

M/3A. W.F. POOLIK

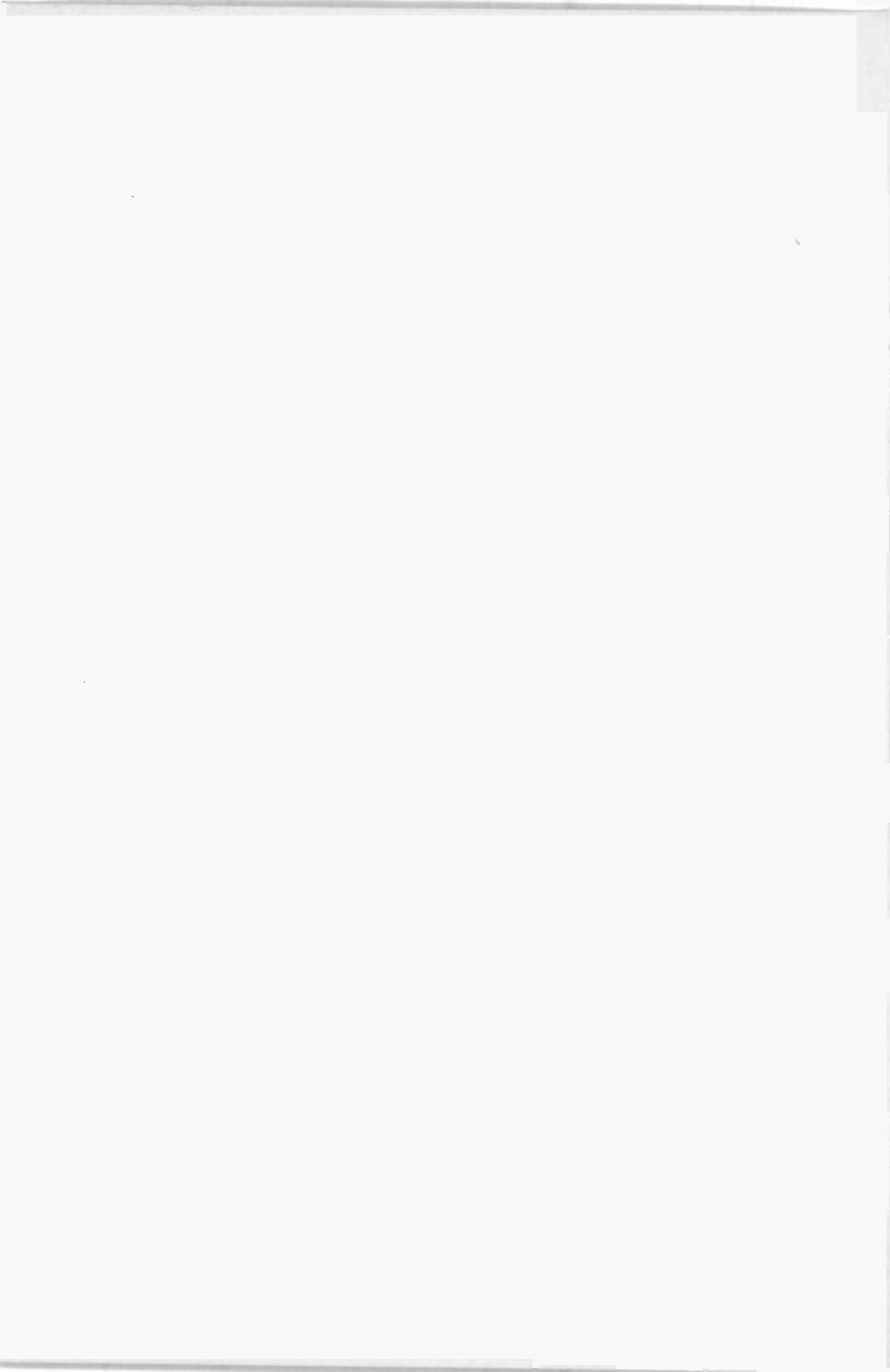
PHILCO

TECHREP DIVISION BULLETIN



1. TRANSISTOR III

SEPTEMBER
1952



PHILCO TECHREP DIVISION BULLETIN

Published Monthly by
The TechRep Division of Philco Corporation
Philadelphia, Pennsylvania

Volume II

SEPTEMBER, 1952

Number 9

CONTENTS

Editorial	2
<i>By John E. Remich</i>	
Constant-Current Voltage Regulation	3
<i>By Robert G. Nevitt</i>	
Negative Feedback and Its Application to Philco Microwave Radio Relay Equipment ..	10
<i>By Gail W. Woodward</i>	
What's Your Answer?	16
Introduction to Transistor Electronics (Installment III)	17
<i>By John Buchanan</i>	
Solution to Last Month's "What's Your Answer?"	Inside Back Cover

Editor
John E. Remich

Managing Editor
Robert L. Gish

Technical Editors
Frederick S. Mohr
Francis R. Sherman
Gail W. Woodward

EDITORIAL OFFICE:
PHILCO CORPORATION
TECHREP DIVISION
22nd and LEHIGH AVENUE
PHILADELPHIA 32, PA.

If any information contained herein
conflicts with a technical order, man-
ual, or other official publication of
the U.S. Armed Forces, the informa-
tion in the official publication should
be used.

Editorial

Training—The Economical Way to Increase the Effectiveness of Electronic Equipment

By John E. Remich, Manager, Technical Department

Much has been said and written about the vast quantities of money, time, and effort being expended to improve the characteristics of our electronic equipment. This expenditure is certainly justified, if America is to hold her place as the leader in technical progress. However, improvement in the design of equipment without proper training in its use and operation, does not allow the improved equipment to be utilized to its full effectiveness, and many times the results fall short of expectations.

It has been adequately proved that few electronic equipments are operated at top efficiency at all times—a survey described in Volume I of the MIT Radiation Laboratory Series, shows that in one instance, where the performance of approximately 100 different radar sets was carefully measured with test equipment of known accuracy, the tests revealed that *on the average* the maximum effective range of the sets under test was only one-half the maximum range possible, had the equipment been operating at peak efficiency. In fact, five of the sets were found to be operating at less than 10% of their possible maximum range, which means, in effect, that these radars were protecting only 1% of their assigned tactical areas. (Yet, in each case, the set under test was thought to be in normal operating condition by the maintenance personnel involved.)

These tests were made some time ago, of course, and the electronics maintenance situation is much improved since that time; however, the survey does point out the need for improvement in operating efficiency—a need which can best be alleviated by adequate training of maintenance personnel.

It is well known that military electronic equipment particularly is designed to continue to perform under even extremely undesirable operating conditions. However, the use which can be made of any piece of electronic equipment is a direct function of the state of operating efficiency of that equipment, and we sincerely believe that the technical training of the electronics maintenance man ranks equally in importance with the potentialities of the equipment for which he is responsible.

CONSTANT-CURRENT VOLTAGE REGULATION

By Robert G. Nevitt
Philco Field Engineer

How constant-current devices can be used for regulation of high-voltage, low-current, power supplies.

CONVENTIONAL voltage regulators are not readily adaptable to high-voltage circuits—especially those where the current requirement is small. The regulator shown in figure 1 is good for low-voltage, low-current requirements, but it becomes impractical for high-voltage, low-current applications. Neon tubes, V-R tubes, or corona tubes can be connected in series to obtain high-voltage regulation, but the bleeder current flowing through these tubes is often large in comparison to the current drawn by the load; therefore the design would be uneconomical.

Figure 2A is the schematic of a conventional electronic voltage-regulated supply. One of the disadvantages of this type of regulator is that the variation in voltage at point B (the grid of the amplifier tube) is only a fractional part of the total variation which exists between point A and ground. This difficulty can be minimized by replacing the voltage divider with the circuit shown in figure 2B.^o Practically the total voltage variation is applied to the grid of the amplifier tube, but again we are confronted with the need for a constant-voltage device such as a voltage-regulator tube. A number of regulator tubes would have to be connected in series if this circuit were used for a high-voltage supply, and the current drain through these tubes would result in heavy loading.

For a given percentage of voltage

^oBousquet, A. B., *Electronics*, July, 1938, p. 26.

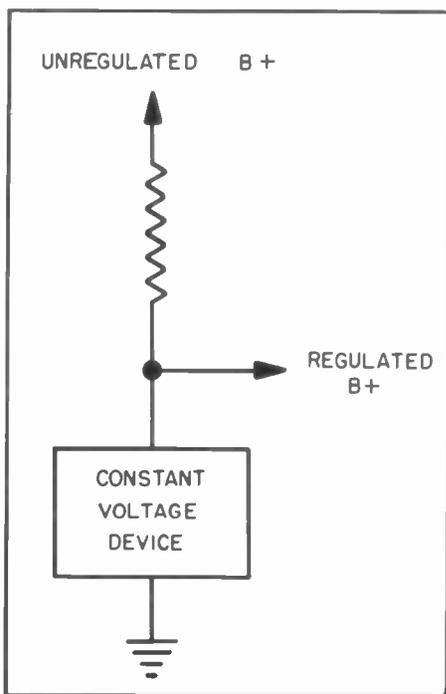


Figure 1. Constant-Voltage Regulator Circuit

regulation, the variations in output voltage for a high-voltage supply will be much larger than they would be for a low-voltage supply. Thus, it can be readily seen that if the total variation in the output voltage of a high-voltage supply could be fed back to the control tube, sufficient regulation could be obtained for most purposes without the need of an amplifier.

CONSTANT-CURRENT ACTION

If a constant-current device is wired in series with a resistor (figure 2C), the voltage drop across the resistor

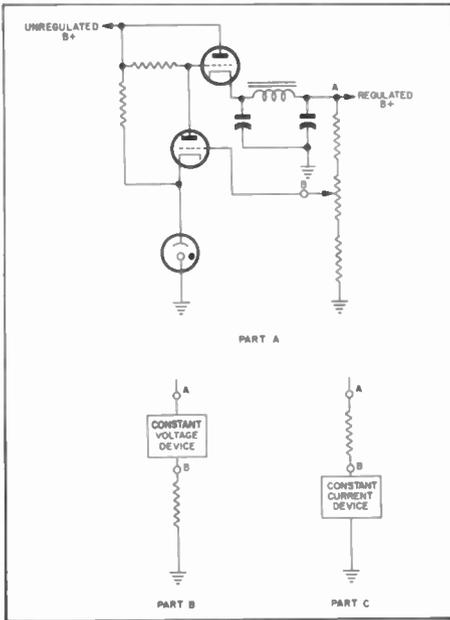


Figure 2. Part A—Typical Voltage-Regulator Circuit
Part B—Bleeder with Constant-Voltage Device
Part C—Bleeder with Constant-Current Device

will tend to remain constant even though the voltage between point A and ground varies. Any variation in voltage will appear across the terminals of the constant-current device (point B to ground). To insure that mainly the *changes* in voltage appear between point B and ground, it is necessary that the constant-current device have a low d-c resistance and a high a-c resistance in comparison with the resistor. The action of the circuit in figure 2C can be improved if the resistor is of the thyrite type (silicon carbide), since thyrite tends to act as a constant-voltage element which would further assist in causing all of the voltage variation to appear across the constant-current element.

One common constant-current device is the pentode tube. From the E_p-I_p curves for a pentode, it can be seen that as the plate voltage is increased, with all other voltages held

constant, the plate current remains essentially constant. However, the triode tube, which is normally considered to be a constant-voltage device, can be made to act as a constant-current device if the effective a-c plate resistance is made very large in comparison with the load resistor, by use of current-type negative feedback. The easiest way to obtain current-type negative feedback is to use an unbypassed cathode resistor as shown in figure 3A. The changes in plate voltage (represented by the a-c generator) produce changes in the bias voltage which appear, when amplified, as a series generator in the plate circuit. The equivalent circuit (figure 3B) shows the series components which will develop the following voltages:

1. The plate resistance will develop a voltage, $E_p = I_p r_p$. (1)

where:

$$I_p = \text{a-c plate current}$$

$$r_p = \text{a-c plate resistance}$$

2. The cathode resistor will develop a voltage, $E_c = I_p R_c$. (2)

where:

$$R_c = \text{cathode-bias resistance}$$

3. The tube will amplify the cathode voltage, thus producing an opposing voltage, $-\mu E_c = \mu I_p R_c$.

Since these three voltages are in series, the total value of a-c plate voltage will be:

$$E_p = I_p r_p + I_p R_c + \mu I_p R_c \quad (3)$$

This reduces to:

$$E_p = I_p (r_p + R_c + \mu R_c) \quad (4)$$

The total opposition represented by the tube's circuit would then be:

$$r'_p = \frac{E_p}{I_p} = r_p + R_c + \mu R_c \quad (5)$$

This reduces to:

$$r'_p = r_p + R_c(\mu + 1) \quad (6)$$

Thus the effective a-c plate resistance is determined by the plate resistance, by the μ of the tube, and by the cathode resistance.

If the 7B6 (high- μ triode section) is taken as an example, as shown in figure 4, the following conditions may be established:

$$E_b = 100 \text{ volts}$$

(E_b = plate-to-cathode voltage)

$$E_c = -1 \text{ volt}$$

$$I_b = 0.4 \text{ ma.}$$

$$r_p = 110\text{K ohms}$$

$$\mu = 100$$

(To produce the indicated bias, a

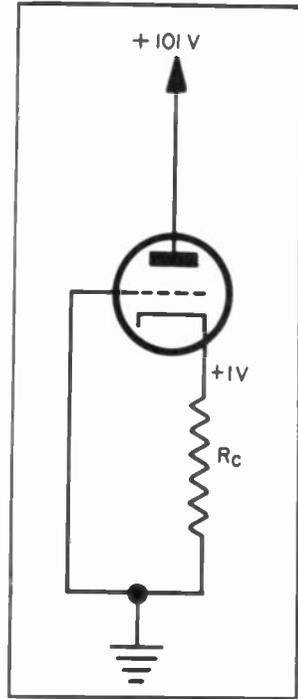


Figure 4. Constant-Current Triode Circuit. Showing Voltages

resistor of 2500 ohms must be used for R_c .)

Using equation 6, it will be found that the effective a-c plate resistance is 362.5K ohms. Since the tube passes 0.4 ma. at 100 volts, the d-c plate resistance is 252.5K ohms. Thus, the a-c plate resistance is only about 1.5 times the d-c plate resistance—this is not a sufficiently high ratio for constant-current operation.

Upon examination of equation 6, it can be seen that increasing r_p , μ , or R_c will increase r'_p . R_c is, of course, the obvious choice. Figure 5 shows the revised version of figure 4. The cathode resistor value is raised to 252.5K ohms, which results in a cathode voltage of 101 volts. To provide the same operating conditions indicated for the circuit in figure 4, the grid is operated at a 100-volt level, and the plate is operated at a 201-volt level.

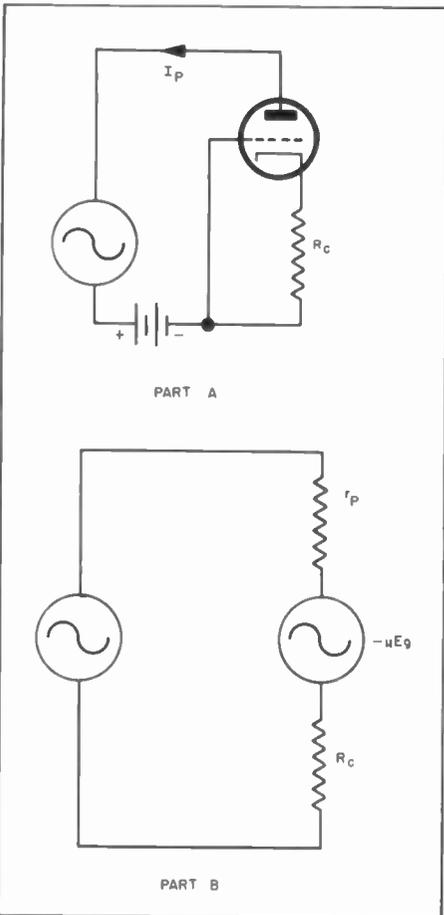


Figure 3. Part A—Triode Tube Connected for Constant-Current Operation
Part B—Equivalent Circuit of Figure 3A

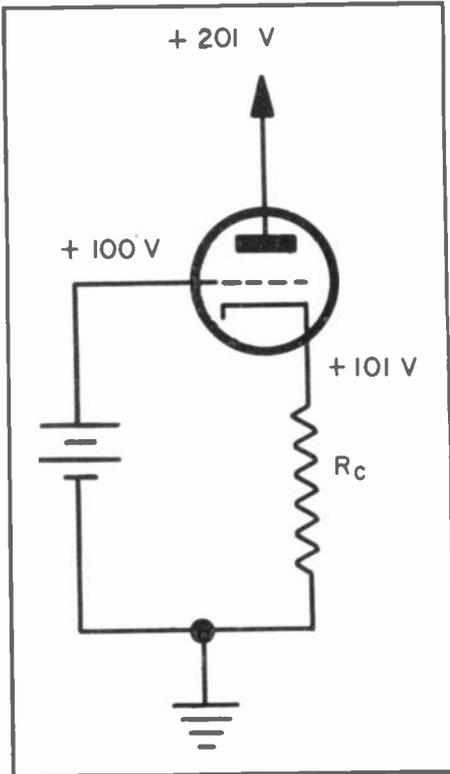


Figure 5. Modified Constant-Current Triode Circuit, Showing Voltages

Using equation 6, r'_p is found to be 25.6 megohms, whereas the d-c resistance of the circuit is now 502.5K ohms. In this case, the ratio of a.c.-to-d.c. resistance is over 50 to 1. This ratio is quite suitable for constant-current applications. (The typical ratio for a pentode is in the order of 30 to 1.)

If a still-higher ratio is desired, the circuit shown in figure 6 can be used. Here, the circuit shown in figure 5 provides the cathode resistance for a second tube in a cascode-type circuit. To use equation 6, it is merely necessary to assume a cathode resistance (R_c) for the upper tube of 25.6 megohms. The resulting a-c plate resistance for the entire circuit will then be 2586 megohms. Since the resistance to d.c. is only 752.5 ohms,

the a.c.-to-d.c. resistance ratio for the cascode circuit is over 3000 to 1.

APPLICATIONS

Figure 7 shows a typical use of the constant-current device in a d-c amplifier. Here, the coupling circuit acts as a voltage divider to d.c. but not to a.c. Thus, the signal-coupling loss associated with the resistor-type divider is eliminated.

A regulated, high-voltage power supply with a negative output is

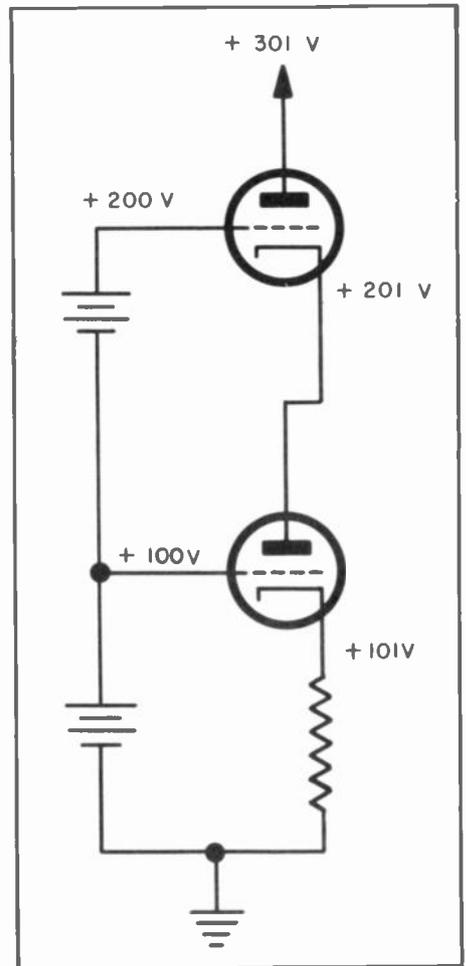


Figure 6. Cascode Constant-Current Circuit, Showing Voltages

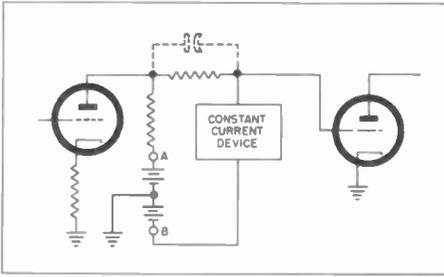


Figure 7. D-C Amplifier with Constant-Current Coupling

shown in figure 8. This circuit is similar to a conventional regulated supply except for the use of the constant-current circuit and the lack of an amplifier stage. The 6BG6G acts as the control tube, and has its operating bias established by the variable resistor R. (Changing the setting of R will also vary the value of regulated output voltage over a limited range.)

Figure 9 shows an r-f type of high-

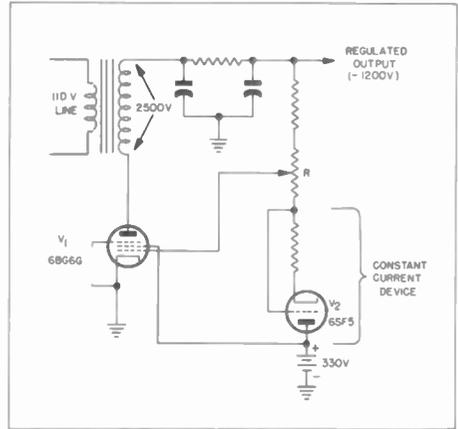


Figure 8. High-Voltage Power Supply with Constant-Current Voltage Regulation

voltage power supply with constant-current voltage regulation. In this case, the r-f oscillator bias is determined, in part, by the voltage across the constant-current device. As was the case in figure 8, this power supply has a negative output and uses no amplifier tube.

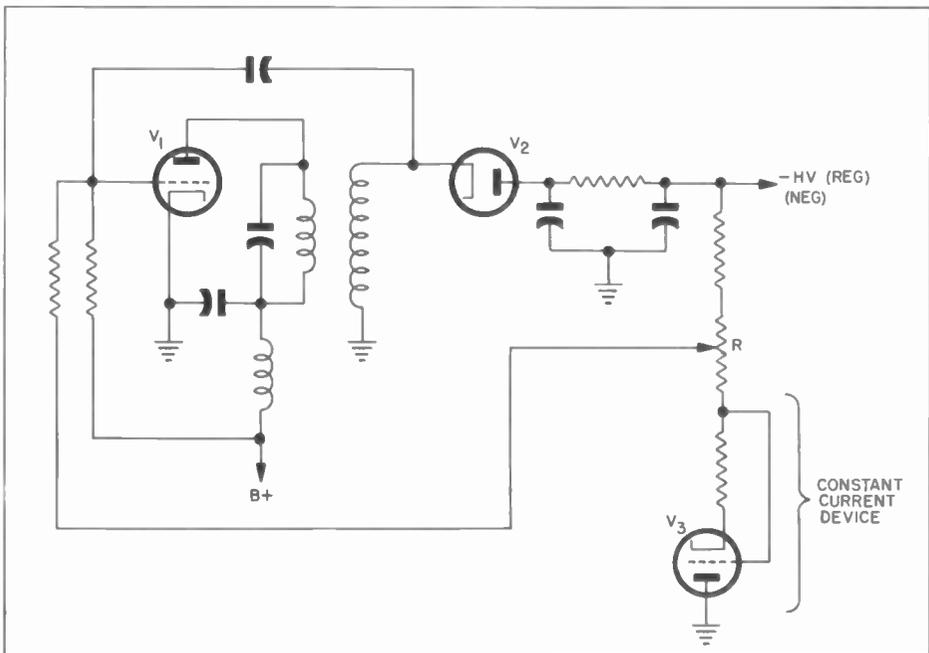


Figure 9. R-F Power Supply with Constant-Current Voltage Regulation (Negative Output)

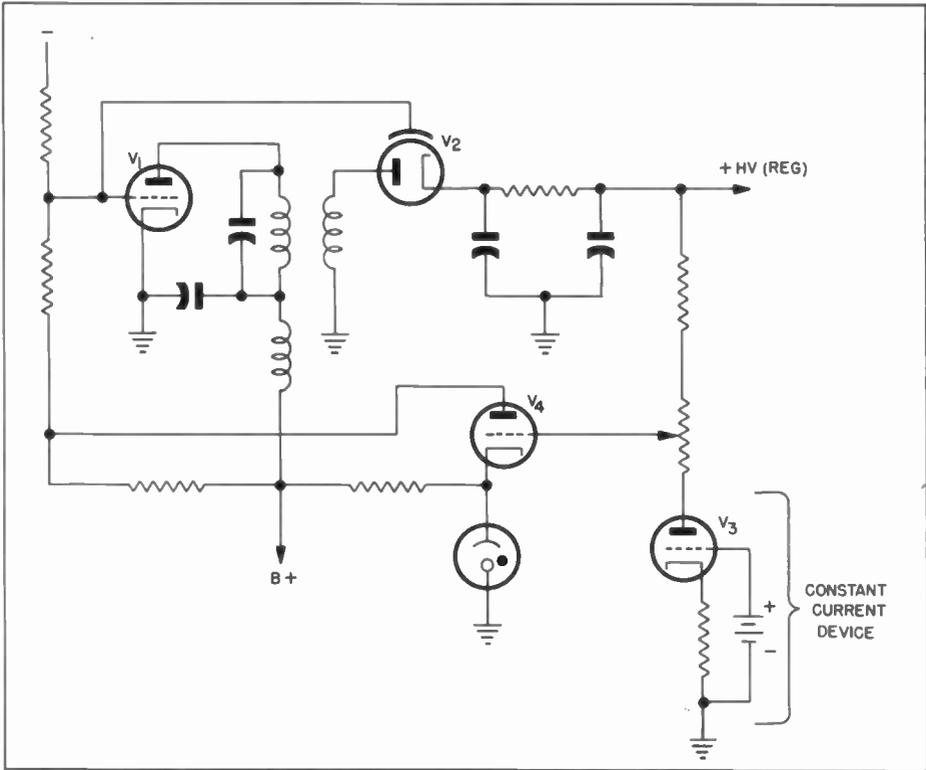


Figure 10. R-F Power Supply with Constant-Current Voltage Regulation (Positive Output)



ROBERT G. NEVITT was born on September 25, 1923, in Meyersdale, Pa. but received his early education in Louisville, Kentucky. Upon graduating from St. Xavier School in Louisville, in 1941, he attended the University of Louisville for three years during which he majored in

electrical engineering until he was called into the U.S. Navy in 1944.

In the service, he attended various electronic schools, and upon graduation was assigned to maintenance work on Navy communications equipment, overseas.

In 1946, with the War over, he returned to the University of Louisville to complete his requirements for the degree of Bachelor of Science in Electrical Engineering, which he received the following year. A position as a television instructor with the Louisville Radio School occupied him until he joined the Philco Tech-Rep Division in July, 1951. He has been working on a Strategic Air Command assignment since that time. He is a member of the A.I.E.E., the A.R.R.L., and the I.R.E.

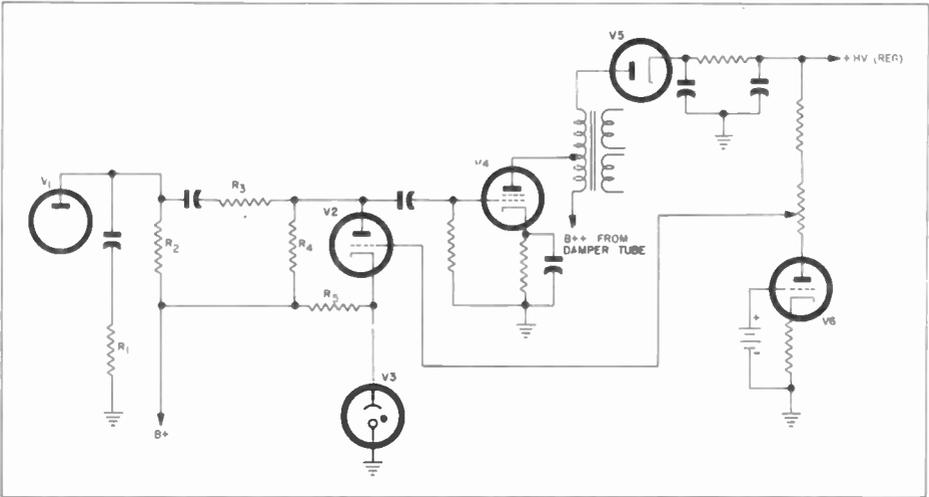


Figure 11. TV High-Voltage Power Supply with Constant-Current Voltage Regulation

Where a positive output voltage is required, the r-f power supply circuit shown in figure 10 can be used. In this case, a phase-inverting stage (V_4) must be used to provide proper control action. An interesting feature of this circuit is found in the oscillator feedback capacitor which consists of a metal ring surrounding V_2 . The capacitance existing between the ring and the plate of V_2 is sufficient to maintain oscillation.

The conventional TV, high-voltage power supply can be adapted for constant-current voltage regulation by means of the circuit shown in figure 11. The variations in high voltage appear across the constant-current device (V_4), and are used to control

the conduction of V_2 . The output of V_1 , the sweep generator, is divided by the action of R_3 and V_2 before it is applied to V_4 . If the output voltage tends to increase, V_2 becomes a better conductor (because of the more-positive grid voltage) and the drive to V_4 is thus reduced. This action will restore the output voltage to its previous value.

CONCLUSION

Constant-current voltage regulation can be applied to any high-voltage, low-current power supply. The advantages are: low current drain in the bleeder section, low d-c voltage level for the control circuit, and availability of practically the entire output voltage variation at the control tube.



NEGATIVE FEEDBACK AND ITS APPLICATION TO PHILCO MICROWAVE RADIO RELAY EQUIPMENT

By Gail W. Woodward
Headquarters Technical Staff

How the negative feedback technique is applied to the Philco CLR-6, and the advantages that result.

A MICROWAVE-RADIO-RELAY SYSTEM must, by nature, produce a tolerably small amount of distortion. It is apparent that in a system which uses a large number of repeaters, each individual repeater must have an extremely small value of distortion.

The most effective known method for reducing distortion is found in the use of negative (or inverse) feedback—the greater the feedback, the less the distortion. Aside from distortion considerations, a negative-feedback system has several other advantages which will be discussed later.

The Philco CLR-6 microwave radio relay equipment incorporates 100% negative feedback to minimize dis-

tortion. This means that the entire output signal is fed back to the input circuit, thus resulting in an exceptionally stable system.

MODULATION

The Philco CLR-6 utilizes an FM carrier for two reasons. The reflex klystrons, used at the radio frequency, are readily susceptible to direct FM, and the noise-reducing feature inherent in an FM system can provide an effective signal-to-noise-ratio improvement of about 20 db (with a reasonable modulation index).

The block diagram of a typical repeater is shown in figure 1. The r-f oscillator is frequency-modulated by the output of the FM receiver which

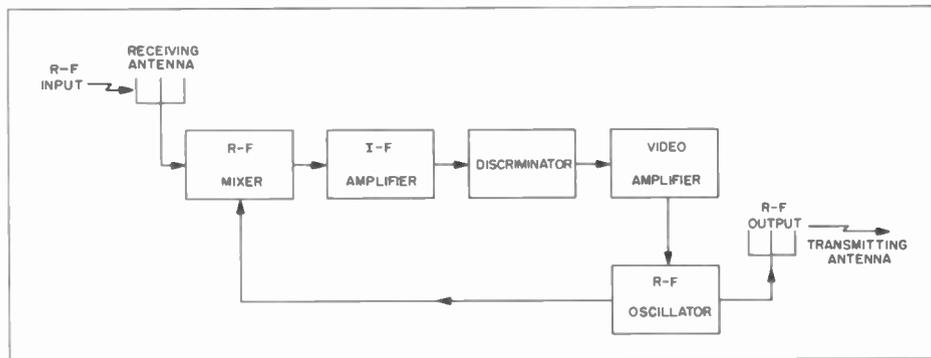


Figure 1. Block Diagram of Typical Repeater

is connected directly to the klystron repeller. A small portion of the r-f output is fed to the receiver mixer—thus, the r-f oscillator also serves as the local oscillator for the receiver r-f sections. Of course, the frequency of r-f output must differ from the frequency of the r-f input by the value of the i.f. (90 mc.). Successive repeaters alternately raise and lower the r-f signal frequency so that only two carrier values are encountered in a single system. It can be seen that a frequency-modulated input must result in a frequency-modulated output.

FEEDBACK

The feedback loop operates as follows: The frequency-modulated, r-f input beats in the mixer, with the injection signal from the r-f oscillator, thus producing a frequency-modulated, i-f signal which appears at the discriminator output as a varying voltage (this voltage is a reproduction of the original intelligence). This signal is amplified by the video amplifier and fed to the repeller plate of the r-f oscillator which then reproduces a frequency-modulated, r-f output signal. The discriminator diodes are connected so that the r-f oscillator modulating voltage will swing the r-f output frequency in the direction of the r-f input modulation. Thus, as far as the i-f signal is concerned, the modulation of the r-f oscillator will tend to cancel the modulation of the r-f input—this means negative feedback in terms of modulation. The feedback is considered to be 100% because the entire value of output deviation acts upon the r-f input deviation. In the CLR-6, an r-f input signal deviation of 3 mc. will result in an i-f signal deviation of only 30 kc. (or 1%), and an r-f output signal deviation of 2.97 mc.

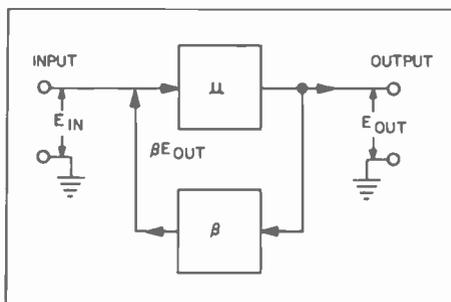


Figure 2. Block Diagram of Basic Feedback System

If the r-f oscillator frequency were to vary in exact accordance with the changes in r-f input frequency, the resultant i.f. would have zero deviation. Actually, the modulation of the r-f oscillator must be somewhat less than that of the r-f input signal; thus, it is obvious that the repeater gain (in terms of modulation) *must be slightly less than 1*. For the CLR-6, the loss in modulation amplitude is about 1% in a single repeater. If 10 repeaters were used, the overall loss would be about 9.5%, and with 100 repeaters, a 63% loss would occur. It is evident that if the gain of individual repeaters were appreciably less than unity, a large number of repeaters would be impossible.

ANALYSIS OF FEEDBACK

Since the feedback loop encompasses the entire repeater (except those r-f components which precede the r-f mixer), it can be treated mathematically as a single amplifier.

The basic feedback system is shown in figure 2, and:

$$E_{out} = \mu E_{in} - \mu \beta E_{out}$$

where:

- E_{out} = output signal voltage
- E_{in} = input signal voltage
- μ = amplifier gain
- β = feedback in terms of fraction of output fed back to input.

Simplifying:

$$E_{out}(1 + \mu\beta) = \mu E_{in}$$

or:

$$\frac{E_{out}}{E_{in}} = \frac{\mu}{1 + \mu\beta} \quad (1)$$

Since the entire output is applied back to the input circuit (100% feedback), β is unity, and equation 1 simplifies to:

$$\frac{E_{out}}{E_{in}} = \frac{\mu}{1 + \mu} \quad (2)$$

This equation is called the *basic transmission equation*. It should be remembered that E_{out} and E_{in} are the modulation values, not the carrier levels, and that μ is the repeater gain in terms of modulation.

Upon examination of equation 2, it can be seen that in order for $\frac{E_{out}}{E_{in}}$ to approach unity, μ must be large—if μ is large, the feedback is correspondingly large.

FREQUENCY-RESPONSE LIMIT AND GAIN

In any amplifier system, a phase shift is encountered as the upper limit of the bandpass is approached. This phase shift is cumulative, and in a multiple-stage repeater can easily approach 180 degrees. If the repeater using 100% inverse feedback suffers an additional 180-degree phase shift, the feedback becomes positive, and oscillation can occur. Here, then, is a very important limiting factor—the gain of an amplifier to be used with 100% inverse feedback must be less than unity without feedback at the frequency for which the phase shift is 180 degrees (this is called the Nyquist criterion).

The repeater shown in figure 1 involves both AM and FM components.

The FM components include the r-f oscillator and the discriminator (the i-f section merely has sufficient gain to produce the required limiting action), whereas the video amplifier is the AM component. The gain of the amplifier section is expressed as:

$$\mu = \left[\frac{A_{osc}}{A_{disc}} \right] G \quad (3)$$

where:

- G = video-amplifier gain
- A_{osc} = r-f oscillator modulation sensitivity in terms of megacycles per volt (about 0.35 for the CLR-6)
- A_{disc} = discriminator sensitivity in terms of megacycles per volt (at maximum i-f output this is about 0.1)

By virtue of the extreme overall bandwidth (10 mc.) the A terms are not dependent upon the modulating frequency—this leaves the video-amplifier circuits as the determining factor in the variation of μ with modulating frequency.

Using typical components, the overall phase shift approaches 180 degrees at about 2 mc. Therefore, at this frequency, the μ factor must be less than unity. At the lower frequencies (mid-audio range), the μ factor must be made as large as possible to obtain a repeater gain of almost unity (equation 2). In the CLR-6, the low-frequency gain (μ) is made 40 db, which means essentially 40 db of feedback (100%). The video amplifier has a cutoff frequency of 4 kc. (which means a gain of 37 db at this frequency) and a 6-db-per-octave decrease in gain as the frequency is increased (see figure 3). Thus, at 300 kc., the gain (and hence the feedback) is 0 db. At the point of 180-degree phase shift, the gain is sufficiently low that no instability is encountered.

It may be questioned as to how a response of 300 kc. can be obtained

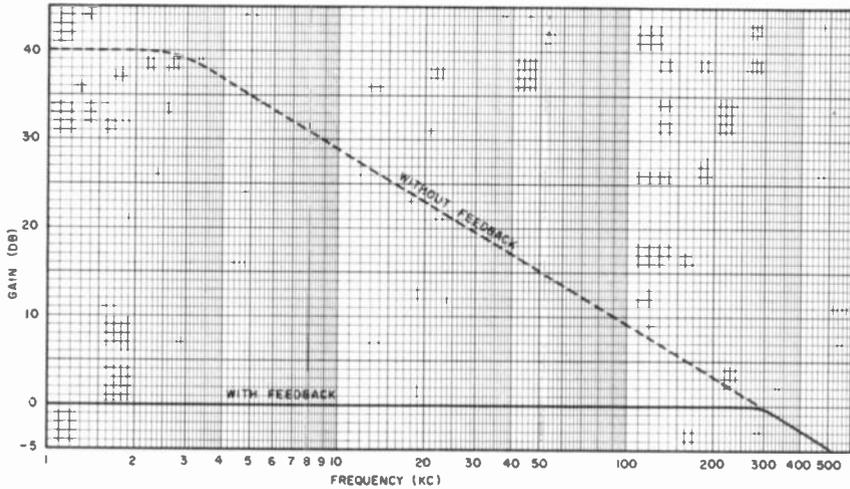


Figure 3. Graph of Frequency Response of CLR-6, Showing Effect of Feedback

with a 4-kc. amplifier. Actually, at the lower frequencies, the gain is suppressed to slightly less than unity by means of inverse feedback. As the gain falls off at higher frequencies, the feedback signal reduces in the same proportion, thus keeping the actual gain at slightly less than unity. This action continues to a point where the feedback is 0 db—at progressively higher frequencies, the flat response can no longer be held by means of feedback, and the response curve therefore falls off. This action is shown graphically in figure 3. The solid line shows the response with feedback, while the dotted line shows the response without feedback.

A simplified schematic of the CLR-6 is shown in figure 4. The two-stage video amplifier (which determines the overall frequency response) is broadbanded except for the plate-circuit components of the second stage—thus, the frequency response is mainly determined by the time constant of R_F and C_F . The value of feedback can be varied (and therefore, to a limited degree, the frequency response) by varying the size of R_G , which is field adjusted so that

a particular repeater will meet the specified bandwidth requirements.

In the discussion of repeater gain and frequency-response characteristics, it should be noted that the expressions apply only to the modulation. R-f gain and r-f signal level represent an entirely different problem. It has been found that a 15-db signal-to-noise ratio is adequate for communication. In practice, Philco systems incorporate a 30-db fading margin, which means that under normal conditions (no fade), the r-f signal level is at least 45 db above the receiver noise level.

Another point of interest in the CLR-6 is the use of the cascode input amplifier. This circuit has the gain of a pentode with the noise figure of a triode. Since this circuit is the major factor in noise-figure determination, the use of the cascode circuit provides a very desirable improvement in noise figure.

SERVO SYSTEM

The servo system shown in figure 5 forms a second, minor, feedback loop within the CLR-6. This circuit

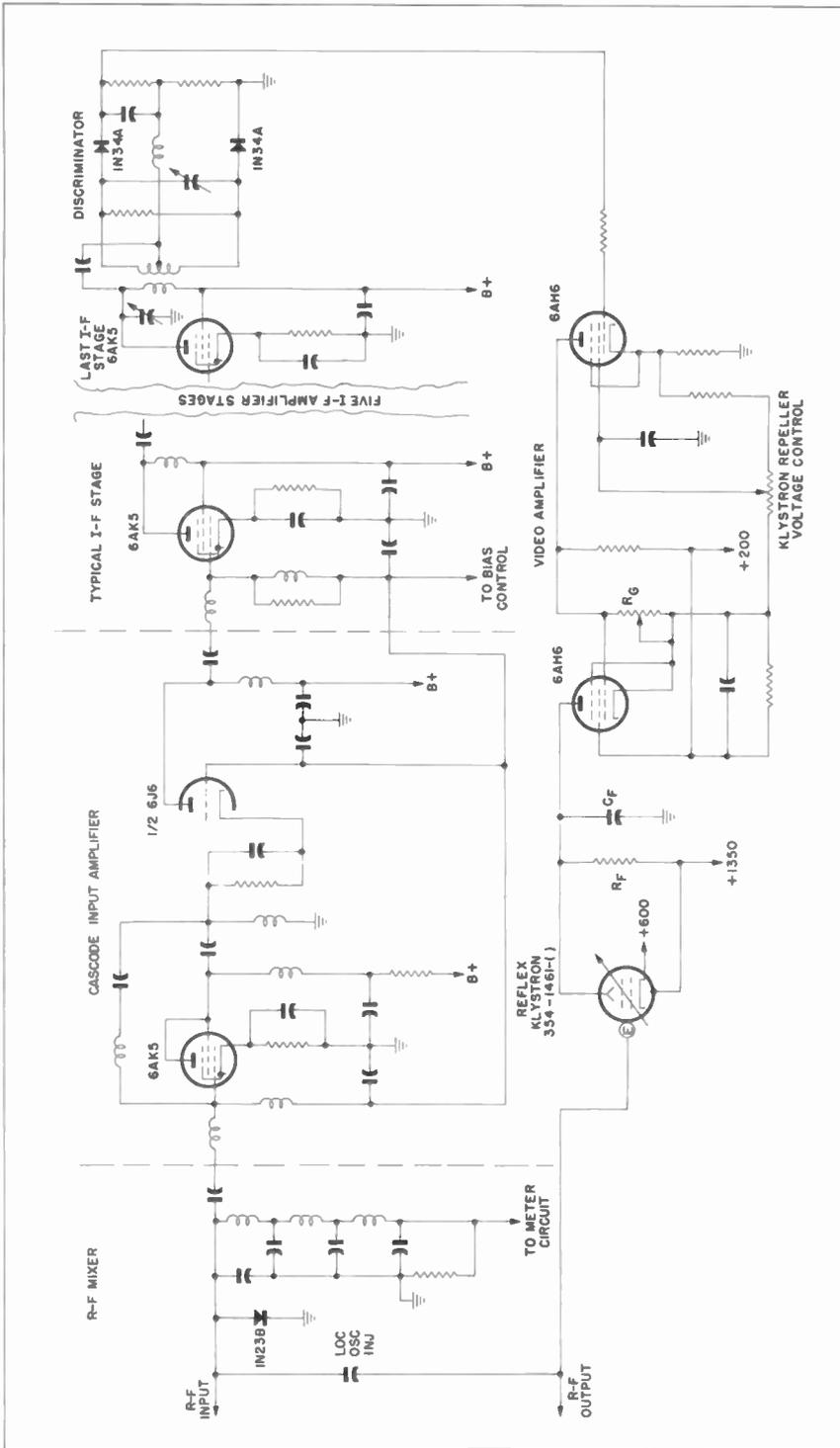


Figure 4. Simplified Schematic of CLR-6

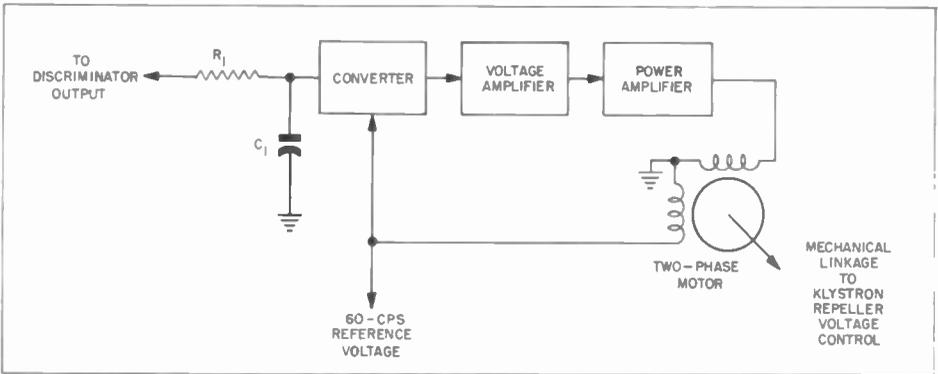


Figure 5. Block Diagram of Servo-Amplifier Feedback Loop

bridges the video-amplifier system, as far as frequency control is concerned. R_1 and C_1 act as an integrator which allows only the d-c component of the discriminator output to be applied to the servo system. Of course, the d-c component will be present only when the r-f oscillator is not exactly on frequency (average conditions). The d.c. applied to the converter is converted to a 60-c.p.s. pulsating signal which has a reference phase determined by the reference voltage which actuates the converter. The converter output is amplified and used to drive one phase winding of a two-phase motor (the other phase winding is supplied by the reference voltage). Thus, if the average discriminator output is positive, the two-phase motor will turn in one direction, while a negative output will cause reversed rotation. Of course, if the r-f oscillator is on frequency, the average discriminator output is zero, and the motor will not be excited. The motor is mechanically linked to the klystron-repeller-voltage control (see figure 4) in such a manner that the motor will run until the average discriminator output is zero, which means that it runs until the r-f oscillator is on frequency.

The importance of this automatic-

frequency-control function is obvious when one considers the number of components in the video amplifier. If any component were to change value, the d-c output level would change and the r-f oscillator would be driven off frequency. The servo amplifier has no critical parts because it functions at the 60-c.p.s. rate. Thus, the critical a-f-c function is handled in a very stable manner, and each repeater will follow any drift tendency that might originate at the insertion terminal.

Late models of the CLR-6 are using a simplified type of servo system. The converter and amplifiers have been replaced by a relay system consisting of a sensitive, meter-type relay and a reset relay. The sensitive relay has its armature centered between two contacts which switch the power required to drive a small reversible motor (from this point on, the action is as was described above). This relay is actuated by the discriminator output—when the average d-c output from the discriminator exceeds 0.05 volt, the relay will swing to one side or the other depending upon the polarity of the voltage. This causes the motor to run in a direction that will tune the klystron for the proper i.f. After a suitable time interval, as determined by an RC time

constant, the second relay terminates motor rotation and resets the sensitive relay. If balance is not restored, the system will recycle the required number of times until the discriminator average output is less than 0.05 volt.

CONCLUSION

The feedback actions which have been described provide several major advantages not found in ordinary systems. These advantages are:

1. Each repeater has an extremely low value of distortion; thus, a large number of repeaters can be cascaded.
2. The large value of feedback tends to minimize microphonic tendencies of any stage in the system to a point where they present no problem at all.
3. By the same reasoning, noise produced in an individual stage will also be suppressed.
4. The frequency response of any individual repeater can be precisely adjusted by a slight variation in the value of feedback.
5. The servo-amplifier feedback system provides rigid control of r-f output frequency in terms of r-f input frequency.



ERRATA

July issue, last paragraph on page 31 should read:
"A new 2C39 should exhibit 18,000 μ mhos, . . ."
Same issue, page 32—The byline should read:
"LCDR Weber, Electronics Officer, USS Sicily, and
O. E. Adams, Philco Field Engineer."

What's Your Answer?

Here is a practical problem as submitted by George A. Brown, Philco Field Engineer.

A 100-wire armored cable has been laid across the bottom of a river. A cableman is confronted with the task of identifying and labeling each wire at both ends of the cable. However, he has only a ringer (battery and buzzer), and he is allowed only one round trip across the river. How did he do it?

(Solution next month)

INTRODUCTION TO TRANSISTOR ELECTRONICS

Installment III Classical Background for the Free-Electron View

By **John Buchanan**
Technical Publications Dept.

A continuation of the discussion of the methods and theories of classical physics as applied to systems of ideal particles.

SYNOPSIS

This is the third of a series of articles covering the fundamentals of transistor electronics. The first article of the series, which appeared in the January, 1952, issue of the BULLETIN, covered the atomic view of solids. This view maintains the particle nature of the electron, but introduces the quantum to restrict the electron's motion within the atom to definite shells, or energy levels. Each shell, in turn, permits a finite number of planetary orbits, or electron states, at the shell's respective energy level. Outside an atom, an electron is assumed to be completely free to move about with random thermal motions. Electronic conduction is visualized

as a drift of free electrons, or as a drift of empty electron energy states (electron holes), or as a combination of the two.

The second article, published in the April issue, laid the groundwork for the more involved free-electron view of solids. The first elements of the kinetic theory of gases were presented as a point for beginning the transition from the classical laws of ideal gases to the wave-mechanical laws of electron gases.

The current article continues the discussion of the methods and theories of classical physics as applied to systems of ideal particles.

NEW CONCEPTS ON THE HORIZON

Four concepts, each peculiar to our normal habits, will meet as one in the wave mechanics of free electrons. The first involves the loss of individuality among the electrons—instead of thinking of electrons as particles, we will think of them as wave functions. The second way of thinking, however, recovers somewhat the identity of an electron, by permitting us to think of it as a particle in an imaginary

phase space. The third concept has the broader and more subtle implications of a philosophical point of view, for it involves the entrance of chance in the scenes normally played by mechanical cause and effect. The fourth turn in our thinking will be an unbending of the mind sufficient to accept the electron simultaneously as both a particle and a wave, or at least as a quantity which can exhibit either type of property.

Except for the fourth, these modes of thought do not mark a drastic break with the concepts of classical mechanics. Rather they are essentially extensions of the methods and points of view arising from the earlier approaches to the problems of statistical mechanics. However, inasmuch as these concepts may be new to us, their qualitative meanings may float about unanchored unless we retrace their lineage to the concrete particles of common-sense physics.

This problem does not arise in the atomic view of solids, for we are never called upon to abandon the electron as a particle, and without the clumsiness of special luggage, we are immediately free to follow the electron's escapades with the quantum. To be sure, these patterns of particle behavior answer only to mathematical equations, but there is always a little body along to act out the part on our mental stage. Not so, in the free-electron view—here the particle disappears, and the function must go it alone as something of a sound effect off stage. Reason, however, revolts at accepting variable physical properties without a fixture of some sort to pin them upon. On this account we risk an overemphasis of the concept of *degree of freedom*, for when the individual particles disappear, these degrees of freedom, acting as stand-ins, can mark their places on the stage. They too, may seem to vanish, but in one costume or another they can always be discovered.

It is thus that our present discussion will center upon the transition from thinking in terms of particles to thinking in terms of degrees of freedom. In the process, the degrees of freedom will lead us along a path through various spaces, including phase space, that are inhabited by ideal molecules, and will introduce

us to those functions, including wave functions, that govern the behavior of the molecules. At the end of the journey, a practical application of the concept of degrees of freedom is attempted, utilizing free electrons on the assumption that the atomic theory of solids is correct. However, when this is done, the degrees of freedom apparently vanish, and instead we find the quantum in the wings, waiting for his cue to make a triumphant entrance once again. At this point the present article ends.

STARTING OUT WITH A DEGREE OF FREEDOM

It seems that the term *degree of freedom* was originally adopted to express a functional characteristic in the field of statistics. In determining a statistical estimate, the expression "number of degrees of freedom" means the number of variables which are independent of each other in a random sampling. For example, if a deck of cards were to be dealt randomly into four equal piles, three of the piles could be partially independent of each other; however, no choice arises for the fourth pile—it must consist of the remaining cards. Suppose we are tempted to test for the presence of psychic influences on the deals, and thus wish to compare the empirical with the theoretical probabilities. If it is desired to estimate the expectation of a particular distribution, say, of the aces and kings, by observing a large number of random deals, then the distribution only for three of the piles needs to be considered. Out of four variables, only those three that are independent, or free, enter into the statistical estimate. Thus, the distribution-probability function can be said to have three degrees of freedom.

We now borrow this term, and apply it to all mathematical func-

tions.* In the example above, note that the degrees of freedom have a twofold meaning.

Not only do they define the number of independent variables, but, equally, they define the *minimum* number of variables that must be determined before a particular distribution of the cards can be known. Each of these interpretations implies the other. Thus, when *degrees of freedom* is used as a general term for describing any function, it refers to the number of independent variables the function contains, and likewise the minimum number of conditions necessary to give the function a fixed value.

Suppose a function is equal to $x + y$, where x and y represent quantities of atoms, amperes, apples, or anything. If x and y are independent of each other, in any way, then the function has two degrees of freedom in which to vary. On the other hand, if y is constant, or is, itself, a function of x alone, then there is but one degree of freedom. For instance, if $y = x + 5$, then $x + y = x + x + 5 = 2x + 5$, which obviously has only the one independent variable, x .

Now that we have taken the "degree of freedom" from its native field of statistics, and placed it in the service of independent variables everywhere, the next step is to see how it performs its duties in the special case of geometrical functions.

FINDING POSITIONS WITH A DEGREE OF FREEDOM

Let us suppose that the particular function is to define the position of a point. In all cases, a convenient point of origin must be chosen as a zero reference. With the origin fixed, a system of measuring must be agreed upon that will give both the distance and the direction of a point from the origin. If a point is restricted to a given line (straight or curved), only one measurement is necessary, and thus there is only one degree of freedom. If a point is restricted to a given surface (plane or curved), two measurements are necessary, and thus there are two degrees of freedom. And for a point in space, three measurements are necessary, so that the position function for such a point has three degrees of freedom.

Each measurement, as well as each predetermined dimension, is called a *coordinate*, and the point is defined when all its coordinates are stated in a conventional sequence—e.g., point (4, 5, -3).

Theoretically, there are an unlimited number of systems for measuring a position in space. We shall discuss the two most commonly used. One employs the rectangular coordinates, and the other uses polar coordinates. In the rectangular system, the measurements are made along three perpendicular axes, called the X, Y, and Z axes, which represent

* The word *function*, when used mathematically, has a way of cloaking its meaning in mysterious shrouds. It simply means a quantity which can be determined when definite values are assigned to certain other quantities. These other quantities are the independent variables of the function. Now the variables do not necessarily cause, *in a physical sense*, the function to take on a specific value. As likely as not they may be the effects of the function itself, or of totally external

causes. For instance, the Ohm's Law equation, $R = V/I$, reveals R as a resistance function of V and I . This fact can be symbolized in a general way as $R(V, I)$. Physically, however, we consider the resistance, R , to be a property of the conductor itself, and normally not actually dependent upon the voltage or current. On the other hand, the current function, $I(V, R)$, equal to V/R , is usually considered dependent physically, as well as mathematically, upon V and R .

the three dimensions of space. If a rectangular box is drawn with its edges parallel to the three axes as shown in figure 1C, so that the origin is at one corner and point P is at the corner diagonally opposite, then the coordinates that define the point are the lengths and directions of the three dimensions of the box. The distance, r , of the point from the origin is considered positive, and is equal to $\sqrt{x^2 + y^2 + z^2}$, where x , y , and z are the coordinates of the point. (In figure 1A, y and z are fixed at zero, and in figure 1B, z is zero.) Actually, r involves only one degree of freedom, since it can be determined by one measurement as well as any of its coordinates. The additional degrees of freedom are required to fix the direction. This fact becomes more apparent on examination of the system of polar coordinates, as shown in figure 2.

The polar system uses the X, Y, and Z axes, but only as reference lines for measuring of directions and not distances. The polar r , like the rectangular r , represents the distance from the origin to the point. It differs, however, by being a primary coordinate itself, with both magnitude and sign,^o rather than simply an undirected resultant of the other coordinates. In figure 2A, r is constant, and the only degree of freedom is the direction, represented by angle θ which r makes with the positive X axis. The point is defined, however, by both the constant coordinate and the variable coordinate; hence P (10, $\pi/4$), or P (10, 45°), depending upon whether θ is measured in radians or degrees, respectively. If θ were predetermined, and only r were per-

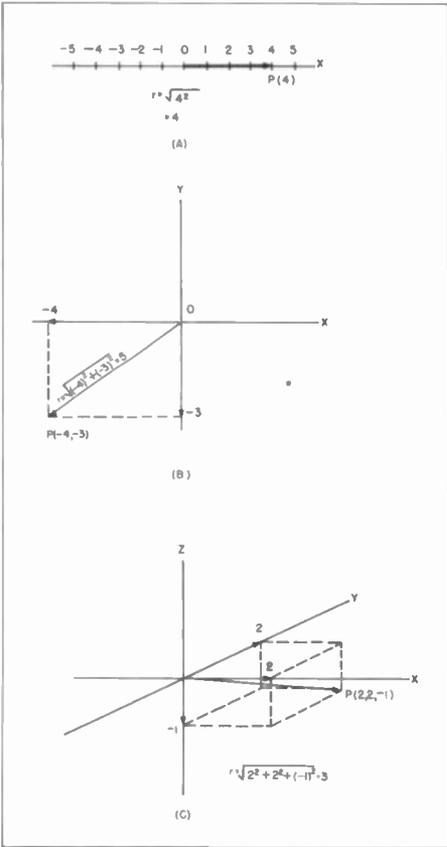


Figure 1. Rectangular-Coordinate System for Measuring Positions

mitted to vary, then the polar system would be identical with the rectangular system of figure 1A. Both r and θ are variable in figure 2B, giving two degrees of freedom to the position function in the XY plane. Figure 2C shows the polar, or, more specifically, the spherical, coordinates for measuring a point position in space. Each of the three variables, r , θ , and ϕ , are degrees of freedom for the position function. The measurement of ϕ defines the angle from the positive Z axis to r .

Note that θ still applies to the XY plane, but is now the angle from the positive X axis to the projection of r on the XY plane. In both the rectangular and polar systems of co-

^o The polar r , called the *radius vector*, may assume positive or negative values. A negative r would have the same magnitude as its positive value, but an opposite direction.

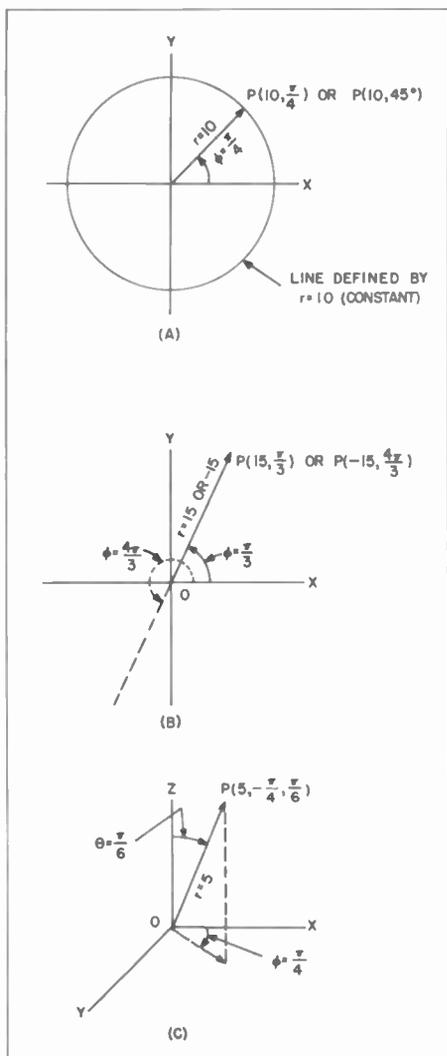


Figure 2. Polar-Cordinate System for Measuring Positions

ordinates, point P is completely defined by the vector OP (not to be confused with r , which gives the distance but not the angular direction).

So far, the degrees of freedom of a single point have been our concern. If the position function of a whole system of points, representing, say, the molecules of a gas at a given instant, is desired, three coordinates will be required for each point. As in the case of the single point, the number of degrees of freedom will equal the total number of independ-

ent variables. Mathematically, the entire system may be treated as a single position. Thus, the function $q_x(x_1, x_2, x_3, \dots, x_N)$, $q_y(y_1, y_2, y_3, \dots, y_N)$, $q_z(z_1, z_2, z_3, \dots, z_N)$, would define the position of a system of N molecules in three-dimensional space. The rectangular coordinates (x_1, y_1, z_1) are those for the first molecule, and so on, to (x_N, y_N, z_N) for the N th molecule. The terms $q_x(x_1, \dots, x_N)$, $q_y(y_1, \dots, y_N)$, and $q_z(z_1, \dots, z_N)$, may be abbreviated to q_x, q_y , and q_z , which symbolizes the three position coordinates of the system as a whole. If we assume that each molecule has three independent coordinates, then the system position function has $3N$ degrees of freedom. If $N = 1$, then the position coordinates reduce to those of figure 1C.

IN STRANGE SPACES WITH A DEGREE OF FREEDOM

The degree of freedom slips with ease from the position functions to those that define the behavior of moving bodies. Fundamentally, its role does not change, but its costumes of motion are attractive to the mind's eye, for these permit a description in physical terms.

Suppose we wished to define the velocity of a moving point. The methods for representing velocity are identical with those for representing position; in fact, a given velocity may actually be conceived as a fixed point in an imaginary *velocity space*. For example, figure 1 would illustrate velocity if the position axes (X, Y, Z) were changed to represent velocity axes ($\dot{X}, \dot{Y}, \dot{Z}$), ° and the position coordinates (x, y, z) were changed

° The dot notation is commonly used to represent a change per interval of time. For example, \dot{x} means a change of position in the X direction per change of time, and \ddot{x} means a change in \dot{x} per change of time.

to velocity coordinates (\dot{x} , \dot{y} , \dot{z}). The origin in this case would represent zero velocity, rather than a reference point in geometrical space. The velocity vector, OV , or simply V , would point in the direction of motion, and its length, $c = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$, would be the speed. In this manner we consider a velocity function, $V_1(\dot{x}, \dot{y}, \dot{z})$, as defining a point in velocity space. Similarly, the acceleration of a point, which is the change of its velocity per change in time, can be treated as a point in an imaginary *acceleration space*. If the velocity or acceleration were multiplied by an associated mass, the point could represent, respectively, a position in *momentum space* or a position in *force space*.

The number of degrees of freedom for the functions involving motion once again equal the minimum number of variables required to define a particular function. Now, in number, these variables are the same as those of corresponding position functions, for the number of ways in which a body is free to move is the same, regardless of whether we are interested in "photographing" its position, its rate of change of position, its rate of change of rate of change of position, or some other relationship. In other words, the degrees of freedom are the same for both the static and the dynamic quantities which describe the state of a body. Differences arise in the nature of the variables, rather than in their number, with the fundamental distinction being that the dynamic functions include the element of time, whereas the static functions do not.

To describe the state of an ideal molecule, its position, velocity, and mass must be defined for some given instant. Assuming the mass to be constant, it may be combined with the

velocity, and the two quantities expressed jointly as mv , the momentum. In this case, the phase (state) of the molecule is defined by a function with six degrees of freedom—three for position, and three for momentum. For this purpose, we postulate the existence of an imaginary *phase space* having six dimensions. A conceptual coordinate system for this space requires a point of origin at the intersection of six mutually perpendicular axes. A point darting back and forth in this space would be the continuously changing state of our molecule. Other points could represent the states of other ideal molecules. If it is desired to represent the state of a whole system of N molecules as a single point, then a phase space of $6N$ dimensions may be conceived. The phase function would then require $6N$ coordinates— $3N$ for position, and $3N$ for momentum. Since the particular system of coordinates that might be selected would not change the number of variables, the coordinates may be represented by the general symbols q and p , position and momentum, respectively. Thus, $P(q_1, q_2, q_3, q_4, \dots, q_{3N}, p_1, p_2, p_3, p_4, \dots, p_{3N})$ would be a position function in phase space denoted by *generalized coordinates*, which are so named because they do not specify a particular reference system of measuring.

The pursuit of our subject will not lead into an extended journey through phase space. However, the six-dimensional space will prove more than an interesting fantasy when we return with the quantum and electron, for here these two little demons behave as if they were in their true abode, with our conventional space but a curtained window for peeping in. At the present, however, we continue our adventures with the degree of freedom, for it is now at the door of the principle of equipartition of energy—

and this is where it assumes its physical airs.

DEGREES OF FREEDOM THAT PARTITION THE ENERGY

First, we recall that the equipartition principle is a statistical law which applies as an average only when a large number of molecules are considered collectively. The law declares that at thermal equilibrium the thermal energy of an average molecule will be equally divided among its different degrees of freedom of motion. We shall see that different types of molecules have different degrees of freedom of motion. For instance, an ideal monatomic molecule has three degrees of freedom of motion, and an ideal, rigid, diatomic molecule has five. Let us assume that a system at thermal equilibrium contains both monatomic and diatomic molecules. According to the equipartition principle, the average molecule of each type will have the same quantity of energy—which we shall call one energy unit, and to which we shall assign the symbol E —for each degree of freedom of motion. Thus, the average monatomic molecule will have an energy equal to $3E$, and the average diatomic molecule will have an energy of $5E$. The equipartition of energy principle, then, does not mean that the average molecules of all types have the same energy per molecule, but, rather, the same energy per degree of freedom of motion. In general, the more complex the molecule, the more its degrees of freedom, and hence the greater its total energy. If all the molecules within a system at equilibrium have the same degrees of freedom, regardless of whether they have the same mass, the total thermal energy of the system may be assumed to be divided equally into as many parts as there are degrees of freedom

of motion per molecule. Thus the total thermal energy of an ideal monatomic gas is divided into three equal components, whereas that of an ideal diatomic gas is divided into five. In a mixture of gases, however, the simplicity may vanish, for a proper weighting must be made for the number of molecules and the degrees of freedom of each type. In this event, the equipartition of energy principle is not directly applicable to a few degrees of freedom for the system as a whole, but applies only at the microscopic level. To estimate the thermal energy, as well as its distribution, we first must determine the total number of degrees of freedom of motion of all the molecules. Or, conversely, if we measure the heat capacity of a system, it should tell us something about the nature of the particles that compose the system.

There are three types of motion by which a molecule may exhibit thermal energy—translational, rotational, and vibrational. Now a molecule composed of several atoms can tumble about with vibrational motion as complex as that of a tangle of bed springs tossed in the air. However, by applying in turn the concepts of center of mass; translational, rotational, and vibrational degrees of freedom; and the principle of equipartition of energy, we can systematically arrive at a *theory* of what the average thermal energy of even the most complex molecule should be at a given temperature.

THE CENTER OF MASS AND MOLECULAR MOTION

In order to isolate the translational energy from the rotational and vibrational components, we select some reference position within a molecule, and assume that all the mass is concentrated at that point. The trans-

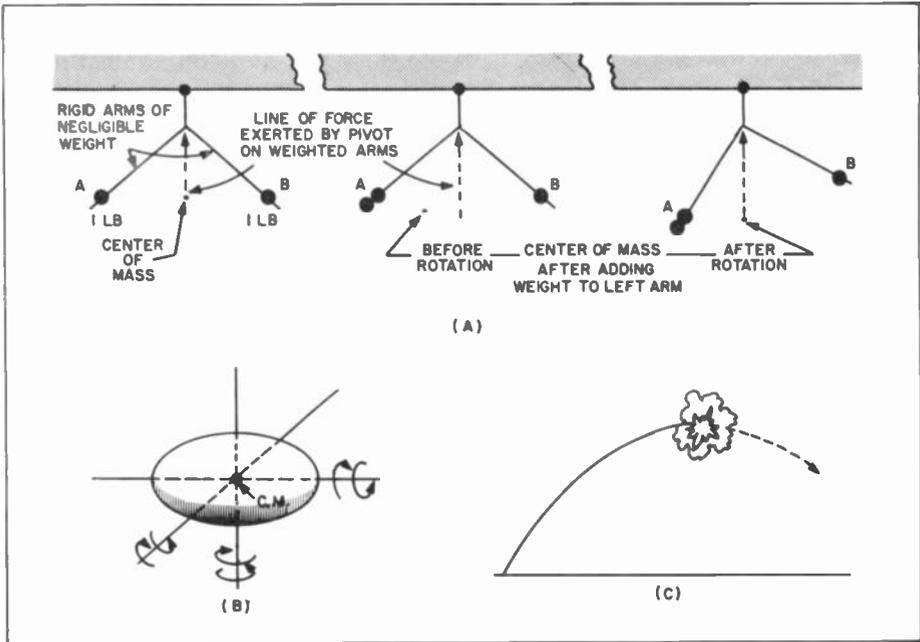


Figure 3. A. Line of Force Exerted by Pivot in a Suspended System at Equilibrium Always Passes Through Center of Mass
B. Three Possible Axes of Rotation About Center of Mass of a Football
C. Trajectory of Center of Mass of Projectile is Unchanged by Explosion of Projectile

lational energy we define to be $\frac{1}{2} mv^2$, where $m =$ the total mass and $v =$ the velocity of the point at a given instant. The problem is greatly simplified if we choose a point that does not wobble with the rotation or internal vibration of the molecule, nor change its speed or direction unless acted upon by external forces. There is only one point in a body that will meet this condition, and that is the *center of mass*—also called the *center of gravity*.

The center of mass is the point at which a body theoretically can be pivoted in a gravitational field and be perfectly balanced regardless of the angle at which it is supported. We say “theoretically” for the center of mass may not lie within the body. For instance, the center of mass of the balanced weights shown in figure 3A is external to the system itself.

Equilibrium is reached when the upward force exerted by the pivot is in a direction that passes through the center of mass. Adding the extra weight to arm A shifts the center of mass toward A, and the angle arms will rotate counterclockwise in the plane of the paper until the center of mass is again directly beneath the point of suspension.

The rotation of a *free* body is always about an axis that passes through its center of mass; thus, this point is unaffected by free rotation. Figure 3B shows three possible axes of rotation for a football. The ball, however, can rotate equally freely about any other axis that passes through the center of mass. The flight of a football’s center of mass is as smooth as that of a baseball’s, even though the football may appear to be moving in a wobbly spiral. If the ball were tossed straight upward with

varying angles of rotation, and allowed to fall on a frictionless surface, its rebounds would also be straight upward (even though its rotation might be affected), for the center of mass could not veer from its vertical line in the absence of external horizontal forces. Thus, by assuming all the mass to be concentrated at its center, the translational motion of any object may be treated as simply as that of an ideal monoatomic molecule, regardless of what shape or angle of rotation the object may have.

The center of mass is also a point independent of the purely internal forces of a system—whether the system is composed of one molecule or trillions. This is because every force has its equal and opposite reaction. When these are completely confined within a system, there can be no net force to act upon the mass. Thus, the trajectory of the center of mass of a projectile (see figure 3C) is not altered in the event of the projectile's explosion in mid-air—assuming the air resistance to be negligible. Likewise, a man on a frictionless surface would be helpless in trying to change the horizontal coordinates of his center of mass, and even his most inspired acrobatics would fall flat in this attempt. The same would be true of a system of a hundred people on the frictionless surface, even if a platform of planking were provided for their maneuvers. In this latter case, the mass of the platform would be part of the system; but neither the ticking of a watch, nor the most impulsive caper could take nature by surprise, and the horizontal position of the system's center of mass would never fluctuate. If the frictionless surface is removed, leaving the system suspended in space, not even the vertical coordinate of the center of mass can be changed. In a like man-

ner, a molecular center of mass is independent of the free vibrations among the atoms which constitute the molecule, and thus such a center of mass serves as a point for concentrating the translational energy, while excluding the internal vibrational energy, as well as the rotational energy.

DEGREES OF FREEDOM OF TRANSLATIONAL MOTION

We now wish to subdivide a molecule's thermal energy according to the degrees of freedom of each kind of motion. Figure 4A illustrates how the translational energy of a particle may be conceived mathematically as being divided into three directions of motion. The vector, however, represents the velocity of the center of mass, and not energy. The energy is $\frac{1}{2}mv^2$. However, since

$$v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$$

$$\text{then } \frac{1}{2}mv^2 = m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

$$= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 + \frac{1}{2}m\dot{z}^2$$

The last three terms divide the energy according to the three degrees of freedom of translation. At thermal equilibrium these are equal on the average, so that, for the average molecule,

$$\frac{1}{2}mv^2 = \frac{3}{2}m\dot{x}^2 =$$

$$\frac{3}{2}m\dot{y}^2 = \frac{3}{2}m\dot{z}^2.$$

DEGREES OF FREEDOM OF ROTATIONAL MOTION

Figure 4B illustrates the rotational energy of a rigid body divided according to the three degrees of freedom of angular rotation. The origin is the center of mass, and the vectors coincide with the axes of rotation—their length being proportional to the

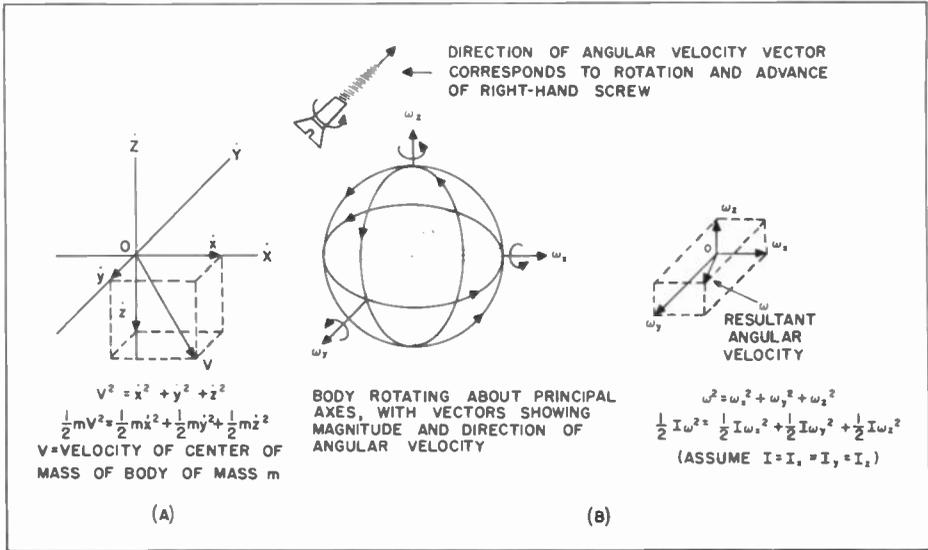


Figure 4. A. Division of Translational Energy into Three Degrees of Freedom
 B. Division of Rotational Energy into Three Degrees of Freedom

angular velocity (radians, or degrees, per unit time), and their direction indicating the direction of rotation. The direction of rotation is the same as that in which a right-hand screw would be rotated to advance it in the direction of the vector. Though vectors of angular velocity are at right angles to their respective planes of rotation, they may be added in the same manner as other vectors, and their resultant will represent the magnitude and direction of the resultant angular velocity. Conversely, the vector of any angular velocity can be resolved into rectangular components. The reference axes for rotation, however, normally do not remain fixed in space, as do those used to describe translational motion; rather, they are fixed with reference to the body, and thus follow its rotational as well as its translational movements. For instance, a shaft can be revolving about an axis directed along its length, but if it is simultaneously being turned to point in another direction, then obviously the longitudinal axis follows this second rotation.

Where the angular velocity, ω ,^o is the rotational analogue of the translational velocity, v , the moment of inertia, I ,^{oo} is the rotational analogue of the translational mass, m . However, unlike m , the moment of inertia of a body varies with the position and direction of the axis of rotation. A slender rod, for example, would offer less inertia on being rolled between the fingers, than when twirled like a baton. The moments of inertia, I_x , I_y , and I_z , will all be equal only if the body has spherical symmetry. The principle of equipartition of

^o $\omega = 2\pi f$ = the number of radians per second, where f = the number of cycles per second.

^{oo} When a rigid body rotates about an axis, regardless of whether the axis passes through or is entirely external to the body, every point in the body (except those that lie on the axis) moves in the path of a circle about the axis. If r is the radius of the circle described by a small element of mass m , mr^2 is the moment of inertia of m about the axis. The total moment of inertia, I , about the axis is the sum of the moments of inertia of all the elements of mass.

energy applies, nevertheless, to the degrees of freedom of rotation. Thus, the molecular axes having the smallest moments of inertia should, on the average, have the greatest angular velocities. The total rotational energy

$$\frac{1}{2} I \omega^2 = \frac{1}{2} I_x \omega_x^2 + \frac{1}{2} I_y \omega_y^2 + \frac{1}{2} I_z \omega_z^2.^\circ$$

For the average molecule,

$$\frac{1}{2} I \omega^2 = \frac{3}{2} I_x \omega_x^2, \text{ etc.}$$

If the rotational energy of a body is to be changed, a turning force, or torque, must be exerted about an axis of rotation. This cannot be accomplished by a force that is directed so that it passes through the axis, whatever the angle, for a torque implies a force with leverage about an axis. Now our definition of an ideal monatomic molecule described it as a particle having a perfectly spherical and frictionless outer surface. The frictionless surface prevents the possibility of a component of force at right angles to the radius of the sphere. All collision forces on such a particle are necessarily directed toward its center of mass and hence toward every possible axis of free rotation, since they all pass through the center of mass.

Any rotational energy that might be present would exist as an inherent property of the particle, and not be subject to change. As an example, suppose that billiard balls were ideal molecules. Instead of rolling, they would slide about like caroms, and their collisions with one another would affect only the translational velocities. The rotational velocities, if any, would remain unchanged—even if we should pick up a ball, we could neither start nor stop its spinning. To

inspect both sides of a nonspinning ball held in the hand, the observer himself would have to turn around. Ideal monatomic molecules, then, do not have rotational degrees of freedom. Neither do actual monatomic molecules such as helium, neon, and argon. One might expect their electron shells to rotate, if it were not for the fact that this would require electron orbits forbidden by the quantum. Because of the quantum, however, all atoms may be considered as frictionless spheres, or better, as merely dimensionless points of mass, for none has rotational degrees of freedom about its own nucleus—at least not from a thermal point of view at ordinary temperatures. The free rotation of a particle requires a combination of at least two atoms, and an axis of rotation that passes through their combined center of mass, but is external to at least one of the nuclei. This latter requirement is illustrated in the following example.

Figure 5 shows three types of molecules, classified according to their rotational degrees of freedom. Figure 5A is simply the equivalent mass point of a monatomic molecule which, as explained above, has no rotational degrees of freedom.^{°°} Diatomic molecules, such as H₂, O₂, etc., have two rotational degrees of freedom, as shown at the left in figure 5B. The distance between the mass points is approximately 10⁻⁸ cm., and the center of mass is at the midway point of the line of centers. The two axes of rotation are perpendicular to each other and to the line of centers. A molecule of three or more atoms also has only two axes of rotation if all the mass points lie in the same straight line. However, if one of the

[°] This equation strictly holds only for certain reference axes, called the *principal axes of inertia*.

^{°°} Actually, atomic nuclei, as well as elec-

trons, appear to spin about their centers of mass. However, both nuclear and electron spins are fixed by the quantum, and hence are not free variables of thermal energy.

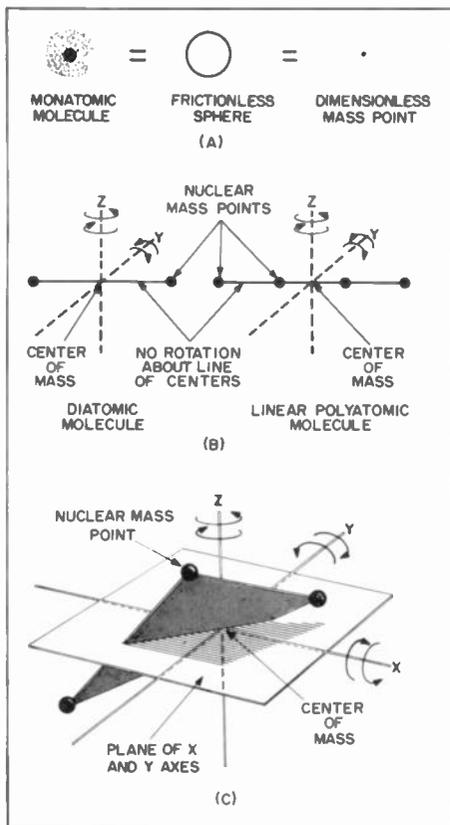


Figure 5. Rotational Degrees of Freedom of (A) Monatomic Molecule, (B) Diatomic Molecule, and (C) Polyatomic Molecule

nuclear mass points of a polyatomic molecule is not collinear with the other mass points, the molecule will have three axes of rotation, and hence three rotational degrees of freedom, as illustrated in figure 5C.

From the foregoing discussion we conclude that a *rigid* gas molecule—that is, a gas molecule with no internal vibrational energy—may have a total of 3, 5, or 6 degrees of freedom of motion, depending upon whether the molecule is respectively monatomic, diatomic, or noncollinearly polyatomic. Three of the degrees of freedom in each instance are those of translational motion, and the remaining degrees of freedom are those of rotational motion. Note that six de-

grees of freedom of motion is the maximum number for a rigid gas molecule, whether the molecule comprises three atoms, or three hundred.

DEGREES OF FREEDOM OF VIBRATIONAL MOTION

If a gas molecule is not rigid, new degrees of freedom must be added to account for vibrational motion. The actual modes of vibration within large molecules can be quite complex, but the essential problem is to determine the number of these modes, and their distribution of energy. Remember, that to completely define at a given instant the motion of a mass point in space, three coordinates are required. To completely define at a given instant the motion of N mass points, $3N$ coordinates are required. If all these coordinates are independently variable, the system will have $3N$ degrees of freedom in which to move. Now consider a single gas molecule composed of N noncollinear atoms, all of which are free to vibrate. There will be a total of $3N$ degrees of freedom of motion. Of these, three will be translational, and three will be rotational, leaving $3N - 6$ degrees of freedom for vibrational motion. If all the N atoms were collinear, there would be only two rotational degrees of freedom, and hence $3N - 5$ vibrational degrees of freedom. In other words, one additional mode of vibration is possible if all the atoms are arranged in a single, straight line.

In a manner similar to the way in which the velocity of a molecule is divided into components according to a system of reference axes, the vibrational motion may also be described. Here, again, the choice of a system of reference is simply one of convenience. Since all the atoms in a molecule are coupled to each other, a gain in energy by any one is soon spread throughout the molecule. For

this reason, the principle components of vibration, equal to the number of vibrational degrees of freedom, are usually selected so that each is a vibrational mode involving all the atoms in the molecule. (Even though a vibrational mode may apply to the molecule as a whole, a particularly strong valence bond may effectively control its frequency. For example, a hydrogen-carbon, or hydrogen-oxygen bond has a characteristic frequency of approximately 9×10^{13} cycles per second, changing but little with the size and shape of the rest of the molecule.) These principal modes are called the *normal vibrations*. A normal vibration is recognized as one in which the atoms oscillate in phase—that is, the atoms pass through their equilibrium positions simultaneously, and reach their peak displacements simultaneously. Each normal vibration has its characteristic frequency. If two or more normal vibrations of the same molecule have the same frequency, they are called *degenerate vibrations*. Later, we shall see that the term *degenerate* carries a broader implication when applied to molecular activities.

Normal vibrations may be divided into two types: *valence vibrations*, and *deformation vibrations*. Valence vibrations are extensional (stretching) oscillations of atoms in the direction of their valence bonds. Deformation vibrations are flexural (bending) oscillations of atoms at right angles to their valence bonds.

It can be shown that a molecule of N atoms, having $3N - 6$ vibrations, will have $N - 1$ valence vibrations, with the remaining normal vibrations being of the deformation type.

Figures 6, 7, 8, and 9 show the normal vibrations of a few of the simpler molecules. A diatomic molecule, as shown in figure 6, has but

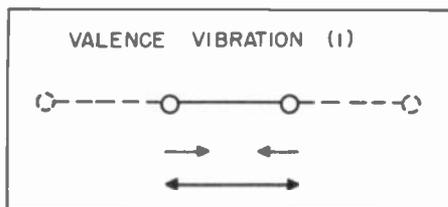


Figure 6. Normal Vibration of a Diatomic Molecule

one vibrational degree of freedom—a valence vibration. A linear, triatomic molecule, such as carbon dioxide, CO_2 , has $3N - 5 = 4$ vibrational degrees of freedom, as shown in figure 7. Of these, $N - 1 = 2$ are valence vibrations. The two deformation vibrations are simply the two components of the flexural oscillation shown in figure 7C. (One component is in the plane of the paper, with the other component perpendicular to the plane of the paper.) Since these can be considered as similar components of a single, combined mode of oscillation, the two deformation modes have the same frequency, and hence are classed as degenerate. Figure 8 shows the $3N - 6 = 3$ vibrational modes of a non-linear triatomic molecule, such as the water molecule, H_2O . There are two valence vibrations (figures 8A and 8B), as in the case of the linear triatomic molecule, but only one deformation vibration (figure 8C). A second deformation vibration is effectively traded for a

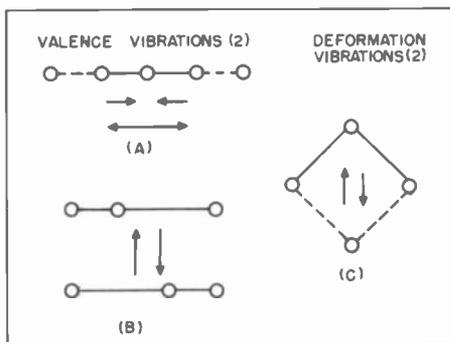


Figure 7. Normal Vibrations of a Linear Triatomic Molecule

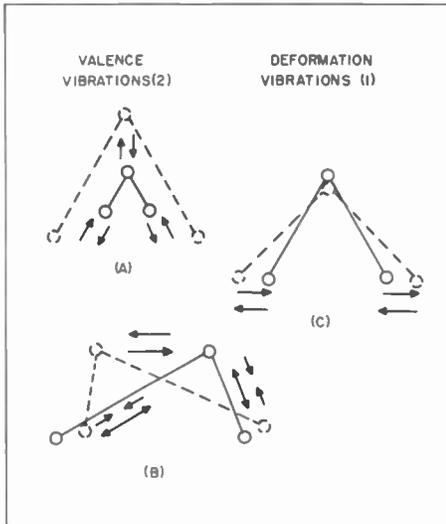


Figure 8. Normal Vibrations of a Non-Linear Triatomic Molecule

third rotational degree of freedom. Figure 9 shows the $3N - 5 = 7$ vibrations of a linear molecule with four atoms. Of these, $N - 1 = 3$ are valence vibrations. This leaves four vibrations for the deformation modes, two of which are the perpendicular components of the oscillations illustrated in figures 9D and 9E.

It should be remembered that the vibrational motions are so proportioned that they do not disturb the center of mass, nor produce a net angular momentum. For example, in figure 9E, if all four atoms were of the same mass, the maximum displacements of the two center atoms would necessarily be greater than those of the end atoms, else there would be a net rotation about the center of mass—clockwise at the instant shown in the left-hand side of figure 9E.

Figure 9 also suggests the modes of vibration that can be expected for a linear molecule containing a large number of atoms. For example, the first deformation oscillation, with its two degrees of freedom, would

correspond to the first harmonic of a vibrating string; the second deformation oscillation to the second harmonic, and so on to $N - 2$ harmonics.

Now imagine a molecule in the shape of a rectangular block containing a trillion trillion atoms. This molecule would be a crystal of approximately the size of a package of cigarettes. $N = 10^{24}$ is such a large number that for all practical purposes we may neglect the translational and rotational degrees of freedom, and assume that the crystal has $3N$ normal vibrations. Of these, we may assume N to be valence vibrations and $2N$ to be deformation vibrations. Instead of assigning but one degree of freedom to each mode, we may assign three. Thus, there would be N normal modes of vibration, each with 3 "coordinates," corresponding to the X,

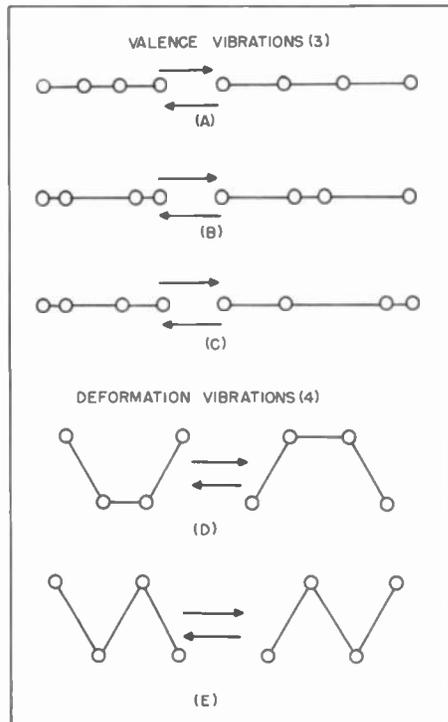


Figure 9. Normal Vibrations of a Linear Molecule of Four Atoms

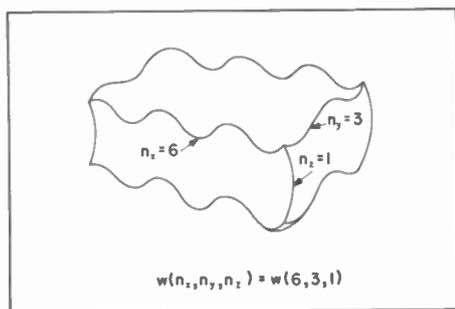


Figure 10. Representative Deformation Mode of Vibration of a Crystal

Y, and Z axes. To define one of these modes, a wave function with three degrees of freedom is required. Thus, the wave function $W(n_x, n_y, n_z) = W(6, 3, 1)$ would define the deformation mode illustrated in figure 10, where n_x , n_y , and n_z represent the standing wave harmonics along their respective axes.^o

It is interesting to note that the cyclic motions normal to a given crystal can be defined by sets of integers which are limited to a finite number of values. In other words, the wave function is effectively quantized, and has definite upper and lower limits. It will be recalled that the cyclic motion of the electron in the Bohr hydrogen atom was also quantized, this motion having a lower orbital limit, though not an upper limit since no outer boundary was assumed. Also, the motion of the electron was expressed by three quantum numbers, each representing a degree of freedom of orbital motion, and a fourth quantum number to represent its degree of rotation.

^o This example is an ideal case assumed for the sake of simplicity. The behavior of actual crystals reveals a great complexity of interlocking vibrations between modes of nearly the same frequency, with vibrational modes resulting which would not occur with crystals of different shape or structure.

We should not carry the analogy too far, however, for with our vibrating atoms and molecules we are still within the laws of classical mechanics, and the quantum mechanics does not begin until the appearance of Planck's constant, h . Though the crystal wave function represents a definite band of frequencies, no restrictions have been introduced which would limit the energy a particular mode could have. Thus, whereas the frequencies are quantized according to the vibrational degrees of freedom, the energy, as yet, we assume to be continuously variable.

When we apply the equipartition of energy principle to the vibrational degrees of freedom, we find that these degrees must be doubled if the principle is to hold. Thus the one vibrational degree of freedom of the diatomic molecule in figure 6 effectively becomes two degrees of freedom. The reason for this is that the energy of a vibrating system is partly kinetic and partly potential. At the instant of maximum displacement, the energy is entirely potential; at the equilibrium point, the energy is entirely kinetic; and at all intermediate points, the energy is divided between the two forms. This is illustrated by the swing of a pendulum, as shown in figure 11. In the case of a simple harmonic oscillator, the energy averaged over a complete cycle is one-half potential and one-half kinetic. The equipartition principle requires the average kinetic energy for all degrees of freedom of motion to be the same at thermal equilibrium. But a vibrational degree of freedom cannot meet this requirement unless a corresponding average quantity of potential energy is also stored. Thus, a vibrational degree of freedom is effectively equivalent to two translational or two rotational degrees of freedom, insofar as its energy content is concerned.

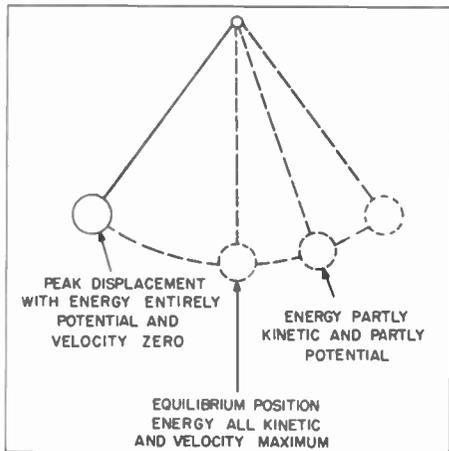


Figure 11. Division of Energy into Potential and Kinetic Forms in a Simple Harmonic Oscillator, as Represented by a Pendulum

This apparently contradicts our previous usage of the term *degree of freedom*. At this point, in order to keep our basic concept of a degree of freedom from eluding us and eloping with a degree of freedom of its own, it might be wise to pause and swat at a semantic gnat.

DEPARTMENT FOR KEEPING THE RECORD STRAIGHT

We defined the degrees of freedom of motion of an N mass-point system to be $3N$. On the other hand, we declared that the equipartition principle applies to all the degrees of freedom of motion. However, if the equipartition of energy principle is true, we see from the discussion above that the number of degrees of freedom becomes greater than $3N$ if vibrational modes are present. The apparent contradiction lies in the fact that if a mass point moves through a field of force—as does an atom when vibrating about a valence bond—the state of its energy varies with its position as well as with its velocity. In geometrical space, the three directions in which a position may vary are the same as the three directions in which

its velocity may vary. Thus, the energy of a vibration is effectively shared by the same degree of freedom of motion twice. We could assume a phase space with six dimensions for each mass point—three for position, and three for momentum—then each vibrational degree of freedom would actually equal two degrees of motion in phase space. But rather than force a meeting of words, it seems the better to disrobe the degree of freedom of its physical masquerade, and keep to its original meaning.

A degree of freedom is simply a unit of freedom, and the degrees of freedom of a function are simply its number of independent variables, and not the variables themselves. The degrees of freedom of an energy function will depend upon the particular function. Thus, the function $E = 1/2 (mc^2)$, where the speed, c , can be defined by one coordinate, has one degree of freedom if m is constant. On the other hand, $E = 1/2 (mv^2)$, where the velocity, v , requires three coordinates to be defined, will have three degrees of freedom. If we so desire, further subdivisions can be made, and we can call the respective coordinates degrees of freedom of one thing or another. From this point of view, the additional degrees of freedom are not extra “baskets” that actually increase the energy-carrying capacity of a system, but are extra ways in which to divide the energy; and the fact that the thermal energy of an average molecule increases with an increase in its degrees of freedom is to be traced to the probabilities of energy exchange governing its collisions. If the statement of the equipartition principle is to have a reasonably concrete meaning, it seems that possibly the expression *degrees of freedom of energy* can be favored to replace *degrees of freedom of motion*. In the discussions to follow,

where the equipartition principle is implied, we shall make this substitution. It should be understood, however, that the number of degrees of freedom of energy is a particular function of, but not necessarily equal to, the number of degrees of freedom of motion.

While straightening the records, we may as well review the log up to this point, the better to fix our present position. We began with the degree of freedom as it appeared in the field of statistics, and followed it from there to its role as a free variable unit applying to all variable functions. From that point we traced its special application in functions defining positions in geometrical space. Here the degree of freedom became identified with the directions in which a point is free to move. We specified these as *degrees of freedom of motion*, and saw how they could be used to construct imaginary spaces similar to geometrical space. Next, the centers of mass of atomic nuclei were substituted for the points in space, and the energy of the molecules was found to divide itself according to the degrees of freedom of motion, each division of which we have just agreed to call a *degree of freedom of energy*. By this division of energy, the degrees of freedom thus become effective substitutes for the particles. Instead of imagining individual particles as carriers of energy, it is actually less complicated to imagine the energy stored in a fixed number of degrees of freedom. Furthermore, in the vibrations among large numbers of atoms or molecules, as in a crystal, the individual particles begin to lose their identity anyway, for when all are coupled together, each mode of vibration involves every particle in the system, and the energy of a particular vibration cannot be confined to a particular particle. Thus, we found

that a wave function, expressing the normal vibrations of a system, assumes a macroscopic nature, where the individuality of the particles is secondary to the number of degrees of freedom of energy. We should note, however, that the degrees of freedom of a wave function, such as n_x, n_y, n_z , of the function $W(n_x, n_y, n_z)$, are not the same as the degrees of freedom of vibrational energy of the system in question. These latter appear as the *number of solutions* which the wave function may have, or rather as twice the number of solutions, if we subdivide them into potential and kinetic components.

If a system were composed of a completely continuous substance, then theoretically it would have an infinite number of degrees of freedom. It is a curious fact that the quantization of the mass into particles limits the degrees of freedom of a system, yet the more this quantization is carried out per unit volume, the more the degrees of freedom should multiply. We shall find, however, that the entrance of the quantum will appear to reverse this prediction.

HEAT CAPACITY

The degrees of freedom of energy are now to be put to a practical test, by using them to evaluate the heat capacity of a system. The common practice is to measure and express thermal energy in calories,^o British thermal units,^{oo} joules, ergs, or foot-pounds. It is principally in equations of a theoretical nature that the units

^o One gram-calorie is the quantity of heat required to raise the temperature of one gram of water 1° C. (or 1° abs.). One kilogram-calorie equals 1000 gram-calories.

^{oo} One British thermal unit (BTU) is the quantity of heat required to raise the temperature of one pound of water 1° F.

k and R, the Boltzmann and universal gas constants, respectively, have a practical use. Whenever a quantity is dependent upon the temperature, k and R are apt to be involved in the formulas that define or predict the value of that quantity. An important example arises when we seek to explain the heat capacities of various substances in accordance with the kinetic theory of molecules.

The term *heat capacity* is defined as the heat absorbed per degree of temperature rise. The heat capacity of a system is the thermal energy absorbed (ΔE_T) divided by the change in temperature (ΔT).

Thus,

$$\text{heat capacity} = \Delta E_T / \Delta T$$

HEAT CAPACITY AT CONSTANT VOLUME

What is the molar heat capacity of an ideal monatomic gas at constant volume?

Since,

$$k_T = \frac{2}{3} \text{ of the average energy per molecule,}$$

then,

$$\frac{3}{2} k_T = \text{average energy per molecule,}$$

and,

$$\frac{3}{2} k = \text{average energy per molecule per degree.} \quad (1)$$

In other words, for each absolute-degree change in temperature, the energy of each molecule will change by an average amount equal to $\frac{3}{2}k$.

Thus, for N molecules, $\Delta E_T = \frac{3}{2} Nk \Delta T$. Let $C_V =$ the heat capacity of an ideal gas at constant volume; then,

$$C_V = \frac{\Delta E_T}{\Delta T} = \frac{\frac{3}{2}(Nk \Delta T)}{\Delta T} = \frac{3}{2} Nk.$$

For one mole of an ideal gas:

$$C_V = \frac{3}{2} N_A k = \frac{3}{2} R,$$

where N_A is Avogadro's number, and R is the universal gas constant. Since R is approximately equal to 2 gram-calories/mole/degree, then the molar $C_V = 3/2 \times 2 = 3$ gram-calories/mole/degree. This is the theoretical value for an ideal gas, but it checks very closely with the observed molar C_V 's of the inert, monatomic gases. For example, the observed value for helium is 3.02 gram-calories/mole/degree, and for argon is 3.07 gram-calories/mole/degree.

When the temperature of an ideal gas is increased one degree, each mole will have gained three calories of energy. This is true regardless of whether the volume is held constant, or not. However, if the gas is allowed to expand while being heated, work is done. The heat required to raise the temperature must then equal $C_V \Delta T + W$, where W is the work of expansion.

HEAT CAPACITY AT CONSTANT PRESSURE

What is the molar heat capacity of an ideal monatomic gas at constant pressure? Suppose the box in figure 12 contains 1 mole of gas, and that the east wall is the face of a piston which moves outward when the gas is heated in such a manner that the pressure at all times remains constant.

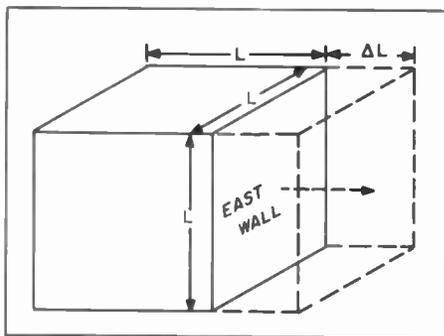


Figure 12. Expansion of Volume of a Box by an Amount $\Delta V = L^2 \Delta L$

If the temperature is increased by ΔT , the east wall moves outward a distance, ΔL , so that the work of expansion is

$W = \text{force} \times \text{distance} = \text{pressure} \times \text{area} \times \Delta L = PL^2 \Delta L = P \Delta V$,
where:

ΔV is the change in volume.

Now, $PV = nRT$ (equation of the state of an ideal gas),

and, $n = 1$ (by the assumption of only 1 mole of gas),

so, $PV = RT$ (before heating). (2)

Also, $P(V + \Delta V) = R(T + \Delta T)$ (after heating),

or, $(PV - RT) + P \Delta V = R \Delta T$.

But, $PV - RT = 0$ (by equation 2).

Thus, $W = P \Delta V = R \Delta T$

Then, total heat = $\Delta E_T = C_V \Delta T$

$$+ W = C_V \Delta T + R \Delta T$$

and the heat per mole per degree =

$$\frac{\Delta E_T}{\Delta T} = \frac{C_V \Delta T + R \Delta T}{\Delta T} = C_V + R$$

The quantity, $C_V + R$, is the molar heat capacity at constant pressure, and is given the symbol C_p . The first term C_V is the added energy actually retained by a mole for each degree of temperature, while R is the energy expended in increasing the volume. Since $C_V = 3/2 R$, then $C_p = 5/2 R = 5$ gram-calories/mole/degree, for an ideal gas.

It is C_V , the heat capacity at constant volume, however, that is our chief interest, since the solids in which electron gases are contained do not appreciably expand, as compared with constant-pressure gas systems.

We originally defined k as a unit of energy per molecule, and then interpreted the unit as a "degree" of temperature. Unfortunately, because k is only a partial measure of a molecule's energy per degree of temperature, it can be awkward to use where

the total energy is to be expressed. The same may be said for the gas constant, R . However, from the point of view of system mechanics, the energy per particle need not be of fundamental importance. What is important is the energy per degree of freedom.

A unit for expressing the total energy of an ideal system per degree of temperature is C_V . Since for an ideal monatomic gas, $C_V = 3/2 Nk$, then, $\frac{C_V}{N} = \frac{3}{2} k$, is a constant expressing the total energy/molecule/degree. Now the ideal monatomic molecule has three degrees of freedom of energy, and its total energy, $3/2 kT$, is divided equally among them. This means that the average ideal molecule has an energy of $1/2 kT$ per degree of freedom. This quantity expresses the total energy per degree of freedom—not some fractional part—where $k/2$ is the unit of energy, and T is the number of units.

PRELUDE TO DISASTER

The principle of equipartition of energy states that at thermal equilibrium the molecular energy is divided equally among all the degrees of freedom of energy.

For a system of ideal molecules, C_V would remain constant and not change with temperature. Thus, the total thermal energy could be expressed as $C_V T$, as accurately as a change in thermal energy is expressed as $C_V \Delta T$. The problem now is to establish the rules for predicting the value of C_V for a system of N molecules.

First, since a change in thermal energy is shared equally by each degree of freedom of energy, then $C_V = n_f c_f$,

where:

$n_f =$ number of degrees of freedom

$$\begin{aligned}
 c_f &= \text{"heat capacity" per degree} \\
 &\quad \text{of freedom} \\
 &= k/2 \text{ (energy/degree of} \\
 &\quad \text{freedom/degree of} \\
 &\quad \text{temperature)}
 \end{aligned}$$

Next, we must determine the number of degrees of freedom of energy of the particular system in question. This, of course, will vary, not only with the number, but the kind of molecules. As a specific example, suppose that we attempt to estimate the heat capacity of a block composed of 1 mole of copper. Since the block is effectively a single molecule, the translational and rotational degrees of freedom of the atoms do not enter the problem. However, the N_A atoms may be assumed to have $3N_A$ vibrational degrees of freedom, giving a total of $2 \times 3N_A = 6N_A$ degrees of freedom of energy. But this is not all, for according to the atomic theory, a metal will also have a gas of monatomic molecules flying about within it in the form of free electrons. Empirical data, such as that of the thermoelectric and the Hall effects, indicates approximately one free electron per atom. Thus, the electron gas should add $3N_A$ degrees of freedom of translational energy.

$$\begin{aligned}
 \text{Thus, } C_V &= n_f c_f \\
 &= (6N_A + 3N_A) \frac{k}{2} \\
 &= 3 N_A k + \frac{3}{2} N_A k
 \end{aligned}$$

$$\begin{aligned}
 &= 3R + \frac{3}{2}R \\
 &= 6 \text{ calories/deg. absolute} \\
 &\quad + 3 \text{ calories/deg.} \\
 &\quad \text{absolute} = 9 \text{ gram-} \\
 &\quad \text{calories/deg. absolute} \\
 &\quad \text{(per mole)}.
 \end{aligned}$$

If the atoms and electrons in a copper bar behave in the manner of ideal particles, then 1 atomic weight of copper should require 9 gram-calories to raise its temperature 1 degree absolute (or 1 degree centigrade)—6 calories for the atoms, and 3 calories for the electrons.

Behind this prediction is arrayed the power and dignity of the whole body of Newtonian mechanics, as well as that of the atomic theory of solids. But wait! A peek at the data of an actual experiment reveals a molar C_V for copper of 6.01 gram-calories/deg. absolute/mole. This is little more than the molar C_V of an average insulator. Where are the free electrons? The dastardly traitors have abandoned their posts in the hour of need. Without a struggle, $3N$ degrees of freedom are surrendered, and after them must roll the revered crowns of classical mechanics. Unless we please to look the other way, the fact must now be faced that the free electrons of the atomic theory of solids do not exist.



Solution to . . .

Last Month's "What's Your Answer?"

The output voltage at E_a will consist of the rectified output of G_1 , the 50-volt generator.

The generator with the largest output voltage will cause the selenium rectifiers associated with the other generators to appear as open circuits because of the inverse nature of the applied voltage. Thus, only CR_1 will conduct, thereby rectifying the output of G_1 .

Mr. Tefius Dectrow, the contributor of this problem, says the circuit is used as a maximum-source trigger device. (The generators are vacuum-tube circuits with varying outputs, of course.) Another application could be found in a diversity reception system in which each generator represented the detector stage of a corresponding space-diversity receiver.

