1968

Preliminary Investigation of a new Frequency Modulator

David A. Cleveland

PRELIMINARY INVESTIGATION OF A NEW FREQUENCY MODULATOR

BY

DAVID A. CLEVELAND

A thesis submitted
in partial fulfillment of the requirements for the
degree Master of Science, Department of
Electrical Engineering, South Dakota
State University

1968

PRELIMINARY INVESTIGATION OF A

NEW FREQUENCY MODULATOR

This thesis is approved as a creditable and independent investigation by a candidate for the degree, Master of Science, and is acceptable as meeting the thesis requirements for this degree, but without implying that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Thesis Advisor	Date
Head, Electrical Engineering Department	Date

ACKNOWLEDGMENTS

The author wishes to express his appreciation and gratitude to Dr. F. C. Fitchen whose guidance and advice made this investigation possible, and to Albert Wang for his help in recording part of the data.

D.A.C.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

A. History

Communication is an inherent characteristic of the human species. Man has continually strived to improve his methods of communicating with his fellow man. With the understanding of electricity and its subsequent control, a powerful tool became available; one important use of this tool is in the field of communications. First applications of electricity to communications involved physical connections between sending and receiving instruments. When scientists desired to eliminate the direct connections between the communicators, new methods had to be devised.

The new methods which were developed involved the transmission of electromagnetic energy through space. One of the basic electrical waveforms used was a sine wave; the sine wave by itself, however, contains no information. Therefore, some characteristic of the wave must be varied in order that the wave can carry information. There are three fundamental qualities of a wave that can be altered--the amplitude, the frequency, and the phase. "Modulation" is the word which was adopted to denote the variation of one of these characteristics of a wave.

Early work was directed toward the perfection of amplitude modulation systems. But amplitude modulation has certain limitations,

among which is a noise problem. Therefore, in the early 1920's frequency modulation (F-M) was proposed. In 1922 J. R. Carlson published a paper which was unfavorable towards frequency modulation, and the result was a decline of interest for several years. E. H. Armstrong soon discovered that F-M would actually result in less noise, and it was not long before F-M came into wide use.[5]

B. Frequency Modulation

In any of the three types of modulation the waveform which contains the information, the modulating signal, is combined with another waveform, the carrier, such that a characteristic of the carrier is varied in accordance with the amplitude of the modulating signal. The modulating signal is usually in the audio frequency range, and the carrier signal is at some higher frequency that can be efficiently transmitted.

A typical carrier signal is of the form

$$v = V_c \cos \theta \tag{1}$$

θ is a displacement angle dependent upon time such that

$$\theta = \int \omega_c dt + \theta_o \tag{2}$$

and therefore

$$v = V_c \cos \left[\int \omega_c dt + \theta_o \right] \tag{3}$$

where V_c represents the amplitude of the carrier, ω_c is the angular

frequency, and θ_o is the phase angle. For frequency modulation, then, the carrier frequency is varied about some center frequency f_o such that

$$\omega_c = \omega_o + k_f V_m \cos \omega_m t \tag{4}$$

where $\omega_o = 2\pi f_o$, V_m and ω_m are respectively the amplitude and angular frequency of the modulating signal, and k_f is the degree of frequency variation and is determined by the amplitude of the modulating signal. Substituting Eq. 4 into Eq. 3 gives:

$$v = V_c \cos\left[\int (\omega_o + k_f V_m \cos \omega_m t)dt + \theta_o\right] \tag{5}$$

If θ_o is considered to be a constant angle, it may be dropped, and upon integration Eq. 5 becomes:

$$v = V_c \cos(\omega_o t + \frac{k_f V_m}{\omega_m} \sin \omega_m t) \tag{6}$$

the term $k_f V_m/\omega_m$ is defined as the modulation index m_f. Substituting m_f into Eq. 6 yields:

$$v = V_c \cos(\omega_o t + m_f \sin \omega_m t) \tag{7}$$

Eq. 7 is the general equation for a frequency-modulated wave.

The amount of frequency change that the carrier exhibits is determined only by the amplitude of the modulating signal. The frequency of the modulating signal controls the rapidity with which these changes occur but has no effect on the extent of these changes. For commercial F-M stations the F.C.C. has set a limitation of

± 75 kHz on the amount that the carrier may be shifted during modulation. An additional ± 25 kHz is provided as a guardband to make a total bandwidth of 200 kHz. The purpose of the guardbands is discussed in the next section.

C. F-M Sidebands

In Eq. 7 it was shown that a frequency-modulated signal may be expressed as

$$v = V_c \cos(\omega_o t + m_f \sin \omega_m t) \tag{8}$$

This equation is of the form cos (a + b) and, therefore, can be expanded as

$$\begin{aligned} &\cos(\omega_o t + m_f \sin \omega_m t) \\ &\quad = \cos \omega_o t \cos(m_f \sin \omega_m t) \\ &\quad\quad - \sin \omega_o t \sin(m_f \sin \omega_m t) \end{aligned} \tag{9}$$

These functions may be evaluated by means of Bessel functions where

$$\begin{aligned} \cos(m_f \sin \omega_m t) &= J_o(m_f) + 2J_2(m_f) \cos 2\omega_m t \\ &\quad + 2J_4(m_f) \cos 4\omega_m t + \ldots \end{aligned} \tag{10}$$

and

$$\begin{aligned} \sin(m_f \sin \omega_m t) &= 2J_1(m_f) \sin \omega_m t \\ &\quad + 2J_3(m_f) \sin 3\omega_m t + 2J_5(m_f) \sin 5\omega_m t + \ldots \end{aligned} \tag{11}$$

The J_n's are Bessel functions of the first kind and order n and are defined by an infinite series given by

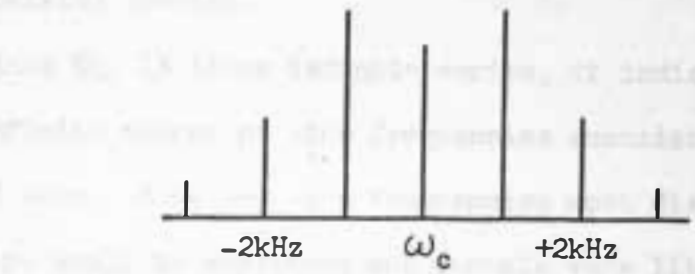
$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{x}{2}\right)^{n+2k} \tag{12}$$

By using the trigonometric identities and Eqs. 10 and 11, the frequency modulated wave becomes

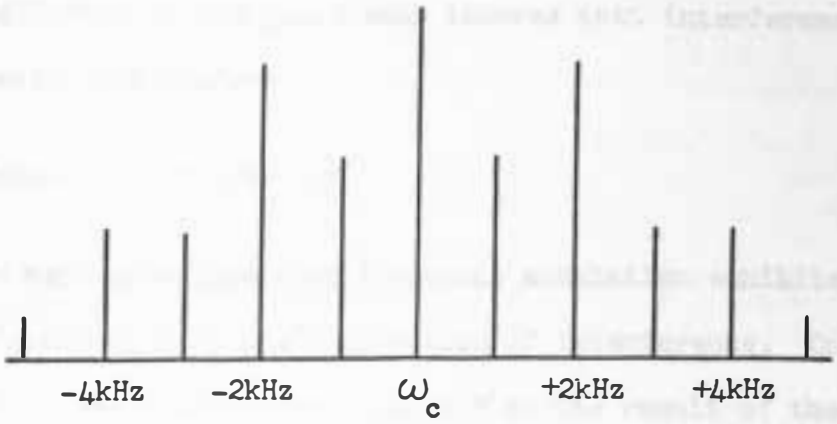
$$\begin{aligned} v = V_c \Big\{ & J_o(m_f) \cos \omega_o t \\ & - J_1(m_f) \left[\cos(\omega_o - \omega_m)t - \cos(\omega_o + \omega_m)t \right] \\ & + J_2(m_f) \left[\cos(\omega_o - 2\omega_m)t + \cos(\omega_o + 2\omega_m)t \right] \\ & - J_3(m_f) \left[\cos(\omega_o - 3\omega_m)t + \cos(\omega_o + 3\omega_m)t \right] + \ldots \Big\} \end{aligned} \tag{13}$$

It thus becomes apparent that a frequency modulated wave consists of a center frequency $\omega_o/2\pi$ and an infinite set of side frequencies equally spaced on both sides of the carrier at distances determined by the frequency of the modulating signal. Figure 1 shows a typical frequency spectrum caused by a 1 kHz audio signal. The amplitude of the audio signal determines the number of significant side frequencies.

The amplitudes of each pair of side frequencies can be determined from the Bessel coefficients of Eq. 13. For a time-varying modulating signal these amplitudes are continually changing; however, the total energy contained in the carrier and the sidebands remains



(a)



(b)

Figure 1. Frequency Spectrum Caused By (a) A Small 1 kHz Audio Signal and (b) A Large 1 kHz Audio Signal

constant at all times. The sidebands obtain all of their energy from the carrier. This is a logical effect when one considers that no energy is added to the system by the modulating signal. The amplitude of the modulated carrier is, therefore, always less than that of the unmodulated carrier.

Since Eq. 13 is an infinite series, it indicates that there are an infinite number of side frequencies associated with a frequency modulated wave. However, side frequencies most distant from the carrier are small in amplitude and contain very little power. These distant side frequencies are the reason for the additional guardbands about each broadcasting station. Each commercial station is limited to use a band of $\pm$ 75 kHz about its center frequency, and the extra $\pm$ 25 kHz allotted to the guardbands insures that interference between stations will be minimized.

D. Advantages of F-M over A-M

The main advantage that frequency modulation exhibits over amplitude modulation is the suppression of interference. One reason that F-M has less interference than A-M is the result of the frequency range that is involved. F-M came into prominence after A-M and was, therefore, allocated the higher frequencies where fewer natural disturbances are to be found. However, this is not the only reason why F-M results in less noise.

The principal reason why F-M is subject to less interference is based upon inherent characteristics of the frequency modulation

process. In amplitude modulation, reproduction of the modulating signal at the receiver depends upon the amplitude of the carrier wave. Interference, whether man-made or natural, changes the amplitude of the carrier and thus produces distortion of the information. In fact, with amplitude modulation the signal-to-noise ratio must be of the order of 100 to 1 to obscure the noise.[7]

Should the amplitude of a frequency-modulated wave be altered by interference, such amplitude variations may be eliminated without destroying information. This is done in the receiver by a limiter stage which limits the amplitude of the incoming signal to a constant level. Therefore, with frequency modulation the signal-to-noise ratio need be only 2 to 1, resulting in a much clearer signal at the receiver.[7]

Another advantage of F-M is the result of operation at the higher frequencies. Because of overcrowding in the allowed amplitude modulation bands, the A-M stations are limited to a narrow bandwidth of 10 kHz each. This limits the audio frequencies that can be used. In contrast, F-M stations are allotted a much wider bandwidth, and much higher audio frequencies can be handled, resulting in high-fidelity transmission and reception.[5]

E. Methods of Generating F-M

Usual generation of a frequency-modulated signal involves variation of the resonant frequency of the tank circuit of an oscillator. Different methods of accomplishing this have been

developed. One of the most common procedures involves the use of a reactance tube which can be made to act as an inductance or a capacitance. By having plate voltage lead grid voltage by 90°, the tube acts as an inductor; and by having plate voltage lag grid voltage, it acts as a capacitor. By properly connecting the reactance tube to the tank circuit of an oscillator, frequency modulation is produced.[14] Other frequency modulators have been developed which make use of changes in capacitance between the terminals of bipolar transistors. Connecting the proper terminals across all or part of the tank circuit of an oscillator produces the desired F-M when an audio signal is applied.

It is the purpose of the study reported here to investigate yet another method of producing frequency modulation. It has been found that the input capacitance of a type of transistor, the MOSFET, varies with changes in input voltage. This varying capacitance shows promise as being the frequency modulating reactance of an F-M transmitter. A MOSFET F-M generator would be superior to other methods if it would make use of commercially available components, and, in addition, it would be more compact in size than transmitters employing reactance tubes. Such a generator could be used in light-weight, portable transmitters.

A d-c analysis is used in this dissertation to study the frequency modulation produced by the input capacitance, and an a-c analysis is used to study the linearity of the capacitance characteristic. In Chapter II an explanation of the input capacitance is

given. In Chapter III the experimental results of the input capacitance as a function of the gate bias are examined, and in Chapter VI the frequency versus gate bias characteristics of a frequency modulator stage are presented. In addition, Chapter IV presents a mathematical analysis of the differences between a linear and a nonlinear relationship of input capacitance as a function of input voltage.

CHAPTER II

THE MOSFET

A. Physical Description and Characteristics

The MOSFET (Metal-Oxide Semiconductor Field-Effect Transistor) is one of the more recent developments in the area of transistor technology. It differs from conventional transistors because it is unipolar instead of bipolar. Instead of employing both holes and electrons as charge carriers, MOSFET devices may use either holes or electrons as carriers but not both at the same time. The MOSFET also differs from junction field-effect transistors because it utilizes a thin layer of insulation between the gate and the channel. This insulating layer gives the MOSFET its characteristically high input resistance which may be greater than 10^{14} ohms.

The cross section of a typical p-channel MOS structure is shown in Figure 2. The device consists of an n-type bulk or substrate into which two p-type areas have been diffused. An insulating layer that is from 800 to 2,000 Å thick is then grown over the area between the two p-type areas. The insulation is generally made of silicon dioxide which is highly nonconductive. Next, metal contacts are overlayed on the two p-type areas to form the source and drain electrodes and on the insulating layer to form the gate electrode. The drain, gate and source may be compared to the collector, base and emitter, respectively, of a bipolar transistor.

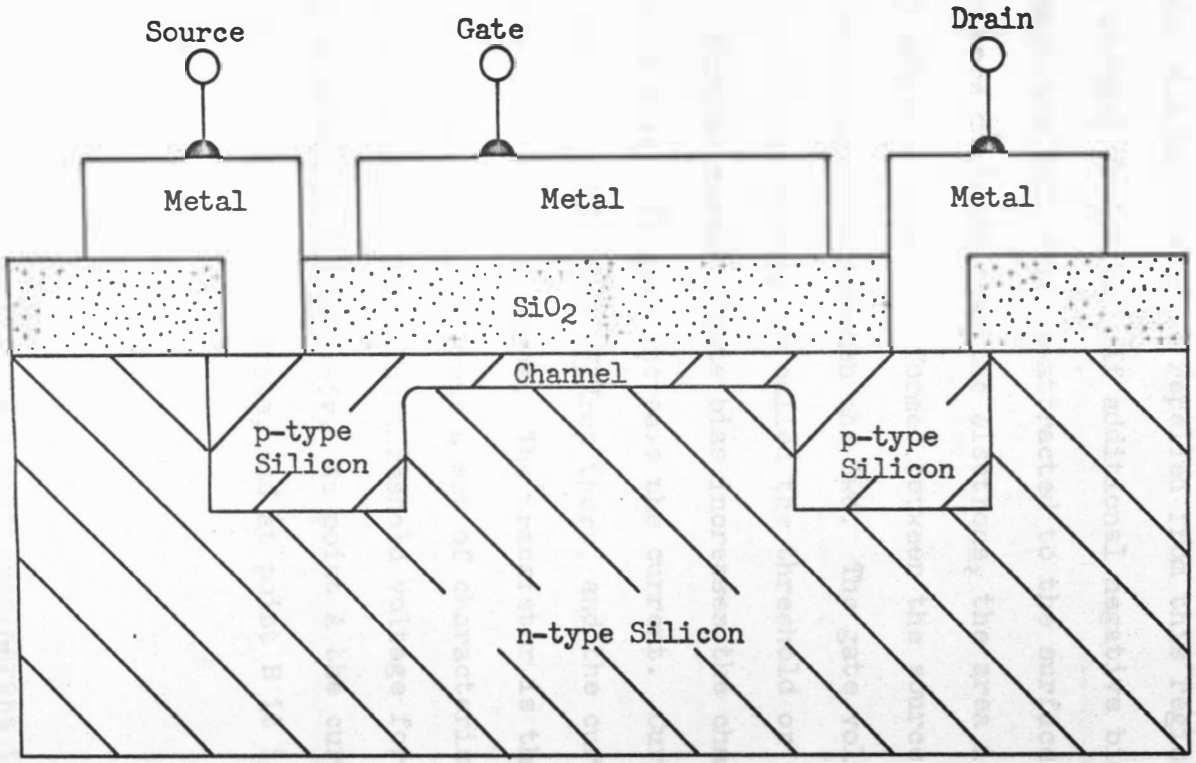


Figure 2. Cross Section of a p-channel MOSFET

A MOSFET is a voltage-controlled device. When bias is applied to the gate, the region under the insulation and between the two p-type areas, known as the channel, is affected. If negative bias is applied, electrons will be repelled from this region. This is known as channel depletion. If additional negative bias is applied, holes within the bulk will be attracted to the surface. When the holes are more dominant than the electrons, the area is said to be inverted; and a p-channel is formed between the source and the drain which allows conduction between the two. The gate voltage which just allows significant current is called the threshold or pinchoff voltage. Further negative gate bias increases the channel current. Proper drain bias will also increase the current. Current in the channel, however, causes an IR drop there, and the current levels off to an almost constant value. The transistor is then in a saturation state. Figure 3 shows a set of characteristic curves for a typical p-channel MOSFET. The threshold voltage for this transistor is approximately $V_{GS} = -5v$. At point A the current for $V_{GS} = -8v$ is just beginning to saturate, and at point B it is well into saturation.

B. Detailed Theory of Operation

In order to understand the capacitance across the gate, a more detailed analysis must be made of the interaction between the gate electrode and the semiconductor bulk. The following discussion

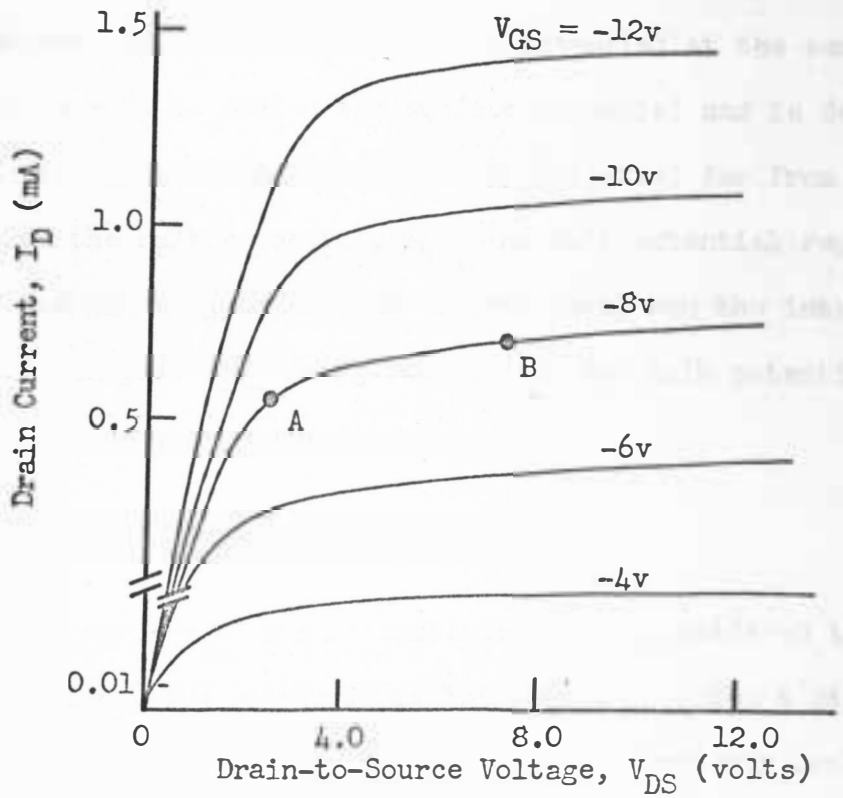


Figure 3. Drain Characteristics for a Typical p-channel MOSFET

pertains to a p-channel MOSFET although an analogous discussion could be made for an n-channel device.

Figure 4 shows an energy band diagram for the gate-oxide-semiconductor interface. The reference level for all of the electrostatic potentials is arbitrarily chosen as the Fermi level, E_F, of the semiconductor, and these potentials are positive towards the valence band. The electrostatic potential at the semiconductor surface, $x = 0$, is called the surface potential and is designated as ϕ_s; the value of the electrostatic potential far from the surface is called the bulk potential, ϕ_b. The bulk potential represents the difference in energy between the Fermi level and the intrinsic Fermi level or mid-gap, E_i. Using the surface and bulk potentials, the surface barrier is thus defined as

$$V_s = \phi_s - \phi_b \tag{14}$$

The surface of a semiconductor is not considered to be a distinct plane, but rather it is the region in which a given property differs from the bulk. An example of such a property could be the energy levels of the semiconductor. Whenever the periodicity of a crystal is disturbed or interrupted as at the surface, energy states are introduced which can be either in an allowed or a forbidden band. These energy states are called surface states. A surface state is generally positive; that is, it usually acts as an ionized donor.

Because of the positive surface states, electrons from within the bulk are attracted to the surface. This results in an accumulation

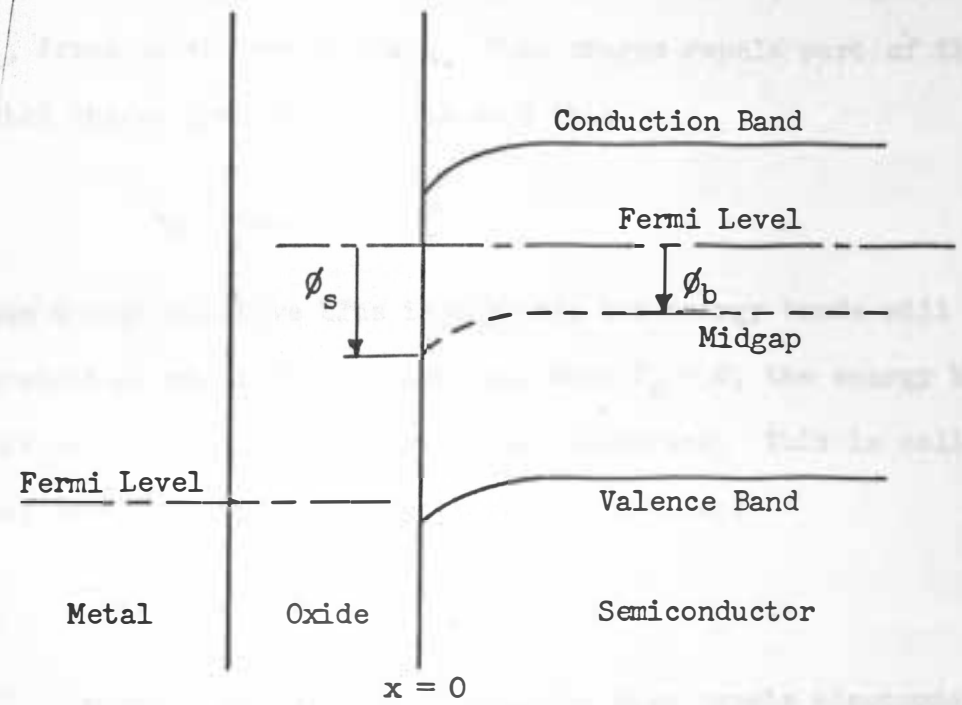


Figure 4. Energy Band Diagram for an n-type MOSFET

of electrons at the surface and explains why the energy bands in the semiconductor are bent downward at the surface. For n-type materials this accumulation condition exists if V_s is greater than zero. The charge which accumulates, Q_a, is equal to the charge of the surface states, Q_{ss}, such that the net charge is zero.

When negative bias is applied to the gate, a negative charge, Q_m, forms on the metal plate. This charge repels part of the accumulated charge from the surface such that

$$Q_m + Q_{ss} + Q_a = 0 \tag{15}$$

When enough negative bias is applied, the energy bands will straighten out at the surface; and when $V_s = 0$, the energy bands are flat out to the oxide-semiconductor interface. This is called the flat band condition. For this case $Q_a = 0$, and

$$Q_{ss} + Q_m = 0 \tag{16}$$

Further application of negative bias repels electrons from the surface which are associated with donor impurities. The donor impurities are uncompensated and leave a net positive charge, Q_d, at the surface in an area called the depletion region. The energy bands at the surface are now turned upward. When the bias becomes large enough, the intrinsic Fermi level will intersect the Fermi level at the surface, and the surface will have gone from n-type to intrinsic.

Additional bias attracts mobile holes to the surface and forms an inversion layer there. The charge which forms is positive and is

labeled Q_i. The total charge due to the depletion and the inversion process is called the space charge, Q_{sc}, and the region which includes the depletion region and the inversion region is called the space-charge region. When the holes become the dominant carriers at the surface, conduction can take place between the source and the drain. This first occurs when ϕ_s is just greater than ϕ_b; however, appreciable conduction does not take place until $\phi_s = 2\phi_b$. The bias voltage at this point is the aforementioned threshold voltage, and the summation of charges is now

$$Q_m + Q_{ss} + Q_{sc} = 0 \tag{17}$$

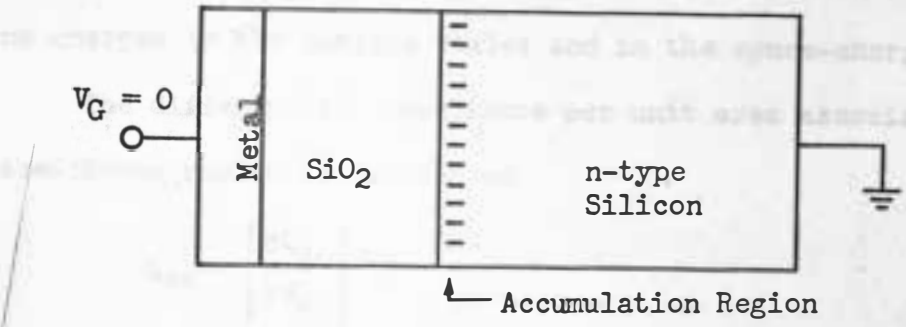
where Q_m is negative and Q_{ss} and Q_{sc} are positive. Figure 5 shows the changes in charge carriers for different gate biases.

C. The MOS Capacitor

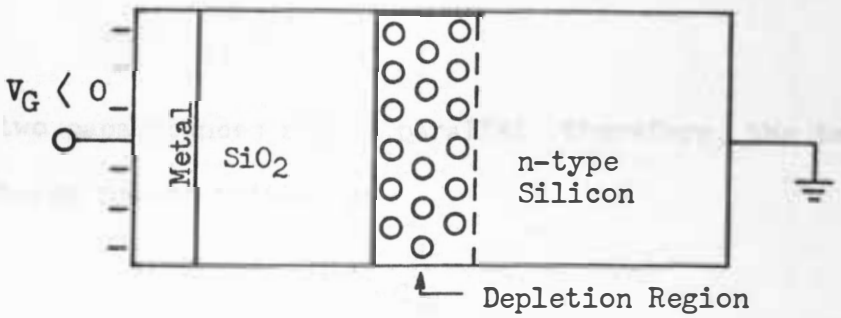
The region between the gate electrode and the semiconductor bulk forms a parallel plate capacitor. If the thickness of the oxide insulation, x_d, is assumed to be much greater than the thickness of the space-charge region, then the total capacitance, C, can be considered as a series combination of the oxide capacitance, C_o, and the surface capacitance, C_s. The oxide capacitance is a constant value and is given by

$$C_o = k_o \epsilon_o / x_d \tag{18}$$

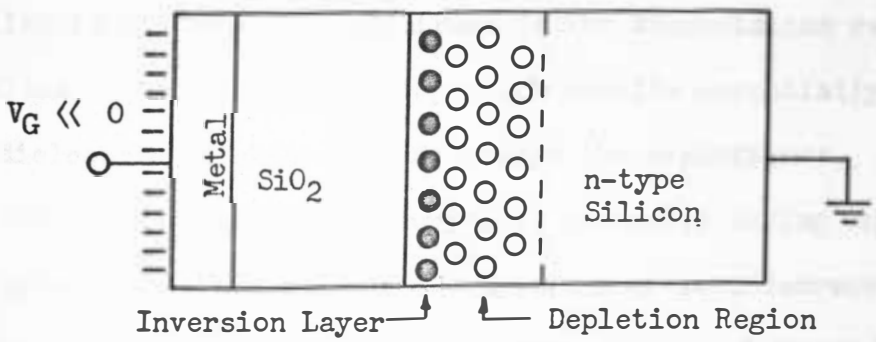
where k_o is the dielectric constant of the oxide layer and ϵ_o is



(a) Accumulation



(b) Depletion



(c) Inversion

Figure 5. The Changes in Charge Carriers in a MOSFET for Different Gate Biases. The dark and light circles represent positive holes in the inversion and depletion regions, respectively.

the permittivity of free space. The surface capacitance results from charges in the surface states and in the space-charge region.

The differential capacitance per unit area associated with the space-charge region is defined as

$$C_{sc} = \left| \frac{\partial Q_{sc}}{\partial V_s} \right| \tag{19}$$

Similarly, the differential capacitance associated with a change of charge in the surface states is defined by

$$C_{ss} = \left| \frac{\partial Q_{ss}}{\partial V_s} \right| \tag{20}$$

These two capacitances are in parallel; therefore, the total surface capacitance may be written as

$$C_s = \left| \frac{\partial Q_T}{\partial V_s} \right| = C_{sc} + C_{ss} \tag{21}$$

Although the oxide capacitance is a constant, the surface capacitance is affected by the gate bias. As negative bias is applied to the gate, the electrons in the accumulation region are repelled. The depletion region which results essentially increases the dielectric thickness and decreases the capacitance. Further application of negative bias produces inversion during which holes are attracted to the interface. This, in effect, decreases the oxide thickness and increases the capacitance. Figure 6 shows the variation of total capacitance as a function of the gate bias.

The capacitance versus gate bias characteristic shows a difference in the capacitance depending upon the frequency of the

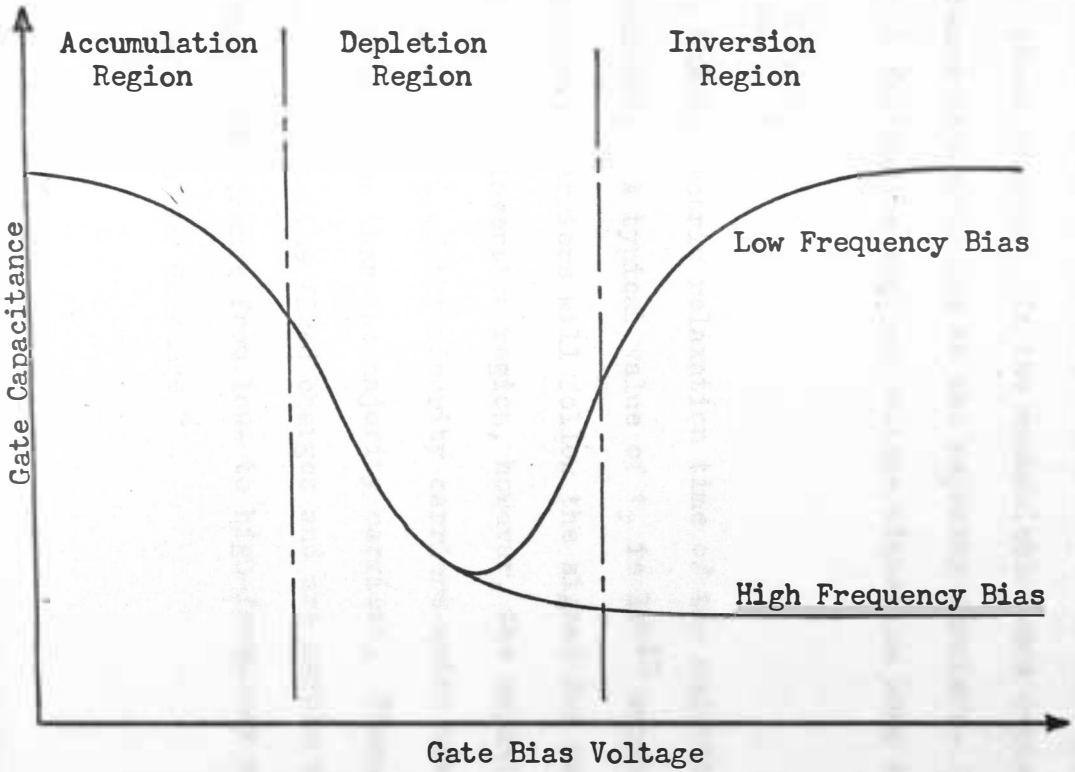


Figure 6. Gate Capacitance Across an MOS Structure as a Function of Gate Bias

voltage signal. In the accumulation and depletion regions the high and low frequency curves are the same, but in the inversion region the high frequency curve has a smaller capacitance than the low frequency curve. This effect can be explained by the type of carriers involved in these regions. In the accumulation and depletion regions the capacitance is mainly due to the majority carriers. These carriers will follow the applied voltage signal as long as

$$\frac{1}{\omega} \gg \tau_r \tag{22}$$

where τ_r is the dielectric relaxation time of the majority carriers in a semiconductor. A typical value of τ_r is 10^{-12} seconds;[2] therefore, the majority carriers will follow the signal for most frequencies of concern. In the inversion region, however, the capacitance is predominantly a result of the minority carriers which have a much longer relaxation time than the majority carriers. Therefore, the minority carriers act like fixed charges and are unable to follow the a-c signal. The change from low- to high-frequency behavior usually occurs in the 100 Hz range.[2]

D. Hysteresis

If the capacitance of the MOSFET is to be used in a frequency modulator, it is important to consider the possibility of hysteresis in the capacitance curve. There are two types of hysteresis which have been noted.[16] The first type is observed only for p-type semiconductors and is caused by charge migration on the surface of the

oxide layer. It produces a large distortion in the capacitance versus bias characteristic in the inversion region.

The second type of hysteresis may be found in either p-type or n-type units. It is caused by differences in the depth to which the surface states are filled and emptied, and it produces a small shift in the capacitance curve to either the left or right. The depth to which surface states or traps in the oxide can be filled, Z_F, or emptied, Z_E, is dependent upon the density of free electrons or holes, respectively, at the surface. For a p-type unit it has been found that an inversion layer exists at the semiconductor surface; therefore, the density of electrons at the surface is much greater than the density of the holes, and Z_F is greater than Z_E. Because of this there will be a difference in the net charge trapped for increasing and decreasing bias. C_{ss} will therefore be changed, and different values of surface capacitance will be found for increasing and decreasing bias. Figure 7 shows the hysteresis characteristic that will result from the above effect. It can be seen that the curve for increasing bias lies to the left of that for decreasing bias. For some cases these two curves will be reversed. This condition is caused by a situation similar to that described above, only for these cases the traps involved are at the metal-oxide interface instead of at the oxide-semiconductor interface.

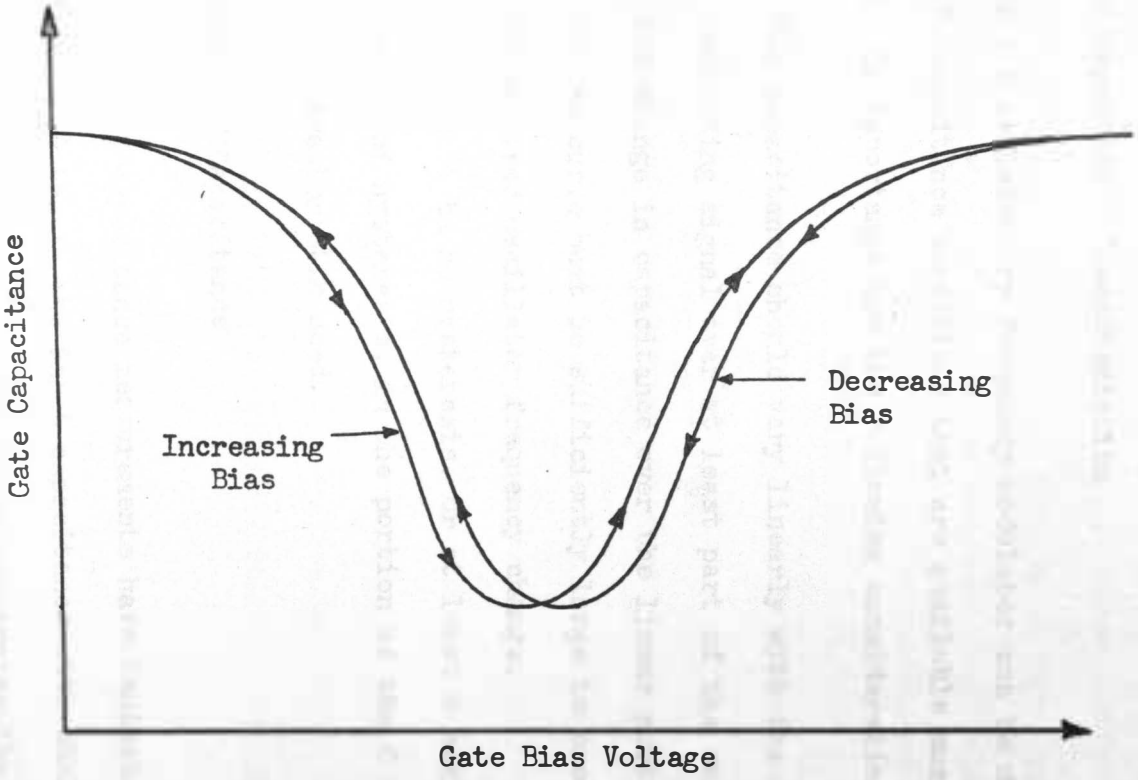


Figure 7. Hypothetical Hysteresis Effect due to Charges in the Surface States of a MOSFET

CHAPTER III

EXPERIMENTAL STUDY OF MOSFET INPUT CAPACITANCE

A. Desired Capacitance Characteristics

Before a satisfactory frequency modulator can be developed, the type of capacitance variations that are available must be determined. Of importance are the following considerations:

1. The capacitance should vary linearly with the modulating signal over at least part of the curve.
2. The change in capacitance over the linear portion of the curve must be sufficiently large to produce the desired oscillator frequency change.
3. There must be no hysteresis, or at least a negligible amount of hysteresis, in the portion of the C versus V_G characteristic used.

B. Measurement of Capacitance

Previous MOS capacitance measurements have indicated as much as a four-to-one change in the input capacitance with changes in the gate voltage.[1] Unfortunately, information concerning the type of device used to obtain those results is not available. Two different MOSFET types were available for the present study. The RCA 3N128 is an n-channel, depletion type, and the Motorola 2N3796 is an n-channel type which may be used in either the enhancement or the depletion

mode. Two type 3N128 samples and one type 2N3796 device were studied. The drain characteristics of the 3N128 are shown in Figure 8.

The circuit used for the capacitance measurements was the same for all of the MOSFET samples; a diagram of the biasing circuit is shown in Figure 9. The circuit was designed as it would actually be later used in the frequency modulator. The input capacitance was then measured on a Boonton RX Meter Type 250A with a 13510A Transistor Test Jig attachment. The ground at the gate is internal to the RX Meter.

The 7.5 megohm resistor across the gate and source is for the protection of the MOSFET. The thin insulating layer of the MOSFET makes the device very vulnerable to damage due to accumulated static charge. If sufficient static charge builds up between the gate and the source, the charge can puncture the insulating layer; this causes the input resistance of the transistor to be greatly decreased. The resistor across the input, therefore, prevents such static charge buildup during handling and operation.

The two 2.2μF capacitors, C_S and C_D, have two functions. First, they act as an a-c bypass for the carrier frequency. C_S is in series with C_{gs} and C_D is in series with C_{ds}; therefore, the equation for the input capacitance, C_i, is

$$C_i = \frac{C_S C_{gs}}{C_S + C_{gs}} + \frac{C_D C_{ds}}{C_D + C_{ds}} \qquad (23)$$

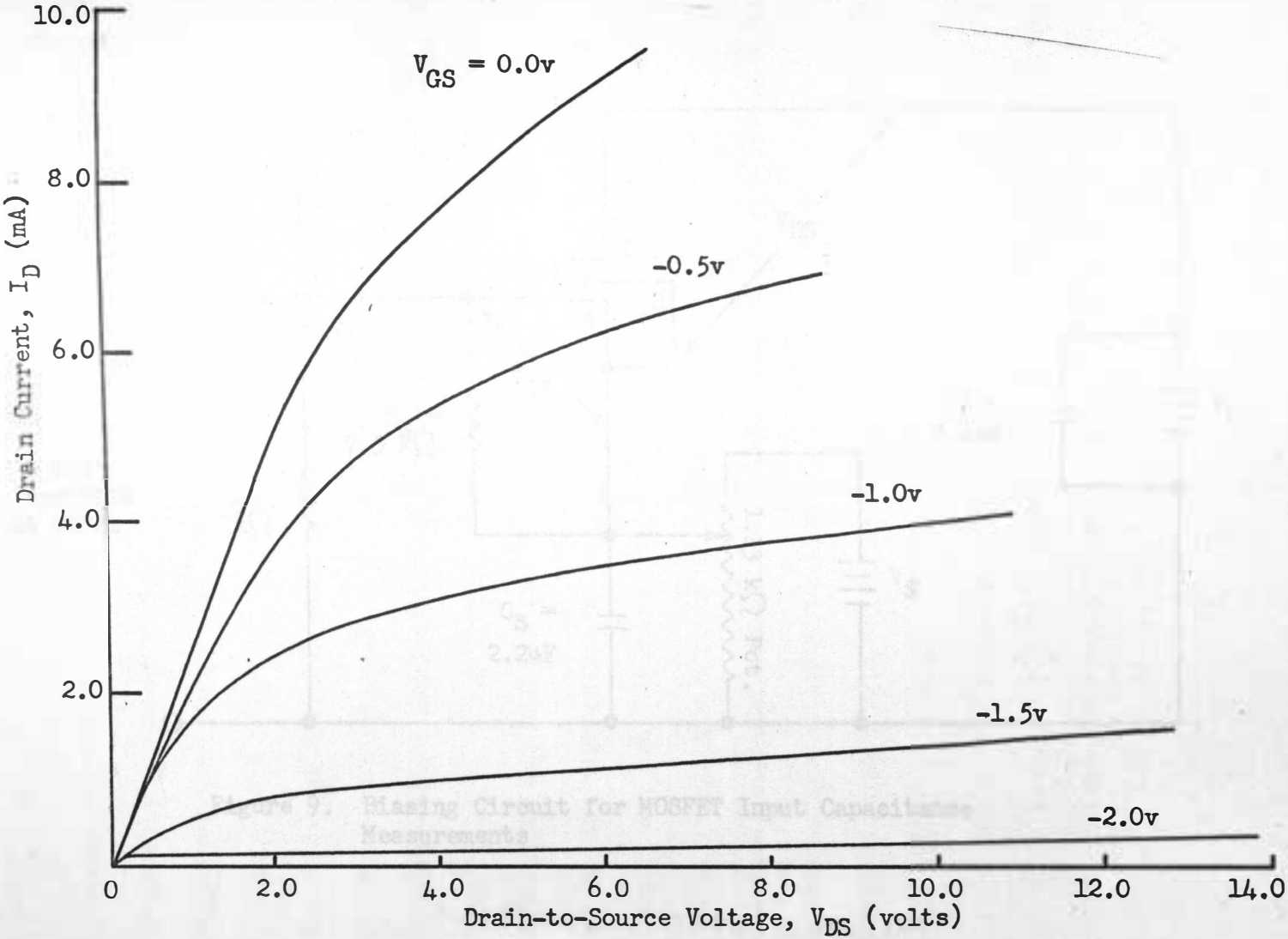


Figure 8. Drain Characteristics of the 3N128 MOSFET

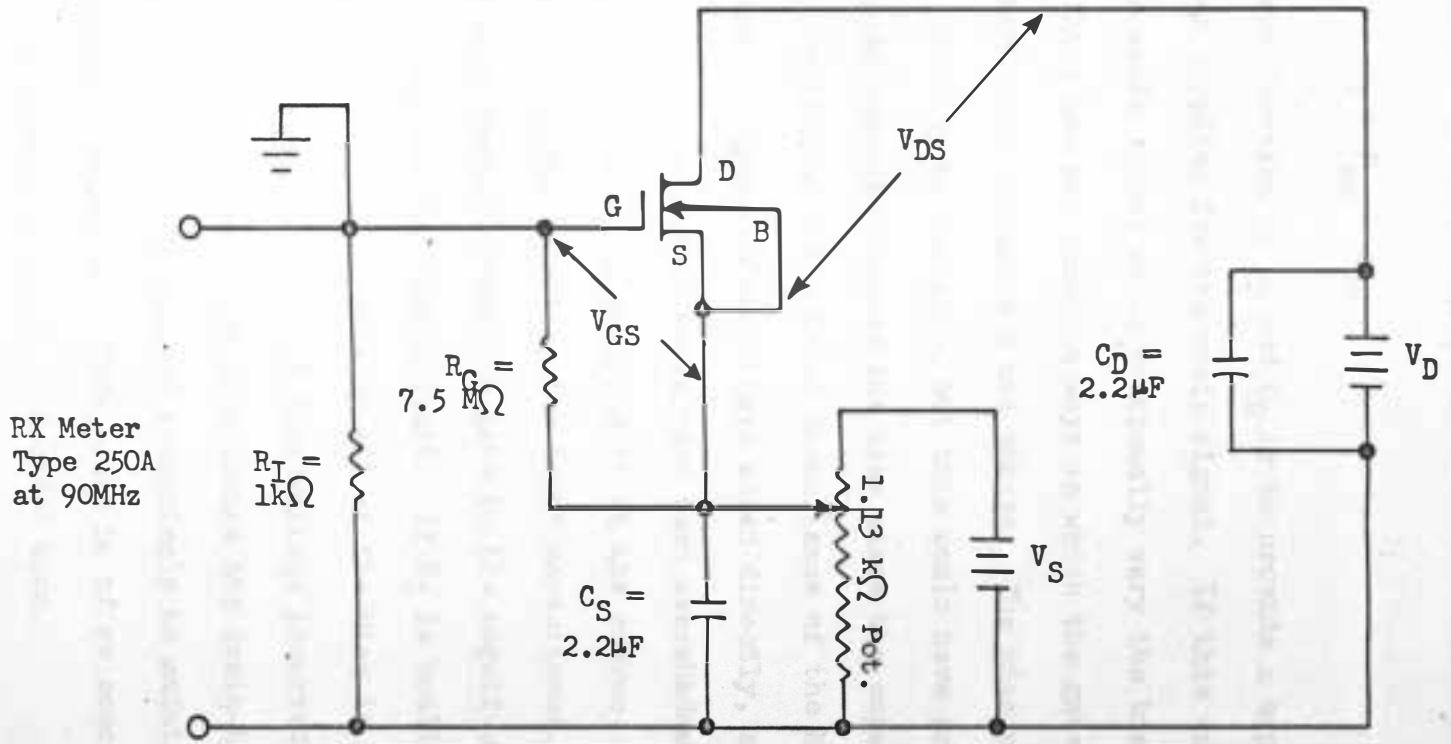


Figure 9. Biasing Circuit for MOSFET Input Capacitance Measurements

but since C_S and C_D are much greater than C_{gs} and C_{ds}, the input capacitance can be written as

$$C_i = C_{gs} + C_{ds} \tag{24}$$

The second function of C_S and C_D is to provide a bypass path across the power supplies for the audio signal. If this was not allowed for, the audio signal would continually vary the bias point.

There are two possible ways in which the gate could be biased; the method shown in Figure 9 was chosen. The bias could have been applied directly to the gate, but this would have prohibited the use of a bypass capacitor across the bias since the capacitor would have been in parallel with the input capacitance of the MOSFET. Values of the two capacitances would have added directly, and the input capacitance of the MOSFET would have been overshadowed. By reversing the sign of the bias and applying it at the source, the supply can be bypassed without affecting the input capacitance. A resistor, R_I, has been connected from the gate to the negative terminal of the supply to complete the bias circuit. If R_I is small compared to the 7.5 megohm resistor, essentially all of the bias is applied across the gate and source. The only disadvantage incurred with this method is that changes in gate bias also change the drain-to-source voltage; therefore, V_D has to be changed accordingly to maintain a constant drain-to-source potential. This would be of no concern, however, once a fixed gate bias point is decided upon.

The potentiometer was placed across V_S to give better control over the gate bias. The resistance of this potentiometer has to be relatively small in order that all desired bias points can be reached. If the potentiometer resistance is of the order of tens of kilohms, then a large drain current will cause a large voltage drop across the potentiometer, and the drain bias supply will have to be increased considerably.

The a-c signal generated within the RX Meter was set at 90 MHz since the meter gave more accurate readings at higher frequencies and this was the approximate frequency that the input capacitance would be subjected to in the oscillator tank circuit.

C. Experimental Results

When the input capacitance of the sample transistors was measured, the drain-to-source voltage was kept constant and the gate bias was incrementally varied. This gave curves of input capacitance as a function of gate bias.

Figure 10 shows a comparison between the 2N3796 and the two 3N128 MOSFET's for $V_{DS} = 2.0$ volts. The 3N128 transistors are labeled 3N128A and 3N128B to distinguish between them. The 3N128A and 3N128B samples have very similar characteristics; however, the 2N3796 does not exhibit as much change in capacitance in the accumulation and depletion regions. All three transistors show the expected changes in capacitance with increasing negative gate voltage. The accumulation region extends to approximately $V_{GS} =$

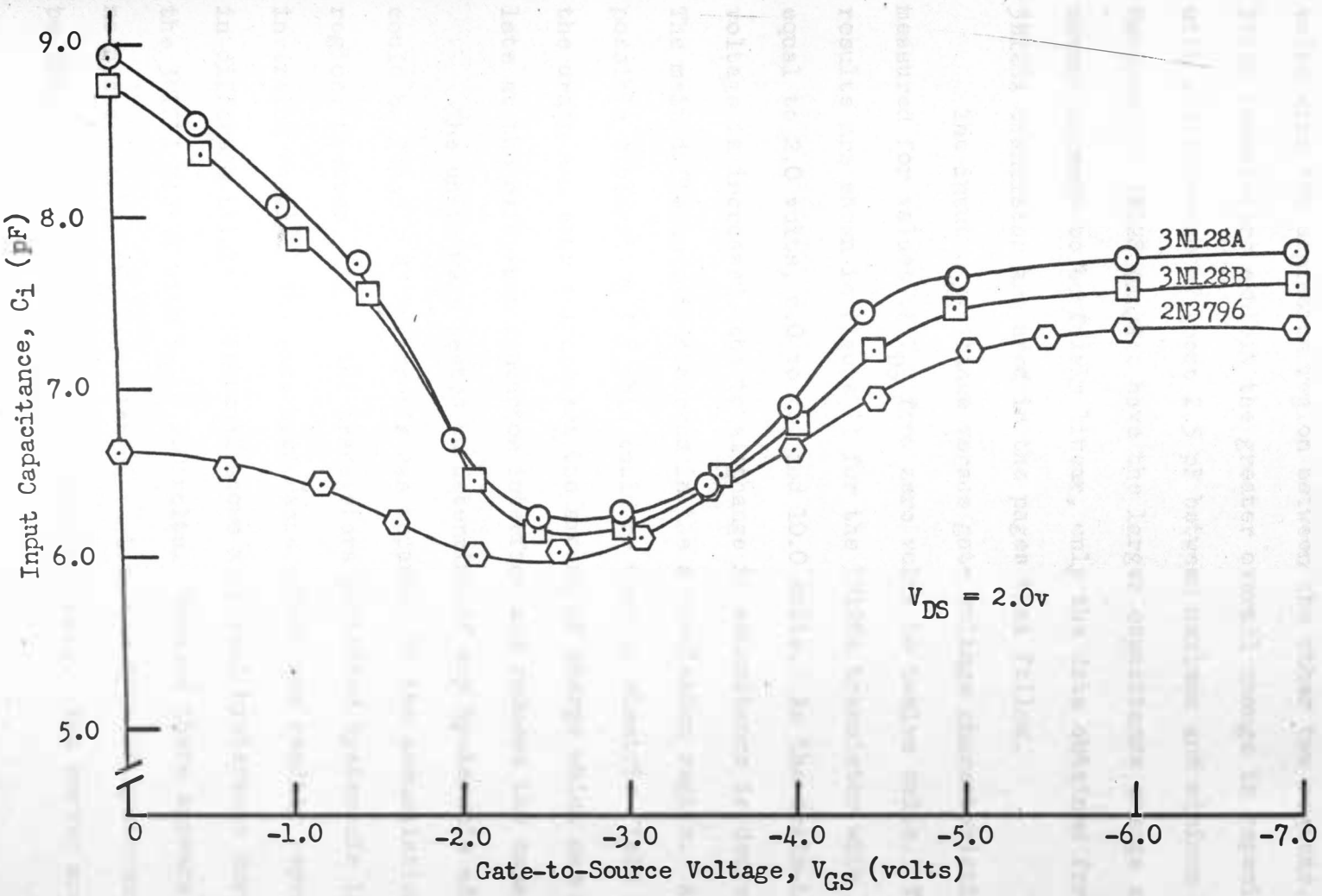


Figure 10. Input Capacitance as a Function of Gate-to-Source Voltage for Three MOSFET's

-1.5 volts, and the inversion region begins at about V_{GS} = -3.8 volts with the depletion region between the other two regions. The 3N128 transistors exhibit the greater overall change in capacitance with a difference of about 2.5 pF between maximum and minimum points. Because the 3N128 samples have the larger capacitance change and this change appears to be fairly linear, only the data obtained for the 3N128A transistor are used in the pages that follow.

The input capacitance versus gate voltage characteristics were measured for values of V_{DS} from zero volts to twelve volts. The results are shown in Figure 11 for the 3N128A transistor with V_{DS} equal to 2.0 volts, 6.0 volts, and 10.0 volts. As the drain-to-source voltage is increased, the total change in capacitance is decreased. The main difference is observed in the accumulation region. A possible explanation for this could be that an electric field between the drain and source increases the amount of charge which can accumulate at the oxide-semiconductor interface and reduces the capacitance.

The units were tested to determine if any hysteresis effects could be found. No hysteresis was apparent in the accumulation region; however, all of the transistors exhibited hysteresis in the inversion region. The same hysteresis effect was readily reproduced in different trials. Figure 12 shows a typical hysteresis curve for the 3N128A MOSFET with V_{DS} = 2.0 volts. Because there appears to be no hysteresis in the accumulation region, the type of hysteresis may be that which is found in p-type MOSFET's only. The curves are

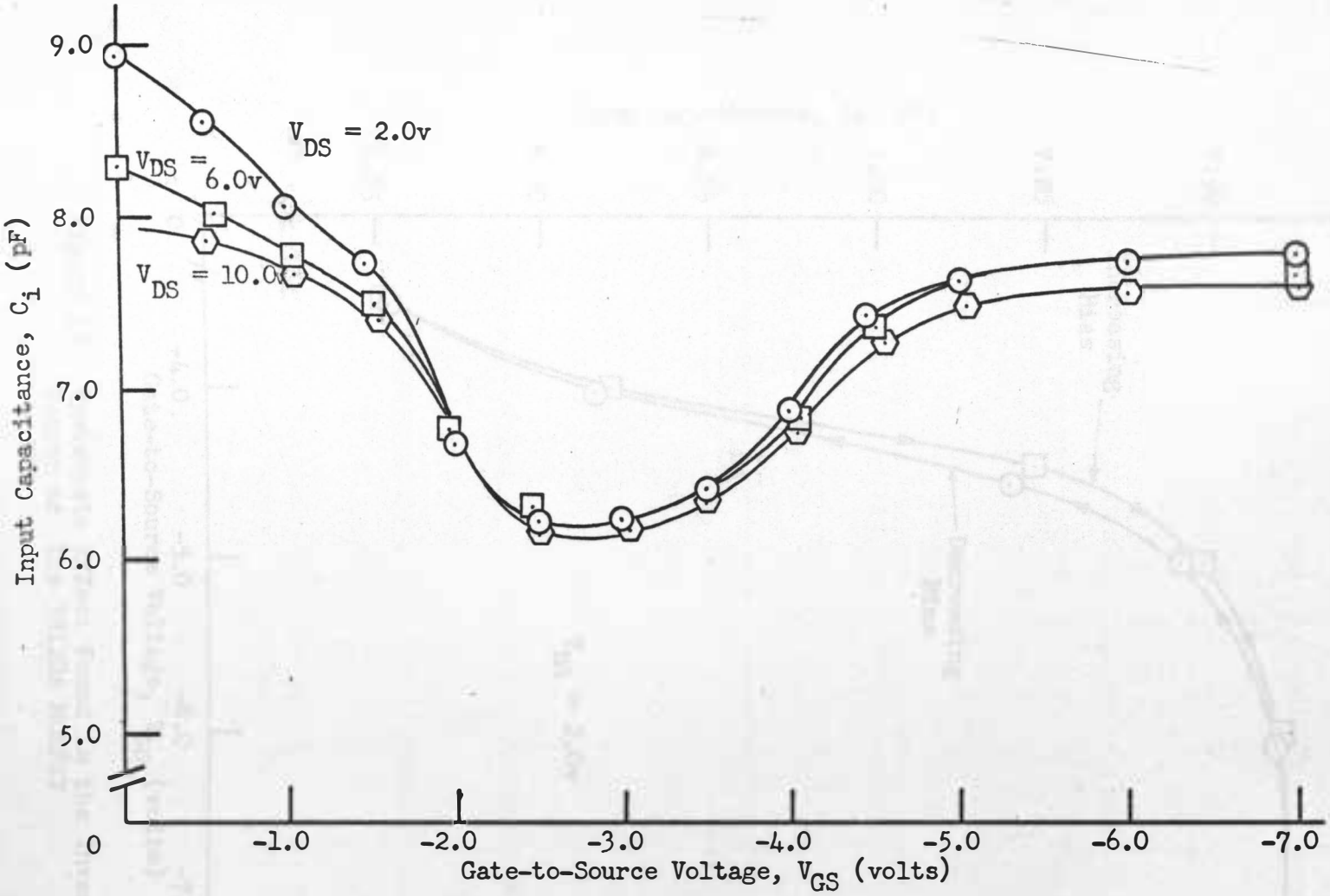


Figure 11. Input Capacitance Versus Gate-to-Source Voltage Characteristics of the 3N128A MOSFET for Three Values of Drain-to-Source Voltage

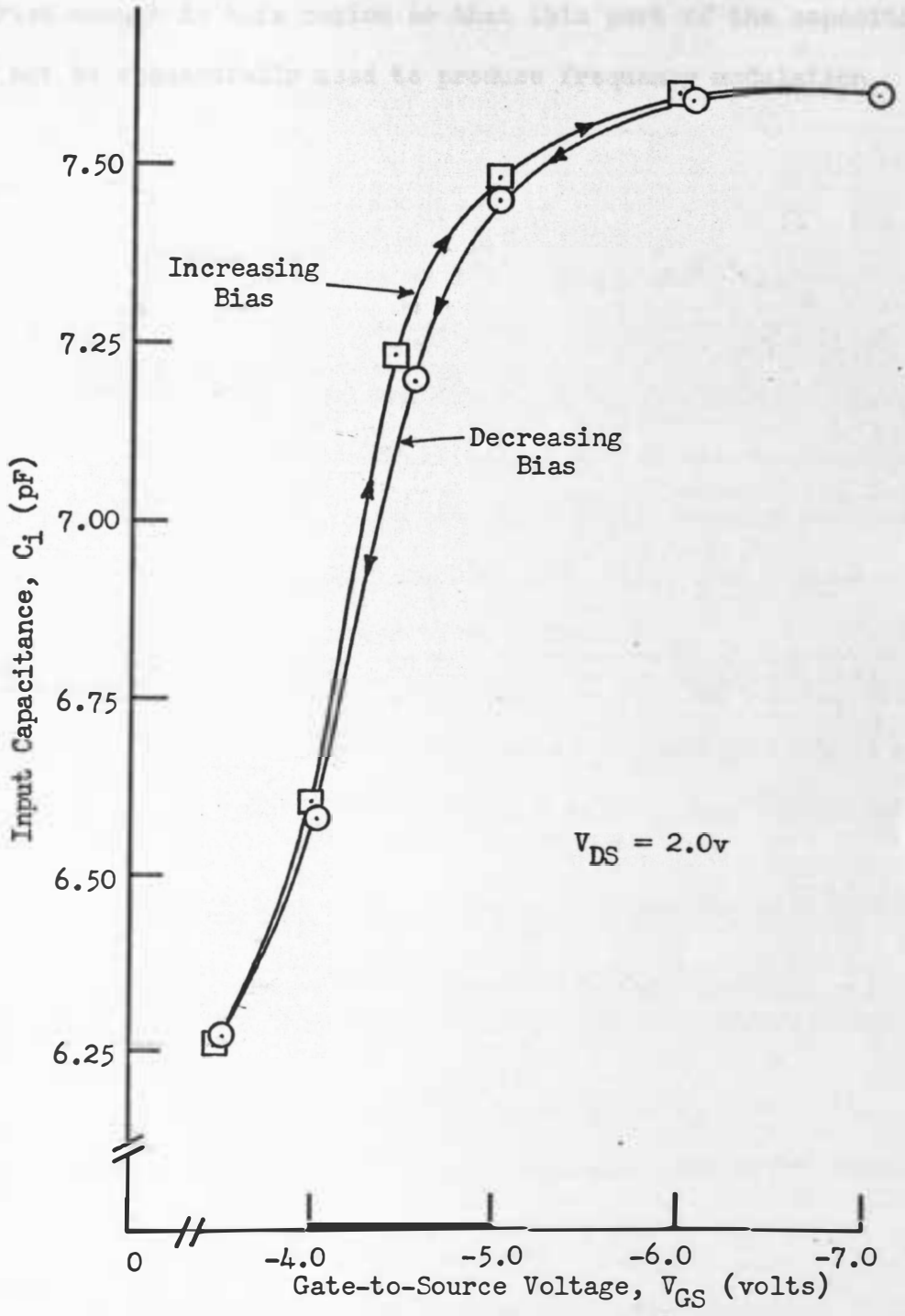


Figure 12. Hysteresis Effect Found in the Inversion Region of the 3N128A MOSFET

distorted enough in this region so that this part of the capacitance could not be successfully used to produce frequency modulation.

CHAPTER IV

MODULATOR ANALYSIS

A. Linear Analysis

In the beginning of Chapter III it was stated that the capacitance should vary linearly with the modulating signal over at least part of the curve. But how important is it that the characteristic be linear, and what effect does a nonlinear relationship between the capacitance and gate bias have on the carrier and sidebands produced during frequency modulation? It is the purpose of the following analysis to answer these questions.

Consider a C-versus-V_{GS} characteristic such as the one shown in Figure 13. If the section between points A and B is assumed to be linear, then the relationship between C and V_{GS} may be written as

$$C = a' + b\left|V_{GS}\right| \tag{25}$$

where a' is the C intercept of the extended linear portion, and b is the slope of the linear section.

If the d-c gate-to-source voltage is fixed at some point, such as Q, in the middle of the linear portion of the curve, and if a time-varying signal, v_{gs}, is applied at the gate, then C may be redefined as

$$C = a' + b\left(\left|V_{GS}\right| + v_{gs}\right) \tag{26}$$

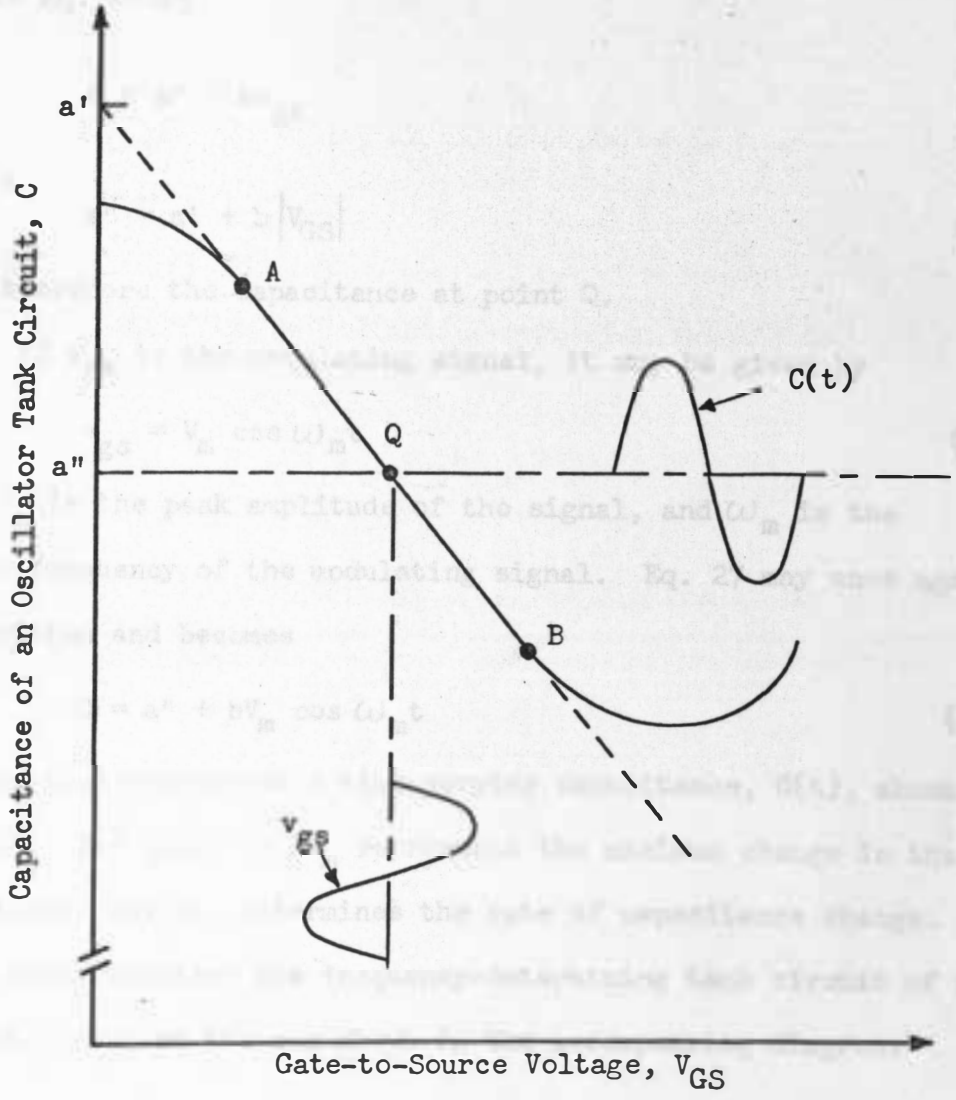


Figure 13. Capacitance Versus Gate-to-Source Voltage Characteristic to which a Modulating Signal, v_{gs}, is Applied

Since $b\left|V_{GS}\right|$ will be a constant in practical applications, we may rewrite Eq. 26 as

$$C = a'' + bv_{gs} \tag{27}$$

where

$$a'' = a' + b\left|V_{GS}\right| \tag{28}$$

a" is therefore the capacitance at point Q.

If v_{gs} is the modulating signal, it may be given by

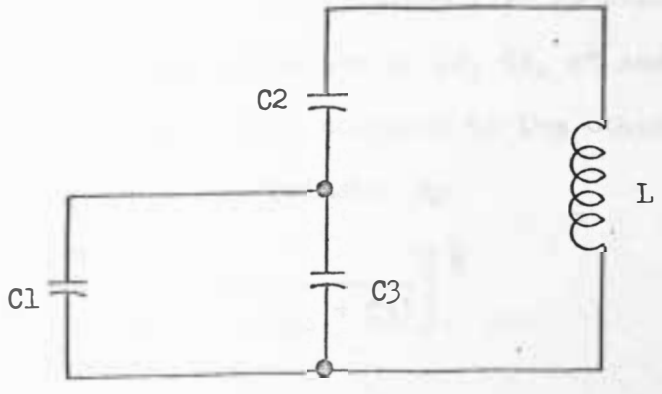
$$v_{gs} = V_m \cos \omega_m t \tag{29}$$

where V_m is the peak amplitude of the signal, and ω_m is the angular frequency of the modulating signal. Eq. 27 may once again be rewritten and becomes

$$C = a'' + bV_m \cos \omega_m t \tag{30}$$

This equation represents a time-varying capacitance, C(t), about the value a". The quantity bV_m represents the maximum change in the capacitance, and ω_m determines the rate of capacitance change.

Next, consider the frequency-determining tank circuit of an oscillator such as the one shown in the accompanying diagram.



The total capacitance of the tank circuit is

$$C = \frac{(C1 + C3)(C2)}{C1 + C2 + C3} \tag{31}$$

The center angular frequency of the oscillator is

$$\omega_c = \frac{1}{(LC)^{\frac{1}{2}}} \tag{32}$$

Substituting Eq. 31 into the above expression results in

$$\omega_c = \frac{1}{\left[\frac{LC2(C1 + C3)}{C1 + C2 + C3}\right]^{\frac{1}{2}}} \tag{33}$$

If C1 represents a time-varying capacitance given by Eq. 27, then Eq. 33 may be written as

$$\omega_c = \frac{1}{\left[\frac{LC2(a'' + bv_{gs} + C3)}{a'' + bv_{gs} + C2 + C3}\right]^{\frac{1}{2}}} \tag{34}$$

or

$$\omega_c = \left[\frac{a'' + bv_{gs} + C2 + C3}{LC2(a'' + bv_{gs} + C3)}\right]^{\frac{1}{2}} \tag{35}$$

In order for the resonant frequency to be primarily affected by variations in C1, typical values of C2, C3, a" and bv_{gs} would dictate that C2 be large-valued compared to the other elements. Therefore, Eq. 35 can be approximated by

$$\omega_c = \left[\frac{C2}{LC2(a'' + bv_{gs} + C3)}\right]^{\frac{1}{2}} \tag{36}$$

Eq. 36 reduces to

$$\omega_c = \frac{1}{\left[L(a + bV_m \cos \omega_m t)\right]^{\frac{1}{2}}} \tag{37}$$

where

$$a = a'' + C3 \tag{38}$$

Eq. 37 may be expanded in a binomial series, and if

$$bV_m \ll a \tag{39}$$

then

$$\omega_c = \frac{1}{(La)^{\frac{1}{2}}} \left(1 - \tfrac{1}{2} \frac{bV_m}{a} \cos \omega_m t\right) \tag{40}$$

This equation for ω_c is analogous to Eq. 4 in Chapter I for ω_c which was

$$\omega_c = \omega_o + k_f V_m \cos \omega_m t \tag{41}$$

Therefore, it follows that

$$\omega_o = \frac{1}{(La)^{\frac{1}{2}}} \tag{42}$$

$$k_f = \frac{-b}{2a(La)^{\frac{1}{2}}} \tag{43}$$

and

$$m_f = \frac{k_f V_m}{\omega_m} = \frac{-bV_m}{2a\,\omega_m (La)^{\frac{1}{2}}} \tag{44}$$

The equation for the carrier and sidebands can thus be written as

$$v = V_c \left\{ J_o \left(\frac{-bV_m}{2a\ \omega_m (La)^{\frac{1}{2}}} \right) \cos \frac{1}{(La)^{\frac{1}{2}}} t \right.$$

$$- J_1 \left(\frac{-bV_m}{2a\ \omega_m (La)^{\frac{1}{2}}} \right) \left[\cos \left(\frac{1}{(La)^{\frac{1}{2}}} - \omega_m \right) t - \cos \left(\frac{1}{(La)^{\frac{1}{2}}} + \omega_m \right) t \right]$$

$$+ J_2 \left(\frac{-bV_m}{2a\ \omega_m (La)^{\frac{1}{2}}} \right) \left[\cos \left(\frac{1}{(La)^{\frac{1}{2}}} - 2\omega_m \right) t + \cos \left(\frac{1}{(La)^{\frac{1}{2}}} + 2\omega_m \right) t \right]$$

$$\left. - + \ldots \right\} \tag{45}$$

B. Nonlinear Analysis

If the variable capacitance is not a linear function of V_{GS}, but is some nonlinear relation such as a binomial, then

$$C = a' + b(|V_{GS}| + v_{gs}) + c(|V_{GS}| + v_{gs})^2 \tag{46}$$

This may be expanded into

$$C = a' + b|V_{GS}| + bv_{gs} + c|V_{GS}|^2$$

$$+ 2c|V_{GS}|v_{gs} + cv_{gs}^2 \tag{47}$$

and rewritten as

$$C = m + nv_{gs} + cv_{gs}^2 \tag{48}$$

where

$$m = a' + b|V_{GS}| + c|V_{GS}|^2 \tag{49}$$

and

$$n = b + 2c\left|V_{GS}\right| \tag{50}$$

Substituting Eq. 29 for v_{gs} into Eq. 48 gives

$$C = m + nV_m \cos \omega_m t + cV_m^2 \cos^2 \omega_m t \tag{51}$$

Therefore, for this case, from Eq. 32, we obtain

$$c = \frac{1}{\left[L(m + nV_m \cos \omega_m t + cV_m^2 \cos^2 \omega_m t)\right]^{\frac{1}{2}}} \tag{52}$$

This expression may also be expanded by using the binomial series, and if

$$nV_m \ll m \tag{53}$$

and

$$cV_m^2 \ll m \tag{54}$$

then

$$\omega_c = \frac{1}{(Lm)^{\frac{1}{2}}}\left(1 - \tfrac{1}{2}\frac{nV_m}{m}\cos \omega_m t - \tfrac{1}{2}\frac{cV_m^2}{m}\cos^2 \omega_m t\right) \tag{55}$$

Since

$$\cos^2 \omega_m t = \tfrac{1}{2}(1 + \cos 2\omega_m t) \tag{56}$$

ω_c may be written as

$$\omega_c = \frac{1}{(Lm)^{\frac{1}{2}}}\left[1 - \tfrac{1}{2}\frac{nV_m}{m}\cos \omega_m t - \tfrac{1}{4}\frac{cV_m^2}{m}(1 + \cos 2\omega_m t)\right] \tag{57}$$

In Eq. 3 of Chapter I it was shown that

$$v = V_c \cos\left[\int_o^t \omega_c dt\right] \tag{58}$$

Substituting Eq. 57 into this expression yields

$$v = V_c \cos\left\{\int_o^t \left[\frac{1}{(Lm)^{\frac{1}{2}}}\left(1 - \tfrac{1}{2}\frac{nV_m}{m}\cos\omega_m t - \frac{cV_m^2}{4m} - \frac{cV_m^2}{4m}\cos 2\omega_m t\right)\right]dt\right\} \tag{59}$$

Upon integration Eq. 59 becomes

$$v = V_c \cos\left[\frac{1}{(Lm)^{\frac{1}{2}}}t - \frac{nV_m}{2m\,\omega_m(Lm)^{\frac{1}{2}}}\sin\omega_m t - \frac{cV_m^2}{4m(Lm)^{\frac{1}{2}}}t - \frac{cV_m^2}{8m\,\omega_m(Lm)^{\frac{1}{2}}}\sin 2\omega_m t\right] \tag{60}$$

By making the following substitutions

$$u' = \frac{1}{(Lm)^{\frac{1}{2}}}t \tag{61}$$

$$u = \frac{-nV_m}{2m\,\omega_m(Lm)^{\frac{1}{2}}} \tag{62}$$

$$v' = \frac{cV_m^2}{4m(Lm)^{\frac{1}{2}}}t \tag{63}$$

$$v = \frac{-cV_m^2}{8m\,\omega_m(Lm)^{\frac{1}{2}}} \tag{64}$$

v may be simplified to

$$v = V_c \cos(u' + u \sin \omega_m t + v' + v \sin 2\omega_m t) \tag{65}$$

The cosine term may be rewritten by using the trigonometric identity for cos(x + y) where

$$x = u' + u \sin \omega_m t \tag{66}$$

and

$$y = v' + v \sin 2\omega_m t \tag{67}$$

This results in

$$\cos(u' + u \sin \omega_m t) \cos(v' + v \sin 2\omega_m t)$$

$$- \sin(u' + u \sin \omega_m t) \sin(v' + v \sin 2\omega_m t) \tag{68}$$

Once more using trigonometric identities for the sum of two angles, Eq. 68 may be rewritten as

$$\left[(\cos u')(\cos u \sin \omega_m t) - (\sin u')(\sin u \sin \omega_m t)\right]$$

$$\times \left[(\cos v')(\cos v \sin 2\omega_m t) - (\sin v')(\sin v \sin 2\omega_m t)\right]$$

$$- \left[(\sin u')(\cos u \sin \omega_m t) + (\cos u')(\sin u \sin \omega_m t)\right]$$

$$\times \left[(\sin v')(\cos v \sin 2\omega_m t) + (\cos v')(\sin v \sin 2\omega_m t)\right] \tag{69}$$

These functions may be evaluated by using Bessel functions.

Therefore, Eq. 69 becomes

$$\left\{\cos u' \left[J_o(u) + 2J_2(u) \cos 2\omega_m t + 2J_4(u) \cos 4\omega_m t + \ldots\right]\right.$$
$$\left. - \sin u' \left[2J_1(u) \sin \omega_m t + 2J_3(u) \sin 3\omega_m t + \ldots\right]\right\}$$
$$\times \left\{\cos v' \left[J_o(v) + 2J_2(v) \cos 4\omega_m t + 2J_4(v) \cos 8\omega_m t + \ldots\right]\right.$$
$$\left. - \sin v' \left[2J_1(v) \sin 2\omega_m t + 2J_3(v) \sin 6\omega_m t + \ldots\right]\right\}$$
$$- \left\{\sin u' \left[J_o(u) + 2J_2(u) \cos 2\omega_m t + 2J_4(u) \cos 4\omega_m t + \ldots\right]\right.$$
$$\left. + \cos u' \left[2J_1(u) \sin \omega_m t + 2J_3(u) \sin 3\omega_m t + \ldots\right]\right\}$$
$$\times \left\{\sin v' \left[J_o(v) + 2J_2(v) \cos 4\omega_m t + 2J_4(v) \cos 8\omega_m t + \ldots\right]\right.$$
$$\left. + \cos v' \left[2J_1(v) \sin 2\omega_m t + 2J_3(v) \sin 6\omega_m t + \ldots\right]\right\} \quad (70)$$

By expanding Eq. 70 the following terms are obtained.

$$J_o(u)J_o(v) \left[\cos u' \cos v' - \sin u' \sin v'\right]$$
$$- 2J_o(v)J_1(u) \sin\omega_m t \left[\sin u' \cos v' + \cos u' \sin v'\right]$$
$$+ 2J_o(v)J_2(u) \cos 2\omega_m t \left[\cos u' \cos v' - \sin u' \sin v'\right]$$
$$- 2J_o(u)J_1(v) \sin 2\omega_m t \left[\cos u' \sin v' + \sin u' \cos v'\right]$$
$$+ 4J_1(u)J_1(v) \sin \omega_m t \sin 2\omega_m t \left[\sin u' \sin v' - \cos u' \cos v'\right]$$
$$+ - \ldots \quad (71)$$

Eq. 71 may be reduced to

$$J_o(u)J_o(v)\cos(u' + v')$$

$$- 2J_o(v)J_1(u)\sin\omega_m t \sin(u' + v')$$

$$+ 2J_o(v)J_2(u)\cos 2\omega_m t \cos(u' + v')$$

$$- 2J_o(u)J_1(v)\sin 2\omega_m t \sin(u' + v')$$

$$- 4J_1(u)J_1(v)\sin\omega_m t \sin 2\omega_m t \cos(u' + v') + - \ldots \quad (72)$$

If $c = 0$, then the terms in Eq. 72 should be the same as in the linear case. For $c = 0$

$$v' = 0 \quad (73)$$

and

$$v = 0 \quad (74)$$

Eq. 72 reduces to

$$J_o(u)\cos u' - 2J_1(u)\sin\omega_m t \sin u'$$

$$+ 2J_2(u)\cos 2\omega_m t \cos u' + \ldots \quad (75)$$

This may be expanded as

$$J_o(u)\cos u' - J_1(u)\left[\cos(u' - \omega_m t) + \cos(u' + \omega_m t)\right]$$

$$+ J_2(u)\left[\cos(u' - 2\omega_m t) - \cos(u' + 2\omega_m t)\right]$$

$$+ \ldots \quad (76)$$

Also when c = 0,

$$m = a' + bV_{GS} = a \tag{77}$$

and

$$n = b \tag{78}$$

Therefore,

$$u' = \frac{1}{(La)^{\frac{1}{2}}}\, t \tag{79}$$

and

$$u = \frac{-bV_m}{2a\,\omega_m(La)^{\frac{1}{2}}} \tag{80}$$

Using these expressions for u' and u, Eq. 76 becomes

$$J_o\left(\frac{-bV_m}{2a\,\omega_m(La)^{\frac{1}{2}}}\right)\cos\frac{1}{(La)^{\frac{1}{2}}}\,t$$

$$- J_1\left(\frac{-bV_m}{2a\,\omega_m(La)^{\frac{1}{2}}}\right)\left[\cos\left(\frac{1}{(La)^{\frac{1}{2}}} - \omega_m\right)t + \cos\left(\frac{1}{(La)^{\frac{1}{2}}} + \omega_m\right)t\right]$$

$$+ J_2\left(\frac{-bV_m}{2a\,\omega_m(La)^{\frac{1}{2}}}\right)\left[\cos\left(\frac{1}{(La)^{\frac{1}{2}}} - 2\omega_m\right)t - \cos\left(\frac{1}{(La)^{\frac{1}{2}}} + 2\omega_m\right)t\right]$$

$$+ - \ldots \tag{81}$$

which is the same as the terms in Eq. 45 for the linear case.

Let us examine the expression obtained for the nonlinear case. By using trigonometric identities, Eq. 71 may be written as

$$J_o(u)J_o(v)\cos(u' + v')$$

$$- J_o(v)J_1(u)\left[\cos(u' + v' - \omega_m t) - \cos(u' + v' + \omega_m t)\right]$$

$$+ J_o(v)J_2(u)\left[\cos(u' + v' - 2\omega_m t) + \cos(u' + v' + 2\omega_m t)\right]$$

$$- J_o(u)J_1(v)\left[\cos(u' + v' - 2\omega_m t) - \cos(u' + v' + 2\omega_m t)\right]$$

$$- J_1(u)J_1(v)\left[\cos(u' + v' - \omega_m t) + \cos(u' + v' + \omega_m t)\right]$$

$$+ J_1(u)J_1(v)\left[\cos(u' + v' - 3\omega_m t) + \cos(u' + v' + 3\omega_m t)\right]$$

$$+ - \ldots \qquad (82)$$

If the above expression is multiplied by V_m, it represents the frequency modulated wave with its carrier and sidebands. The frequencies of the carrier and sidebands are determined by the cosine angles, and their peak amplitudes are determined by V_m times the Bessel functions.

The last two terms in Eq. 82 are of particular interest. These two terms were obtained from the expansion of the last term in Eq. 71 which resulted in two separate side frequencies. This means that every time two terms which have different frequencies are multiplied together in Eq. 70, there will be two sidebands created at frequencies equal to $(u' + v')/t \pm (j\omega_m + k\omega_m)$ and $(u' + v')/t \pm (j\omega_m - k\omega_m)$ where $j\omega_m$ and $k\omega_m$ represent the frequencies of the multiplied terms. Therefore, the amplitudes of the side frequencies are infinite series of Bessel functions times

V_m. For example,

$$V_m(-J_o(v)J_1(u) - J_1(u)J_2(v) + \ldots)$$

is the peak amplitude of the first side frequency below the carrier. When $j = k$, then the frequency obtained is that of the carrier. Hence, the amplitude of the carrier also consists of a series of Bessel functions. Although the carrier and side frequencies for the nonlinear case appear to be different in amplitude than the carrier and corresponding side frequencies for the linear case, this is not a definite conclusion because the infinite Bessel function series has not been evaluated using circuit values.

It should be noted that the amplitudes of the side frequencies equally spaced above and below the carrier are no longer the same in the nonlinear case. For example, the amplitude of the first side frequency below the carrier is given by Eq. 83, but the amplitude of the first side frequency above the carrier is equal to

$$V_m(+ J_o(v)J_1(u) - J_1(u)J_2(v) + \ldots) \qquad (84)$$

Note the difference of the sign on the first Bessel function product. This means that even if one of the side frequencies for the nonlinear case has the same amplitude as the corresponding side frequency in the linear case, there will still be distortion because the sidebands are not symmetrical about the carrier for the nonlinear case.

Another striking difference also exists. The frequencies of the carrier and all of the side frequencies are shifted for the

nonlinear case. For the linear case the carrier frequency was equal to

$$\omega_c = \frac{1}{(La)^{\frac{1}{2}}} \tag{85}$$

But for the nonlinear case the carrier frequency is

$$\omega_c = u' + v' \tag{86}$$

or

$$\omega_c = \frac{1}{(Lm)^{\frac{1}{2}}}\left[1 - \frac{cV_m^2}{4m}\right] \tag{87}$$

However, it was assumed earlier that

$$cV_m^2 \ll m \tag{88}$$

Therefore, ω_c for the nonlinear case is

$$\omega_c = \frac{1}{(Lm)^{\frac{1}{2}}} \tag{89}$$

In other words, the frequency of the carrier has been reduced by the amount

$$\Delta\omega_c = \frac{1}{(La)^{\frac{1}{2}}} - \frac{1}{(Lm)^{\frac{1}{2}}} \tag{90}$$

Each of the side frequencies has been reduced by the same amount.

This shift in frequency, although unexpected, can be explained. The frequency of the carrier wave is determined by the capacitance of the tank circuit with no modulating signal plus the average capacitance change caused by the modulating signal. If the C-versus-V_{GS} curve is linear, a sinusoidal modulating signal will cause a

sinusoidal change in capacitance, and the average value of this capacitance change will be equal to zero. Therefore, the carrier frequency will simply be the frequency of the tank circuit with no modulating signal applied. However, if the C-versus-V_{GS} characteristic is not linear, the change in capacitance will not be sinusoidal. The average capacitance change caused by a modulating signal will, therefore, no longer be equal to zero. As a result the carrier frequency is shifted.

To determine the extent of the frequency shift of the carrier and sidebands, the following example is hypothesized. Let L = 0.1μH, a = 30 pF, b = -0.5, c = 0.01 and V_{GS} = -5.0 volts. Then, for the linear case,

$$f_c = \frac{\omega_c}{2\pi} = \frac{1}{2\pi(La)^{\frac{1}{2}}} \tag{91}$$

or

$$f_c = 92.0 \text{ MHz} \tag{92}$$

For the above conditions m = 30.25 pF. If V_m = 1.0 volts, then for the nonlinear case

$$f_c = \frac{\omega_c}{2\pi} = \frac{1}{2\pi(Lm)^{\frac{1}{2}}} \tag{93}$$

or

$$f_c = 91.65 \text{ MHz} \tag{94}$$

This represents a shift in frequency of 350 kHz which is definitely significant.

CHAPTER V

THE OSCILLATOR

The purpose of the oscillator is to create a radio-frequency signal that can be frequency modulated. The frequency of oscillation should be relatively stable compared to the desired variations in frequency caused by the audio signal, and the frequency should be dependent upon some variable capacitance in the circuit.

Of the various types of oscillator designs available, a Colpitts oscillator was selected because of its simplicity and because its frequency of oscillation is determined mainly by two capacitors and an inductor forming the tank circuit. A basic Colpitts oscillator design was given in Transistor Circuit Design;[8] the circuit was revised to coincide with the components that were available. Figure 14 shows the oscillator circuit and component values.

The active device in the oscillator is the 2N743 transistor. This N-P-N transistor is a high-speed switching device capable of operating at frequencies to 400 MHz. The collector characteristics for this transistor are shown in Figure 15. The operating point, Q, was chosen at $V_{CE} = 5$ volts, $I_C = 4.5$ mA. At this point $I_B = 0.2$ mA.

The d-c circuit for the oscillator includes the four resistors designated as R1, R2, R3 and R4. The voltage across R4 is

$$V_{R4} = I_C R4 = 2.43 \text{ volts} \qquad (95)$$

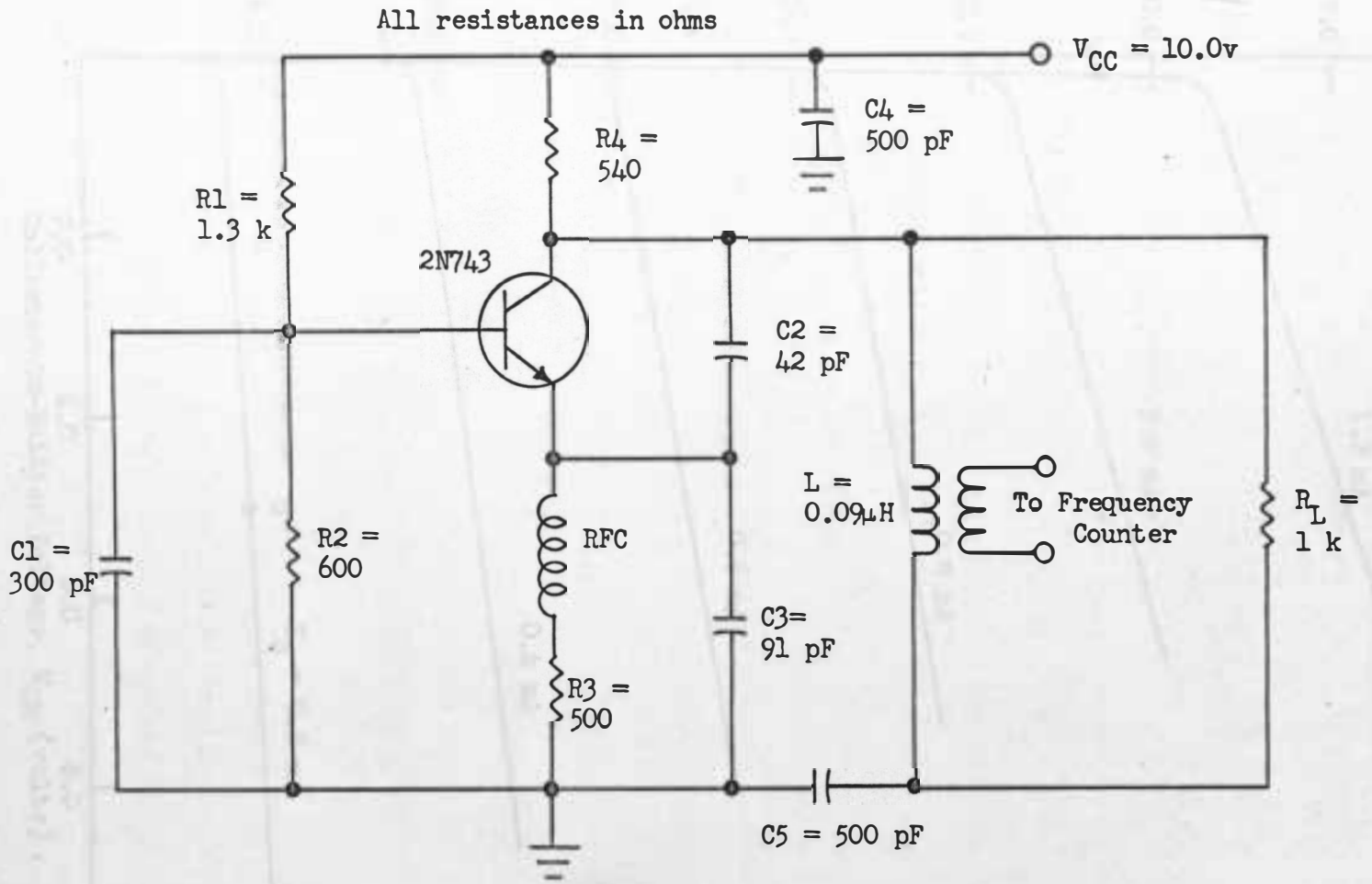


Figure 14. The Oscillator Circuit

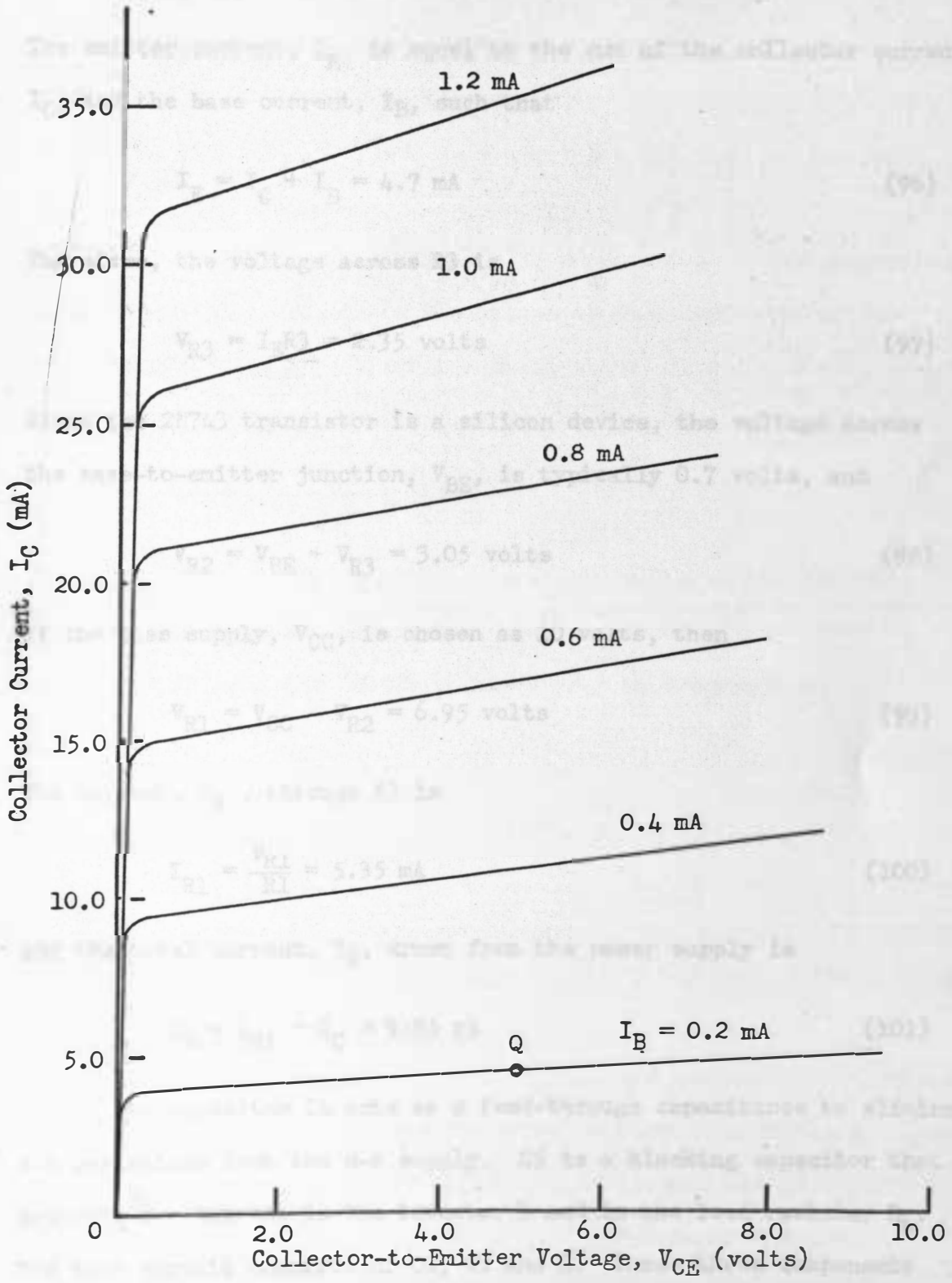


Figure 15. Collector Characteristics of the 2N743 Transistor

The emitter current, I_E, is equal to the sum of the collector current, I_C, and the base current, I_B, such that

$$I_E = I_C + I_B = 4.7 \text{ mA} \tag{96}$$

Therefore, the voltage across R3 is

$$V_{R3} = I_E R3 = 2.35 \text{ volts} \tag{97}$$

Since the 2N743 transistor is a silicon device, the voltage across the base-to-emitter junction, V_{BE}, is typically 0.7 volts, and

$$V_{R2} = V_{BE} + V_{R3} = 3.05 \text{ volts} \tag{98}$$

If the bias supply, V_{CC}, is chosen as 10 volts, then

$$V_{R1} = V_{CC} - V_{R2} = 6.95 \text{ volts} \tag{99}$$

The current, I_{R1}, through R1 is

$$I_{R1} = \frac{V_{R1}}{R1} = 5.35 \text{ mA} \tag{100}$$

and the total current, I_S, drawn from the power supply is

$$I_S = I_{R1} + I_C = 9.85 \text{ mA} \tag{101}$$

The capacitor C4 acts as a feed-through capacitance to eliminate a-c variations from the d-c supply. C5 is a blocking capacitor that prevents d-c current in the inductor L and in the load resistor R_L. The tank circuit consists of C2, C3 and L. These three components are mainly responsible for the frequency of the generated signal.

The transistor provides sufficient power gain at this frequency to overcome circuit losses and maintain unity gain around the feedback loop. The capacitor C1 provides a short-circuit path to the base for the feedback signal.

The frequency of oscillation is dependent upon the parameters of the transistor as well as the values of the components of the tank circuit. For this circuit a small-signal linear analysis yields the following complete expression for the frequency of oscillation:[3]

$$\omega_o^2 = \frac{h_{oe}}{C3C3h_{ie}} + \frac{(C2 + C3)}{LC2C3} \qquad (102)$$

where h_{oe} is the output admittance and h_{ie} is the short-circuit input impedance of the transistor. For practical values of the parameters the angular frequency may simply be expressed as

$$\omega_o^2 = \frac{C2 + C3}{LC2C3} \qquad (103)$$

Since

$$f_o = \omega_o/2\pi$$

the frequency of oscillation for the values of the tank circuit shown in Figure 13 should be 96 MHz.

When the oscillator was constructed, the measured d-c voltages were very nearly equal to the calculated values. The only significant difference was that V_{BE} measured 0.45 volts instead of the assumed value of 0.7 volts. The frequency of oscillation, however, was considerably different than the calculated value. The inductor of

the tank was inductively coupled to a Hewlett-Packard 524A Frequency Counter, and the frequency was measured to be approximately 82 MHz. The discrepancy between measured and calculated frequency may have been partly due to stray capacitances and to inductance in the leads between the circuit components.

The center frequency of oscillation, however, is not as crucial for the final frequency modulator analysis as is the stability of the oscillator. A ten-hour test was made, and during this period the frequency varied 6.2 kHz. This represents a change of 0.0076 percent of the reference oscillator frequency. For short periods of time, such as one-half hour, the total observed change was between 1 and 2 kHz. If the carrier signal was varied a total of 150 kHz, a 2 kHz drift would result in only a 1.3 percent total error. This would be satisfactory for a d-c analysis of the frequency as a function of the gate voltage.

CHAPTER VI

THE FREQUENCY MODULATOR

A. The Circuit

The two essential items for frequency modulation are now available: a radio-frequency oscillator to produce the carrier; and a device, the MOSFET, the capacitance of which changes with variations in input voltage. These two systems are connected together as in Figure 16 to produce the frequency modulator. The gate of the MOSFET is connected into the tank circuit of the oscillator through a 500 pF capacitor. The 500 pF capacitor blocks the d-c of the MOSFET circuit from the oscillator circuit. It is in series with the input capacitance of the MOSFET, but it is large enough so as to have a negligible effect on the capacitance network.

The input capacitance of the MOSFET can be connected to the tank circuit either in parallel with both C2 and C3 or exclusively in parallel with C3. It is desirable to connect it at the location that will produce the most practical change in frequency. To determine this, the partial derivatives of f_o with respect to C_i were taken for both connections.

When C_i is in parallel with both C2 and C3, the expression for f_o is

$$f_o = \frac{1}{2\pi \left[L\left(C_i + \frac{C2C3}{C2 + C3} \right) \right]^{\frac{1}{2}}} \qquad (105)$$

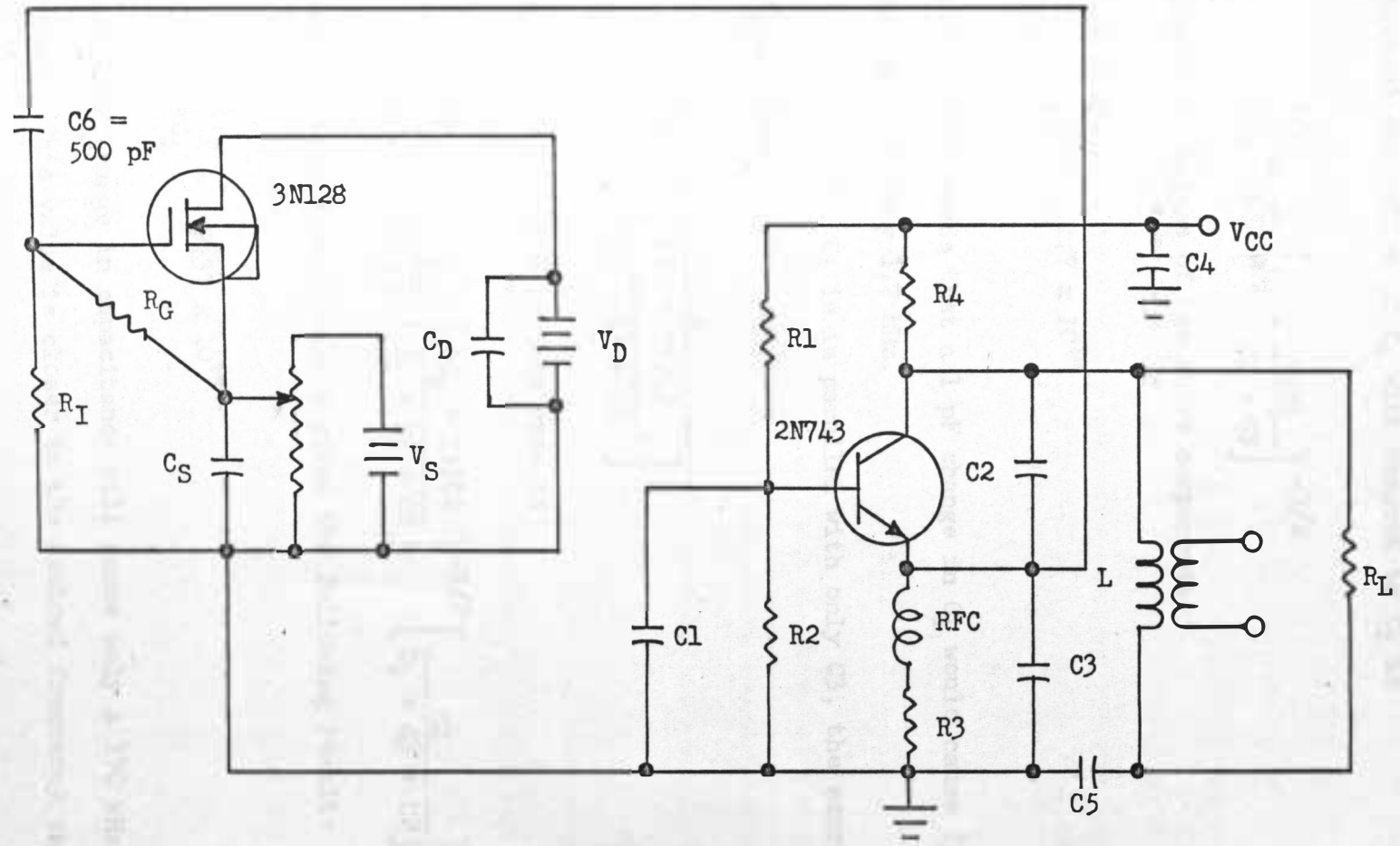


Figure 16. The Frequency Modulator Consisting of the MOSFET Biasing Circuit Coupled Through C6 into the Tank Circuit of the Oscillator

The partial derivative of f_o with respect to C_i is

$$\frac{\partial f_o}{\partial C_i} = \frac{1}{4\pi L}\left[C_i + \frac{C2C3}{C2 + C3}\right]^{-3/2} \tag{106}$$

For practical values of the above components

$$\frac{\partial f_o}{\partial C_i} = -1.7 \times 10^{18} \tag{107}$$

This derivative means that a 1 pF change in C_i would cause f_o to change approximately 1.7 MHz.

If, however, C_i is in parallel with only C3, the expression for f_o is then

$$f_o = \frac{1}{2\pi\left[\frac{(C_i + C3)C2}{C_i + C3 + C2} L\right]^{\frac{1}{2}}} \tag{108}$$

The partial derivative in this case is

$$\frac{\partial f_o}{\partial C_i} = -\frac{1}{2\pi\sqrt{L}}\left[\frac{(C_i + C3)C2}{C_i + C3 + C2}\right]^{-3/2}\left[\frac{C2}{C_i + C3 + C2}\right]^2 \tag{109}$$

and substituting circuit values gives the following result:

$$\frac{\partial f_o}{\partial C_i} = -0.37 \times 10^{18} \tag{110}$$

Thus, a 1 pF change in capacitance will cause only a 370 kHz in frequency. This value is closer to the desired frequency change

for an F-M transmitter. It should be noted that for both cases the change in frequency is opposite in sign to the change in capacitance.

The value of C3 was changed to 82 pF to keep a total capacitance of approximately 90 pF in series with C2. When the MOSFET circuit was connected to the oscillator, however, the frequency of the output signal dropped to about 71.5 MHz. This discrepancy may again be partly due to additional stray capacitances and inductances.

B. Experimental Results

The frequency of the oscillator was measured as the gate-to-source voltage was varied. Figure 17 shows the characteristic obtained for the 3N128A transistor in the accumulation and depletion regions. The curve is relatively linear between - 0.2 volts and - 0.9 volts, but it becomes nonlinear as the MOSFET enters the depletion region. The total change in oscillator frequency is 195 kHz; however, over the linear region the change is only about 90 kHz.

Figure 18 shows the f_o-versus-V_{GS} characteristic in the inversion region. This part of the curve was examined for hysteresis, but no noticeable difference was observed for increasing and decreasing biases. The hysteresis found in the C_i-versus-V_{GS} curve is probably present, but its effect is overshadowed by the other capacitors in the tank circuit.

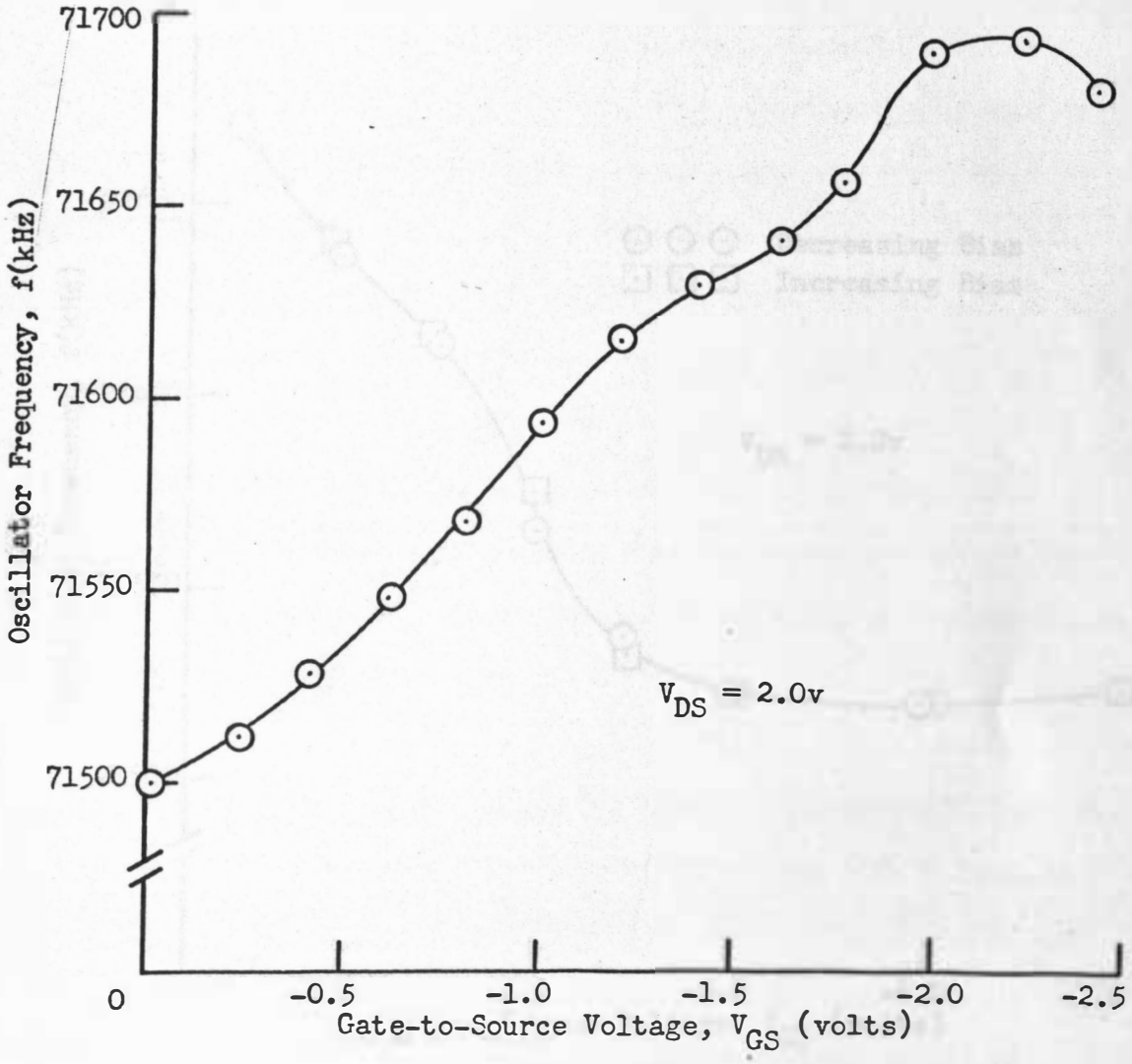


Figure 17. Oscillator Frequency as a Function of Gate-to-Source Voltage for the 3N128A MOSFET in the Accumulation and Depletion Regions

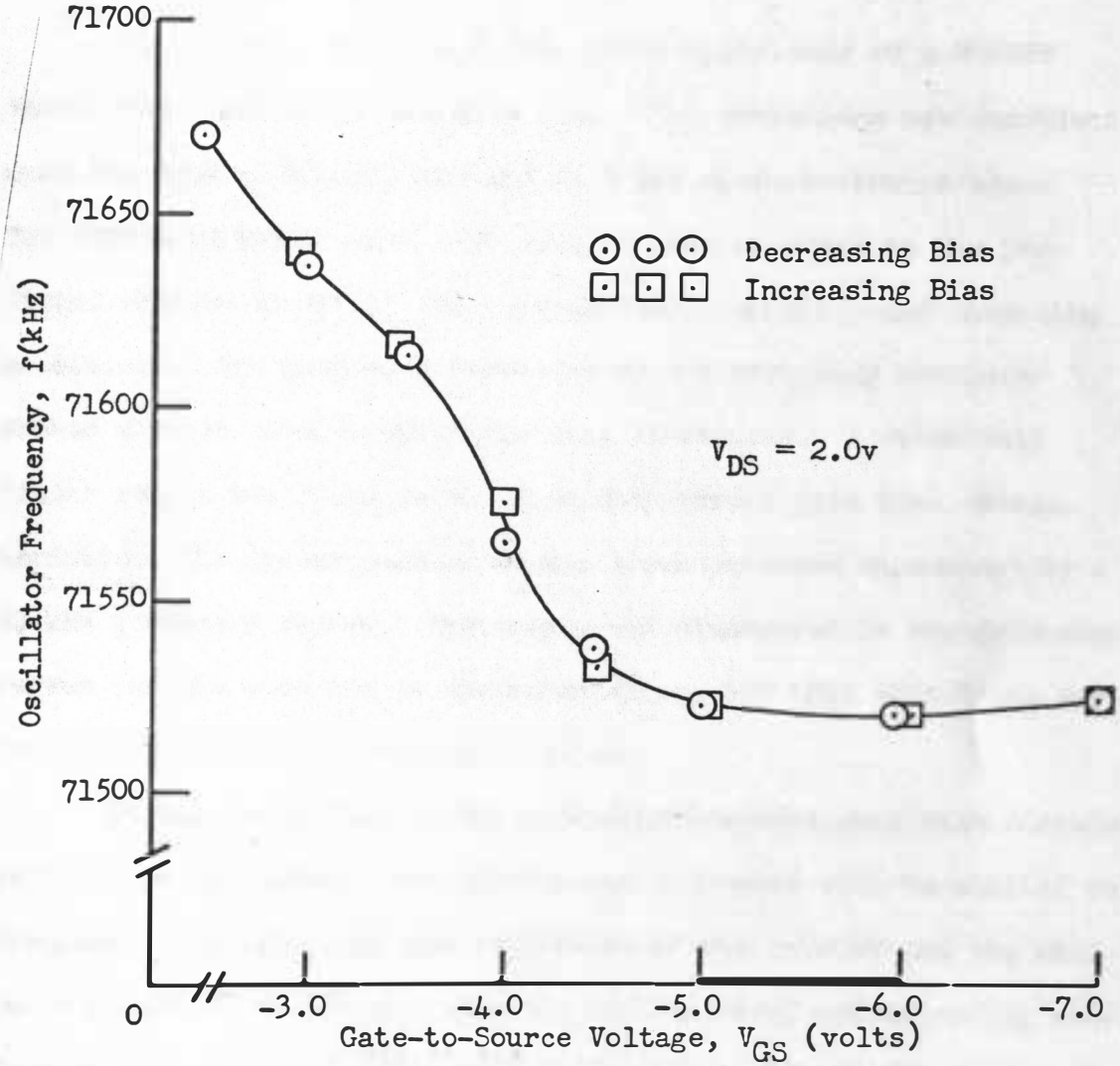


Figure 18. Oscillator Frequency as a Function of Gate-to-Source Voltage for the 3N128A MOSFET in the Inversion Region

CHAPTER VII

CONCLUSIONS

It has been shown that the input capacitance of a MOSFET varies with changes in the gate bias. The variations are dependent upon the type of MOSFET used and upon the drain-to-source bias. The nature of the changes that occurred corresponded to the predicted changes caused by the accumulation, depletion and inversion conditions. The generated frequency of the frequency modulator showed similar changes with gate bias variations. A relatively linear region was found in the frequency versus gate bias characteristic. The linear portion of the curve produced approximately a 90 kHz frequency change. Hysteresis was discovered in the inversion region for the capacitance characteristics, but this effect was not found in the frequency characteristics.

It was found that if the capacitance versus gate bias characteristic is not linear, the carrier and sidebands will be shifted in frequency. In addition, the amplitudes of the carrier and the sidebands appear to be changed, and the amplitudes of corresponding side frequencies above and below the carrier are no longer the same.

This work represents an introductory analysis of the MOSFET as a frequency modulator. The results obtained were generally favorable as far as the actual construction of an F-M transmitter is concerned. Further study in this area should include some of the following considerations:

1. Investigation of the input capacitance characteristics of other MOSFET types.

2. Design and construction of a frequency modulator with a lower carrier frequency. This would allow the frequency spectrum of the generated waveform to be measured on a spectrum analyzer.

3. Measurement of amplitude modulation with an oscilloscope when the F-M generator is operating.

4. Determination of what constitutes a good frequency modulator and comparison of the MOSFET frequency modulator to other types of frequency modulators.

REFERENCES

1. Crawford, Robert H., MOSFET in Circuit Design, McGraw-Hill Book Co., Inc., 1967, pp. 3-8, pp. 21-40.

2. Deal, B. E., Sah, C. T. and Snow, E. H., "Simple Physical Model for the Space-Charge Capacitance of MOS Structures," Journal of Applied Physics, Vol. 35, No. 8, August 1964, pp. 2458-2460.

3. Fitchen, F. C., Transistor Circuit Analysis and Design, D. Van Nostrand Co., Inc., 1966, p. 332.

4. Grove, A. S. and Lamond, P., "Stable MOS Transistors," Electro-Technology, Vol. 76, No. 6, December 1965, pp. 40-43.

5. Heising, Raymond A., "Modulation Methods," Proc. IRE, Vol. 50, 1962, pp. 896-901.

6. Kane, James F., Understanding Field-Effect Transistors, Motorola Technical Information Note AN 206, 1965.

7. Kiver, Milton S., F-M Simplified, D. Van Nostrand Co., Inc., 1947, pp. 3-57.

8. Miller, John R. and Walston, Joseph A., ed., Transistor Circuit Design, McGraw-Hill Book Co., Inc., 1963, pp. 316-318.

9. Moore, Gordon E., The MOS Transistor as an Individual Device and in Integrated Arrays, Fairchild Semiconductor Application Note, 1965, pp. 44-45.

10. Ramey, Robert L., Physical Electronics, Wadsworth Publishing Co., 1961, pp. 141-151.

11. Ryder, John D., Electronic Fundamentals and Applications, Prentice-Hall, Inc., 1964, pp. 443-473.

12. Seely, J. Leeland, Designing with MOS Semiconductors, General Instrument Corporation, Microelectronic Technical Bulletin, 1966.

13. Sproull, Robert L., Modern Physics, John Wiley and Sons, Inc., 1956, p. 75, pp. 305-306.

14. Tepper, M., Basic Radio Vol. 6, John F. Rider Publisher, Inc., 1961, pp. 86-88.

15. Van Der Ziel, Aldert, Electronics, Allyn and Bacon, Inc., 1966, pp. 430-440.

16. Wallmark, J. T. and Johnson, H., ed., "Semiconductor Surface Physics," (Karl H. Zaininger), Field-Effect Transistors, Prentice-Hall, Inc., 1966, pp. 17-45.