kiloradians per second* the reactances are $X_L = 180$ ohms and $X_C = 20$ ohms. We therefore cannot use the procedure of part (b), because we have neither a "balanced" pi network nor matched conditions at this frequency. We must therefore resort to the method of loop currents, as follows.

Returning to the figure, disregard the 300 kilorad/sec generator and draw three loop currents, I_1, I_2, I_3, from left to right in the cw sense. Doing this, the three loop-voltage equations for the network at 900 kilorad/sec are (after simplifying each equation as much as possible)

$$\begin{aligned} (9 - j5)I_1 + j5I_2 + \quad 0I_3 &= 5 \\ jI_1 + j7I_2 \qquad jI_3 &= 0 \\ 0I_1 + \ jI_2 + (5 - j)I_3 &= 0 \end{aligned}$$

thus

$$I_3 = \frac{\begin{vmatrix} (9 - j5) & j5 & 5 \\ j & j7 & 0 \\ 0 & j & 0 \end{vmatrix}}{\begin{vmatrix} (9 - j5) & j5 & 0 \\ j & j7 & j \\ 0 & j & (5 - j) \end{vmatrix}} = \frac{-5}{272 + j270}$$

the magnitude of which is

$$|I_3| = \frac{5}{\sqrt{(272)^2 + (270)^2}} = 0.013\,05 \text{ amperes}$$

and this current, flowing through the 100-ohm load, produces 1.305 volts at 900 kilorad/sec, *answer*.

The above problem illustrates two benefits that can be realized by inserting a pi network between generator and load. First, the network enables us to get increased power to the load by providing the proper impedance match to the generator. Second, it can greatly reduce the percentage of unwanted harmonic energy in the load.

164. (a) By eq. (314), $P = 2$; thus, applying the calculator to eq. (315), we have

$$10 \log 2 = 3.01 \text{ decibels, } \textit{answer}.$$

(b) Here $P = 0.26/1.65 = 0.157\,58$ approx., and thus, applying the calculator to eq. (315), we have

$$10 \log(0.157\,58) = -8.025 \text{ decibels, } \textit{answer} \text{ (a } \textit{loss} \text{ of 8.025 dB).}$$

(c) Substituting the given values into eqs. (314) and (315), we have $1.8 = \log(P_{\text{out}}/0.75)$, which (using the inverse key on the calculator) gives

$$P_{\text{out}} = (0.75)(63.0957) = 47.322 \text{ watts approx., } \textit{answer}.$$

165. Since $\omega_1 = 2\pi f_1 = 1/RC$, we have

$$R = \frac{1}{C 2\pi f_1} = \frac{1}{(5)10^{-8}(2\pi)(7.2)10^3} = 442.1 \text{ ohms, } \textit{answer}.$$

166. Since $\omega/\omega_1 = f/f_1$, eq. (321) becomes $0.6 = \log[1 + (f/f_1)^2]$, then, using the inverse key, $3.9811 = 1 + (f/f_1)^2$, then

$$f = f_1\sqrt{2.9811} = 7.2\sqrt{2.9811} = 12.4314 \text{ kHz } \textit{answer}.$$

167. We'll use eqs. (319), (322), and (325). First note that $\omega_1 = 1/RC = 400$ rad/sec. Then, since $2\pi f_1 = \omega_1$, we have that $f_1 = \omega_1/2\pi = 400/2\pi =$ half-power frequency in *hertz* (Hz). Thus we'll deal with the ratio

$$(\omega/\omega_1) = (f/f_1) = (2\pi f/400) = (0.015\,71 f)$$

(a) For $f = 30$ Hz, $(\omega/\omega_1) = 0.4713$; thus, by eqs. (319) and (322), with $V_i = 10$ volts, we have

$$|\bar{V}_o| = 9.046 \text{ volts, } \textit{answer}.$$

Then, by eq. (325),

$$\phi = -\arctan(0.4713) = -25.235^\circ, \textit{ answer}.$$

(b) For $f = 300$ Hz, $(\omega/\omega_1) = 4.713$; thus, by eqs. (319) and (322), with $V_i = 10$ volts, we have

$$|\bar{V}_o| = 2.076 \text{ volts, } \textit{answer}.$$

Then, by eq. (325),

$$\phi = -\arctan(4.713) = -78.021^\circ, \textit{ answer}.$$

168. In terms of steady-state sinusoidal theory, the explanation is as follows. In order for an OUTPUT WAVE to have exactly the same waveshape as the INPUT WAVE, the fundamental and harmonics of the input wave *must* ALL *be treated the same* in passing through the circuit.

In the present case the ‘‘half-power’’ frequency of the given low-pass filter was found to be $400/2\pi = 63.66$ Hz, while the fundamental frequency of the input square wave is 30 Hz; thus the higher harmonics of the square wave would be severely discriminated against, causing the output wave shape to be considerably different from the input square wave.

169. $C = \dfrac{1}{R\omega_1} = \dfrac{1}{2\pi R f_1} = \dfrac{1}{2\pi(1.2)10^3(2.2)10^3} = 0.0603(10^{-6})\ \mathrm{F} = 0.0603\ \mu\mathrm{F}$, *answer*.

170. (a) In the equation immediately following eq. (328) set dB $= -6$. Then, since if $\log x = y$, then $x = 10^y$, we have that

$$\frac{h}{\sqrt{1+h^2}} = 10^{-0.3}$$

Now square both sides; then, remembering that $(a^x)^2 = a^{2x}$, we have that

$$h^2 = (1+h^2)10^{-0.6}$$

thus, by calculator,

$$h = \sqrt{0.335\,449\,7} = 0.579\,180\,2$$

then, since $h = \omega/\omega_1 = f/f_1$, we have

$$f = hf_1 = 1.274\,196\,4\ \mathrm{kHz} = 1274\ \mathrm{Hz\ approx.},\ \textit{answer}.$$

(b) As before, $h^2 = (1+h^2)10^{-0.2}$ which gives $h = \sqrt{1.709\,713\,7} = 1.307\,560\,2$, so that

$$f = hf_1 = 2.876\,632\,5\ \mathrm{kHz} = 2877\ \mathrm{Hz\ approx.},\ \textit{answer}.$$

171. Applying Ohm’s law and eq. (333) to Fig. 197, we have

$$\bar{V}_o = R\bar{I} = \frac{RV_i}{R + j\omega L}$$

thus

$$\frac{\bar{V}_o}{V_i} = \bar{G} = \frac{R}{R + j\omega L}$$

which gives *eq. (334)*.

Next, upon rationalizing the last expression for $\bar{G}$ above,

$$\bar{G} = \frac{R^2}{R^2 + (\omega L)^2} - j\frac{-\omega LR}{R^2 + (\omega L)^2} = a + jb$$

hence,

$$\phi = \arctan(b/a) = \arctan(-\omega L/R) = -\arctan(\omega L/R)$$

thus verifying *eq. (335)*.

172. Applying Ohm's law and eq. (333) to Fig. 198, we have

$$\bar{V}_{\mathrm{o}} = j\omega L\bar{I} = \frac{j\omega L V_{\mathrm{i}}}{R + j\omega L}$$

thus

$$\frac{\bar{V}_{\mathrm{o}}}{V_{\mathrm{i}}} = \bar{G} = \frac{j\omega L}{R + j\omega L}$$

which gives *eq. (336)*.

Next, upon rationalizing the last expression for G above,

$$\bar{G} = \frac{(\omega L)^2}{R^2 + (\omega L)^2} + j\frac{\omega LR}{R^2 + (\omega L)^2} = a + jb$$

hence

$$\phi = \arctan(b/a) = \arctan(R/\omega L)$$

thus verifying *eq. (337)*.

173. The easiest way is to set $\omega = R/L$ in eq. (334); doing this, eq. (334) becomes

$$|\bar{G}| = \frac{1}{\sqrt{2}} = (2)^{-1/2}$$

thus, by eq. (320) (and note 19 in Appendix), we have

$$\mathrm{dB} = 20\log(2)^{-1/2} = -10\log 2 = -3 \text{ decibels}$$

hence (see definition following Fig. 191) $\omega = R/L$ is the half-power frequency for Fig. 197.

174. By eq. (340),

$$\bar{Z}_0 = \sqrt{(2 + j5)(4 - j3) + (2 + j5)^2/4}$$

thus

$$\bar{Z}_0 = \sqrt{17.75 + j19.00} = [26.001\underline{/46.948^\circ}]^{1/2}$$

Hence, since $[A\underline{/\theta}]^{1/2} = A^{1/2}\underline{/\theta/2}$, we have

$$\bar{Z}_0 = 5.099\underline{/23.474^\circ} = 4.68 + j2.03, \text{ } \textit{answer} \text{ in rectangular form.*}$$

The physical meaning of the above value of $\bar{Z}_0$ is that, if an impedance of $4.68 + j2.03$ ohms is placed across the OUTPUT terminals of this symmetrical T network, the impedance LOOKING INTO THE INPUT TERMINALS is ALSO $4.68 + j2.03$ ohms.

175. In general terms, if $\bar{Z}_{\mathrm{p}} = |\bar{Z}_{\mathrm{p}}|\epsilon^{j\theta}$ and $\bar{Z}_{\mathrm{s}} = |\bar{Z}_{\mathrm{s}}|\epsilon^{j\phi}$, then eq. (341) becomes

$$\bar{Z}_0 = [|\bar{Z}_{\mathrm{p}}||\bar{Z}_{\mathrm{s}}|\epsilon^{j(\theta+\phi)}]^{1/2} = \sqrt{|\bar{Z}_{\mathrm{p}}||\bar{Z}_{\mathrm{s}}|}[\cos(\theta + \phi)/2 + j\sin(\theta + \phi)/2]$$

hence, here,

$$\bar{Z}_0 = \sqrt{375}\,[\cos(-20) + j\sin(-20)] = (18.20 - j6.62) \text{ ohms, } \textit{answer}.$$

* By the basic eq. (176) (section 6.7), there is also a *second* solution, $\bar{Z}_0 = -4.68 - j2.03$; while this is mathematically correct, it has no practical meaning here because of the requirement of a "negative" value of resistance.

176. The best procedure is to first find a general equation for ω, then substitute in the given values of L, C, and R. To do this, we'll make use of eq. (340), because the condition for which $\bar{Z}_{\text{in}} = \bar{Z}_{\text{L}}$ is given by this equation; the steps are as follows.

First, using the terminology of Fig. 201, $\bar{Z}_1/2 = j\omega L$, thus $\bar{Z}_1 = j2\omega L$. Next, since $\bar{Z}_2 = 1/j\omega C$, we have, putting these values of $\bar{Z}_1$ and $\bar{Z}_2$ into eq. (340), that

$$\bar{Z}_0 = \sqrt{\frac{2L}{C} - \frac{4\omega^2 L^2}{4}} = \sqrt{\frac{2L}{C} - \omega^2 L^2}$$

Now, upon squaring both sides then solving for ω, you should find that

$$\omega = \pm\sqrt{\frac{2}{LC} - \frac{\bar{Z}_0^2}{L^2}}$$

where we've shown the "plus or minus" sign just for the sake of mathematical completeness; actually, however, since "negative frequencies" don't exist in the real world, we'll disregard the minus possibility in this case. If, now, you replace L and C with the given values and also set $\bar{Z}_0 = \bar{Z}_{\text{in}} = 2$, you should find that

$$\omega = 100\sqrt{3} = 173.21 \text{ rad/sec}, \ \textit{answer}.$$

177. Here we're dealing with the condition in which $\bar{Z}_{\text{in}} = \bar{Z}_{\text{L}} = \bar{Z}_0$, so that eq. (342) applies. Thus, in general terms we have, in this case

$$\frac{\bar{I}_1}{\bar{I}_2} = \frac{2 + j\omega L}{2 - j\omega L} = \frac{(4 - \omega^2 L^2) + j4\omega L)}{4 + \omega^2 L^2}$$

Now substituting in the values $L = (2)10^{-2}$ and $\omega = 100\sqrt{3}$, you should find that

$$\bar{I}_1/\bar{I}_2 = (-1 + j\sqrt{3})/2 = -0.500 + j0.866, \ \textit{answer}.$$

178. In Fig. 207, the left- and right-hand components labeled $\bar{Z}_0$ would be replaced by $\bar{Z}_{\text{in}}$ and $\bar{Z}_{\text{L}}$ respectively. Thus we would have that $\bar{V}_1 = \bar{Z}_{\text{in}}\bar{I}_1$ and $\bar{V}_2 = \bar{Z}_{\text{L}}\bar{I}_2$; hence the second equation following Fig. 207 would become

$$\bar{Z}_{\text{in}}\bar{I}_1 = \bar{Z}_1\bar{I}_1/2 + \bar{Z}_1\bar{I}_2/2 + \bar{Z}_{\text{L}}\bar{I}_2$$

which gives *eq. (343)*.

179. First, for $\omega = 200$, we have $\bar{Z}_1/2 = j\omega L = j4$ and $\bar{Z}_2 = -j2$ and, as before, $\bar{Z}_{\text{L}} = 2$. Next, put these values into eq. (338); doing this, you should find that $\bar{Z}_{\text{in}} = 1 + j$. Now, upon substituting all the foregoing values into eq. (343), we have that

$$\frac{\bar{I}_1}{\bar{I}_2} = \frac{2 + j4}{1 - j3} = \frac{2(1 + j2)(1 + j3)}{10} = (-1 + j), \ \textit{answer}.$$

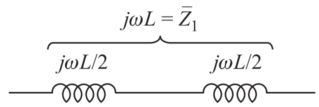
180. Equation (342) applies, because the network is terminated in $\bar{Z}_0$. Hence, all we need do is substitute the known values of $\bar{Z}_1$ and $\bar{Z}_0$ into eq. (342); thus

$$\frac{\bar{I}_1}{\bar{I}_2} = \frac{5.68 + j4.53}{3.68 - j0.47} = \frac{7.27\underline{/38.57^\circ}}{3.71\underline{/-7.28}} = 1.96\underline{/45.85^\circ}$$

Thus, $\bar{I}_1 = (1.96\underline{/45.85^\circ})\bar{I}_2$, meaning that the magnitude of $\bar{I}_1$ is 1.96 times the magnitude of $\bar{I}_2$ and that $\bar{I}_1$ *leads* $\bar{I}_2$ by 45.85°, *answer*.*

* Just as an example, if $\bar{I}_2$ were equal to $2\underline{/30^\circ}$, then $\bar{I}_1 = 3.92\underline{/75.85^\circ}$.

181. (a) Since we're using the notation of Fig. 201,



thus $L/2 = 800\,\mu\text{H}$, we have $L = 1600\,\mu\text{H} = (16)10^{-4}$ H and $C = (4)10^{-8}$ F, and upon putting these values into eq. (250) we find that $R_L = 200$ ohms, *answer*.

(b) Set $\bar{Z}_0 = 0$ in eq. (349); then, squaring both sides and solving for ω gives

$$\omega = \frac{2}{\sqrt{LC}}$$

which, upon substituting $L = (16)10^{-4}$ and $C = (4)10^{-8}$, gives

$$\omega = (2.5)10^5 \text{ radians/second, } \textit{answer}.$$

or, if we wish,

$$f = \omega/2\pi = 39{,}789 \text{ Hz}$$

(c) From the above, $\omega = (25)10^4, L = (16)10^{-4}, C = (4)10^{-8}$, thus $\bar{Z}_1 = j\omega L = j400, \bar{Z}_2 = -j/\omega C = -j100$, and also $\bar{Z}_L = R_L = 200$. Upon substituting these values into eq. (338) you should find that

$$\bar{Z}_{in} = j200 + \frac{200(1-j)}{2+j}$$

which, after rationalizing the last fraction to the right, becomes

$$\bar{Z}_{in} = j200 + 40 - j120 = 40(1 + j2) \text{ ohms, } \textit{answer}.$$

(d) The value of $\bar{Z}_0$ can be calculated for any frequency from eq. (340), but the generator will *see* $\bar{Z}_0$ only if $\bar{Z}_L$ is made *equal* to $\bar{Z}_0$, which is not the case in part (c) because $\bar{Z}_L$ is given to be 200 ohms (and *not* zero ohms).

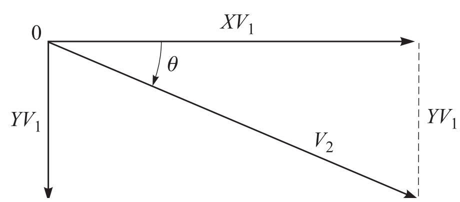
182. We use eq. (359), in which ω_c, L, and C are in radians per second, henrys, and farads. Thus, upon solving eq. (359) for L, we have

$$L = \frac{4}{\omega_c^2 C} = \frac{4}{4\pi^2(10^{10})(1.5)10^{-8}} = (6.7547)10^{-4} \text{ H} = 675.47\,\mu\text{H}$$

hence inductance of each coil $= L/2 = 337.7\,\mu\text{H}$, *answer*.

183. By eq. (355), $\bar{V}_2 = \left(\dfrac{1}{A + jB}\right)V_1 = \left(\dfrac{A}{A^2 + B^2} - j\dfrac{B}{A^2 + B^2}\right)V_1 = (X - jY)V_1$

Graphically,



hence

$$\phi = -\arctan(Y/X) = -\arctan(B/A)$$

or, since $A = (1 - 2h^2)$ and $B = 2h(1 - h^2)$, we have

$$\phi = -\arctan[2h(1 - h^2)/(1 - 2h^2)], \; \textit{answer}.$$

Note: When using the above formula we should remember that "arctan x" is a multiple-valued function (for example, arctan $1.0 = 45^\circ = 225^\circ$) and, for this reason, uncertainty about the correct value of ϕ can sometimes arise. That is, we may be uncertain about which of the four quadrants the angle ϕ terminates in. In such a case we must refer back to the basic complex relationship from which the formula for ϕ was derived, which in the above case is

$$\bar{V}_2 = (X - jY)V_1$$

in order to be sure which quadrant ϕ terminates in.

184. (a) The purpose here is to emphasize that the value of an induced voltage does not depend on HOW MUCH current is flowing, but only on how fast the current is changing, that is, upon the *rate of change of current*, $\mathrm{d}i/\mathrm{d}t$. Thus here, if i_1 is *constant* in value, then $\mathrm{d}i_1/\mathrm{d}t = 0$; hence, by eq. (368), the voltage induced into the secondary coil is *zero* (*answer*).

(b) By eqs. (368) and (371),

$$\begin{aligned} v_2 &= k\sqrt{L_1 L_2}\; \mathrm{d}i_1/\mathrm{d}t \\ &= 0.85\sqrt{13.5}(25) = 78.0775 \text{ volts}, \; \textit{answer}. \end{aligned}$$

185. (a) Inspection of eqs. (368) and (369) shows that the statement *is true*, because, if $\mathrm{d}i_1/\mathrm{d}t = \mathrm{d}i_2/\mathrm{d}t$, and since M has the same value in both equations, inspection of the equations shows that, for this condition, $v_1 = v_2$, as proposed, *answer*.

(b) By eqs. (369) and (371),

$$v_1 = 0.85\sqrt{13.5}(25) = 78.0775 \text{ volts}, \; \textit{answer}.$$

186. (a) Applying the right-hand rule to Fig. 217, note that if the primary and secondary currents were BOTH *reversed in direction* their magnetic effects would still be additive. Thus reversing *both* reference current arrows would have *no effect* on eq. (389).

(b) In this case $\bar{I}_1$ and $\bar{I}_2$ would have *opposite effects magnetically*, and eq. (385) would be $-j\omega M\bar{I}_1 + \bar{Z}_2\bar{I}_2 = 0$; thus eq. (389) would become $\bar{I}_2 = j\omega M\bar{I}_1/\bar{Z}_2$, showing that this would *reverse the sign* of $\bar{I}_2$. Thus, in simple circuits in which we need to find only the *magnitude* of the secondary current, it's not necessary to be concerned about the polarity dots at all.

187. First, $\omega L_1 = \omega L_2 = (200)(45)10^{-3} = 9$ ohms. Next, making use of eq. (371), you should find $M = (4.05)10^{-2}$ henrys, and thus $\omega M = (200)(4.05)10^{-2} = 8.1$ ohms. Next, by eq. (387), $\bar{Z}_{\text{ref}} = \dfrac{(8.1)^2}{2 + j9}$ ohms, which, after "rationalizing" (multiplying numerator and denominator by the conjugate of the denominator), gives $\bar{Z}_{\text{ref}} = (1.5438 - j6.9469)$, and hence, by eq. (388),

$$|\bar{I}_1| = \left|\frac{36}{6.5438 + j2.0531}\right| = \frac{36}{6.8583} = 5.249 \text{ amperes}, \; \textit{answer}.$$

Discussion Note: In the solution to the above problem, note that the impedance reflected into the primary coil is *capacitive* in nature, although the secondary circuit is, by itself, *inductive* in nature, being $(2 + j9)$ ohms. This illustrates the fact that if $\bar{Z}_2$ is *inductive*, then the impedance reflected into the primary coil will be *capacitive* in form. The reason for this can be seen from eq. (387), for $\bar{Z}_2 = R + jX$; thus

$$\bar{Z}_{\mathrm{ref}} = \frac{\omega^2 M^2}{R + jX} = \frac{\omega^2 M^2}{R^2 + X^2}(R - jX)$$

which is a *capacitive* type impedance. Or, if the net reactance of $\bar{Z}_2$ is *capacitive* in nature, then the impedance reflected into the primary coil will be *inductive* in nature. Thus, for $\bar{Z}_2 = R - jX$, eq. (387) gives

$$\bar{Z}_{\mathrm{ref}} = \frac{\omega^2 M^2}{R - jX} = \frac{\omega^2 M^2}{R^2 + X^2}(R + jX)$$

which is *inductive* in nature.

Of course, if the secondary circuit is *purely resistive* (because of series resonance on the secondary side), the reflected impedance will also be a pure resistance.

188. (a) Let us first calculate the reactances; thus

$$X_{L1} = X_{L2} = 4(10^5)(60)10^{-6} = 24 \text{ ohms},$$

$$X_C = 1/4(10^5)(5)10^{-7} = 5 \text{ ohms},$$

$$\omega M = 20 \text{ ohms}$$

Now let us look back to Fig. 222 and eq. (383). From Fig. 222 we see that $\bar{Z}_{\mathrm{b}}$ is the impedance *connected to the terminals of the secondary coil*. Thus, in our problem here, $\bar{Z}_{\mathrm{b}}$ is the parallel combination of R and $-jX_C$, so that for Fig. 225 we have

$$\bar{Z}_{\mathrm{b}} = \frac{-jX_C R}{R - jX_C} = \frac{-j50}{10 - j5} = (2 - j4) \text{ ohms}$$

hence

$$\bar{Z}_2 = \bar{Z}_{\mathrm{b}} + JX_{L2} = (2 + j20) \text{ ohms}$$

Now, putting all the known values into eq. (387) you should find that

$$\bar{Z}_{\mathrm{ref}} = (1.980 - j19.80) \text{ ohms}$$

Since there is no resistance in the primary circuit, $\bar{Z}_1 = jX_{L1} = j24$ ohms. Thus, by eq. (388), we have

$$|\bar{I}_1| = \left|\frac{28}{1.980 + j4.20}\right| = 6.030 \text{ amperes, } \textit{answer}.$$

(b) By eq. (389),

$$|\bar{I}_2| = \left|\frac{-j\omega M \bar{I}_1}{\bar{Z}_2}\right| = \frac{\omega M |\bar{I}_1|}{|\bar{Z}_2|} = \frac{120.6}{\sqrt{404}} = 6.000 \text{ amperes, } \textit{answer}.$$

189. (a) First, $\omega = 2\pi f = 10^5$ radians/second; thus, the values of $\bar{I}_1, \bar{I}_2$, and $\bar{I}_3$ are

$$\bar{I}_1 = \frac{40}{j8} = -j5\,\text{A}, \qquad \bar{I}_2 = \frac{40}{10} = 4\,\text{A}, \qquad \bar{I}_3 = \frac{40}{-j4} = j10\,\text{A}$$

Next, the value of $\bar{I}_4$ is found by making use of eqs. (388), (387), and (371); thus

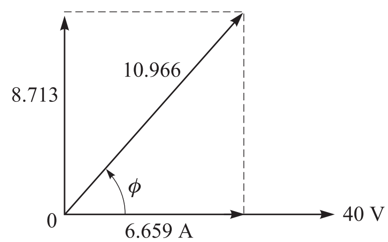
$$\bar{I}_4 = \frac{40}{jX_{L1} + Z_{\text{ref}}} = \frac{40}{j12 + (5.099 - j19.120)} = \frac{40}{5.099 - j7.120} = (2.659 + j3.713)\text{ A}$$

Hence, $\bar{I} = \bar{I}_1 + \bar{I}_2 + \bar{I}_3 + \bar{I}_4 = (6.659 + j8.713)$, thus

$$|\bar{I}| = 10.966 \text{ amperes, } \textit{answer}.$$

(b) $$|\bar{I}_s| = \left|\frac{-j\omega M I_4}{Z_s}\right| = \frac{\omega k\sqrt{L_1 L_2}|2.659 + j3.713|}{|2 + j7.5|} = 7.292 \text{ amperes, } \textit{answer}.$$

(c) The principle referred to here is that TRUE POWER is equal to the voltage



times the component of current IN PHASE with the voltage. From the solution to (a), we have the vector diagram shown below.

Thus, $P = (40)(6.659) = 266.36$ watts approx., *answer*.

(d) Note that the value of ϕ can be found from the above figure. However, since we're instructed to use Fig. 155, we must find the value of $\bar{Z} = R + jX$, and to do this we'll use the basic relationship

$$\bar{Z} = \frac{\bar{V}}{\bar{I}} = \frac{40}{6.659 + j8.713} = (2.2149 - j2.8981) \text{ ohms}$$

and hence, by Fig. 155,

$$\phi = \arctan X/R = \arctan(-2.8981/2.2149) = -52.611^\circ$$

and thus, by eq. (228),

$$P = (40)(10.966)\cos(-52.611^\circ) = 266.35 \text{ watts, } \textit{answer}.$$

(e) As explained in section 8.5, there is no net energy loss in an ideal inductor or capacitor; that is, energy is removed from a network only through resistive elements which, in Fig. 226, are the devices having 10 ohms and 2 ohms of resistance. Since the current in the 10-ohm load is 4 amperes and the magnitude of the current in the 2-ohm load is 7.292 amperes, we have that

$$P = (4)^2(10) + (7.292)^2(2) = 266.35 \text{ watts, } \textit{answer}.$$

As must be the case, all three procedures give the same value of P, after taking into account a slight round-off difference.)

190. Since the two coils have equal inductances of L henrys, let us, for convenience, refer to them as the ''left-hand'' and ''right-hand'' (LH and RH) inductors. Next note that there are *five* voltage drops to consider, as follows.

$$j\omega L\bar{I} \quad \text{and} \quad j\omega L\bar{I} = \text{the ordinary voltage drops across the inductors,}$$
$$j\omega M\bar{I} = \text{voltage induced to LH coil from RH coil,}$$
$$j\omega M\bar{I} = \text{voltage induced into RH coil from LH coil,}$$
$$R\bar{I} = \text{voltage drop across the pure resistance of } R \text{ ohms.}$$

Whether or not the two mutually induced voltages AID each other or OPPOSE each other depends upon the relative senses in which the coils are wound, and this information is given us by the placement of the dots. Thus, upon taking the placement of the dots into consideration, the equations are

$$\text{(a)} \quad \bar{I} = \frac{\bar{V}}{R + j2\omega(L+M)}, \ \textit{answer}.$$

$$\text{(b)} \quad \bar{I} = \frac{\bar{V}}{R + j2\omega(L-M)}, \ \textit{answer}.$$

191. First, $\omega L_1 = 25$ ohms, $\omega L_2 = 56.25$ ohms, and $\omega M = \omega k\sqrt{L_1 L_2} = 30$ ohms. Then, for the given directions of the reference arrows and the placement of the dots, the loop voltage equations for Fig. 229 are

$$\text{around LH loop:} \quad (10 + j25)\bar{I}_1 - (10 + j30)\bar{I}_2 = 50$$
$$\text{around RH loop:} \quad -(10 + j30)\bar{I}_1 + (15 + j56.25)\bar{I}_2 = 0$$

Since $\bar{V}_o$ is the voltage drop across the 10-ohm resistance, we have $\bar{V}_o = 10(\bar{I}_1 - \bar{I}_2)$; thus

$$|\bar{V}_o| = 10|\bar{I}_1 - \bar{I}_2| \tag{A}$$

Now apply the method of determinants to the above two simultaneous equations; doing this, you can verify that

$$\bar{I}_1 = \frac{50(15 + j56.25)}{-456.25 + j337.5} \quad \text{and} \quad \bar{I}_2 = \frac{50(10 + j30)}{-456.25 + j337.5}$$

thus

$$\bar{I}_1 - \bar{I}_2 = \frac{50(5 + j26.25)}{-456.25 + j337.5} = \frac{0.50(5 + j26.25)}{-4.5625 + j3.375}$$

and hence

$$|I_1 - I_2| = \frac{0.5\sqrt{714.0625}}{\sqrt{32.207031}} = 2.354 \text{ approx.}$$

and thus, by eq. (A),

$$|V_o| = 23.54 \text{ volts}, \ \textit{answer}.$$

192. The dot for the secondary coil would move to the opposite end from that shown in Fig. 229. Thus the two simultaneous equations would now be

$$\text{around the LH loop:} \quad (10 + j25)\bar{I}_1 + (-10 + j30)\bar{I}_2 = 50$$
$$\text{around the RH loop:} \quad (-10 + j30)\bar{I}_1 + (15 + j56.25)\bar{I}_2 = 0$$

which upon solving for $\bar{I}_1$ and $\bar{I}_2$ gives

$$\bar{I}_1 = \frac{50(15 + j56.25)}{-456.25 + j1537.50} \quad \text{and} \quad \bar{I}_2 = \frac{-50(-10 + j30)}{-456.25 + j1537.50}$$

thus

$$\bar{I}_1 - \bar{I}_2 = \frac{50(5 + j86.25)}{-456.25 + j1537.50}$$

Hence, by eq. (A),

$$|\bar{V}_o| = 10|\bar{I}_1 - \bar{I}_2| = 26.94 \text{ volts approx., } \textit{answer}.$$

193. If $\bar{N}$ is a complex number in the form of $\bar{N} = \dfrac{a + jb}{c + jd}$, then

$$\bar{N} = \frac{a + jb}{c + jd} = \frac{(a + jb)(c - jd)}{c^2 + d^2} = \left[\frac{ac + bd}{c^2 + d^2}\right] + j\left[\frac{bc - ad}{c^2 + d^2}\right]$$

and thus $\bar{N} = |\bar{N}|\underline{/\phi}$, where $\phi = \arctan\dfrac{bc - ad}{ac + bd}$; thus, from the solution to problem 191, we have that

$$a = b, \; b = 26.25, \; c = -456.25, \; d = 337.5$$

hence

$$\bar{V}_o = 23.54\underline{/\arctan(-2.077\,20)} = 23.54\underline{/-64.29^\circ}, \; \textit{answer}.$$

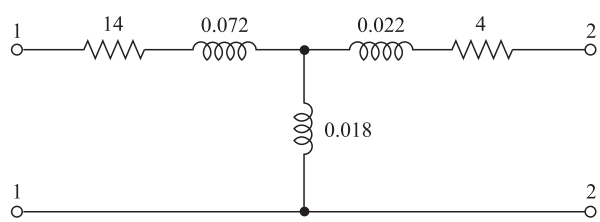
194. From the solution to problem 192 we have that

$$a = 5, \; b = 86.25, \; c = -456.25, \; d = 1537.50$$

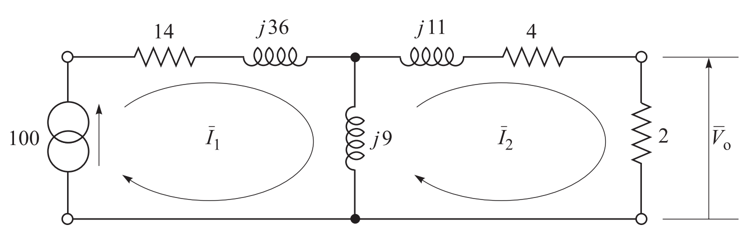
hence

$$\bar{V}_o = 26.94\underline{/\arctan(-0.36093)} = 26.94\underline{/-19.85^\circ}, \; \textit{answer}.$$

195. First, by eq. (371), $M = (0.3)\sqrt{(0.09)(0.04)} = 0.018$ H (henry). Thus $(L_1 - M) = 0.072$ H, $(L_2 - M) = 0.022$ H, and, since $R_1 = 14$ ohms and $R_2 = 4$ ohms, the equivalent T is as shown below, the values being in ohms and henrys.



196. First:



Thus, by inspection, the loop-voltage equations are

$$(14 + j45)\bar{I}_1 \qquad - j9\bar{I}_2 = 100$$
$$-j9\bar{I}_1 + (6 + j20)\bar{I}_2 = 0$$

from which we find that

$$\Delta = 5(147 + j110)$$

Solving for I_2,

$$\bar{I}_2 = j900/\Delta$$

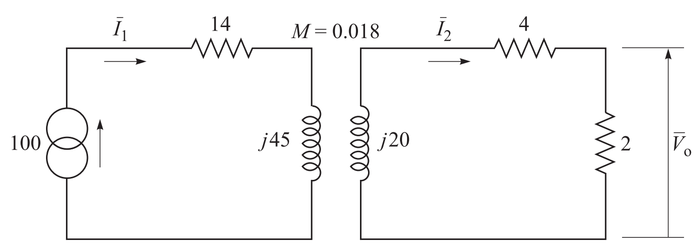
then

$$|\bar{I}_2| = \left|\frac{j900}{5(147 + j110)}\right| = \frac{180}{\sqrt{(147)^2 + (110)^2}} = 0.98039\,\text{A}$$

hence

$$|\bar{V}_o| = 2|\bar{I}_2| = 1.961 \text{ volts approx., } \textit{answer}.$$

197. The circuit will be that of Fig. 230, with 100-volt generator applied to terminals (1, 1) and 2-ohm load connected to terminals (2, 2), where, from problem 195, $M = 0.018$ H; thus



By eqs. (382) and (383), $\bar{Z}_1 = 14 + j45$ and $\bar{Z}_2 = 6 + j20$, and upon substituting these values into eq. (386) you should find that

$$\bar{I}_1 = \frac{2(3 + j10)}{-7.35 + j5.50}$$

and thus, by eq. (389),

$$\bar{I}_2 = \frac{-j(500)(0.018)(2)(3 + j10)}{(6 + j20)(-7.35 + j5.5)} = \frac{18(10 - j3)}{-(154 + j114)}$$

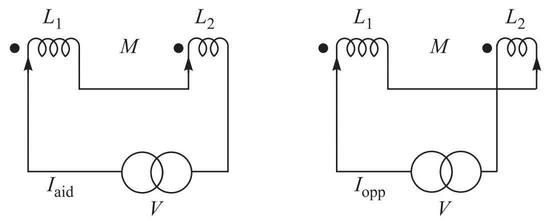
hence

$$|\bar{I}_2| = 18\sqrt{\frac{109}{36{,}712}} = 0.9808\,\text{A}$$

thus,

$$|\bar{V}_o| = (2)(0.9808) = 1.962 \text{ volts approx., } \textit{answer}.$$

198. Noting the dot-marked terminals, the two connections are as follows, in which the same voltage V is applied to both connections.



Note that the current in L_1 is equal to the current in L_2. Thus, from our work in sections 10.2 and 10.3, the equations are

$$V = j(\omega L_1 + \omega L_2 + \omega M + \omega M)I_{\text{aid}} = j\omega(L_1 + L_2 + 2M)I_{\text{aid}}$$

and

$$V = j(\omega L_1 + \omega L_2 - \omega M - \omega M)I_{\text{opp}} = j\omega(L_1 + L_2 - 2M)I_{\text{opp}}$$

hence

$$(L_1 + L_2 + 2M) = \textit{inductance} \text{ measured in AIDING case} = L_{\text{aid}}$$

and

$$(L_1 + L_2 - 2M) = \textit{inductance} \text{ measured in OPPOSING case} = L_{\text{opp}}$$

Thus, subtracting the second equation from the first, then solving for M, we have

$$M = (1/4)(L_{\text{aid}} - L_{\text{opp}})$$

hence, by eq. (371),

$$k = \frac{L_{\text{aid}} - L_{\text{opp}}}{4\sqrt{L_1 L_2}}, \ \textit{answer}.$$

The above result is of considerable practical importance because it allows us, working in the laboratory, to physically adjust the spacing between two coils until a required value of k is obtained.

199. (a) Here $T = 4$; hence, by eq. (418)

$$\bar{Z}_{\text{in}} = T^2 \bar{Z}_{\text{L}} = (16)(3 - j5) = (48 - j80) \text{ ohms}, \ \textit{answer}.$$

(b) By eq. (419), $V_2 = V_1(N_2/N_1) = 240/4 = 60$ volts; hence, by Ohm's law, the secondary current is

$$\bar{I}_2 = \frac{60}{3 - j5} = \frac{60(3 + j5)}{34} = (5.294 + j8.824) \text{ amperes}, \ \textit{answer}.$$

(c) One way is to use the formula $P = RI^2$, where I is the magnitude of the current. Thus we can take the square of the magnitude of the secondary current times the resistance in the secondary circuit,

$$P = (3)(10.290)^2 = 317.65 \text{ watts}, \ \textit{answer}.$$

A second way is to multiply the secondary voltage by the "in-phase" component of the secondary current (section 8.5). Thus

$$P = (60)(5.294) = 317.65 \text{ watts}, \ \textit{answer}.$$

A third way is to use the power factor, $\cos\phi = R/Z = 3/\sqrt{34} = 0.5145$, thus

$$P = V_2 I_2 \cos\phi = (60)(10.290)(0.5145) = 317.65 \text{ watts}, \ \textit{answer}.$$

200. $\bar{Z}_{in}$ is equal to the denominator of eq. (386), where $M^2 = k^2 L_1 L_2$, and where, in this case, $\bar{Z}_1 = j\omega L_1$ and $\bar{Z}_2 = j\omega L_2$. Thus

$$\bar{Z}_{in} = j\omega L_1 + \frac{\omega^2 k^2 L_1 L_2}{j\omega L_2} = j\omega L_1(1 - k^2),\ \textit{answer}.$$

201. Setting $\omega = 2\pi f = 376.99, L_1 = 4$, and $k = 0.995$ in the answer to problem 200, we find that $\bar{Z}_{in} = j15.04$ ohms. Thus by Ohm's law,

$$\bar{I}_1 = 120/j15.04 = -j7.98 \text{ amperes},\ \textit{answer}.$$

(Note that in the ''ideal'' case, $k = 1$, the shorted secondary would entirely neutralize the primary inductance, causing theoretically ''infinite'' primary line current to flow.)

202. Because the INSTANTANEOUS SUM of the three generator voltages around the loop is always equal to *zero*. This can be shown as follows. Let us, in Fig. 247, select one of the three generators to be the ''reference generator,'' and let the equation of this voltage wave be given by $V \sin \omega t$, where V is the peak voltage (V and ω having the same values in all three generators). Now, for convenience, let $V = 1$ volt and $\omega = 1$ rad/sec. Then, letting v be the *sum* of the three voltages at any time t, we would have that, at any instant,

$$v = \sin t + \sin(t - 120^\circ) + \sin(t - 240^\circ)^*$$

Now let us make use of the following trigonometric identity

$$\sin(x + y) = \sin x \cos y + \cos x \sin y \qquad \text{(note 7 in Appendix)}$$

Now, in the above identity set $x = t, y = -120^\circ$ in one case, and $x = t, y = -240^\circ$ in the second case. Doing this, and remembering that $\sin(-h) = -\sin h$ and $\cos(-h) = \cos h$, you should find that the preceding equation for v becomes

$$v = \sin t + \sin t \cos 120^\circ - \cos t \sin 120^\circ + \sin t \cos 240^\circ - \cos t \sin 240^\circ$$

thus

$$v = (\sin t)(1 - 0.5 - 0.5) + (\cos t)(-0.8660 + 0.8660) = 0$$

showing that the instantaneous sum of the voltages around the closed loop of Fig. 247 is always equal to *zero*. Thus NO CURRENT can flow under such conditions.

Note that the same conclusion is reached if we are thinking in terms of sinusoidal steady-state vector representation. In that case we would be dealing with the *vector sum* of the three voltages around the loop, which (as is evident from inspection of Fig. 252) will add up to *zero*.

203. (a) By eq. (435),

$$V_p = 3300/1.732 = 1905.3 \text{ volts},\ \textit{answer}.$$

(b) First, the phase voltage on the Y-connected secondary side would, by eq. (435), be equal to $V_p = 66{,}000/1.732 = 38{,}106.2$ volts. This voltage would then, going from right to left in Fig. 254, be ''stepped down'' by a factor of 12; thus the line voltage on the delta-connected primary side would be $38{,}106.2/12 = 3175.5$ volts, making the generator phase voltage equal to $3175.5/1.732 = 1833.4$ volts, *answer*.

* Here $\omega t = t$ radians, so that 120° and 240° should really be expressed in radians instead of degrees. For the special demonstration here, however, no difficulty arises.

204. One way is to make use of eq. (440), as follows. First, by eq. (435)

$$V_L = (1.732)(330) = 571.56 \text{ volts}$$

Next,

$$I_L = I_p = V_p/|\bar{Z}| = 330/\sqrt{306} = 18.865 \text{ amperes}$$

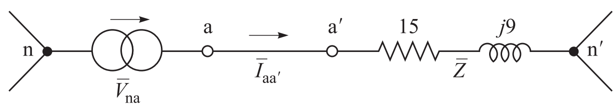
Then

$$\cos\phi = R/|\bar{Z}| = 15/\sqrt{306} = 0.8575$$

Hence

$$P_T = 1.732 V_L I_L \cos\phi = 16014.03 \text{ watts, } \textit{answer}, \text{ or,}$$
$$= 16.014 \text{ kilowatts, } \textit{answer}.$$

205. First consider the phase voltage $\bar{V}_{na}$; note that the phase current here is the same as the line current $\bar{I}_{aa'}$ as shown in the figure below.



Since the system is given to be completely balanced, the voltage at junction point n′ is the same as the voltage at point n (a wire connected from n to n′ would show zero current). Hence, by Ohm's law,

$$\bar{I}_{aa'} = \frac{\bar{V}_{na}}{\bar{Z}} = \frac{V_{na}}{15 + j9} = 0.0098(5 - j3)V_{na}$$

thus $\phi = \arctan(-3/5) = -30.96^\circ$, showing that line current $\bar{I}_{aa'}$ LAGS phase voltage $\bar{V}_{na}$ by approximately 31°. Hence, as inspection of Fig. 253 shows, line current $\bar{I}_{aa'}$ lags line voltage $\bar{V}_{AB}$ by approximately 61°. Likewise, the same procedure will show that $\bar{I}_{bb'}$ and $\bar{I}_{cc'}$, respectively, will lag line voltages $\bar{V}_{BC}$ and $\bar{V}_{CA}$ by 61°.

206. Remember that V' and I' in eq. (445) are *maximum* (peak) values. Then note that eq. (445) can be written as

$$p = \frac{3}{2} V'I' \cos\phi = 3 \frac{V'}{\sqrt{2}} \frac{I'}{\sqrt{2}} \cos\phi = 3VI \cos\phi$$

where V and I are now *rms* values (rms values are used in AVERAGE POWER calculations). Next, inspection of Fig. 257 will show that $V = V_L$, and also that $I = I_L/\sqrt{3}$, where V and I are the rms values for each of the three impedances $\bar{Z}$; thus

$$p = 3V_L \frac{I_L}{\sqrt{3}} \cos\phi = \sqrt{3} V_L I_L \cos\phi$$

showing that in a balanced three-phase system the instantaneous power p is the SAME as the total average power P_T.

207. To express the SUM of any two sets we first label the sets in A, B, and C notation. We then go in the ccw sense, in both sets, first finding the sum of the two A vectors, then the sum of the two B vectors, then the sum of the two C vectors, the SUM of the two sets then being expressed in the form of eq. (449).

In the particular case here, it's given that both sets are to be labeled in the sequence "ABC"; hence, after applying eq. (446), eq. (449) becomes*

$$\bar{S}_1 + \bar{S}_2 = (A_1 + A_1') + (A_1 + A_1')\epsilon^{j120} + (A_1 + A_1')\epsilon^{j240}$$

or, letting $A_1 + A_1' = A_1''$, the above becomes

$$\bar{S}_1 + \bar{S}_2 = A_1'' + A_1''\epsilon^{j120} + A_1''\epsilon^{j240}$$

which is exactly the basic form of eq. (446), showing that the SUM OF TWO POSITIVE-SEQUENCE SETS is a BALANCED set of three vectors. Thus an unbalanced set of three vectors (Fig. 261) cannot be expressed as the sum of two positive-sequence sets (or the sum of two negative-sequence sets).

208. The same general discussion, given at the start of the solution to problem 207 above, applies here also. Then, referring to Figs. 262 and 263, we have, in this case

BY FIG. 262:	BY FIG. 263:
$\bar{B}_1 = \bar{A}_1\epsilon^{j120}$	$\bar{B}_2 = \bar{A}_2\epsilon^{j240}$
$\bar{C}_1 = \bar{A}_1\epsilon^{j240}$	$\bar{C}_2 = \bar{A}_2\epsilon^{j120}$

and, upon substituting these values into eq. (449), you can show that

$$\bar{S}_1 + \bar{S}_2 = (\bar{A}_1 + \bar{A}_2) + (\bar{A}_1 + \bar{A}_2\epsilon^{j}120)\epsilon^{j120} + (\bar{A}_1 + \bar{A}_2\epsilon^{-j120})\epsilon^{j240} \qquad \text{(A)}$$

However, since $(\bar{A}_1 + \bar{A}_2)$ does *not*, in general, equal either $(\bar{A}_1 + \bar{A}_2\epsilon^{j120})$ or $(\bar{A}_1 + \bar{A}_2\epsilon^{-j120})$, it follows that eq. (A) is not of the general form of eq. (446) and thus must represent an *unbalanced* set of three vectors.

209. We wish to express the given unbalanced set as the sum of positive, negative, and zero sequence sets. To do this we'll make use of eqs. (460) through (466), where, in this case (angles in degrees), it's given that

$$\bar{A} = 15\underline{/0} = 15\epsilon^{j0} \quad \bar{B} = 9\underline{/100} = 9\epsilon^{j100} \quad \bar{C} = 24\underline{/215} = 24\epsilon^{j215}$$

Let us begin by substituting the above values into eq. (460), which gives us

$$\bar{A}_1 = 5 + 3\epsilon^{j340} + +8\epsilon^{j335}$$

Now, in order to find the value of the indicated sum of the above complex numbers, they must first be put into the rectangular form $(a + jb)$, which is done by applying Euler's formula $\epsilon^{j\theta} = \cos\theta + j\sin\theta$. Upon doing this, you should find that

$$\bar{A}_1 = 15.070 - j4.407$$

showing that the vector $\bar{A}_1$ lies in the fourth quadrant, thus (continuing to use degrees)

$$\bar{A}_1 = 15.701\underline{/-16.301}$$

or, in exponential form,

$$\bar{A}_1 = 15.701\epsilon^{-j16.301}$$

* In the solution here we'll omit the "overscore" (vector) notation on the As because of the presence of so many "prime marks" that will be used with them.

and hence, upon applying eqs. (461) and (462), we have that

$$\bar{S}_1 = 15.701(\epsilon^{-j16.301} + \epsilon^{j103.699} + \epsilon^{j223.699)}) = 0$$

this being the POSITIVE SEQUENCE set. (The vector sum is zero because the set is balanced.) Thus, in this particular problem, the vectors depicted in Fig. 262, in the ccw sense, have the values

$$\bar{A}_1 = K\epsilon^{-j16.301} \quad \bar{B}_1 = K\epsilon^{j103.699} \quad \bar{C}_1 = K\epsilon^{j223.699}$$

where $K = 15.701$.

Now we must turn our attention to finding the values in the negative sequence set, the first step being to find the value of $\bar{A}_2$. To do this we use the same procedure as was used to find $\bar{A}_1$, except that now we use eq. (463) instead of (460). Doing this, you should find that

$$\bar{A}_2 = 5 + 3\epsilon^{j220} + 8\epsilon^{j455}$$

hence

$$\bar{A}_2 = 2.005 + j6.041$$

and thus we find that $\bar{A}_2$ lies in the first quadrant, having magnitude of 6.365 and angular displacement of arctan $(6.041/2.005) = 71.639°$; that is, in exponential notation,

$$\bar{A}_2 = 6.365\epsilon^{j71.639}$$

and thus, applying eqs. (464) and (465), we have that

$$\bar{S}_2 = 6.365(\epsilon^{j71.639} + \epsilon^{j311.639} + \epsilon^{j191.639}) = 0$$

this being the NEGATIVE SEQUENCE set, the vector sum again being zero, because this is a balanced set with equal magnitudes and equal phase displacements of 120 degrees.

Thus, in this particular problem, the vectors depicted in Fig. 263, in the ccw sense, have the values

$$\bar{A}_2 = K\epsilon^{j71.639} \quad \bar{C}_2 = K\epsilon^{j191.639} \quad \bar{B}_2 = K\epsilon^{j311.639}$$

where $K = 6.365$.

All that remains now is to find the value of the zero-sequence set, which is done by substituting the given values of $\bar{A}$, $\bar{B}$, and $\bar{C}$ into eq. (466). Doing this, you should find that

$$\bar{A}_0 = \bar{B}_0 = \bar{C}_0 = 5 + 3\epsilon^{j100} + 8\epsilon^{j215} = -(2.074 + j1.634),$$

a third-quadrant vector that let us express in the forms $\bar{A}_0 = \bar{B}_0 = \bar{C}_0 = 2.640\underline{/218.233°} = 2.640\epsilon^{j218.233}$. Thus the total value of the ZERO-SEQUENCE vector is

$$\bar{S}_0 = \bar{A}_0 + \bar{B}_0 + \bar{C}_0 = -3(2.074 + j1.634) = 7.920\epsilon^{j218.233}.$$

210. Remember that the $\bar{V}_0$s are identical voltages in both magnitude and phase. Then note that just three closed loops exist in the circuit. Now note that the net generator voltage is ZERO around each of these loops. Thus *no zero-sequence current* can flow in such a circuit, regardless of whether the load is balanced or unbalanced.

211. Using the same notation as in section 10.10 we have, in exponential notation, using degrees,

$$\bar{A} = 90\epsilon^{j0} = 90 \quad \bar{B} = 72\epsilon^{j120} \quad \bar{C} = 54\epsilon^{j240}$$

and thus, by eqs. (460) through (462), we find that

$$\bar{A}_1 = \tfrac{1}{3}(90 + 72 + 54) = 72 \text{ volts} \quad (\epsilon^{j360} = 1)$$
$$\bar{B}_1 = 72\epsilon^{j120}$$
$$\bar{C}_1 = 72\epsilon^{j240}$$

these being the POSITIVE-SEQUENCE set of vectors. Next, by eqs. (463) through (465) we find that

$$\bar{A}_2 = 6(5 + 4\epsilon^{j240} + 3\epsilon^{j120}) \qquad (480^\circ = 120^\circ)$$
$$\bar{B}_2 = 6(5\epsilon^{j240} + 4\epsilon^{j120} + 3)$$
$$\bar{C}_2 = 6(5\epsilon^{j120} + 4 + 3\epsilon^{j240})$$

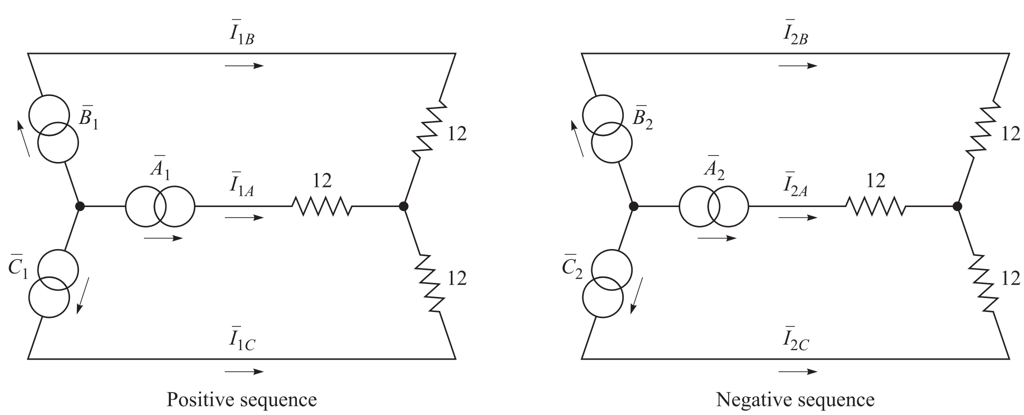
these being the NEGATIVE-SEQUENCE set of vectors. Lastly, by eq. (466),

$$\bar{A}_0 = \bar{B}_0 = \bar{C}_0 = 6(5 + 4\epsilon^{j120} + 3\epsilon^{j240}) \text{ volts,}$$

these being the ZERO-SEQUENCE set of vectors. However, since no zero-sequence current can flow in a balanced Y-connected load (problem 210), we'll not need to make any further reference to the zero-sequence voltage in this problem.

The procedure now is to SEPARATELY find *first* the value of the phase currents due to the POSITIVE-sequence voltages acting alone and *then* the phase currents due to the NEGATIVE-sequence voltages acting alone. The *total* phase current is then, by the principle of superposition, equal to the *sum* of the two currents found separately. The procedure is advantageous because the positive-sequence and negative-sequence sets are separately both *balanced* three-phase systems, which are individually easy to deal with. The procedure can be applied to the present problem as follows.

Let us imagine that the generator in Fig. 266 is composed of the sum of two separate generators, each generator working into the same common balanced load of 12 ohms per phase, as shown in the two figures below.



Since both figures represent completely balanced systems, it follows that each phase voltage appears separately as a voltage drop across each corresponding 12-ohm load. Hence, since "phase current" is the same as "line current" in this case, we have that, in the positive-sequence figure,

$$\bar{I}_{1A} = \bar{A}_1/12 \quad \bar{I}_{1B} = \bar{B}_1/12 \quad \bar{I}_{1C} = \bar{C}_1/12$$

and in the negative-sequence figure,

$$\bar{I}_{2A} = \bar{A}_2/12 \quad \bar{I}_{2B} = \bar{B}_2/12 \quad \bar{I}_{2C} = \bar{C}_2/12$$

and thus, by the principle of superposition, we have that the three line currents in Fig. 266 are equal to

$$\bar{I}_A = \bar{I}_{1A} + \bar{I}_{2A} = (\bar{A}_1 + \bar{A}_2)/12$$
$$\bar{I}_B = \bar{I}_{1B} + \bar{I}_{2B} = (\bar{B}_1 + \bar{B}_2)/12$$
$$\bar{I}_C = \bar{I}_{1C} + \bar{I}_{2C} = (\bar{C}_1 + \bar{C}_2)/12$$

Now substitute, into the above three equations, the values of the positive-sequence and negative-sequence voltages previously found. Doing this, the above three equations become

$$\bar{I}_A = (102 + 18\epsilon^{j120} + 24\epsilon^{j240})/12$$
$$\bar{I}_B = (18 + 96\epsilon^{j120} + 30\epsilon^{j240})/12$$
$$\bar{I}_C = (24 + 30\epsilon^{j120} + 90\epsilon^{j240})/12$$

Recall now that the value of an indicated *sum* of a number of complex numbers can be found only if the numbers are expressed in rectangular form, because

$$\text{(sum of complex numbers)} = \text{(sum of real parts)} + j\text{(sum of imaginary parts)}$$

and thus, upon converting the above exponential forms into their equivalent rectangular forms,* then adding, you should find that the above three equations become

$$\bar{I}_A = (6.750 - j0.433), \text{ hence } |\bar{I}_A| = 6.764 \text{ amperes, } \textit{answer}.$$
$$\bar{I}_B = (-3.750 + j4.763), \text{ hence } |\bar{I}_B| = 6.062 \text{ amperes, } \textit{answer}.$$
$$\bar{I}_C = -(3.000 + j4.330), \text{ hence } |\bar{I}_C| = 5.268 \text{ amperes, } \textit{answer}.$$

212. In Fig. 266, $\bar{A} = \bar{V}_{\text{na}} = 90\underline{/0^\circ}$, $\bar{B} = \bar{V}_{\text{nb}} = 72\underline{/120^\circ}$, $\bar{C} = \bar{V}_{\text{nc}} = 54\underline{/240^\circ}$, where $\bar{A} = \bar{V}_{\text{na}}$ is the reference vector. Then, using the same basic procedure as in section 10.8, we have (angles in degrees)

$$\bar{V}_{\text{ab}} = -\bar{B} + \bar{A} = -72\underline{/120} + 90$$
$$\bar{V}_{\text{bc}} = -\bar{C} + \bar{B} = -54\underline{/240} + 72\underline{/120}$$
$$\bar{V}_{\text{ca}} = -\bar{A} + \bar{C} = -90 + 54\underline{/240}$$

In order to find the values of the sums indicated above, the polar forms must first be converted to rectangular forms; to do this, note that

$$1\underline{/120} = \cos 120 + j \sin 120 = -0.5 + j0.8660$$
$$1\underline{/240} = \cos 240 + j \sin 240 = -0.5 - j0.8660$$

* Euler's formula, $A \in^{j\theta} = A(\cos\theta + j\sin\theta)$.

thus

$$\bar{V}_{ab} = |126 - j62.352| = 140.58 \text{ volts, } \textit{answer}.$$

$$\bar{V}_{bc} = |-9.000 + j109.12| = 109.49 \text{ volts, } \textit{answer}.$$

$$\bar{V}_{ca} = |-117 - j46.76| = 126.00 \text{ volts, } \textit{answer}.$$

213. The TOTAL power P is the sum of the powers produced in each of the 12-ohm loads. Hence, using the values of current found in problem 211, we have that

$$P = 12[(6.764)^2 + (6.062)^2 + (5.268)^2] = 1323 \text{ watts approx., } \textit{answer}.$$

214. The given values of the generator voltages, and the values of the generator currents $\bar{I}_A$, $\bar{I}_B$, and $\bar{I}_C$, were found in problem 211. Using these values in eq. (467) gives the following answers.

$$P_A = \text{r.p.}[90(6.750 + j0.433)] = 607.5 \text{ watts}$$

$$P_B = \text{r.p.}[72(-0.5 + j0.8660)(-3.750 - j4.763)]$$

$$= 72(1.875 + 4.125) = 432 \text{ watts}$$

$$P_C = \text{r.p.}[54(-0.5 - j0.8660)(-3.000 + j4.330)]$$

$$= 54(1.500 + 3.750) = 283.5 \text{ watts}$$

and thus the total power output of the generator is

$$P_T = P_A + P_B + P_C = 1323\,\text{W}, \text{as found in problem 213.}$$

215. First, a comparison of the notation used in eqs. (460) and (463) with the notation used in Fig. 267 shows that $\bar{A} = \bar{I}_A$ (thus $\bar{A}_1 = \bar{I}_{A1}$ and $\bar{A}_2 = \bar{I}_{A2}$), and also that $\bar{B} = \bar{I}_B$ and $\bar{C} = \bar{I}_C$: thus, by eqs. (460) and (463), we have that

$$\bar{I}_{A1} = \tfrac{1}{3}(\bar{I}_A + \bar{a}^2\bar{I}_B + \bar{a}\bar{I}_C)$$

and

$$\bar{I}_{A2} = \tfrac{1}{3}(\bar{I}_A + \bar{a}\bar{I}_B + \bar{a}^2\bar{I}_C)$$

Now substitute these values into eqs. (468) through (470); upon doing this, and noting that $\bar{a}^3 = 1$, $\bar{a}^4 = \bar{a}$, and also that $(\bar{a} + \bar{a}^2) = -1$, you should arrive at the following three simultaneous equations

$$2\bar{I}_A - \bar{I}_B - \bar{I}_C + 3\bar{Y}_A\bar{V}_N = 3\bar{Y}_A V$$

$$-\bar{I}_A + 2\bar{I}_B - \bar{I}_C + 3\bar{Y}_B\bar{V}_N = 3\bar{a}\bar{Y}_B V$$

$$-\bar{I}_A - \bar{I}_B + 2\bar{I}_C + 3\bar{Y}_C\bar{V}_N = 3\bar{a}^2\bar{Y}_C V$$

Note that we have four unknowns $(\bar{I}_A, \bar{I}_B, \bar{I}_C, \bar{V}_N)$ but only three equations; this situation can be overcome by making the substitution $\bar{I}_C = -\bar{I}_A - \bar{I}_B$, and upon doing this you should find that the above three equations can be written in terms of just three unknowns, $\bar{I}_A$, $\bar{I}_B$, $\bar{V}_N$, as follows:

$$3\bar{I}_A + 0\bar{I}_B + 3\bar{Y}_A\bar{V}_N = 3\bar{Y}_A V$$

$$0\bar{I}_A + 3\bar{I}_B + 3\bar{Y}_B\bar{V}_N = 3\bar{a}\bar{Y}_B V$$

$$-3\bar{I}_A - 3\bar{I}_B + 3\bar{Y}_C\bar{V}_N = 3\bar{a}^2\bar{Y}_C V$$

The above set of equations can be solved either by the process of elimination or by determinants; if we choose the more straightforward method of determinants the procedure is as follows:

First,

$$\bar{\Delta} = 27\begin{vmatrix} 1 & 0 & \bar{Y}_A \\ 0 & 1 & \bar{Y}_B \\ -1 & -1 & \bar{Y}_C \end{vmatrix} = 27(\bar{Y}_A + \bar{Y}_B + \bar{Y}_C)$$

The solution for $\bar{I}_A$ is then as follows.

$$\bar{I}_A = \frac{1}{\bar{\Delta}}\begin{vmatrix} 3\bar{Y}_A V & 0 & 3\bar{Y}_A \\ 3\bar{a}\bar{Y}_B V & 3 & 3\bar{Y}_B \\ 3\bar{a}^2\bar{Y}_C V & -3 & 3\bar{Y}_C \end{vmatrix} = \frac{27\bar{Y}_A V}{\bar{\Delta}}\begin{vmatrix} 1 & 0 & 1 \\ \bar{a}\bar{Y}_B & 1 & \bar{Y}_B \\ \bar{a}^2\bar{Y}_C & -1 & \bar{Y}_C \end{vmatrix}$$

It will be easiest, now, to expand the above determinant in terms of the elements of the second column; doing this, we have that

$$\bar{I}_A = \frac{27\bar{Y}_A V}{\bar{\Delta}}\left[\bar{Y}_C\begin{vmatrix} 1 & 1 \\ \bar{a}^2 & 1 \end{vmatrix} + \bar{Y}_B\begin{vmatrix} 1 & 1 \\ \bar{a} & 1 \end{vmatrix}\right] = \frac{27\bar{Y}_A V}{\bar{\Delta}}\,[\bar{Y}_C(1-\bar{a}^2) + \bar{Y}_B(1-\bar{a})]$$

Now, upon substituting in the value of $\bar{\Delta}$, and also noting that

$$(1-\bar{a}^2) = 1 - \cos 240 - j\sin 240 = 1.732\underline{/30^\circ} = 1.732\epsilon^{j30^\circ}$$

and

$$(1-\bar{a}) = 1 - \cos 120 - j\sin 120 = 1.732\underline{/-30^\circ} = 1.732\epsilon^{-j30^\circ}$$

you can verify that the final answer for $\bar{I}_A$ can be written in the form of eq. (471). Following the same procedure will likewise produce eq. (472) for the value of $\bar{I}_B$. Then, lastly, eq. (473) is most easily verified by making use of the relationship $\bar{I}_C = -\bar{I}_A - \bar{I}_B$.

216.

$$\bar{Y}_A = 1/\bar{Z}_A = 1/5\epsilon^{j53.130} = 0.2\epsilon^{-j53.130}\ \text{mhos}$$
$$\bar{Y}_B = 1/\bar{Z}_B = 1/8 = 0.125\ \text{mhos}$$
$$\bar{Y}_C = 1/\bar{Z}_C = 1/5 = 0.200\ \text{mhos}$$

from which we find that

$$\bar{Y}_A + \bar{Y}_B + \bar{Y}_C = (0.445 - j0.160) = 0.473\epsilon^{-j19.776}$$

Now substituting all of the above values and $V = 125$ into eqs. (471) through (473), you should find that

$$\bar{I}_A = 18.309\epsilon^{-j3.354} + 11.443\epsilon^{-j63.354} = 23.410 - j11.299$$
$$\bar{I}_B = 11.443\epsilon^{j109.776} - 11.443\epsilon^{-j63.354} = -9.004 + j20.996$$
$$\bar{I}_C = -18.309\epsilon^{-j3.354} - 11.443\epsilon^{j109.776} = -(14.406 + j9.697)$$

which, after applying the Pythagorean theorem in the usual manner, gives the required magnitudes of current.

217. In Fig. 267 the magnitude of the line voltage is 1.732 times the magnitude of the phase voltage. Hence, all we'd need to do, in all of the equations, is to replace V with V_{AB} and delete the 1.732 factors. Then, in problem 216, instead of $V = 125$ we would use the value $V_{AB} = 1.732(125) = 216.5$, thus producing the same current values as before.

218. 45 elements—5 rows—9 columns—row four—column six—square—6×5.

219. No; the sum $\mathbf{A} + \mathbf{B}$ can exist only if both have the same number of rows and the same number of columns.

220.

$$\begin{bmatrix} (3+0) & (2+2) & (-4+1) \\ (1+1) & (-7+2) & (5+4) \end{bmatrix} = \begin{bmatrix} 3 & 4 & -3 \\ 2 & -5 & 9 \end{bmatrix}, \text{ answer.}$$

221.

$$\begin{bmatrix} 0 & 4 \\ 3 & -2 \end{bmatrix} + (-1)\begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 3 & -2 \end{bmatrix} + \begin{bmatrix} -2 & 3 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -2 & 7 \\ 2 & -4 \end{bmatrix}, \text{ answer}$$

222.

$$\text{(a)} \quad \begin{bmatrix} (6+2) & (2-2) & (-1-3) \\ (4+4) & (0-3) & (0+0) \\ (3+1) & (0+1) & (5-6) \end{bmatrix} = \begin{bmatrix} 8 & 0 & -4 \\ 8 & -3 & 0 \\ 4 & 1 & -1 \end{bmatrix}, \text{ answer.}$$

$$\text{(b)} \quad \begin{bmatrix} (1+2-7) & (3-9+4) \\ (4+3-0) & (-6+6-10) \end{bmatrix} = \begin{bmatrix} -4 & -2 \\ 7 & -10 \end{bmatrix}, \text{ answer.}$$

223. Two matrices can be EQUAL only if (1) they have the same number of rows and the same number of columns and (2) all corresponding elements are equal. Hence, by inspection, $a = 6$, $b = -3$, $c = 4$, $e = 5$, $f = 3$, and $g = -6$.

224. This is a (2×2) matrix times a (2×1) matrix, so the product does exist in the order as shown. The product will be a (2×1) matrix, thus

$$\begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix} = \begin{bmatrix} (24+32) \\ (72-32) \end{bmatrix} = \begin{bmatrix} 56 \\ 40 \end{bmatrix}, \text{ answer.}$$

225. Since $\mathbf{A}$ is a (2×3) matrix and $\mathbf{B}$ a (3×2) matrix, the product in the order $\mathbf{AB}$ does exist. The result will be a (2×2) matrix, as follows:

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} (4+3+10) & (-2+4-4) \\ (6+0-20) & (-3+0+8) \end{bmatrix} = \begin{bmatrix} 17 & -2 \\ -14 & 5 \end{bmatrix}, \text{ answer}$$

226. First take the sum of the three (2×2) matrices inside the parentheses, as indicated. We then have the product of a (3×2) and a (2×2) matrix, which gives a (3×2) matrix; thus

$$\begin{bmatrix} 0 & 2 \\ 0 & 4 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} 7 & -4 \\ 6 & 11 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} (0+12) & (0+22) \\ (0+24) & (0+44) \\ (42+6) & (-24+11) \end{bmatrix} = \begin{bmatrix} 12 & 22 \\ 24 & 44 \\ 48 & -13 \end{bmatrix}, \ \textit{answer}$$

227. This is a matrix product of the form $\mathbf{ABC} = \mathbf{D}$. The procedure is to first find the product $\mathbf{AB}$, then take that result times $\mathbf{C}$. The work is as follows.

First,

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 8 & 14 \\ -16 & -18 \end{bmatrix},$$

a 2×2 matrix, which we now multiply the 2×3 matrix $\mathbf{C}$ by. The result is a 2×3 matrix $\mathbf{D}$; thus

$$\begin{bmatrix} 8 & 14 \\ -16 & -18 \end{bmatrix} \begin{bmatrix} -7 & 6 & 0 \\ -2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

$$= \begin{bmatrix} (-56-28) & (48+14) & (0+56) \\ (112+36) & (-96-18) & (0-72) \end{bmatrix}$$

$$= \begin{bmatrix} -84 & 62 & 56 \\ 148 & -114 & -72 \end{bmatrix}, \ \textit{answer}.$$

228. The first factor is a 4×4 matrix and the second factor is a 4×1 matrix, so the product of the two, in the order shown, does exist and will be a 4×1 matrix; thus

$$\begin{bmatrix} 1 & 2 & -2 & 3 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 7 & 2 \\ 4 & 2 & -5 & 10 \end{bmatrix} \begin{bmatrix} 6 \\ -2 \\ 0 \\ 9 \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ c_{31} \\ c_{41} \end{bmatrix} = \begin{bmatrix} (6-4+0+27) \\ (0+0+0+9) \\ (0+0+0+18 \\ (24-4+0+90) \end{bmatrix} = \begin{bmatrix} 29 \\ 9 \\ 18 \\ 110 \end{bmatrix}, \ \textit{answer}.$$

229. The square of any $(n \times n)$ square matrix is also an $(n \times n)$ square matrix; thus, the square of the given 3×3 matrix is a 3×3 square matrix, $\mathbf{A}^2 = \mathbf{AA} = \mathbf{C}$; thus

$$\begin{bmatrix} 1 & 1 & 4 \\ 0 & -3 & 2 \\ 9 & 6 & -5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 4 \\ 0 & -3 & 2 \\ 9 & 6 & -5 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

$$= \begin{bmatrix} (1+0+36) & (1-3+24) & (4+2-20) \\ (0+0+18) & (0+9+12) & (0-6-10) \\ (9+0-45) & (9-18-30) & (36+12+25) \end{bmatrix}$$

$$= \begin{bmatrix} 37 & 22 & -14 \\ 18 & 21 & -16 \\ -36 & -39 & 73 \end{bmatrix}, \ \textit{answer}.$$

230. Since **A** is a (3×4) matrix and **B** is a (4×2) matrix, their product **AB** does exist and is a (3×2) matrix, $\mathbf{C} = \mathbf{AB}$, whose value is found as follows.

$$\begin{bmatrix} 2 & 0 & 1 & -3 \\ 1 & 2 & 0 & 4 \\ 3 & 2 & -6 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & 0 \\ 2 & 5 \\ -6 & 1 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

$$= \begin{bmatrix} (2+0+2+18) & (-4+0+5-3) \\ (1+6+0-24) & (-2+0+0+4) \\ (3+6-12-6) & (-6+0-30+1) \end{bmatrix}$$

$$= \begin{bmatrix} 22 & -2 \\ -17 & 2 \\ -9 & -35 \end{bmatrix}, \text{ answer.}$$

231. The first step is to find Δ, the value of the third-order determinant formed from the elements of the given matrix. Using, for example, the elements of the second row, we find that

$$\Delta = -5\begin{vmatrix} 0 & 4 \\ -1 & 2 \end{vmatrix} + 6\begin{vmatrix} 2 & 4 \\ -3 & 2 \end{vmatrix} = 76$$

Next, in the manner of eq. (483), we replace each element in the given matrix with its cofactor, to get

$$\mathbf{A}_0 = \begin{bmatrix} \begin{vmatrix} 6 & 0 \\ -1 & 2 \end{vmatrix} & -\begin{vmatrix} 5 & 0 \\ -3 & 2 \end{vmatrix} & \begin{vmatrix} 5 & 6 \\ -3 & -1 \end{vmatrix} \\ -\begin{vmatrix} 0 & 4 \\ -1 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ -3 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 0 \\ -3 & -1 \end{vmatrix} \\ \begin{vmatrix} 0 & 4 \\ 6 & 0 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 5 & 0 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 5 & 6 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 12 & -10 & 13 \\ -4 & 16 & 2 \\ -24 & 20 & 12 \end{bmatrix}$$

Now, in the last expression, interchange the rows and columns; that is, let the first row become the first column, the second row become the second column, and so on. Doing this, and remembering to multiply by $1/\Delta$, the final *answer* can be written in either of the forms

$$\mathbf{A}^{-1} = \frac{1}{76}\begin{bmatrix} 12 & -4 & -24 \\ -10 & 16 & 20 \\ 13 & 2 & 12 \end{bmatrix} = \begin{bmatrix} 12/76 & -4/76 & -24/76 \\ -10/76 & 16/76 & 20/76 \\ 13/76 & 2/76 & 12/76 \end{bmatrix}$$

the second form being in accordance with the rule for multiplication of a matrix by a constant, as laid down at the end of section 11.1.

232. First, $\Delta = (-20 + 21) = 1$. Now replace each element in the given matrix with its cofactor to get $\mathbf{A}_0$; thus

$$\mathbf{A}_0 = \begin{bmatrix} -5 & -3 \\ 7 & 4 \end{bmatrix}$$

which, after interchanging rows and columns and then multiplying by $1/\Delta$ (which in this case is $1/1 = 1$), we have

$$\mathbf{A}^{-1} = \begin{bmatrix} -5 & 7 \\ -3 & 4 \end{bmatrix}, \textit{ answer}.$$

233. First,

$$\Delta = -6\begin{vmatrix} 2 & 4 & -2 \\ 0 & 2 & 1 \\ 0 & -5 & 3 \end{vmatrix} = (-6)(2)\begin{vmatrix} 2 & 1 \\ -5 & 3 \end{vmatrix} = -132$$

Next, in the manner of eq. (483), the value of $\mathbf{A}_0$ is

$$A_0 = \begin{bmatrix} \begin{vmatrix} 0 & 6 & 0 \\ 2 & 0 & 1 \\ -5 & 0 & 3 \end{vmatrix} & -\begin{vmatrix} 3 & 6 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -5 & 3 \end{vmatrix} & -\begin{vmatrix} 3 & 0 & 6 \\ 0 & 2 & 0 \\ 0 & -5 & 0 \end{vmatrix} \\ -\begin{vmatrix} 4 & 0 & -2 \\ 2 & 0 & 1 \\ -5 & 0 & 3 \end{vmatrix} & \begin{vmatrix} 2 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{vmatrix} & -\begin{vmatrix} 2 & 4 & -2 \\ 0 & 2 & 1 \\ 0 & -5 & 3 \end{vmatrix} & \begin{vmatrix} 2 & 4 & 0 \\ 0 & 2 & 0 \\ 0 & -5 & 0 \end{vmatrix} \\ \begin{vmatrix} 4 & 0 & -2 \\ 0 & 6 & 0 \\ -5 & 0 & 3 \end{vmatrix} & -\begin{vmatrix} 2 & 0 & -2 \\ 3 & 6 & 0 \\ 0 & 0 & 3 \end{vmatrix} & \begin{vmatrix} 2 & 4 & -2 \\ 3 & 0 & 0 \\ 0 & -5 & 3 \end{vmatrix} & -\begin{vmatrix} 2 & 4 & 0 \\ 3 & 0 & 6 \\ 0 & -5 & 0 \end{vmatrix} \\ -\begin{vmatrix} 4 & 0 & -2 \\ 0 & 6 & 0 \\ 2 & 0 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 0 & -2 \\ 3 & 6 & 0 \\ 0 & 0 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 4 & -2 \\ 3 & 0 & 0 \\ 0 & 2 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 4 & 0 \\ 3 & 0 & 6 \\ 0 & 2 & 0 \end{vmatrix} \end{bmatrix}$$

Note that, in this particular example, the values of the determinants are all easy to find because each can be expanded in terms of a row or column in which all the elements except one are zero. (Note that three of the determinants have the value zero by inspection, since, if the elements of any row or column are all equal to zero, the value of the determinant is zero.) You can now verify that the above matrix reduces to the form

$$A_0 = \begin{bmatrix} -6\begin{vmatrix} 2 & 1 \\ -5 & 3 \end{vmatrix} & -3\begin{vmatrix} 0 & 1 \\ 0 & 3 \end{vmatrix} & 3\begin{vmatrix} 2 & 1 \\ -5 & 3 \end{vmatrix} & -3\begin{vmatrix} 2 & 0 \\ -5 & 0 \end{vmatrix} \\ 0 & 0 & -2\begin{vmatrix} 2 & 1 \\ -5 & 3 \end{vmatrix} & 0 \\ 6\begin{vmatrix} 2 & -2 \\ -5 & 3 \end{vmatrix} & -6\begin{vmatrix} 2 & -2 \\ 0 & 3 \end{vmatrix} & -3\begin{vmatrix} 4 & -2 \\ -5 & 3 \end{vmatrix} & 6\begin{vmatrix} 2 & 4 \\ 0 & -5 \end{vmatrix} \\ -6\begin{vmatrix} 4 & -2 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 3 & 6 \end{vmatrix} & 3\begin{vmatrix} 4 & -2 \\ 2 & 1 \end{vmatrix} & -2\begin{vmatrix} 2 & 0 \\ 3 & 6 \end{vmatrix} \end{bmatrix}$$

thus

$$\mathbf{A}_0 = \begin{bmatrix} -66 & 0 & 33 & 0 \\ 0 & 0 & -22 & 0 \\ 12 & -36 & -6 & -60 \\ -48 & 12 & 24 & -24 \end{bmatrix}$$

Now, in the last expression, interchange the rows and columns. Doing this, and remembering to multiply the result by $1/\Delta$, we have that

$$\mathbf{A}^{-1} = -\frac{1}{132}\begin{bmatrix} -66 & 0 & 12 & -48 \\ 0 & 0 & -36 & 12 \\ 33 & -22 & -6 & 24 \\ 0 & 0 & -60 & -24 \end{bmatrix}, \textit{ answer}.$$

234. First (we'll expand in terms of the elements of the second row) we have

$$\Delta = 2\begin{vmatrix} 8 & -1 \\ -6 & 3 \end{vmatrix} - 3\begin{vmatrix} 8 & 0 \\ -6 & 4 \end{vmatrix} = -60$$

Then

$$A_0 = \begin{bmatrix} \begin{vmatrix} 2 & 3 \\ 4 & 3 \end{vmatrix} & -\begin{vmatrix} 0 & 3 \\ -6 & 3 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ -6 & 4 \end{vmatrix} \\ -\begin{vmatrix} 0 & -1 \\ 4 & 3 \end{vmatrix} & \begin{vmatrix} 8 & -1 \\ -6 & 3 \end{vmatrix} & -\begin{vmatrix} 8 & 0 \\ -6 & 4 \end{vmatrix} \\ \begin{vmatrix} 0 & -1 \\ 2 & 3 \end{vmatrix} & -\begin{vmatrix} 8 & -1 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} 8 & 0 \\ 0 & 2 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -6 & -18 & 12 \\ -4 & 18 & -32 \\ 2 & -24 & 16 \end{bmatrix}$$

therefore,

$$\mathbf{A}^{-1} = -\frac{1}{60}\begin{bmatrix} -6 & -4 & 2 \\ -18 & 18 & -24 \\ 12 & -32 & 16 \end{bmatrix} = \begin{bmatrix} 1/10 & 1/15 & -1/30 \\ 3/10 & -3/10 & 4/10 \\ -1/5 & 8/15 & -4/15 \end{bmatrix}$$

either way is a correct *answer*.

235. First, in the manner of eq. (476), write the first set of equations in the matrix form

$$\begin{bmatrix} 3 & -4 & 1 \\ -2 & 1 & -5 \\ 4 & 6 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix} \tag{A}$$

In the above, now verify that $\Delta = 164$. Then, in the manner of eq. (480), eq. (A) can be written as follows, where the superscript "-1" indicates that the inverse of the matrix is to be taken

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & -4 & 1 \\ -2 & 1 & -5 \\ 4 & 6 & -2 \end{bmatrix}^{-1} \begin{bmatrix} r \\ s \\ t \end{bmatrix} \tag{B}$$

To find the indicated inverse of the matrix, let us next find the value of $\mathbf{A}_0$ which is, in this case, using eq. (483),

$$\mathbf{A}_0 = \begin{bmatrix} 28 & -24 & -16 \\ -2 & -10 & -34 \\ 19 & 13 & -5 \end{bmatrix}$$

and thus

$$\begin{bmatrix} 3 & -4 & 1 \\ -2 & 1 & -5 \\ 4 & 6 & -2 \end{bmatrix}^{-1} = \frac{1}{164} \begin{bmatrix} 28 & -2 & 19 \\ -24 & -10 & 13 \\ -16 & -34 & -5 \end{bmatrix}$$

therefore, eq. (B) becomes

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{164} \begin{bmatrix} 28 & -2 & 19 \\ -24 & -10 & 13 \\ -16 & -34 & -5 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix}$$

Now take the product of the two matrices on the right-hand side, as indicated; the right-hand side then becomes a 3×1 matrix; thus

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{164} \begin{bmatrix} (28r - 2s + 19t) \\ (-24r - 10s + 13t) \\ (-16r - 34s - 5t) \end{bmatrix}$$

Now carry out the indicated multiplication of the right-hand side by (1/164) (section 11.1). We then have the equality of two 3×1 matrices, and therefore, by the definition of equal matrices from section 11.1, the last matrix equation is the equivalent of the following three simultaneous equations

$$\begin{aligned} (28/164)r - (2/164)s + (19/164)t &= x \\ -(24/164)r - (10/164)s + (13/164)t &= y \\ -(16/164)r - (34/164)s - (5/164)t &= z \end{aligned}$$

Now compare the coefficients in the last three equations with the corresponding ones in set 2; doing this, you'll find that $a = (28/164)$, $b = -(2/164)$, $c = (19/164)$, and so on, to the final value, $i = -(5/164)$, *answers*.

236.

$$\mathbf{AA}^{-1} = -\frac{1}{37}\begin{bmatrix} 2 & 1 & -4 \\ 1 & 5 & 3 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 5 & -1 & 23 \\ -7 & -6 & -10 \\ 10 & -2 & 9 \end{bmatrix}$$

$$= -\frac{1}{37}\begin{bmatrix} (10-7-40) & (-2-6+8) & (46-10-36) \\ (5-35+30) & (-1-30-6) & (23-50+27) \\ (-10+0+10) & (2+0-2) & (-46+0+9) \end{bmatrix}$$

$$= -\frac{1}{37}\begin{bmatrix} -37 & 0 & 0 \\ 0 & -37 & 0 \\ 0 & 0 & -37 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

which indicates that the work *is* correct.

237.

$$\text{(a)} \quad \mathbf{A}_t = \begin{bmatrix} 4 & 3 \\ -4 & 2 \end{bmatrix}, \ \textit{answer}.$$

$$\text{(b)} \quad \mathbf{A}_t = \begin{bmatrix} 6 & 0 & 4 & 0 \\ 9 & 2 & 5 & 7 \\ -1 & 1 & 9 & 0 \\ 4 & 3 & -2 & 8 \end{bmatrix}, \ \textit{answer}.$$

238. First

$$[\mathbf{A}+\mathbf{B}]_t = \begin{bmatrix} (3+5) & (-3+2) \\ (-7+4) & (6-3) \end{bmatrix}_t = \begin{bmatrix} 8 & -1 \\ -3 & 3 \end{bmatrix}_t = \begin{bmatrix} 8 & -3 \\ -1 & 3 \end{bmatrix}$$

then

$$\mathbf{A}_t + \mathbf{B}_t = \begin{bmatrix} 3 & -7 \\ -3 & 6 \end{bmatrix} + \begin{bmatrix} 5 & 4 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} 8 & -3 \\ -1 & 3 \end{bmatrix}$$

which is the same answer as above.

239. Since **A** is a (3×2) and **B** a (2×3) matrix, the product in the order **AB** does exist and will be a (3×3) matrix. From the definition of matrix multiplication laid down in section 11.2, we have that

$$\mathbf{AB} = \begin{bmatrix} 2 & -4 \\ 0 & 6 \\ -7 & 3 \end{bmatrix}\begin{bmatrix} 3 & 0 & 3 \\ 2 & -1 & 9 \end{bmatrix} = \begin{bmatrix} (6-8) & (0+4) & (6-36) \\ (0+12) & (0-6) & (0+54) \\ (-21+6) & (0-3) & (-21+27) \end{bmatrix}$$

thus

$$\mathbf{AB} = \begin{bmatrix} -2 & 4 & -30 \\ 12 & -6 & 54 \\ -15 & -3 & 6 \end{bmatrix}$$

and hence

$$(\mathbf{AB})_t = \begin{bmatrix} -2 & 12 & -15 \\ 4 & -6 & -3 \\ -30 & 54 & 6 \end{bmatrix}$$

then, next,

$$\mathbf{B}_t\mathbf{A}_t = \begin{bmatrix} 3 & 2 \\ 0 & -1 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} 2 & 0 & -7 \\ -4 & 6 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} (6-8) & (0+12) & (-21+6) \\ (0+4) & (0-6) & (0-3) \\ (6-36) & (0+54) & (-21+27) \end{bmatrix} = \begin{bmatrix} -2 & 12 & -15 \\ 4 & -6 & -3 \\ -30 & 54 & 6 \end{bmatrix}$$

thus proving that $(\mathbf{AB})_t = \mathbf{B}_t\mathbf{A}_t$ for the given matrix.

240. g_{11} has the dimension of admittance (mhos).
g_{22} has the dimension of impedance (ohms).
g_{12} and g_{21} are dimensionless ratios.

a_{12} has dimensions of impedance (ohms).
a_{21} has dimensions of admittance (mhos).
a_{11} and a_{22} are dimensionless ratios.

241.

$$y_{22} = \frac{I_2}{V_2}_{(V_1=0)} \qquad g_{21} = \frac{V_2}{V_1}_{(I_2=0)}$$

242. Let us begin by writing eqs. (512) and (513) with the minus sign inside the **a** matrix, so that eq. (514) takes the equivalent form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & -a_{22} \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

Now multiply both sides of the above equation by the *inverse* of the **a** matrix, which gives

$$\begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & -a_{22} \end{bmatrix}^{-1} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & -a_{22} \end{bmatrix}^{-1} \begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & -a_{22} \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \qquad \text{(A)}$$

Note that the right-hand side of the equation is of the form

$$[\mathbf{a}]^{-1}[\mathbf{a}] \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \mathbf{I} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

by eqs. (489), (486), where **I** is the unit matrix of section 11.4. Therefore the preceding equation (eq. A) becomes

$$\begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & -a_{22} \end{bmatrix}^{-1} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

which *is* a correct way of writing the answer.

A more detailed answer, can, however, be written by taking the inverse of the 2×2 matrix as indicated. Following the procedure of problem 232, you should find that the inverse of the above 2×2 matrix is

$$\begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & -a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{12}a_{21} - a_{11}a_{22}} \begin{bmatrix} -a_{22} & a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

and therefore the preferable *final answer*, free of inverse notation, is

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \frac{1}{a_{12}a_{21} - a_{11}a_{22}} \begin{bmatrix} -a_{22} & a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

243. Multiplication of both sides of eq. (482) by the constant c gives

$$c\mathbf{A} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nn} \end{bmatrix}$$

We now wish to take the inverse of both sides of the above equation. The first step in doing this is to find the value of the *determinant* formed from the elements of the matrix. Note that, when regarded as a determinant, c factors from every row of the determinant, and thus we have

$$\text{determinant of } cA = c^n \Delta$$

where Δ is the determinant value of the original square matrix $\mathbf{A}$. Next, note that c^{n-1} will factor from every cofactor in $\mathbf{A}_0$, eq. (483), and therefore from the transpose of $\mathbf{A}_0$. It thus follows that

$$[c\mathbf{A}]^{-1} = \frac{c^{n-1}}{c^n \Delta} \begin{bmatrix} \text{transpose} \\ \text{of} \\ \text{matrix } \mathbf{A}_0 \end{bmatrix} = \frac{1}{c}\,\mathbf{A}^{-1}$$

244. As mentioned following eq. (504), $[\mathbf{z}] = [\mathbf{y}]^{-1}$, so our problem is basically *to find the inverse of the* 2×2 *admittance matrix*. Following the procedure of problem 232, you should find that

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}^{-1} = \frac{1}{y_{11}y_{22} - y_{12}y_{21}} \begin{bmatrix} y_{22} & -y_{12} \\ -y_{21} & y_{11} \end{bmatrix} = \begin{bmatrix} y_{22}/\Delta & -y_{12}/\Delta \\ -y_{21}/\Delta & y_{11}/\Delta \end{bmatrix}$$

where $\Delta = y_{11}y_{22} - y_{12}y_{21}$. The $\mathbf{z}$-matrix is equal to the last matrix to the right above, and thus, from the definition of equal matrices in section 11.1, the *answers* to the problem are

$$z_{11} = y_{22}/\Delta, \quad z_{12} = -y_{12}/\Delta, \quad z_{21} = -y_{21}/\Delta, \quad z_{22} = y_{11}/\Delta$$

245. First, remembering that $j^2 = -1$, you should find that $\Delta = 20(4 - j)10^{-6}$. Then $1/\Delta = 2.941(4 + j)10^3$ approximately, and using this value with the other given values, in the equations found in problem 244, gives the following values in ohms:

$$z_{11} = -17.65 + j70.59 \qquad z_{12} = 20.59 + j17.65$$

$$z_{21} = -735.3 + j941.1 \qquad z_{22} = 191.2 + j235.3$$

246. Going in the ccw sense around the two secondary circuits, the sums of the voltage drops are (letting $Z_2 = R + j\omega L$)

$$\text{for Fig. 281:} \qquad j\omega M I_1 + Z_2 I_2 = 0$$

$$\text{for Fig. 282:} \qquad -j\omega M I_1 + Z_2 I_2 = 0$$

hence, by eq. (517) for Fig. 281,

$$h_{21} = I_2/I_1 = -j\omega M/Z_2), \ \textit{answer}.$$

then, by eq. (517), for Fig. 282,

$$h_{21} = I_2/I_1 = j(\omega M/Z_2), \ \textit{answer}.$$

The results simply show that, for the particular network of Figs. 281 and 282, the sign of h_{21} depends upon the sense in which the secondary turns are wound relative to the primary turns.

247. First verify that $dh = 0.1320$. Next recall, from section 11.1, that two matrices can be equal only if *all corresponding elements are equal*. With this in mind, inspection of the fourth row of the conversion chart then gives the approximate *answers*

$$g_{11} = h_{22}/dh = 0.0030\,\text{mhos} \qquad g_{12} = -h_{12}/dh = -0.0606$$

$$g_{21} = -h_{21}/dh = -197.0 \qquad g_{22} = h_{11}/dh = 6439.4 \text{ ohms}$$

248. One procedure is as follows. First solve eq. (507) for V_2, then substitute that value of V_2 into eq. (506), which then becomes

$$V_1 = (h_{11} - h_{12}h_{21}/h_{22})I_1 + (h_{12}/h_{22})I_2$$

which, since $(h_{11} - h_{12}h_{21}/h_{22}) = (h_{11}h_{22} - h_{12}h_{21})/h_{22} = dh/h_{22}$, becomes

$$V_1 = \frac{dh}{h_{22}} I_1 + \frac{h_{12}}{h_{22}} I_2$$

Next, by eq. (507),

$$V_2 = -\frac{h_{21}}{h_{22}} I_1 + \frac{1}{h_{22}} I_2$$

Or, in matrix form, the last two simultaneous equations become

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} dh/h_{22} & h_{12}/h_{22} \\ -h_{21}/h_{22} & 1/h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

Comparison of the above equation with eq. (519) shows it has to be true that

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} dh/h_{22} & h_{12}/h_{22} \\ -h_{21}/h_{22} & 1/h_{22} \end{bmatrix}$$

thus proving that the relationship given in the conversion table *is correct*.

249. From the first row of the conversion chart we have that

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} 1/g_{11} & -g_{12}/g_{11} \\ g_{21}/g_{11} & dg/g_{11} \end{bmatrix}$$

thus

$$z_{11} = 1/g_{11} = 14.71 \text{ ohms} \qquad z_{12} = -g_{12}/g_{11} = 1.074 \text{ ohms}$$

$$z_{21} = g_{21}/g_{11} = -3353 \text{ ohms} \qquad z_{22} = dg/g_{11} = 8510 \text{ ohms}$$

250. By eq. (511),

$$[\mathbf{g}]\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$$

which can be written

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = [\mathbf{g}]^{-1}\begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$$

because $[\mathbf{g}]$ is a square matrix. Comparison of the last equation with eq. (508) shows that

$$[\mathbf{h}] = [\mathbf{g}]^{-1}$$

as proposed.

251. First we have

$$dh = h_{11}h_{22} - h_{12}h_{21} = 0.1320$$

Then, from the matrix conversion chart, the *answers* are

$$z_{11} = dh/h_{22} = 330 \text{ ohms} \qquad z_{12} = h_{12}/h_{22} = 20 \text{ ohms}$$

$$z_{21} = -h_{21}/h_{22} = -65{,}000 \text{ ohms} \qquad z_{22} = 1/h_{22} = 2500 \text{ ohms}$$

252. First, from inspection of Fig. 286, note that (where "e" refers to the single equivalent two-port)

$$V_{1e} = V_{1a} + V_{1b} \qquad V_{2e} = V_{2a} = V_{2b}$$

$$I_{1e} = I_{1a} = I_{1b} \qquad I_{2e} = I_{2a} + I_{2b}$$

Next, from eq. (521) we have
for two-port a:

$$\begin{bmatrix} V_{1a} \\ I_{2a} \end{bmatrix} = \begin{bmatrix} h_{11a} & h_{12a} \\ h_{21a} & h_{22a} \end{bmatrix}\begin{bmatrix} I_{1a} \\ I_{2a} \end{bmatrix} \qquad \text{(x)}$$

for two-port b:

$$\begin{bmatrix} V_{1b} \\ I_{2b} \end{bmatrix} = \begin{bmatrix} h_{11b} & h_{12b} \\ h_{21b} & h_{22b} \end{bmatrix}\begin{bmatrix} I_{1b} \\ V_{2b} \end{bmatrix}$$

$$= \begin{bmatrix} h_{11b} & h_{12b} \\ h_{21b} & h_{22b} \end{bmatrix}\begin{bmatrix} I_{1a} \\ V_{2a} \end{bmatrix} \qquad \text{(y)}$$

Now note that the sum of eqs. (x) and (y) can be written as

$$\begin{bmatrix} V_{1e} \\ I_{2e} \end{bmatrix} = \begin{bmatrix} (h_{11a} + h_{11b}) & (h_{12a} + h_{12b}) \\ (h_{21a} + h_{21b}) & (h_{22a} + h_{22b}) \end{bmatrix}\begin{bmatrix} I_{1e} \\ V_{2e} \end{bmatrix}$$

which, upon reference to eq. (521), shows that the h-parameters of the single equivalent two-port are equal to the *sum of the h-parameters of the individual two-ports* when the two-ports are connected in the series-parallel mode of Fig. 286.

253. Making use of eq. (536) and the conversion chart in section 11.8 we have, for three identical two-ports in parallel,

$$\begin{bmatrix} 3y_{11} & 3y_{12} \\ 3y_{21} & 3y_{22} \end{bmatrix} = 3\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} 3z_{22}/dz & -3z_{12}/dz \\ -3z_{21}/dz & 3z_{11}/dz \end{bmatrix}$$

and thus the *answers* are that parameters of the equivalent two-port, in terms of the z-parameters of the individual two-ports, are

$$z_{11\mathrm{e}} = 3z_{22}/dz, \quad z_{12\mathrm{e}} = -3z_{12}/dz, \quad z_{21\mathrm{e}} = -3z_{21}/dz, \quad z_{22\mathrm{e}} = 3z_{11}/dz$$

254. Since we're dealing with two identical two-ports, eq. (541) becomes

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} V_\mathrm{o} \\ I_\mathrm{o} \end{bmatrix}$$

which, upon using the conversion chart in section 11.8, becomes

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \frac{1}{h_{21}^2}\begin{bmatrix} dh & h_{11} \\ h_{22} & 1 \end{bmatrix}\begin{bmatrix} dh & h_{11} \\ h_{22} & 1 \end{bmatrix}\begin{bmatrix} V_\mathrm{o} \\ I_\mathrm{o} \end{bmatrix}$$

Now make use of the procedure for matrix multiplication defined in section 11.2; doing this gives the *final answer*

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} (d^2h + h_{11}h_{22}) & (dhh_{11} + h_{11}) \\ (dhh_{22} + h_{22}) & (h_{11}h_{22} + 1) \end{bmatrix}\begin{bmatrix} V_\mathrm{o} \\ I_\mathrm{o} \end{bmatrix}$$

(From Fig. 288, note that, for two two-ports in cascade, $V_\mathrm{o} = V_4 = V_5$ and $I_\mathrm{o} = I_5$.)

255. Note that the answer to problem 254 is of the form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} V_\mathrm{o} \\ I_\mathrm{o} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix}$$

and therefore

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} AV_\mathrm{o} + BI_\mathrm{o} \\ CV_\mathrm{o} + DI_\mathrm{o} \end{bmatrix}$$

which says that

$$V_1 = AV_\mathrm{o} + BI_\mathrm{o} = (A + B/Z_\mathrm{L})V_\mathrm{o}$$

and

$$I_1 = CV_\mathrm{o} + DI_\mathrm{o} = (C + D/Z_\mathrm{L})V_\mathrm{o}$$

because, from Fig. 288, $I_\mathrm{o} = I_7 = V_\mathrm{o}/Z_\mathrm{L}$. Now solve the first of the two simultaneous equations for V_o, then substitute the result in place of V_o in the second equation. Doing this gives us

$$I_1 = \frac{(C + D/Z_\mathrm{L})V_1}{A + B/Z_\mathrm{L}} = \frac{(D + CZ_\mathrm{L})V_1}{B + AZ_\mathrm{L}}$$

Now find the values of A, B, C, D, by making use of, from problem 254, the matrix equation,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} (d^2h + h_{11}h_{22}) & (dhh_{11} + h_{11}) \\ (dhh_{22} + h_{22}) & (h_{11}h_{22} + 1) \end{bmatrix}$$

256. As noted in the solution to problem 255, the network matrix in the answer to problem 254 is of the form

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

the inverse of which is (see problem 232 if you wish)

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \frac{1}{AD - BC}\begin{bmatrix} D & -B \\ -C & A \end{bmatrix}$$

Now substitute, into the last expression, the values of A, B, C, D, found from inspection of the matrix equality that appears in the solution to problem 255. Doing this gives, after some simplification, the *answer*

$$[\mathbf{h}]^{-1} = \frac{1}{(dh - h_{11}h_{22})^2}\begin{bmatrix} (1 + h_{11}h_{22}) & -(1 + dh)h_{11} \\ -(1 + dh)h_{22} & (d^2h + h_{11}h_{22} \end{bmatrix}$$

257. Note that eq. (545) is of the form

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} (AI_1 + BI_2) \\ (CI_1 + DI_2) \end{bmatrix}$$

which, by the definition of equal matrices, shows that

$$AI_1 + BI_2 = V_1$$
$$CI_1 + DI_2 = V_2$$

Thus

$$I_1 = \frac{\begin{vmatrix} V_1 & B \\ V_2 & D \end{vmatrix}}{AD - BC} = \frac{DV_1 - BV_2}{AD - BC}$$

which, upon substituting in the values of $A, B, C,$ and D, gives the required answer.

258. In eq. (545) we must first write the transistor z-parameters $(Z_{11}, Z_{12}, Z_{21}, Z_{22})$ in terms of h-parameters, which is most easily done by making use of the conversion chart in section 11.8. Doing this, eq. (545) becomes

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \left(\dfrac{dh}{h_{22}} + Z\right) & \left(\dfrac{h_{12}}{h_{22}} + Z\right) \\ \left(-\dfrac{h_{21}}{h_{22}} + Z\right) & \left(\dfrac{1}{h_{22}} + Z\right) \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix}, \ \textit{answer}.$$

259. *No*, because $[\mathbf{Z}]$ in eq. (543) is a "singular" matrix. (See discussion just prior to eq. (484) in section 11.3.)

260. First write eq. (548) in the inverse form

$$\begin{bmatrix} V_o \\ I_o \end{bmatrix} = \begin{bmatrix} 1 & Z' \\ 1/Z & (1 + Z'/Z) \end{bmatrix}^{-1} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

in which (as in the solution to problem 256) $A = 1$, $B = Z'$, $C = 1/Z$ and $D = (1 + Z'/Z)$; thus, using the special formula noted in the solution to problem 256, we find that

$$\begin{bmatrix} V_o \\ I_o \end{bmatrix} = \begin{bmatrix} (1 + Z'/Z) & -Z' \\ -1/Z & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}, \textit{ answer}.$$

261. We can make use of eqs. (506) and (507) in section 11.6 as follows. First, from Fig. 307, note that

$$V_2 = I_L Z_L \quad \text{and} \quad I_2 = -I_L$$

and thus eqs. (506) and (507) become

$$V_1 = h_{11} I_1 + h_{12} Z_L I_L$$
$$0 = h_{21} I_1 + (1 + h_{22} Z_L) I_L$$

We must now solve the foregoing two simultaneous equations for I_L. This can be done by using either determinants or the method of elimination. Thus, if you multiply the first equation by h_{21} and the second equation by $-h_{11}$, then add the two equations together, you should find that

$$I_L = \frac{h_{21} V_1}{h_{12} h_{21} Z_L - h_{11}(1 + h_{22} Z_L)}$$

which, upon substituting in the given h-values, $Z_L = 150$ ohms, $V_1 = 12$ volts, should give you the required answer.

262. Since $V_2 = R_L I_L$ and $-I_2 = I_L$, eq. (514) becomes

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} R_L I_L \\ I_L \end{bmatrix}$$

thus

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} (a_{11} R_L I_L + a_{12} I_L) \\ (a_{21} R_L I_L + a_{22} I_L) \end{bmatrix}$$

hence, by the definition of equal matrices (section 11.1),

$$V_1 = (a_{11} R_L + a_{12}) I_L$$

thus

$$I_L = \frac{V_1}{a_{11} R_L + a_{12}}$$

We must now make use of the conversion chart in section 11.8 to express the required a-parameters in terms of the given h-parameters. Carefully doing this, you should find that $a_{11} = 0.14$ and $a_{12} = 20$, and now, substituting these a values into the last equation above, along with $V_1 = 12$ and $R_L = 150$, you should find that $I_L = 0.2927$ amperes, as before.

263. Here we have a cascade connection of three two-ports in the manner of Fig. 288 and eq. (541) in section 11.9, in which, for Fig. 308, in terms of a coefficients,

$$[\mathbf{a}_1] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \text{ for first transistor}$$

$$[\mathbf{a}_2] = \begin{bmatrix} 1 & 0 \\ 1/R & 1 \end{bmatrix}, \text{ shunt impedance of } Z = R \text{ ohms in terms of } a \text{ coefficients*}$$

$$[\mathbf{a}_3] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \text{ for second transistor (same values as for first transistor)}$$

Thus we now have, for Fig. 308, in the manner of eq. (540) ($I_7 = I_\mathrm{L}$, $V_7 = R_\mathrm{L} I_\mathrm{L}$),

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1/R & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} R_\mathrm{L} I_\mathrm{L} \\ I_\mathrm{L} \end{bmatrix}$$

We must now take the first matrix times the second matrix, then take that result times the third matrix, then that result times the fourth matrix. Upon doing this, carefully following the rule for matrix multiplication from section 11.2, you should find that the result is the matrix expression

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} [(a_{11} + a_{12}/R)a_{11} + a_{12}a_{21}]R_\mathrm{L} I_\mathrm{L} + [(a_{11} + a_{12}/R)a_{12} + a_{12}a_{22}]I_\mathrm{L} \\ [(a_{21} + a_{22}/R)a_{11} + a_{21}a_{22}]R_\mathrm{L} I_\mathrm{L} + [(a_{21} + a_{22}/R)a_{12} + a_{22}^2]I_\mathrm{L} \end{bmatrix}$$

Now, using the same procedure as in the solution to problem 262 (noting that $R_\mathrm{L} I_\mathrm{L} = V_\mathrm{L}$), you can verify that

$$V_\mathrm{L} = \frac{V_1 R_\mathrm{L}}{[(a_{11} + a_{12}/R)a_{11} + a_{12}a_{21}]R_\mathrm{L} + (a_{11} + a_{22} + a_{12}/R)a_{12}}$$

We must now turn to the conversion chart in section 11.8 to find the values of the above a-parameters in terms of the given h-parameters. To do this we note that, from the fifth row of the chart,

$$a_{11} = -dh/h_{21} = -0.0085 \qquad a_{12} = -h_{11}/h_{21} = -25$$
$$a_{21} = -h_{22}/h_{21} = -0.000\,012\,5 \quad a_{22} = -1/h_{21} = -0.025$$

Now carefully substitute these values (including $V_1 = 0.001$, $R = 500$, $R_\mathrm{L} = 900$, into the above formula for V_L. Doing this, you should find that

$$V_\mathrm{L} = \frac{0.9}{2.816\,275} = 0.319\,571$$
$$= 0.3196 \text{ volts approx., } \textit{answer}.$$

In regard to the above problem it should be noted that the total phase shift produced by two CE stages in cascade (resistive loads) is equal to $180^\circ + 180^\circ = 360^\circ$, which effectively amounts to zero degrees of phase shift between input and output signals; thus we can disregard the 180° of phase shift produced in each stage.

* Upon applying the conversion chart of section 11.8 to eq. (543).

264. By the rules for matrix multiplication (and dispensing with the "overscore" notation) we have

$$\begin{bmatrix} 1 & 1 & 1 \\ a & a^2 & 1 \\ a^2 & a & 1 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_0 \end{bmatrix} = \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

265. The basic problem here is to find the inverse of the (3×3) matrix, as indicated. This can be done by following the procedure summarized in connection of eq. (482) in section 11.3.

In the present case you should find, for the first step, noting that $a^4 = a^3 a = a$, that

$$\Delta = 3a(a-1)$$

Next, you can check that the "cofactor" form of the (3×3) matrix is (again noting that $a^4 = a$)

$$\begin{bmatrix} (a^2-a) & -(a-a^2) & (a^2-a) \\ -(1-a) & (1-a^2) & -(a-a^2) \\ (1-a^2) & -(1-a) & (a^2-a) \end{bmatrix}$$
$$= \begin{bmatrix} a(a-1) & a(a-1) & a(a-1) \\ (a-1) & -(a-1)(a+1) & a(a-1) \\ -(a-1)(a+1) & (a-1) & a(a-1) \end{bmatrix}$$
$$= (a-1)\begin{bmatrix} a & a & a \\ 1 & -(a+1) & a \\ -(a+1) & 1 & a \end{bmatrix}$$

Now *transpose* the above matrix, then multiply by $1/\Delta$; doing this, the statement of problem 265 becomes

$$\begin{bmatrix} A_1 \\ A_2 \\ A_0 \end{bmatrix} = \frac{1}{3a}\begin{bmatrix} a & 1 & -(a+1) \\ a & -(a+1) & 1 \\ a & a & a \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

Now, on the right-hand side of the above equation, take the product of the (3×3) matrix and the (3×1) matrix, then multiply the result by $1/3a$. You now have the equality of two (3×1) matrices and thus, in accordance with the law of equal matrices, you have found that

$$A_1 = \frac{1}{3a}[aA + B - (a+1)C]$$
$$A_2 = \frac{1}{3a}[aA - (a+1)B + C]$$
$$A_0 = \frac{1}{3a}(aA + aB + aC)$$

Actually (though it doesn't seem true at first glance) the above three equations *are* the same as eqs. (460), (463), and (466). To see this, first note that

$$a^3 = a^2 a = 1 \qquad \text{thus also, } 1/a = a^2$$

thus the last three equations above can be written in the form

$$A_1 = \tfrac{1}{3}[A + a^2 B - (1 + a^2)C]$$

$$A_2 = \tfrac{1}{3}[A - (1 + a^2)B + a^2 C]$$

$$A_0 = \tfrac{1}{3}(A + B + C)$$

For the final step note that

$$-(1 + a^2) = -(1 + \epsilon^{j240}) = -(1 + \cos 240 + j \sin 240) = -0.500 + j0.8660 = \epsilon^{j120} = a$$

and thus, upon setting $-(1 + a^2) = a$, the last three equations become exactly the same as eqs. (460), (463), and (466).

266. As in example 1:

$$\begin{aligned}
67 \text{ divided by } 2 &= 33, \text{ plus remainder } 1, \text{ the LSD}\\
33 \text{ divided by } 2 &= 16, \text{ plus remainder } 1, \text{ next LSD}\\
16 \text{ divided by } 2 &= 8, \text{ plus remainder } 0, \text{ next LSD}\\
8 \text{ divided by } 2 &= 4, \text{ plus remainder } 0, \text{ next LSD}\\
4 \text{ divided by } 2 &= 2, \text{ plus remainder } 0, \text{ next LSD}\\
2 \text{ divided by } 2 &= 1, \text{ plus remainder } 0, \text{ next LSD}\\
1 \text{ divided by } 2 &= 0, \text{ plus remainder } 1, \text{ MSD}
\end{aligned}$$

Hence

$$67 \text{ dec} = 1000011 \text{ binary}, \textit{ answer}.$$

267. As in example 1:

$$\begin{aligned}
383 \text{ divided by } 2 &= 191, \text{ plus remainder } 1, \text{ the LSD}\\
191 \text{ divided by } 2 &= 95, \text{ plus remainder } 1, \text{ next LSD}\\
95 \text{ divided by } 2 &= 47, \text{ plus remainder } 1, \text{ next LSD}\\
47 \text{ divided by } 2 &= 23, \text{ plus remainder } 1, \text{ next LSD}\\
23 \text{ divided by } 2 &= 11, \text{ plus remainder } 1, \text{ next LSD}\\
11 \text{ divided by } 2 &= 5, \text{ plus remainder } 1, \text{ next LSD}\\
5 \text{ divided by } 2 &= 2, \text{ plus remainder } 1, \text{ next LSD}\\
2 \text{ divided by } 2 &= 1, \text{ plus remainder } 0, \text{ next LSD}\\
1 \text{ divided by } 2 &= 0, \text{ plus remainder } 1, \text{ MSD}
\end{aligned}$$

thus

$$383 \text{ dec} = 101111111 \text{ binary}, \textit{ answer}.$$

268. First, for the *whole part*, 118 dec = 1110110 binary. Next, the binary equivalent of the decimal fraction 0.182 is found as follows (see example 2):

$$0.182 \times 2 = 0.364,\ \text{hence},\ 0.0$$
$$0.364 \times 2 = 0.728,\ \text{hence},\ 0.00$$
$$0.728 \times 2 = 1.456,\ \text{hence},\ 0.001$$
$$0.456 \times 2 = 0.912,\ \text{hence},\ 0.0010$$
$$0.912 \times 2 = 1.824,\ \text{hence},\ 0.00101$$
$$0.824 \times 2 = 1.648,\ \text{hence},\ 0.001011$$
$$0.648 \times 2 = 1.296,\ \text{hence},\ 0.0010111$$
$$0.296 \times 2 = 0.592,\ \text{hence},\ 0.00101110$$
$$0.592 \times 2 = 1.184,\ \text{hence},\ 0.001011101$$

thus

$$118.182 \text{ dec} = 1110110.001011101 \text{ binary},\ \textit{answer}$$

the *check* on the binary fraction part showing that

$$0.001011101 \text{ bi} = 1/8 + 1/32 + 1/64 + 1/128 + 1/512 = 0.182 \text{ dec. approx.}$$

269. In the same way as in example 3 we have

$$2 \times 1 = 2, +1 = 3$$
$$2 \times 3 = 6, +1 = 7$$
$$2 \times 7 = 14, +0 = 14$$
$$2 \times 14 = 28, +1 = 29$$
$$2 \times 29 = 58, +0 = 58$$
$$2 \times 58 = 116, +1 = 117,\ \textit{answer}.$$

270. First, for the *whole part*, 1001 bi = 9 dec. Next, to find the decimal equivalent of the binary fractional part 0.01101, use the procedure of example 4, thus

$$1 \text{ divided by } 2 = 0.5$$
$$(0.5 + 0) \text{ divided by } 2 = 0.25$$
$$(0.25 + 1) \text{ divided by } 2 = 0.625$$
$$(0.625 + 1) \text{ divided by } 2 = 0.8125$$
$$(0.8125 + 0) \text{ divided by } 2 = 0.40625$$

Hence the equivalent decimal number is 9.40625, *answer*.

271. (a) Noting that 1 + 1 = "zero, carry 1 to next column to left," we have

$$\begin{array}{rl} 0\,1\,0\,1\,1\,0\,1 & (45 \text{ dec}) \\ \underline{1\,0\,1\,1\,0\,0\,1} & (89 \text{ dec}) \\ 1\,0\,0\,0\,0\,1\,1\,0,\ \textit{answer} & (134 \text{ dec}) \end{array}$$

(b) Noting $1+1+1=(1+1)+1=1$ and carry 1 to next column to left, we have

$$\begin{array}{rl} 1\,0\,1\,1 & \text{(11 dec)} \\ 0\,1\,0\,1 & \text{(5 dec)} \\ \underline{1\,1\,0\,1} & \text{(13 dec)} \\ 1\,1\,1\,0\,1, \; \textit{answer} & \text{(29 dec)} \end{array}$$

(c)

$$\begin{array}{rl} 1\,1\,.\,0\,1\,1 & \text{(3.375 dec)} \\ \underline{0\,1\,.\,1\,0\,1} & \text{(1.625 dec)} \\ 1\,0\,1\,.\,0\,0\,1, \; \textit{answer} & \text{(5.000 dec)} \end{array}$$

272. (a) Letting N_c denote the 1's complement of the subtrahend N, we take the *sum* of Y and N_c, then make the overflow "end-around carry," thus

$$\begin{array}{rl} & 1\,1\,0\,1\,0\,1\,1 \\ + & \underline{0\,1\,1\,0\,0\,0\,0} \\ 1 & 0\,0\,1\,1\,0\,1\,1 \\ \hookrightarrow & \underline{\qquad\qquad\;\; 1} \\ & 0\,0\,1\,1\,1\,0\,0, \; \textit{answer} \text{ (28 dec)} \end{array}$$

(b) Here the above procedure produces *no overflow 1*, which signals that the minuend Y is *less* than the subtrahend N, meaning that the answer is *negative*. In such a case we use the 1's complement of the *minuend* Y, and proceed as before; thus

$$\begin{array}{rl} & 0\,1\,1\,0\,0\,0\,0 \\ + & \underline{1\,1\,0\,1\,0\,1\,1} \\ 1 & 0\,0\,1\,1\,0\,1\,1 \\ \hookrightarrow & \underline{\qquad\qquad\;\; 1} \\ & 0\,0\,1\,1\,1\,0\,0, \; \textit{answer} \text{ (}-28\text{ dec)} \end{array}$$

Note: In the following problems "item numbers" refer to the table of "theorems" in section 12.2.

273. By repeated application of item (6), $A+A+A+A=A$ and $B+B+B=B$, the *answer* is

$$\text{"}A+B\text{," that is, } A \text{ or } B.$$

274. By repeated application of item (5),

$$AAABBCCCC = ABC, \textit{answer}.$$

275. $A+B+C+ABC+1=1$, *answer*, by the basic item (9). That is, a network composed of a closed switch *in parallel* with anything else constitutes a closed network.

276. We use the basic item (17); thus

$$\overline{(A+B)+C} = \overline{A+B}\,\bar{C} = \bar{A}\,\bar{B}\,\bar{C}.$$

277. We use the basic item (18); thus

$$\overline{ABC} = \overline{(AB)C} = \overline{AB} + \bar{C} = \bar{A} + \bar{B} + \bar{C}.$$

278. $Z = A(1 + B) = A$, *answer*, by items (9) and (7).

279. $Z = AAB(1 + \bar{C}) = AB$, *answer*, by items (5), (9), and (7).

280. One way is to first note that $A(A + B) = A + AB = A(1 + B) = A$, so that the given expression simplifies to

$$Z = A(B + C), \ \textit{answer}.$$

Or instead, if we wish, we can begin by first taking the product of the two binomials; thus

$$\begin{aligned} Z = A(AB + AC + BB + BC) &= A(B + C) + AB(1 + C) \\ &= AB + AC + AB \end{aligned}$$

which, since $AB + AB = AB$, by item (6), gives

$$Z = AB + AC = A(B + C), \ \textit{answer}, \ \text{as before.}$$

281. One procedure is to first note that, by items (14) and (7), we have $(BC + \bar{B}C) = (B + \bar{B})C = C$ and $AC + A\bar{C} = A(C + \bar{C}) = A$ so that the given expression becomes

$$Z = A + C + \bar{B}\bar{C} = A + (C + \bar{C}\bar{B})$$

Now, by item (16), note that $C + \bar{C}\bar{B} = C + \bar{B}$, so that the last expression above becomes.

$$Z = A + \bar{B} + C, \ \textit{answer}.$$

282. One way is to first apply item (17) then item (5), thus getting

$$Z = B + \bar{B}\bar{C}\bar{D}\bar{D} = B + \bar{B}\bar{C}\bar{D}$$

Now apply the generic form of item (16), that is, $A + \bar{A}Y = A + Y$, then apply item (17). Doing this, the last answer above becomes

$$Z = B + \bar{C}\bar{D} = B + \overline{C + D}, \ \textit{answer}.$$

Comment. We have thus found, by applying the laws of Boolean algebra, that

$$B + \bar{C}\bar{D}(\overline{B + D}) = B + \overline{C + D}$$

It should be noted that, although the two sides of the equation represent *physically different switching arrangements*, both sides perform the *same electrical switching operation*. Thus the right-hand side does the same electrical switching job as the left-hand side, but has the advantage of being *physically simpler*. Also, from a practical viewpoint, it's important to note that the switching network represented by the *right-hand side* requires just *one inverter* (one transistor), while the electrically equivalent network represented by the left-hand side requires *three* inverters.

283. Note first that $ABC + \bar{A}BC = (A + \bar{A})BC = BC$, by items (14) and (7), and upon making use of this relationship the problem becomes

$$Z = \overline{BC + \bar{A}\bar{B}\bar{C}} = \overline{BC + \overline{A + B + C}}$$

because $\bar{A}\bar{B}\bar{C} = \overline{A + B + C}$ (from problem 276). Now apply, to the last expression above, the generic form of item (17), that is, $\overline{X+Y} = \bar{X}\,\bar{Y}$, to get

$$Z = \overline{BC}\,(\overline{\overline{A+B+C}}) = \overline{BC}\,(A+B+C),\ \textit{answer},\ \text{by item (15).}$$

Note that we've greatly simplified the original switching network because now only two "and" circuits, one "or" circuit, and one "not" circuit are needed.

284. Using the result of problem 276, the given problem can be written as $Z = \bar{A}\,\bar{B}\,\bar{C}\,\bar{D} + \bar{A}\,\bar{C}\,\bar{D} = \bar{A}\,\bar{C}\,\bar{D}(\bar{B}+1) = \bar{A}\,\bar{C}\,\bar{D}$, by items (9) and (7). Hence, again from problem 276,

$$Z = \overline{A+C+D},\ \textit{answer}.$$

285. By items (17) and (15),

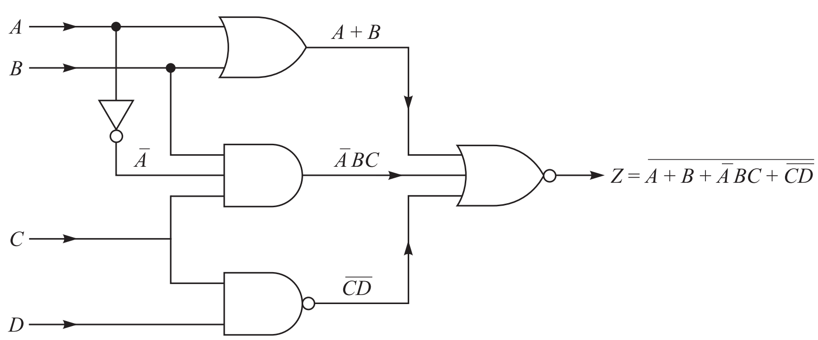
$$A + \bar{A}\,B = A + B,\ \textit{answer}.$$

286. $\overline{Z = (A+\bar{A})\bar{B}\,C + A\,\bar{B}\,\bar{C}} = \overline{\bar{B}\,C + A\,\bar{B}\,\bar{C}}$, by items (14) and (7); then

$Z = \overline{\bar{B}(C+\bar{C}\,A)} = \overline{\bar{B}(C+A)}$, by item (16); thus

$$Z = B + \overline{A+C},\ \textit{answer},\ \text{by item (18).}$$

287.



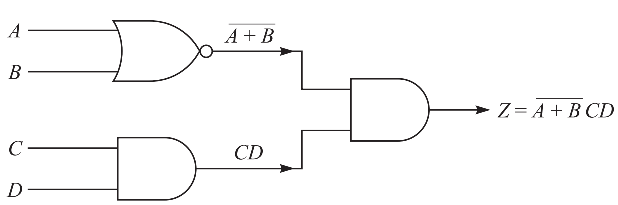
288. We are asked to simplify the expression for the output signal Z in problem 287. To do this, we can begin with the simplification $A + \bar{A}\,BC = A + BC$, by item (16), which puts Z into the form

$$Z = \overline{A+B+BC+\overline{C\,D}} = \overline{A+B+\overline{C\,D}}$$

where we used the simplification $B + BC = B(1+C) = B$. Now apply items (17) and (15) to get

$$Z = \overline{A+B}\,CD,\ \textit{answer}$$

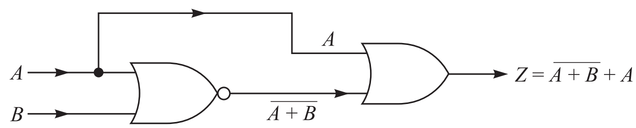
(*not* the quantity A or B *and* the quantity C and D). A sketch of the simplified network is shown below.



289. First, the elemental equation is $Z = \bar{A}\,\bar{B} + A\,\bar{B} + AB = \bar{A}\,\bar{B} + A(\bar{B} + B)$, and thus

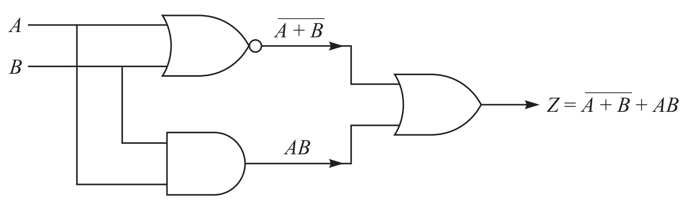
$$Z = \overline{A + B} + A, \ \textit{answer}, \text{ by items (14), (7), (17).}$$

The above answer is accomplished by the network below.



290. (a) $Z = \bar{A}\,\bar{B} + AB$, *answer*.

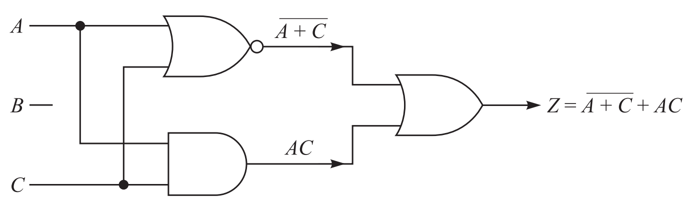
(b) First, by item (17), $Z = \overline{A + B} + AB$, which is produced by the network



Note that $Z = \overline{A + B} + AB$ *does* satisfy the given truth table.

291. (a) $Z = \bar{A}\,\bar{B}\,\bar{C} + \bar{A}\,B\,\bar{C} + A\,\bar{B}\,C + ABC$, *answer*.
Check: inspection of the answer shows that $Z = 1$ (which denotes the presence of a signal on the output line) if any of the input signal combinations 000, 010, 101, and 111 appear on the input lines, and $Z = 0$ for any other combination of input signals.

(b) First note that $Z = \bar{A}\,\bar{C}(\bar{B} + B) + AC(\bar{B} + B)$, which, after applying items (14), (7), and (17), becomes $Z = \overline{A + C} + AC$, which can be generated by the hardware arrangement



Note that in this case the signal B need not be connected to the network, because the state of signal B, 1 or 0, does not affect the state of output signal Z.

292. (a) $Z = \bar{A}\,\bar{B}\,C\,\bar{D} + \bar{A}\,B\,\bar{C}\,\bar{D} + \bar{A}\,B\,C\,\bar{D} + A\,\bar{B}\,C\,D + ABC\bar{D} + ABCD$, *answer*.

(b) One way to proceed is as follows. First, noting that the quantity $\bar{A}\,\bar{D}$ factors from the first three terms, and AC factors from the last three terms, we can take the following steps:

$$Z = \bar{A}\,\bar{D}(\bar{B}\,C + B\,\bar{C} + BC) + AC(\bar{B}\,D + B\,\bar{D} + BD)$$

then

$$Z = \bar{A}\,\bar{D}\,\bar{B}\,C + B(\bar{C} + C) + AC\,\bar{B}\,D + B(\bar{D} + D)$$

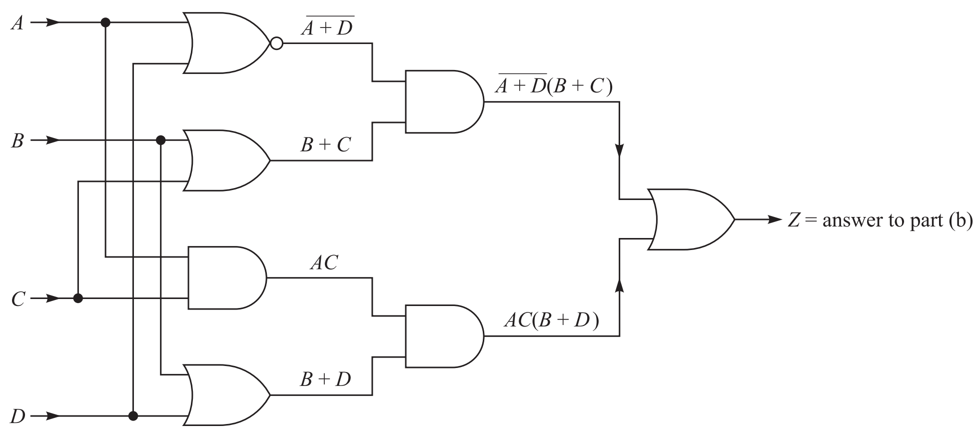
which, after applying items (14) and (7), becomes

$$Z = \bar{A}\,\bar{D}(B + \bar{B}\,C) + AC(B + \bar{B}\,D)$$

which, after applying items (17) and (16), becomes

$$Z = \overline{A + D}(B + C) + AC(B + D),\ \textit{answer}.$$

(c)



293. (a) $Z = \bar{A}\,\bar{B}\,C\,\bar{D} + \bar{A}\,\bar{B}\,C\,D + \bar{A}\,B\,C\,D + A\bar{B}\,C\,\bar{D} + ABCD$, *answer*.

(b) $Z = C[\bar{A}\,\bar{B}(\bar{D} + D) + (\bar{A} + A)BD + A\,\bar{B}\,\bar{D}]$, then

$$Z = C(\bar{A}\,\bar{B} + BD + A\,\bar{B}\,\bar{D})$$

by items (14) and (7). Next,

$$Z = C[BD + \bar{B}(\bar{A} + A\,\bar{D})]$$

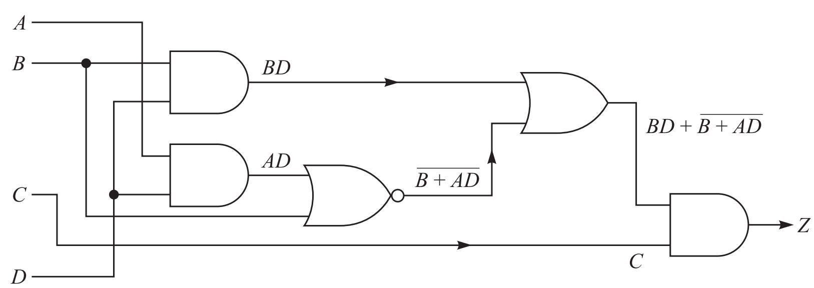
and since $\bar{A} + A\,\bar{D} = \bar{A} + \bar{D}$ (example 11 on p. 349, we have

$$Z = C[BD + \bar{B}(\bar{A} + \bar{D})] = C(BD + \bar{B}\,\overline{AD})$$

by item (18). Now, making use of the basic relationship $\bar{X}\,\bar{Y} = \overline{X + Y}$, the last result becomes

$$Z = C(BD + \overline{B + AD}),\ \textit{answer}.$$

(c)



294. (a) Here, each block of information fed into the encoder will consist of 12 binary digits, each such block representing one of the 12 possible values of V_q, including $V_q = 0$.

Hence, letting $A, B, C, \ldots, I, J, K$, denote each input block to the encoder, these blocks will be in the form of 12 binary numbers from 0000 to 1011 (binary eleven). Thus the TRUTH TABLE for the encoder is as follows:

V_q	A	B	C	D	E	F	G	H	I	J	K	X_4	X_3	X_2	X_1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0
3	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1
4	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0
5	1	1	1	1	1	0	0	0	0	0	0	0	1	0	1
6	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0
7	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1
8	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0
9	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1
10	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0
11	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1

(b) Close inspection of the above table will show that the output "X digits" will have the value "1" (signal "is present") only if the following Boolean equations are satisfied. (Otherwise it's understood that $X = 0$, meaning the signal is "not present" for any other arrangements of the $A, B, C, \ldots, I, J, K$ input signals.)

$$X_1 = 1 = A\,\bar{B} + C\bar{D} + E\,\bar{F} + G\,\bar{H} + I\,\bar{J} + K$$

$$X_2 = 1 = B\,\bar{D} + F\,\bar{H} + J$$

$$X_3 = 1 = D\,\bar{F} + F\,\bar{H}$$

$$X_4 = 1 = H$$

Thus, in terms of hardware, it would be necessary to provide 5 "not," 7 "and," and 3 "or" devices.

295. First note that as we go from RIGHT TO LEFT in the series of eq. (576) the values of the exponents increase by $+1$ from term to term. Let us show this more clearly by showing a few more terms at the right-hand end of the series, thus

$$F(z) = 1 + z^{-1} + z^{-2} + z^{-3} + \cdots + z^{-n}z^{3} + z^{-n}z^{2} + z^{-n}z^{1} + z^{-n} \tag{A}$$

where we made use of the law of exponents, $z^{-n+3} = z^{-n}z^3$, $z^{-n+2} = z^{-n}z^2$ and so on.

Now multiply both sides of eq. (A) by $-z^{-1}$, as suggested. Doing this, eq. (A) becomes

$$-z^{-1}F(z) = -z^{-1} - z^{-2} - z^{-3} - z^{-4} - \cdots - z^{-n}z^{2} - z^{-n}z^{1} - z^{-n} - z^{-n}z^{-1} \tag{B}$$

Now take the algebraic sum of eqs. (A) and (B), as suggested. Upon doing this, note that on the right-hand side all of the terms except two will cancel out, leaving

$$(1 - z^{-1})F(z) = 1 - z^{-n}z^{-1}$$

which upon solving for $F(z)$ gives eq. (577), as explained in the text.

296. First, for this problem, $v(nT) = \cos \omega nT$. Then, by Euler's formulas, we have

$$\epsilon^{j\omega nT} = \cos \omega nT + j \sin \omega nT$$
$$\epsilon^{-j\omega nT} = \cos \omega nt - j \sin \omega nT$$

which, upon taking the algebraic sum of the two equations, shows that

$$\cos \omega nt = \tfrac{1}{2}(\epsilon^{j\omega nT} + \epsilon^{-j\omega nT}) = v(nT)$$

and upon substituting this value of $v(nT)$ into the basic eq. (573) we have

$$F(z) = \tfrac{1}{2}\sum_{n=0}^{n=\infty}(\epsilon^{j\omega nT}z^{-n} + \epsilon^{-j\omega nT}z^{-n})$$

or, if we wish,

$$F(z) = \tfrac{1}{2}\sum_{n=0}^{n=\infty}[(\epsilon^{-j\omega T}z)^{-n} + (\epsilon^{j\omega T}z)^{-n}] \qquad \text{(A)}$$

in which we made use of the fact that since $(X^{ab}) = (X^a)^b$, then also $(X^{ab}) = (X^{-a})^{-b}$.

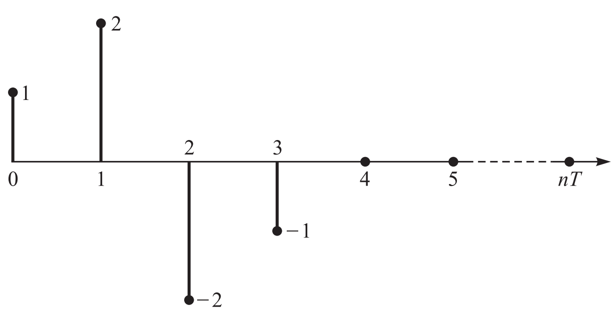
Note, now, that eq. (580) applies to eq. (A), where $b = -j\omega$ in the *first term* and $b = j\omega$ in the *second term*.

Thus eq. (581) applies to eq. (A), where $k = \epsilon^{j\omega T}$ for the first term and $k = \epsilon^{-j\omega T}$ for the second term. Hence, by eq. (581), eq. (A) becomes

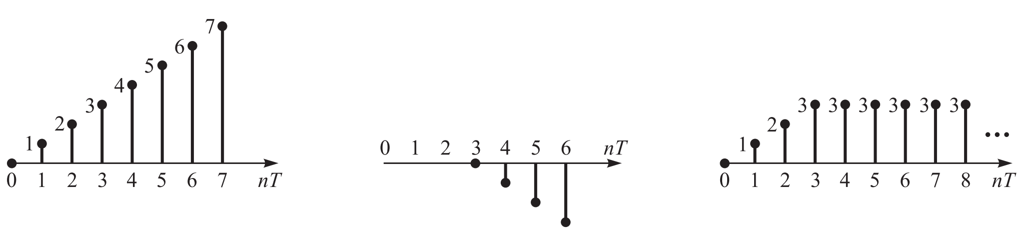
$$F(z) = \frac{z}{2}\left[\frac{1}{z - \epsilon^{j\omega T}} + \frac{1}{z - \epsilon^{-j\omega T}}\right]$$

Now, inside the brackets, combine the two fractions together over the common denominator (the product of the two denominators), then apply Euler's formulas to the result. Carefully doing this, you should find that the z-transform of $v(nT) = \cos \omega nT$ is truly given by (7) in the table.

297.

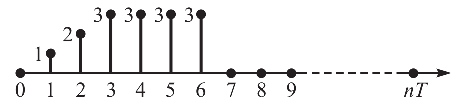


298. The *first two sequences*, and their algebraic sum, are shown below.



Sequence nT + sequence $-(nT-3T)$ = sum of $nT-(nT-3T)$

Therefore the addition of the delayed step sequence, $-3U(nT-7T)$, produces the *final answer* shown to the right.



299. IN SAMPLED form $U(t)$ becomes $U(nT)$ and t becomes nT. Thus we have that

$$v(nT) = 3U(nT) - 20nT$$

Hence, making use of (1) and (2) in Table 2, then (2) and (4) in Table 1, we have that

$$Zv(nT) = F(z) = \frac{3z}{z-1} - \frac{20Tz}{(z-1)^2}$$

By eq. (91), Chap. 5, $T = 1/f = 1/100 = 0.01$. Making use of this fact, then combining the two fractions over a common denominator, should produce the *answer* given with the problem.

300. Applying (2) of Table 1 and (4) of Table 2, we have

$$F(z) = \frac{z}{z-1} + \frac{z^{-1}z}{z-1} + \frac{z^{-2}z}{z-1} = (z+1+z^{-1})/(z-1), \ \textit{answer}.$$

301. Using (4) and (2) of Table 1 and (4) of Table 2, we have

$$F(z) = \frac{Tz}{(z-1)^2} - \frac{Tz^{-2}}{(z-1)^2} - \frac{z^{-6}}{(z-1)}, \ \textit{answer},$$

or, if you wish,

$$F(z) = \frac{Tz}{(z-1)^2} - \frac{T}{z^2(z-1)^2} - \frac{1}{z^6(z-1)}, \ \textit{answer}.$$

302.

$$\begin{array}{r|l}
 & z^{-1} + 0.5z^{-2} + 0.25z^{-3} + 0.125z^{-4} + 0.0625z^{-5} \\ \hline
z - 0.5 & 1 \\
 & \underline{-1 + 0.5z^{-1}} \\
 & \quad 0.5z^{-1} \\
 & \quad \underline{-0.5z^{-1} + 0.25z^{-2}} \\
 & \qquad 0.25z^{-2} \\
 & \qquad \underline{-0.25z^{-2} + 0.125z^{-3}} \\
 & \qquad\quad 0.125z^{-3} \\
 & \qquad\quad \underline{-0.125z^{-3} + 0.0625z^{-4}} \\
 & \qquad\qquad 0.0625z^{-4}
\end{array}$$

hence the *answers* are

$$\begin{array}{lll}
y(0) = 0.000 & y(2T) = 0.500 & y(4T) = 0.125 \\
y(T) = 1.000 & y(3T) = 0.250 & y(5T) = 0.0625
\end{array}$$

303. For convenience, let's temporarily write "a" in place of "0.45." Thus we have $F(z) = \dfrac{z}{z^2 - a}$, and our problem is to put $F(z)$ into the form of eq. (583). To do this we can use algebraic long division; thus,

$$
\begin{array}{r|l}
 & z^{-1} + az^{-3} + a^2z^{-5} + a^3z^{-7} + a^4z^{-9} + \cdots \\
\hline
z^2 - a & z \\
 & \underline{-z + az^{-1}} \\
 & \quad az^{-1} \\
 & \quad \underline{-az^{-1} + a^2z^{-3}} \\
 & \qquad a^2z^{-3} \\
 & \qquad \underline{-a^2z^{-3} + a^3z^{-5}} \\
 & \qquad\qquad a^3z^{-5} \\
 & \qquad\qquad \underline{-a^3z^{-5} + a^4z^{-7}} \\
 & \qquad\qquad\qquad a^4z^{-7}
\end{array}
$$

Thus we have that

$$
\begin{aligned}
F(z) = {} & 0 + (1)z^{-1} + (0)z^{-2} + az^{-3} + (0)z^{-4} + a^2z^{-5} + (0)z^{-6} + a^3z^{-7} \\
& + (0)z^{-8} + a^4z^{-9} + \cdots
\end{aligned}
$$

which is in the form of eq. (583). The general statement in the time domain, eq. (584), is

$$v_s(t) = v(0)\delta(t) + v(T)\delta(t - T) + v(2T)\delta(t - 2T) + v(3T)\delta(t - 3T) + \cdots$$

and thus, by direct comparison, the *answers* are

$v(0) = 0.000$	$v(3T) = a = 0.450$	$v(6T) = 0.000$	$v(9T) = a^4 = 0.041$
$v(T) = 1.000$	$v(4T) = 0.000$	$v(7T) = a^3 = 0.091$	
$v(2T) = 0.000$	$v(5T) = a^2 = 0.203$	$v(8T) = 0.000$	

304. Referring to the notation in the generalized Fig. 346, we see that, in Fig. 345, $b_0 = 2$, $b_1 = 3$, $b_2 = 7$, and $a_1 = 10$. Hence, for Fig. 345, eq. (589) becomes

$$Y(z)/X(z) = H(z) = (2 + 3z^{-1} + 7z^{-2})/(1 - 10z^{-1}), \ \textit{answer}.$$

305. The a coefficients in Fig. 346 would all be equal to *zero*, because a non-recursive processor uses no feedback; hence, for this type of processor, eq. (589) reduces to

$$Y(z)/X(z) = H(z) = b_0 + b_1z^{-1} + b_2z^{-2} + \cdots + b_pz^{-p}, \ \textit{answer}.$$

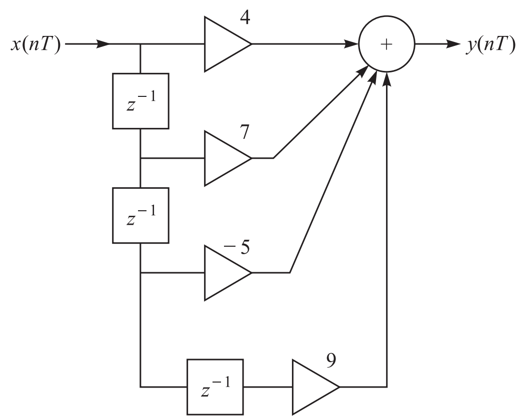
306. From eq. (588) and Fig. 346, a processor is non-recursive if all the a coefficients are zero. Hence the *answers* here are

(a) non-recursive,

(b) non-recursive,

(c) recursive.

307. (a) Comparison with eq. (588) shows that this is a purely non-recursive processor (no "$y(nT - qT)$" terms on the right-hand side). Thus this is a non-recursive processor such as is illustrated in Fig. 343, but here requiring three delays. Hence the block diagram is as follows.



(b) Since this is a purely non-recursive processor, all the a coefficients in eq. (589) are zero. Thus setting $b_0 = 4$, $b_1 = 7$, $b_2 = -5$, and $b_3 = 9$ in eq. (589), the *answer* is

$$H(z) = 4 + 7z^{-1} - 5z^{-2} + 9z^{-3}$$

308. For $q = 4$ in eq. (597) we have

$$z^4 - a_1z^3 - a_2z^2 - a_3z - a_4 = 0$$

which, as explained in connection with eq. (602), possesses four poles. Since complex poles can occur only in the form of pairs of conjugate complex numbers, we have that the possibilities are

(a) four real poles, or
(b) two real and one pair of conjugate poles, or
(c) two pairs of conjugate poles.

309. For $q = 5$ in eq. (597) we have

$$z^5 - a_1z^4 - a_2z^3 - a_3z^2 - a_4z - a_5 = 0$$

and hence, since a fifth-degree equation possesses five roots, the possibilities are

(a) five real poles, or
(b) three real and one conjugate pair, or
(c) one real and two conjugate pairs.

310. First, setting the *numerator* equal to zero, $4z + 9 = 0$, shows that $H(z)$ has one *zero*, for $z = -9/4 = -2.25$.

Next, to find the poles we set the *denominator* equal to zero and solve for z; thus

$$z(z-9)(z^2 + 5z + 7) = 0 \tag{A}$$

which let us now put into the form of eq. (602), thus

$$z(z-9)(z-h_a)(z-h_b)=0 \tag{B}$$

in which h_a and h_b are the two *roots* of the quadratic factor, which are, in this case, found by setting $a=1$, $b=5$, and $c=7$, into the "standard quadratic formula" which, as you should verify, yields the results

$$h_a=-2.5+j0.87 \text{ approx.,}$$
$$h_b=-2.5-j0.87 \text{ approx.}$$

The advantage of putting eq. (A) into the form of (B) is that we can then see, by direct inspection of (B), that the FOUR POLES of $H(z)$ are located at

$$z=0 \qquad z=h_a=-(2.5-j0.87)$$
$$z=9 \qquad z=h_b=-(2.5+j0.87)$$

The above demonstrates that the most time-consuming part of such solutions lies in the necessity of *factoring* higher-degree expressions that may be present.

311. (a) The roots of the denominator, that is, the "poles" of $H(z)$, are, by inspection, located at $z=0.46$ and $z=0.22$. Thus, since both poles lie *inside* the unit circle, the processor *is stable*.

(b) The roots (poles) of $H(z)$ are those values of z for which the denominator of $H(z)$ is equal to zero.

In this case, by inspection, we see that the first pole is at $z=0.61$. Next, setting $z^2-1.6z+0.48=0$ gives the two poles $z=1.2$ and $z=0.4$. Since the pole at $z=1.2$ lies *outside* the unit circle, the processor is *unstable*.

(c) Let us first write

$$H(z)=\frac{z-1.2}{z^2-1.37z+0.305}$$

which was obtained by multiplying the numerator and denominator of the given fraction by z^2. Then, setting the denominator $z^2-1.37z+0.305=0$ gives the two roots (poles) $z=1.09$ and $z=0.28$. Since the pole at $z=1.09$ lies *outside* the unit circle, the processor is *unstable*.

312. Multiply the numerator and denominator by z^{-3}, thus getting,

$$H(z)=\frac{2+4z^{-1}+z^{-2}}{1-5z^{-1}+6z^{-2}+9z^{-3}}=\frac{2+4z^{-1}+z^{-2}}{1-(5z^{-1}-6z^{-2}-9z^{-3})}, \ \textit{answer}.$$

313. The "unit pulse response" of a processor is the same as the "transfer function" of the processor (section 13.6). Since the processors are connected in cascade (series), we have (section 13.8) that

$$\begin{aligned}H(z)&=H_1H_2H_3\\ &=\frac{z^3}{(z-0.2)(z-0.4)(z^2-0.8z+0.15)}=\frac{z^3}{z^4-1.4z^3+0.71z^2-0.154z+0.012}\end{aligned}$$

This is a "first answer" which, upon multiplying the numerator and denominator of the right-hand fraction by z^{-4}, is thereby put in the equivalent form of eq. (606); thus

$$H(z) = \frac{z^{-1}}{1 - (1.4z^{-1} - 0.71z^{-2} + 0.154z^{-3} - 0.012z^{-4})}, \text{ answer.}$$

314. This is a parallel connection of processors, and thus (section 13.8)

$$H(z) = H_1 + H_2 + H_3 = \frac{z}{z - 0.2} + \frac{z}{z - 0.4} + \frac{z}{z^2 - 0.8z + 0.15}$$

which, upon combining the three fractions into a single fraction, gives the equivalent *answers*

$$\begin{aligned} H(z) &= \frac{2z^4 - 1.2z^3 + 0.53z^2 + 0.31z}{z^4 - 1.4z^3 + 0.71z^2 - 0.154z + 0.012} \\ &= \frac{2 - 1.2z^{-1} + 0.53z^{-2} + 0.31z^{-3}}{1 - (1.4z^{-1} - 0.71z^{-2} + 0.154z^{-3} - 0.012z^{-4})} \end{aligned}$$

The second answer is in the form of eq. (606), and is found by multiplying the numerator and denominator of the first answer by z^{-4}.

315. As explained in section 13.6, for *unit-pulse input* the output $Y(z)$ of a processor is numerically *equal* to the transfer function $H(z)$. Thus, in this particular case for unit-pulse input voltage, we can use either of the equivalent answers found in problem 313.

The problem now is to put $Y(z)$ into the *form of eq. (583) in section 13.4*. To do this we can apply algebraic long division to either of the equivalent answers found in problem 313. If we choose to use the "first answer," the details of the long division are as follows.

$$\begin{array}{r|l} & z^{-1} + 1.4z^{-2} + 1.25z^{-3} + \cdots \\ \hline z^4 - 1.4z^3 + 0.71z^2 - 0.154z + 0.012 & z^3 \\ & \underline{-z^3 + 1.4z^2 - 0.71z + 0.154 - 0.012z^{-1}} \\ & \quad 1.4z^2 - 0.71z + 0.154 - 0.012z^{-1} \\ & \quad \underline{-1.4z^2 + 1.96z - 0.994 + 0.2156z^{-1} - 0.0168z^{-2}} \\ & \qquad 1.25z - 0.840 + 0.2036z^{-1} - 0.0168z^{-2} \end{array}$$

This is as far as we need to go to find the required answer. If, however, you wish to continue the above division for a couple of more terms, you'll find that

$$Y(z) = 0 + z^{-1} + 1.4z^{-2} + 1.25z^{-3} + 0.91z^{-4} + 0.5901z^{-5} + \cdots$$

which is the output of the network, in the z-domain, for unit pulse input. By comparison with eq. (583) we see that,

$$\begin{array}{lll} y(0) = 0.000 & y(2T) = 1.400 & y(4T) = 0.910 \\ y(T) = 1.000 & y(3T) = 1.250 & y(5T) = 0.5901 \end{array}$$

thus

$$y(3T) = 1.25\,\text{V}, \text{ answer.}$$

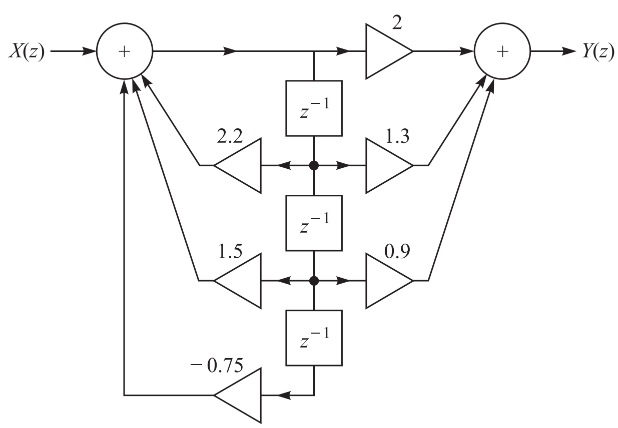
316. The notation in Fig. 356 is tied to that of eq. (606). Let us therefore put the given $H(z)$ into the equivalent form of eq. (606) by multiplying the numerator and denominator of $H(z)$ by z^{-3}, thus getting

$$H(z) = \frac{2 + 1.3z^{-1} + 0.9z^{-2}}{1 - (2.2z^{-1} + 1.5z^{-2} - 0.75z^{-3})}$$

Note that $H(z)$ now has the form of eq. (606), and thus comparison with eq. (606) shows that

$$b_0 = 2 \quad b_1 = 1.3 \quad b_2 = 0.9 \quad a_1 = 2.2 \quad a_2 = 1.5 \quad a_3 = -0.75$$

and hence, from comparison with Fig. 356, the required block diagram is as follows:



317. Comparing the given network with Fig. 356 shows that

$$a_1 = 1.66 \quad b_0 = 1.9 \quad b_2 = 1.6$$
$$a_2 = -1.5353 \quad b_1 = -2.2$$

and hence, by eq. (606), we have that

$$H(z) = \frac{1.9 - 2.2z^{-1} + 1.6z^{-2}}{1 - 1.66z^{-1} + 1.5353z^{-2}} = \frac{1.9z^2 - 2.2z + 1.6}{z^2 - 1.66z + 1.5353}$$

Now setting the denominator equal to zero, then making use of the formula for the roots of a quadratic function, you should find that

$$z = \frac{1.66 \pm j1.84}{2} = \begin{matrix} 0.83 + j0.92, & \text{first root,} \\ 0.83 - j0.92, & \text{second root} \end{matrix}$$

hence

$$|z| = \sqrt{(0.83)^2 + (0.92)^2} = 1.2391^*$$

which means that the points $0.83 \pm j0.92$ lie *outside* the unit circle; thus the given processor is *unstable*, *answer*.

* The *radius* of a circle with center at the origin of the complex plane is $r = |z| = \sqrt{x^2 + y^2}$.

318. From comparison of the given network with Fig. 356

$$a_1 = \quad 0.54 \qquad b_0 = 0.92$$
$$a_2 = -0.7453 \qquad b_1 = 0.82$$

Hence, by eq. (606),

$$H(z) = \frac{0.92 + 0.82z^{-1}}{1 - 0.54z^{-1} + 0.7453z^{-2}} = \frac{0.92z^2 + 0.82z}{z^2 - 0.54z + 0.7453}$$

To find the poles of $H(z)$, set the denominator equal to zero and solve for z. Doing this, making use of the quadratic formula, gives two poles, $0.27 \pm j0.82$, both of magnitude 0.8633. Thus all poles lie *inside* the unit circle, so the processor is *stable, answer*.

319. First, by eq. (608)

$$H(z) = 1 - z^{-1}$$

then,

$$\begin{aligned} H(r) &= 1 - \epsilon^{-j2\pi r} \\ &= 1 - (\cos 2\pi r - j \sin 2\pi r) \end{aligned}$$

that is,

$$H(r) = (1 - \cos 2\pi r) + j \sin 2\pi r$$

or, if we wish to work in degrees instead of radians,

$$H(r) = (1 - \cos 360r) + j \sin 360r$$

Hence, remembering that $\sin^2 x + \cos^2 x = 1$, we have that

$$|H(r)| = \sqrt{2(1 - \cos 360r)}$$

and

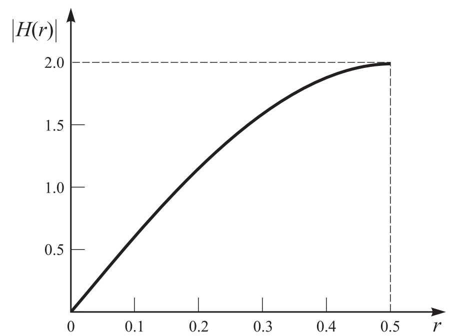
$$\phi = \arctan\left(\frac{\sin 360r}{1 - \cos 360r}\right)$$

Now, using the last two equations, you can verify that the following "table of values" of magnitudes and phase angles versus r is correct.

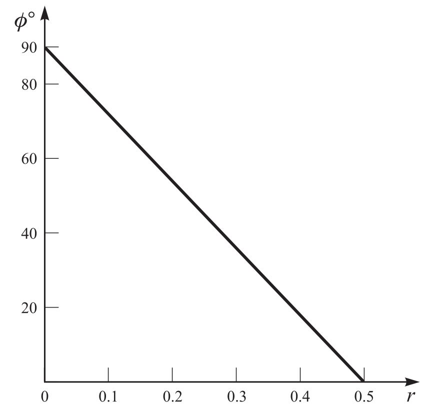
r	$\lvert H(r)\rvert$	ϕ°
0.0	0.00	90*
0.1	0.62	72
0.2	1.18	54
0.3	1.62	36
0.4	1.90	18
0.5	2.00	0

* Arctan $\infty = 90^\circ$

In graphical form these values appear as follows:



"High-pass" filter



Note that merely CHANGING THE SIGN of the multiplier b_1 converted the LOW-PASS filter of Fig. 358 into the HIGH-PASS filter of Fig. 365. Also note that phase shift ϕ is exactly proportional to analog frequency $\omega (r = \omega/\omega_s)$, which is characteristic of non-recursive digital filters.

320. Note that the processor is of the non-recursive type, corresponding to the left-hand side of Fig. 354 in section 13.8. Hence the "sinusoidal frequency response" of the processor is found by setting $z = \epsilon^{j2\pi r}$ in eq. (608), along with the given values of the "b" multiplier coefficients. Doing this, and setting $2\pi r = 360r$ if you prefer to work in degrees instead of radians, eq. (608) becomes, after applying Euler's formula,

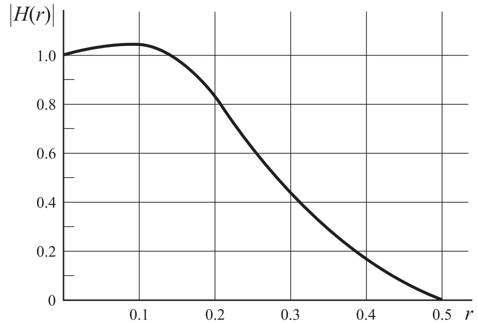
$$H(r) = [0.40 + 0.55\cos(360r) + 0.13\cos 2(360r) - 0.05\cos 3(360r) - 0.03\cos 4(360r)]$$
$$- j[0.55\sin(360r) + 0.13\sin 2(360r) - 0.05\sin 3(360r) - 0.03\sin 4(360r)]$$

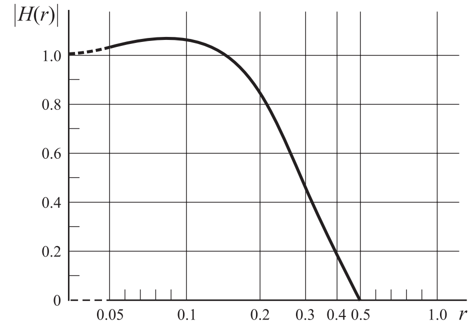
The first step now is to substitute, into the above equation for $H(r)$, the value $r = 0$; doing this, you should find that $H(0) = 1 + j0 = 1$.

Next substitute, into the above equation for $H(r)$, the value $r = 0.1$; doing this, you should find that, approximately, $H(0.1) = 0.93 - j0.38$. Continuing on in this way, for $r = 0.2$, 0.3, 0.4, and 0.5, should give you the following table of values:

r	$H(r)$	$\lvert H(r)\rvert$
0.0	$1.00 + j0.00$	1.00
0.1	$0.93 - j0.38$	1.05
0.2	$0.50 - j0.66$	0.83
0.3	$0.08 - j0.45$	0.46
0.4	$0.00 - j0.17$	0.17
0.5	$0.00 + j0.00$	0.00

The final step, now, is to plot the above values of $|H(r)|$, using either linear or semi-log graph paper (section 9.5). The result, for both types of paper, is shown as follows.





321. Fig. 366 is a recursive filter in which $a_1 = 0$, $a_2 = -0.25$, $b_0 = 1.0$, $b_1 = 0$, $b_2 = -1$, and upon substituting these values into eq. (606) we have that

$$H(z) = \frac{1 - z^{-2}}{1 + 0.25z^{-2}} = \frac{z^2 - 1}{z^2 + 0.25}$$

(a) Setting $z^2 + 0.25 = 0$ gives the poles, $z = \pm\sqrt{-0.25} = \pm j0.50$, both of which lie inside the unit circle; hence the filter *is stable, answer*.

(b) Setting $z = \epsilon^{j360r}$ in the above equation for $H(z)$, then applying Euler's formula, gives

$$H(r) = \frac{\epsilon^{j720r} - 1}{\epsilon^{j720r} + 0.25} = \frac{(-1 + \cos 720r) + j \sin 720r}{(0.25 + \cos 720r) + j \sin 720r}$$

One way to proceed now is to first rationalize the equation (section 6.3); doing this, then making use of the trigonometrical identity $\sin^2 x + \cos^2 x = 1$, you'll find the above equation becomes

$$H(r) = \frac{0.75(1 - \cos 720r) + j1.25 \sin 720r}{(1.0625 + 0.5 \cos 720r)}$$

Now substituting, into the above equation, the required values of r, you can verify the following table of values (rounded off, here, to two decimal places):

r	$H(r)$	$\lvert H(r) \rvert$
0.00	$0.00 + j0.00$	0.00
0.05	$0.10 + j0.50$	0.51
0.10	$0.43 + j1.00$	1.09
0.20	$2.06 + j1.12$	2.35
0.25	$2.67 + j0.00$	2.67
0.30	$2.06 - j1.12$	2.35
0.40	$0.43 - j1.00$	1.09
0.45	$0.10 - j0.50$	0.51
0.50	$0.00 + j0.00$	0.00

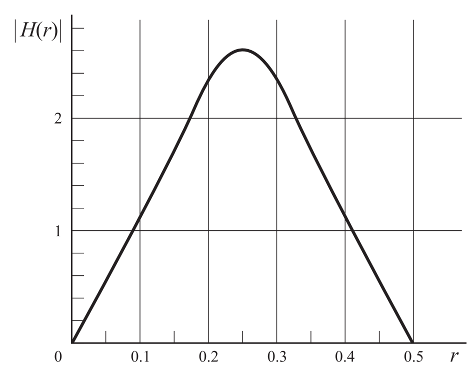


Figure 366 is thus a BAND-PASS type of digital filter

INDEX

Absolute value, 402
Active device, 28
Adder, full, 353
Admittance, 167
Algebra, review, 401–404
Algebraic long division, 437
Alternating current (ac), 76
Ammeter, 30
Ampere, 15
Amplifier, definition, 432
  transistor, 432
Analog signal, 324
Arctan, notation, 101
Associative, 401

Battery, 20, 21
Bilateral, 72
Binary arithmetic, 325–328
Binary signal, 324
  bandwidth needs, 357
Boolean algebra, 338–347

Capacitive reactance, 161
Capacitors, capacitance, 144–148
  series and parallel, 148
Characteristic impedance, 216
Coefficient of coupling, 229
Commutative, 401
Comparators, 362
Conductance, 66
Conductor, 2
Constant-current generator, 70
Constant-k filter, 219–223
Constant-voltage generator, 69
Conventional current, 18
Conversion factors, 404
Conversions, pi to T, T to pi, 196
Coulomb, 8
Coulomb's law, 8
Cramer's rule, 54
Critical coupling, 247
Current, 15

Decibel, 203
De Moivre's theorem, 131
Determinants, 38–57
Dielectric constant, 145
Digital filters, 393–400
Direct current (dc), 31
Discrete-time (DT) processors, 377–383
  introduction, 377
  recursive, non-recursive, 379
  stability, 383
  structure, 378, 389
  transfer function $H(z)$, 377, 382

Distortion (of signal), 423
Distributive, 401
Division, notation, 402, 404
Dot-marked terminals, 230
Double-subscript notation, 430

Electric charge, 1–8
Electric field strength, 11
Electromagnetism, 138
Electromotive force (emf), 20
Electron, 5, 16
Elemental equation, 349
Epsilon, ϵ, 125
Euler's formula, 127
Exponents, laws of, 403
Exponent zero, 403

Factorial, 126
Farad, 147
Frequency, 88
Frequency response, 201
Filters
  constant-*k*, 219, 223
  digital, 393–400
  RC type, 204, 209
  RL type, 212

Generator, 28
  ac and dc, 96
Graph paper, logarithmic, 425
Greek letters, 412
Ground, 37, 73

Harmonics, 419–421
Henry, 143
Hertz, 88

Identity, 80
Imaginary and complex numbers, 114–122
  as vectors, 134
  exponential form, 125, 127
  powers and roots of, 131
  trigonometric form, 124
Impedance
  series *RC* circuit, 162
  series *RL* circuit, 154
  series *RLC* circuit, 165
Impedance transformation
  by L networks, 187
  by T and pi networks, 198
  by transformer, 235
Impulse notation, 365
Indeterminate condition, 56
Indeterminate value, 56
Inductance, 140
  unit of, 143
Inductive reactance, 153
Infinitely great, 79
Insulator, 2
Internal resistance, 30
Ion, 5
Irrational number, 414

Kirchhoff's laws, 58–61

L-networks, 188
Lenz's law, 141
Linear resistance, 65, 72
Logarithms, 421
Logic network, 347
  and, or, not, nor, nand, 348
  elemental equation, 349
Loop-currents procedure,
  for ac circuits, 108, 159
  for dc circuits, 62–65

Matrix, matrices, 277–294
Mho, 66
Millman's theorem, 67
Mutual inductance, 229

Networks
  conversion, T to pi, pi to T, 196
  L-type, 188
  symmetrical T, 213
  T and pi, 190
Node, nodal point, 59
Node voltages, 73
Non-recursive processor, 379
Norton's theorem, 71

Ohm, 23

Ohm's law, 23
  for ac circuits, 104, 154

PAM, 359
Parallel circuits, 32
Passive device, 28
PCM, 360
Phase angle, 92, 169
Phase shift, 92
pi (π) networks, 193
Potential difference, 13
Power, 24
  apparent, 172
  average, 96, 106
  factor, 169
  in ac circuits, 106
  reactive, 172
  true, 171, 432
Power series, 415
Powers of ten, 405
Pythagorean theorem, 403

Q, quality factor, 179
Quadratic formula, 403
Quantization, 360

Radian, 89
Rational number, 414
Recursive processor, 379
Resistance R, 22, 136
Resistivity, 25
Resistor, 25
Resonance,
  parallel, 180
  series, 175
rms (root-mean-square), 95
  as vector quantities, 96

Sampling theorem, 359
Series circuit, 27
Series-parallel circuit, 35
Shifting theorem, 434
Sideband frequencies, 428
Signs, laws of, 402
Similar triangles, 410
Simultaneous equations, 52–57
Sinusoids as vectors, 99–105, 413
Solutions to problems, 440–550
Stability of digital processors, 383
Superposition, 11, 108
Switching algebra, 338–347
Symmetrical components, 265–272
Symmetrical-T networks, 213

Thevenin's theorem, 69
Three-phase system, 255
  balanced case, 256, 261
  unbalanced case, 265
Time constant
  L/R, 416
  RC, 417
Time delay, 423
Time rate of change, 410
T-networks, 213
Transformers, 227–229
  construction, 228
  dot-marked terminals, 230
  double-tuned, 241
  ideal iron-core, 250
  T-equivalent, 239
Transistor amplifier, 432
Trigonometric functions, 77–87
Truth table, 342
Two-port networks, 294–323
  conversion chart, 304
  interconnected, 306
  some applications, 316–323

Unit impulse, 435
Unit pulse $p(nT)$, 367
Unit-step sequence $U(nT)$, 368

Vectors, 405–409
  rms values as vectors, 96
  sinusoids as vectors, 99–105
Volt, 13
Voltage drop, 28
Voltmeter, 30

Watt, 23

z-transform, 366
  inverse, 373