THESIS

Ferize GÖZÜM

FOR REFERENCE

NOT TO BE TAKEN FROM THIS ROOM

FIELD EFFECT TRANSISTORS

THEORY AND APPLICATIONS

PREFACE

This thesis was prepared in accordance with the requirements of the Engineering Graduate School program of Robert College. The subject of the thesis is the Field-effect Transistor . This is a new electronic device which unites in itself the advantages of conventional transistors and valves.

This thesis is prepared in the form of a text-book which combines in its body the knowledge necessary to understand and use this device. It consists of twelve chapters .The first three chapters give the theory , static and dynamic characteristics and equivalent circuits of the junction gate field effect transistors. The succeeding six chapters deal with the insulated gate field effect transistor. Chapter nine discusses noise in field effect transistors. The following three chapters are devoted to the various applications of the device.

I would like to express my deep gratitude to Prof. Dr. Mustafa Santur who proposed the subject of the thesis,provided me with valuable suggestions and very useful reference books and also enabled me to use the library of the Istanbul Technical University. I am grateful to Mr. Fonger who provided me with the necessary books which he ordered specially for my thesis. I am very much indebted to Mrs. Hofland who prepared the photocopies for me.

Ferize Gözüm

May 15, 1967

Bogazici University Library
39001100540494
14
124123

INTRODUCTION

1. History

Altough it has been possible for more than thirty years to make field effect devices in the laboratory, it is only recently that the developments in the semiconductor technology have made possible the manifacture of reliable units with useful characteristics.

In its most fundamental form the field effect transistor is a decendant of the electrical capacitor. In the case of an ordinary capacitor, the modulation of the charge on the metal plate of the capacitor results in a small change in the sheet conductivity and the charge moved before breakdown of the dielectric is only about 10^{13} atoms/cm^2 at a field of 10^7 volt/cm. For a useful electronic device, the charge moved must be a large fraction of the total available charge and the mobility of the moving charge must be high. J.E.Lilienfeld and Oscar Heil were the first to observe these two characteristics in semiconductor materials which are not obtained with metallic materials.

In 1933, J.E.Lilienfeld took out a patent for what can be regarded as the progenitor of the insulated gate field effect transistor in that the conductivity between two electrodes was modulated by a potential applied to a third electrode close to yet insulated from a semiconductor layer. In 1935, Oscar Heil obtained a British patent for a device which was nothing else but a unipolar field effect transistor with insulated gate(Fig.1)

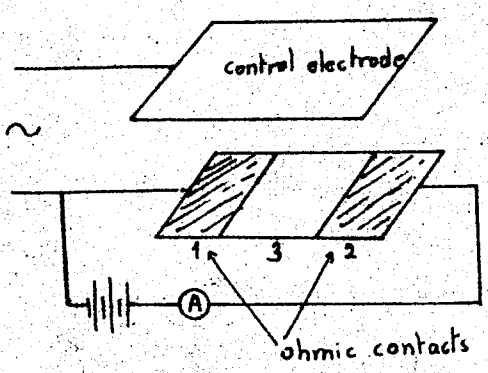


Fig 1. Drawing from British patent 439457 inventor O.Heil

In this device the resistance of a semiconducting layer made of tellurium iodine, cuprous oxide, or vanadium pentoxide, could be varied by the application of a varying voltage to an adjacent control electrode. The control electrode, although close to the semiconducting layer, was electrically insulated from it. Unfortunately, due to the limited knowledge of the physics of surfaces and thin films at that time, these devices could not be perfected.

In 1948, Shockley and Pearson reexamined the modulation of surface charge in semiconductor films. They placed a thin insulating strip between the semiconductor and evaporated metal film forming a parallel plate capacitor. They measured the change in the conductance of the semiconductor as a function of the voltage applied across the capacitor. They found that only 10% of the excess charge placed in the germanium was mobile. 90% of the charges resided on the surface in states which rendered them immobile for purposes of conduction. This was found to be due to surface conductions such as adsorbed gases, surface strains, mechanical damage, etc.

Finally in November 1952, Shockley revealed theory and details of a practical unipolar field effect transistor which by-passed the problem of surface states. A reverse p-n junction was used to modulate the mobile charge in the lightly doped n-type portion of the junction. The electric field being immersed in the semiconductor the surface problem played no part in the device operation. However the use of the p-n junction allowed only depletion of the charge in the n-type channel and no enhancement type of operation was possible.

In 1960, Kahng and Atalla proposed a silicon structure using an insulated gate to induce conduction between two back biased surface diodes. In 1961, P.K. Weimer described the thin film insulated gate transistor(TFT). Finally in 1962, S.R.Hofstein and F.P.Heiman described the MOSFET, another insulated gate field effect transistor. This device allowed enhancement as well as depletion of the charge in the channel and also made possible operation with zero bias. The technique developed by this time

for producing clean, passivated silicon surfaces and control of fine dimensions made the MOSFET possible. We have, in the field effect transistor, a device of outstanding importance, and it is not therefore suprising that in the last few years increasing emphasis has been placed on research into the physics and application of such devices.

2. Evaluation of Field Effect Transistors

The field effect transistors now being sold are competing directly with conventional transistors. They are becoming popular because they combine the advantages of both the conventional transistors and valves. They permit operation at high-input impedances with very low noise levels. They are less sensitive to radiation than the conventional transistors and are especially useful in integrated circuits because they can be operated in directly coupled arrays, thus eliminating external capacitive coupling . However, the highest frequency at which a commercial FET can operate as an amplifier is 300 megacycles per second (KMC type K1001), wheras commercial conventional junction transistors as amplifiers go to 1000 megacycles per second with useful gain.

The highest continous current rating specified for a FET is 10 amperes at 25°C (Sesco type 15PI) . Conventional silicon power transistors recently became available with dissipation ratings of 300 watts and collector current ratings of 150 amperes at 100°C. However a new type of field effect transistor, the multichannel FET is under development which will enable power field-effect transistors to be designed. The multichannel field effect transistor functions similarly to the single channel devices. The use of several channels allows the handling of higher currents. These devices also avoid the reduction in frequency response that normally occur with increasing the size of a single-channel device to handle higher currents. Germanium devices with 50 ampere ratings are under development. Sesco also

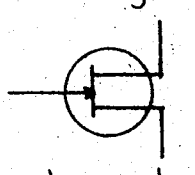
is developing silicon power field-effect transistors that are expected to have 500 volt and 50 ampere ratings. The high power devices are called tecnetrons.

Tetrode and pentode field effect transistors are also available. Texas Instrument, TIXSII , the siliconix 3N89 , RCA 3N98 are examples. These transistors do not have their gates tied to each other and may be operated as tetrodes if separate bias supplies are provided to each gate. However, the fourth lead of the tetrodes is usually connected to ground for increased stability at high frequencies. By eliminating the bulk gate - source connection and by supplying independent bias to the bulk gate, the device can also be used as a pentode.

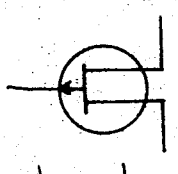
A twin-triode has been brought out by General Micro-Electronics. It consists of two p-channel MOS field-effect transistors diffused into a monolithic chip.

Another important field effect device besides the junction gate (FET) and insulated gate (MOSFET) type is the thin film field effect transistor (TFT) . The TFT is actually an insulated gate field effect transistor but is manifactured differently than the ordinary MOSFETS. They are developed especially for integrated circuits.

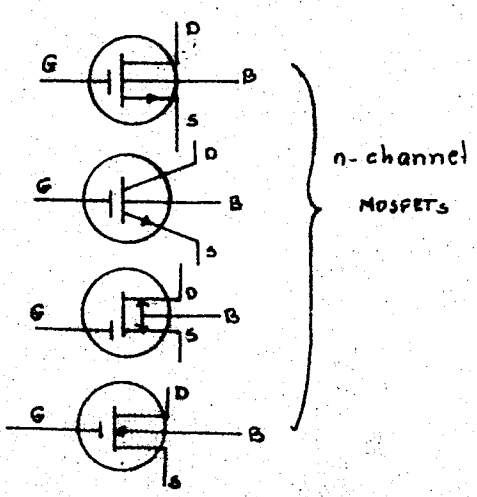
Symbols used for Field-effect transistors:



n-channel junction gate transistor



p-channel junction gate transistor



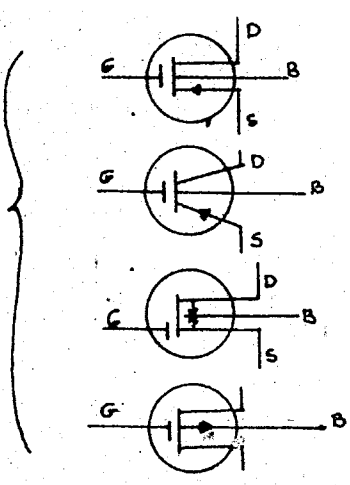


TABLE OF CONTENTS

CHAPTER I

THE REVERSE-BIASED JUNCTION FET

The field effect transistor as conceived by Shockley is essentially a semiconductor current path to either end of which an ohmic connection is made. The conductance of the current in the channel is controlled by applying an electric field perpendicular to the current. The electric field results from reverse biasing a p - n junction which is more or less parallel to the direction of the current flow. Fig.I.I shows schematically the construction of such a device with p-type channel. This is a three terminal device, consists of a p-type material sandwitched between two layers of strongly doped n-type (n_+) material and having ohmic contacts at two ends (the two types of semiconductors may be reversed). Conduction will be by means of majority charge carriers (holes in a p-channel, electrons in an n-channel). The terminal from which they flow (positive for a p-channel, negative for an n-channel) is called the source while the other ohmic connection to which they pass is called the drain. The third terminal used to apply the reverse bias to the p-n junction is called the gate. The gate is the control electrode.

I.I Behavior of Reverse Biased Junctions.

When a p-n junction is formed, mobile current carriers near either side of the junction diffuse across the junction and recombine with carriers of the opposite type leaving equal and opposite electric charges on either side of the junction. Thus the n-type material becomes positively charged and the p-type material becomes negatively charged. This in turn sets up an electrostatic difference ϕ across the junction which tends to retard any further diffusion of free electrons and holes across the junction. This potential is called the barrier potential or contact potential. The region on either side of the junction from which mobile carriers have disappeared is called the space charge or depletion

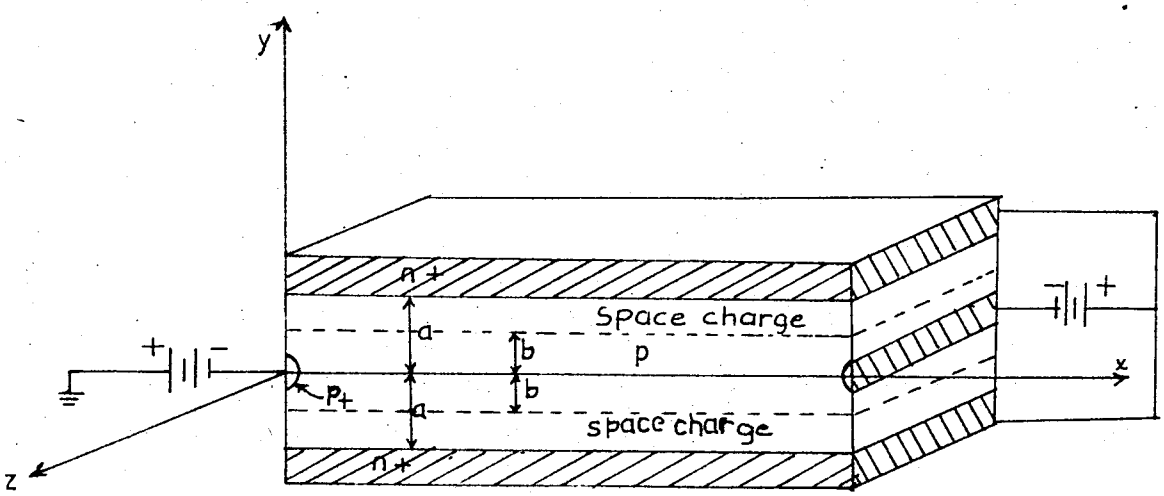


Fig 1.1 Unipolar field effect device with p-channel (uniform channel model)

layer. When a reverse bias, V , is applied across the p-n junction, the potential barrier at the junction will increase to the value $\phi + V$ because the polarity of the external voltage V is in the same sense as the polarity of the contact potential. Thus the depletion layer length will increase. If the impurity density on one side of the junction is made very large compared to the impurity density on the other side, then the depletion layer width on the more dense side can be considered negligible.

1-2 . Conductance of a Semiconductor Channel

Let us consider a semiconductor bar with the following dimensions, impurity and conductivity:

L = length between the ohmic contacts
Z = width of the bar
2a = heigth of the bar
q = electronic charge = 1.6019×10^{-19} coul.
N = impurity density (Na for acceptor, Nd for donor impurity)
μ = carrier drift mobility

We know that a pure semiconductor such as Si or Ge is a poor conductor because it contains few mobile carriers (holes and free electron). The number of free electrons per centimeter cube in intrinsic Si at room temperature is about 1.6×10^{10} . The density of holes is also the same. Certain materials when added to Si or Ge exert a profound influence on their electrical conductivity . By carefully regulating the amount of certain desirable impurities and by utilizing the electrical effects produced by their presence , we can control the conductivity of semiconductors. When such semiconductors are doped with donor or acceptor impurity atoms, it is assumed that above 200° K there will be one free current carrier per impurity atom added. For the Si devices being made today, the impurity concentration of the selected impurity varies from a few tenths of an atom per billion Si atoms (apb) up to about 10000 apb . An impurity level of 1 apb produces a carrier density (holes or electrons per cubic centimeter) of 5×10^{13} . Since electron and hole densities for intrinsic Si at room temperature are 1.6×10^{10} , we are justified in ignoring the intrinsic contribution to the conduc-

tivity of Si containing impurities in the range mentioned above .

Since in an intrinsic semiconductor $p_i = n_i$, and since the impurity atom density is large compared to the density of thermally generated intrinsic electron-hole pairs, the current carriers available can be considered to be of only one type ,e.g., if the impurity atoms are all acceptors, only holes are available to carry current.

The conductance, Go , of the semiconductor bar is proportional to the total number of carriers present, so that if we assume the impurity concentration to be uniform, we can write

$$Go = \frac{2a\, q\mu Z}{L} N \qquad (1\text{-}1)$$

If the impurity is introduced by some diffusion process,then the concentration will not be uniform but will be a function of one or more of the dimensions of the semiconductor bar. If the impurities are diffused uniformly over one of the faces, then the concentration will be a function of the dimension perpendicular to this face. The conductance of such a diffused bar will be given by

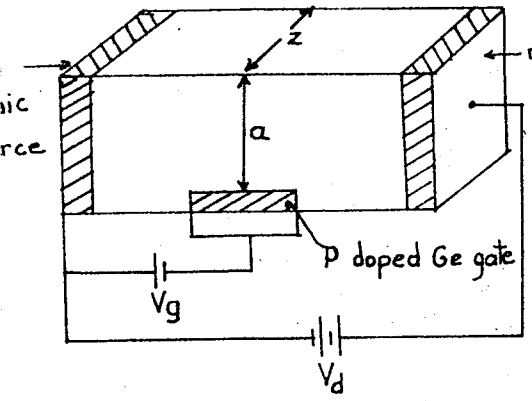
$$Go = q\mu \frac{Z}{L} \int_o^{2a} N(y)dy. \qquad (1\text{-}2)$$

where y is the direction perpendicular to the uniform distribution plane.

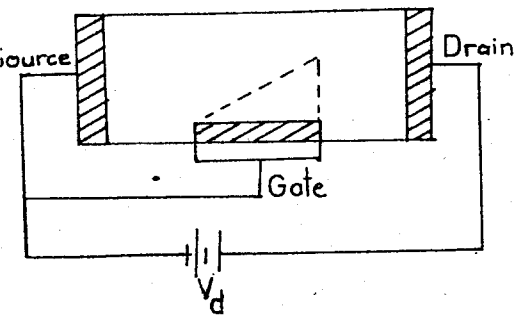
Now if this bar is made of a p-type (contains acceptor impurity) semiconductor and we form p-n junctions on both sides of this bar either by alloying, diffusion or epitaxial growth, we shall get the field effect transistor of Fig.1-1. When we apply reverse bias to the junctions as seen in Fig.1-1., depletion layers will form on both sides of the channel, and therefore the heigth of the conducting channel will no longer be 2a but will be of some smaller value, and its conductance, G , will vary.

1-3. Theory of Modulation of a Conducting Channel by Electric Fields.

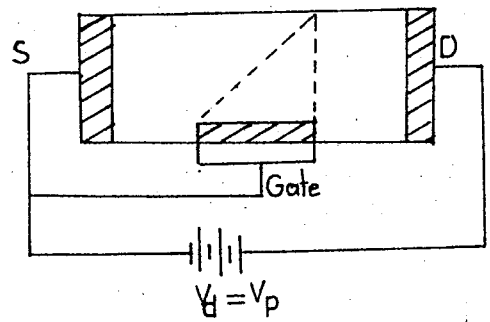
Fig.1-2 shows schematically the modulation of FET channels by applied voltages.



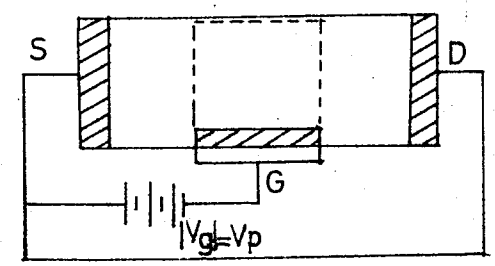
a) Single gate Ge FET with grounded source and conventional polarity



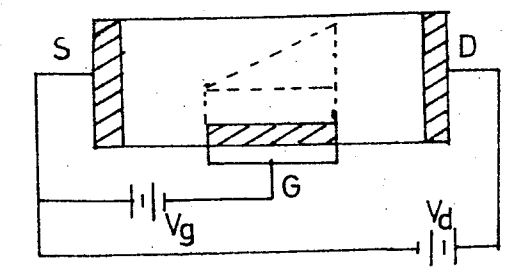
b) Depletion layer with drain voltage



c) Depletion layer with drain voltage equal to Pinch off voltage V_p .



d) Depletion layer with gate voltage



e) Depletion layer with drain and gate voltage.

$V_d + |-V_g| = V_p$

Fig.1-2 Modulation of FET channel by applied voltages . (n-channel)

The theory of the FET as originally formulated by Shockley can be considered as a first order theory. It assumes an abrupt p-n junction, and a uniformly doped channel, and neglects high field mobility effects. But in actual devices, the gate-channel junctions can be either graded or abrupt depending on the method of fabrication. For the abrupt, uniformly doped channel structure, Shockley's theory may be directly applicable . Gate-junction fabrication by diffusion techniques results in impurity density profiles which may to a first order be approximated by one of the distributions shown in Fig. 1-3

R.S.C.Cobbold and F.N.Trofimenkoff have derived expressions for the drain current I_d , and conductance g(v)for field effect devices showing such impurity profiles as those shown in Fig. 1-3 . They have shown that although these expressions are somewhat different from the ones Shockley has found for abrupt junctions, when they are plotted, the curves do not fall off from each other very much. Therefore a theoretical discussion as given by Shockley is satisfactory for understanding the behavior of field effect transistors in general.

The theory is developed in two parts, the first dealing with a uniform width channel where the drain-source voltage(V_{ds}) is small compared with the gate-source voltage (V_{gs}) and the latter being the gradually narrowing channel where small currents flow.

1-3a. The Uniform Width Channel Approximation.

Consider Fig.1.1 again. Note that the two junctions forming the channel are reverse biased and space-charge regions are formed. The current carriers flow in a channel bounded by these regions. In Fig.1.1. it is assumed that the reverse bias at the p_+ terminals are equal so that the channel shown has a uniform width. If the reverse bias at the two terminals (drain and source were not equal, the channel would be narrower at the terminal with larger reverse bias as in Fig. 1-2b.

To discuss the dependence of conductance of the channel upon the magnitude of the reverse bias, let us first assume that

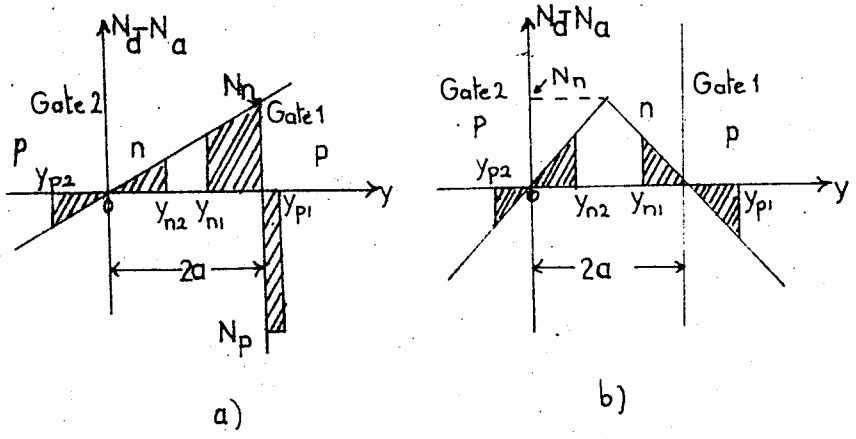


Fig.1-3. First order approximations to an n-channel double diffused FET.

a) An abrupt linear asymmetrical approximation

b) Symmetrical linearly graded approximation

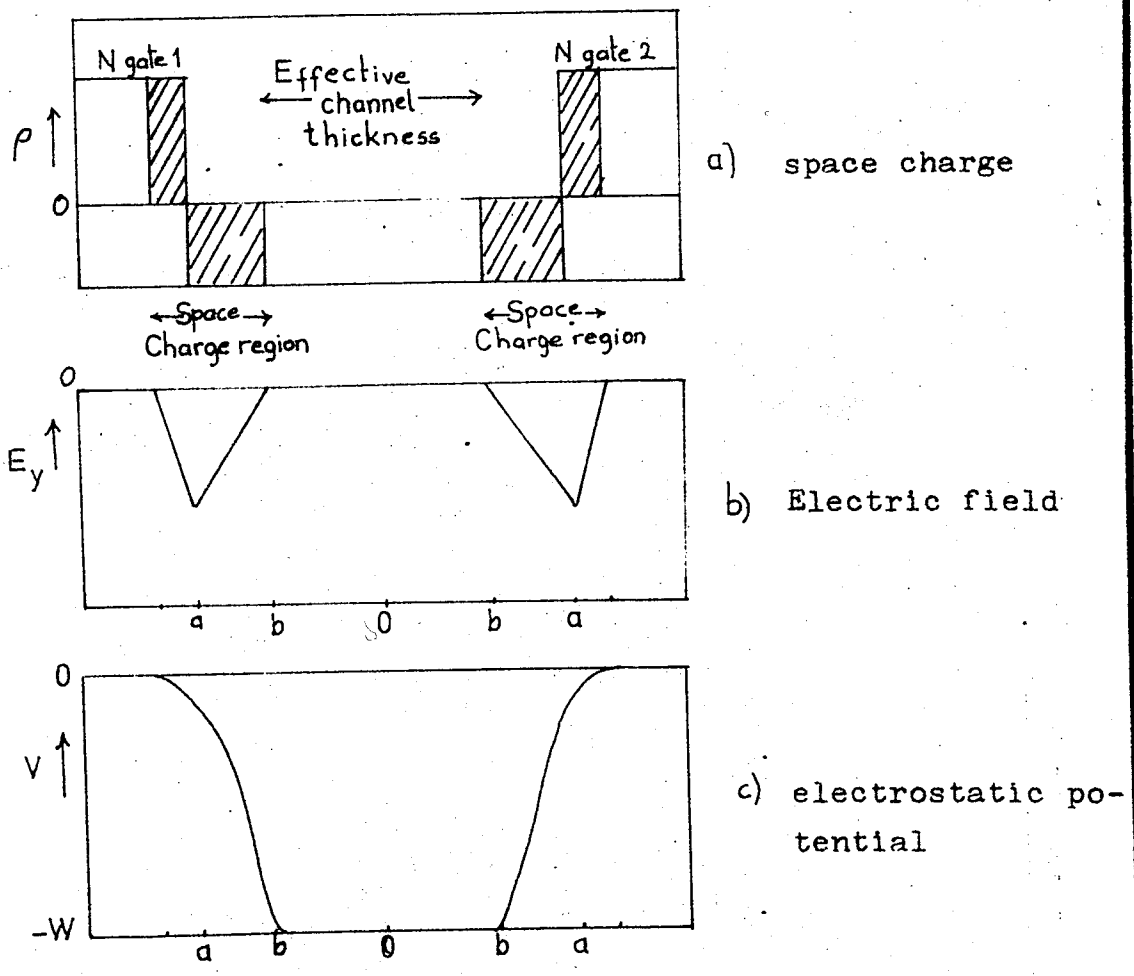


Fig.1-4. Space charge, electric field, and electrostatic potential in a reverse biased p-channel FET.

the difference in the reverse bias at the source and drain end is so small that the width of the channel is almost uniform. In the figure, the space charge region is represented as lying entirely in the p-region. Although this is approximation it is easy to show that it is a good one since n_+ regions are more heavily doped than the p-region.

In semiconductor junction devices the chemical charge density ρ may be defined in terms of the donor and acceptor concentrations such that

$$\rho_o = qN_d \qquad (1\text{-}3)$$

in the donor region and

$$\rho_n = qN_a \qquad (1\text{-}4)$$

in the acceptor region. Since n-region is more heavily doped, $\rho_n > \rho_o$

Let $y=0$ be the middle plane of the p-region
$y=b$ the edge of the space charge layer
$y=a$ be the p-n junction.

The space charge region is a dipole layer where the charges add to zero. Therefore the two areas shown in Fig.1-4a must be equal. Because if it were not so, there would be an electric field beyond the boundaries of the space charge layer in accordance with Gauss' theorem. The same result can also be obtained mathematically from Poisson's equation. Referring to Fig. 1-4b let V(y) represent the electrostatic potential, E(y) the electric field, y the distance from the center of the p-layer. Then from Poisson's equation (assuming E_x is negligible),

$$k\epsilon_o \frac{d^2V}{dy^2} = -k\epsilon_o \frac{dE_y}{dy} = -\rho(y) \qquad (1\text{-}5)$$

It is considered that

$$E(y)=0 \quad \text{for} \quad y \leq b \qquad (1\text{-}6)$$

and the small transition region where $\rho(y)$ changes from zero to $-\rho_o$ is neglected, so that

$$\frac{dE_y}{dy} = \frac{\rho(y)}{k\epsilon_o} = -\frac{\rho_o}{k\epsilon_o} \quad \text{for} \quad a>y>b \qquad (1\text{-}7)$$

Intergrating Eq.(1-7) we get

$$E_y = -\frac{\rho_o}{k\epsilon_o}(y-b) \qquad (1\text{-}8)$$

The field Ey will vanish when

$(a-b)\rho_o$ = thickness of space charge layer in n-region $x\rho_n$ (1-9)

For Eq. (1-9) to be true, the two areas (1 and 2) in Fig.1-4a

must be equal.

Almost all the voltage drop between gate and source occurs across the depletion region so that

$$\int_{y=0}^{y=a} dV = -\int_{b}^{a} E_y\,dy$$

$$= \int \frac{\rho_o}{k\epsilon_o}(y-b)\,dy \qquad (1-10)$$

$$= \frac{\rho_o}{2K}(y-b)^2 + C$$

where $K = k\epsilon_o$ dielectric constant

for Ge $K = 1.42 \times 10^{-12}$ farad/cm

for Si $K = 1.06 \times 10^{-12}$ farad/cm

Using boundury conditions

When $y=a$ $V=0$ we find that $C = -\frac{\rho_o}{2K}(a-b)^2$, so that

$$V = \frac{\rho_o}{2K}\left[(y-b)^2 - (a-b)^2\right] \qquad (1-11)$$

The potential in the channel is found when y is equated to b

$$V(b) = -(\rho_o/2K)(a-b)^2 \qquad (1-12)$$

Let $W = -V(b)$ = the magnitude of reverse bias

The space charge penetrates the entire p-region when the reverse bias is equal to Wo . Then $b = 0$, so that from Eq.(1-12)

$$W_o = \rho_o a^2 / 2K \qquad (1-13)$$

and

$$W = \left[1 - (b/a)\right]^2 W_o \qquad (1-14)$$

Assuming that the reverse biases at the source and drain of Fig.1-1 are W and $W+\Delta W$ respectively where ΔW is small compared to Wo - W (necessary condition in order to keep the channel width substantially uniform). There will be a current I_d flowing from source to drain. The conductivity of the channel is given by

$$\sigma_o = \mu_o \rho_o \qquad (1-15)$$

where μ_o = mobility of holes

σ_o = semiconductor conductivity

If the electric field in the conducting channel due to small V_{ds} $(= V_d - V_s)$ is E_x , the current density per unit length in the Z direction is $\sigma_o E_x$, and

$$I_d = 2b\sigma_o E_x = 2b\mu_o \rho_o E_x$$

$$= g(W)E_x \qquad (1-16)$$

Here g(W) is the conductance per unit square of the layer 2b thick. Conductance per cm in the Z direction of a channel of length L in the x - direction will be

$$G = g(W)/L \quad \text{for uniform channel with } \Delta W \ll Wo - W \qquad (1-17)$$

Since ΔW is not always small, the channel width will not be uniform always. A more general relation for g(W) can be derived making use of Eq.(1-14)

$$g(W) = 2\sigma_o b(W) = \left[1-(W/Wo)^{1/2}\right] g_o \qquad (1-18)$$

where
$$g_o = 2\sigma_o a \qquad (1-19)$$

Making use of Eq.(1-16) and (1-18), the current I_d flowing in the channel can be calculated when there is a large voltage between the source and drain.

Before going into further discussion some terms must be made clear:

V_s, V_g, V_d = dc votages of the source, gate and drain electrodes respectively

V_{ds} = source to drain voltage = $V_s - V_d$

$V_{dg} = W_d = V_g - V_d$ = gate to drain bias

$V_{gs} = W_s = V_g - V_s$ = gate to source bias

I_s, I_g, I_d = dc currents flowing into source, gate, drain electrode respectively

$v_s, v_g, v_d, i_s, i_g, i_d$ = a.c vottage and current values.

If the channel is of length L, then the current is just

$$I_d = g(W) V_{ds} / L \qquad (1-20)$$

but from Eq(1,14)

$$b = a\left[1 - (V_{gs}/Wo)^{1/2}\right] \qquad (1-21)$$

therefore

$$g(W) = 2a\mu_o\rho_o Z \left[1 - (V_{gs}/Wo)^{1/2}\right] \qquad (1-22)$$

and hence

$$I_d = \frac{2a\mu_o\rho_o Z}{L} \left[1 - (V_{gs}/W_o)^{1/2}\right] \qquad (1-23)$$

Eq.(1-23) suggests that for small values of V_{ds} ($V_{ds} \ll V_{gs}$), the device behaves like an ohmic resistor, the value of which is determined by V_{gs}. Note that this equation is valid only for $|V_{gs}| < W_o$ and $b > 0$.

1-3b. Gradually Narrowing Channel.

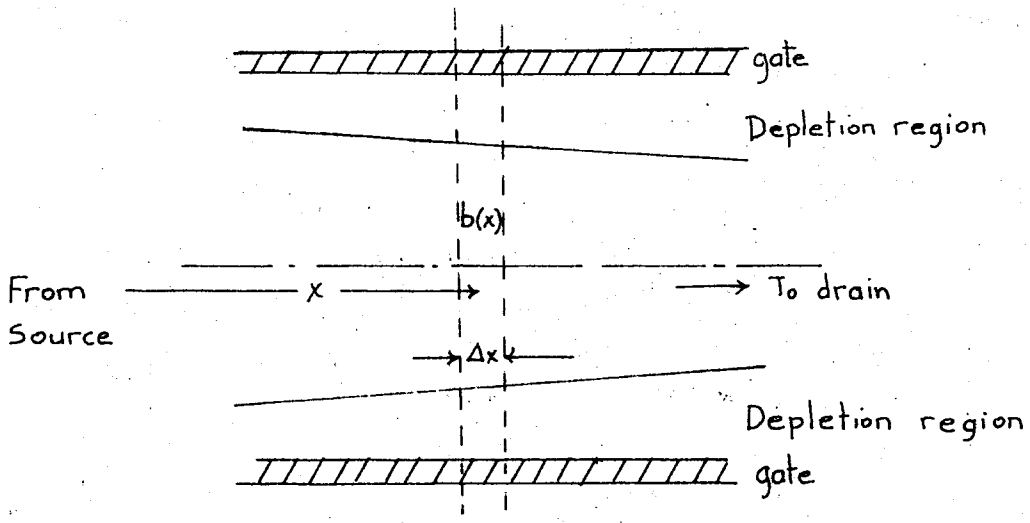


Fig. 1-5a Section of the gradually narrowing p-channel FET

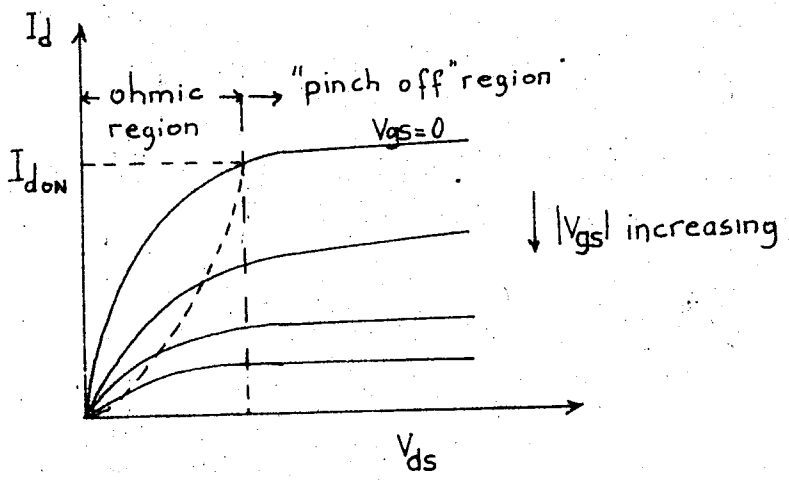


Fig. 1-5b. Drain characteristics

Now consider an FET with a larger field E_x along the x-axis (Fig. 1-5a). Under such conditions, a current I_d flows along the channel from source to drain between the gate electrodes. Because of the IR drop along the channel, the reverse bias on the gate-to-channel junction will not be uniform along the whole length of the channel, the greater the distance from the source, the stronger the reverse bias on the junction will be, this causes the depletion regions to assume a wedge shape. However, provided that this convergence is fairly gentle and that the depletion region never extends right through the channel, uniform channel analysis can be applied to a thin slice of the FET cut perpendicular to the x-axis and of thickness Δx . If the potential at the center of this element is V_x relative to the source, which is taken at a distance x from the element, then the voltage V_x replaces V_{ds} , and Δx replaces L in equation (1-23).

If we treat the gate junction as an equipotential, we can replace V_{gs} of eq(1-23) by $(V_{gs} + V_x)$. Then

$$I_d \Delta x = 2a\mu_o\rho_o Z \left\{ 1 - \left(\frac{V_{gs}+V_x}{W_o}\right)^{1/2} \right\} \Delta V_x$$

$$I_d = \frac{2a\mu_o\rho_o Z}{L} \left\{ V_{dg}\left[1 - \frac{2}{3}\left(\frac{V_{dg}}{W_o}\right)^{1/2}\right] - V_{gs}\left[1 - \frac{2}{3}\left(\frac{V_{gs}}{W_o}\right)^{1/2}\right] \right\} \qquad (1\text{-}24)$$

$$I_d = G_o \left\{ V_{dg}\left[1 - \frac{2}{3}\left(V_{dg}/W_o\right)^{1/2}\right] - V_{gs}\left[1 - \frac{2}{3}\left(V_{gs}/W_o\right)^{1/2}\right] \right\}$$

where

$$G_o = \frac{1}{R_o} = 2a\mu_o\rho_o Z / L \qquad (1\text{-}25)$$

The transconductance of the device is :

$$g = \partial I_d / \partial V_g \Big|_{V_d} = G_o \left[\left(V_{gd}/W_o\right)^{1/2} - \left(V_{gs}/W_o\right)^{1/2} \right] \qquad (1\text{-}26)$$

and the saturated transconductance is given by :

$$g_m = \partial I_{ds} / \partial V_g \Big|_{V_d} = G_o \left[1 - \left(V_{gs}/W_o\right)^{1/2} \right] \qquad (1\text{-}27)$$

and the output conductance g_d of the device is

$$g_d = -\partial I / \partial V_d \Big|_{V_g} = G_o \left[1 - \left(V_{gd}/W_o\right)^{1/2} \right] \qquad (1\text{-}28)$$

It will be noted that Eq.(1-24) is only valid for values of V_{dg} up to Wo at which point the current is "pinched off", that is the depletion region extends right through the semiconductor channel, and the gradual approximation used in deriving

the current expression (Eq. 1-24) is no more valid. According to the first order theory just developed, the conductance of the channel and therefore the current I_d would drop to zero when pinch off occurs. However, this does not occur. The concentration of the charge carriers at the center of the channel has a screening effect and the drain current continues to flow at a substantially constant value, called the saturation current, as V_{dg} increases beyond Wo. Thus Eq.(1-24) can be used to plot the family of I_d/V_{ds} curves with V_{gs} as parameter up to $V_{ds}\,(=V_{dg}-V_{gs}) = Wo-V_{gs}$, and at larger values of V_{ds}, the curves are approximated by

$$I_{ds} = G_o \left\{ Wo/3 - V_{gs} \left[1 - \tfrac{2}{3} (V_{gs}/Wo)^{1/2}\right] \right\} \qquad (1\text{-}29)$$

where I_{ds} = saturated drain current. The value of the saturated drain current when $V_{gs} = 0$ is usually designated by I_{don} or simply I_{do}, and is given by

$$I_{do} = \frac{Wo}{3Ro} \qquad (1\text{-}30)$$

The saturated transconductance occurs when $V_{gs} = 0$ so that

$$g_{max} = \frac{1}{Ro} \qquad (1\text{-}31)$$

Therefore

$$I_{do} = \frac{Wo}{3}\, g_{max} \qquad (1\text{-}32)$$

Thus for a given FET device Wo can be experimentally calculated from Eq.(1-32).

The frequency response of the device can be estimated by the following simple argument. In order to change the gate voltage, the capacity of its p-n junction must be charged trough the resistance of the channel. This process has an associated time constant which limits the frequency response.

If a wedge-shaped channel, completely pinched off at the drain end and comletely open at the source end ($V_g = V_s = 0$) is assumed, the capacity for unit length in the Z direction can be approximated by

$$C = 4KL/a \qquad (1\text{-}33)$$

The factor 4 arises because the average width of the space-charge region is approximately a/2 and because there are two such regions, one on either side. This capacity on the average charges

through half the resistance of the channel. Thus a limiting frequency f exists given by

$$f = \frac{1}{2\pi RC} = \frac{1}{2\pi}\left(a^2\sigma_o / 2L^2K\right) \qquad (1\text{-}34)$$

1-4 Discussion of the Theory.

As was mentioned before, Trofimenkoff has derived current and conductance expressions for three different distributions : linear, abrupt and abrupt linear. Derivations are given in appendix A. Fig. 1-6a shows the variation of the normalized I_d/I_{do} as a function of V_{gs}/Wo, and Fig. 1-6b shows the variation of the normalized transconductance g_m/Go with V_{gs}/Wo for these three distributions. As might be expected, the variation for the abrupt linear distribution lies in between those for the linear and abrupt symmetrical distributions. It is clear therefore that the form of the channel doping profile does not have very great effect on the dc characteristics of the device. If the ratio $I_d/g_m \,/\, I_{do}/G_o$ is plotted as shown in Fig. 1-7 it will be observed that over quite an appropriate range the ratio I_d/g_m varies almost linearly with V_{gs}/Wo . That is

$$\left(I_d/g_m\right)\Big/\left(I_{do}/G_o\right) = K\left(1 - V_{gs}/W_o\right) \qquad (1\text{-}35)$$

where K is a constant. So that

$$I_d/g_m = \frac{1}{n}\left(W_o - V_{gs}\right) \qquad (1\text{-}36)$$

where

$$n = G_o W_o / I_{do} K \;,\; W_o = n\, I_{do} K / G_o \;,\; I_{do} = 4qZ\mu W_o^2 / n\epsilon L \qquad (1\text{-}37)$$

From various experimental results and from discussions by Middlebrook, it is found out that for the majority of the transistors $n \cong 2.0$.

Thus a simpified expression for the drain current characteristics can be derived by integrating Eq. (1-36).. This yields

$$I_d = I_{do}\left(1 - \frac{V_{gs}}{W_o}\right)^n \qquad (1\text{-}38)$$

This expression is found to be well obeyed by most of the FETS and is very useful for design work. The values of n can be found from a graph such as Fig. 1-8, and the value of I_{do} can be measured directly.

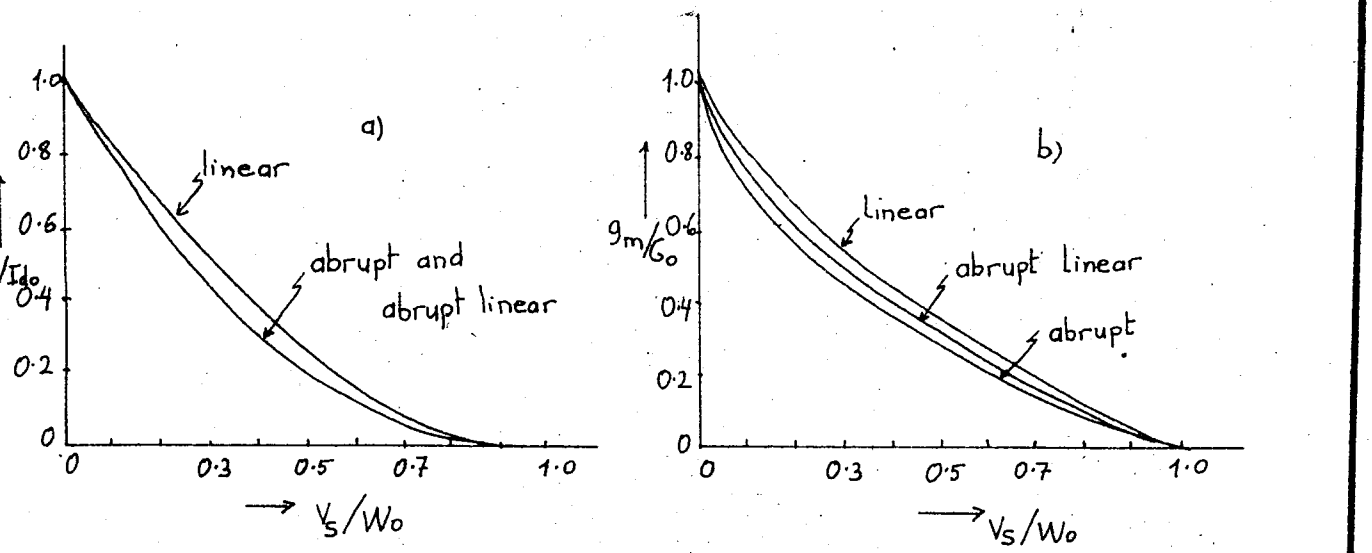


Fig. 1-6. Theoretical curves for three impurity-density profiles
a) Variation of normalised drain current as a function of V_s/W_o
b) Variation of normalized transconductance as a function of V_s/W_o

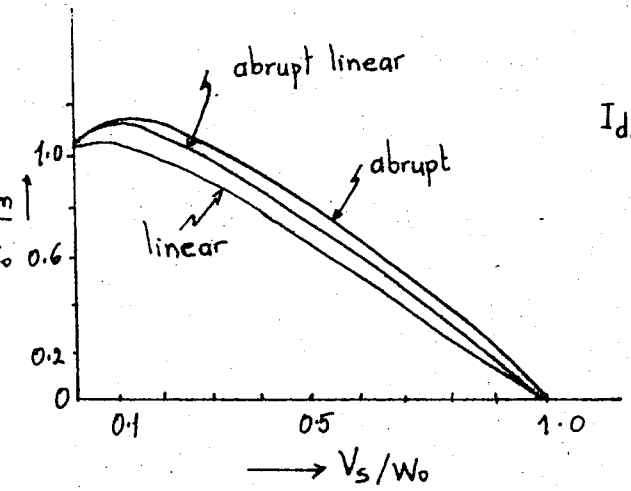


Fig.1-7. Normalized I_d/g_m plot for three impurity-density profile approximations. The linearity for V_s/W_o near to unity should be noted.

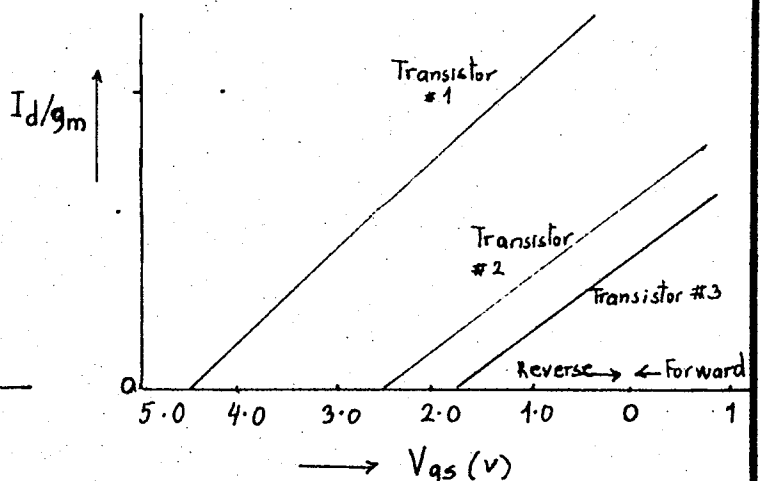


Fig.1-8. Experimental I_d/g_m plot for three FETs. The intercepts and the slopes give W_o, and n respectively.

CHAPTER 2

MODIFICATION OF IDEAL THEORY and DEVICE CHARACTERISTICS

Various experiments carried out by specially prepared units and actual FET devices on the market have shown that although the general features of the d-c characteristics of some units are adequately explained by the first order of Shockley, a modified theory is necessary for some others. Dacey and Ross and others have modified the original theory to include the effects of variable mobility. They have also considered other effects such as series resistance, negative gate resistance and temperature effects. These effects have to be considered in understanding the divergence of the actual characteristics from the ideal and also important in manufacturing the devices and designing circuits with them.

2-1. Series Resistance

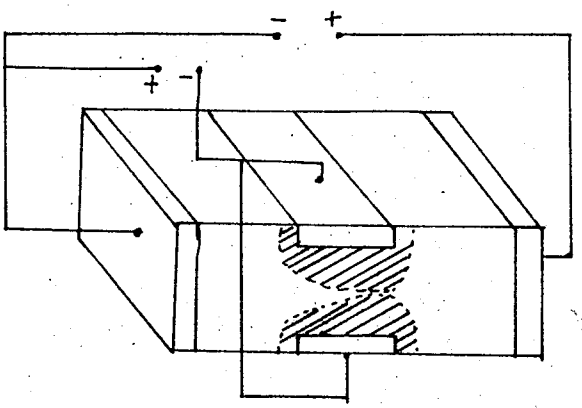
In Shockley's ideal model, the source and drain connect directly onto the channel between the gates. The actual units with which experiments have been carried on and the units later manifactured commercially look like Fig. 2-1a and 2-1b. In such units some semiconductor materil is left between the actual contact and the gate. Thus series resistances exist between the electrodes to which voltages are applied and the working part of the structure. These resistances can be taken into account in the already derived expressions of chapter 1 by simple circuit theory.

The following changes must be made in the theoretical expressions:

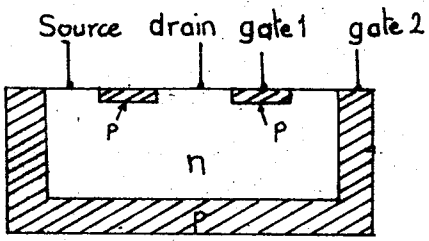
$$\begin{aligned} V_g &\longrightarrow V_g \\ V_s &\longrightarrow IR_s \\ V_d &\longrightarrow V_d - IR_d \end{aligned} \qquad (2\text{-}1)$$

where R_s and R_d are source and drain resistances. Once the transconductance is calculated according to the first order theory, than the apparent transconductance must be calculated according to the formula :

$$g'_m = \frac{g_m}{1 + R_s g_{max} + R_d g_d} \qquad (2\text{-}2)$$



(a)



(b)

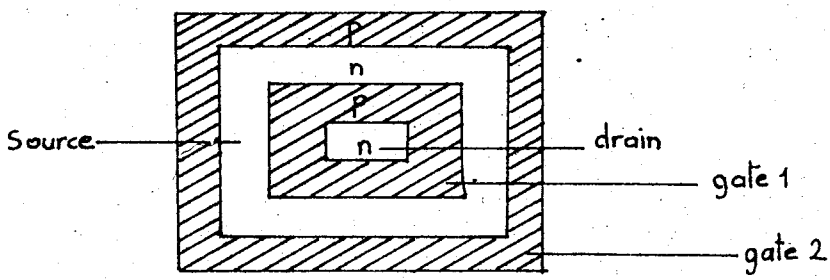


Fig. Experimental and commercial FET devices.
a) Shockley experimental unit with an n-channel
b) a commercial FET (n-channel)

Since at saturation $g_d=0$ (See Eq. 1-28), the apparent maximum transconductance becomes

$$g'_{max} = \frac{g_{max}}{1+R_s g_{max}} \qquad (2-3)$$

If $g_{max} R_s \gg 1$, then the transconductance is simply $1/R_s$, and to have a device with high conductance the source resistance must be made small. The main disadvantage of the drain resistance, R_d, is that it makes necessary a higher supply voltage and causes $I_d R_d$ heat which must be dissipated. The source resistance R_s also affects the frequency response of the device. If the gates are situated half way between the source and the drain, then $R_s = R_d$. To reduce the value of R_d, the gates are placed as close as possible to the drain.

A method for the determination of source, channel, and drain resistances of an FET device is given in appendix B.

2-2. High Field Mobility Effects

For devices with relatively short channels and/or high pinch off voltages the electric field, Ex, under pinch off conditions can be very high. For example a Si device which has a channel length of approximately 6×10^{-4} cm and a pinch off voltage of 2V has an average channel field of about 3300 v/cm. Under such high field conditions, the current density can no longer be assumed to be proportional to the electric field because the mobility has become field dependent. Many measurements have been made with Si and Ge semiconductors to show the field dependency. Fig. 2-2 shows the result of such measurements. These results show that the smallest field dependency occurs for p-type Si, while in n-type Ge, for a field of approximately 5×10^3 V/cm, the mobility becomes almost inversely proportional to the field.

Many laborious calculations have been carried out and various expressions have been obtained to show the effect of field dependent mobility on the dc characteristics of the FETS which have different impurity density profiles. The results shown in Fig. 2-3 for the normalized I_d/g_m ratio enable the effects of various degrees of field dependent mobility to be compared.

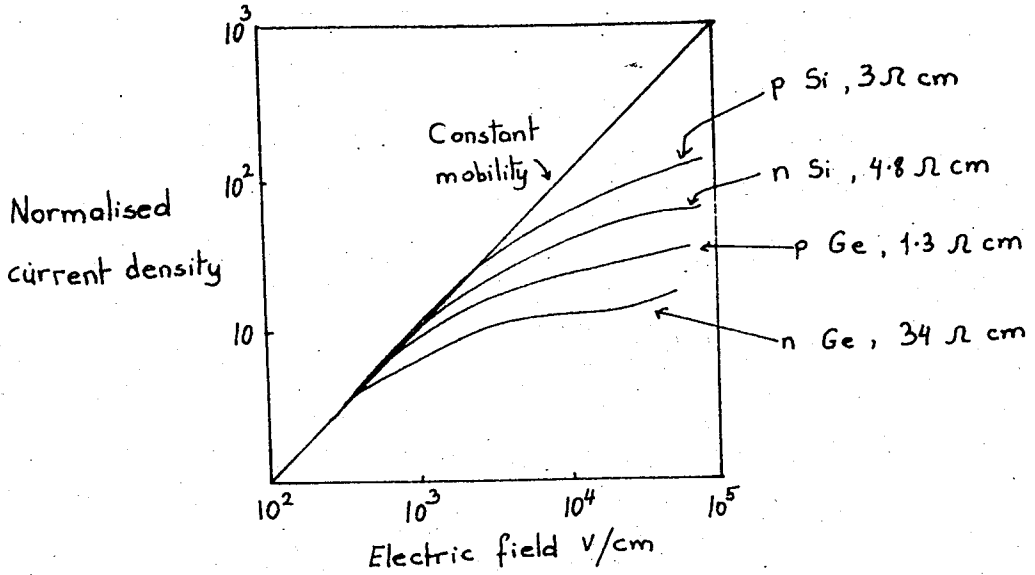


Fig 2-2. Variation of the current density at room temperature with electric field. For n-and p-type Si and Ge.

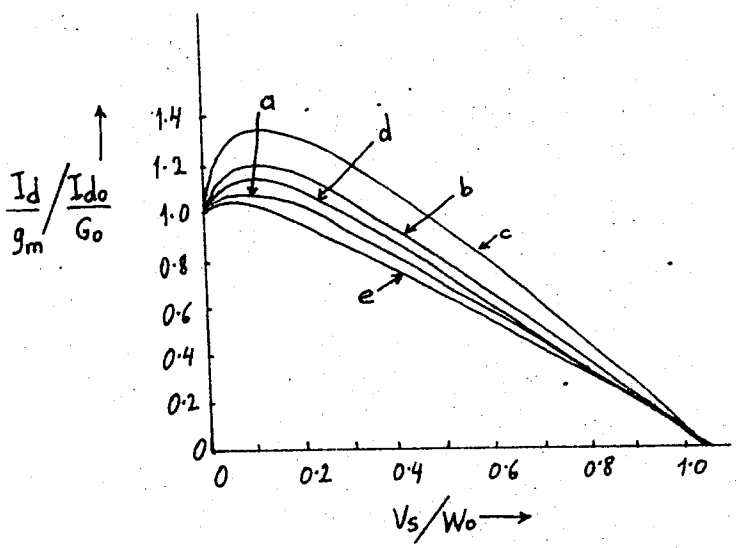


Fig. 2-3. Effects of various degrees of field dependent mobility on the normalized I_d/g_m plot.

a- μ constant (for abrupt linear)
b- $\mu \propto E^{-1/4}$ (for abrupt linear)
c- $\mu \propto E^{-1/2}$ (for abrupt linear)
d- $\mu \propto E^{-1/2}$ (for linear)
e- μ constant (for linear)

2-3 Temperature Effects

The field effect transistor has two independent temperature varying parameters. These are the channel conductivity and the equilibrium junction potential ϕ. To study these variations, we consider an FET biased in the pinch-off region with a constant applied gate-channel bias V_{gs}. V_{gs} is an external voltage and can be assumed held constant while dI_d/dT is investigated.
From appendix A, Eq.(A-4), $I_d = -\int_{V_s}^{V_d} g(V)dV$ and $g(V) = \frac{Z}{L}\int q\mu N(y)dy$,
So that we can write for I_d,

$$\frac{dI_d}{dT} = \frac{\partial I_d}{\partial \phi}\frac{d\phi}{dT} + \frac{\partial I_d}{\partial \sigma}\frac{d\sigma}{dT} \qquad (2\text{-}4)$$

Then

But $$\frac{\partial I_d}{\partial \phi} = g_m \quad \text{, so that} \qquad (2\text{-}6)$$

$$\frac{dI_d}{dT} = g_m\frac{d\phi}{dT} + \frac{I_d}{\sigma}\frac{d\sigma}{dT} \qquad (2\text{-}7)$$

Eq. (2-7) can be written in terms of the FET static quantities I_{do} and Wo by making use of the square law approximation of Eq. (1-38). Differentiating Eq.(1-38) with respect to V_{gs}, we get

$$g_m = \frac{dI_d}{dV_{gs}} = 2\frac{I_{do}}{W_o}\left(1 - \frac{V_{gs}}{W_o}\right) \qquad (2\text{-}8)$$

$$g_m/I_d = 2/W_o\sqrt{I_{do}/I_d} \qquad (2\text{-}9)$$

$$\frac{1}{I_d}\frac{dI_d}{dT} = -\frac{g_m}{I_d}\frac{d\phi}{dT} + \frac{1}{\sigma}\frac{d\sigma}{dT} \qquad (2\text{-}10)$$

$$= -2/W_o\sqrt{I_{do}/I_d}\; d\phi/dT + \frac{1}{\sigma}\frac{d\sigma}{dT}$$

Mobility versus temperature curves are given in several texts, the most familar being Shockley's "Electron and holes in Semiconductors". The mobility vs. temperature curves given for Si show that, over the temperature range of practical interest, a reasonable approximation to the mobility-temperature function is

$$\mu = \mu_o\left(\frac{T}{T}\right)^{-n} \qquad (2\text{-}11)$$

where n depends on the impurity concentration. Differentiating Eq.(2-11) we get

$$\frac{1}{\mu}\frac{d\mu}{dT} = -\frac{n}{T} \qquad (2-12)$$

Since $\sigma = \mu q N$

$$\frac{1}{\sigma}\frac{d\sigma}{dT} = -\frac{n}{T} \qquad (2-13)$$

Substituting Eq(2-13) into Eq.(2-10)

$$\frac{1}{I_d}\frac{dI_d}{dT} = -\frac{2}{W_o}\sqrt{\frac{I_{do}}{I_d}}\frac{d\phi}{dT} + \frac{n}{T} \qquad (2-14)$$

Equation (2-14) is the temperature coefficient of the drain current for any bias point. The zero of the temperature coefficient occurs when Eq.(2-14) equals zero, at some bias current I_d'. If the transconductance for zero temperature coefficient is designated by g_m', making use of Eq.(2-7), we can write,

$$\frac{I_d'}{g_m'} = -\frac{d\phi}{dT} \Big/ \left(\frac{1}{\sigma}\frac{d\sigma}{dT}\right) \qquad (2-15)$$

We can find an expression for I_d' as follows:
Replacing I_d by I_d' in Eq.(2-14), and equating it to zero, we get

$$I_d' = I_{do}\,\frac{4T^2}{n^2 W_o^2}\left(\frac{d\phi}{dT}\right)^2 \qquad (2-16)$$

Ordinarily one would expect this bias point to change with temperature, but if we assume that the free-carrier density is constant in the active channel and therefore the square-law approximation valid, we can express I_{do} in terms of mobility and Wo :

$$I_{do} = \frac{2qZ}{\epsilon L}\mu W_o^2 = \frac{2qZ}{\epsilon L}\mu_o\left(\frac{T}{T_o}\right)^{-n} W_o'^2\left[1 - \frac{1}{\phi}\frac{d\phi}{dT}(T - T_o)\right]^2 \qquad (2-17)$$

where μ_o, To, and Wo' are reference temperature values of these quantities.
If we call $\frac{2qZ}{\epsilon L}\mu_o W_o'^2 = I_{do}'$, then we can write Eq.(2-17) as

$$I_{do} = I_{do}'\left(\frac{T}{T_o}\right)^{-n}\left[1 - \frac{1}{\phi}\frac{d\phi}{dT}(T - T_o)\right]^2 \qquad (2-18)$$

Substituting Eq.(2-18) into Eq.(2-16), we find that the bias point for zero temperature coefficient of I_d is relatively independent of temperature:

$$I_d' = I_{do}\,\frac{4T_o^n T^{(2-n)}}{n^2 W_o'^2}\left(\frac{d\phi}{dT}\right)^2 \qquad (2-19)$$

But $n \cong 2$, so that the temperature T practically drops out of the

equation. Fig 2-4 shows a set of transfer curves for a particular FET device at three different temperatures. This figure shows that the crossover current I_d is practically independent of temperature.

In practice, one is often concerned with the equivalent input drift dV_{gs}/dT at some operating point I_d, g_m. This can be obtained by dividing Eq.(2-7) by g_m, noting that $dI_d = g_m dV_g$, and substituting Eq.(2-15), yielding

$$\frac{dV_{gs}}{dT} = \frac{d\phi}{dT}\left[1 - \left(\frac{I_d}{g_m} \Big/ \frac{I'_d}{g'_m}\right)\right] \qquad (2-20)$$

or in terms of n, using Eq. (1-36) and (1-38), we obtain

$$\frac{dV_{gs}}{dT} = \frac{d\phi}{dT}\left[1 - \left(\frac{I_d}{I'_d}\right)^{1/n}\right] \qquad (2-21)$$

Eq. (2-21) gives the gate to source voltage change with temperature when the drain current is kept constant.

Substituting the value of I_d as given by Eq.(2-16), we get

$$\frac{dV_{gs}}{dT} = \frac{d\phi}{dT} - \sqrt{\frac{I_d}{I_{do}}} \cdot \frac{n}{2T} W_o \qquad (2-22)$$

Eq.(2-22) show that either a positive, zero or negative temperature coefficient can be obtained depending on whether I_d is greater, equal to , or less than I'_d.

It is found out that in the usual impurity concentration range

$$\frac{d\phi}{dT} = 2.1 \ldots 2.5 \text{ mv/deg C} \qquad (2-23)$$

and

$$\frac{1}{\sigma}\frac{d\sigma}{dT} = -5\times10^{-3} \ldots -8\times10^{-3}/\text{deg C} \qquad (2-24)$$

Taking the approximate values of Eq.(2-24) and substituting into Eq.(2-15), the condition for zero temperature coefficient becomes

$$\frac{I'_d}{g'_m} = 0.313 \text{ V}$$

Fig.2-5 shows the I_{dss} $(=I_{do})$ versus temperature characteristic of a number of devices of the Amelco planar FE200 n-channel family. I_d, g_m, and Wo were measured over a temperature range of -54^oC to $+197^oC$. The drain to source voltage is kept at 5V in all measurements of the experiment. It can be seen that high current devices

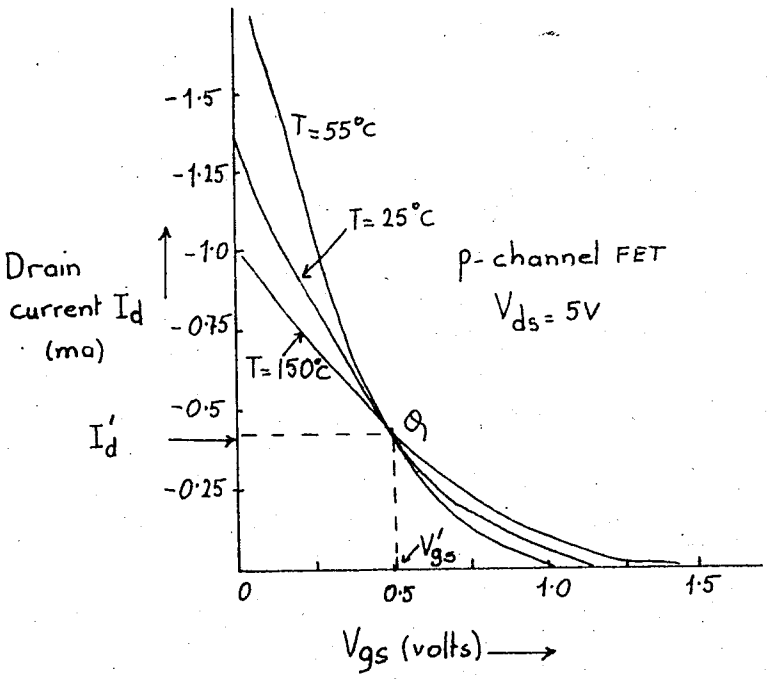


Fig. 2-4. Effect of temperature on transfer curves.

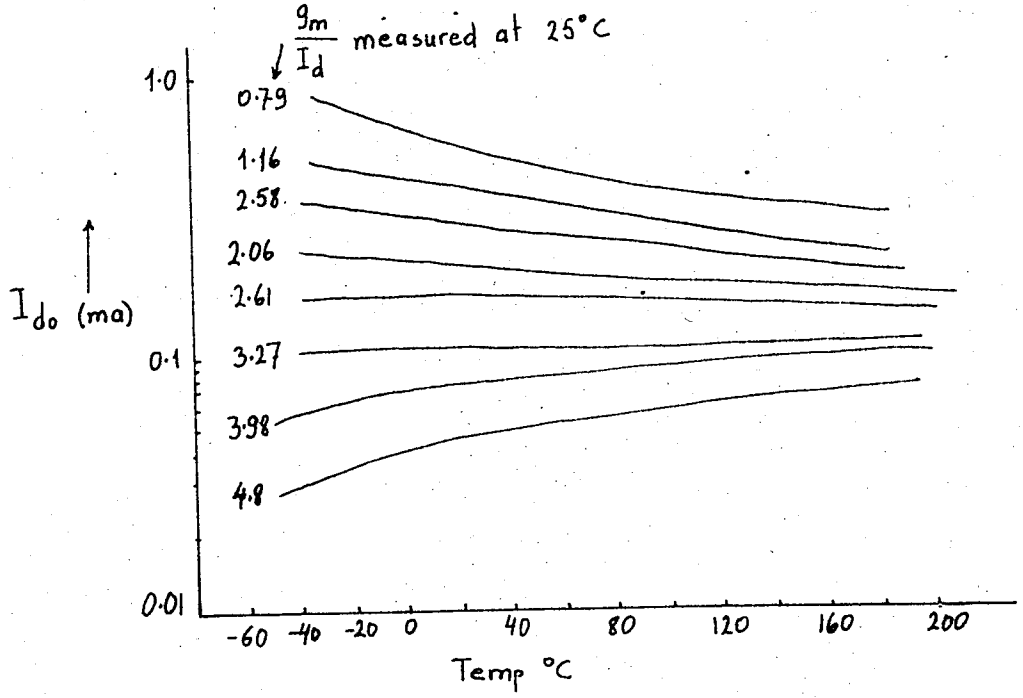


Fig. 2.5. I_{do} vs temp. characteristics of a number of Amelco planar FE200 n-channel devices.
Note that $\frac{g_m}{I_d} = 3.27$ has smallest temp. coefficient.

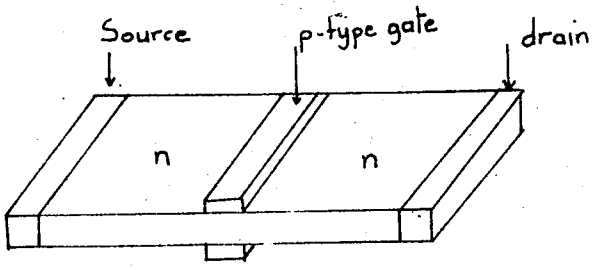
have a negative temperature coefficient and low current devices have positive coefficients. This is due to the fact that the temperature dependence of conductivity is dominant for high current devices. Any FET device can be made to have a nearly zero temperature coefficient if the gate is biased in such a way that I_d and g_m both measured at the operating point are related as in Eq.(2-15)

2-4. Experimentally Obtained Static Characteristics

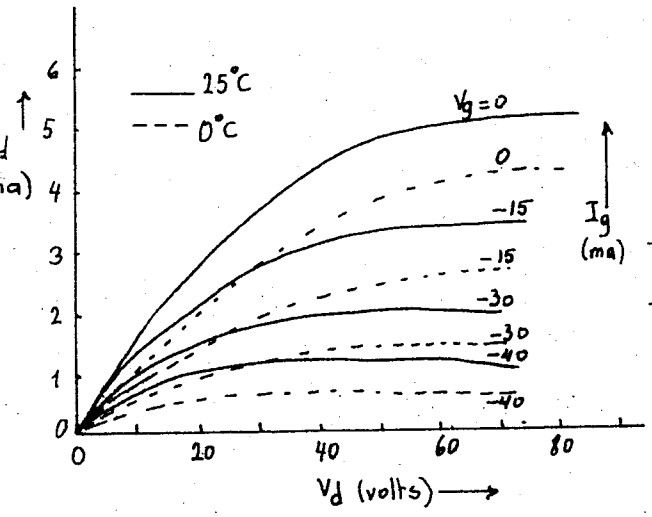
Fig.2-6 shows the shematic diagram of an experimental unit and its various characteristics. This is an n-channel G_e device. It has ohmic contacts at the source and drain ends. The gates are situated half way between source and drain. Therefore $R_d = R_s$. R_d, and R_s can be calculated by a method described in Appendix B. The experimental value of Wo can then be calculated with the help of Fig. 2-6b. From the curve, we find that for $V_g=0$, $I_{ds}=5.5$ mA. Therefore using the value of V_d at which the current saturates, and using the equation $Wo = V_d - I_{ds}R_d$, we can determine Wo.

Fig. 2-6c is the gate current (I_g) vs drain voltage (V_d) characteristics. The gate current consists of holes (for n-channel FETS) which may arise in three ways : thermal generation in the G_e between gate and drain, generation at the surface of the Ge, and injection at the drain. The gate current component due to the first two causes is temperature dependent but constitutes only a small portion of the I_g current. This can be seen from Fig. 2-6d. A change of temperature from 0° to 25°C results only 30% change in I_g. Thus we conclude that the gate current results predominantly from hole injection at the drain.

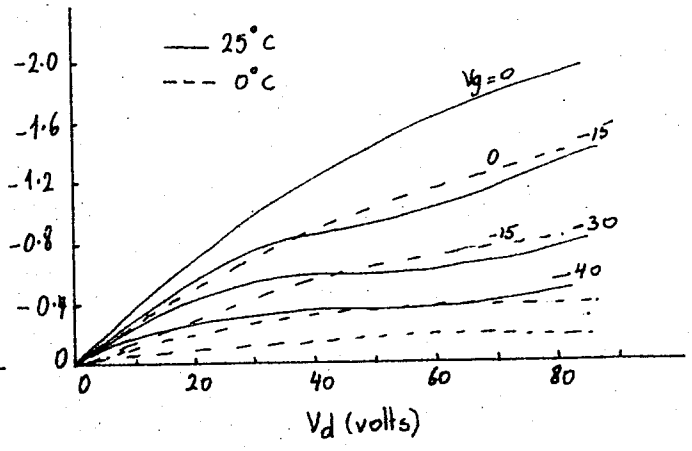
Fig.2-7 shows another unit and its characteristics. The value of R_d is reduced in this specimen since the gate is placed very close to the drain. Therefore the IR_d drop can be neglected so that the value of V_d at saturation is equal to Wo (Fig.2-7b). Fig(2-7c) shows that the gate current due to the injection of holes at the drain has been reduced in comparison to that in Fig 2-6c due to the n-n_+ juction used for the drain contact. The fact that the current I_g increases suddenly at high values of V_d when $V_g=0$



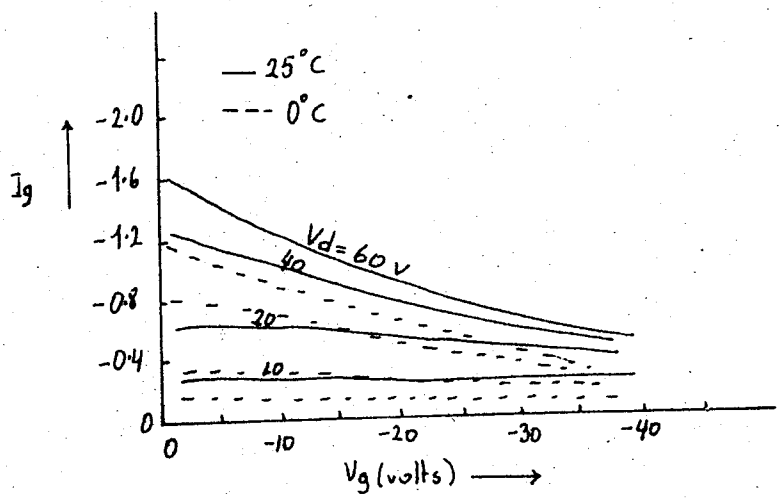
(a)



(b)



(c)



(d)

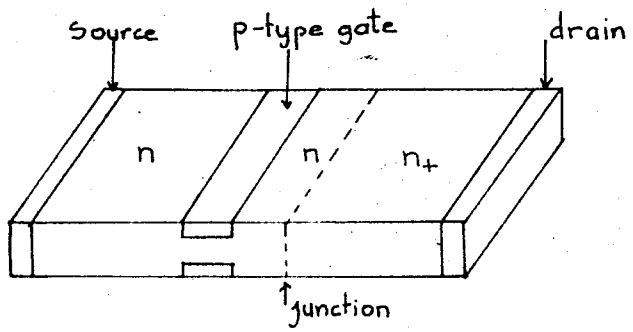
Fig . 2-6 Experimental n-channel FET and its characteristics.

b) Output characteristics

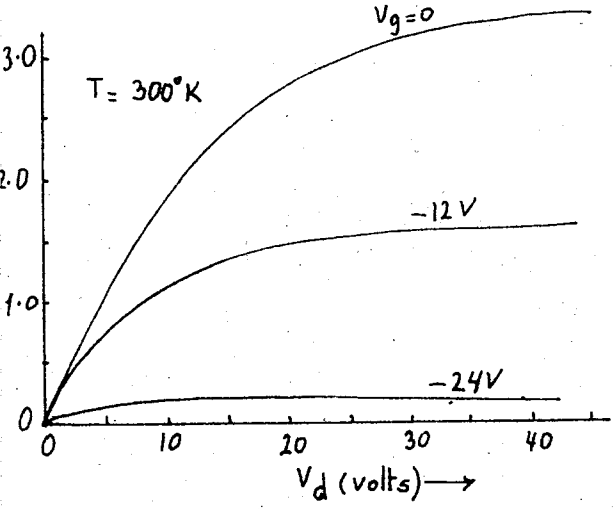
c) Gate current vs drain voltage

d) Gate characteristics

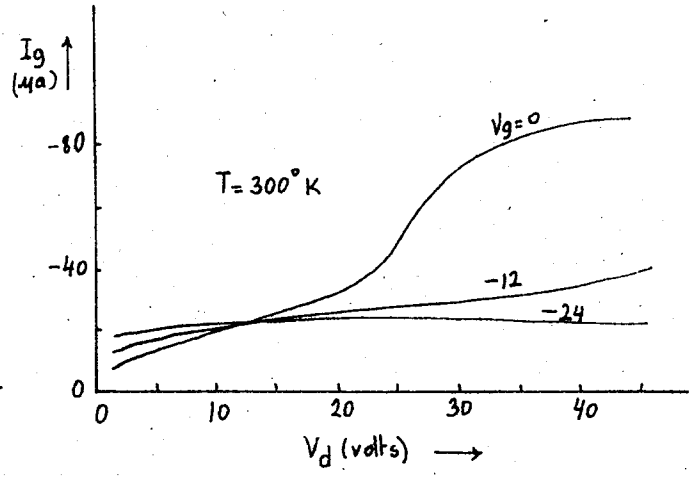
BOĞAZİÇİ ÜNİVERSİTESİ KÜTÜPHANESİ



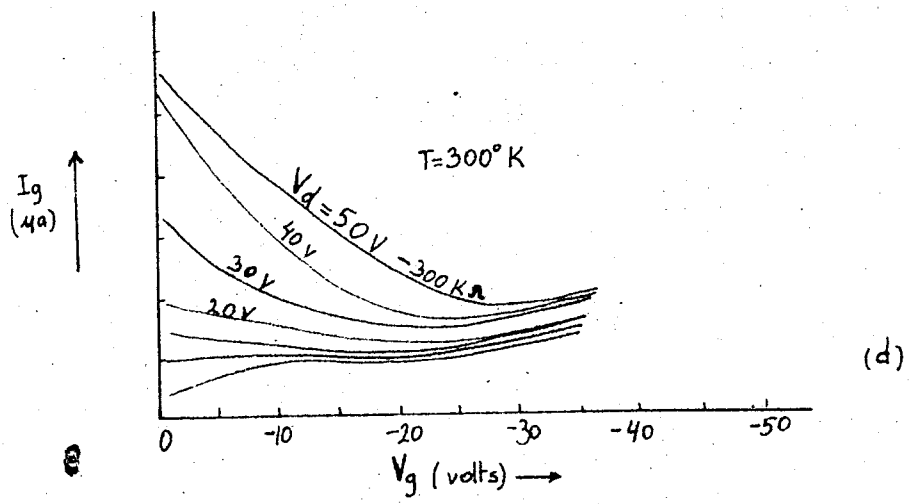
(a)



(b)



(c)



(d)

Fig. 2-7. Experimental n-channel FET with n-n+ junction at the drain

b) drain characteristics

c) gate current vs drain volt

d) gate characteristics

shows that the source of current at such conditions is the thermal generation of holes since these are the conditions under which the power dissipated in the unit is the greatest and hence the temperature highest.

Another feature of the field-effect transistors can also be seen from the results of the experiments. Fig 2-6d, and Fig 2-7d show that FET devices exhibit negative gate resistance at high V_d voltages. However the nature of the gate resistance can be made either positive or negative depending on the nature of the drain contact.

2-5. Breakdown Phenomena and Punch through

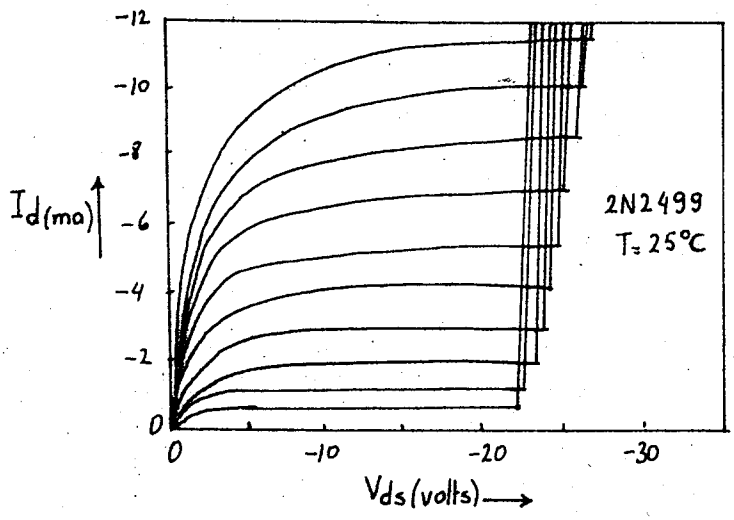
Fig.2-8a represent a typical output characteristic for a 2N2499 FET. These are the curves of drain current, I_d, versus drain voltage, V_{ds}, for the common source configuration with gate to source voltage, V_{gs}, as a parameter. The gate bias voltage is of polarity opposite to that of the drain supply voltage. Therefore for ordinary bias conditions, a greater potential difference exists across the gate-drain diode than exists across the gate-source diode. As a consequence gate-drain diode breakdown will occur before gate source diode breakdown. The break in the drain characteristic curves occur at lower drain voltages as gate voltage is increased, that is, the drain-to-gate breakdown voltage (BV_{dg}) is almost constant and independent of drain-source current I_d. This relationship can be expressed as,

$$BV_{dg} = BV_{dsx} + V_{gs} \cong \text{constant} \qquad (2\text{-}25)$$

In Eq.(2-25) the subscript x denotes the value of BV_{ds} (drain to source voltage at breakdown) for a particular value of V_{gs}. Since $BV_{dg} \cong$ constant, it may be designated by BV_{dgo}. BV_{dgo} may be found by a simple experiment. Using the set up shown in Fig. 2-9.
For the 2N2499 FET, $BV_{dgo} = -20$ volts was found.
Solving Eq.(2-25) for BV_{dsx}

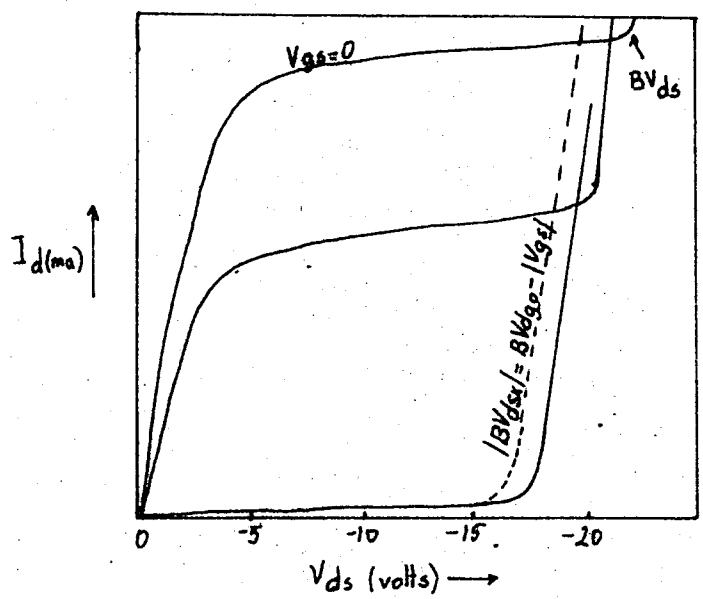
$$|BV_{dsx}| = BV_{dgo} - V_{gs} \cong 20 - |V_{gs}| \qquad (2\text{-}26)$$

Thus a curve can be plotted on the drain characteristics of the FET with the help of Eq.(2-26). In the area to the right of this curve



(a)

Fig.2-8a I_d vs drain voltage V_{ds} for the common source configuration



(b)

Fig. 2-8b. I_d vs drain voltage showing BV_{dsx} locus

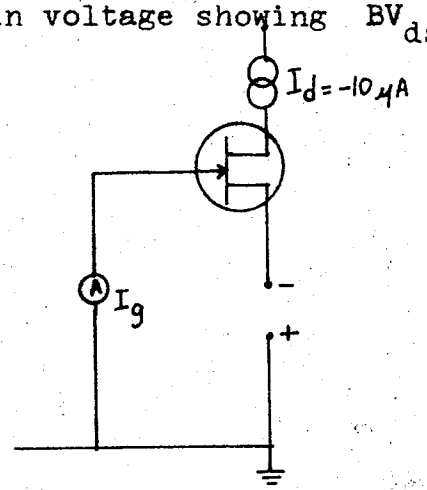


Fig. 2-9 Experimental set-up for measuring BV_{dgo}

breakdown may occur. The location of the breakdown depends on the structure and doping profile of the device. The gate-drain breakdown does not occur at the same time for both of the gates. Let us designate the gate for which breakdown occurs first as gate 1 and the gate for which breakdown occurs at a much higher voltage as gate two. The gate-drain breakdown for gate 2 cannot be observed when the two gates are connected. Even when separate gate connections are made the second breakdown voltage cannot be observed since punch through between the two gates occurs first. To understand the punch through phenomena well, a double diffused n-channel field effect transistor which has a doping profile similar to that shown in Fig. 1-3a, is considered. Letting the source, drain and gate 2 be at zero potential, a reverse bias is applied to gate 1. When the reverse bias reaches a certain value the depletion layer due to gate 1 will touch the depletion layer due to gate 2 as seen in Fig.2-10. As the reverse bias on gate 1 is increased further, the channel will become fully depleted. The space charge region due to gate 1 will get larger at the expense of that due to gate 2. This will produce an effect on gate 2 as though it were forward biased. Finally an appreciable hole current is injected into the space charge region by gate 1 and will flow to gate 2. This phenomena is known as punch through. Under these conditions, while the intergate impedance has dropped to a low value, the gate to source, and gate-to-drain impedance remain large. Fig 2-11a shows the breakdown and punch trough phenomena for an n-channel device when gate 2 is grouded and the reverse bias on gate 1 is increased.

2-6. Negative Resistance in Drain Characteristics

In Si field effect transistors with high pinch off voltages, variations in internal temperature will produce negative resistance. The magnitude of the negative resistance depends on three factors: amount of drain current ,drain current temperature coefficient, and thermal resistance between internal drain and ambient.

Negative resistance becomes evident only when the drain current in an FET is measured with slow discrete changes in drain

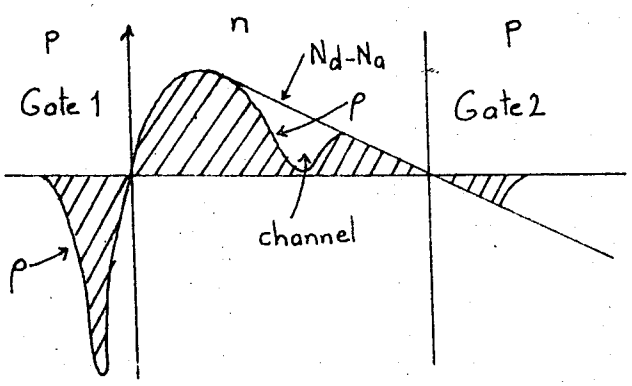


Fig.2-10. Depletion layers when gate 1 is reverse biased (source , drain and gate 2 are grounded)

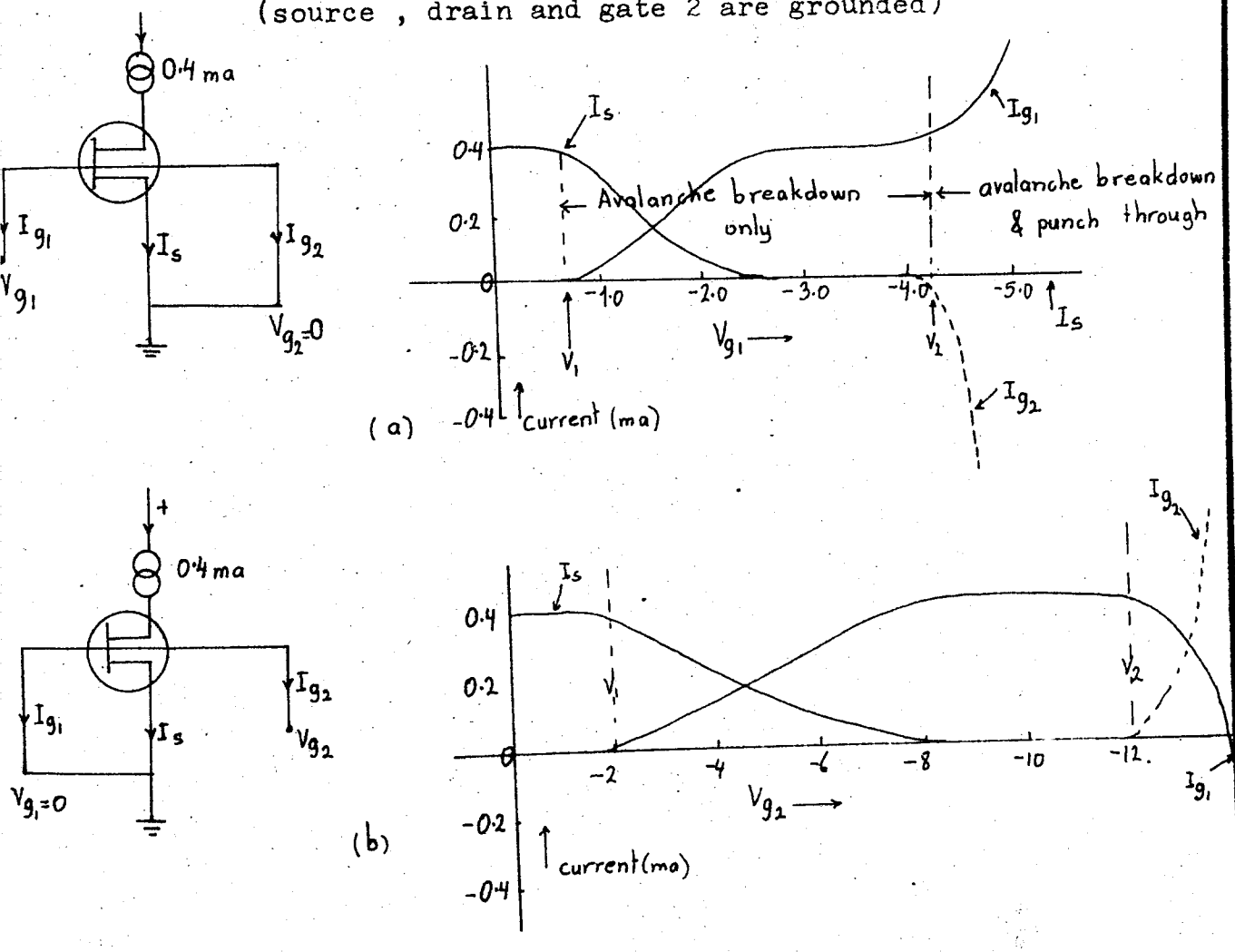


Fig 2-11. Avalanche breakdown and punch through in a double diffused n-channel FET

a) Gate 2 grounded, gate 1 varied

b) Gate 1 grounded, gate 2 varied

to source voltage and the data plotted on a graph.Fig 2-12 shows the V_{ds} - I_d characteristics of a 2N3368 FET measured with an x-y recorder at slow sweep. A very slow rate is necessary to obtain the characteristics of Fig 2-12 representing the thermal equilibrium condition where each point has a chance to reach its stabilized operating temperature before further variation occurs. The thermal time constant for a typical FET is about 25 seconds.

The negative resistance effect for any FET may be enhanced or diminished by controlling its thermal resistance. When negative resistance effect is undesirable, it can be eliminated by mounting the FET on a heat sink to decrease its thermal resistance. Fig.2-13 illustrates the effectiveness of the heat sink in reducing the negative resistance. On the other hand, by reducing the leakage of heat away from the FET, using thermal insulation and lead wires with low thermal conductivity, the thermal resistance may be made as high as possible for applications where negative resistance characteristic is desirable. But in such case, care must be taken not to exceed the FET's rated maximum junction temperature.

Field effect transistor which have low g_m/I_{ds} (a high pinch off voltage) will exhibit negative resistance effect while the FET's which have a high g_m/I_{ds} ratio (a low pinch-off voltage) will not exhibit negative resistance.

The negative resistance in the FET characteristic may produce detrimental effects such as a bistable bias condition in circuits designed for linear operation. On the other hand the effect may be utilized in certain unique circuits. For examples a bistable memory practically immune to noise may be made with the negative resistance characteristics. But in such applications, since changes in external temperature will also shift the FET characteristic curves, substantial fluctuations in the ambient temperature must be avoided.

2-7. Forward Biased Gate-channel Junction

Since forward biasing of an FET causes large gate current, in representing the drain characteristics of a forward biased FET

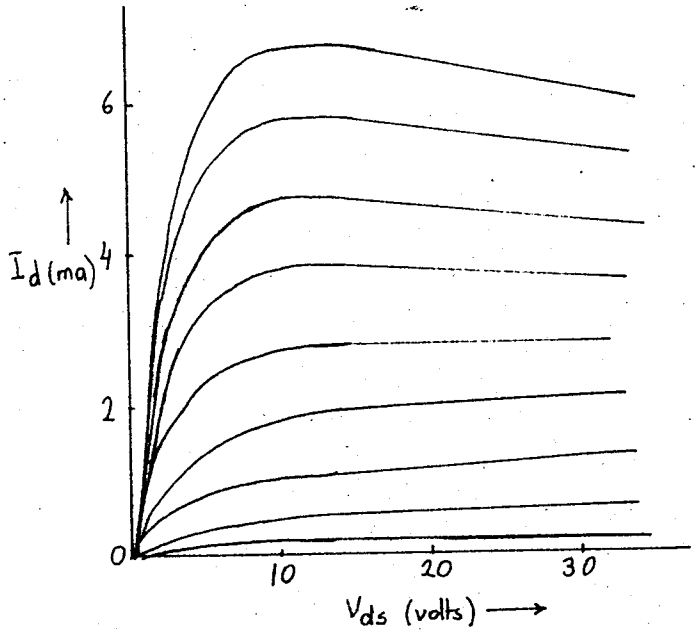


Fig.2-12. V_{ds} vs I_d characteristics for a 2N3368 FET measured with an x-y recorder at slow sweep .

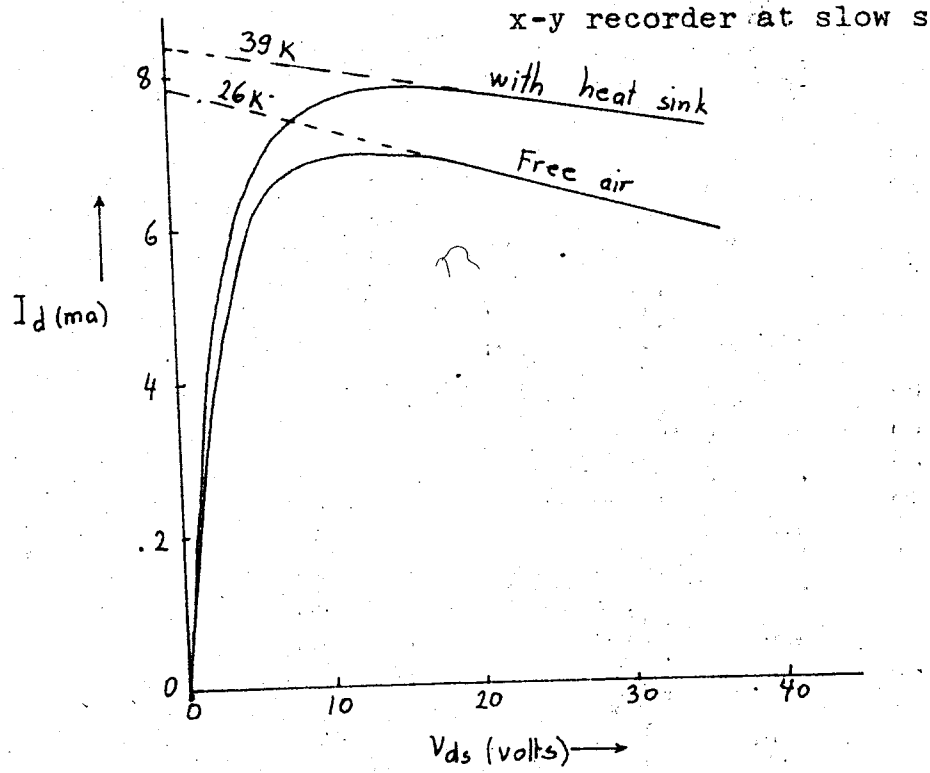


Fig.2-13 V_{ds} vs I_d characteristics for a 2N3368 showing the effect of a heat sink on negative drain resistance. Both curves were measured with $V_{gs} = 0$

it is sometimes prefered to use the gate current as a parameter rather than the gate to source voltage. Fig.2-14a shows the complete drain characteristics of a p-channel FET. In this figure , the gate to source voltage is used as a parameter when the gate to channel junctions are reverse biased, and the gate current is used as a parameter when the junctions are forward biased. The boundary between the two conditions of operation is the zero gate to source voltage characteristic. The dotted curve, defined by the pinch-off voltage forms a dividing line between the active and bottomed regions of operation.

R.S. Cobbold and F.N.Trofimenkoff point out the similarity between the drain characteristics of the gate-to-channel junctions forward biased FET and the collector characteristics of a conventional transistor operating in the common emitter configuration. They define a new parameter for the forward biased FET which they designate as h_{fs} in a manner analogous to h_{fe} for a conventional transistor. In Fig. 2-14b a plot of h_{fs} for the same transistor used in Fig.2-14a is given. It can be seen that h_{fs} is a function of the gate current and decreases as the gate current increases.

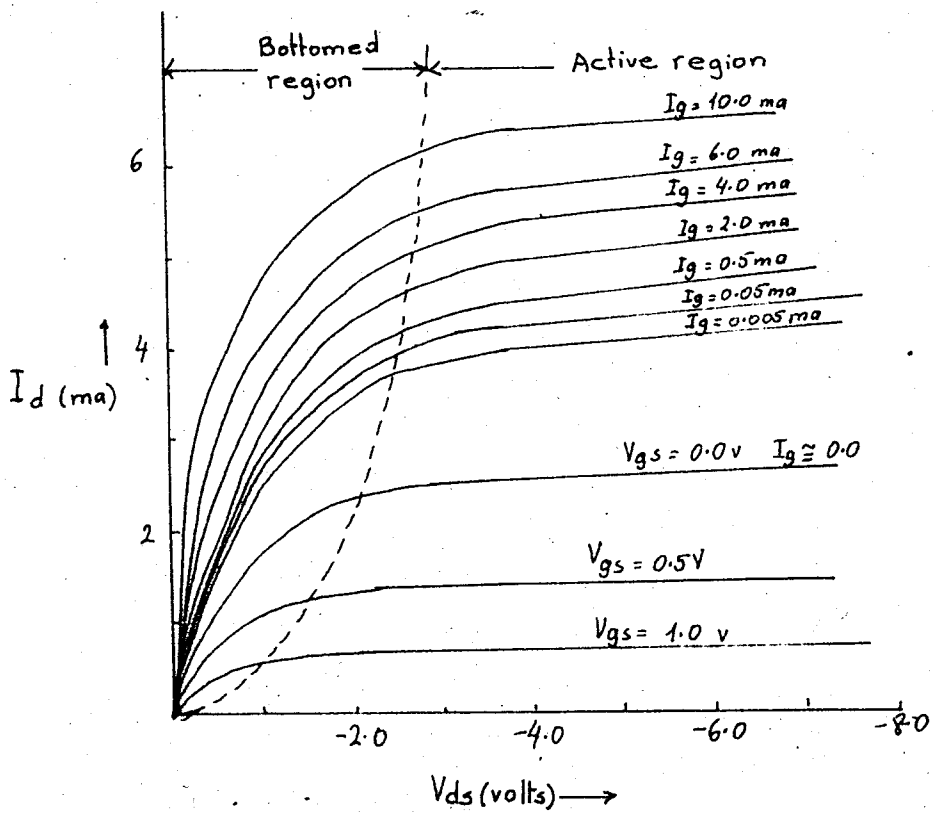


Fig. 2-14a . Drain characteristics of a 2N2498 FET Showing forward biased region.

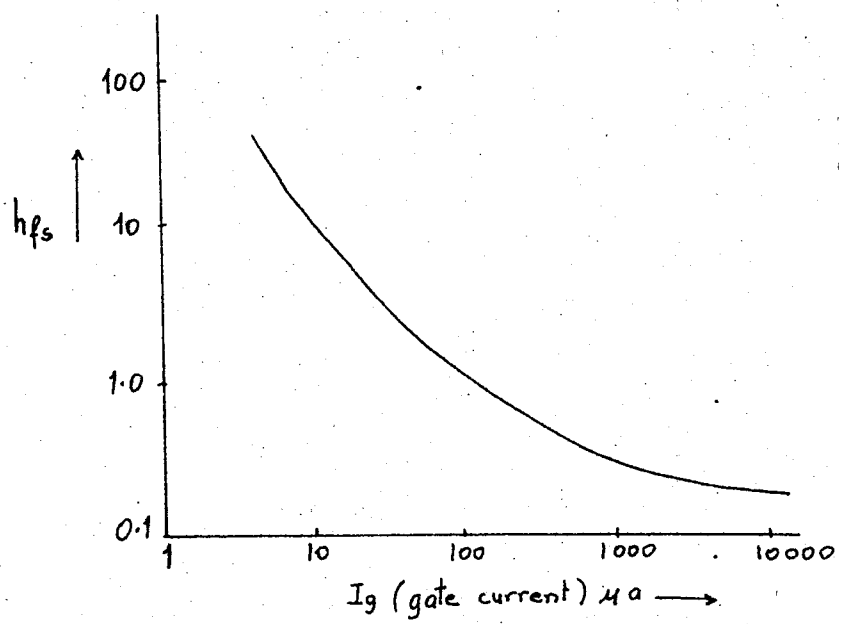


Fig. 2-14b. h_{fs} as a function of I_g for 2N2498 FET.

CHAPTER 3

SMALL SIGNAL EQUIVALENT CIRCUIT OF JUNCTION-GATE FETS

3-1. Introduction.

It is possible to derive an equivalent circuit for the FET at very low frequencies based on the observed or calculated characteristics. The commonest equivalent dc-circuit used is the one shown in Fig.3-1a wherein R_{gs} and R_{gd} represent the leakage resistance of the reverse biased gate junction. The parasitic resistances in the gate, source and drain leads are not shown in Fig.3-1a and hence this equivalent circuit is a representation of the intrinsic FET structure only.

When the FET bias voltages are changed, a redistribution of charge within the device takes place via the external circuits. At high frequencies this charge movement becomes important and can be represented in an equivalent circuit by using capacitive elements. The capacitance C_{gd} between the gate and the drain, and the capacitance C_{gs} between the gate and the source can be computed using the space charge approximation to the potential function in the vicinity of the gate-channel p-n junction. $Q'dx$ is the distributed charge per length dx stored in the space charge regions (a-b) of the Shockley unit (Fig. 1-1), this charge will be compensated by an equal and opposite charge Qdx per length dx at the gate contact so that

$$Qdx = 2a\,\rho_o\left(1-\frac{b}{a}\right)\,dx \qquad (3\text{-}1)$$

By intergrating Eq.(3-1) Van der Ziel has shown that the total charge Q expression becomes

$$Q = \frac{Q_o\left[3\left\{\left(\frac{W_d}{W_o}\right)^2-\left(\frac{W_s}{W_o}\right)^2\right\}-4\left\{\left(\frac{W_d}{W_o}\right)^{\frac{3}{2}}-\left(\frac{W_s}{W_o}\right)^{\frac{3}{2}}\right\}\right]}{\left[\frac{I_d}{I_{do}}-\frac{6L}{a}\right]} \qquad (3\text{-}2)$$

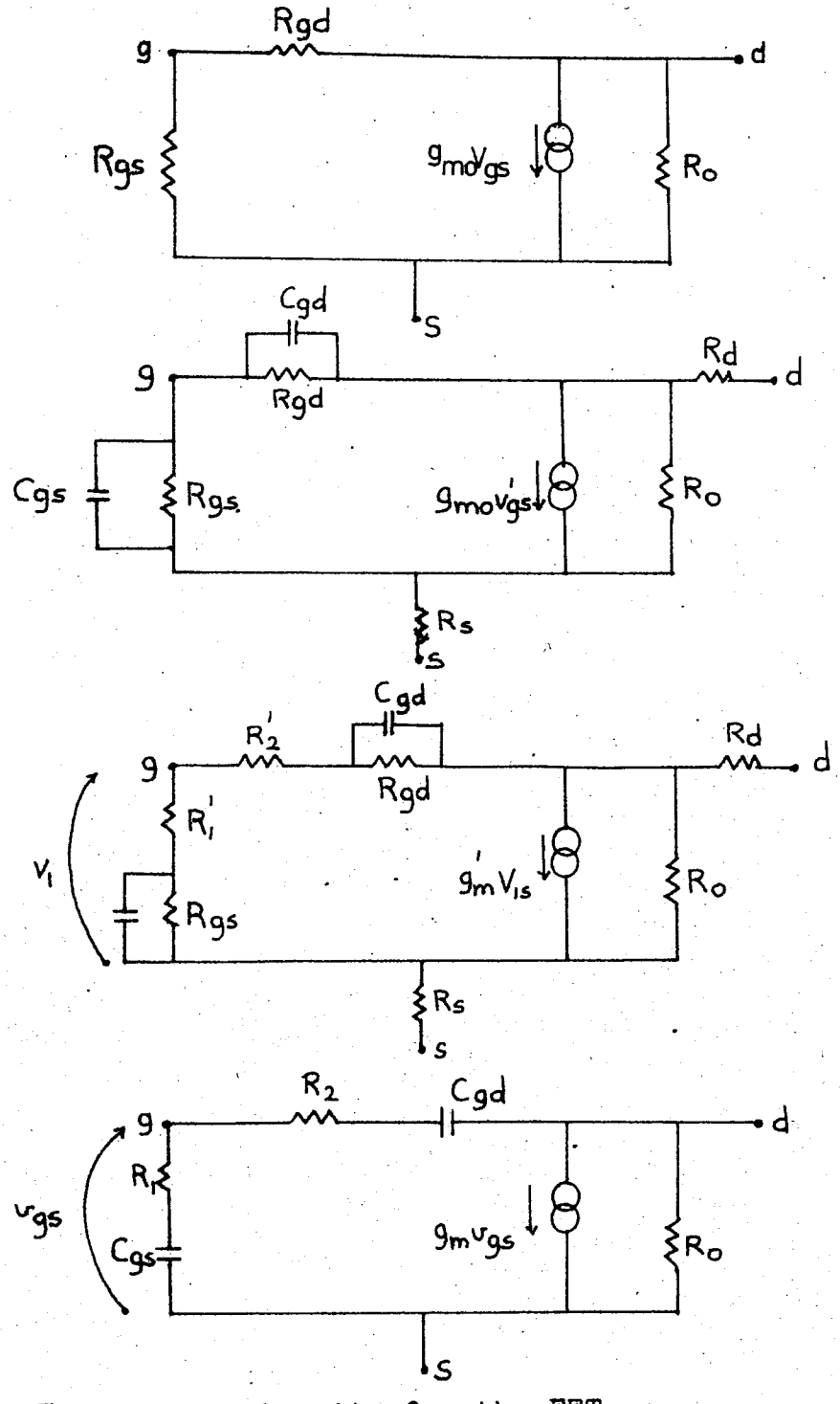


Fig.3-1. Equivalent circuits for the FET

a) Original Shockley low-frequency circuit

b) Shockley equivalent circuit with the addition capacitances C_{gd} and C_{gs}

c) Complete intrinsic equivalent circuit including the parasitic resistances R_d and R_s.

d) Simplified equivalent circuit

where $Q_o = qN_c aLZ$ (3-3)

N_c = ionized impurity in the channel (N_a or N_d)

If the drain is a.c short-circuited to the source, a capacitance C_{gg} can be defined as

$$C_{gg} = -\frac{\partial Q}{\partial V_{gs}} \quad (3-4)$$

where $$C_{gg} = C_{gd} + C_{gs} \quad (3-5)$$

Similarly if the gate is a.c short-circuited to the source, a capacitance C_{gd} can be defined as

$$C_{gd} = -\frac{\partial Q}{\partial V_{ds}} \quad (3-6)$$

Then C_{gs} becomes

$$C_{gs} = C_{gg} - C_{gd} \quad (3-7)$$

Expressions for these capacitances are derived by Van der Ziel, and will be given in appendix C.

The equivalent circuit of Fig. 3-1a can then be modified by inserting C_{gd} and C_{gs} to yield the equivalent circuit shown in Fig. (3-1b). Values R_d and R_s are the parasitic bulk resistances of the semiconductor path from the channel edges to the drain and source contacts respectively, they will be in the order of 100 ohms or less, depending somewhat on geometry and manifacturing process. At low frequencies, the effect of R_d is quite negligible, it can be considered as only a very small part of any practical load resistance. The value of R_s has a slight, generally negligible effect on the apperent transconductance of the device, the voltage v'_{gs} in Fig.3-1b is related to the therminal voltage v_{gs} by

$$v'_{gs} = \frac{v_{gs}}{1+g_{mo}R_s} \quad (3-8)$$

At very high frequencies the equivalent circuit of Fig.3-1b will be inadequate.The resistivity of the channel due to charge redistribution in response to rapid changes in the terminal voltages must be considered. In order to solve the pro-

blem exactly the differential equation of a tapered resistance capacitance transmission line must be solved. The derivations of such analysis have shown that Fig. 3-1b must be modified as shown in Fig. 3-1c by the addition of resistance R_1', and R_2'. The current generator must also be modified by $g_m' v_{gs}$, where

$$g_m' = \frac{g_{mo}}{1 + j\omega\tau} \qquad (3\text{-}9)$$

g_{mo} = low frequency value of transconductance

$$\tau = \gamma R_1' C_{gs} \qquad (3\text{-}10)$$

Theoretical calculations carried out for the case of uniform channel FET have shown that, the value of constant γ for the pinch off case varies between 0 and 1 depending on the value of the gate bias voltage.

The circuit of Fig.3-1c can be modified to that shown in Fig.3-1d in order to make high-frequency calculutions simpler. In this circuit

$$R_1 = R_1' + R_s$$
$$R_2 = R_2' + R_d \qquad (3\text{-}11)$$
$$g_m = g_m' \frac{(1 - j\omega C_{gs} R_1')}{(1 + j\omega C_{gs} R_1)} \cdot \frac{(1 + j\omega C_{gd} R_2')}{(1 + j\omega C_{gd} R_2)}$$

For this simplified equivalent circuit to be valid

$$j\omega C_{gs} \gg 1/R_{gs} \; ; \; j\omega C_{gd} \gg 1/R_{gd} \text{ and } g_{mo} R_d, R_d/R_o, R_s/R_o \ll 1 \qquad (3\text{-}12)$$

3-2. Equivalent Circuit Parameter Determinations

3-2A. Admittance parameters :

It is common practice among manufacturers to specify the FET by its equivalent short-circuit admittance parameters. A Texas Instruments data sheet is shown in the next page. The general two-part y-parameter network is shown in Fig 3-2. The terminal small-signal voltages and currents are referred to the common source connection. The parametric equations of the network are :

$$i_g = y_{is} \; v_{gs} + y_{rs} \; v_{ds}$$

Types TIX 880 , TIX 881, TIX 882, TIX 883 N-channel alloy-junction Ge FETs. Texas Instruments Data Sheet. Electrical characteristics at 25°C free air temperature (unless otherwise noted)

Parameter	Test Conditions	TIX 880			TIX 881			TIX 882			TIX 883			UNIT
		MIN	TYP	MAX	MIN	TYP	MAX	MIN	TYP	MAX	MIN	TYP	MAX	
BV_{dgo} drain-gate breakdown voltage	I_d=100 μa I_s=0	25	60	—	25	60	—	30	60	—	35	60	—	V
I_{gss} gate cutoff current	V_{gs} = -10 V V_{ds} =0	—	-2	-5	—	-2	-5	—	-2	-5	—	-2	-5	μa
I_{gss} gate cutoff current	V_{gs}=-10 V T_A=55°C V_{ds}=0	—	-12	-25	—	-12	-25	—	-12	-25	—	-12	-25	μa
$I_{d(on)}$ Zero gate voltage drain current	V_{ds}=20 v V_{gs}=0	0·5	1·0	1·5	1·0	1·7	2·5	2·0	3·5	5·0	4·0	6·0	7·5	ma
$I_{d(off)}$ Pinchoff drain current	V_{ds} = (see note 4) V_{gs} = (see note 5)	—	2	10	—	2	10	—	20	100	—	20	100	μa
r_{ds} static drain-source resist.	I_d=100 μa V_{gs}=0	—	1400	2500	—	1000	1800	—	700	1400	—	500	1200	ohm
$\|Y_{is}\|$ small signal common source input adm.	I_d see note 6 V_{ds}= 15 v f= 1kc	—	0·2	1·0	—	0·2	1·0							μmho
$\|Y_{rs}\|$ Small signal common source reverse transfer adm.		—	0·05	0·5	—	0·05	0·5							μmho
$\|Y_{fs}\|$ small signal common source forward transfer adm.		225	500	1125	350	700	1400							μmho
$\|Y_{os}\|$ small signal common source output adm.		—	9	12	—	10	15							μmho
$\|Y_{fs}\|$ Small signal common source forward transfer adm.	V_{ds}= 15 v f=10mc I_d (see note 6)	200	450	—	300	625	—							μmho
C_{iss} Common source short-circuit input capacitance	V_{gs}=0 V_{ds}= 20 v f= 1mc	—	26	—	—	26	—	—	26	—	—	26	→	pf

notes					
4	V_{ds}= 15 v	V_{ds}= 15 v	V_{ds} = 20 v	V_{ds} = 20 v	
5	V_{gs}= -8 v	V_{gs}= -10 v	V_{gs} = -10 v	V_{gs} = -15 v	
6	5 I_d = 0·5 ma	I_d = 1ma	I_d = 2 ma	I_d = 4ma	

THESIS
ROBERT COLLEGE GRADUATE SCHOOL
BEBEK, ISTANBUL

$$i_d = y_{fs} v_{gs} + y_{os} v_{ds} \qquad (3\text{-}13)$$

The terminal conditions for determining the parameters are :

output shorted:

$$y_{is} = \frac{i_g}{v_{gs}}$$

$$y_{fs} = \frac{i_d}{v_{gs}} \qquad (3\text{-}14)$$

input shorted

$$y_{rs} = \frac{i_g}{v_{ds}}$$

$$y_{os} = \frac{i_d}{v_{ds}}$$

When these conditions are applied to the physical equivalent circuit of Fig. 3-16 , the y-parameters at low and medium frequencies can be written in terms of the lumped pysical elements:

$$y_{is} = j\omega(C_{gd} + C_{gs})$$

$$y_{rs} = -j\omega C_{gd} \qquad (3\text{-}15)$$

$$y_{fs} = g_m - j\omega C_{gd}$$

$$y_{os} = G_o + j\omega C_{gd}$$

Note that all the diode conductances and bulk resistances are neglected and also that these parameters are all bias dependent. Making use of the square-low approximation of I_d given by Eq. (1-38) and differentiating I_d with respect to v_{gs}, we can get an approximate idea about g_m.

$$\frac{dI_d}{dV_{gs}} = g_m = \frac{2I_{do}}{W_o}\left(1 - \frac{V_{gs}}{W_o}\right) \qquad (3\text{-}16)$$

$$= 2/W_o\sqrt{I_{do} I_d}$$

As for the y-parameters at very high frequencies, we make use of Fig. 3-id as follows :

$$y_1 = \frac{j\omega C_{gs}}{1 + j\omega C_{gs} R_1}$$

$$y_2 = \frac{j\omega C_{gd}}{1 + j\omega C_{gd} R_2} \qquad (3\text{-}17)$$

It can be shown that

$$\begin{aligned} y_1 &= y_{11} + y_{12} \\ y_2 &= -y_{12} \\ \frac{1}{R_o} &= y_{22} + y_{12} \\ g_m &= y_{21} - y_{12} \end{aligned} \qquad (3\text{-}18)$$

or that

$$\begin{aligned} y_{11} &= y_1 + y_2 = y_{is} \\ y_{12} &= -y_2 = y_{rs} \\ y_{21} &= -y_2 + g_m = y_{fs} \\ y_{22} &= y_{21} + \frac{1}{R_o} = y_{os} \end{aligned} \qquad (3\text{-}19)$$

The elements of the proposed equivalent circuit are thus simply and conveniently related to the admittance parameters, and a measurement of the admittance parameters as a function of frequency can be readily used to obtain the equivalent circuit elements.

3-2B. Results of Admittance parameter measurements:

Some measurement technique of the admittance parameters is given in the appendix D.

Here we shall discuss the variation of these parameters with frequency.

1. $y_{11} = y_{is} = g_{11} + jb_{11}$

From Eqns. (3-18) and (3-19)

$$y_{is} = \frac{(\omega C_{gs})^2 R_1}{1 + (\omega C_{gs} R_1)^2} + \frac{(\omega C_{gd})^2 R_2}{1 + (\omega C_{gd} R_2)^2} + j\left[\frac{\omega C_{gs}}{1 + (\omega C_{gs} R_1)^2} + \frac{\omega C_{gd}}{1 + (\omega C_{gd} R_2)^2}\right] \qquad (3\text{-}20)$$

For low frequencies :

$$(\omega C_{gs} R_1)^2 \ll 1 \; ; \; (\omega C_{gd} R_2)^2 \ll 1$$

Thus Eq.(3-20) reduces to

$$\begin{aligned} g_{11} &\cong \omega^2 (C_{gs}^2 R_1 + C_{gd}^2 R_2) \\ b_{11} &\cong \omega (C_{gs} + C_{gd}) \end{aligned} \qquad (3\text{-}21)$$

Fig. 3-3 shows the plot of g_{11} and b_{11} as a function of frequency.

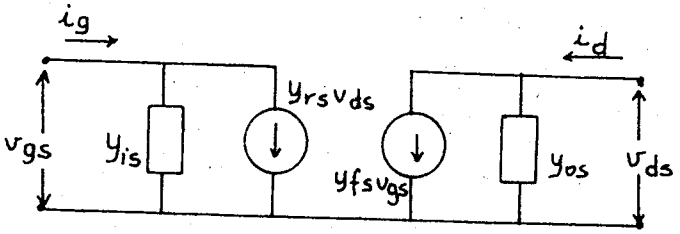


Fig. 3-2. General two port y-parameter network

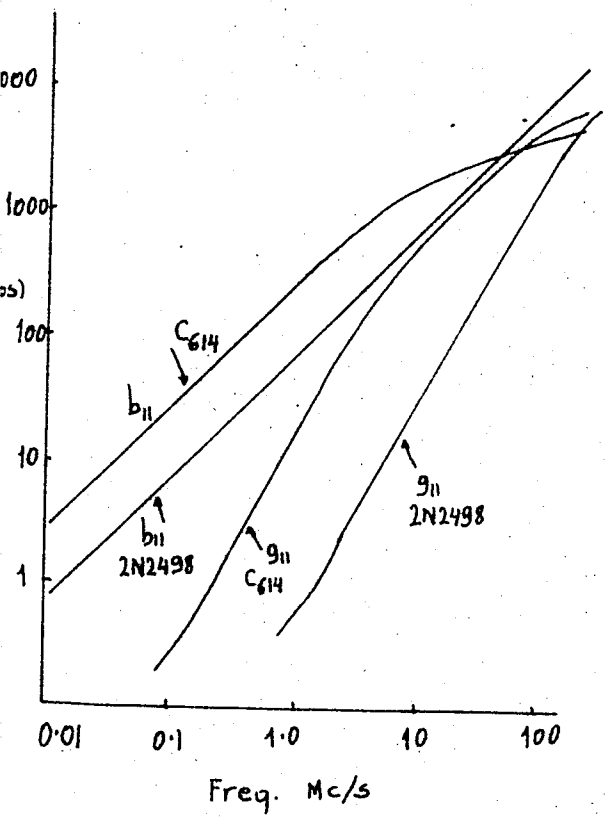


Fig. 3-3 Plot of g_{11} and b_{11} as a function of frequency

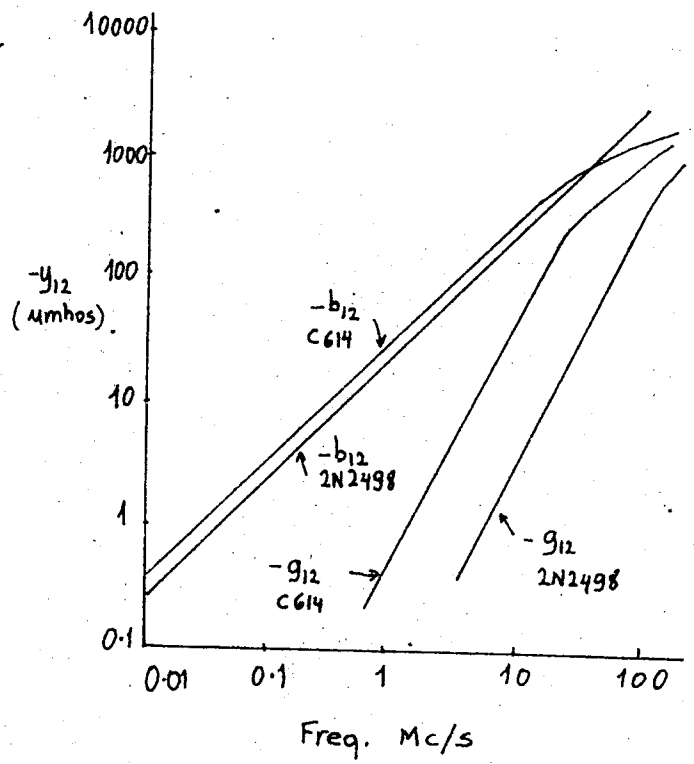


Fig.3-4. Plot of g_{12} and b_{12} as a function of frequency

We can see that g_{11} is very small at low frequencies so that the expression given by Eq. 3-15 and therefore the low frequency circuit of Fig. 3-1a is good approximation.

It can be seen from Fig. 3-3 that the plots of b_{11} and g_{11} are linear for frequencies below 100 mc/s. Thus the slopes of these lines can be used to calculate the values of $(C_{gd} + C_{gs})$ and $(C_{gs}^2 R_1 + C_{gd}^2 R_2)$.

2. $y_{12} = y_{rs} = g_{12} + jb_{12}$

From Eqns. (3-18) and (3-19)

$$y_{rs} = -y_2 = -\left[\frac{(\omega C_{gd})^2 R_1}{1 + (\omega C_{gd} R_1)^2} + \frac{j\omega C_{gd}}{1 + (\omega C_{gd} R_1)^2}\right] \qquad (3\text{-}22)$$

at low frequencies

$$g_{12} = -(\omega C_{gd})^2 R_2$$
$$b_{12} = -\omega C_{gd} \qquad (3\text{-}23)$$

at very high frequencies

$$g_{12} \cong -\frac{1}{R_2}$$
$$b_{12} \cong -(\omega C_{gd})^{-1} \qquad (3\text{-}24)$$

Fig. 3-4 shows the plot of g_{12} and b_{12} as a function of frequency. If this information is combined with that derived from y_{is} data, C_{gs}, C_{gd}, R_1, and R_2 can all be calculated.

3. $y_{21} = y_{fs}$

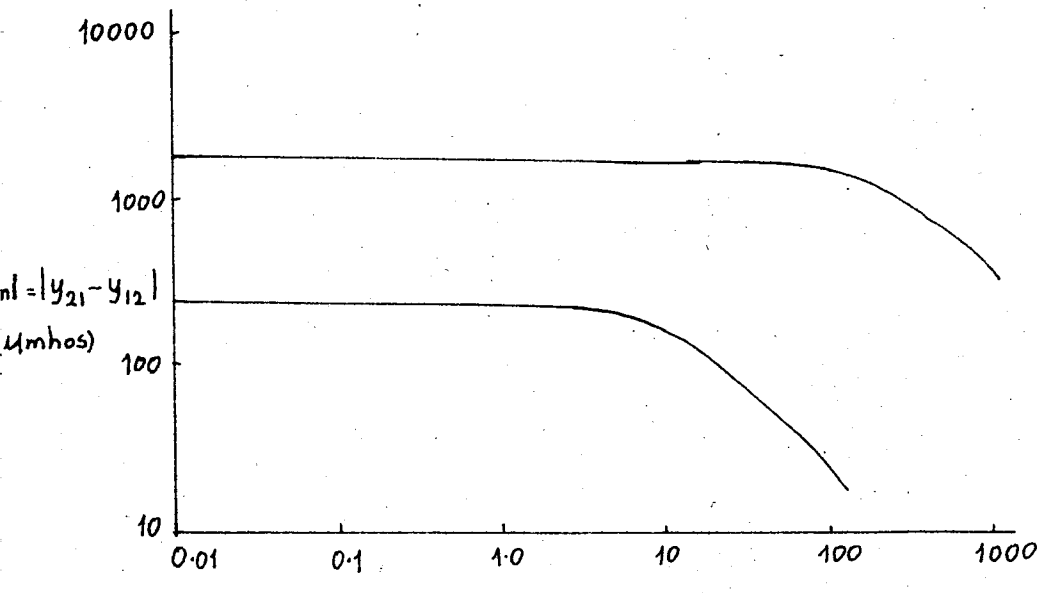
The measurements of the forward- transfer admittance presents a difficult problem, but methods of measuring $y_{11} + y_{21}$ is possible as discussed in appendix E. From these measurements and y_{11} measurements, the values of g_m at different frequencies can be calculated using the relation

$$g_m = y_{21} - y_{12} \qquad (3\text{-}25)$$

Fig. 3-5 gives the plot of the modulus of g_m. From this figure, the modulus of g_m can be fitted with an equation of the form

$$g_m = \frac{g_{mo}}{1 + j\omega\tau} \qquad (3\text{-}26)$$

Fig. 3-5 can be used to obtain the value of g_{mo} and τ . The ratio



ig 3-5 Plot of $g_m = |y_{21} - y_{12}|$ as a function of frequency

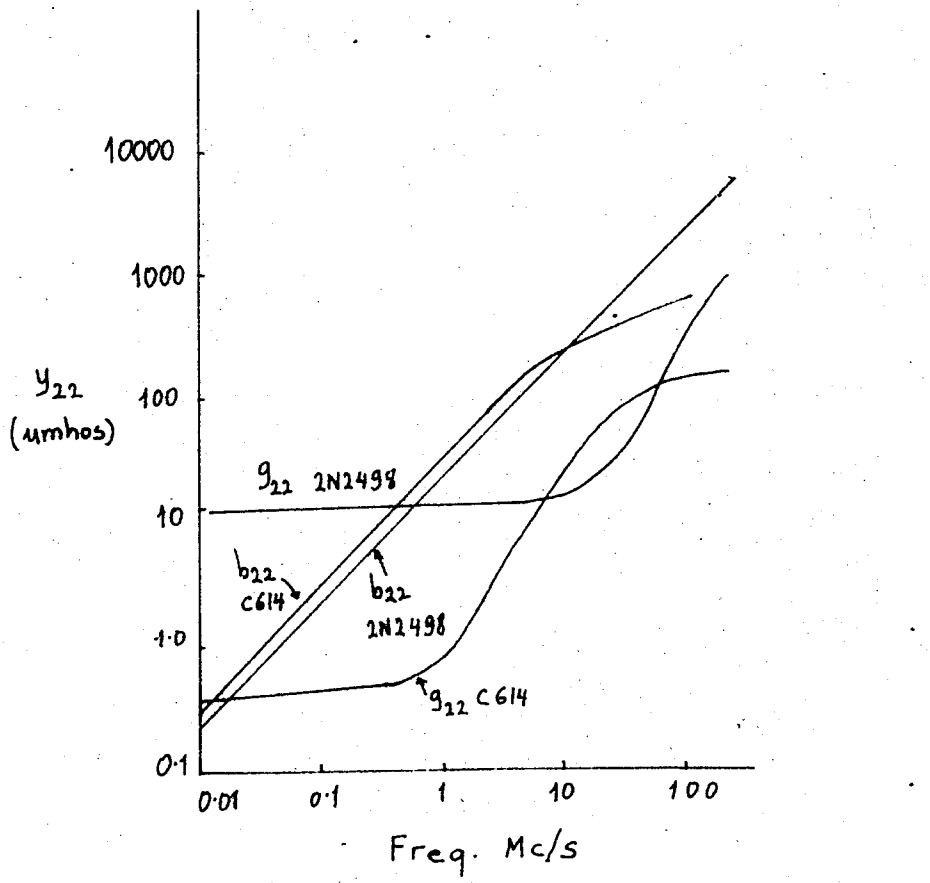


Fig 3-6 Plot of g_{22} and b_{22} as a function of frequency.

of $\tau/R_1 C_{gs}$ must give th value of γ used in Eq.(3-10) . The theoretical and experimental values of the constant γ, which is called the impurity profile constant, is due to additional capacitance and resistance between the source and drain and lead inductance effects.

4. $y_{22} = y_{os} = g_{22} + jb_{22}$

From Eqns. (3-18) and (3-19)

$$y_{os} = \frac{1}{R_o} + \frac{(\omega C_{gd})^2 R_2}{1 + (\omega C_{gd} R_1)^2} + j\frac{\omega C_{gd}}{1 + (\omega C_{gd} R_2)^2} \qquad (3-27)$$

Three characteristic frequency ranges can be recognised. At very low frequencies:

$$g_{22} \cong \frac{1}{R_o}$$
$$b_{22} \cong \omega C_{gd} \qquad (3-18)$$

At medium frequencies

$$g_{22} \cong (\omega C_{gd})^2 R_2$$
$$b_{11} \cong \omega C_{gd} \qquad (3-29)$$

At very high frequencies

$$g_{22} \cong \frac{1}{R_2}$$
$$b_{22} \cong (\omega C_{gd} R_2^2)^{-1} \qquad (3-30)$$

Fig. 3-6 shows the plot of g_{22} and b_{22} as a function of frequency for a number of sample FETS.

CHAPTER 4

PHYSICAL CHARACTERISTICS OF SEMICONDUCTOR SURFACES

The band theory of solids is derived with the assumption of perfect crystals which are infinite in all directions. But real crystals are finite and surface imperfection problems are unavoidable. The surface of a crystal may be defined as a dividing plane between two bulk phases. It is usually treated as a region since the mutual interaction of the two phases is not localized in a plane.

A major crystal imperfection is localized electronic energy states. Such states are localized at the surface and therefore they are called surface states. In order to understand surface states, we must first investigate the physics of semiconductor surfaces in the absence of surface states.

4-1 Semiconductor Surfaces in the Absence of Surface States for gas or vacuum-semiconductor systems:

To simplify the analysis certain assumptions are made :

a) Normal temperature exists so that all shallow donors and acceptors are ionized.

b) The semiconductor is not degenerate and Maxwell-Boltzman statics can be applied.

c) The donor and acceptor concentrations are uniform throughout the semiconductor up to the interface.

d) N_s stationary (trapped) charges other than ionized donors or acceptors in the semiconductor space charge region exist.

e) The crystal is semi-infinite, homogeneous and in thermal equilibrium. That is, if z-axis is taken perpendicular to the surface , the surface conditions are uniform in planes normal to z .

4-1A. Carrier Densities and Band Model

In thermal equilibrium, the following equations hold true

for the semiconductor bulk.

$$n = N_c \exp\left[-(E_c - E_f)/kT\right]$$
$$= n_i \exp\left[(E_f - E_i)/kT\right] \qquad (4-1)$$
$$= n_i \exp\left[q\phi/kT\right]$$

$$p = N_v \exp\left[(E_v - E_f)/kT\right]$$
$$= n_i \exp\left[(E_i - E_f)/kT\right] \qquad (4-2)$$
$$= n_i \exp\left[-q\phi/kT\right]$$

$$np = n_i^2 = N_c N_v \exp -(E_c - E_v)/kT = \text{a constant for a given temperature.} \qquad (4-3)$$

where n = concentration of free electrons

p = concentration of holes

$$N_c, N_v = 2\left(\frac{2\pi m_{e,h} kT}{h^2}\right)^{3/2} = 2.5 \times 10^{19}\left(\frac{m_{e,h}}{m}\right)^{3/2} cm^{-3}$$

q = electronic charge

ϕ = electrostatic potential

E_f = Fermi energy level

E_i = Fermi energy for intrinsic material

$q\phi = E_i - E_f$

$\frac{m_e}{m} = 0.55$ for Ge and 1.10 for Si

$\frac{m_h}{m} = 0.37$ for Ge and 0.57 for Si

Since for an intrinsic material $n = p = n_i$ and $E_f = E_i$

$$E_i = \frac{1}{2}(E_c - E_v) + \frac{1}{2} kT \ln \frac{N_v}{N_c} \cong \frac{1}{2}(E_c - E_v) \qquad (4-4)$$

Eq.(4-4) is only an approximation since $N_v \neq N_c$ due to the difference between the effective masses of the holes and electrons.

Fig 4-1 is an energy level diagram indicating the various parameters used to characterize the surface, in doped semiconductors. E_i is taken as the reference energy level.

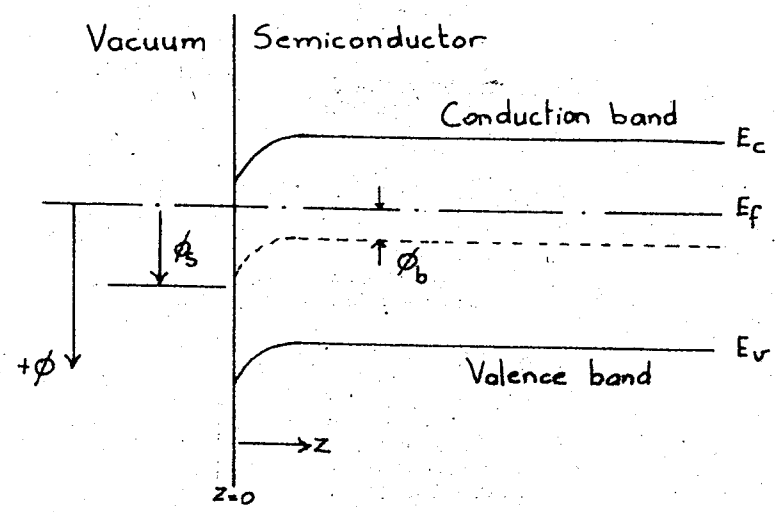


Fig.4-1 Energy-level diagram indicating the various parameters used to characterize the surface.

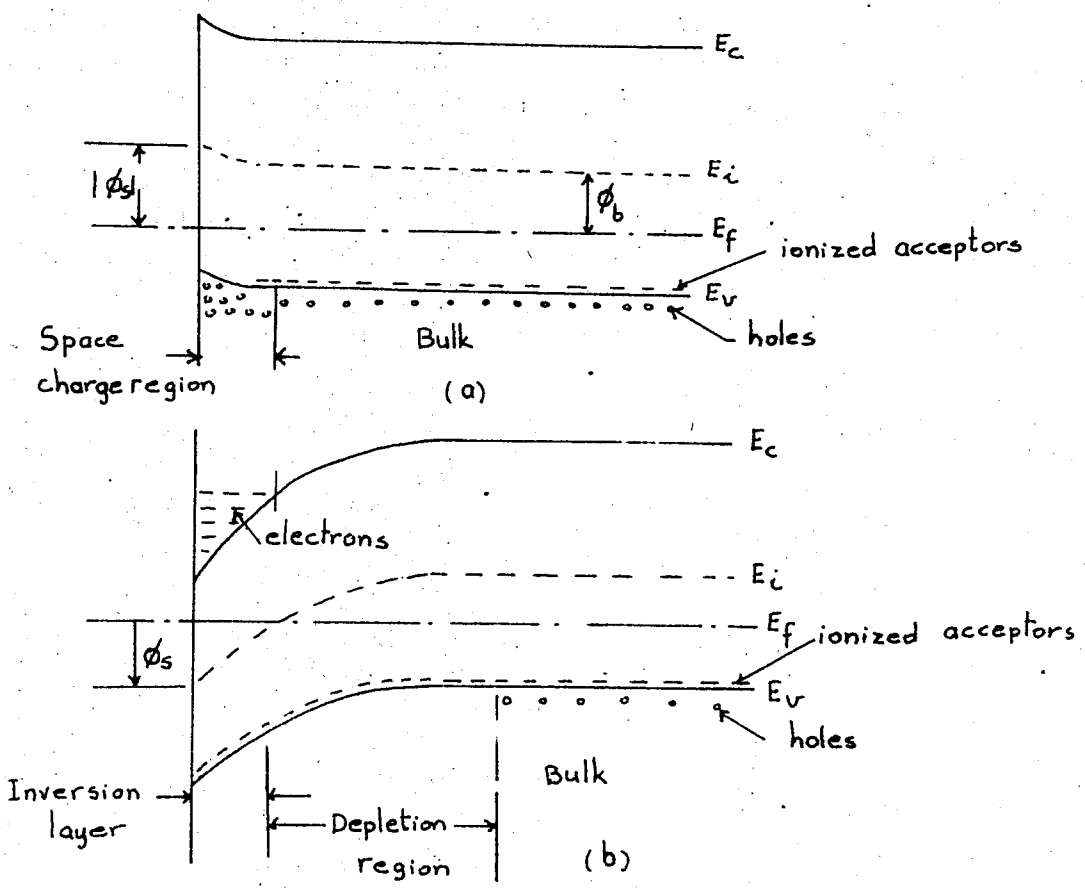


Fig. 4-2 Energy level diagram for p-type semiconductor
a) with accumulation layer
b) with inversion layer and depletion region.

ϕ_b = the electrostatic potential in the bulk of the material far from surface

ϕ_s = electrostatic potential at the surface.

Using these parameters, the bulk carrier densities n_b , p_b , and the surface carrier densities n_s , p_s can be expressed as follows :

$$n_b = n_i \exp\left[q\phi_b/kT\right]$$
$$p_b = n_i \exp\left[-q\phi_b/kT\right] \qquad (4\text{-}5)$$
$$n_s = n_i \exp\left[q\phi_s/kT\right] = n_b \exp\left[q(\phi_s-\phi_b)/kT\right]$$
$$p_s = n_i \exp\left[-q\phi_s/kT\right] = p_b \exp\left[-q(\phi_s-\phi_b)/kT\right]$$

n_b , p_b , and ϕ_b are determined by the doping of the material. The potential at any point in the bulk of the semiconductor is given by

$$V = \phi - \phi_b \qquad (4\text{-}6)$$

and the surface barrier potential is defined by :

$$V_s = \phi_s - \phi_b \qquad (4\text{-}7)$$

The value and polarity of the surface potential ϕ_s and the surface barrier , V_s , are used to classify the surface conditions, such that

If $\phi_s > 0$, the surface is n -type (4-8)

$\phi_s < 0$, the surface is p -type

Surface conditions can be classified as follows:

a) When an enhancement of majority carriers exists at the surface , the surface is referred to as an enhancement or accumulation layer. This condition exists when $V_s<0$ for p-type and $V_s > 0$ for n-type material.

b) Flat band condition exists when $V_s = 0$. In such a case , the energy band is flat out to the surface.

c) A depletion layer is formed when some of the majority carriers are repelled from the surface , leaving donor or acceptor ions uncompensated. This occurs when the sign of V_s is opposite to that of ϕ_b and $|V_s| < |2\,\phi_b|$.

d) An inversion layer is obtained when the minority carrier density at the surface equals or exceeds the majority carrier bulk density. This occurs when the sign of V_g is opposite to that of

ϕ_b and $|V_s| \geq |2\phi_b|$. Fig 4-2 illustrates these surface layers.

4-1B. Surface Conductance

The conductivity of a semiconductor is given by

$$\sigma(z) = q\left[n(z)\mu_n(z) + p(z)\mu_p(z)\right] \qquad (4-9)$$

where $\mu_n(z), \mu_p(z)$ = electron and hole mobilities respectively. $n(z)$, $p(z)$ = free electron and hole concentrations at thermal equilibrium. Deep inside the bulk, the mobilities and carrier densities are constant so that Eq.(4-9) becomes

$$\sigma_b = q\left(n_b\mu_b + p_b\mu_b\right) \qquad (4-10)$$

But at the surface, since usually $V_s \neq 0$, a large positive or negative surface barrier (a potential well) exists for one or the other type of carrier resulting in a reduced surface mobility If we define surface excesses as the additional number of free carriers per unit area of surface due to the presence of the surface barrier, we can express these excesses as :

$$\Delta N = \int_{\infty}^{0} (n-n_b)\,dz = -\int_{0}^{V_s} \frac{n-n_b}{E}\,dV \qquad (4-11)$$

and

$$\Delta P = \int_{\infty}^{0} (p-p_b)\,dz = -\int_{0}^{V_s} \frac{p-p_b}{E}\,dV \qquad (4-12)$$

Then the surface conductance g_s due to this change becomes

$$g_s = q\left(\mu_p \Delta P + \mu_n \Delta N\right) \quad \text{mhos} \qquad (4-13)$$

Surfaces with accumulation or inversion layers have high surface conductance due to large number of majority carriers respectively but for depletion layer surfaces conductance is smaller and passes through a minimum value $g_{s\,min}$ when very few excess free carriers are present in the space charge region. The surface barrier height at which this minimum occurs can be derived as follows . From Eq.(4-13) the surface conductance becomes zero when

$$\Delta N = \Delta p = 0 \qquad (4-14)$$

and when

$$\Delta p = -\left(\frac{\mu_n}{\mu_p}\right)\Delta N \qquad (4-15)$$

Differentiating (4- 15) with respect to u_s , where $u_s = q\phi/kT$,

$$\frac{\partial \Delta p}{\partial u_s} = -\frac{\mu_n}{\mu_p} \frac{\partial \Delta N}{\partial u_s} \qquad (4\text{-}16)$$

and referring back to Eq. (4-11) and (4-12)

$$v_{smin} = \frac{qV_{smin}}{kT} \cong -2u_b - \ln \frac{\mu_n}{\mu_p} \quad \text{for} \quad v_s \gg 0 \qquad (4\text{-}17)$$

Surface conductance versus v_s for intrinsic , p-type and n-type materials is shown in Fig 4-3 .

4-2 Semiconductor Surfaces in the Presence of Surface States.

Surface states are localized allowed energy states some of which may be in the forbidden gap . The distribution and properties of these bound states are a function of the chemical and mechanical treatment of the surface and the environment to which it is exposed. There are two kinds of surface states:

4-2A. Slow surface States :

They exchange charge very slowly with the bulk material. The time constants of charge exchange range from a few seconds to months. (These states are associated with the oxide films on Ge or Si surfaces). The density of these states depends on the adsorption equilibrium with the ambient and is not a constant of the surface except in ultrahigh vacuum.

4-2B. Fast surface states .

They exchange charge with the bulk material with time constants ranging from millisecond to microseconds or less. They exist near the interface between the semiconductor and the oxide and are relatively unaffected by the atmospheric ambient.These states are mainly responsible for the generation and recombination of electrons and holes at the surface.

4-2C. Probability of occupation of surface states.

Probability of occupation of a surface state at energy E_t is given by the Fermi-Dirac statistics as :

$$f(E_t) = \frac{1}{1 + g \exp\left[(E_t - E_f)/kT\right]} \qquad (4\text{-}18)$$

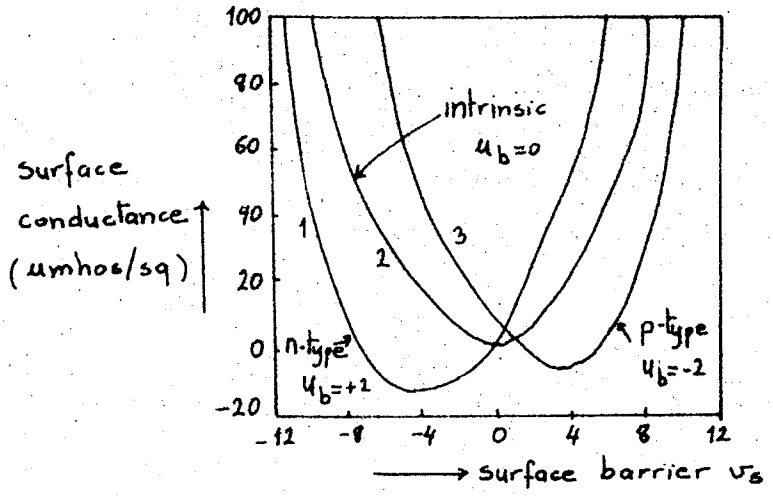


Fig. 4-3 Surface conductance vs v_s for various bulk resistivities
n-type ---- 15 ohm-cm (curve 1)
p-type ---- 10 ohm-cm (curve 3)

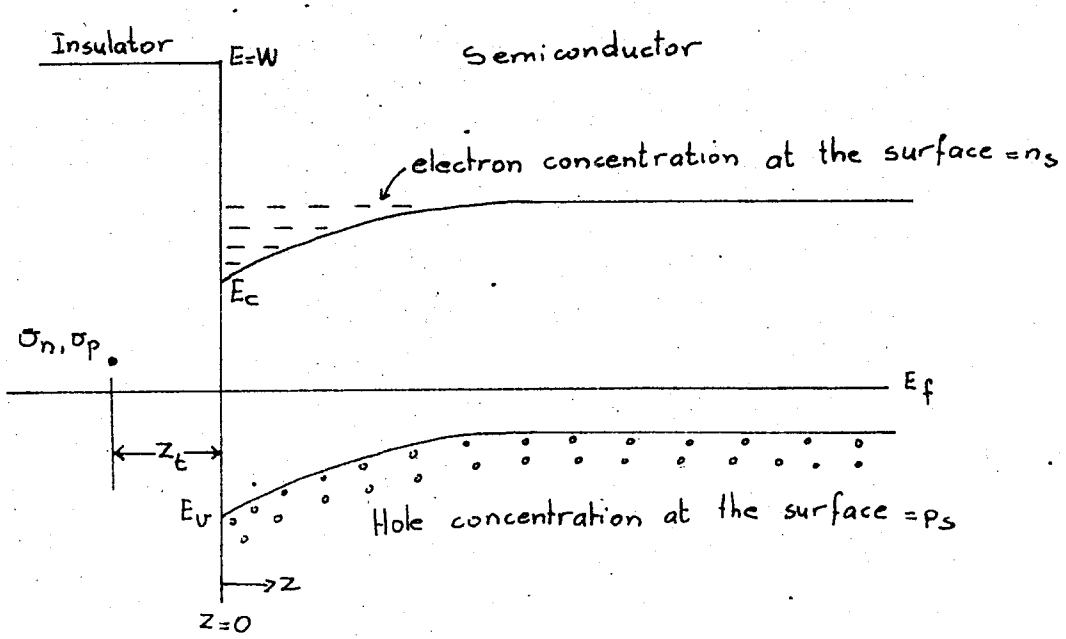


Fig. 4-4 Trapping site imbedded in the insulator

$g = 1/2$ for acceptors, and $g = 2$ for donors . The constant g takes care of spin degeneracy.

At equilibrium, the surface potential will adjust itself in such a way that overall charge neutrality exists while the surface traps are filled in accordance with Eq. (4-18) , that is the surface potential is such that it produces the necessary charge in the space charge layer to neutralize the charge in surfaces states :

$$Q_{ss} + Q_{sc} = 0 \qquad (4\text{-}19)$$

Surface states are neutral when not not occupied.

4-2D Surface State Capacitance

The differential capacitance C_{ss} associated with change of charge in surface states is defined by :

$$C_{ss} = \frac{\partial Q_{ss}}{\partial V_s} \qquad (4\text{-}20)$$

Thus the total surface capacitance is given by

$$C_s = \frac{\partial Q_I}{\partial V_s} = C_{sc} + C_{ss} \qquad (4\text{-}21)$$

Thus the semiconductor surface capacitance is a parallel combination of space charge and surface-state-capacitance.

4-3. Semiconductor-Insulator Interface

4-3A. MOS capacitor.

The MOS capacitor consists of an n-type semiconductor covered by a thin film ($50\text{Å} < T_{ox} < 5000\text{Å}$) of insulating material with a metal electrode on top of the insulator. The assumptions governing the following discussion will be that the film is perfectly insulating and is characterized by a certain trap distribution . Experiments have shown that all ranges of response time can be expected for surface states in such a system . But fast surface states can be reduced by controlling the conditions under which the insulating film is produced by oxidation of the semiconductor.

Fig 4-5a and 4-5b show a MOS capacitor and its band picture when an external voltage V is applied. The picture is drawn for an n-type semiconductor, but the general voltage equation can be written as :

$$V = \phi_m + V_{ox} - X + V_s - V_c \qquad (4\text{-}22)$$

where
- V = total applied voltage across the MOS capacitor
- ϕ_m = work function of the metal
- V_{ox} = voltage across the oxide film
- X = electron affinity of the semiconductor
- V_c = electostatic potential difference between Fermi level and conduction band.

For a certain given unit V_c , X , and ϕ_m will be constants so that the effective applied voltage will be

$$V = V - \phi_m + X + V_c = V_{ox} + V_s \qquad (4\text{-}23)$$

A further assumption is that the voltage drop across the oxide layer due to the distributed surface states in the oxide is small compared to the overall oxide voltage. Thus the electric field in the oxide may be considered uniform. But this uniform field is valid only when

1) The space charge regions at the metal-SiO_2 interfaces are widely separated , that is

$$\frac{T_{ox}}{\text{Total space charge regions}} \gg 1 \qquad (4\text{-}24)$$

2) When $$\frac{T_{ox}}{\text{total space charge regions}} \ll 1 \qquad (4\text{-}25)$$

In the second case, the two space-charge regions overlap and the intergrated space-charge in the oxide can be neglected with respect to the surface charge either in the metal or in the semiconductor.

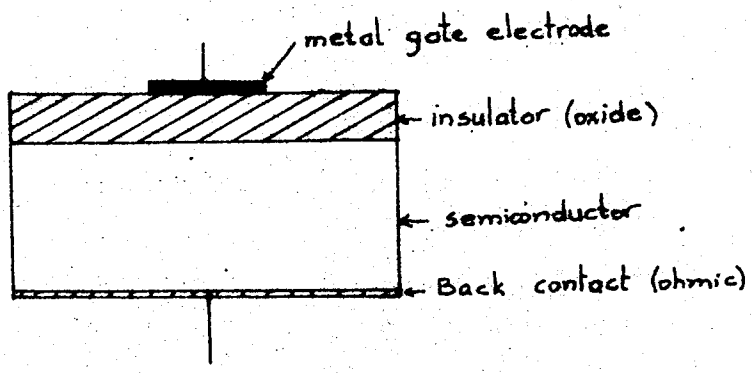


Fig. 4-5a Schematic representation of an MOS capacitor

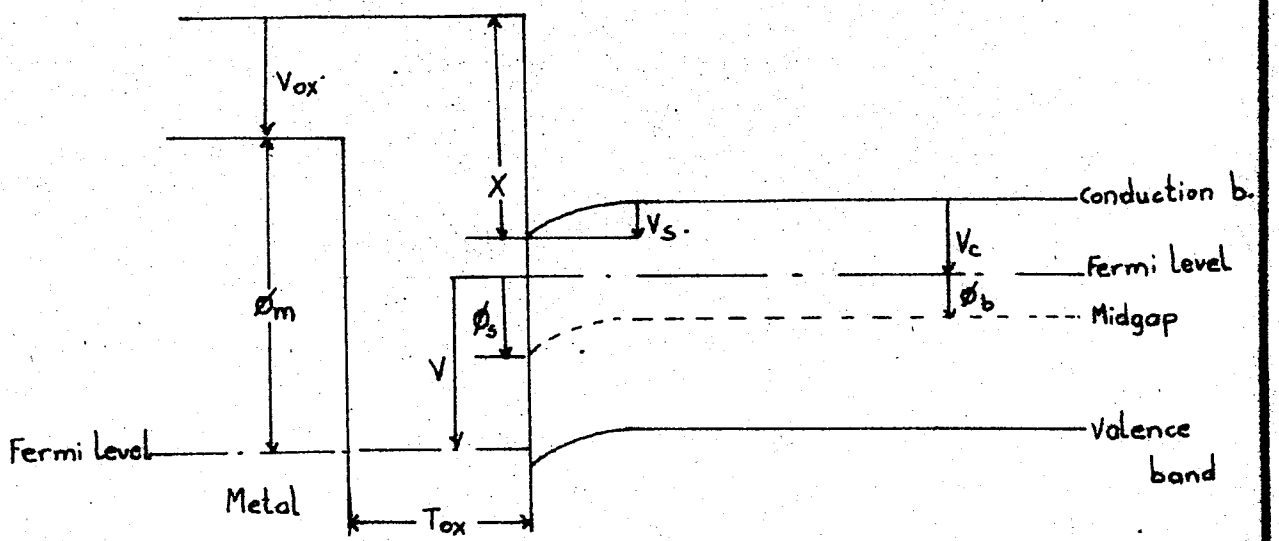


Fig. 4-5b Band picture of MOS capacitor with an applied voltage V_a . N-type semiconductor is used and an accumulation layer is formed.

In the intermediate case, when

$$\frac{T_{ox}}{\text{Total space charge regions}} \cong 1$$

, The assumption of uniform field is not valid.

4-3B. MOS Capacitance

When a voltage is applied across the device, assuming uniform field in the oxide layer , the electric field terminates on two kinds of charges: The charge in the semiconductor space charge region, Q_{sc} , and the charge in the surface states, Q_{ss} . Both Q_{ss} and Q_{sc} are of same sign and add to give

$$Q_T = Q_{sc} + Q_{ss} \qquad (4\text{-}26)$$

The charge Q_m on the metal plate is equal and opposite in sign to Q_T. Then

$$|Q_m| = Q_T = C_{ox} V_{ox} \qquad (4\text{-}27)$$

where $C_{ox} = \epsilon_{ox}/T_{ox}$ oxide capacitance per unit area
If we call total surface capacitance of the semiconductor C_s , the total capacitance per unit area of the MOS structure will be

$$C = \frac{C_{ox} C_s}{C_{ox} + C_s} \qquad (4\text{-}28)$$

Thus the total capacitance corresponds to a series combination of the oxide and the surface capacitance.
The oxide capacitance is constant and frequency independent . Therefore the frequency dependence of the MOS capacitance results from the frequency dependence of the semiconductor surface capacitance C_s only .
The complete functional form of the overall capacitance and its frequency dependence can be analyzed as follows:
From Eq. 4-23, the effective voltage across the device was given as:

$$V = V_{ox} + V_s = (V_{ss} + V_{sc}) + V_s \qquad (4\text{-}29)$$

and the total capacitance per unit area of the MOS is:

$$C = \frac{\partial Q_T}{\partial V} \qquad (4\text{-}30)$$

Making use of Eqns. (4-27) and (4-29) , Eq.(4-30) can be written as

$$C = C_{ox}\left(1 - \frac{\partial V_s}{\partial V} \right) \qquad (4\text{-}31)$$

Now the dependence of the capacitance C on the surface barrier can be obtained by making use of Eq.(4-29) and (4-31)

$$C(V_s) = C_{ox}\left[\frac{(\partial V_{sc}/\partial V_s) + (\partial V_{ss}/\partial V_s)}{1 + (\partial V_{sc}/\partial V_s) + (\partial V_{ss}/\partial V_s)}\right] \qquad (4\text{-}32)$$

If the frequency of the ac signal is so high that none of the surface states can follow (i.e., $w \gg 1/\tau_c$) , $\partial V_{ss}/\partial V_s \rightarrow 0$. In such a case , the MOS capacitance reduces to

$$C = C_{ox}\left[\frac{(\partial V_{sc}/\partial V_s)}{1 + (\partial V_{sc}/\partial V_s)}\right] \qquad (4\text{-}33)$$

Normalized MOS capacitance versus gate voltage for no surface states is shown in Fig 4-6 and 4-7. In Fig 4-7 high frequency and low-frequency capacity curves are presented for various SiO_2 thickness.

4-3C. Surface Conductivity

The equation for the conductance of the semiconductor surface was given by

$$g_s = q(\mu_p \Delta P + \mu_n \Delta N) \qquad (4\text{-}34)$$

Since ΔP, and ΔN were shown to be functions of the electric field at the surface to modulate this conductance , the electric field at the surface must be changed. When a voltage is applied, the surface potential adjusts itself according to the equation

$$V = V_{sc} + V_{ss} + V_s \qquad (4\text{-}35)$$

in order to maintain charge neutrality. When the surface potential changes, two things happen :

1) The Fermi level at the surface is altered so that the occupancy of the surface states which was given by the Fermi -Dirac

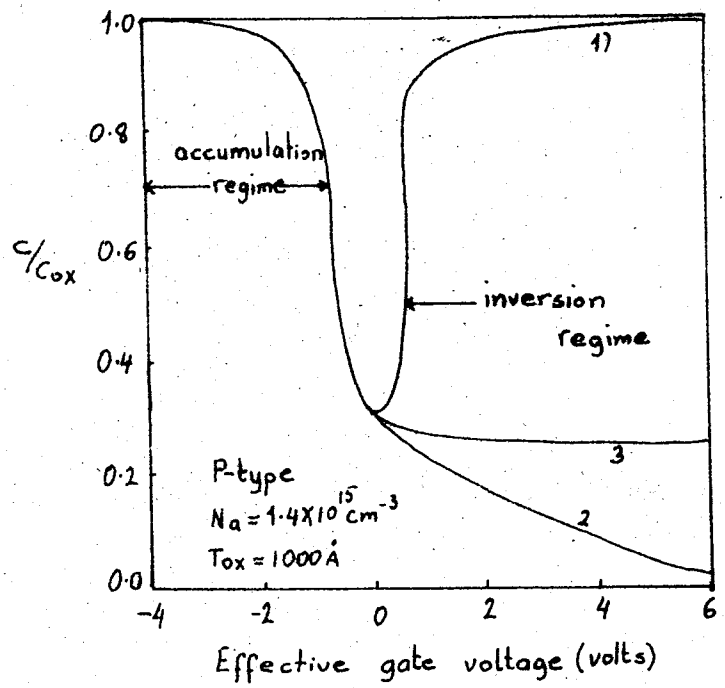


Fig. 4-6 Normalized MOS capacitance vs gate voltage for no surface states. In the inversion regime curve : 1-minority carriers follow ac and dc voltages. 2- minority carriers cannot accumulate at surface. 3- minority carriers follow dc but not ac signals

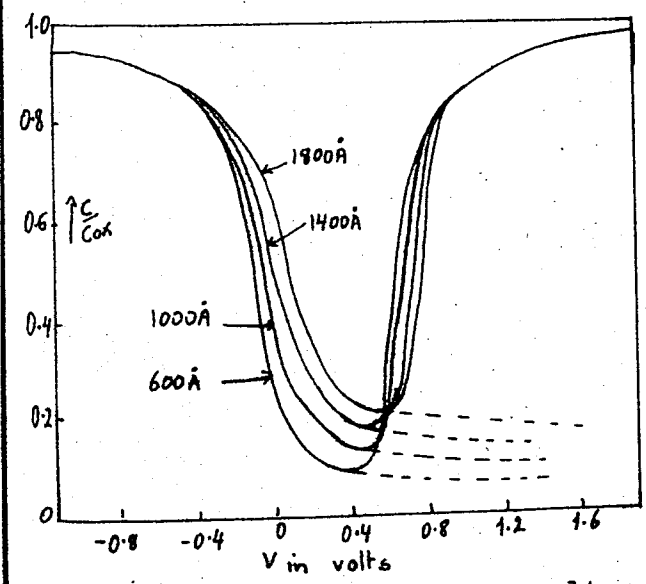


Fig.4-7a MOS capacity vs voltage. Oxide thickness 600-1800 Å . (Na 1.0 x 10^{14} cm^{-3})

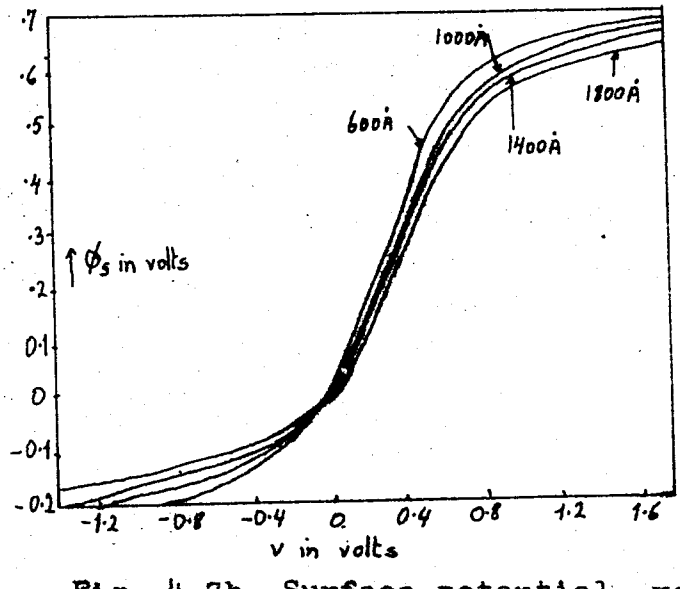


Fig. 4-7b Surface potential vs voltage. Oxide thickness 600-1800Å (Na 1.0 x 10^{14} cm^{-3})

statistics (Eq.4-18) is changed.

2) The surface excesses of holes and electrons (ΔP, ΔN) are changed resulting in a change in the surface conductance. This is called the field effect.

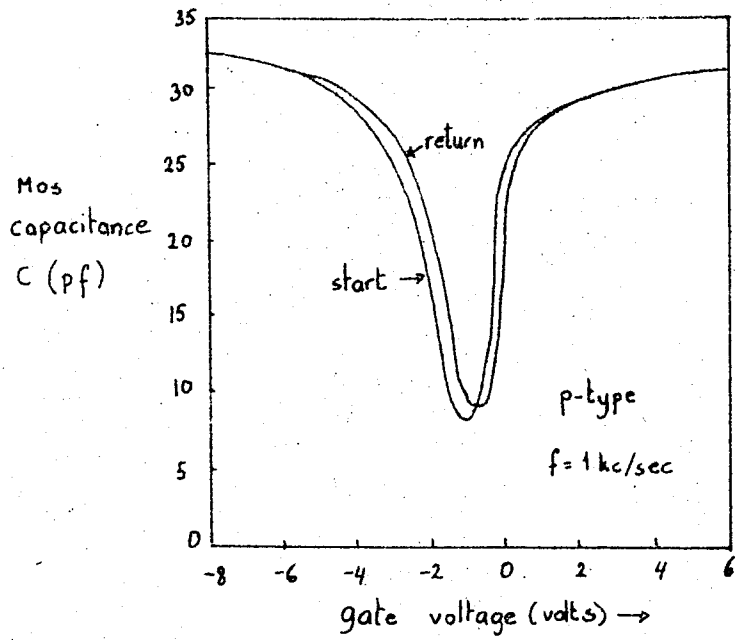
4-3D Hysteresis due to Carrier Trapping.

Measurements made for MOS capacitance show hysteresis effects of two kinds:

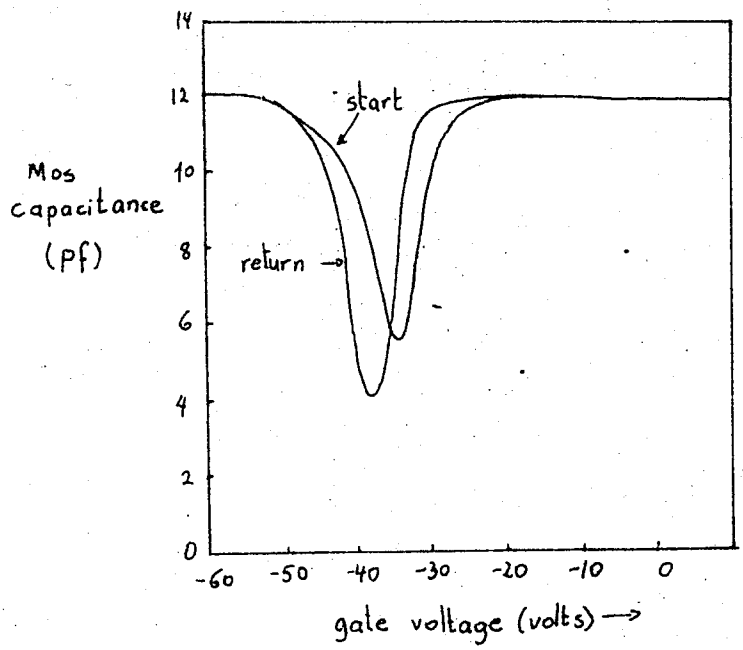
1) Hysteresis effect dependent on ambient and time. This is observed only for p-type units under inversion layer operation conditions. It is due to the charge migration on the surface of the oxide layer. No change of capacitance minimum is observed in this case.

2) Hysteresis due to trapping effects at the two interfaces. This effect is endependent of ambient and causes a shift of the C vs V characteristics to the left or to the right. This effect is due to trapping of free carriers at the two interfaces. For example, for p-type materials which usually have an inversion layer and therefore an excess of n_s at the surface, there will be a negative charge trapped in the oxide (at low negative bias). The trapped charge will be a part of V_{ss}. But for the same applied voltage V, V_{ss} will be different for increasing and decreasing bias. Since $V_{sc} + V_s = V - V_{ss}$, this requires that $V_{sc} + V_s$ also be different. Therefore the surface capacitance C_s will also be different.

When surface state concentration is greater in the oxide-semiconductor interface, the C vs V curve for increasing bias lies to the left of that for decreasing bias, but if the surface state concentration is greater at the metal-oxide interface then the hysteresis may be in the opposite direction. Fig 4-8a and 4-8b show MOS capacitance versus gate voltage plots. Trapping at the oxide-semiconductor interface is dominant in Fig 4-8a while trapping at the metal oxide interface is dominant in 4-8b..



(a)



(b)

Fig. 4-8 MOS capacitance vs Gate electrode voltage.
a) showing hysteresis because of trapping at the oxide-semiconductor interface.
b) showing hysteresis because of trapping at the metal-oxide interface.

4-4. MOS Characteristics and Equivalent Circuit.

Fig 4-9 shows typical differential capacitance and equivalent parallel conductance vs field plate bias curves measured at 100 kc with an admittance bridge. The capacitance curve has three distinct regions , two of high nearly constant capacitance and a third where the capacitance changes very rapidly with bias and goes through a minimum. Behavior at high negative bias (left of the minimum) and the decrease toward the minimum can be explained as follows: At high negative bias, a heavy accumulation layer is formed in the Si surface under the field plate. Because this layer has now a high charge density only the oxide capacitance is measured. (See Eq.4-28). When bias is reduced, charge in the accumulation layer decreases causing a decrease in the differential capacitance measured. With further bias reduction, the accumulation layer turns into a depletion layer. The capacitance of this layer now appears in series with the oxide capacitance reducing overall capacitance. This range may be influenced by the capacitance associated with fast surface states which may also cause the corresponding conductance peak.

The simplified equivalent circuit of the MOS capacitor in the inversion bias range is shown in Fig. 4-10 . The minority carrier sources for the inversion regime are represented by resistances. R_d is associated with electron diffusion current. R_{gd} is associated with volume-generated current within the depletion region . R_{gs} is associated with a surface generated current related to the surface states at the insulator-semiconductor interface. With each of these sources a time constant is associated so that the response time for the inversion layer will be

$$1/\tau = \left[1/\tau_d + 1/\tau_{gs} + 1/\tau_{gd}\right] \qquad (4\text{-}36)$$

or

$$1/\tau = \left[1/R_d + 1/R_{gs} + 1/R_{gd}\right]\frac{1}{C_d}$$

where C_d = capacitance of the depletion region.

The capacitance minimum occurs when the net field (the difference between applied and surface state fields which are in opposite direction when the field plate is negative) causes the channel under the field plate to disappear or pinch off. At this moment the applied field divides between the oxide and space charge region in the silicon.

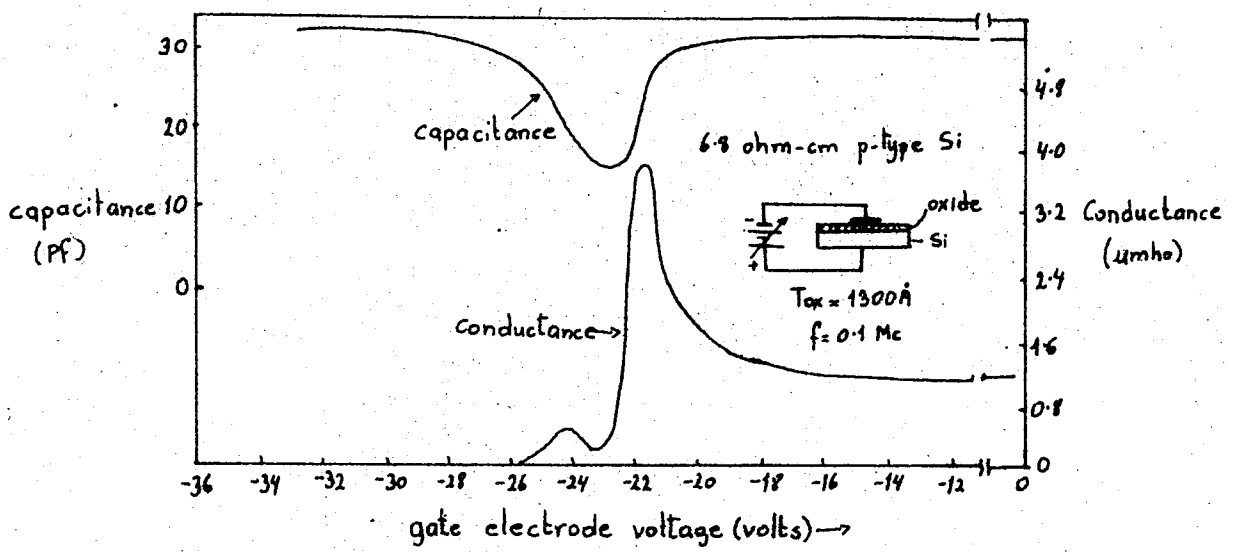


Fig.4-9 Capacitance and equivalent parallel conductance as a function of gate electrode voltage for a MOS structure shown in the insert.

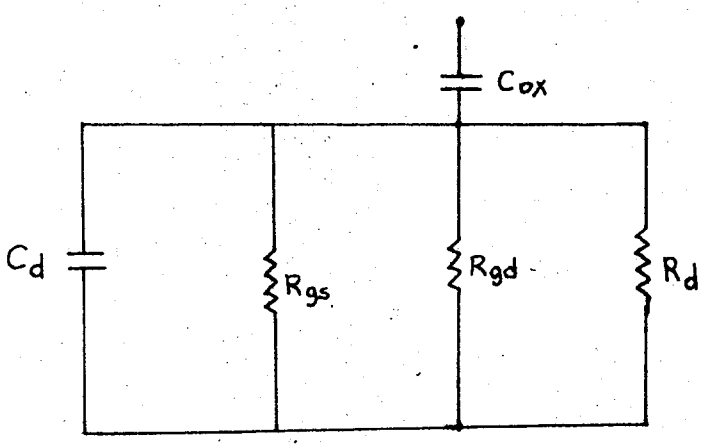


Fig. 4-10 The ac equivalent circuit for the MOS structure which operates in the inversion layer. It includes the effects of surface, bulk, and junction-generated currents

CHAPTER 5

INSULATORS IN MOS STRUCTURES

In the MOS transistor , the gate is seperated from the conducting channel by an insulating layer. An ideal gate insulator should have zero conductance and very small capacitance to minimize dissipation and storage effects in the gate circuit. Real insulators however show conduction at considerably lower applied fields than the ideal insulators. The conduction characteristic of real insulators may be linear or nonlinear. It may be uniquely defined, show hysteresis effects, lead to permanent changes or destructive break down. These effects may be due to electronic or ionic conduction or to surface currents which are greatly affected by impurities and defects.

5-1 Electronic Conduction

5-1A Injection over a Barrier: Schottky Emission

Fig 5-1 shows a simple schematic energy band diagram of the insulator when it is placed between the metal gate and the semiconductor. ϕ_o represents the work function for the insulator when no external voltage is applied. The gradients in ϕ are due to differences in work functions between the different materials. As can be seen from Fig 5-1 , even in the absence of an external voltage , a small built in electric field may exist in the insulator. If we designate this by E_{int} ,

$$E_{int} = \frac{\partial \phi_o}{\partial z} \qquad (5\text{-}1)$$

If a voltage is applied to the insulator the resulting field in the insulator will become

$$E = E_{int} + E_{ext} = \frac{\partial \phi_o}{\partial z} + \frac{\partial E_f}{\partial z} = \frac{\partial \phi}{\partial z} \qquad (5\text{-}2)$$

This situation is shown in Fig 5-2 . The voltage drop across the semiconductor is always less than E_g/q, where E_g is the band gap of the semiconductor. Therefore we can consider that

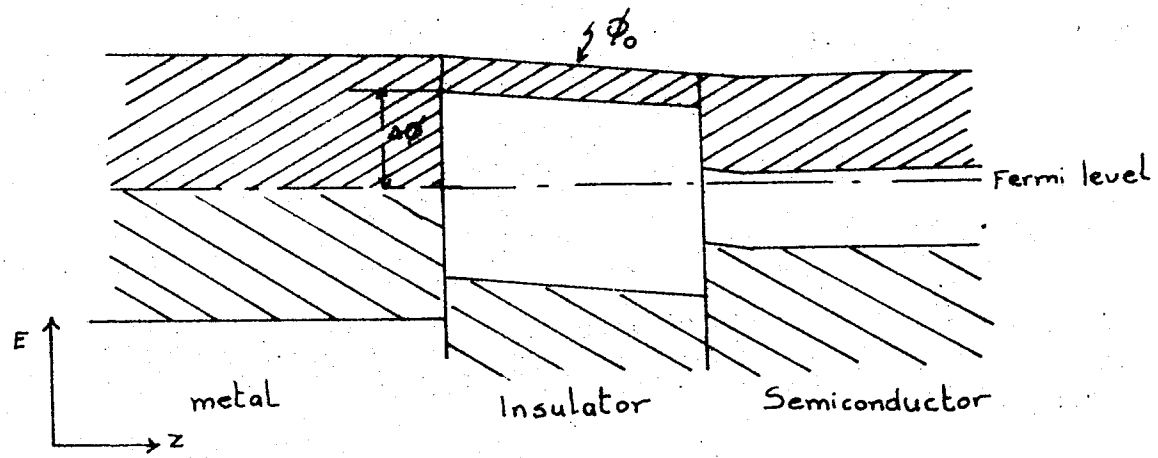


Fig.5-1. One dimensional energy-band diagram of MOS transistor at equilibrium.

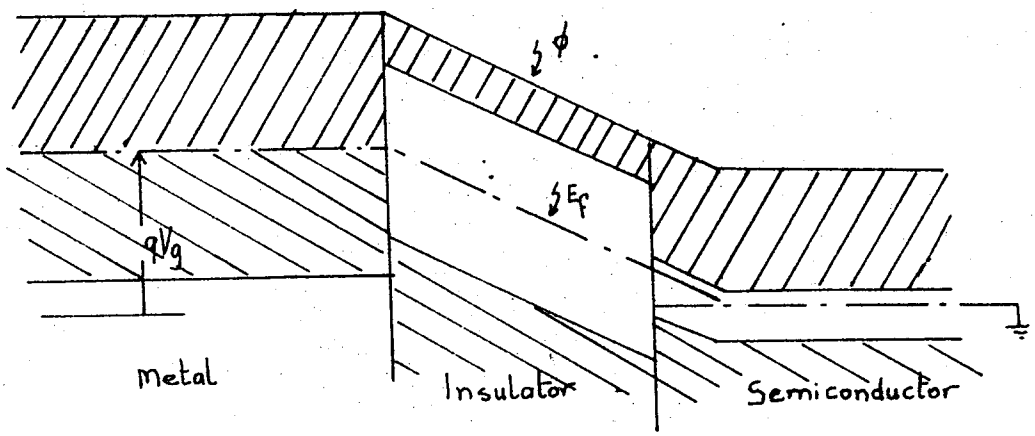


Fig. 5-2. Energy band-diagram of MOS transistor under applied gate voltege V_g.

all of the applied gate voltage V_g appears across the insulator, and approximate the insulator field E by

$$E \cong \frac{V_g}{T_{ox}} \qquad (5\text{-}3)$$

where V_g = gate electrode voltage
T_{ox} = insulator thickness

In Fig 5-2 , the field across the insulator is shown as constant. This is a simplified situation which assumes that there is no net charge density present. However real situations are more complex and will be discussed as we go on .

When the potential barrier $\Delta\phi$ between the metal and the insulator is small, or when the temperature is high , some electrons with sufficient energies will pass over the barrier and flow into the insulator conduction band. At equilibrium an equal number of electrons will flow from the insulator to the metal, but this may be eliminated by applying a field. The current flowing in the z-direction (into the insulator), due to thermionic emission is obtained by intergrating the charge flow over all the electrons in the metal with sufficient momentum in the z-direction and

$$J_z = \frac{1}{(2\pi)^2} \frac{2q m_{eff} (kT)^2}{h^3} \exp\left(-\frac{\Delta\phi}{kT}\right) \qquad (5\text{-}4)$$

Eq(5-4) is known as the Richardson equation. This thermionic current is small for the type of insulators and the range of temperatures used in MOS transistors. The currents may be increased by high fields which modify the shape of the barrier and lower its heigth. This causes more thermionic carriers to flow over the barrier. This effect was first described by Schottky.

To calculate the current through the insulator when a field is applied , we must modify the metal-insulator interface as seen in Fig (5-3) to include the image force. The image force results from the positive charge induced in the metal by the emitted electron. This force tends to attract the electron back into the metal and is given by

$$F = qE_{im} = \frac{q^2}{4\pi\epsilon_i (2z)^2} \qquad (5\text{-}5)$$

where z = distance of the electron in the insulator from the metal interface.

e_i = permittivity of the insulator

Correponding to the image force, there is a potential given by $q^2/16\pi\epsilon_i z$ so that the net potential barrier is reduced to

$$\phi = \phi_o - \frac{q^2}{16\pi\epsilon_i z} \qquad (5\text{-}6)$$

The resulting potential is shown in Fig 5-3 . If now an electrical potential E_o is applied to the emitting surface, then the potential energy barrier will be

$$\phi = \phi_o - \frac{q^2}{16\pi\epsilon_i z} + qE_o z \qquad (5\text{-}7)$$

where E_o = the external electrcal field. The barrier will have a maximum at

$$z_m = \sqrt{\frac{q}{16\pi\epsilon_i E_o}} \qquad (5\text{-}8)$$

with a maximum ϕ given by

$$\phi = \phi_o - \frac{q}{2}\sqrt{\frac{qE_o}{\pi\epsilon_i}} \qquad (5\text{-}9)$$

so that the lowering of the barrier height by an applied field E_o is given by

$$\Delta\phi_{max} = \sqrt{\frac{q^3 E_o}{4\pi\epsilon_i}} \qquad (5\text{-}10)$$

Inserting Eq (5-10) into Richardson's Equation (Eq.5-4), we get the modified Schottky equation which gives us the current flowing into the insulator :

$$J_z = \frac{1}{(2\pi)^2}\,\frac{2qm_{eff}(kT)^2}{h^3}\,\exp\left[-\frac{\Delta\phi - (q^3E_o/4\pi\epsilon_i)^{1/2}}{kT}\right] \qquad (5\text{-}11)$$

5-1B Tunneling or Field Emission

The mechanism of field effect emission is different from thermionic emission. In the latter case the electrons must possess sufficient energy to pass over the surface barrier. With field effect emission however the electron can tunnel through the barrier

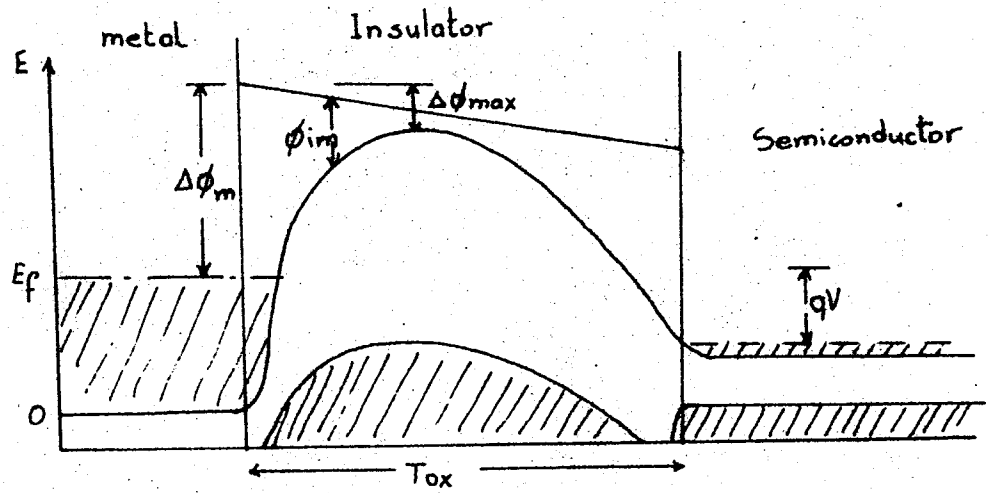


Fig.5-3. Band diagram of the MOS transistor with the metal-insulator interface modified by the image force(ϕ_{im}), and lowered barrier height due to an applied field E.

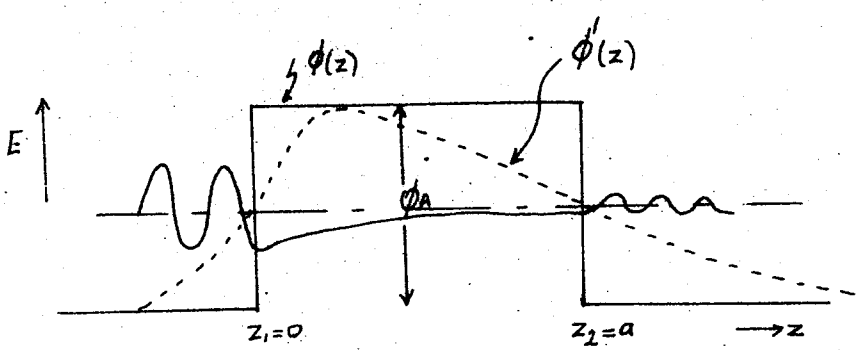


Fig.5-4. One dimensional tunneling through a potential barrier $\phi(z)$ or through $\phi'(z)$.
$\phi(z)$ is a rectangular barrier
$\phi'(z)$ is a smoothed-out (more realistic) barrier.

To find an expression for this probability, the one-diminsional Schrödinger equation given by Eq(5-12) must be solved :

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dz^2} + \phi(z)\,\psi = E\,\psi \qquad (5\text{-}12)$$

Eq.(5-12) can be solved exactly for a square potential barrier giving

$$\begin{aligned} \psi_I &= A_1 e^{ik_1 z} + A_2 e^{-ik_1 z} && z<0 \\ \psi_{II} &= B_1 e^{k_2 z} + B_2 e^{-k_2 z} && 0<z<a \\ \psi_{III} &= C_1 e^{ik_1 z} && z>a \end{aligned} \qquad (5\text{-}13)$$

where

$$k_1 = \sqrt{\frac{2mE}{\hbar^2}} \qquad k_2 = \sqrt{\frac{2m(\phi_0 - E)}{\hbar^2}}$$

The coefficients are determined by applying the continuity condition on the wave functions given in Eq.(5-13) and their first derivatives

at $z=0$ and $z=a$:

$$\begin{aligned} A_2 + A_1 &= B_2 + B_1 \\ ik_1(A_1 - A_2) &= k_2(B_1 - B_2) \\ B_1 e^{k_2 a} + B_2 e^{-k_2 a} &= C_1 e^{ik_1 a} \\ k_2(B_1 e^{ak_2} - B_2 e^{-k_2 a}) &= i C_1 k_1 e^{ik_1 a} \end{aligned} \qquad (5\text{-}14)$$

let $u = k_1/k_2$ and solve Eq (5-14) for C_1

$$C_1 = \frac{e^{-ik_1 a}\, 2u}{2u\cosh k_2 a - i(1-u^2)\sinh^2 k_2 a} \qquad (5\text{-}15)$$

and the transmission coefficient is found to be

$$|T| = \frac{4u^2}{(1-u^2)\sinh^2 k_2 a + 4u^2 \cosh^2 k_2 a} \qquad (5\text{-}16)$$

It is apparent from this result that transmission or tunneling through the insulator barrier can occur even though the energy E of the electron in the metal is less than the barrier height. However , the transmission through the barrier decreases very rapidly with increasing $k_2 a$, i.e., with increasing barrier height and thickness. But a large electric field narrows the

barrier and permits electrons to tunnel through.

Real barriers are never infinitely sharp, but rather look like the dotted $\phi'(z)$ function shown in Fig 5-4. The solution of such a problem is complicated. But if $\phi'(z)$ varies smoothly (which is usually the case) the WKB approximation may be applied. The transmission probability for such a curved barrier is given by WKB method as :

$$T(E) = \exp\left\{-\int_{z_1}^{z_2} \left[2m\left(\phi'(z) - E\right)\right]^{1/2} dz\right\} \qquad (5\text{-}17)$$

The tunnel current flowing through the insulator is now calculated by intergrating over all electrons

$$J_z = q \int N(E) f(E)\, v_z\, T(E_z)\, dE_z \qquad (5\text{-}18)$$

where E = Kinetic energy of the electron in the cathode(metal) material, measured from the bottom of the conduction band.

$N(E)$ = the density of states corresponding to an energy in the interval E to $E + \Delta E$.

$$f(E) = \left\{1 + \exp\left[\left(E_x^2 + E_y^2 + E_z^2\right) - E_f\right)/kT\right]\right\}^{-1}$$

There have been many attempts to evaluate Eq.(5-18). Some have obtained answers by making approximations and simplifying assumptions which are valid only within certain limits.

5-1C . Conduction in High Electric Fields

The electric fields applied to the insulating layer in the transistor are large . When the field is very large, electrons attain high average energies and give rise to optical phonon scattering, impact ionization and pair production. If the insulating layer is thin, the electron will probably pass through without being trapped and if it is not larger than the mean-free-path between ionizing collisions (~ 20 Å) multiplication of carriers will not occur because the electron will leave the crystal before colliding. But if the insulator is thick and the rate of collision and creation of additional carriers exceeds the

rate of recombination , then the conductivity will increase. More current starts flowing and destructive breakdown takes place.This is known as electronic breakdown to be distinguished from the thermal breakdown which is caused by the heating of the material in the high current density regions.

5-1D Effect of Impurities and Defects

Imperfections introduce localised energy states. In the insulators, these energy states generally lie deep in the forbidden gap. The states may be donor or acceptor states. If the states are empty in equilibrium, they may trap conduction carriers or if they are electrically charged (ionized impurities), they may scatter free charge carriers and reduce their mobility.

If when an electric field is applied the density of the electrons flowing (thermionic emission) into an insulator is smaller than the density of the traps, most of the electrons will be trapped leaving only a small number of excited free carriers. As the voltage is increased , and the number of electrons injected becomes larger than the number of traps, the number of carriers left free increases. Therefore the current through an insulator which contains traps will rise slowly with the application of an increasing electric field until a certain value is reached, and then it will increase very strongly.

Traps have a strong effect on tunneling currents. They produce space charge which reduces the field near the cathode (metal - insulator junction) and thus make the barrier thicker. This results in reduced injection.

The filling and emptiying of the traps may take place by transition and absorption of radiation. High fields may increase the impurity ionization rate by lowering the barrier surrounding the trap . These are all vertical transitions in the band picture. Horizontal transitions due to traps may also occur. This happens between the contacts and impurity states. All these processes cause the current flowing in the insulator to be different from what is expected at equilibrium. The processes by

which the electrons enter and leave trap states have long time constants. Therefore thermal equilibrium will exist only under dc conditions. Under ac conditions, depending on the frequency of operation, different phenomena will make themselves felt. Hysterisis effects is an example .

5-2. Ionic Conduction

The forces holding an insulator material are ionic in character. Insulator materials with combined ionic and covalent binding may also exist . The thermodynamic equilibrium lattice is not a perfect lattice above absolute zero. It contains a certain number of vacancies and other imperfections. There is a certain probability that the lattice ions may move from one vacancy to another. The probability of this diffusion depends exponentially on the temperature.

$$D = D_o \exp\left[-\Delta E/kT\right] \qquad (5\text{-}19)$$

and by Einstein's , the low field conductivity is given as:

$$\sigma = \frac{q^2 N}{kT} D = \frac{q^2 N}{kT} D_o \exp\left[-\Delta E/kT\right] \qquad (5\text{-}20)$$

where E = activation energy = the hight of the potential barrier over which the ion must pass to move from one lattice position to the next.

N = number of mobile charge carriers (vacancies or interstitial ions)

The insulator may also be doped likee a semiconductor by means of which its ionic conductivity may be increased above its intrinsic conductivity. But even when doped , ionic conduction will be small at low fields. However , conductivity is strongly field dependent and increases at high fields which are of the order of 10^6 V/cm.

When a dc field is applied across an insulator , there will be an initial current flow, but since the ions cannot be injected or extracted from the material , negative and positive space charges start to build up near the two electrodes. The potential distribution within the insulator is thus distorted as

shown in Fig 5-5. When the external field is then removed, large internal fields remain (some of the ions flow back toward their equilibrium positions, but not all). Thus for slowly varying high fields large hysteresis effects will appear. As the ionic charges accumulate near the insulator interfaces, they affect the surface conductance of the semiconductor . They act just like the electronic surface charges discussed in the surface physics chapter. They bend the bands in the semiconductor. Thus ionic conductance may modify both the channel conductance and the transconductance of the field effect transistor. The difference between the ionic and electronic surface charges is that the ionic surface charges respond only to very low frequencies.

5-3 Surface Conduction

Electronic states in the forbidden gap due to dangling surface bonds , and adsorption impurities on the surface may give rise to surface conduction phenomena in the insulator . These effects are negligible in the actual device because the field along the surface is smaller than that across the bulk , and surface imperfections can be reduced by advanced technology.

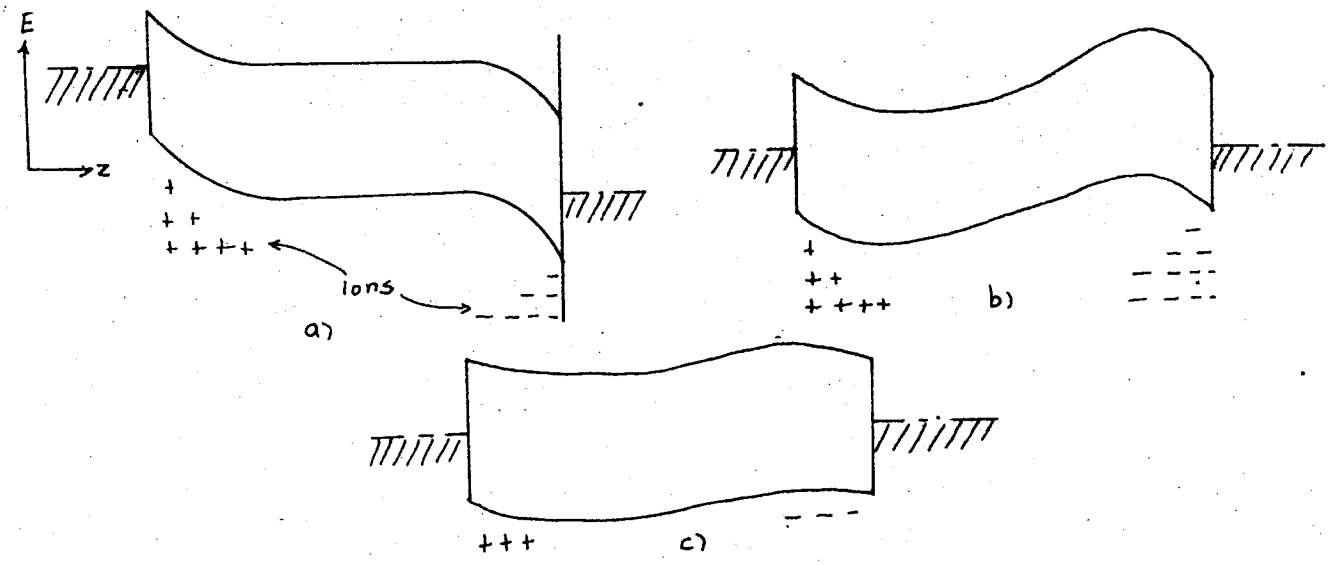


Fig.5-5 Schematic field distribution in an insulator
a) when ionic conduction has saturated
b) after removal of the field
c) after final equilibrium has been reached.

CHAPTER 6

INSULATED-GATE FIELD EFFECT TRANSİSTOR THEORY

There are two main types of insulated-gate field effect transistors. These are illustrated in Figs 6-1 and 6-2. The figures and the discussions to follow refer to electron conduction devices(n-type units)with n-type source and drain contacts,and p-type substrate. Fig. 6-1 is an induced channel type transistor. This unit is fabricated with source and drain contacts of opposite conductivity type from the channel. Back to back diodes are formed between source and drain contacts and the channel current is essentially zero for zero gate bias. If a positive voltage is applied to the gate of this unit holes will be depleted from the surface of the semiconductor . A further increases in bias will produce an accumulation of electrons at the surface. The surface channel goes from p-type through intrinsic, to an inverted n-type layer at which point ohmic conduction from source to drain commencences. Enhancement mode operation is obtained for a further increase in gate bias. Fig. 6-2 is a depletion type transistor. It is fabicated with the channel of the same conductivity type as the source and drain contacts. This may be operated in either the depletion or enhancement modes. To operate in the depletion mode, the gate is reverse biased so that carriers are depleted from the channel. Therefore maximum channel current flows in this mode of operation for zero gate bias. This type of operation is analogous to the operation of the Shockley junction FET. For enhancement mode of operation, the gate is forward biased so that carriers are drawn into the channel. Minimum current flows in this mode of operation for zero gate bias. In contrast to the Shockley unit , no gate current flows due to the insulating SiO_2 .

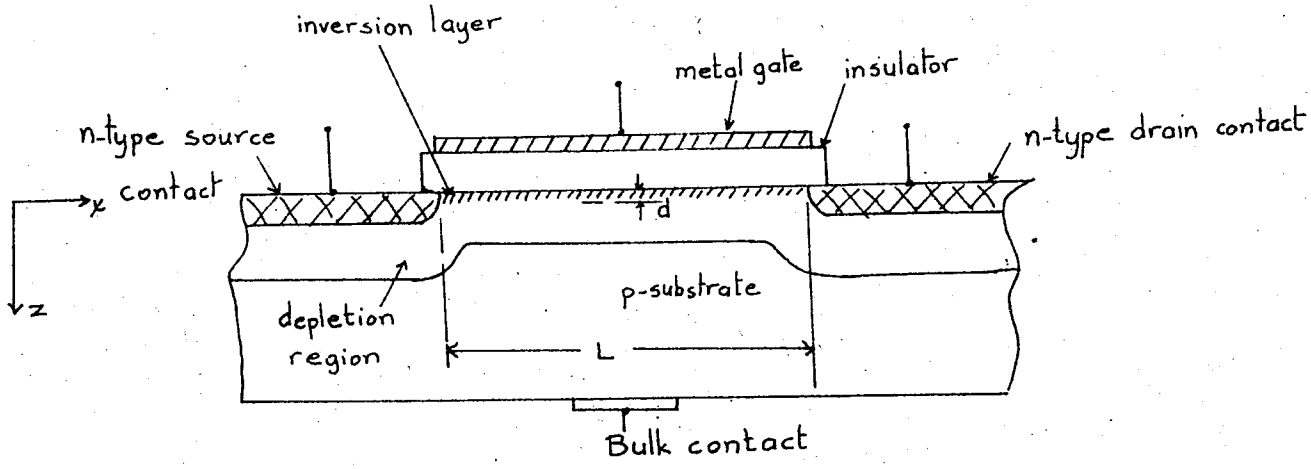


Fig 6-1. Schematic representation of an n-channel most utilizing a surface inversion layer channel.

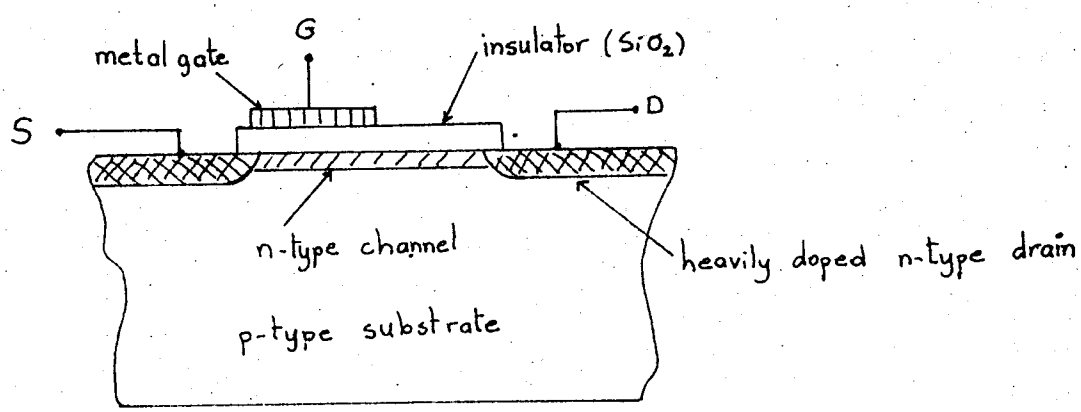


Fig 6-2. Depletion type MOS transistor with doped channel

6-1. Mobility of Charge Carriers.

To derive the important device characteristics such as current-voltage characteristics, transconductance, and frequency response one must know the drift mobility of the charge carriers in the channel region.

6-1A. Dependence on Gate Field :

In a large class of the MOS transistors, the channel consists of an inversion layer induced by a field normal to the surface of the semiconductor. There are two primary sources for this surface field.

a) Charge on the metal gate electrode due to the applied voltage.

b) Charged sites present in the oxide. These may be produced by the exposure of the oxide to hydrogen at elevated temperatures.

As the gate field is increased, the carriers in the channel move closer to the surface and their mobility will be expected to decrease due to increased surface scattering. However in many units, defects in the surface region act independently to reduce the electron mobility and this effect dominates over the surface scattering. Thus very little dependence of mobility on the gate field is observed. For this reason , in the following analysis we will assume that the carrier mobility does not depend on the magnitude of the gate field.

6-1B. Dependence on the Applied Field between Source and Drain.

For low fields less than 10^3 V/cm in n-type Si and less than 5×10^3 V/cm in p-type Si, the mobility is constant and the carrier velocity is directly proportional to the field.For moderate fields, 10^3-10^4 V/cm in n-type Si and 5×10^3-5×10^4 V/cm in p-type Si, the mobility is approximately proportional to the inverse square root of the applied field and the carrier velocity is proportional to the square root of the field. For very high fields, the carrier velocity becomes nearly constant and mobility is inversely proportional to the applied field.

6-2. Commonly used Approximations:

6-2a Gradual channel approximation. It is assumed that the rate of change of the drift field along the channel is very small compared to the rate of the change of the gate field in the channel normal to the surface. This approximation permits us to solve Poisson's equation in one dimension, leading to a channel conductivity which is a function of the applied gate field only. This approximation can be expressed as

$$\frac{dE_x}{dx} \ll \frac{dE_z}{dz} \qquad (6\text{-}1)$$

If the effective depth, d, of the channel in Fig.6-1 is very small compared to the length, l, of the channel between source and drain, then the gradual channel approximation will be valid over a substantial portion of the channel extending from the source to some point near the drain. Let this length of the source region of the channel be l_s. From the point l_s to the drain, a drain region of length l_d exists where the space charge effects cannot be neglected and the gradual channel approximation is no longer valid.

6-2B. Shallow Channel Approximation : The maximum potential drop across a surface inversion layer will not exceed half the band gap potential of the semiconductor by more than a few kT even if the inversion layer is degenerated (e.g.) o.5 volts or less for Si . If the potential drop between the gate and the channel is several volts, we can assume that most of the total potential drop is in the oxide, and that the total channel sheet charge density Q(x), at the point x , can be expressed as

$$Q(x) \cong -C_{ox}\left[V_g - V(x)\right] \quad \text{for} \quad V(x) \ll V_g \qquad (6\text{-}2)$$

where C_{ox} = capacitance per unit area across the insulator
ϵ_{ox} = dielectric constant of the oxide
V_g = the potential of the gate electrode
$V(x)$ = the potential at the point x in the channel.

The validity of this assumption depends on the substrate doping level and/or the magnitude of the applied gate field. Eq. (6-2)

assumes that the channel depth. d , is very small compared to the thickness of the oxide Tox. This approximation is referred to as the shallow channel approximation and it is sufficiently accurate for a first order analysis.

6-3. First Order Analysis of the Current Voltage Characteristics. The assumptions on which the analysis is based :

a) Uniformly doped p-type substrate with n-type contacts. The n-type substrate case may be readily extrapolated.

b) Zero surface doping (zero pinch off voltage). If the pinch-off voltage is not zero due to the surface charge which may be present in the absence of an applied potential difference , then effective gate potential V_g' must be used.

$$V_g' = V_g - V_T \qquad (6\text{-}3)$$

where V_T = off set gate potential

c) constant mobility.

d) gradual and shallow channel throughout the entire channel length

c) The substrate will be considered so lightly doped that its gating action on the channel may be neglected.

When it is assumed that the presence of depletion charge together with the effect of work function differences (between the gate metal and the oxide, and between the semiconductor and the oxide) in interface states, and of fixed charges in the oxide, results in a constant surface-charge density $C_{ox}V_T$, which is present irrespective of the potential difference $V_g - V(x)$, the mobile charge density in the channel may be given by

$$Q_m(x) = -C_{ox}\left[V_g - V_T - V(x)\right] \qquad (6\text{-}4)$$

The subscript m stands for mobile.

The conductance at location x in the channel for a channel width of W is

$$G(x) = Q_m(x)\,\mu W = \mu W\, C_{ox}\left[V_g - V_T - V(x)\right] \qquad (6\text{-}5)$$

where μ = carrier mobility.

and for the channel current from source to drain we have

$$I = G(x)\ \frac{dV(x)}{dx} \qquad (6\text{-}6)$$

When I, μ, and V_T are constants along the channel, integration of Eq. 6-6 is carried out as follows:

$$I_d = \mu W C_{ox}\left[V_g - V_T - V(x)\right]\frac{dV}{dx} \qquad (6\text{-}7)$$

$$I_d \int_0^L dx = \mu W C_{ox} \int_{V_s'}^{V_d'} \left[V_g - V_T - V(x)\right] dV \qquad (6\text{-}8)$$

$$I_d = \frac{C_{ox}\mu W}{L}\left[(V_g - V_T)(V_d' - V_s') - 1/2(V_d'^2 - V_s'^2)\right] \qquad (6\text{-}9)$$

where $V_d' = V_d - I_d R_d$

$V_s' = I_d R_s$

V_d = applied drain potential

V_s = applied source potential zero for grounded source configuration

R_d, R_s = parasitic drain and source resistances

For a grounded source case $V_s = 0$ and usually R_s is small and negligible and Eq. 6-9 becomes

$$I_d = \frac{C_{ox}\mu W}{L}\left[(V_g - V_T)V_d' - 1/2\ V_d'^2\right] \qquad (6\text{-}10)$$

$$= \beta\left[(V_g - V_T)V_d' - 1/2\ V_d'^2\right]$$

where $\beta = \frac{\mu C_{ox} W}{L}$

When V_d is increased while keeping V_g constant, I_d attains a maximum when $V_d' = V_g - V_T$. According to Eq.(6-4), Eq.(6-10) the mobile charge in the channel becomes zero when $V(x) = V_g - V_T$. Hence we must conclude that I_d is maximum when the mobile charge in the channel near the drain has decreased to zero. The potential in the channel at which this situation occurs is called the pinch-off potential V_p.

$$V_p = V_g - V_T = V_d' \text{ at } V_p \qquad (6\text{-}11)$$

Substitution of $V_d' = V_g - V_T$ into Eqn's (6-9) and (6-10) yields for the maximum drain curret

$$I_{ds} = \beta \frac{(V_g - V_T)^2}{1+\beta R_s (V_g - V_T) + \sqrt{1 + 2\beta R_s (V_g - V_T)}} \qquad 6\text{-}12$$

For $R_s \cong 0$

$$I_{ds} = \beta \frac{(V_g - V_T)^2}{2} \qquad (6\text{-}12a)$$

The drain current as given by Eq.(6-12) is caused by a potential drop V_p along the conductive channel. A further increase of V_d beyond the pinch off value V_p will not alter this potential drop. If furthermore the channel conductance (i.e., the charge distribution inside the channel), and the effective channel length remain constant, the drain current will not alter when $V_d' > V_p$. It is assumed that beyond pinch off the potential difference $V_d' - V_p$ appears across a very small, high-ohmic region, depteted of electrons, adjoinig the drain. The electrons supplied by the conductive channel are transported towards the drain by the high electric field in the pinch off region.

We can calculate the low-frequency transconductance in common source connection.

$$g_g = \left.\frac{\partial I_d}{\partial V_g}\right|_{V_d} = \beta V_d' \quad \text{for } V_d' < V_p, R_s \cong 0 \qquad (6\text{-}13)$$

The transconductance increases linearly with V_d until $V_d' = V_p$ and g_g attains the saturation value g_m,

$$g_m = \left.\frac{\partial I_{ds}}{\partial V_g}\right|_{V_d} = \beta (V_g - V_T) = \beta V_p \quad \text{for } V_d' > V_p \qquad (6\text{-}14)$$

If we do not neglect the parasitic source resistance then Eq.(6-14) becomes

$$g_m = \frac{\partial I_{ds}}{\partial V_g} = 2\beta \frac{V_g - V_T}{1 + 2\beta R_s (V_g - V_T) + \sqrt{1 + 2\beta R_s (V_g - V_T)}} \qquad (6\text{-}15)$$

We note that g_m is equal to the channel conductance at small drain voltages. From Eq.(6-10)

$$G_o = \left(\frac{I_d}{V_d}\right)_{V_d' \to 0} = \beta (V_g - V_T) = \beta V_p \qquad (6\text{-}16)$$

The same equality for $g_m = G_o$ was found for the junction FET [Eq. 1-27 and 1-37].

These are the equations of a square-law device. The quantitative aspects of the characteristic curves are illustrated in Fig 6-3 . The gain bandwidth product of this device is

$$\frac{g_m}{C_{in}} = \frac{1}{\tau_{gb}} = \left(\frac{2\mu}{L^2}\right) \frac{V_g - V_T}{1 + 2\beta R_s (V_g - V_T) + \sqrt{1 + 2\beta R_s (V_g - V_T)}} \qquad (6\text{-}17)$$

where C_{in} = Total input capacitance $= (C_{ox} W)L$

neglecting R_s, this reduces to

$$\frac{g_m}{C_{in}} = \frac{1}{\tau_{gb}} = \frac{\mu}{L^2}(V_g - V_T) \qquad (6\text{-}17a)$$

The transit time τ_T of a carrier in the channel as it moves from source to drain is :

$$\tau_T = \int_0^L \frac{dx}{\mu E_x} \quad \text{for} \quad V_d > V_g - V_T \qquad (6\text{-}18)$$

Then $\tau_{gb} \cong \tau_T$, and the gain-band width is

$$GB \cong (2\pi\tau_T)^{-1} \qquad (6\text{-}19)$$

At this point we should remark that this first order MOS transistor theory is the simplest way to describe the observed I - V characteristics quantitatively, though quantitatively this theory is not always correct. The gradual channel approximation is valid only when the electric field Ex in the channel is small compared with the field E_z in the oxide. When V_d is increased, this approximation becomes less justified, especially when V_d is increased beyond pinch off. Apart from the use of the gradual and shallow channel approximations and the assumption of constant mobility, the validity of the simple MOS-transistor theory described above depends on the assumption of a constant threshold voltage V_T in Eq.(6-3). V_T represents different physical entities of which the oxide charge, the interface states charge, and the contact potential seem to be constant under relevant experimental conditions. But the charge due to acceptor ions in the depletion region is a function of the channel potential V(x) which varies along the channel. Deviations from this first - order approximations are important because they determine the finite drain resistance in saturation and hence the voltage gain of the transistor. In the sections to follow these effects will be discussed.

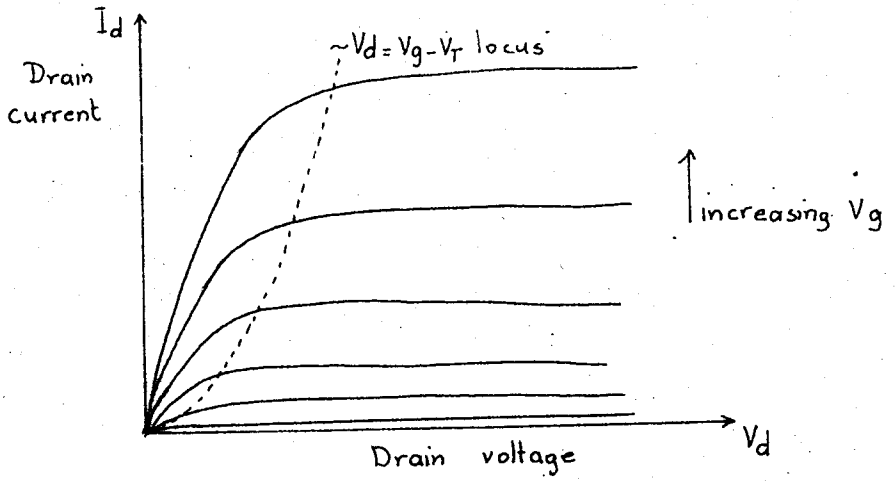


Fig 6-3. MOS-transistor characteristics for an inversion type unit. I_d vs V_d with gate voltage as a parameter.

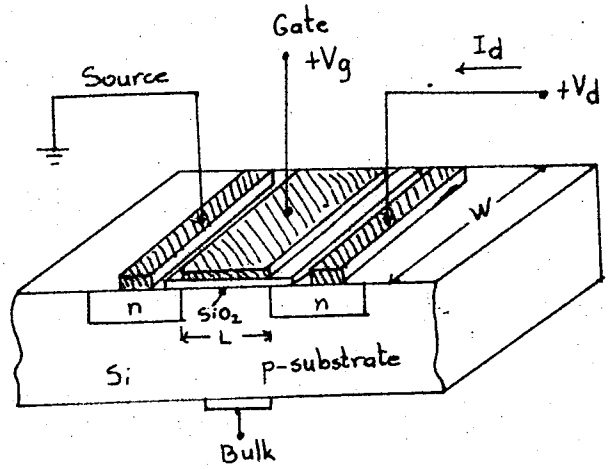


Fig 6-4. Schematic representation of an n-channel MOST with outside connections.

6-4 MOS Transistor Theory with Depletion Charge effect Included.

Consider Fig.6-4. Let $V_d = 0$ and $V_b = 0$. Let the gate potential V_g increase slowly. At a certain gate potential, the total surface density, V_s, in the Si becomes zero everywhere between source and drain ; this means that the electron energy bands in the crystal are straight at the surface (flat-band condition). When V_g is increased further, holes will be driven away from the surface , leaving behind a depletion layer consisting of acceptor ions.The energy bands become curved at the surface.At sufficient band bending inversion will occur at the surface. Thus an n-channel is formed which is separated from the p-type bulk by a depletion region It has been mentioned already that after an inversion layer is formed at a certain gate potential, a further increase of V_g only results in an increase of electron charge in the inversion layer, but the depletion charge remains almost constant. This is caused by the fact that the electron density at the surface increases exponentially with the band bending while the depletion charge is proportional only to the square root of the same quantity. It is assumed that the depletion charge remains constant after the surface electron density n_s becomes equal to N_a of the bulk material. In such a case the diffusion potential , V_o , between channel and bulk becomes

$$V_o = 2 \frac{kT}{q} \ln \frac{N_a}{n_i} \qquad (6\text{-}20)$$

and the depletion charge Q_d is

$$Q_d = -(2 \epsilon_o \epsilon_{si} q N_a)^{1/2} V_o^{1/2} \qquad (6\text{-}21)$$

If now an additional external potential V_b is applied between the channel and the substrate, an additional depletion charge results. Under such circumstances the total depletion charge density at a point x where the channel potential is V(x) will become

$$Q_d(x) = -a \left[V(x) - V_b + V_o \right]^{1/2} \qquad (6\text{-}22)$$

where $a = (2 \epsilon_o \epsilon_{si} q N_a)^{1/2}$

The charge density in the channel as given by Eq.(6-4) must be modified now for

$$Q_d + Q_m = -C_{ox}\left[V_g - V_T - V(x)\right] \qquad (6\text{-}23)$$

so that

$$Q_m(x) = -C_{ox}\left[V_g - V_T - V(x)\right] + a\left[V(x) - V_b + V_o\right]^{1/2} \qquad (6\text{-}24)$$

and the differential current equation becomes:

$$I = \mu C_{ox} W \left\{ V_g - V_T - V(x) - \frac{a}{C_{ox}}\left[V(x) - V_b + V_o\right]^{1/2} \right\} \frac{dV(x)}{dx} \qquad (6\text{-}25)$$

Intergrating Eq.(6-25), we get

$$\int_0^L I\,dx = \int_0^{V_d} \mu C_{ox} W \left\{ V_g - V_T - V(x) - \frac{a}{C_{ox}}\left[V(x) - V_b + V_o\right]^{1/2} \right\} dV(x) \qquad (6\text{-}26)$$

and assuming constant I, μ, V_T and V_o

$$I_d = \beta\left\{(V_g - V_T)V_d - 1/2 V_d^2 - 2/3 \frac{a}{C_{ox}}\left[(V_d - V_b + V_o)^{3/2} - (-V_b + V_o)^{3/2}\right]\right\}$$
$$\text{for } V_d < V_p \qquad (6\text{-}27)$$

At pinch off $Q_m \longrightarrow 0$ and from Eq. (6-24)

$$C_{ox}\left[V_g - V_T - V_p\right] = a\left[V_p - V_b + V_o\right]^{1/2} \qquad (6\text{-}28)$$

so that $V_p = (V_g - V_T) - 2\phi\Theta$

where $\Theta = \left[1 + \left\{(V_g - V_T) + (-V_b + V_o)\right\}/\phi\right]^{1/2} - 1$

$\phi = \frac{1}{4}(a/C_{ox})^2$

Since $C_{ox} = \epsilon_o \epsilon_{ox} / T_{ox}$ and $a = (2\epsilon_o \epsilon_{si} q N_a)^{1/2}$, we can consider ϕ as a device constant which is determined by the oxide thickness and the bulk acceptor density N_a. As V_T and V_o are constants, the factor varies with the potential difference between gate and bulk $V_g - V_b$.

Differentiating Eq. 6-27 with respect to V_d, we see that maximum current is again obtained when $V_d = V_p$.

This maximum saturated current will be :

$$I_{ds} = \beta\left[\frac{1}{2}(V_g - V_T)^2 + \frac{4}{3}\phi^{1/2}(-V_b + V_o)^{3/2} - \phi^2\Theta^2(2 + 4/3\,\Theta)\right] \qquad (6\text{-}29)$$

The transconductance, with the gate as input electrode can be obtained from Eq. 6-27 :

$$g_g = \left.\frac{\partial I_d}{\partial V_g}\right|_{V_d, V_b} = \beta V_d \quad \text{for} \quad V_d < V_p \tag{6-30}$$

and from Eq.(6-29) , The saturated transconductance becomes

$$g_m = \beta V_p = \beta\left[(V_g - V_T) - 2\phi\theta\right] \quad \text{for} \quad V_d \gg V_p \tag{6-31}$$

But since I_d and I_{ds} depend on the bulk voltage as well as the gate voltage, the bulk of the MOS transistor can also be used as a control electrode. Thus with the bulk electrode as the signal input electrode :

$$g_b = \left.\frac{\partial I_d}{\partial V_b}\right|_{V_d, V_g} = \beta 2\phi^{1/2}\left\{\left[V_d + (-V_b + V_o)\right]^{1/2} - (-V_b + V_o)^{1/2}\right\} \quad \text{for} \quad V_d < V_p \tag{6-32}$$

$$g_{mb} = \beta\left\{2\phi\theta - 2\phi^{1/2}(-V_b + V_o)^{1/2}\right\} \quad \text{for} \quad V_d \gg V_p \tag{6-33}$$

when the gate and bulk are connected in parallel and used as the signal input electrode :

$$g_{g,b} = \beta\left[V_d + 2\phi^{1/2}\left\{(V_d + (-V_b + V_o))^{1/2} - (-V_b + V_o)^{1/2}\right\}\right] \text{ for } V_d < V_p \tag{6-34}$$

$$g_{mgb} = \beta\left\{(V_g - V_T) - 2\phi^{1/2}(-V_b + V_o)^{1/2}\right\} \text{ for } V_d \gg V_p \tag{6-35}$$

Note that $g_{g,b} = g_g + g_b$ (6-36)

and $g_{mgb} = g_m + g_{mb}$

For smalldrain voltages the channel conductance becomes

$$G_o = (I_d/V_d)_{V_d \to 0} = \beta\left\{(V_g - V_T) - 2\phi^{1/2}(-V_b + V_o)^{1/2}\right\} \tag{6-37}$$

Thus $g_{mgb} = G_o$

From Eqns: (6-29),(6-31),(6-33) and (6-35) we can see that I_{ds}, g_m , g_{mb} , and g_{mgb} all reduce to zero at the same threshold gate voltage. This is the gate voltage which causes the mobile charge Q_m near the source to become zero and is given by

$$(V_g - V_T) = 2\phi^{1/2}(-V_b + V_o)^{1/2} = \frac{a}{C_{ox}}(-V_b + V_o)^{1/2} \tag{6-39}$$

Note that in the MOS transistor theory just derived β, V_T , ϕ and V_o are device parameters. The value of V_o can be estimated from Eq.(6-20), When the acceptor density N_a at the surface is known. Methods for determining β, V_T and ϕ experimentally will be discussed in the next chapter.

6-5. Relation of charge mobility to saturation

6-5A. Constant mobility case : The concept of complete pinch off at the drain for $V_d \geq V_g$ with the paradox of continuity of current flow has led to proposals that phenomena such as saturation of carrier velocity or minority carrier flow are responsible for current saturation. In actuality there is never complete pinch off at the drain, but as V_d approaches V_g ,the gradual channel approximation is not valid any more and the drain region must be treated differently from the source region. First let us show that actually there is not a complete pinch off when a drain region forms; assuming that $V_T = 0$, we can write the current expression of Eq.(6-10) as

$$I_d = \frac{C_{ox} \mu W}{2L} \left[V_g^2 - (V_g - V_d)^2 \right] \quad \text{for } V_d < V_g \qquad (6\text{-}40)$$

But if we first assume that the drain region starts to form at a point x, we can write Eq. 6-40 as

$$I_d = \frac{C_{ox} \mu W}{2x} \left[V_g^2 - (V_g - V(x))^2 \right] \qquad (6\text{-}41)$$

$$V_g - V(x) = \left[-\frac{2x}{\mu C_{ox} W} I_d + V_g^2 \right]^{1/2} \qquad (6\text{-}42)$$

differentiating Eq.(6-42) twice with respect to x

$$\frac{-d^2V(x)}{d^2x} = \frac{dE(x)}{dx} = -\frac{1}{4} \left(V_g^2 - \frac{2x}{\mu C_{ox} W} I_d \right)^{-3/2} \left(2I_d / \mu C_{ox} W \right)^2 \qquad (6\text{-}43)$$

As V_d approaches some value V_{dc} , the gradual channel approximation fails, the rate of change of the drift field $\frac{dEx}{dx}$ exceeds the rate of change of the gate field $\frac{dEz}{dz}$ in some part of the channel, so that the charge density in that part of the channel begins to be controlled by the drift field. This is nothing else but a space-charge limited current flow condition where the appied field determines both the drift velocity and the space charge density.

Since as the drain approaches pinch off $V_d \rightarrow V_g$, and $x \rightarrow L$, imposing these conditions on Eq. 6-41 and rearranging terms, we get

$$\left(\frac{2I_d}{\mu C_{ox} W}\right)^2 \cong \frac{V_g^4}{L^2} \qquad (6\text{-}44)$$

and $$(V_g^2 - \frac{2x}{\mu C_{ox} W} I_d)^{-3/2} = \left[V_g^2 - V_g^2 + (V_g - V_d)^2\right]^{-\frac{3}{2}} = (V_g - V_d)^{-3}$$

so that $$\left.\frac{dE_x}{dx}\right|_{x=L, V_d \rightarrow V_g} \cong \frac{V_g^4}{4L^2(V_g - V_d)^3} = \frac{V_g}{4L^2}\left(1 - V_d/V_g\right)^{-3} \qquad (6\text{-}45)$$

If we set, for the condition when the drain region just starts to form $$\left.\frac{dE_x}{dx}\right|_{x=L} = \left.\frac{dE_z}{dz}\right|_{x=L}$$

where

$$\frac{dE_z}{dz} \cong \frac{E_z}{d} = \frac{\epsilon_{ox} E_{ox}}{\epsilon_{si}\, d} = \frac{\epsilon_{ox}\,(V_g - V_d)}{\epsilon_{si}\; dT_{ox}} \qquad (6\text{-}46)$$

where ϵ_{ox}, ϵ_{si} are permittivities of the oxide and semiconductor respectively.

Equating Eq.(6-45) and(6-46), and solving for $V_d = V_{dc}$, we get

$$V_{dc} = V_g\,(1 - K_1) \qquad (6\text{-}47)$$

as the actual drain voltage for which the drain region forms. Here

$$K_1 = \left(\frac{\epsilon_{si}\, d\, T_{ox}}{4\epsilon_{ox} L^2}\right)^{1/4} \qquad (6\text{-}48)$$

Thus we see that at $V_d = V_{dc}$ a drain region forms, but the channel charge is not zero and complete pinch off has not occured. For $V_d = V_{dc}$, space charge dominated currents maintain current continuity from the source region to the drain. The excess drain voltage generates a space charge region extending from the drain to the channel a distance l_d. This is called the drain region. From Eq.(6-47) it is apparent that V_{dc} is less than V_g by an amount equal to K_1. In most devices it can be shown that the error introduced when we set $V_d = V_g$ for saturation is small enough to be safely neglected. In a typical Si device $L = 10\,\mu$, $d = 1000$ Å, $T_{ox} = 1000$ Å, $\epsilon_{ox}/\epsilon_{si} = 1/3$. This gives $K_1 \cong 10^{-1}$ so that $V_{dc} \cong 0.9\, V_g$, and

the error is less than 10% . Hofstein and Warfield have shown that the position x at which the mobility begins to become field dependent is about o.81 where l is the channel length. Thus within actual structural and voltage limits the constant mobility approximation is valid over the major portion of the channel source region. Si MOS transistors which have a channel length of about 10 μ show excellent " square law" characteristics associated with the constant mobility approximation for gate voltages up to 10-20 volts. But deviations from constant-mobility approximation become significant as the channel length is reduced or operation voltages are increased.

6-5B. Inverse linear mobility, constant carrier velocity.

The effect of saturation of carrier velocity causing current saturation before channel pinch off at the drain occurs has been experimentally observed for Ge junction - type FETs. In Si junction gate FETs and MOS transistors saturation of current occurs for a value of drain voltage close to the pinch off voltage and therefore the effect of mobility variation with field within the drain region is not a dominant effect. The analysis of the saturation of carrier velocity involves the solution of Poisson's and Gauss' laws and is the same for the FET as well as the MOS transistor :

Let $\bar{v}_l$ = saturated carrier drift velocity.

E_l = field when the semiconductor exhibits a saturation of carrir velocity .

V_{dc} = drain voltage when drain region forms.

$E_{dc} = E_x\ (l) \Big|_{V_d = V_{dc}}$

If we leave the gate voltage at some value V_g , and increase the drain voltage slowly above zero, as the drain voltage rises, current flows from source to drain according to the first order gradual channel analysis. As mentioned before, the drain region of the channel will form when $\frac{dEx}{dx} \geq \frac{dEz}{dz}$ at the drain. If the

field at the drain reaches E_l before $\frac{dEx}{dx} > \frac{dEz}{dz}$, a region of saturated velocity appears. Here we can separate the analysis into two parts :

a) $E_l > E_{dc}$

For this case the drain region forms before the onset of a saturation velocity. Therefore the saturation of drain current results from the formation of a drain region.

b) $E_l < E_{dc}$

In such a case, carrier velocity saturates before drain region forms. If we consider that transition from constant mobility region I to saturated velocity region II is abrupt :

$$\mu = \mu_o \qquad E(x) < E_l \qquad o < x < x_l \quad \text{Region I} \quad (6\text{-}49)$$

$$\mu = \mu_o \frac{E_l}{E(x)} \qquad E(x) > E_l \qquad x_l < x < L \quad \text{RegionII} \quad (6\text{-}50)$$

where μ_o = mobility under constant mobility regime

x_l = The channel distance where carrier velocity becomes constant .

The electric field at the drain end of region I is :

$$E(x_l) = I_d / \mu_o Q(x) = \frac{1}{2l} \cdot \frac{V_g^2 - (V_g - V_l)^2}{V_g - V} , \quad E(x) < E_l \quad (6\text{-}51)$$

Since there is an abrupt jump at $x = x_l$ from $E(l)$ to E_l, we can find the drain voltage V_d when saturation just starts. At such an instant $l_d = 0$, $x_l = l$. Making use of Eqn's (6-50) and (6-51)

$$E_l = \frac{I_d}{\mu_o E_l / E(x_l)} = \frac{1}{2l} \cdot \frac{V_g^2 - (V_g - V_l)^2}{V_g - V_l} \quad (6\text{-}52)$$

By letting $V_l = V_{dl}$ for the instant when saturation has just started, we can solve Eq.(6-52) for V_{dl} which is the drain voltage at which region II just forms at the drain due to saturation of carrier velocity :

$$V_d = V_g + E_l l - \left[E_l^2 l^2 + V_g^2\right]^{1/2} \quad (6\text{-}53)$$

From the continuity of the current in the channel, we can derive

the current expression per unit channel width in the source region of the channel.

$$I_d = \frac{\mu_o C_{ox}}{2l} \left[V_g^2 - (V_g - V_{dl})^2 \right] \qquad (6\text{-}54)$$

Substituting Eq(6-53) into (6-54) we get

$$I_d = \frac{\mu_o C_{ox}}{l} (E_l L)^2 \left\{ \left[1 + \left(\frac{V_g}{E_l L} \right)^2 \right]^{1/2} - 1 \right\} \qquad (6\text{-}55)$$

Then the transconductance per unit channel can be written as :

$$g_m = \frac{\partial I_d}{\partial V_g} = \mu_o C_{ox} E_l V_g' \, (1 + V_g'^2)^{-1/2} \qquad (6\text{-}56)$$

where $V_g' = V_g / E_l L$

Studying Eq. (6-56) we can see that for large values of V_g which makes $V_g' \gg 1$, the transconductance g_m saturates. Neglecting 1 with respect to V_g' in Eq.(6-56), we get

$$g_m = \mu_o C_{ox} E_l = C_{ox} \bar{v}_l \quad \text{for} \quad V_g \gg E_l L \qquad (6\text{-}57)$$

Under such conditions (from Eq.6-53) the channel field is approximately given by V_d/L and is independent of V_g. The channel current per unit channel width is then :

$$I_d = C_{ox} \bar{v}_l V_g \qquad \text{for } V_g \gg E_l L \qquad (6\text{-}58)$$

Fig. 6-5 shows the normalized transconductance versus normalized gate voltage for the case of limited carrier velocity $\bar{v}_l$.

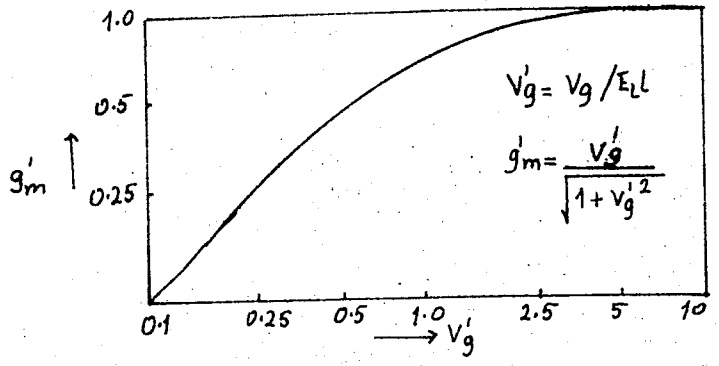


Fig 6-5. Normalized transconductance g_m vs normalized gate voltage

CHAPTER 7

MOS-TRASISTOR CHARACTERISTICS

The theory of operation of insulated gate field effect transistors is discussed in the preceding chapters. Static characteristics and device parameters will be discussed in this chapter.

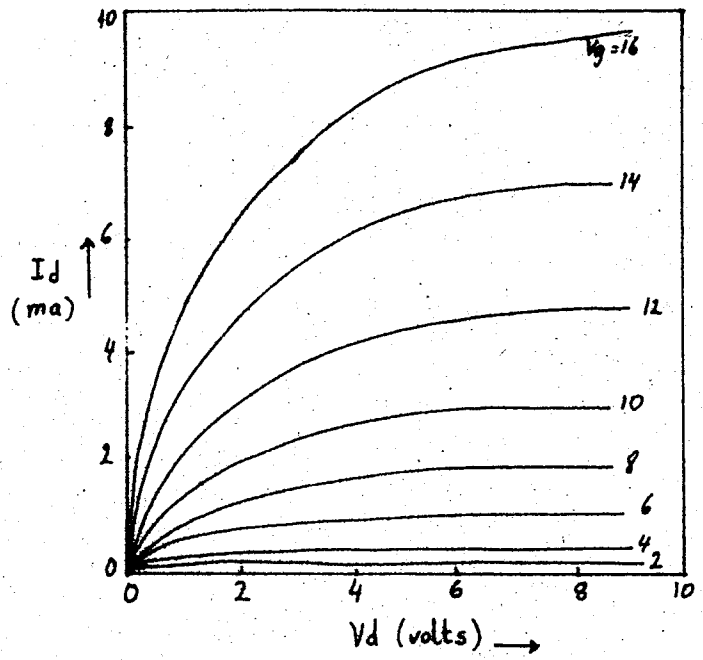
7-1 . Output characteristics

7-1A. Output characteristics of induced channel transistors.

Fig.7-1 shows the output characteristics of an induced channel MOS transistor. This unit is desirable for use in logical switching circuits because direct-coupled inversion is possible without the need for level shifting between stages. The device of Fig.7-1 is made of ptype Si. Its effective pinch off voltage is approximately +2 volts. As mentioned in chapter 6, the device exhibits very low drain current at zero gate bias. The current increases with positive gate voltage. Since conduction carriers for the MOS unit shown in Fig.7-1 are electrons, this is an n-type device. Output characteristics of a p-type unit is shown in Fig. 7-2. The pinch off voltage of this unit is -0.5 volts and requires negative gate voltage for conduction.

An important feature of this type of transistor is that the gate electrode must cover the entire channel and overlap both source and drain regions as shown in Fig.7-3 . Any channel region left exposed will contribute a very high series resistance to the device. Since this overlap results in a substantial capacitance from the gate-to-source (C_{gs}) and gate-to-drain (C_{gd}), the oxide thickness (T_{ox}) is increased over the heavily doped regions as seen in Fig. 7-3 to compensate for it.

If the special case of oxide thickness much greater than channel depth is considered, the dynamic saturation drain resistance may be directly related to the input transconductance of the device and can be derived from the output characteristics.



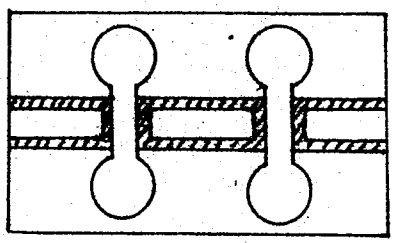
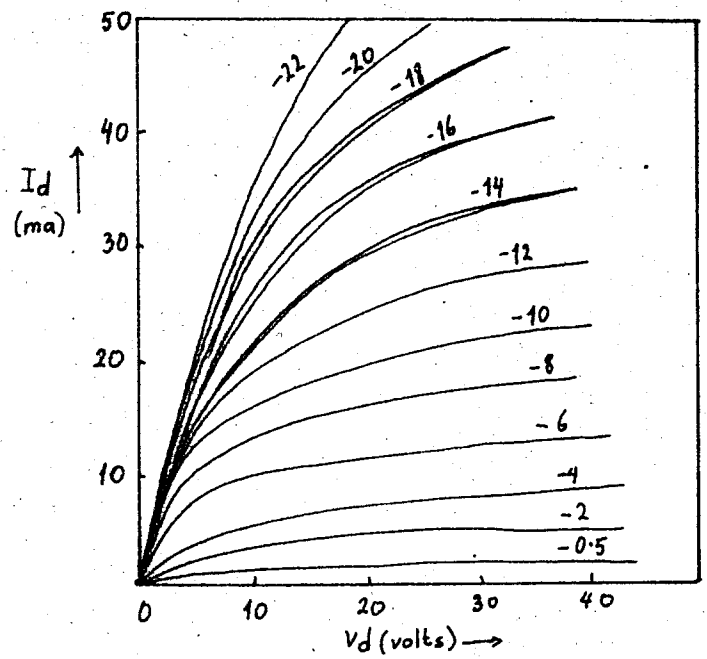


Fig.7-1a. Ouput characteristics of a typical ladder geometry enhancement transistor. Source and substrate are at ground potential.(n-type device)

7.1b. Ladder geometry layout of MOS transistor .



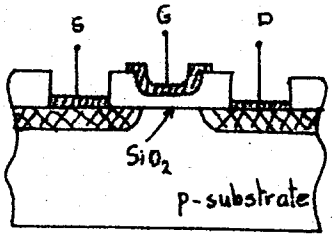


Fig 7.3. Cross sectional view of ladder geometry transistor showing raised oxide over source and drain regions.

Fig.7-2. Output characteristics of a p-type enhancement MOS transistor which has maze geometry.

Defining $$\frac{1}{r_{ds}} = g_{ds} = \left.\frac{I_{ds}}{V_d}\right|_{V_g} \qquad (7\text{-}1)$$

The following can also be deduced from the output characteristics:

$$\mu_A = \left.\frac{\partial V_d}{\partial V_g}\right|_{I_d} \qquad (7\text{-}2)$$

$$g_m = \left.\frac{\partial I_{ds}}{\partial V_g}\right|_{V_d} \qquad (7\text{-}3)$$

hence $$g_m r_{ds} = \mu_A \qquad (7\text{-}4)$$

One finds that

$$g_m/g_{ds} = \mu_A = C_{gc}/C_{dc} \qquad (7\text{-}5)$$

where

$$\frac{C_{dc}}{C_{gc}} = \frac{\text{effective drain-to-channel coupling capacitance}}{\text{gate to channel capacitance}}$$

Later on it will be shown that

$$\frac{C_{dc}}{C_{gc}} \cong \frac{T_{ox}\,\epsilon_{si}}{L\,\epsilon_{ox}} \qquad (7\text{-}6)$$

where ϵ_{si}=dielectric constant of the intrinsic silicon substrate. As we can see from Eqns (7-5) and (7-6), the transistor amplification factor μ_A is dependent solely on the geometry. This result is interesting since it is exactly analogous to that which one obtains for the vaccum tube triode.

Since $g_m = G_o \frac{(V_g - V_T)}{V_p}$, where G_o is the unmodulated channel conductance

$$\frac{g_{ds}}{G_o} = \left(\frac{C_{dc}}{C_{gc}}\right)\left(\frac{V_g - V_T}{V_p}\right) \qquad (7\text{-}7)$$

Fig 7-4 shows measured values of g_m, g_{ds}, I_{ds} and μ_A versus V_g for the ladder type enhancement MOS transistor of Fig. 7-3. When experimental results and theoretical values for the parameters are compared close agreement is found. Fig 7-5 shows the measured values of g_m, g_{ds}, and μ_A versus V_g for the p-type

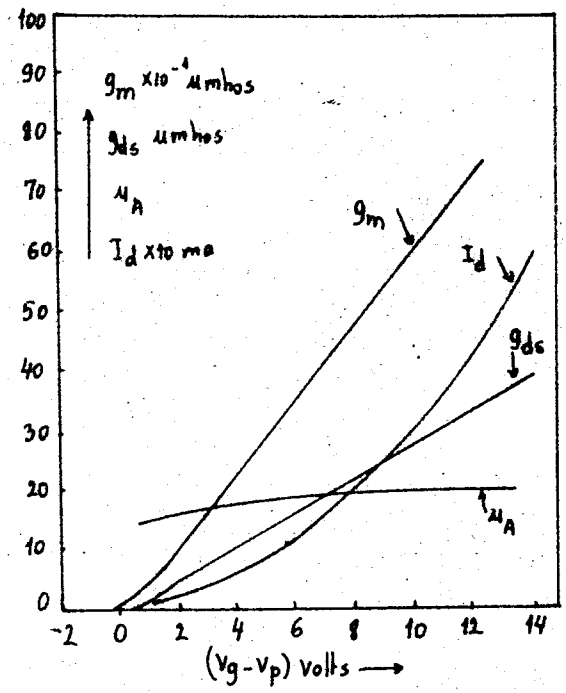


Fig.7.4. Measured values of g_m, g_{ds}, I_{ds}, and voltage gain μ_A vs V_g for the MOS transistor which has ladder geometry shown in Fig 7-3.

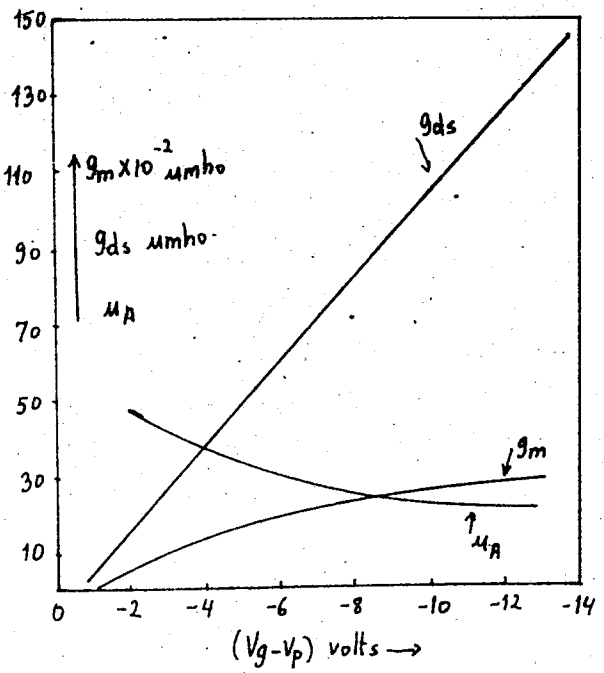


Fig 7.5. Measured values of g_m, g_{ds}, and μ_A vs V_g for the enhancement p-channel MOS transistor which has maze geometry and the output characteristics of which is given in Fig 7-2.

enhancement unit of Fig. 7-2.

Another type of induced channel MOS transistor geometry is the maze geometry. Fig. 7-6 shows the output characteristics of such an enhancement transistor. This is a high current transistor used in memory driver applications. A drain current in excess of 100 ma. can be supplied by a gate voltage of 4.5V while only a drain current of 10 ma is supplied by a gate voltage of 16 V by the ladder geometry unit of Fig. 7-1.

7-1B. Output Characteristics of Depletion-type Transistors

Fig 7-7 gives the output characteristics of a depletion-type transistor. This type of transistor is fabricated with the source, drain and channel of the same conductivity type to yield substantial drain current for zero gate bias. The free carrier concentration in the channel is much lower than the source or drain doping level so that pinch off is obtained with a moderate oxide field. This transistor is widely used as a small signal linear amplifier over a wide frequency range. The geometry used for this type of transistor is seen in Fig 7-8. The depletion unit does not require the gate electrode to overlap both source and drain regions. offset gate geometry is used. The gate electrode is usually made to overlap the source region in order to reduce the source resistance R_s. This is done so because large values of R_s cannot be tolerated since any series source resistance will be multiplied by $\mu_A + 1$, where μ_A is the voltage amplification factor of the transistor, when it is reflected to the output. The effect of increased R_d, on the other hand, is only to increase the drain voltage at which drain current saturates.

7-2. Voltage Breakdown in MOS Transistors:

7-2A. Soft Channel breakdown:

To obtain maximum drift velocity and minimum transit time, it is desirable to have large channel drift fields. For sufficiently small dimensions, internal fields may become large enough to cause avalanche breakdown within the channel or directly from the drain to the substrate. As the channel field is increased ,

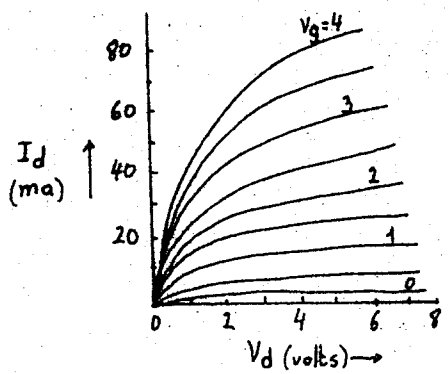


Fig 7.6. Output characteristics of a maze geometry enhancement MOS transistor. (n-type)

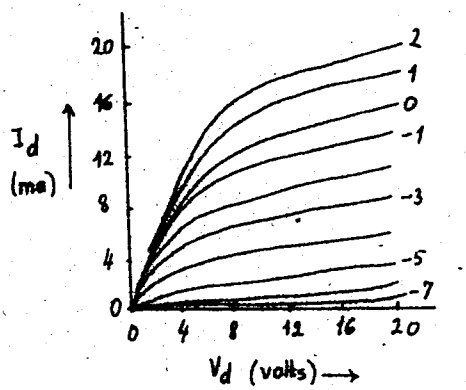


Fig.7.7. Output characteristics of a depletion type MOS transistor.

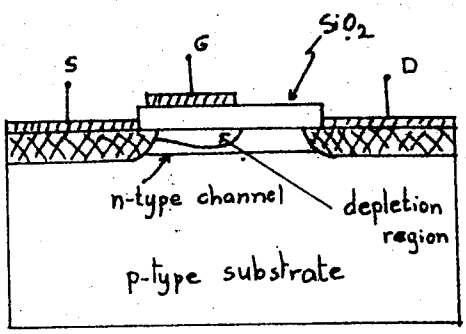


Fig.7.8 Off set gate geometry used in depletion-type MOS transistors.

hole electron pair generation due to impact ionization by hot carrieres increases. This breakdown in the channel usually commences in the region of maximum field (i.e., the space charge limited region near the drain.) For convenience, an n-type channel and n-type source drain contacts will be assumed in the following discussion. However, the discussion may be readily extended to include p-type channel units. A good quantitative understanding may be obtained from the following: As the voltage at the drain end of the channel exceeds $V_g - V_p (= V_g - V_T)$ the depletion region at that point extends completly across the channel. The channel may now be thought of as consisting of three regions: Region 1- The section from the source up to the point where the channel voltage equals V_g , region 2- from the V_g point to the point where the channel voltage is $V_g - V_p$, and Region 3- from the $V_g - V_p$ point to the drian contact. In region 1 , there is an enhancement of the channel charge, in region 2, there is a partial depletion of channel charge and in region 3, the depletion region extends completely across the channel. Hence in region 1 and 2 , the current flow is ohmic while as in region 3 it is space-charge limited. A study of Fig.7-9 will show that when the gate is made more positive to enhance the conductivity of the channel, the length of region 2 and 3 decreases. This yields an increased electric field in these regions resulting in a diminishing of the excess drain voltage required for channel breakdown. For a negative gate voltage, region 1 vanishes and region 2 and 3 are enlarged , yielding an increase in the drain voltage required for breakdown. However as the gate voltage continues to go negative significantly beyond cut off , the gate field begins to directly add to and enhance the effect of the drain field, thereby again reducing the drain voltage required for channel breakdown. Note that the soft channel breakdown caused by carrier generation is greatest for maximum gate voltage and becomes almost negligible (hardening of the knee) at channel

pinch off. The final asymptotic drain current rise represents direct breakdown to the substrate.

7-2B. Hard direct Breakdown :

For a p-type substrate, the breakown voltage of the drain to-substrate diode is nearly the same as that of an n^+- p diode and varies as a function of doping on the low-doped side of the junction as shown in Fig.7-10. However, the concentration of the electric field at the periphery of the junction results in surface region avalanche breakdown at a voltage lower than that required for breakdown of the bulk portion of the diode. Fig. 7-11a and 7-11b show quantatively the field distribution for both the deep and shallow diffused junctions. Note that field crowding occurs when the diffusion depth is less than the width of the depletion region. Since planar MOS transistors are often diffused to depths of less than 0.1 μ, the high fields occuring at the periphery result in much lower breakdown voltages than those measured for deep-diffused junctions. A further reduction of the breakdown voltage occurs due to the insulated metal electrode which overlaps the diode junction as seen in Fig.7-12. The field in the surface portion of the junction depletion region increases due to the collective action of both to the drain - to-substrate and drain to gate fields. Special experiments have been carried out with the unit of Fig.7-9. When the metal gate electrode was removed by chemical etching, the breakdown voltage rose from 40 volts to 70 v. The breakdown occurs directly between the drain and gate with the substrate acting as a collector of the impact ionized carriers. Since the substrate acts only as a collector, the breakdown voltage has only a weak dependence on the resistivity of the substrate.

7-2C. Punch-trough Breakdown: If the drain depletion region extends across the channel region until it reaches the source , then it draws space-charge limited current. This is known as the punch-trough breakdown. This breakdown is less significant

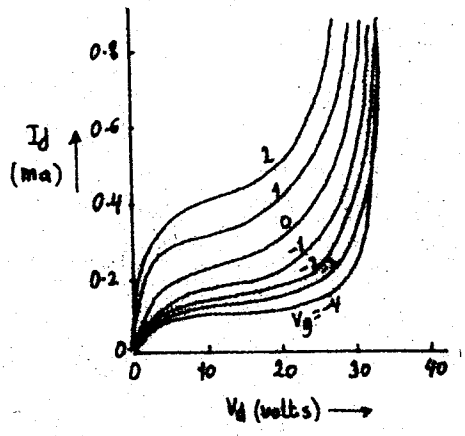


Fig.7.9 Characteristic curves of an n-type transistor showing channel breakdown.

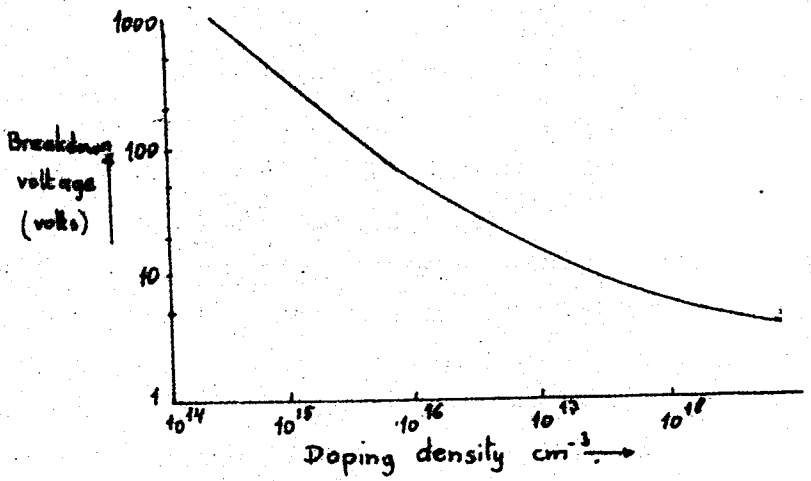


Fig.7.10 Breakdown voltage for p^+-n or n^+-p junctions in Si as a function of doping on the low-doped side of the junction.

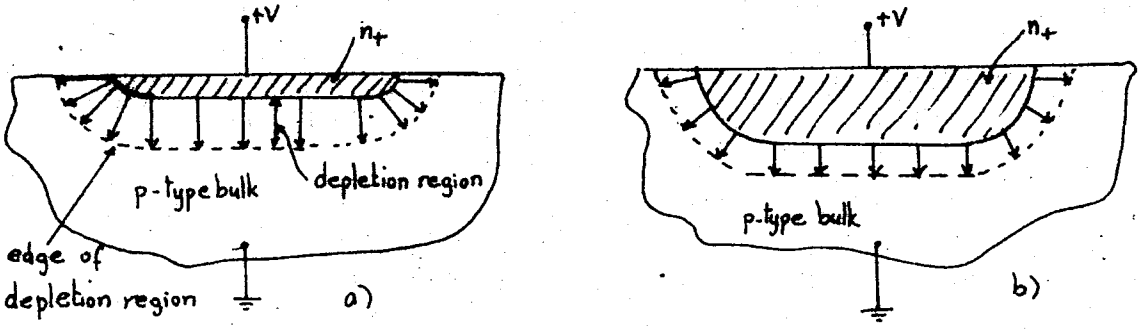


Fig.7.11 Field distribution in deep and shallow diffused junctions.
a) Junction depth ≪ depletion region depth
b) Junction depth ≫ depletion region depth.

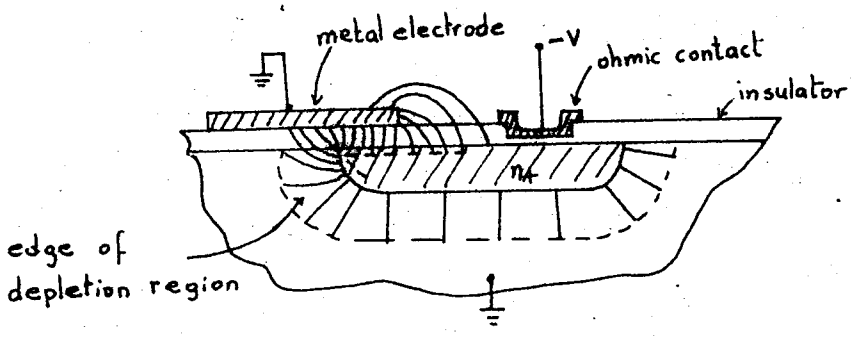


Fig.7.12 Field distortion in a planar p-n junction due to overlapping metal electrode.

in MOS transistors than in bipolar transistors.

7-3. Effect of Substrate Resistivity on Output Characteristics:

A high resistivity (preferably intrinsic) substrate is desirable in MOS transistors in order to keep drain-to-substrate capacitance small. But if we look at Fig. 7-13a and 7-13b which show the effect of substrate resistivity on the amplification factor of a transistor, we see that transistors fabricated on 1 ohm-cm Si wafer demonstrate better saturation and higher amplification factor than the transistors fabricated on a 1000 ohm-cm p-type Si base wafer.

When the field-effect transistor is operated in the saturated portion of the output characteristics, the average drift velocity, $\bar{v}_d$, for the channel carriers (in regions 1 and 2) is approximately

$$\bar{v}_d = \mu \frac{(V_g - V_p)}{L} \qquad (7\text{-}8)$$

The change in channel charge ΔQ per unit channel width may be written as:

$$\Delta Q = C_{gs} \Delta V_g + C_{dc} \Delta V_d \qquad (7\text{-}9)$$

where C_{gc} = gate-to-channel capacitance per unit width

C_{dc} = drain-to-channel capacitance per unit width. Then

$$\Delta I_d = \frac{\Delta Q \; v_d}{L} = (C_{gc} \Delta V_g + C_{dc} \Delta V_d) \frac{V_d}{L} \qquad (7\text{-}10)$$

The amplification factor is defined as:

$$\mu_A = \left. \frac{\Delta V_d}{\Delta V_g} \right|_{\Delta I_d = 0} \qquad (7\text{-}11)$$

Equating Eq. 7-10 to zero and solving for μ_A, we get

$$\mu_A = \frac{C_{gc}}{C_{dc}} \qquad (7\text{-}12)$$

where C_{gc} = gate to channel capacitance $= C_{ox} L = L \frac{\epsilon_{ox}}{T_{ox}}$ (7-13)

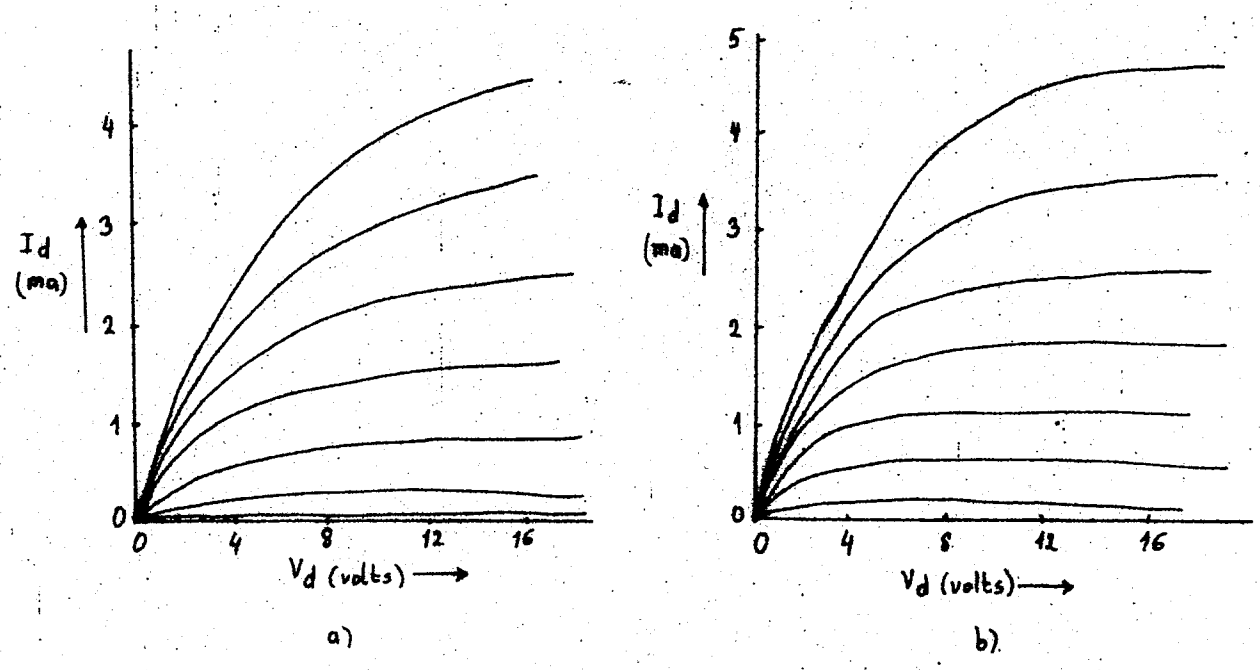


Fig.7-13 Effect of substrate resistivity on the amplification factor of the transistor:
a) 1000 ohm-per-centimeter substrate
b) 1 ohm-per-centimeter substrate

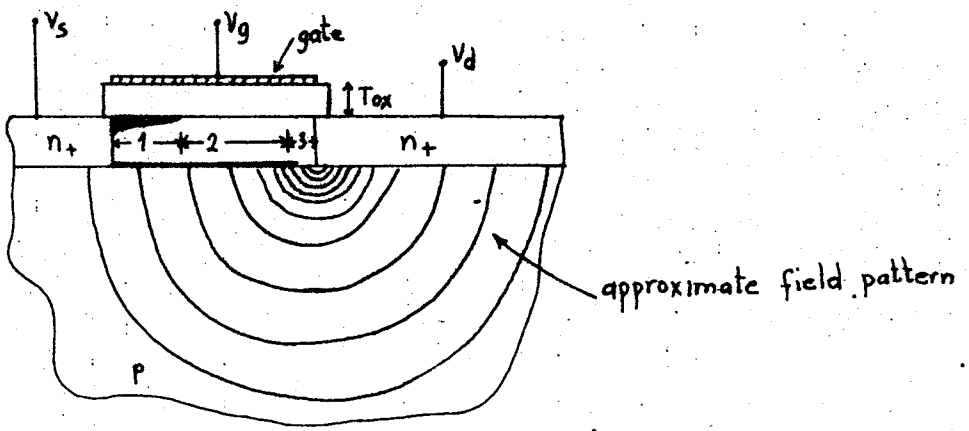


Fig. 7.14 Model for calculating drain-to-channel capacitance C_{dc} .

An expression for C_{dc} can be derived as follows: The model of Fig.7-14 can be used with the assumption that the field lines between the drain and the underside of the channel are semicircles. Then the incremental capacitance from the drain to a section of channel dx in length is simply:

$$d(C_{dc}) = \frac{\epsilon_{si}}{\pi x} dx \qquad (7-14)$$

where x = the distance from the edge of the drain contact to dx. To obtain the total capacitance Eq.(7-14) is intergrated over the ohmic regions of the channel (regionsı and 2) giving

$$C_{dc} = \int_{x_l}^{L} \frac{\epsilon_{si}}{\pi x} dx = \frac{\epsilon_{si}}{\pi} \ln L/x_l \qquad (7-15)$$

where x_l = the beginning of region (3) (= The space charge region). x_l is a function of V_d.

Since C_{dc} varies logarithmically with x_l , we may consider it to be constant. $\frac{L}{x_l}$ varies usually between 10^1 and 10^2,and

$$C_{dc} \approx \epsilon_{si} \text{ per unit channel width} \qquad (7-16)$$

$$\text{and} \quad \mu_A = \frac{L}{T_{ox}} \frac{\epsilon_{ox}}{\epsilon_{si}} \qquad (7-17)$$

Equation (7-17) points out a problem in optimizing the performance of these transistors. If the channel length is reduced to increase the frequency response of the transistor, the voltage gain will drop because of a decrease in the saturation drain resistance.This problem can be remedied by an increase in substrate doping level to the point where the substrate-to-channel depletion region is sufficiently narrow so that penetration of the drain field into the channel region is greatly reduced. Thus the substrate acts as a screen grid in capacitively decoupling the drain from the channel. But this is obtained at the cost of increasing the drain to substrate capacitance. Therefore a compromise must be reached for each application. For Si depletion units, 20-25 ohm-cm material is found to give optimum hich frequency power gain.

7-4 Effect of Substrate doping on Pinch-off Voltage

In the theory of field effect transistors which are based upon surface inversion layers, the effect of the depletion charge is important. Using the pinch off voltage expression derived in chapter 6,

$$V_p = (V_g - V_T) - 2\phi\Theta \qquad (7\text{-}18)$$

and plotting, $V_p = f(V_g - V_T)$ for different substrate doping levels, we get Fig. 7-15 . The oxide thickness was taken to be 0.2μ, $V_b = 0$, and V_o was assumed to be constant. Only when $N_a = 0$ and the depletion charge can be neglected, is V_p equal to $V_g - V_T$ as predicted by the first order theory. Usually substrate doping levels are of the order of 10^{15} to 10^{16} cm^{-3}. It can be seen from Fig 7-15 that in these cases V_p will be considerably smaller than $V_g - V_T$.

7-5. Effect of Substrate bias on MOS Characteristics and Parameters

a) An analysis of the effect of the depletion charge and therefore the substrate was given in the MOS transistor theory (chapter 6) and the possibility of using the bulk as a control electrode was pointed at. The transfer characteristics of a depletion unit controlled from the substrate is shown in Fig 7-16 with the metal gate connected to the grounded source ($V_g = 0$). This curve closely approximates an exponential because the depletion of channel charge proceeds from the low density in the bulk to the high density at the Si surface. The equation governing the dependence of I_d upon substrate bias was derived in chapter 6 .It is repeated here for easy reference.

$$I_d = \beta\left[V'_g V_d - 1/2\ V_d^2 - 2/3\ \frac{a}{C_{ox}}\left\{(V_d + V'_b)^{3/2} - (V'_b)^{3/2}\right\}\right] \qquad (7\text{-}19)$$

where

$$V'_g = V_g - V_T$$

$$V'_b = -V_b + V_o$$

$$a = (2\epsilon_o \epsilon_{si} q N_a)^{1/2}$$

$$\beta = \mu C_{ox} W/L$$

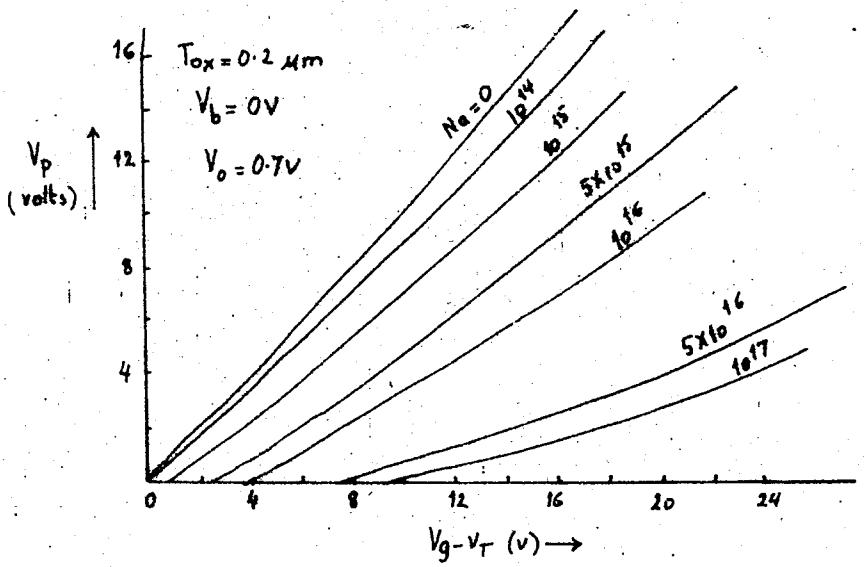


Fig.7-15. $V_p = (V_g - V_T) - 2\phi_o$ for different substrate doping levels.

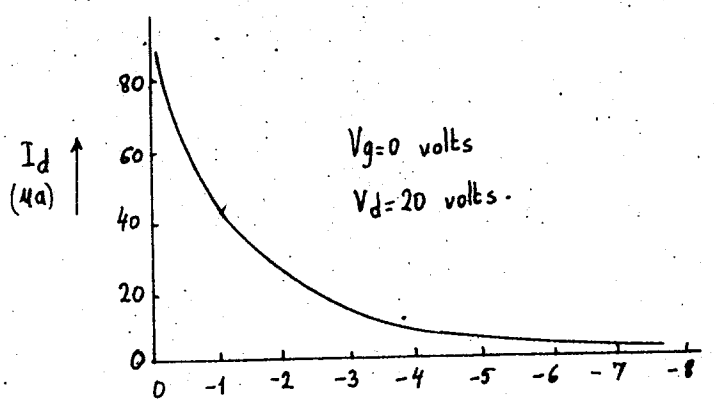


Fig.7-16. Transfer characteristics of a depletion type MOS transistor operated from the 10 ohm-cm-resistivity substrate. The metal gate electrode is at source potential.

V_T = offset gate potential

and V_o is calculated from

$$V_o = 2 \frac{kT}{q} \ln \frac{Na}{ni} \qquad (7\text{-}20)$$

Calculating V_o from equation (7-20), the device parameters β, V_T, and ϕ can be found from the experimentally obtained fig (7-17) which is a plot of the transconductance $g_{m\,g,b}$ measured as a function of V_g for different values of V_b, because

$$g_{m\,g,b} = \beta \{ V_g' - 2\phi^{1/2} (V_b')^{1/2} \} \quad \text{for} \quad V_d \geq V_p \qquad (7\text{-}21)$$

b) Fig 7-18 shows measured and calculated values of the saturuted transconductance g_m as a function of V_g for different values of V_b. Note that

$$g_m = \beta V_p = \beta (V_g' - 2\phi\theta) \quad \text{for} \quad V_d \geq V_p \qquad (7\text{-}22)$$

c) The experimental curves of Fig 7-19 show that the application of reverse bias on the substrate causes a substantial reduction of the drain saturation resistance and voltage gain.

d) The pinch off voltage of a MOS transistor controlled from the metal gate is a function of the substrate bias.

$$V_p = (V_g - V_T) - 2\phi\theta \qquad (7\text{-}23)$$

where θ is a function of V_b

$$\theta = \left[1 + (V_g' - V_b')/\phi \right]^{1/2} - 1 \qquad (7\text{-}24)$$

A convenient way of controlling this parameter is thus established. It is desirable to have a positive pinch off voltage (for an n-type MOS transistor) or threshold voltage in digital applications and this may be conveniently achieved with a negative bias on the substrate. Fig 7-20 shows a graph of the effective pinch off voltage V_g' as a function of substrata bias. For this experiment pinch off was defined as the gate voltage necessary to produce 10 μa of channel current.

7-6. Transfer Characteristics with the Metal Gate as the control electrode

Fig 7-21 compares the transfer characteristics of the depletion unit with that of the induced channel unit. An induced

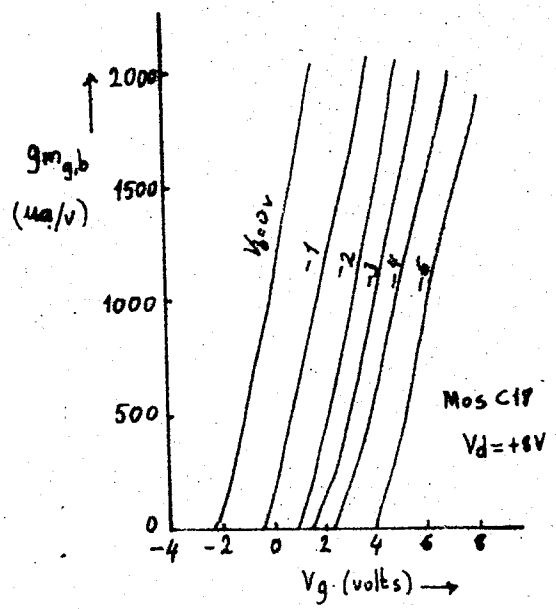


Fig.7-17. Measured values of the saturated transconductance $gm_{g,b}$ as a function of V_g for different values of V_b.

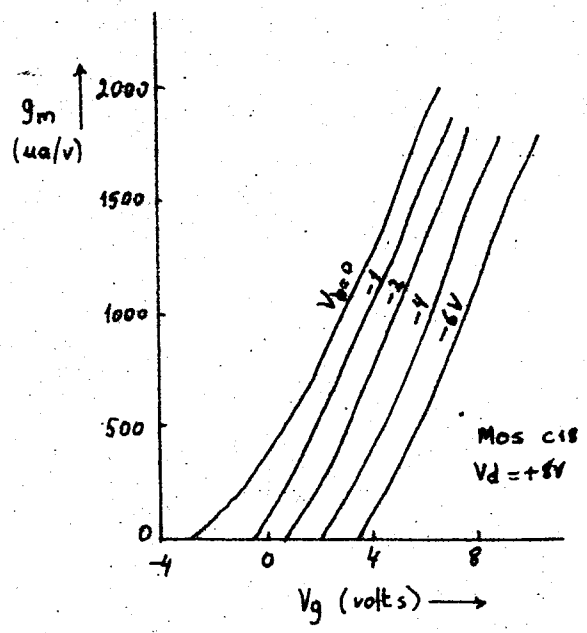


Fig.7-18 Measured values of the saturated transconductance g_m as a function of V_g for different values of V_b.

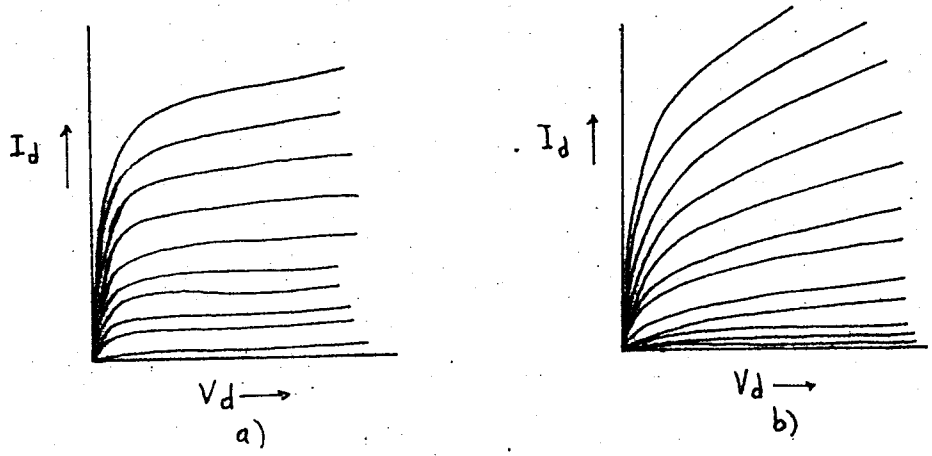


Fig.7-19. Output characteristics of an n-type MOS transistor showing the dependence of the saturation resistance and voltage gain on the substrate bias.
a) substrate bias = 0 v. b) substrate bias = 20 v.

channel unit may be transformed into a depletion unit by baking in dry hydrogen at 400°C for about 15 minutes. This process will introduce positive charge in SiO_2 layer which is compensated by negative charge in the Si surface of the p-type substrate. Prolonged time and/or higher temperature increases the charge density.

7-7. Instability of Characteristics

MOS transistors operating in the presence of high electric fields show instability of characteristics due to the motion of ions and molecules in SiO_2. Instability due to migration of ions in the bulk of the oxide layer is more pronounced in the enhancement transistors because these units are used mostly in digital circuits and operate with oxide fields greater than 10^6 V/cm which is close to the dielectric breakdown strength of the SiO_2 layer. The depletion transistors on the other hand are usually operated near zero gate voltage in small signal amplifier applications. Under such operation conditions the oxide field is on the order of 10^5 V/cm.

Much research is presently underway to understand the details of the observed instabilities in field effect transistors. If we assume that the oxide layer is filled with mobile, polyatomic molecules that may be easily dissociated, then when a positive potential is applied to the gate of an n-type transistor, the negative ions will tend to move toward the gate electrode and the positive ions toward the Si surface as shown in Fig 7-22 With the gate returned to zero potential, there will remain an electric field pointing into the Si which will terminate on negative charge. Thus when the gate bias is removed from an enhancement transistor that has been on for a long time, the transistor does not turn off. Whether the moving charges are positive negative or both is yet uncertain. However the external effect may be understood by examining the transfer characteristic of an n-type enhancement transistor before and after the application of a large positive gate bias for an extended period of time . (See Fig 7-23). After the application of gate voltage, the curve

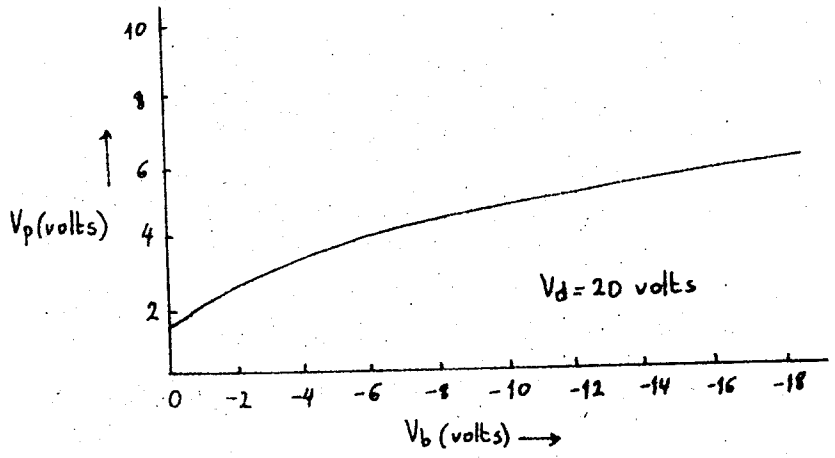


Fig.7-20. Effective pinch-off voltage of an enhancement type MOS transistor as a function of substrate bias.

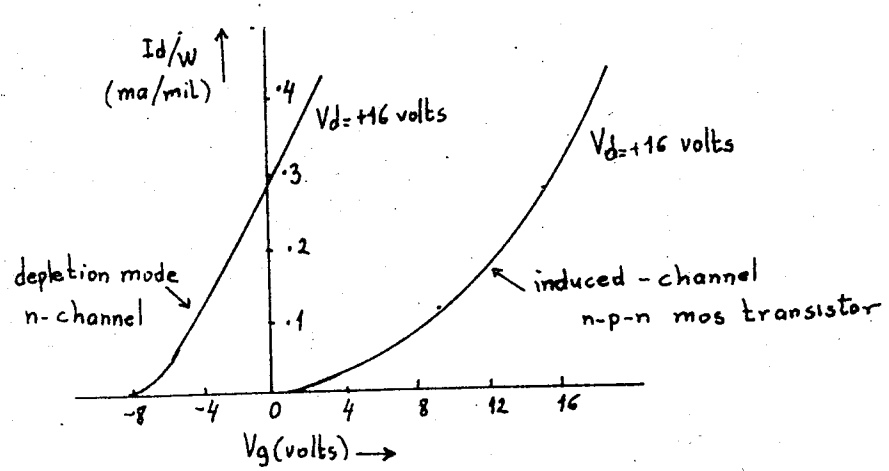


Fig.7-21. Comparison of transfer characteristics of MOS transistors. The drain current is normalized to unit channel width . (n-p-n MOS transistor means n-type source and drain with p-type substrate enhancement unit)

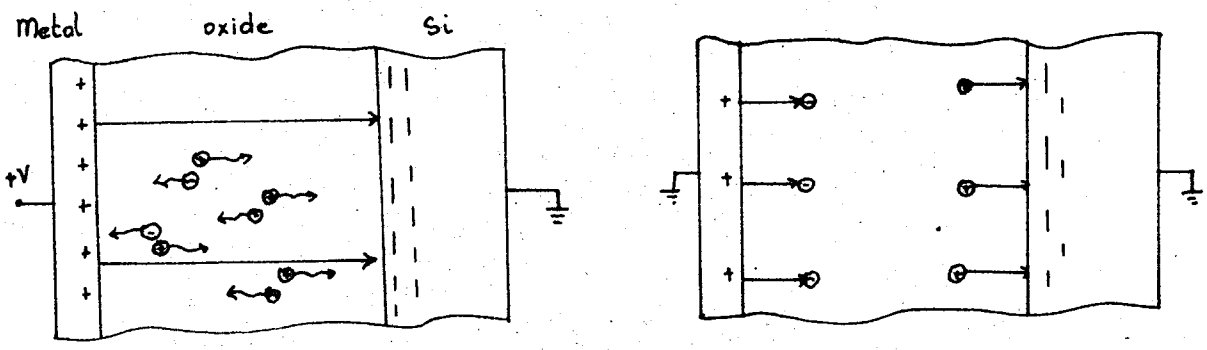


Fig.7.22. Migration of ions in the bulk of the SiO_2 layer cause the transistor characteristics to drift.
a) shows how the ions drift with the application of a potential.
b) The gate potential is reduced to zero but a finite electric field exists.

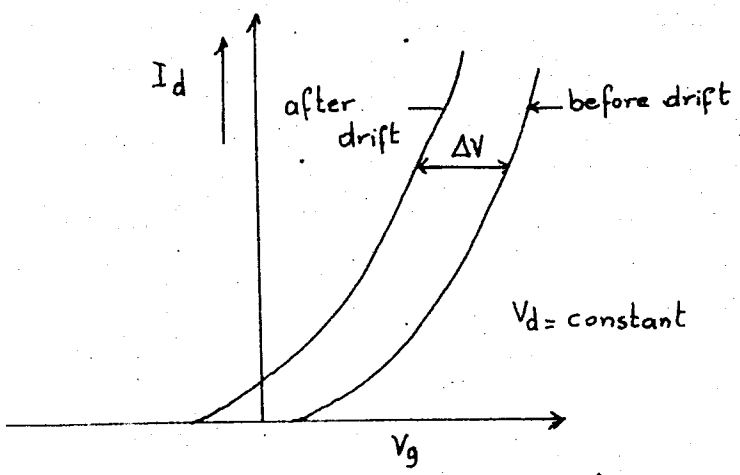


Fig.7.23. Shift in the transfer characteristic of an n-type enhancement transistor caused by the extended application of positive gate bias.

drifts to the left by an amount ΔV. The magnitude of the voltage shift ΔV increases with temperature, time of application of the gate voltage, and electric field strength. In some units the drift voltage ΔV saturates, but in others ΔV seems to continually increase with time. It is found out that the addition of phosphorous to the insulator, gives a glass like structure to the oxide and reduces drift by reducing the mobility of the ions.

A similar drift effect occurs with depletion transistors. However since these units operate at a low oxide field, ion migration is not the problem here but rather the offset gate geometry. An equivalent circuit demonstrating drift characteristics of an offset gate geometry transistor is shown in Fig 7-24. An RC network exist between gate and drain. With the gate negative and the drain positive, a potential gradient exists across the resistive surface and current starts to flow. This charges the distributed oxide capacitance in a time which may be from seconds to weeks depending on the value of the surface resistance. The surface next to the gate electrode is negatively charged. Causing depletion of channel charge beyond the gate geometry. When the gate potential is returned to zero, this charge continues to exist and a portion of the channel remains partially depleted, adding a series resistance to the channel. This lowers the zero bias channel current when the gate potential is reduced to zero after long negative bias. Proper packaging and surface treatments can reduce this form of instability.

7-8. Temperature Dependence of MOS Transistors.

There are two ways in which the temperature can influence the drain current of the field effect transistors.

a) negative coefficient mechanism : The temperature dependence of the surface mobility of electrons in Si has been found by Fang and Triebwasser to be $T^{-.403}$ above 200°K. This effect is highly dependent on surface treatment and oxidation conditions, and a decrease of mobility is to be expected with increasing temperature for a wide variety of preparations. This decrease of mobility

is due to the surface state effects.

b) positive coefficient mechanism: This effect is due to a change in surface potential caused by the change in temperature. This change in surface potential due to a change in the trap population for insulated gate transistors works in opposition to the mobility change. Experimental insulated gate transistors using various diffusion techniques exhibit negative, zero, and positive temperature coefficients. Fig 7-25 shows the transfer characteristics of an n-channel MOS transistor with temperature as the parameter. The drain current of this device exhibits a positive temperature coefficient.If the drain current is held constant by varying the gate voltage while the temperature rises, we see that the device exhibits zero temperature coefficient at some drain current. Fig.7-26 shows the plot of the data of such an experiment. It can be seen that at I_d=704A , this device shows zero temperature coefficient. This can be explained as follows: For higher levels of drain current, the Fermi level is closer to the conduction band.Chemical potential, $\emptyset = (E_c - E_g)$, decreases logarithmically for increasing drain current at a fixed temperature. As temperature increases, the rate of increase in chemical potential is a linear function of its initial value. Thus the rate of increase in chemical potential per unit temperature increase decreases with the drain current level. Since the product of the rate of increase in chemical potential and the surface state density sets the number of electrons donated to the conduction band per unit temperature increase, fewer electrons are donated at higher levels than at lower levels of drain current. The loss in number of electrons per unit increase in temperature due to mobility increases with drain current, the number available per unit temperature increase from surface states decreases with drain current. There exists a surface state density such that below some level of drain current, the transistor exhibits a negative temperature coefficient, above the same level a positive temperature coefficient and zero coefficient at this compensation level. The zero coefficient region is fairly broad since these

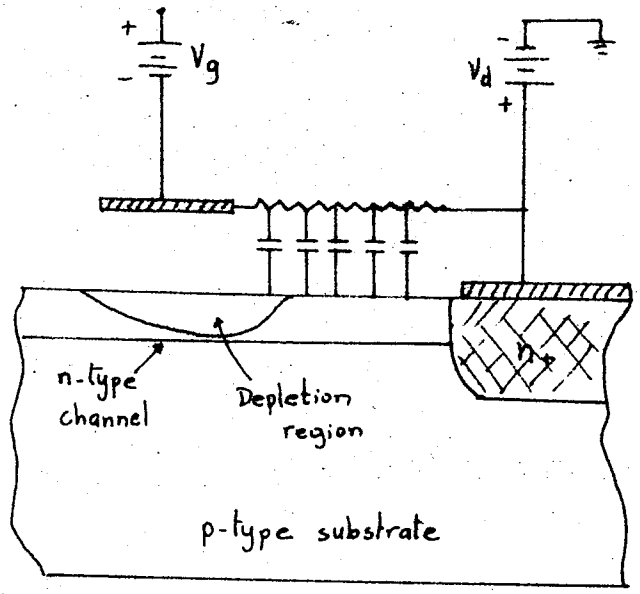


Fig.7-24. Equivalent circuit demonstrating drift characteristics of an off set gate geometry transistor.

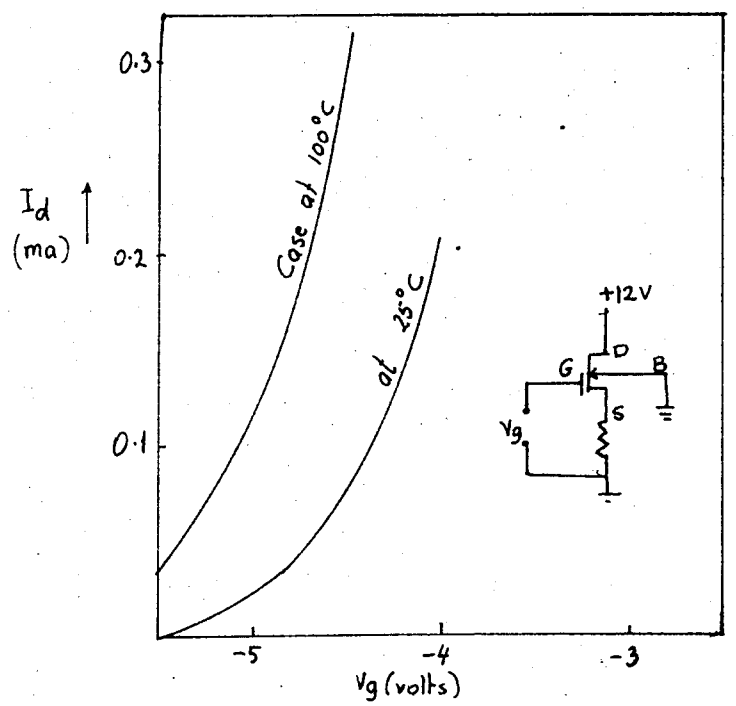


Fig.7-25 Low-current transfer characteristics of an n-channel MOS transistor as a function of temperature.

functions are slowly varying. Fig.7-27 shows the typical linear variation of transconductance with temperature for MOS transistors The gate leakage current is typically much less than 0.1×10^{-12}A with a gate-to-source voltage of 6 V at room temperature for an insulated gate transistor. It remains stable up to about 85-100°C, but after that it may increase to higher values of the order of 2×10^{-9}A.

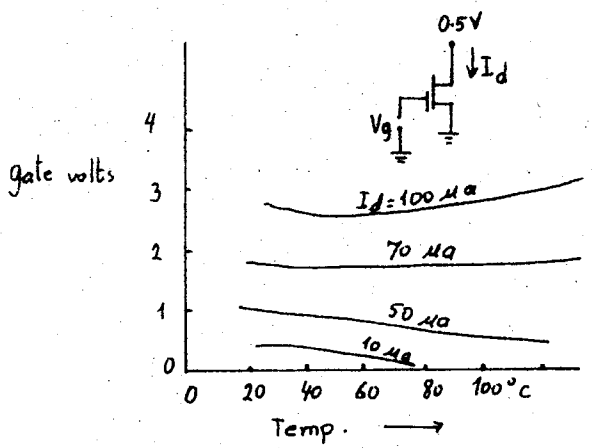


Fig.7-26. Gate voltage as a function of temperature for a device with zero temperature coefficient at 70μA.

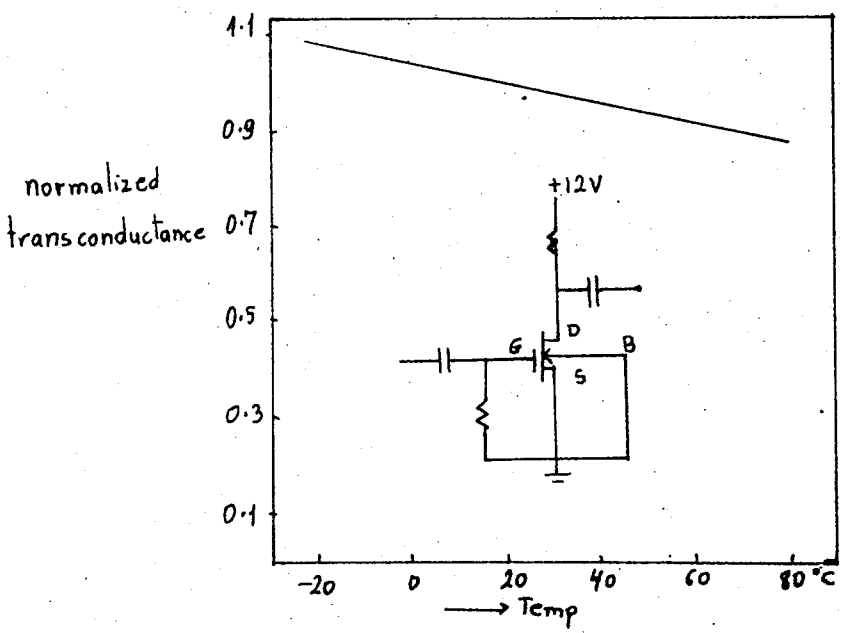


Fig 7-27. Relative transconductance of an n-channel MOS transistor as a function of temperature.

CHAPTER 8

SMALL SIGNAL EQUIVALENT CIRCUIT OF MOS TRANSISTOR

The MOS transistor is most generally used in the common source configuration. A common source equivalent circuit of the MOS transistor can be seen in Fig.8-1 in the most general form . This equivalent circuit is useful to approximately 100 mc/s .At higher frequencies the circuit of Fig. 8-1 must be modified by the addition of inductive elements to account for the lead inductances which may resonate with device capacitances.

The equivalent circuit shown in Fig 8-1 reflects the physics of the device. It is useful to summarize here the device parameters as follows:

8-1 Device parameters:

a) Leakage resistances :

rgs = leakage resistance through the insulating oxide from gate to source.

rgd = leakage resistance through the insulating oxide from gate to drain.

These leakage resistances are very high. They are of the order of 10^{12} to 10^{17} ohms.

b) Intrinsic channel capacitance:

Cc = Sum total of the small capacitors distributed along the active channel. It is equal to Cox discussed in the theory of the MOS transistor intergrated for the total unit $\left[\text{Cox}WL\right]$.

rc = effective gate series resistance

rc and Cc determine the time constant τ_r which governs the high frequency performance.

c) Transconductance

gm = the slope of the Id - Vg transfer characteristic curves at low frequencies. In the pinch off region the drain current is relatively constant and can be represented by a constant generat

gme_s .

$$gm = \beta(Vg - V_T) \qquad (8\text{-}1)$$

since $$Id_s \cong \beta \frac{(Vg - V_T)^2}{2}$$

$$gm = \sqrt{2Id_s\beta} \qquad (8\text{-}2)$$

d) Output resistance

rd_s = the dynamic output resistance of the transistor. Since rd_s is much larger than Rd and Rs (the parasitic drain and source resistances), it can be approximated by the slope of tne output characteristic curves.

e) Parasitic resistances

R'_d = portion of source to drain channel resistance due to the off-set of the insulated gate away from the drain.

R'_s = portion of source to drain resistance due to bulk resistance or contact resistance in the source plus any unmodulated channel caused by misalignment of the insulated gate.

In a typical amplifier type depletion MOS transistor, the sum of R'_d, R'_s and r_{ds} may be varied from a low value of about 300 ohms (in the ohmic region of operation) to a high value of several hundred megohms(in the pinch off region).

f) Input resistance:

r_{gs} = gate to source leakage resistance. This is a large resistance of the order of 10^{14} - 10^{16} ohms.

g) Capacitances:

C_{gs} = the gate to source capacitance is the parallel plate capacitance between the metal gate and the Si. This has a value of about 1-3pf and is independent of bias conditions for normal operations.

C_{gd} = feedback capacitance. It is due to the slight overlap of the metal gate electrode over the drain electrode. This capacitance increases slightly with the drain voltage since at the

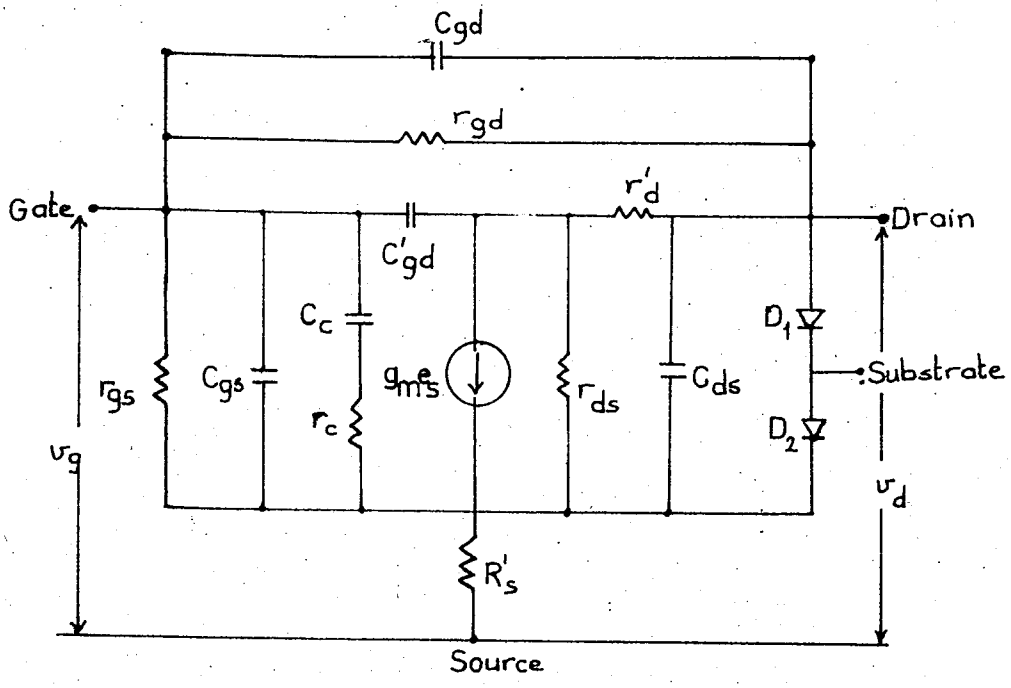


Fig 8-1. Common source equivalent circuit of the MOS transistor

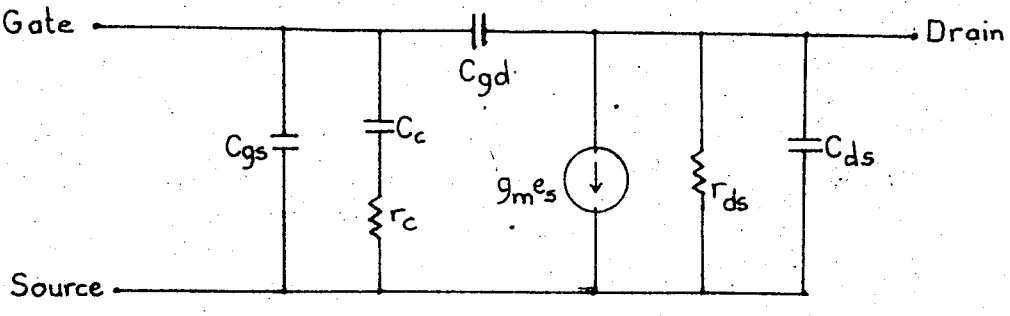


Fig 8-2. Simplified equivalent circuit of the MOS transistor

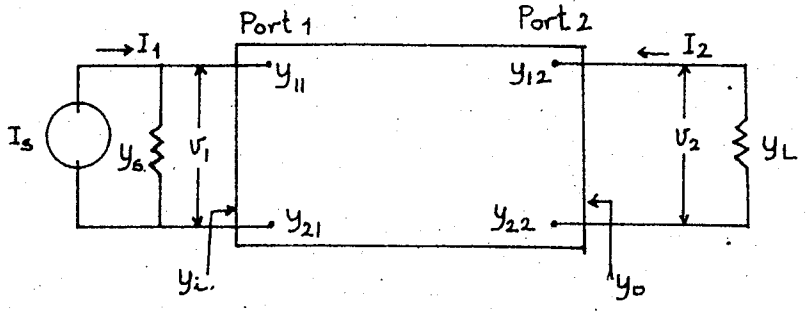


Fig 8-3. Two port admittance representation of a network.

higher voltage the depletion layer spreads further under the gate.
C'_{gd} = intrinsic gate-to - drain capacitance. This capacitance decreases with the increasing drain voltage in an n-channel off set gate MOS transistor. The overall effect is that larger value of stable voltage gain can be achieved in the off set gate transistor than in a full-gate device having comparable value of trans conductance.
C_{ds} = drain to source capacitance. It is mainly determined by the diffused pn junction capacitance between the drain electrode and the substrate and is a function of the voltage. A typical value is 2.5 pf at 20 V. C_{ds} is independent of the drain current.

h) Substrate diodes D_1 and D_2.

The diode D_1 is the junction formed between the heavily diffused source region and the substrate. D_1 and D_2 are back to back diodes in shunt with the channel. In some amplifier applications the anodes of D_1 and D_2 are connected to the source .In such a case D_1 becomes reverse biased while D_2 is shorted. A distributed bulk resistance associated with these diodes prevent their being shorted completely by an external short. D_1 and D_2 have negligible effects at low frequencies. At high frequencies however they contribute a series RC network and thereby change the effective output admittance.

As it can be seen from the characteristic values of the device parameters summarized above, the equivalent circuit of Fig.8-1 can be simplified as seen in Fig. 8-2. The circuit elements of this equivalent circuit are usually independent of frequency over a wide range. The MOS transistor is a three terminal device (when the base electrode is not used). One of the electrodes is common to the input and output circuits. The source is most frequently used as the common electrode in describing the field effect transistor.

8-2. Equivalent Circuit Parameter Determination.

8-2A. Admittance parameters.

A convenient method used for linear circuit analysis is the two port admittance method. Fig.8-3 shows a two-port admittance representation of a network. The relationship between the input and output currents and voltages is given by:

$$i_1 = y_{11}v_1 + y_{12}v_2$$
$$i_2 = y_{21}v_1 + y_{22}v_2 \qquad (8-4)$$

The input admittance y_{11} is the admittance of port 1, when port 2 is short-circuited ($v_2=0$).

y_{12} is the transfer admittance from port 2 to port 1 when port 1 is short-circuited .($v_1 = 0$).

y_{21} is the transfer admittance from port 1 to port 2 when port 2 is short-circuited($v_2=0$) = forward transfer admittance

y_{22} is the output admittance. It is the admittance of port 2 when port 1 is short-circuited ($v_1=0$).

However, when stating the above characteristic set of admittance parameters, the choice of the common electrode must be specified. For example, the set of common-source admittances will be written as :

$y_{is} = g_{is} + jb_{is}$ = input admittance source configuration

$y_{rs} = g_{rs} + jb_{rs}$ = reverse transfer admittance (8-5)

$y_{fs} = g_{fs} + jb_{fs}$ = forward transfer admittance.

$y_{os} = g_{os} + jb_{os}$ = output admittance

where the subscript s denotes the common source. Subscripts g and d will be used to denote common gate and common drain parameters respectively. In the analysis to follow, the general relations will be written without the configuration subscript. By calculating the two-port admittance for the common source circuit of Fig.8-2. we can express the frequency dependent variation of the admittance parameters in terms of the equivalent circuit elements as :

$$y_{is} = \frac{r_c}{r_c^2 + \left(\frac{1}{wC_c}\right)^2} + jw\left(C_{gs} + C_{gd} + \frac{C_c}{1 + (wr_cC_c)^2}\right)$$

$$y_{rs} = -jwC_{gd}$$

$$y_{fs} = \frac{g_m}{1 + jwr_cC_c} - jwC_{gd} \qquad (8\text{-}6)$$

$$y_{os} = \frac{1}{r_{ds}} + jw(C_{gd}+C_{ds})$$

8-2B. Deductions from Admittance data :

Admittance plots obtained by measurements over a wide range of frequencies may be used for analysis at low and high frequencies. Fig.(8-4) shows the two port common source admittances of an experimental insulated gate field effect transistor. A comparison of these curves with those obtained for the junction gate FET show similarities in the general appearance.

1) $y_{is} = g_{is} + jb_{is}$

From Eq. 8-6

$$y_{is} = \frac{r_c(wC_c)^2}{(r_cwC_c)^2+1} + jw\left(C_{gs}+C_{gd}+\frac{C_c}{1+(wr_cC_c)^2}\right)$$

at low frequencies $(r_cwC_c)^2 \ll 1$

$$g_{is} \cong r_cw^2C_c^2 \qquad (8\text{-}7)$$

$$b_{is} = w(C_{gs}+C_{gd}+C_c)$$

The input admittance y_{is} remains suspective up to several hundred mc/s since the absolute value of $b_{is} > g_{is}$.

2) $y_{rs} = g_{rs} + jb_{rs}$

The reverse transfer admittance y_{rs} is dominated by C_{gd} up to about 300 mc/s. The conductive component is neglegible at all frequencies, and the suspective component varies linearly with frequency.

$$jb_{rs} = -jwC_{dg} \qquad (8\text{-}8)$$

The slope of the plot of b_{rs} can be used to evaluate C_{dg}.

3) $y_{fs} = g_{fs} + jb_{fs}$

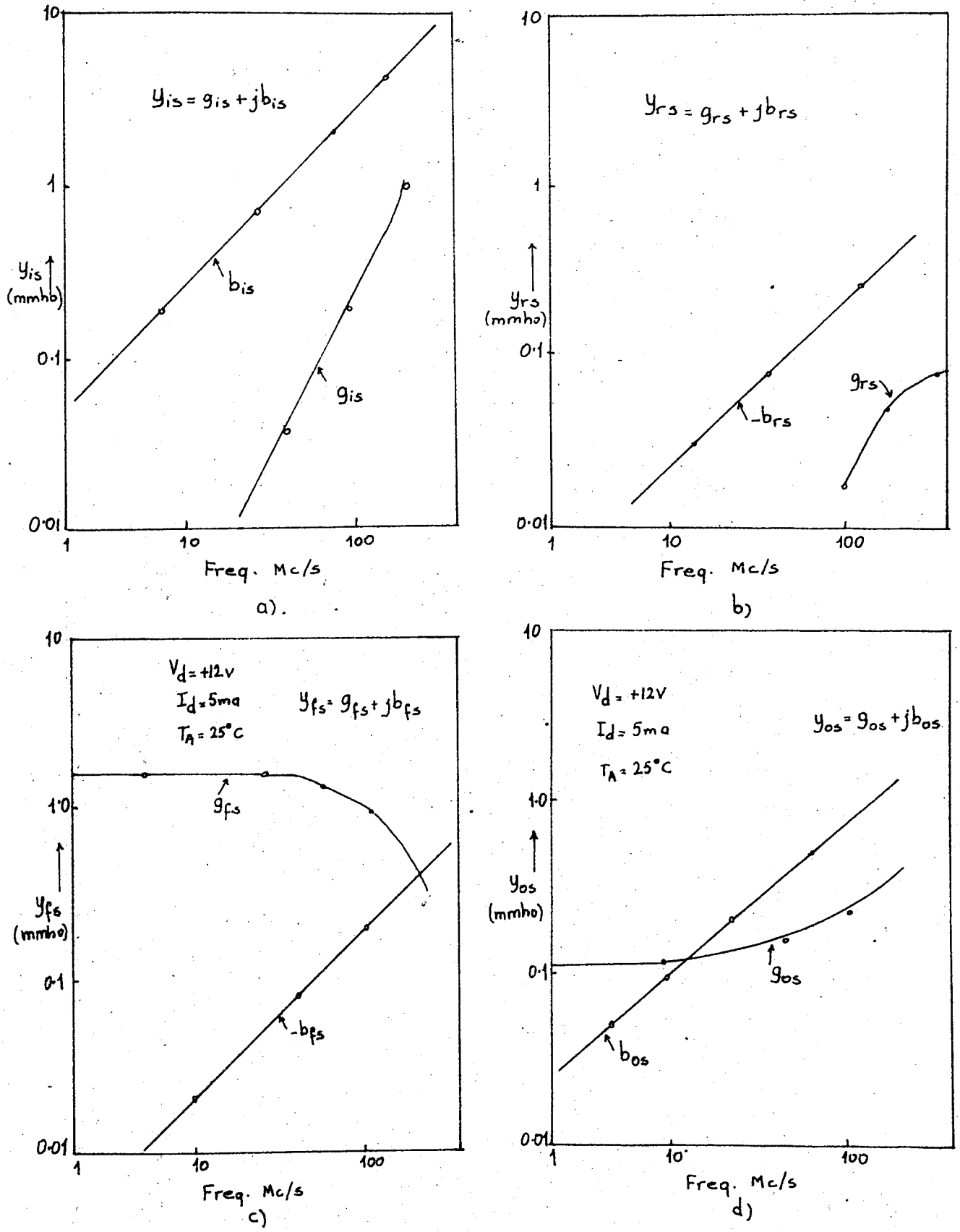


Fig 8-4. Common source admittances of a MOS transistor as a function of frequency

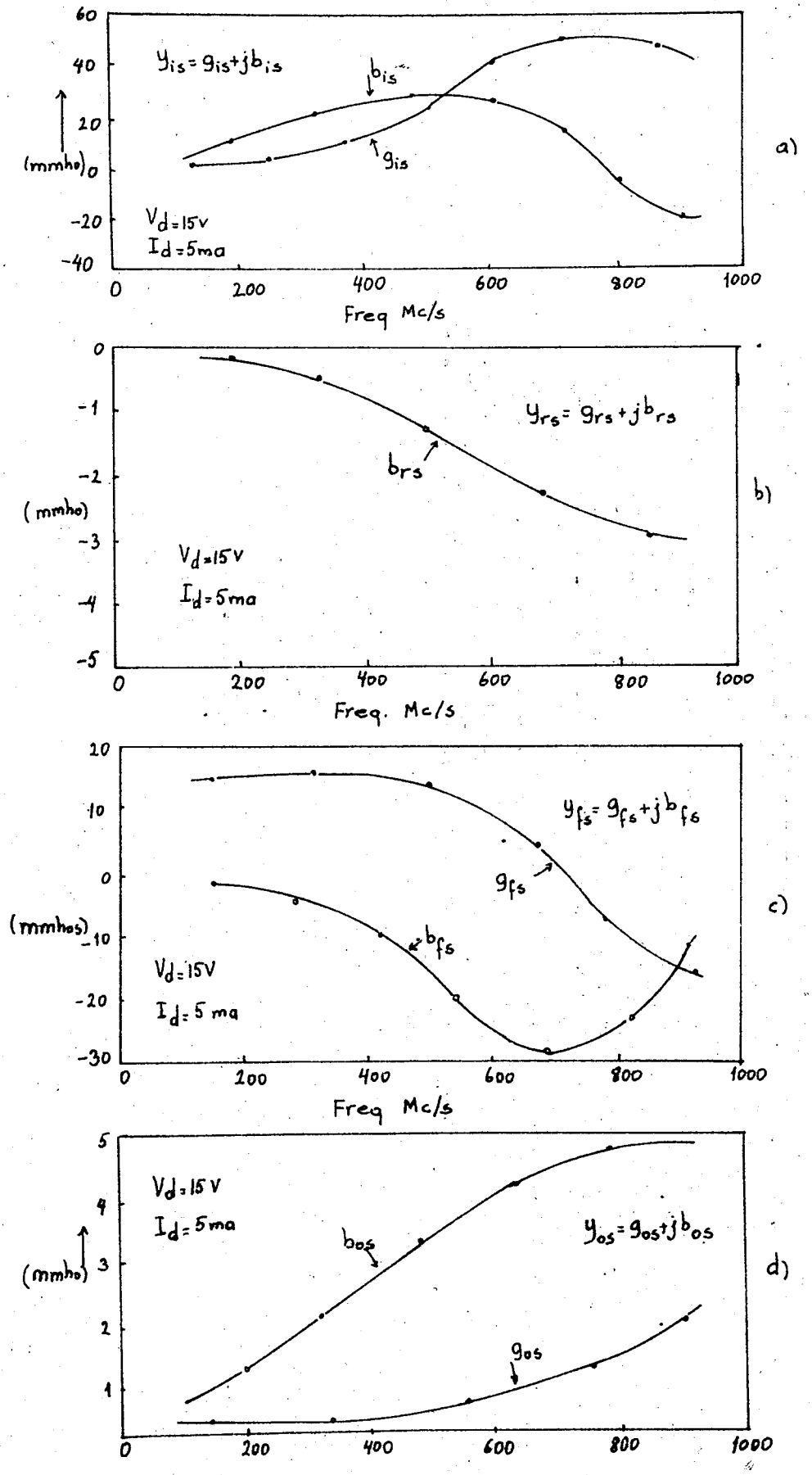


Fig 8-5. Common source admittance of a uhf MOS transistor

The forward transadmittance y_{fs} is essentially constant and real up to about 60 mc/s. At higher frequencies the susceptive part b_{fs} increases. The real part g_{fs} decreases in a fashion which can be fitted with an equation of the form

$$g_m = \frac{g_{mo}}{1 + jw\tau} \qquad (8\text{-}9)$$

where $\tau = r_c C_c$

g_{mo}= value of g_m at low frequencies.

Thus Fig.(8-4) can be used to obtain the value of g_{mo} and

4) $y_{os} = g_{os} + jb_{os}$

The suspective component b_{os} varies linearly with frequency as expected from the model of Fig.8-2. and Eq.8-6. The slope of the plot of b_{os} can be used to calculate the value of $C_{ds} + C_{gd}$. g_{os} is constant at low frequencies and is equal to $1/r_{ds}$. But at higher frequencies it starts to rise.

8-3 MOS transistor Admittance parameters at very high frequencies

At very high frequencies, about 400mc/s, deviations from the simple low frequency theory become evident in the two-port common source admittances. Fig. 8-5 shows the admittance parameters of an experimental MOS transistor as a function of frequency in the uhf range. Only the output admittance y_{os} is well behaved. The experimental data indicates that the simplified equivalent circuit of Fig.8-2 is no longer adequate at these frequencies.

Table 1 gives a good idea about the general characteristics of the three different configurations of the MOS transistor. As can be seen from the values given by this table, the transconductance g_f is essentially independent of the configuration. The common gate configuration shows a high input admittance like the grounded grid vacuum tube. The common-drain configuration has a low input admittance and high output admittance like a cathode follower.

Admittance (μmhos)	f = 10 mc/s		f = 20 mc/s		f = 40 mc/s		f = 80 mc/s	
	g	b	g	b	g	b	g	b
y_{is}	2	150	7	300	25	600	86	1200
y_{ig}	1650	215	1650	390	1700	710	1750	1300
y_{id}	4	150	10	300	28	600	82	1150
y_{rs}	—	−8	—	−16	—	−32	—	−66
y_{rg}	60	55	72	97	93	180	140	320
y_{rd}	5	140	10	280	24	570	76	1120
y_{fs}	1600	−15	1600	−36	1600	−90	1500	−220
y_{fg}	1650	110	1650	110	1650	120	1600	160
y_{fd}	1600	150	1600	290	1600	540	1600	1000
y_{os}	60	65	72	115	94	215	140	390
y_{og}	60	62	70	110	92	210	140	380
y_{od}	1650	200	1650	390	1700	720	1750	1350

Table I : MOS transistor admittances for CS, CG, CD configurations.

THESIS
ROBERT COLLEGE GRADUATE SCHOOL
BEBEK, ISTANBUL

CHAPTER 9

NOISE CHARACTERİSTİCS OF FIELD EFFECT TRANSİSTOR

9-1. Introduction

Thermal noise and shot noise are the main two kinds of noises encountered in the FETs at moderate and high frequencies. Thermal noise is caused by thermal agitation of charge carriers and can be represented by a mean square noise voltage of the form

$$\overline{v^2} = 4kTR\Delta f \qquad (9\text{-}1)$$

where k = Botzmann's constant

T = degrees in Kelvin

R = resistance of the material

Δf = band of the frequencies over which the noise is measured

This is known as "white" noise since it has a constant frequency spectrum, that is , for a constant bandwidth , Δf , $\overline{v^2}$ does not depend on frequency.

Shot noise is produced wherever a current is flowing in the transistor, and it also is " white" noise . The current is carried by individual carriers (holes or electrons) ,the charge of each carrier constituting a current pulse of small duration . Therefore the current fluctuates slightly about its mean value. The mean square fluctuation in the current is

$$\overline{i^2} = 2qI\Delta f \qquad (9\text{-}2)$$

where I is the average current. If this mean square current is flowing through a resistance R , it will produce a mean square voltage across its terminals:

$$\overline{v^2} = 2qIR^2\Delta f \qquad (9\text{-}3)$$

At low frequencies (~1Kc and lower), the thermal and shot noise are overshadowed by another type of noise which is often called excess noise. The mechanism of excess noise involves two step processes with fluctuations characterized by exponential decay with time constant τ_t. Examples of two step processes are: 1) random emission of holes and electrons by SRH centers , 2) The trapping and releasing of conduction electrons by the majority donor impurities causing a fluctuation of the carrier concentration in a neutral n-type semiconductor , 3) a process similar to 2 , but occuring in the transition region of a p-n junction , 4) recombination through SRH centers where the electron and hole capture processes occur. The time constant for each of the above cases is different. For example , in case 1) , the time constant is given by

$$\tau_t = 1 / [\, c_p p_1 + c_n n_1 \,] \qquad (9\text{-}4)$$

where c_p = capture probability of a hole by the SRH center
c_n = capture probability of an electron by the SRH center
$p_1 = n_i \exp [(E_i - E_t)/kT]$
n_i = carrier concentration in an intrinsic semiconductor specimen
E_i = Fermi energy level for an intrinsic specimen
E_t = The energy level of the SRH centers. $E_v < E_t < E_c$
E_v = top of valence band edge
E_c = bottom of the conduction band edge.
τ_t is of the order of 10^{-2} sec.

For cases 2) and 3) the time constant is the trapping time of electrons by the donors and is of the order of 10^{-11} sec at room temperature. For case 4) the time constant is the steady state SRH lifetime of the minority carriers which is of the order of 10^{-6} to 10^{-9} sec , depending on the concentration of the SRH centers. The general frequency spectrum of the mean square fluctuation of the trapped electron concentration in an elemental volume ΔV is calculated from the Weiner-Khintchine theorem , and is :

$$S_{\delta n t}(f)\,\Delta f = 4\Delta f \, \frac{\tau_t}{1+\omega^2\tau_t^2} \cdot \frac{N_t f_t f_{tp}}{\Delta V} \qquad (9\text{-}5)$$

where Δf = bandwidth

N_t = concentration of SRH centers

f_t = fraction of occupied SRH centers(by holes or electrons)

f_{t_p} = fraction of empty SRH centers.

Fig 9-1 shows the electron and hole capture and emission process at the Shockley-Read-Hall-Centers (SRH centers). SRH centers are recombination-generation centers found in the transition region of junction. These centers are the dominant sources of the reverse leakage current of a Si diode. The predominant factor of low - frequency noise comes from the fluctuation of the charge states of the SRH centers due to the random generation of electrons and holes from these centers as shown in Fig 9-1. In the transition region of a reverse biased PN junction, the electron and hole emission processes labeled b) and d) in Fig. 9-1 are the most important ones since the high electric field completely depletes the electrons and holes in this region making the capture process improbable.

At very low frequencies, the excess noise tends to 1/f noise spectrum. Therefore excess noise is usually referred to as 1/f noise. Whatever the cause may be, the effect of 1/f (excess) noise may be analyzed in terms of a single elementary event. A carrier of charge $\pm q$ becomes available to take part in the conduction process at time t_n and is withdrawn at time $t_n + \tau_n$. The change in conduction caused by creation and annihilation of a single carrier is illustrated in Fig 9-2. Here G_o represents the average value of conductance and ΔG the small increment occuring when the additional carrier exists. ΔG will be negative in the case of trapping an available carrier followed by an escape from the trap.

Fortunately the shot noise of the gate current and the excess noises can be reduced by the choise of the transistor material , by improved construction of the device and by proper treatment of the finished product. However the thermal noise of the conducting channel and the capacitive gate current resulting

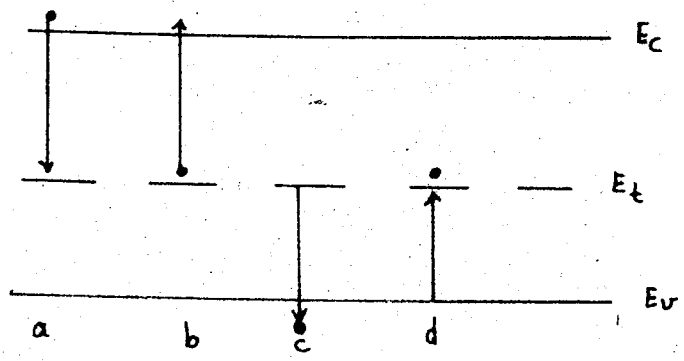


Fig.1 Electron and hole capture and emission processes at the SRH centers b and d show electron and hole emission processes respectively.

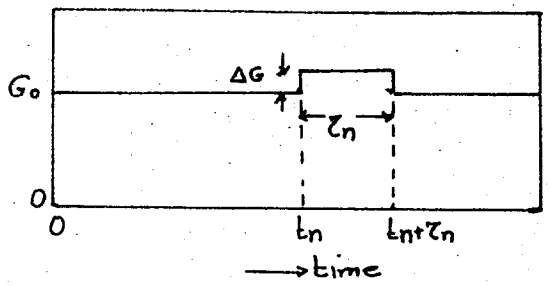


Fig.2 Change in conductance caused by creation and annihilation of a single carrier.

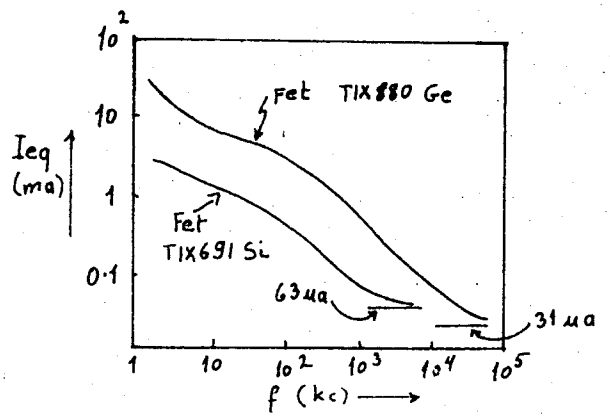


Fig.13(a) Equivalent noise current Ieq vs frequency operating in saturation.

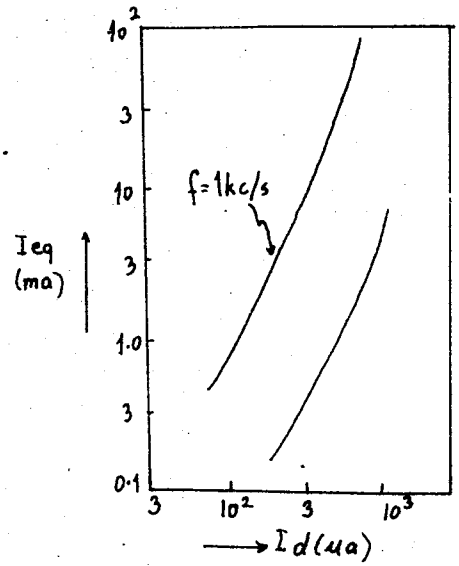


Fig 13(b) Equivalent noise current Ieq vs drain current I_d with frequency as a parameter.

from it are always present. They determine the lowest noise figure that can be obtained with field effect transistors.

When a transistor has to amplify weak signals, one must examine whether the signal to be amplified is sufficiently above the noise level of the transistor itself . In order to describe the noise, either sufficient number of noise sources as voltage or current sources are inserted into an appropriately chosen equivalent circuit (a linear two-port equivalent), or the physical reasons for the noise are examined and introduced into an equivalent circuit. First we shall discuss the second method as developed by Sah and Van der Ziel for the junction-gate FETs and by Jordan and Jordan for the MOS transistors.

9-2. Low-frequency Noise Generation in Junction Gate FETs

A general solution for noise current at low frequencies has been derived by C.T.Sah by making use of Shockley's gradual channel model. Sah has divided the generated noise into two: Gate junction noise and channel noise.

9-2A. Gate Junction Noise

It is mainly excess noise which shows an approximate $K/1+\omega^2\tau^2$ frequency dependence. By detailed analysis and using the Wiener-Khintchine theorem, Sah has derived an exact expression for the short circuit mean square noise current (with both the gate and drain electrodes ac short circuited to the source) in the drain, arising from the SRH centers in the gate-junction transition region which modulate the width of the channel.

$$i_{dn}^2 = \Delta f \left[\tau_t/(1+\omega^2\tau_t^2)\right](q/K\epsilon_o)^2\left(a^3/LZ\right)\left(N_t f_t f_{tp}\right)\left(I_{do}/V_{dso}\right)^2 I \qquad (9\text{-}6)$$

where I = an integral current directly proportional to the drain current I_d

I_{do} = dc drain current in the saturation region for $V_{gs}=0$

$= q^2\mu_n N_d^2 Z a^3/6K\epsilon_o L$

$V_{dso} = qN_d a^2/2K\epsilon_o =$ saturation drain voltage for $V_{gs}=0$

If a short-circuited output equivalent saturated diode noise current at the drain, designated by Ieq , is used,

$$i_{dn}^2 = 2q\,I_{eq}\,\Delta f$$

where

$$I_{eq} = \frac{1}{2q}\left[\tau_t/(1+\omega^2\tau_t^2)\right](q/K\epsilon_o)^2(a^3/LZ)(N_t f_t f_{tp})(I_{do}/V_{dso})^2 I \qquad (9\text{-}8)$$

9-2B. Channel Noise

At low frequencies the channel noise consists of thermal noise and excess noise due to carrier fluctuation. The thermal noise in the channel is best expressed by the equation derived by Van Der Ziel.

$$i_{dn}^2 = 4kT\,g_{max}\,Q\,(V_g, V_d)\,\Delta f \qquad (9\text{-}9)$$

The derivation of Eq.(9-9) will be given later when noise at moderately high frequencies will be discussed.

The excess channel noise originates from the random trapping and releasing of the current carriers (electrons in the n-type channel) at the SRH centers or shallow level majority impurities (donors in the n-type channel) in the channel. The mean-square channel noise due to carrier fluctuation is given by

$$i_{dn}^2 = 12\Delta f\left[\tau_c/(1+\omega^2\tau_c^2)\right](N_c f_c f_{cp}/N_d^2\,aLZ)(I_d/I_{do})(V_d/V_{dso}) \qquad (9\text{-}10)$$

where f_c = fraction of the fluctuating occupied impurity centers in the channel

f_{cp} = fraction of the fluctuating unoccupied impurity centers in the channel.

N_c = concentration of SRH centers in the channel.

9-2C. Other Noise Sources.

In a practical device there are several additional noise sources located at the various regions of a junction gate FET. But the noise contributed from these sources is substantially smaller than that from the channel and is therefore negligible.

The surface noise which is caused by the random generation and trapping of carriers by the surface states is an example. The magnitude of the surface noise is small in the junction gate FETs while it is the dominant low frequency noise source in the MOS transistors. Fig 9-3 shows typical noise curves for high noise units TIX69I Si and TIX880 Ge FETs. The transistors were operated in saturation with the gate dc connected to the source for this experiment . In Fig 9-3a the $K/1+\omega^2\tau^2$ dependence is evident with a tendency toward 1/f spectrum at low frequencies.At high frequencies Ieq for both units approaches a constant level that is in close agreement with the respective theoretical thermal noise level. Fig 9-3b shows the expected drain current dependence of the noise.

9-3. FET Noise at Moderately high Frequencies:

The dominant noises are thermal noise at the channel, shot noise at the gate and the capacitive gate current resulting from the thermal noise of the conducting channel.

9-3A. Thermal Noise in the Channel.

A schematic representation of the junction-gate FET is shown in Fig 9-4. Assuming that the dc current I flowing through the channel is constant, then a thermal noise voltage developed between x and $x + \Delta x$ will modulate the width of the channel between x and the drain and give an amplified noise voltage at the drain. By integrating over all sections Δx , the total output noise voltage can be obtained. Letting

$$W(x) = W_1 + \Delta w$$
$$b(x) = b_1 + \Delta b \qquad (9\text{-}11)$$

where W_1 and b_1 are dc values before perturbation for the bias across the junction and half channel width, respectively.

ΔW and Δb are the fluctuations due to thermal noise developed between x and $x + \Delta x$.

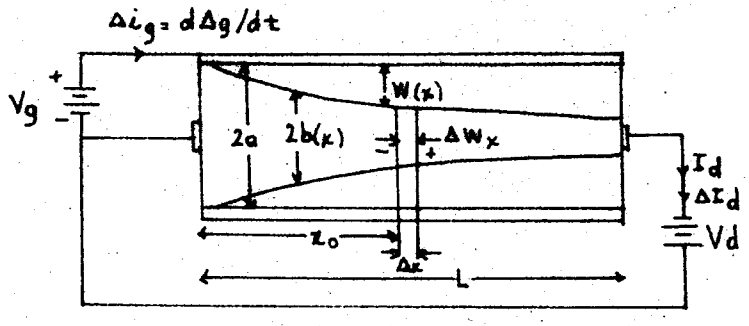


Fig.9-4: Cross section of a planar field effect transistor.

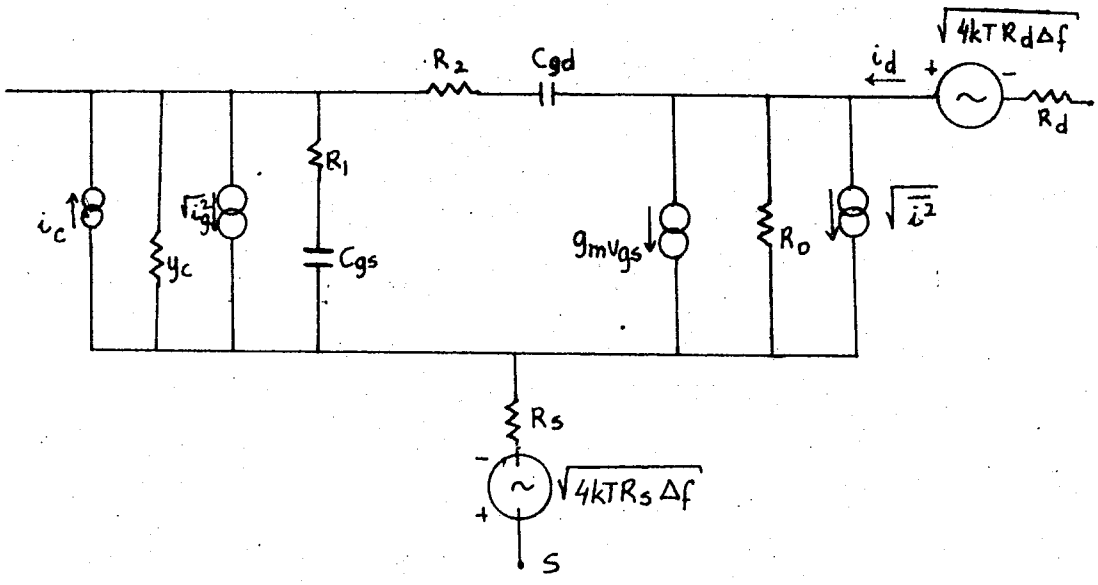


Fig.9-5: Equivalent circuit of a field effect transistor showing the noise sources.

From chapter 1, Eq(1-14).

$$W_1 = W_0 \left(1 - b_1/a\right)^2 \qquad (9\text{-}12)$$

So that

$$(W_1 + \Delta W) = W_0 \left[1 - \frac{(b_1 + \Delta b)}{a}\right]^2 \qquad (9\text{-}13)$$

expanding Eq(9-13) and neglecting the term with Δb^2

$$\Delta W = -2\,\frac{W_0}{a}\left(1 - \frac{b_1}{a}\right)\Delta b = -\frac{2}{a}\left(W_0 W_1\right)^{1/2}\Delta b \qquad (9\text{-}14)$$

from chapter 1, Eq.(1-18)

$$g(W) = g_0 \left[1 - \left(\frac{W}{W_0}\right)^{1/2}\right] \qquad (9\text{-}15)$$

and the general current expression is :

$$I = 2\sigma_0 b\,\frac{dW}{dx} = g(W)\,\frac{dW}{dx} \qquad (9\text{-}16)$$

From the continuity of current

$$I = 2\sigma_0 b_1 \frac{dW}{dx} = 2\sigma_0 (b_1 + \Delta b)\left(\frac{dW_1}{dx} + \frac{d\Delta W}{dx}\right) \qquad (9\text{-}17)$$

Since I is kept constant

$$\Delta b \frac{dW_1}{dx} + b_1 \frac{d\Delta W}{dx} = 0 \qquad (9\text{-}18)$$

Substituting Eq (9-14) into Eq(9-18)

$$\frac{d\Delta W}{dW} = \frac{1}{2}\,\frac{d\left(W_1/W_0\right)}{\left(W_1/W_0\right)^{1/2}\left[1 - \left(W_1/W_0\right)^{1/2}\right]} = \frac{du}{1-u} \qquad (9\text{-}19)$$

where $u = \left(W_1/W_0\right)^{1/2}$

Integrating between the limits x and L_1

$$\frac{\Delta W_d}{\Delta W_x} = \frac{1-u_x}{1-u_L} = \frac{1 - \left[W_1(x)/W_0\right]^{1/2}}{1 - \left(W_d/W_0\right)^{1/2}} \qquad (9\text{-}20)$$

where ΔW_x is the fluctuation in the section Δx and ΔW_d is the resulting fluctuation at the drain. $W_1(x)$ is the value of W_1 at the point x where the fluctuation occurs. Since ΔW_x is caused by thermal noise, its mean square value is:

$$\overline{\Delta W_x^2} = 4kT\,\Delta f\,\frac{\Delta x}{g(W_1)} = \frac{4kT\,\Delta f\,\Delta W_1}{g(W_1)\,dW_1/dx} \qquad (9\text{-}21)$$

$$= \frac{4kT\,\Delta f}{I}\,\Delta W_1$$

Thus

$$\overline{\Delta W_d^2} = \frac{4kT\,\Delta f}{I}\left[\frac{1-(W_1/W_o)^{1/2}}{1-(W_d/W_o)^{1/2}}\right]\Delta W_o \qquad (9\text{-}22)$$

The total mean square open-circuit noise voltage is obtained by integrating over the length of the sample, between the limits W_s and W_d .

$$\overline{e^2} = \frac{4kT\Delta f}{I}\,\frac{\left[(W_d-W_s)-4/3\,(W_d^{3/2}-W_s^{3/2})/W_o^{1/2}+1/2\,(W_d^2-W_s^2)/W_o\right]}{\left[1-(W_d/W_o)^{1/2}\right]} \qquad (9\text{-}23)$$

Then the short-circuit noise current mean square value becomes:

$$\overline{i^2} = \overline{e^2}\, g_d^2 \qquad (9\text{-}24)$$

Substituting the expression for g_d as given by Eq(i-28), and letting $x=W_d/W_o$, $y=W_s/W_o$, we get

$$\overline{i^2} = 4kT\,\frac{g_o}{L}\,\Delta f\,\frac{\left[(x-y)-4/3\,(x^{3/2}-y^{3/2})+1/2\,(x^2-y^2)\right]}{\left[(x-y)-2/3\,(x^{3/2}-y^{3/2})\right]} \qquad (9\text{-}25)$$

when $V_d \longrightarrow 0$ $\quad x \longrightarrow y$, and Eq.(9-25) becomes

$$\overline{i^2} = 4kT\,\frac{g_o}{L}\,\Delta f\,(1-y^{1/2}) = 4kT\,g_{do}\,\Delta f \qquad (9\text{-}26)$$

Eq(9-26) shows that the device gives thermal noise for drain bias since g_{do} is the ac output conductance for zero drain bias. Letting

$$Q(W_d, W_s) = Q(x,y) = \frac{\left[(x-y)-4/3\,(x^{3/2}-y^{3/2})+1/2\,(x^2-y^2)\right]}{(1-y^{1/2})\left[(x-y)-2/3\,(x^{3/2}-y^{3/2})\right]} \qquad (9\text{-}27)$$

Eq.(9-25) can be written as:

$$\overline{i^2} = 4kT\,g_{max}\,\Delta f\,Q(x,y) \qquad (9\text{-}28)$$

At pinch off $W_d = W_o$ and therefore $x=1$, and the current expression becomes

$$\overline{i^2} = 4kT\,g_{max}\,\Delta f\,Q(1,y) \qquad (9\text{-}29)$$

Here g_{max} = saturated value of the apparent transconductance g_m'

$$Q(1,y) \cong 0.6 - 0.67 \quad (\text{usually}) \qquad (9\text{-}30)$$

Eq.(9-29) is correct for the saturated part of the characteristic as long as the field strength in the cut off part of the channel is not large.

The thermal noise resistance R_n of the device may be found easily:

$$\overline{i^2} = 4kT R_n \Delta f g_m^2 \qquad (9-31)$$

Equating Eq.(9-31) and Eq.(9-28)

$$R_n = \frac{g_{max}}{g_m^2} Q(x,y) \qquad (9-32)$$

At saturation $g_{max} \longrightarrow g_m \longrightarrow R_n = \dfrac{Q(1,y)}{g_{max}}$ (9-33)

$$R_n \cong \frac{0.6}{g_{max}} \quad \text{at saturation}$$

The thermal noise current source and resistance expressions were derived using Shockley's ideal model shown in Fig 9-1. But in the actual FETs the gate contacts cover only part of the channel so that the apparent transconductance g_m' and apparent output conductance g_d' of the device should be used. For easy reference, the expressions for these apparent parameters will be given :

$$g_m' = \frac{g_m}{1 + R_s g_{max} + R_d g_d}$$

$$g_{max}' = \frac{g_{max}}{1 + R_s g_{max}} \qquad (9-34)$$

$$g_d' = \frac{g_d}{1 + R_s g_{max} + R_d g_d}$$

$$g_{do}' = \frac{g_{max}}{1 + R_s g_{max} + R_d g_{max}}$$

Fig 9-5 shows the equivalent circuit of a field-effect transistor including the thermal noise of the two series resistances R_s and R_d of the channel, and the mean square noise current source.

9-3B. Shot noise in the gate.

The shot noise is caused by the two currents I_1 and I_2 . I_1 is the current due to the electrons arriving at the gate and holes

leaving the gate . I_2 is the current due to electrons leaving the gate and holes arriving at the gate. Thus

$$\overline{i_g^2} = 2q\,(I_1+I_2)\,\Delta f \qquad (9\text{-}35)$$

9-3C. Capacitive noise current in the Gate

This noise can be represented by a current generator $\sqrt{i_c^2}$ between the gate and the source. This comes about because of the capacitive coupling between the channel and the gate. $\overline{i_c^2}$ is proportional to the square of the frequency and is larger than the shot noise component of i_g over a wide frequency range. The functional dependence of this current source may be expressed as:

$$\overline{i_c^2} = \frac{4kT\,\Delta f}{g_{max}}\,\omega^2 C_{gs}^2\, f(y,z) \qquad (9\text{-}36)$$

9-4 Noise in MOS transistors.

The main noise sources in MOS transistors are located in the channel and gate of the device like in FETs. Fig 9-6 shows a small-signal equivalent circuit representation of the MOS transistor with noise generators $\overline{i_d^2}$ in the output circuit and $\overline{i_g^2}$ in the gate circuit. Like in the FET, the principal contribution to $\overline{i_d^2}$ will be thermal carrier fluctuations in the channel. The principal contributions to $\overline{i_g^2}$ will be thermal fluctuations in the channel coupled into the gate circuit, shot noise of current flowing through the gate, fluctuations in occupation of surface states, and leakage noise. In addition parasitic series resistances will add thermal noise sources as indicated.

9-4A. Channel Noise

The thermal noise source expressions in the MOS channel are derived by a procedure similar to that of the FET channel noise . Fig 9-7 shows a MOS transistor which will be made use of in the following derivations. The current and voltage expressions found in the first-order theory of chapter 6 will be used. The necessary transistor relations which were derived in chapter 6 are

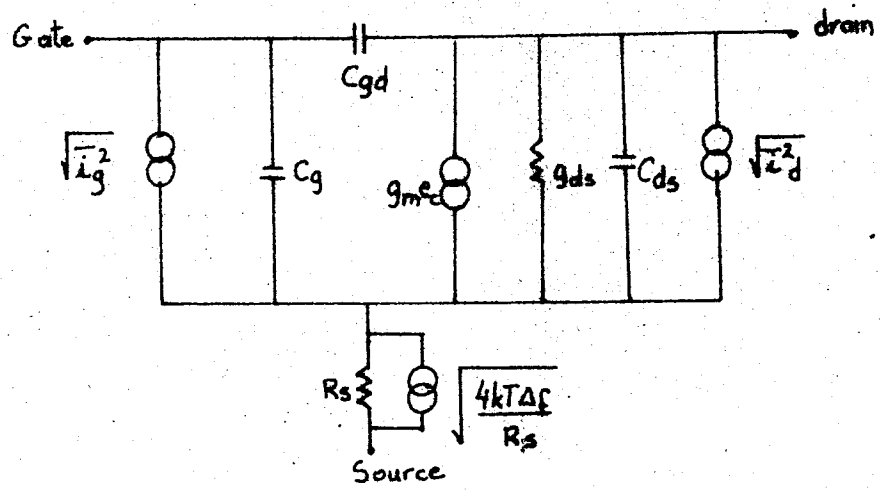


Fig.9-6: Equivalent circuit for the MOS transistor including noise generators.

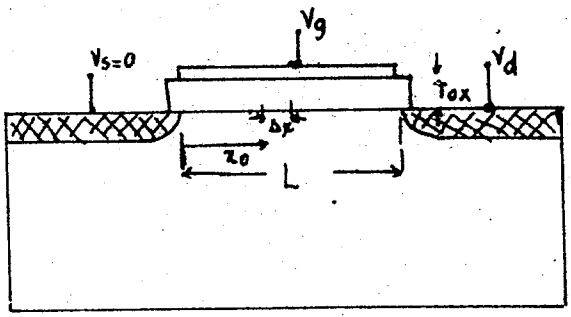


Fig 9-7: Model for an MOS transistor to be used in noise analysis.

repeated here for easy reference. The following quantities are for unit channel width. The surface charge density in the channel is:

$$Q(x) = \frac{\epsilon_{ox}}{T_{ox}} \left[V_g - V_T - V(x) \right] \quad \text{for} \quad V(x) < V_g - V_T \tag{9-37}$$

Note that for an inversion-type MOS device $V_T = V_p$, so that we can writte Eq.9-37 as

$$Q(x) = \frac{\epsilon_{ox}}{T_{ox}} \left[V_g - V_p - V(x) \right] \quad \text{for} \quad V(x) < V_g - V_p \tag{9-37a}$$

The surface conductivity of the channel will be

$$G(x) = \mu Q(x) \tag{9-38}$$

So that the differential channel resistance can be written as

$$dr_c = \frac{dx}{\mu Q(x)} \tag{9-39}$$

From Eq.(6-7) the channel current is

$$i_d = \frac{\mu \epsilon_{ox}}{T_{ox}} \left[V_g - V_p - V(x) \right] \frac{dV}{dx} \tag{9-40}$$

and the drain current is

$$I_d = \frac{\beta_s}{L} \left[(V_g - V_p) V_d - 1/2 V_d^2 \right] \tag{9-41}$$

where $\beta_s = \frac{\epsilon_{ox} \mu}{T_{ox}}$, and the parasitic source and drain resistance are neglected.

When the drain current saturates, $V_d = V_g - V_p$, so that the saturated drain current is:

$$I_{ds} = \frac{\beta_s}{2L} (V_g - V_p)^2 \tag{9-42}$$

and the saturated transconductance is:

$$g_{ms} = \frac{\beta_s}{L} (V_g - V_p) \tag{9-43}$$

The drain conductance dI_d / dV_d when $V_d \rightarrow 0$ becomes

$$g_{do} = \frac{\beta_s}{L} (V_g - V_p) = g_{ms} \tag{9-44}$$

when both $V_g = 0$ and $V_d = 0$, the drain conductance becomes

$$g_{doo} = -\frac{\beta_s}{L} V_p \tag{9-45}$$

To find the change in the drain current due to a voltage perturbation in the channel, we proceed as follows:

Let $V_o(x)$ = The equilibrium potential distribution in the channel

$V_1(x)$ = The potential distribution between $x = 0$ and $x = x_o$ due to a perturbation δV introduced at $x = x_o$

$V_2(x)$ = The potential distribution between $x = x_o$ and $x = L$ due to the perturbation

The current change due to the perturbation will be, by using Eq.(9-40)

$$\Delta i = \beta_s \left\{ \left[V_g - V_p - V_1(x) \right] \frac{dV_1}{dx} - \left[V_g - V_p - V_o(x) \right] \frac{dV_o}{dx} \right\} \qquad 0 < x < x_o$$

$$\Delta i = \beta_s \left\{ \left[V_g - V_p - V_2(x) \right] \frac{dV_2}{dx} - \left[V_g - V_p - V_o(x) \right] \frac{dV_o}{dx} \right\} \qquad x_o < x < L$$

Integrating Eq. 9-46 between the limits indicated, and adding, we get

$$\Delta i = -\delta V \frac{\beta_s}{L} \left\{ V_g - V_p - V(x_o) \right\} \qquad (9\text{-}47)$$

where $\delta V = V_2(x_o) - V_1(x_o)$ = the perturbation voltage.

The mean squared voltage fluctuation caused by the thermal noise fluctuation in the differential channel resistance dr_c is

$$\overline{\delta v^2} = 4kT \Delta f \, dr_c = 4kT \Delta f \frac{dV}{i_d} \qquad (9\text{-}48)$$

Since i_d is constant in the channel, the mean-squared current fluctuation in the drain current due to the voltage perturbation at $x = x_o$ becomes

$$\overline{\Delta i^2} = 4kT \Delta f \frac{\beta_s^2}{I_d L^2} \left\{ V_g - V_p - V(x_o) \right\}^2 dV \qquad (9\text{-}49)$$

Now Eq.(9-49) can be integrated over the entire channel to give the output noise current :

$$\overline{i_d^2} = 4kT \Delta f \, g_{ms} \, f_1(p) \qquad (9\text{-}50)$$

where $p = V_d / (V_g - V_p)$, and

$$f_1(p) = \frac{1 - p + 1/3\, p^2}{1 - 1/2\, p} \qquad (9\text{-}51)$$

The relations do not apply beyond $p = 1$, that is beyond saturation. But experiments have shown that conditions at saturation do not vary much beyond saturation as long as we do not go to very high drain voltages.

At saturation $p = 1$, and Eq.(9-50) becomes

$$\overline{i_d^2} = 2/3 \; 4kT \Delta f \, g_{ms} \qquad (9\text{-}52)$$

At $V_d = 0$

$$\overline{i_d^2} = 4kT \Delta f \, g_{do} = 4kT \Delta f \, g_{ms} \qquad (9\text{-}53)$$

At $V_d = 0$, and $V_g = 0$

$$\overline{i_d^2} = 4kT \Delta f \, g_{doo} \qquad (9\text{-}54)$$

Fig 9-8 shows the variation of the function $f_1(p)$ with p.

9-4B. Noise in the gate:

The channel noise is coupled to the gate circuit by means of the gate to channel capacitance C_{gc}. Thus the input noise generator $\overline{i_d^2}$ and the output noise generator $\overline{i_g^2}$ will be partly correlated. The output noise is not the simple sum of $\overline{i_d^2}$ and $\overline{i_g^2}$ calculated separately, but the correlation existing between the two must be taken into account also.

The contribution of the thermal noise in the channel to the gate noise current can be calculated as follows:

Making use of Eq.(9-37), the incremental change in charge density of the channel due to the voltage perturbation δV at $x = x_o$ may be written as

$$\begin{aligned} \Delta Q(x) &= Q(x) - Q_o(x) \\ &= \frac{\epsilon_{ox}}{T_{ox}} \left\{ V_o(x) - V_1(x) \right\} \qquad 0 < x < x_o \\ &= \frac{\epsilon_{ox}}{T_{ox}} \left\{ V_o(x) - V_2(x) \right\} \qquad x_o < x < L \end{aligned} \qquad (9\text{-}55)$$

A fluctuation in the channel charge will induce an equal and opposite fluctuation in the charge on the gate electrode. The net change in charge per unit channel width will be:

$$\Delta q = \int_o^L \Delta\sigma(x)\, dx$$

$$= \frac{\epsilon_{ox}}{T_{ox}} \left\{ \int_o^L V_o(x)\, dx - \int_o^{x_o} V_1(x)\, dx - \int_{x_o}^L V_2(x)\, dx \right. \qquad (9\text{-}56)$$

Letting i_o and $(i_o+\Delta i)$ be the channel currents at equilibrium and perturbed conditions respectively , we can integrate Eq.(9-56) as follows:

$$\Delta q = \frac{\epsilon_{ox}}{T_{ox}} \beta_s \left\{ \int_o^{V_d} \frac{[V_g - V_p - V_o(x)]\, V_o(x)}{i_o}\, dV - \int_o^{V(x_o)} \frac{[V_g - V_p - V_1(x)]\, V_1(x)}{i_o + \Delta i} dV \right.$$
$$\left. - \int_{V(x_o)}^{V_d} \frac{[V_g - V_p - V_2(x)]\, V_2(x)}{i_o + \Delta i}\, dV \right\} \qquad (9\text{-}57)$$

Neglecting terms with $\delta V \Delta i$, Eq(9-57) can be evaluated to be:

$$\Delta q = \frac{\Delta i}{I_d} \left\{ Q_o - \frac{\epsilon_{ox} L}{T_{ox}} V(x_o) \right\} \qquad (9\text{-}58)$$

where $$Q_o = \frac{\epsilon_{ox} \beta_s}{T_{ox}} \frac{1}{I_d} \left\{ (V_g - V_p) \frac{V_d^2}{2} - \frac{V_d^3}{3} \right\} \qquad (9\text{-}59)$$

The mean-squared charge fluctuation can then be obtained by squaring Eq.(5-58) and substituting Eqns 9-45 and 9-46

$$\overline{q^2} = \left(\frac{\epsilon_{ox} L}{T_{ox}} \right)^2 4kT\Delta f \frac{\beta_s^2}{L^2 I_d^3} \int_o^{V_d} \left\{ V_g - V_p - V(x) \right\}^2 \left\{ \frac{Q_o T_{ox}}{\epsilon_{ox} L} - V(x) \right\}^2 dV \qquad (9\text{-}60)$$

when (9-60) is integrated

$$\overline{q^2} = 4kT\Delta f \frac{(C_{ox} L)^2}{g_{ms}} f_2(p) \qquad (9\text{-}61)$$
$$= 4kT\Delta f \frac{C_{gc}^2}{g_{ms}} f_2(p)$$

where $$f_2(p) = \frac{1}{(1 - 1/2\, p)^3} \left\{ \frac{(1/2 - 1/3\, p)^2}{(1 - 1/2\, p)^2} (1 - p + \tfrac{1}{3} p^2) - \frac{(1/2 - 1/3\, p)}{(1 - 1/2\, p)} (1 - 4/3\, p + 1/2\, p^2) + (\tfrac{1}{3} - \tfrac{1}{2} p + \tfrac{1}{5} p^2) \right\}$$

(9-62)

Fig 9-9 shows the plot of $f_2(p)$.

At saturation $p = 1$, and

$$\overline{q^2} = 0.12 \times 4kT\Delta f \frac{C_{gc}^2}{g_{ms}} \qquad (9\text{-}63)$$

and the fluctuation current in the short-circuited gate circuit will be

$$i_g = j\omega q \qquad (9\text{-}64)$$

so that

$$\overline{i_g^2} = 0.12 \times 4kT\Delta f \frac{\omega^2 C_{gc}^2}{g_{ms}} \qquad (9\text{-}65)$$

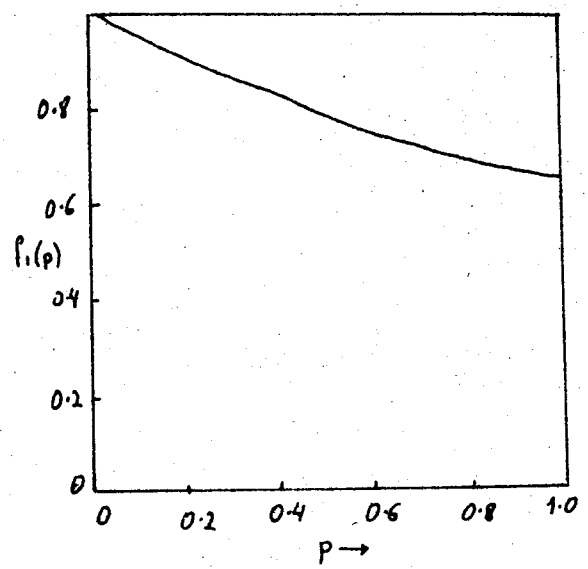


Fig 9-8: Plot of the function $f_1(p)$ where $p = V_d/(V_g - V_p)$

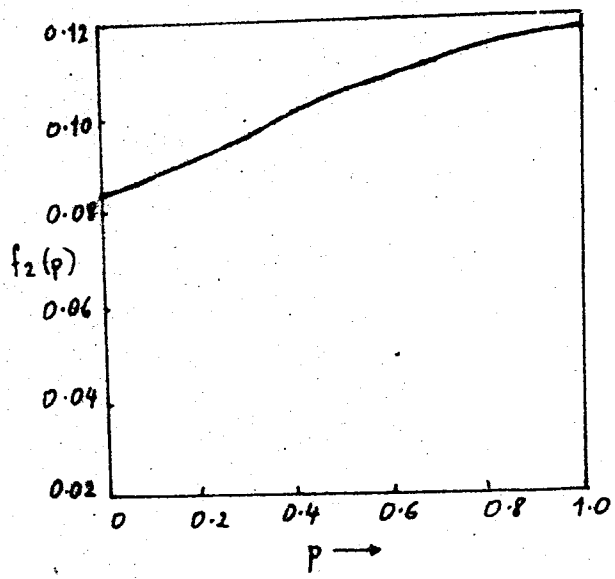


Fig 9-9: Plot of the function $f_2(p)$ where $p = V_d/(V_g - V_p)$

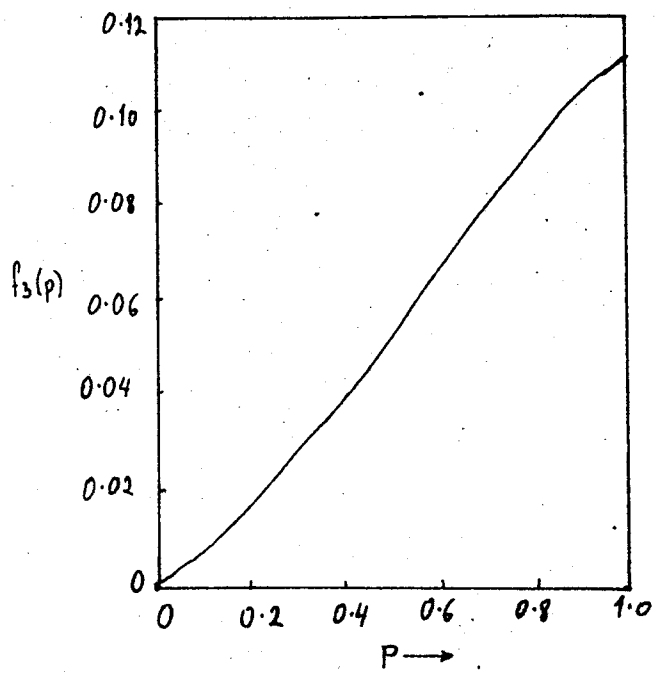


Fig 9-10 : Plot of the function $f_3(p)$ where $p = V_d/(V_g - V_p)$

At $V_d = 0$, $p = 0$, and

$$\overline{q^2} = \frac{1}{12} \cdot 4kT\,\Delta f\,\frac{C_{gc}^2}{g_{ms}} = \frac{1}{12}\,4kT\,\Delta f\,\frac{C_{gc}^2}{g_{do}} \qquad (9\text{-}66)$$

At $V_d = V_g = 0$

$$\overline{q^2} = \frac{1}{12}\,4kT\,\Delta f\,\frac{C_{gc}^2}{g_{doo}} \qquad (9\text{-}67)$$

The cross-product fluctuation $\Delta q \Delta i_d$, becomes by using Eq.(9-58)

$$\Delta q \Delta i_d = 4kT\,\Delta f\,\frac{\beta_s^2}{L^2 I_d^2}\left\{V_g - V_p - V(x_o)\right\}^2\left\{Q_o - \frac{\epsilon_{ox} L}{T_{ox}}\,V_o(x)\right\} dV \qquad (9\text{-}68)$$

Integrating Eq.(9-68) over the channel and multiplying by jw , we get

$$\overline{i_g^* i_d} = 4kT\,\Delta f\,jw\,C_{gc}\,f_3(p) \qquad (9\text{-}69)$$

where

$$f_3(p) = \frac{1}{(1-\frac{1}{2}p)^2}\left\{\frac{(\frac{1}{2} - \frac{1}{3}p)}{(1-\frac{1}{2}p)}\left(1 - p + \frac{1}{3}p^2\right) - \left(\frac{1}{2} - \frac{2}{3}p + \frac{1}{4}p^2\right)\right\} \qquad (9\text{-}70)$$

Fig (9-10) shows a plot of $f_3(p)$

At saturation $p = 1$, and

$$\overline{i_g^* i_d} = 4kT\,\Delta f\,0.11\,jw\,C_{gc} \qquad (9\text{-}71)$$

The correlation coefficient is defined as

$$c = \frac{\overline{i_g^*\, i_d}}{\sqrt{i_g^2\, i_d^2}} \qquad (9\text{-}72)$$

The correlation at saturation is calculated to be

$$c = j0.39 \qquad (9\text{-}73)$$

9-5. Noise Parameters and Methods of Measuring Noise in Field effect Transistors:

To enable calculations on networks which include field effect transistors to be made , an equivalent two-port circuit representation was decided upon. The theorem that forms the basis for this method of representation is that any linear active two-port network can be represented by a series noise voltage generator and a shunt noise current generator at the input of an

ideal noiseless network . Fig 9-11 shows such a circuit.

This representation is a good one for a number of reasons: The ratio of the magnitude of the two noise sources gives the value of the impedance(R_{so}) for a minimum noise factor; the noise factor is determined only by these two generators and by the correlation between them, and is independent of the parameters of the two-port noise-free network; the magnitudes of the two input generators can be separated by measurements involving the variation of the source resistance, and the complex correlation between them can be determined by measurements involving the source reactance.

The magnitude of the mean square noise voltage V^2 arising from these two sources and referred to the input of the device can be obtained easily as follows:

If two noise generators v and iZ_s are considered , the second being the Thévenin voltage equivalent of the current generator i and the source impedance Z_s in Fig 9-11, the magnitude of the total mean square voltage V^2 is given by:

$$V^2 = 4kTR_s + \overline{|v + (iZ_s)|^2} \qquad (9\text{-}74)$$

$$\overline{v+(iZ_s)^2} = \overline{\{v+(iZ_s)\}\{v+(iZ_s)\}^*}$$

$$= \overline{\{v+(iZ_s)\}\{v^*+(iZ_s^*)\}} \qquad (9\text{-}75)$$

$$= \overline{vv^*} + \overline{(iZ_s)(iZ_s)^*} + \overline{v(iZ_s)^*} + \overline{v^*(iZ_s)}$$

$$= |\overline{v}|^2 + \overline{|i|^2}|Z_s|^2 + 2Re\,(Z_s\,\overline{v^*i})$$

If there is no correlation between the two sources $\overline{v^*i} = 0$ and

$$\overline{|v+(iZ_s)|^2} = \overline{|v|^2} + \overline{|i|^2|Z_s|^2} \qquad (9\text{-}76)$$

If a complex correlation coefficient is defined as

$$\gamma = Re(\gamma) + j\,Im(\gamma) = \overline{v^*i} \,/\, \sqrt{|v|^2|i|^2} \qquad (9\text{-}77)$$

and

$$Z_s = R_s + jX_s \qquad (9\text{-}78)$$

Then

$$\overline{|v+(iZ_s)|^2} = \overline{|v|^2} + \overline{|i|^2}Z^2 + 2\sqrt{(|v|^2|i|^2)}\{R_s Re(\gamma) - X_s I_m(\gamma)\} \qquad (9\text{-}79)$$

so that when Z_s is real ($x_s = 0$)

$$V^2 = V_r^2 = 4kTR_s + \overline{|v|^2} + \overline{i^2}R^2 + 2R\sqrt{\overline{|v|^2}\,\overline{|i|^2}}\ \mathrm{Re}\,\gamma \qquad (9\text{-}80)$$

and when Z is imaginary , ($R_s = 0$)

$$V^2 = V_i^2 = \overline{|v|^2} + \overline{|i|^2}X^2 - 2X\sqrt{\overline{|v|^2}\,\overline{|i|^2}}\ \mathrm{Im}\,\gamma \qquad (9\text{-}81)$$

Letting the term due to the correlation $= \overline{|v_c|^2}$, we have

$$\overline{|v_c|^2} = 2\sqrt{\overline{|v|^2}\,\overline{|i|^2}}\ \{R_s\,\mathrm{Re}(\gamma) - X_s\,\mathrm{Im}\,\gamma\} \qquad (9\text{-}82)$$

we can writte Eq.(9-74) in the following form:

$$V^2 = 4kTR_s + \overline{|v|^2} + \overline{i^2}|Z|^2 + \overline{|v_c|^2} \qquad (9\text{-}83)$$

To find the noise factor expression of this two port device Fig 9-12 is considered. The thermal noise generator which is considered purely resistive at the moment is represented by $4kTR_s$. The noise factor is given by:

$$F = \frac{\text{mean-square noise voltage at xy from all sources}}{\text{mean-square thermal noise voltage at xy from the source alone.}}$$

From Fig 9-12b

$$F = \frac{\{\overline{|v|^2} + \overline{|i|^2}R_s^2 + 4kTR_s + 2R_s\sqrt{\overline{|v|^2}\,\overline{|i|^2}}\ \mathrm{Re}\,\gamma\}\left|\frac{Z_{in}}{R_s+Z_{in}}\right|^2}{4kTR_s\left|\frac{Z_{in}}{R_s+Z_{in}}\right|^2} \qquad (9\text{-}84)$$

$$F = 1 + \left\{\frac{\overline{|v|^2} + \overline{|i|^2}R_s^2 + 2R_s\sqrt{\overline{|v|^2}\,\overline{|i|^2}}\ \mathrm{Re}\,\gamma}{4kTR_s}\right\} \qquad (9\text{-}85)$$

If excess noise factor Fe is defined as

$$Fe = F - 1 \text{ , we can write} \qquad (9\text{-}86)$$

$$Fe = \frac{\overline{|v|^2}}{4kT}\frac{1}{R_s} + \frac{\sqrt{\overline{|v|^2}\,\overline{|i|^2}}}{2kT}\ \mathrm{Re}\,\gamma + \frac{\overline{|i|^2}}{4kT}R_s \qquad (9\text{-}87)$$

Eqns(9-85) and (9-87) show that the expressions for noise factor are independent of the two port parameters.

When we differentiate Eq(9-87) with respect to R_s , we find that the source resistance R_{so} giving minimum excess noise factor , Feo , is

$$R_{so}^2 = \frac{\overline{|v|^2}}{\overline{|i|^2}} \qquad (9\text{-}88)$$

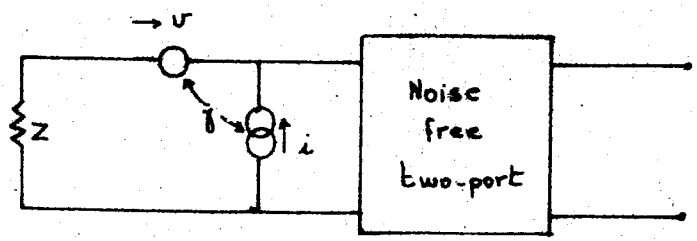


Fig 9-11. General noise representation of a two-port.

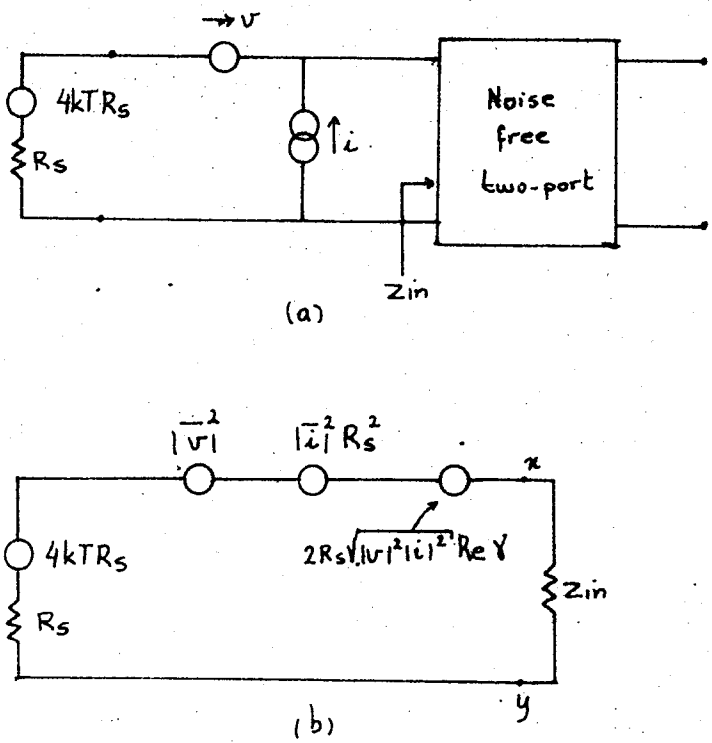


Fig. 9-12 Equivalent circuits for the two-port of Fig 9-11
a) Circuit including the thermal noise of the source
b) Equivalent circuit at the input with respect to the noise.

Eq.9-88 shows that R_{so} is independent of the correlation between the noise sources for a purely resistive source. Substituting Eq.(9-88) into Eq.(9-87), we find

$$F_{eo} = \frac{2|\bar{v}|^2 + 2R_{so}\sqrt{|v|^2 \frac{|v|^2}{R_{so}^2}}\ Re\,Y}{4kTR_{so}} \qquad (9\text{-}89)$$

$$F_{eo} = \frac{|\bar{v}|^2 \{1 + Re(Y)\}}{2kTR_{so}} \qquad (9\text{-}90)$$

Let us now consider the case when the source impedance Z_s is complex. R_{so}^2 is dependent on the imaginary part of the correlation coefficient because for a complex source impedance, we have

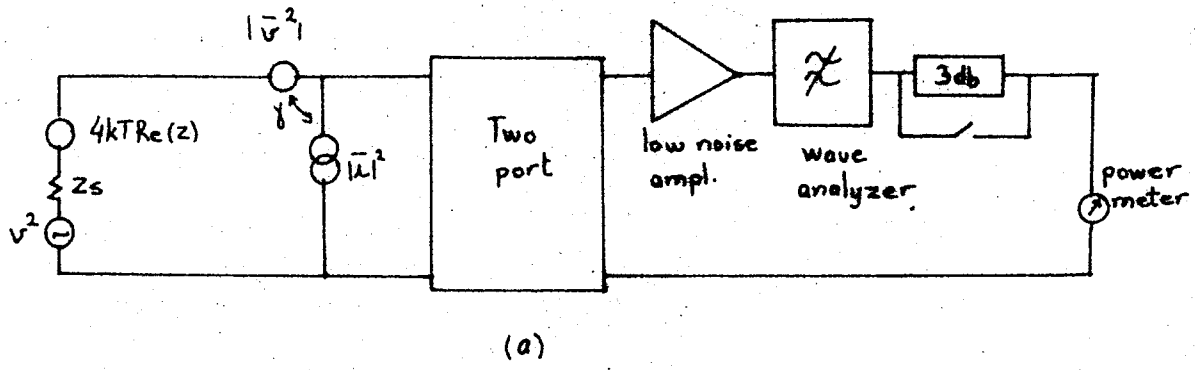
$$F_e = \frac{|\bar{v}|^2 + |\bar{i}|^2 R_s^2 + |\bar{i}|^2 X^2 + 2R_s\sqrt{|v|^2|i|^2}\ Re\,Y - 2X_s\sqrt{|v|^2|i|^2}\ Im\,Y}{4kTR_s} \qquad (9\text{-}91)$$

when Eq(9-91) is differentiated with respect to R_s and equated to zero, the expression for R_{so} is obtained as

$$R_{so}^2 = \frac{|\bar{v}|^2 + |\bar{i}|^2 X^2 - 2X\sqrt{|v|^2|i|^2}\ Im\,Y}{|i|^2} \qquad (9\text{-}92)$$

The total mean square noise voltage can be measured by doubling the noise power output from a selective measuring channel by the introduction of an equal and known mean square voltage in series with the source. Either a calibrated noise generator or a sine wave generator could be used. The principle of the measurement is outlined in Fig 9-13. First the bias condition of the FET or MOS transistor is adjusted with the required source impedance , and with the oscillator set at zero output. The frequency of the generator must be set at the center frequency of the filter.

The transistor circuit power gain must be checked at each bias setting to ensure that the noise from the low noise amplifier and subsequent apparatus is negligible when referred to the transistor input. With the 3 db pad omitted , the noise output of the transistor, after amplification by the low noise amplifier, is applied to the wave analyzer, the output of which is indicated on the powermeter. The average reading is noted . Then the 3db



(a)

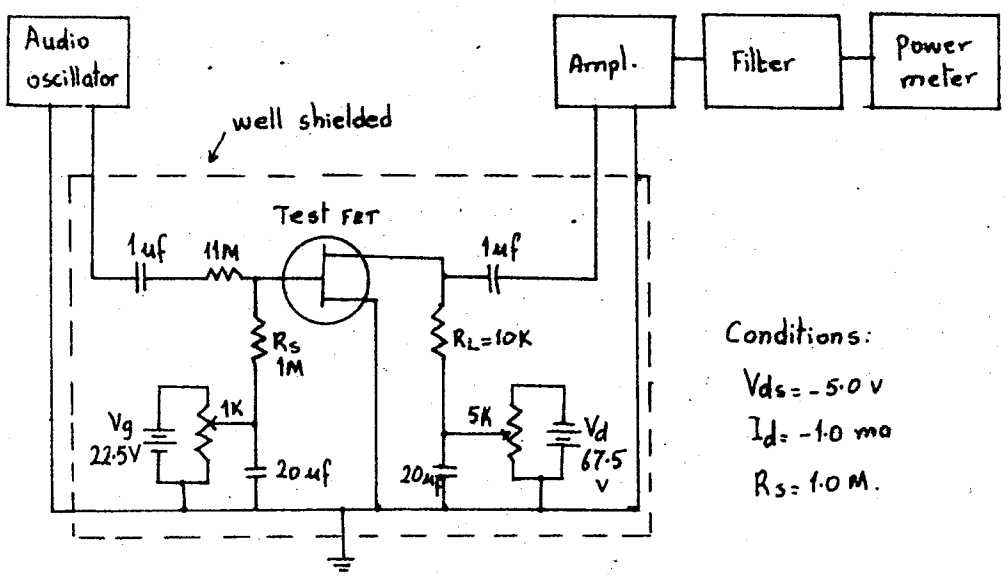


Fig.9-13 a) Principle of noise measurement
b) an actual system used in measuring the noise factor of an FET.

pad is inserted. This causes the milliammeter deflection to reduce The output of the sine wave oscillator is increased until the reading on the powermeter is restored to the initial value. The output of the oscillator at this point is measured on the r_{ms} reading volmeter. V_{cw}^2 is calculated.

V_{cw} = the r_{ms} oscillator voltage

$V_{cw}^2 = v^2 \Delta f$

From relation in (9-93) , the applied voltage density ($v^2 = V_{cw}^2 / \Delta f$) must be calculated. But to do this, separate measurement must be made to find the bandwidth Δf .

9-6. Discussion of Experimental Results.

Fig 9-14 shows a graph of noise factor against frequency for a fixed source resistance of 100 kΩ for two similar MOS transistors. The characteristics divide into three regions. The general form of these noise factor characteristics look like that of a junction gate FET but the excess noise region starts to rise steeply as the frequency is reduced at a frequency which is about 30 times the value for a good junction gate FET . Thus the MOS transistor is an inherently noisy transistor compared with a junction gate FET. At frequencies of about 10 kc and over the device shows white noise. But at frequencies of the order of hundreds of kilocycles per sec , the noise factor starts to rise again.

Figs 9-15 a and b show variations of excess noise factor with source resistance at a frequency of 10 kc per sec and 100 kc per sec respectively.

Fig 9-16 is a graph of the variation of the components of total noise V_r^2 (see Eq.9-80) with source resistance at a frequency of 100 kc/s. Fig 9-17 shows the variation of V_1^2 (see Eq. 9-81) with source resistance at the same frequency. These last two graphs are used in calculating the separate noise parameters. With the help of Eq. 9-80 the generator $|\bar{v}|^2$ can be measured directly by putting $R_s = 0$ because when $R_s = 0$, $V_r^2 = |\bar{v}|^2$. This value can be read directly from the graph of Fig 9-16.

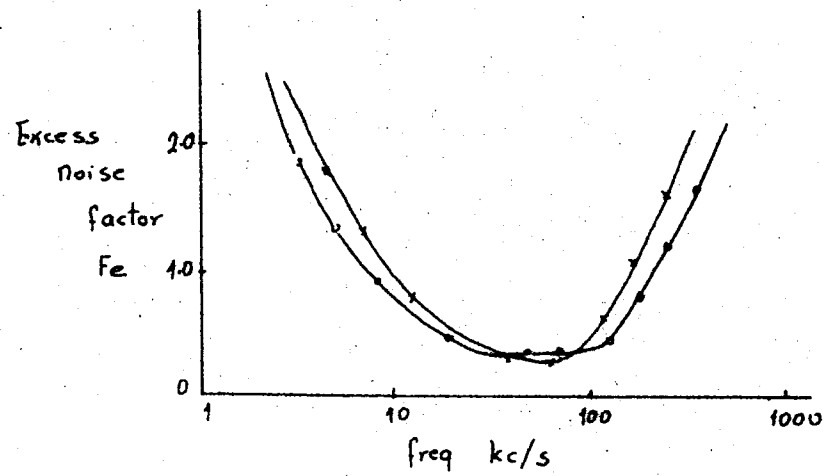


Fig.9-14 Variation of excess noise factor with frequency for two similar metal oxide semiconductor transistors (TIXSII p-channel MOSFET)
Source resistance = 100 kΩ

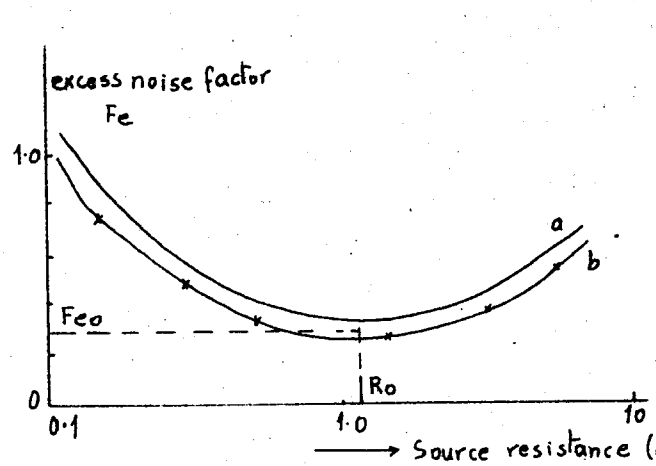


Fig 9-15a Variation of excess noise factor with source resistance measured by:

a) noise diode
b) sinewave oscillator, at a freq. of 10 kc/s

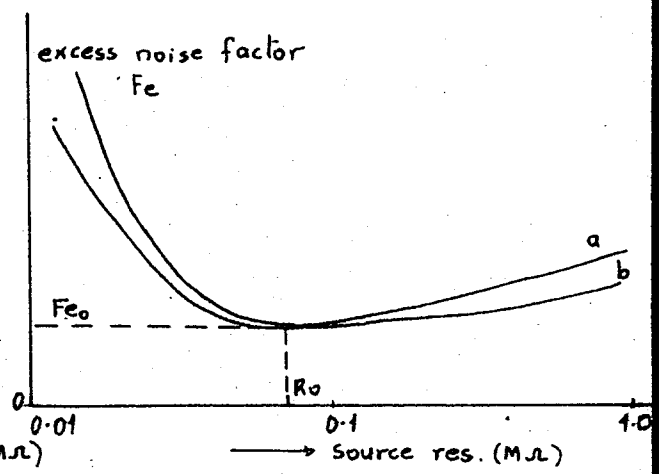


Fig.9-15b Variation of excess noise factor with source resistance measured by

a) noise diode
b) sinewave oscillator at a freq. of 100 kc/s

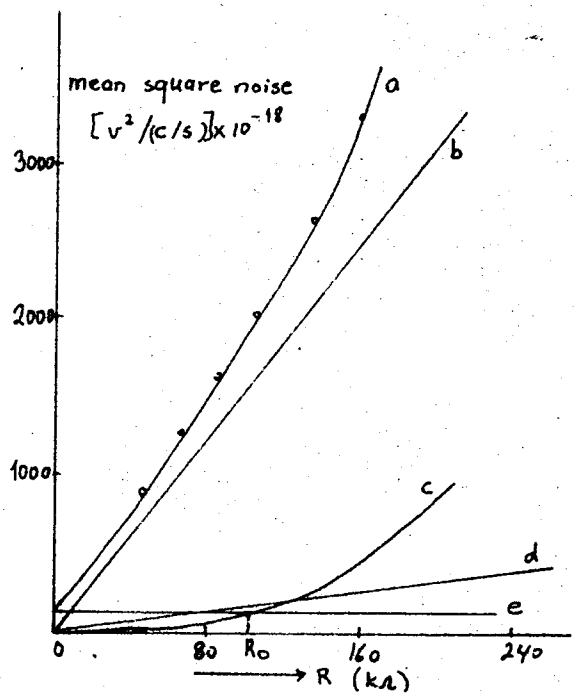


Fig 9-16. Variation of the components of total noise V_T with source resistance, at a freq. of 100 kc/s
a) V_T^2
b) $4kTR$ d) $2R\sqrt{|v|^2|i|^2}\ \mathrm{Re}\,Y$
c) $|\bar{i}|^2R^2$ e) $|\bar{v}|^2$

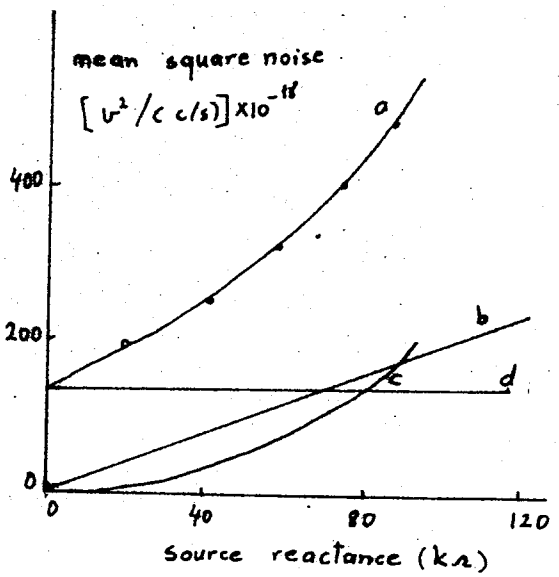


Fig 9-17. Variation of components of total noise V_i^2 with source reactance at freq. 100 kc/s.
a) V_i^2
b) $2X\sqrt{|v|^2|i|^2}\ \mathrm{Im}\,Y$
c) $\overline{|i|^2}X^2$
d) $|\bar{v}|^2$

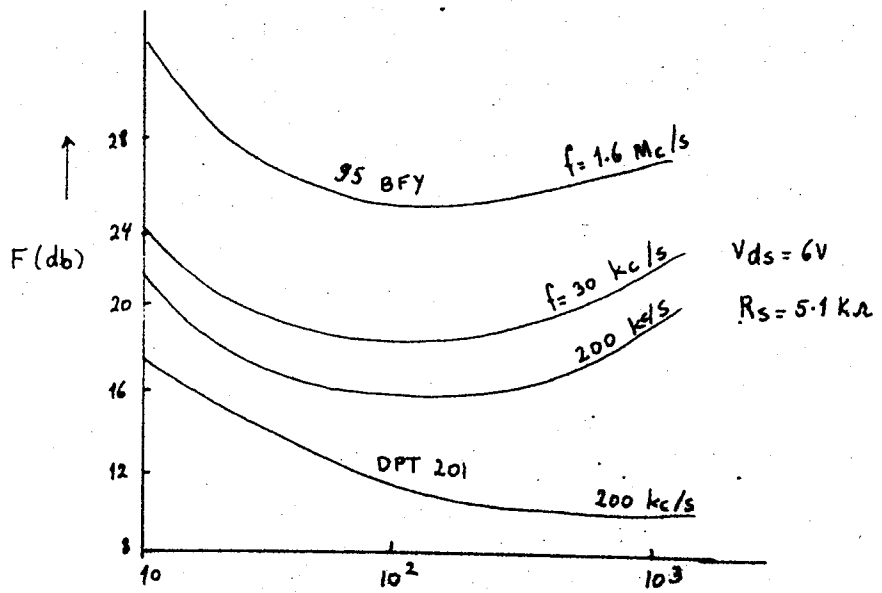


Fig 9-18. Noise figure as a function of dc conditions

The horizontal line e drawn in Fig 9-16 represents this value of $|\bar{v}|^2$. The value of R_{so} can be found from Fig 9-15 as shown. Then by making use of Eq. 9-88, and using the value of $\overline{|v|^2}$ and R_{so} just found we can find $\overline{|i|^2}$. Using Eq.(9-80) again, the residual term $R_e(Y)$ can be calculated. In a similar fashion, we can find the value of $Im(Y)$ from Fig 9-17, making use of Eq.(9-81). The term $R_e(Y)$ can also be calculated by using Eq. 9-90.

Further investigations have been carried out to determine the noise figure behavior as a function of the operating dc conditions. One set of the results $F(I_d, f)$ is presented in Fig 9-18. The fixed frequencies for the 95 BFY unit have been chosen at, below, and above the frequency of the minimum noise figure. The results show that the noise figure has wide minima with respect to the drain current and that the shapes of the curves at the three frequencies (30, 200, and 1600 kc/s) do not differ appreciably. The sample DPT201 shows lower noise than the 95BFY but the general behavior is similar.

Measurements have also shown that the noise figure is independent of the drain voltage V_d for the DPT201.

The conclusion is that in order to achieve a low noise figure with the present types of MOS transistors, they have to be operated from a large source resistance R_s ($> 100\ k\Omega$) and at low frequencies (10 kcls $< f <$ 100 kc/s).

CHAPTER 10

LINEAR FET AMPLIFIER CIRCUITS

10-1. Basic Fet Amplifier Circuits:

The FET can be operated with either the drain, gate or source terminal common to input and output circuits. The most useful configuration is the common source, which has the highest gain. In the following discussions the P-channel FET will be considered unless otherwise stated.

10-1A. The Common Source Amplifier:

The circuit in Fig.10-1 is a basic common-source FET amplifier with two fixed-bias batteries. A signal voltage generator v_g is applied to the input and the output is taken between the drain and the common terminal . Since the FET characteristics vary much with temperature and device selection, fixed bias is not desirable. Several biasing methods may be used. Fig 10-2 and 10-3 show such set-ups. The analysis of Fig 10-2, which shows a self biased scheme, at low and medium frequencies can be made by using the low-frequency equivalent circuit shown in Fig 10-2(b) and 10-2c. For low frequencies the effects of C_{gd} and C_{gs} are negligible . Z_L includes the drain resistor R_d and the input impedance of another stage of amplifier. If v_o is the output voltage,

$$v_o = -g_m v_{gs} \frac{Z_L}{1 + Z_L g_{ds}} \qquad (10\text{-}1)$$

The voltage gain of the amplifier is:

$$A_v = \frac{v_o}{v_{gs}} = - \frac{g_m Z_L}{1 + Z_L g_{ds}} \qquad (10\text{-}2)$$

When the product $Z_L g_{ds} \ll 1$, the gain reduces to

$$A_v = - g_m Z_L \qquad (10\text{-}3)$$

The output is 180 degrees out of phase with the input.

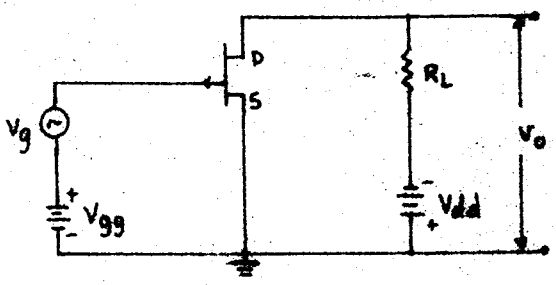


Fig. 10-1. Basic common source (CS) FET amplifier.

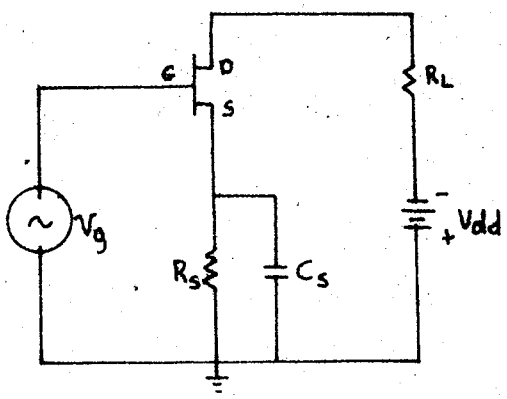


Fig.10-2.a Self biased CS amplifier

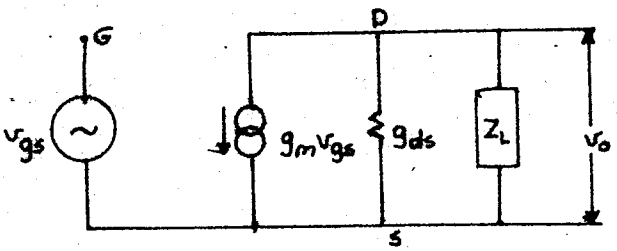


Fig.10-2b. Low frequency equivalent circuit of CS amplifier.

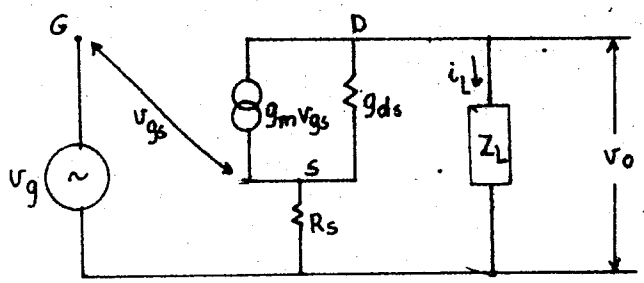


Fig.10-2c. CS amplifier circuit with a source resistor

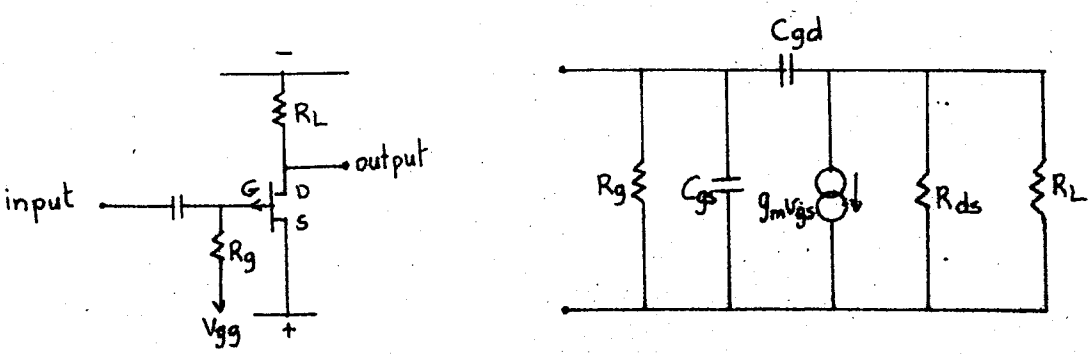


Fig.10-3. CS amplifier and its small signal equivalent circuit.

Eq.(10-3) is similar to the approximate gain of a pentode stage.

The output admittance of the amplifier at low frequencies, without the load Z_L is g_{ds}. But with Z_L, the output impedance becomes:

$$Z_o = \frac{Z_L}{1 + Z_L g_{ds}} \qquad (10\text{-}4)$$

Fig 10-2c shows the equivalent circuit of the self-biased circuit with a source resistor. In this circuit, the load current i_L is

$$i_L = -g_m v_{gs} \frac{1}{1 + g_{ds}(Z_L + R_s)} \qquad (10\text{-}5)$$

where $$v_{gs} = v_g + i_L R_s \qquad (10\text{-}6)$$

so that $$i_L = -\frac{g_m v_g}{1 + g_m R_s + g_{ds}(Z_L + R_s)} \qquad (10\text{-}7)$$

and $$v_o = i_L Z_L = -\frac{g_m v_g Z_L}{1 + g_m R_s + g_{ds}(Z_L + R_s)} \qquad (10\text{-}8)$$

Now the voltage gain of the amplifier becomes

$$A_v = -\frac{g_m Z_L}{1 + g_m R_s + g_{ds}(Z_L + R_s)} \cong -\frac{g_m Z_L}{1 + g_m R_s} \qquad (10\text{-}9)$$

At higher frequencies the situation is more complex. The capacitances C_{gd} and C_{gs} cannot be neglected. In such a case the modified voltage becomes

$$A'_v = A_v \frac{(g_m Z_{gd} + 1)}{(g_m Z_{gd} + A_v)} \qquad (10\text{-}10)$$

In general A'_v depends on frequency both in magnitude and angle. It becomes frequency independent for $A_v = 1$. The value of Z_{gd} is given by

$$Z_{gd} = R - j \frac{1}{wC_{gd}} \qquad (10\text{-}11)$$

But at medium frequencies R can be neglected, and in such a case, the product

$$g_m Z_{gd} = \frac{g_m}{wC_{gd}} \qquad (10\text{-}12)$$

Since $g_m Z_{gd} \gg 1$, Eq.(10-10) can be written as

$$A'_v = A_v \left(1 + \frac{jwC_{gd} A_v)^{-1}}{g_m} \right. \qquad (10\text{-}13)$$

The loading that the FET presents to a generator can be calculated by dividing the input current by the voltage between the input terminals, thus from Fig.10-4

$$(Z_i)^{-1} = \frac{i_s}{v_{gs}} = \frac{1}{v_{gs}} \left(\frac{V_{gs}}{Z_{gs}} + \frac{V_{gs} - V_o}{Z_{gd}} \right). \qquad (10\text{-}14)$$

$$y_{in} = jw \left[C_{gs} + C_{gd} \left(1 + g_m \frac{Z_L}{1 + g_{ds} Z_L} \right) \right] \qquad (10\text{-}15)$$

$$= jw \left[C_{gs} + C_{gd} \left(1 + A_v \right) \right]$$

Eq.(10-15) can be expressed in a more general way, for all frequencies as:

$$y_{in} = \left[\frac{1}{Z_{gs}} + \frac{A'_v + 1}{Z_{gd}} \right] \qquad (10\text{-}16)$$

If the amplifier circuit contains a gate continuity resistor R_g as in Fig. 10-3 , then the input admittance becomes

$$y_{in} = \frac{1}{R_g} + \frac{1}{Z_{gs}} + \frac{A'_v + 1}{Z_{gd}} \qquad (10\text{-}17)$$

In terms of the chort-circuit admittance parameters given in most data sheets

$$y_{in} = jw \left[b_{is} - A'_v \, b_{rs} \right] \qquad (10\text{-}18)$$

If Z_L has a reactive component, y_{in} will contain a real part. If this reactance is capacitive, the real part of y_{in} will be positive and if the reactance is inductive, the real part of y_{in} is negative. In the second case, the amplifier will be unstable and will oscillate at the frequency at which the imaginary part of y_{in} is zero. The situation is particularly serious if the load is a parallel LC circuit, which looks inductive, resistive or capacitive according as the operating frequency is below, at or above the resonance frequency of the circuit. For these reasons

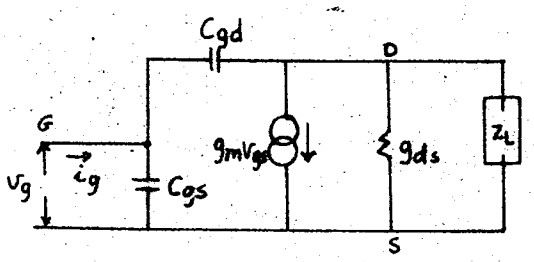


Fig. 10-4. CS equivalent circuit with capacitances included.

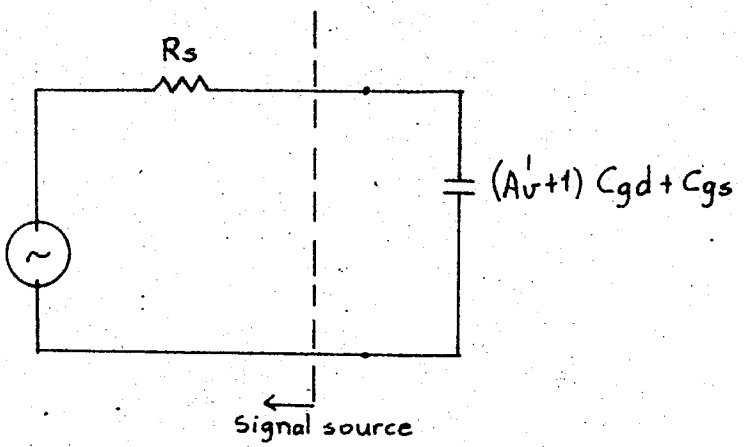


Fig. 10-5. The equivalent input circuit of the CS amplifier Rs is the signal source resistor.

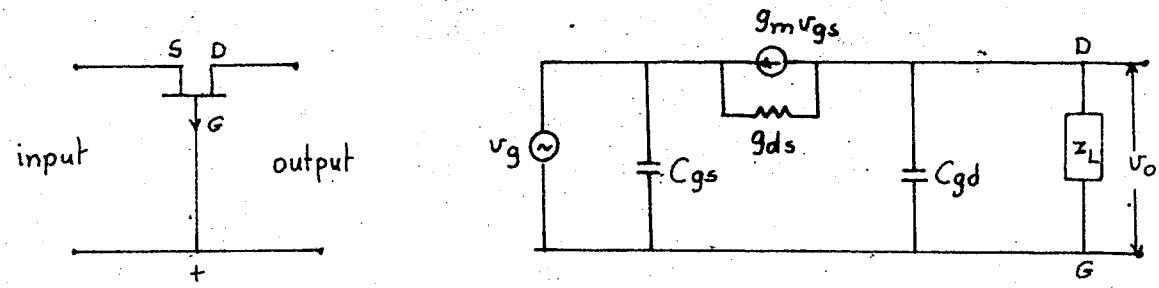


Fig. 10-6 Common gate configuration.
a) CG circuit
b) CG equivalent circuit.

the simple CS amplifier is used primarily at low frequencies and with a resistive load. For higher frequency operations it is required to take precautions to reduce or eliminate the unwanted feedback effects by neutralization or by mismatching. The input capacitance of the FET given by Eq.(10-15) forms a low-pass L-section with the source impedance of the signal generator driving it as shown in Fig 10-5 . The 3 db down frequency of this equivalent circuit is given by

$$w_{co} = \frac{1}{R_s\left[(A_v' + 1)\, C_{gd} + C_{gs}\right]} \qquad (10\text{-}19)$$

It is because of this Miller-effect that FETs cannot be operated in the CS mode to yield high values of A_v' and at the same time give good frequency response.

10-1B. The Common Gate Amplifier

This configuration seldom finds application. Fig 10-6a shows the common gate amplifier in which the gate is the common terminal between input and output circuits. As can be seen from the equivalent circuit, Fig 10-6b, this configuration is unilateral since C_{gd} and C_{gs} are of the same order.

The voltage gain can be calculated by summing currents at the output node D .

$$g_m v_{gs} = - v_{gs} g_{ds} + v_o \left(g_{ds} + \frac{1}{Z_L} + jwC_{gd}\right) \qquad (10\text{-}20)$$

so that the voltage gain is

$$A_v = \frac{v_o}{V_{gs}} = \frac{(g_m + g_{ds})\, Z_L}{1 + Z_L\,(g_{ds} + jwC_{gd})} \qquad (10\text{-}21)$$

At low frequencies

$$A_v \cong \frac{Z_L\,(1 + \mu)}{Z_L + R_{ds}} \qquad (10\text{-}22)$$

where $\mu = g_m R_{ds}$

The magnitude of the common gate gain is approximately equal to that of the common source gain when $g_{ds} \ll g_m$. But unlike the common source gain , there is no signal reversal at frequencies where

reactance are negligible.

The input admittance can be determined by summing currents at the input node of the equivalent circuit of fig. 10-6b.

$$i_g = jwC_{gs}v_{gs} + g_m v_{gs} + g_{ds}\ (v_{gs} - v_o) \qquad (10\text{-}23)$$

therefore

$$y_{in} = \frac{i_g}{v_{gs}} = g_m + g_{ds}\ (1\text{-}A) + jw\ C_{gs} \qquad (10\text{-}24)$$

It can be shown that at low frequencies the input impedance is just

$$Z_{in} = \frac{R_{ds} + R_L}{1 + \mu} \qquad (10\text{-}25)$$

For small R_L and large μ the input impedance is approximately $1/g_m$, and does not exhibit Miller effect. A bipolar transistor in the common emitter configuration will usually give a better gain with an impedance only a little lower than that of the CG FET stage. Therefore the CG configuration is rarely used unless the low noise or radiation resistance properties of the FET are needed.

10-1C. Common - drain Amplifier or source Follower.

The FET used in the common drain configuration is analogous to the vacuum tube cathode follower , and the transistor emitter follower. It is characterized by a high input impedance , low out put impedance and voltage gain slightly less than unity. It is very useful for impedance transformations when FETs are being used with bipolar transistors.

Consider the circuit of Figs 10-7a and 7b. Summing the currents at the source node s , and the voltages around the input loop, we get

$$g_m v_{gs} = v_o\ (g_{ds} + \frac{1}{R_L}) - jw\ C_{gs}(v_{gd} - v_o) \qquad (10\text{-}26)$$

where $v_{gs} = v_{gd} - v_o$

$$A_v = \frac{v_o}{v_{gd}} = \frac{(g_m + jw\ C_{gs})\ R_L}{1 + (g_m + g_{ds} + jwC_{gs})R_L} \qquad (10\text{-}27)$$

At moderate frequencies where $1/wC_{gs} \gg R_{ds}$, or R_L , the voltage

gain becomes

$$A_v = \frac{g_m R_L}{1 + (g_m + g_{ds})R_L} \qquad (10\text{-}28)$$

$$= \frac{g_m R'_L}{1 + g_m R'_L} < 1 \qquad (10\text{-}29)$$

where $R'_L = \frac{R_{ds} R_L}{R_{ds} + R_L}$

The maximum possible source follower gain can be found by taking the limit as $R_L \to \infty$, giving

$$A_{vmax} < \frac{\mu}{\mu + 1} \qquad (10\text{-}30)$$

If $\mu \gg 1$ (which is true for all available FETs), the gain is nearly unity and there is no phase shift, the source follows the gate.

The input admittance of a source follower can be obtained as follows :

$$i_g = v_g jw\,(C_{dg} + C_{gs}) - jw\,C_{gs} v_o \qquad (10\text{-}31)$$

$$y_{in} = \frac{i_g}{v_g} = jw\left[C_{gd} + C_{gs}\,(1 - A)\right] \qquad (10\text{-}32)$$

The source follower output-impedance can be calculated by setting $v_g = 0$ in Fig 10-7b and summing the currents at the source node s.

$$i_o = v_o\,\left(\frac{1}{R_L} + g_{ds} + jw\,C_{gs}\right) + g_m v_o \qquad (10\text{-}33)$$

$$Z_o = \frac{v_o}{i_o} = \frac{R_L}{1 + (g_{ds} + g_m)R_L + jwC_{gs}} \qquad (10\text{-}34)$$

At low frequencies

$$Z_o = \frac{R'_L}{1 + g_m R'_L}$$

when $R_L \to \infty$, the output resistance becomes

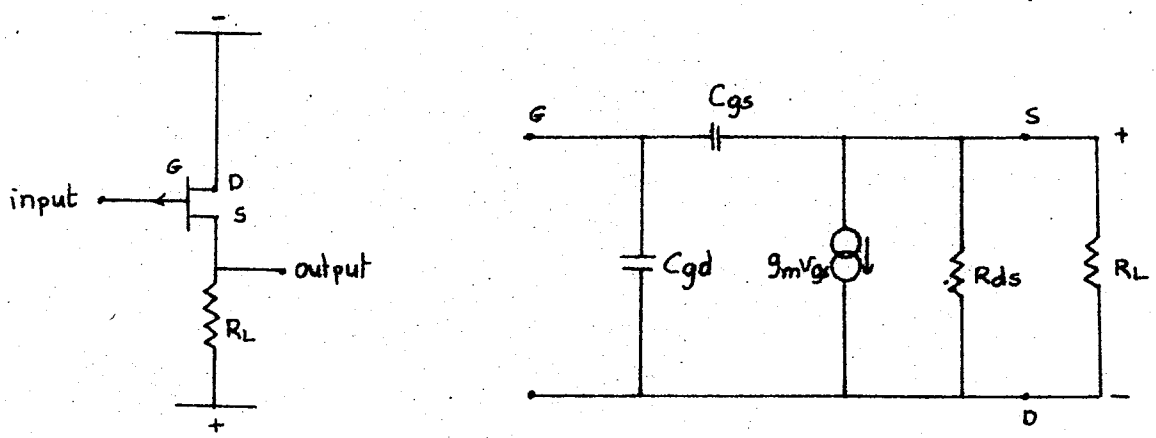


Fig. 7 CD amplifier (a) and its small signal equivalent circuit (b)

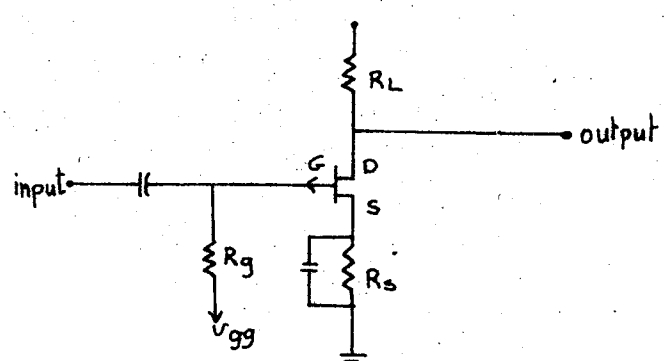


Fig.8. FET amplifier with bias

$$Z_o = \frac{1}{g_m} \qquad (10\text{-}35)$$

10-2. Biasing FET Amplifiers

The characteristics of the FET show variation between different samples of the same nominal type and also a dependence on ambient temperature. Biasing techniques are employed which stabilize the operating conditions against these variations. From Eq. 1-25 the drain current is

$$I_d = G_o \left\{ V_{dg} \left[1 - 2/3 \left(\frac{V_{dg}}{W_o} \right)^{1/2} \right] - V_{gs} \left[1 - 2/3 \left(\frac{V_{gs}}{W_o} \right)^{1/2} \right] \right. \qquad (10\text{-}36)$$

since $I_{do} = \frac{W_o}{3} g_{max}$

and $$g_m = G_o \left[1 - \left(\frac{V_{gs}}{W_o} \right)^{1/2} \right]$$

we can approximate the drain current beyond pinch off by

$$I_d = I_{do} - V_{gs} g_m \qquad (10\text{-}37)$$

The most generally used biasing technique for CS configuration is shown in Fig 10-8 where an external bias source voltage V_{gg} is connected to the gate through a resistance R_g. A resistor R_s may be used in the source circuit.

The overall expression for drain current in Fig 10-8 is

$$I_d = I_{do} - g_m (I_d R_s + I_g R_g - V_{gg}) \qquad (10\text{-}38)$$

$$I_d = \frac{I_{do} + g_m (V_{gg} - I_g R_g)}{1 + g_m R_s} \qquad (10\text{-}39)$$

where I_g = gate current.

The drain current will vary with temperature T as follows:

$$\frac{\partial I_d}{\partial T} = \frac{\partial I_{do}/\partial T + (\partial g_m/\partial T)(V_{gg} - I_g R_g - I_d R_s) - g_m (\partial I_g/\partial T) R_g}{(1 + g_m R_s)} \qquad (10\text{-}40)$$

To make the drain current independent of temperature, we set Eq.(10-40) equal to zero and get the following condition :

$$\frac{\partial g_m}{\partial T} V_{gg} + \frac{\partial I_{do}}{\partial T} = \frac{\partial g_m}{\partial T} (I_g R_g + I_d R_s) + g_m \frac{\partial I_g}{\partial T} R_g \qquad (10-41)$$

Note that the solution of Eq.(10-41) will enable us to choose a certain value of R_g for satisfactory operation with zero temperature coefficient only at a certain given temperature. If R_s and V_{gg} are equal to zero, Eq (10-41) reduces to

$$\frac{\partial I_{do}}{\partial T} = R_g \frac{\partial}{\partial T} (g_m I_g) \qquad (10-42)$$

assuming that $\partial g_m / \partial T$ is small

$$\frac{\partial I_{do}}{\partial T} \cong R_g g_m \frac{\partial I_g}{\partial T} \qquad (10-43)$$

I_g varies exponentially with temperature and its temperature dependence is given by

$$I_g = I_g' \exp\left(\frac{T-T'}{\theta}\right) \qquad (10-44)$$

where I_g' = value of I_g at a temperature T', and θ is constant. For Si, θ is less than $14^{\circ}C$.

Differentiating Eq.(10-44), we get

$$\frac{\partial I_g}{\partial T} = I_g' \frac{1}{\theta} \exp\left(\frac{T-T'}{\theta}\right) = \frac{I_g}{\theta} \qquad (10-45)$$

so that

$$\frac{\partial I_{do}}{\partial T} = \frac{R_g g_m}{\theta} I_g \qquad (10-46)$$

$$R_g = \frac{\theta}{g_m I_g} \frac{\partial I_{do}}{\partial T} \qquad (10-47)$$

For Si devices, $\frac{I_{do}}{T} \cong 0.75$ % and approximately constant over the atmospheric temperature range for transistors of large pinch off voltage. Thus Eq. 10-47 for Si devices may be written as

$$R_g = \frac{14}{g_m I_g} \times I_d \times 0.75 \times 10^{-2} \qquad (10-48)$$

when the value of the gate current at the operation temperature, the drain current I_d and the transconductance g_m are known,

the minimum value of R_g to be used can be calculated from Eq.10-48.

When negative feedback (R_s) is used to reduce current variations "worst-case " transistor parameters method of design is used . If we neglect I_g, Eq.(10-39) becomes

$$I_d = \frac{I_{do} + V_{gg}\, g_m}{1 + g_m R_s} \qquad (10\text{-}49)$$

We can see from Eq.(10-49) that for large values of V_{gg} and R_s , the drain current I_d is practically independent of I_{do} . The worst-case biasing method can best be explained by an example.

Example : Consider a FET type which has the following characteristics:

	Worst transistor	Average	Best transistor
I_{do}(mA)	0.5	1	1.5
g_m (mA/v)	0.5	0.75	1

If this transistor type is to be used to operate at a nominal drain current $I_d = 0.25$ mA , find the value of R_s to be used ?

a) Iff $V_{gg}=0$, we can find R_s using the average values , and Eq.(10-49)

$$1 + g_m R_s = \frac{I_{do}}{I_d}$$

$$(1 + 0.75 R_s) = 4$$

$$R_s = 4\ K\Omega$$

For the minimum case

$$I_d = \frac{0.5}{1 + 0.5x4} = \frac{0.5}{3} = 0.17\text{mA}$$. This varies from 0.25mA by -32% while if R_s were zero , we would have -50% variation.

For the maximum case

$$I_d = \frac{1.5}{1 + 1x4} = 0.3$$ mA. This varies from 0.25 mA by 20% while if R_s were zero , we would have 50% variation from the average.

b) If $V_{gg} \neq 0$

Consider $V_{gg} = 3$ volts .

Calculating R_s from the average case

$$(1 + 0.75\ R_s) = \frac{1+0.75x3}{0.25} = \frac{3.25}{0.25} = 13$$

$$0.75\ R_s = 12$$
$$R_s = 16\ K\Omega$$

For the minimum case

$$I_d = \frac{0.5 + 0.5x3}{1 + 0.5x16} = 0.22mA = -12\% \text{ relative to } 0.25\ mA$$

For the maximum case

$$I_d = \frac{1.5 + 1x3}{1 + 1x16} = \frac{4.5}{17} = 0.26mA = 4\% \text{ relative to } 0.25mA$$

Graphic methods may also be used to bias FET stages. Let us illustrate this method by a hypothetical example. Let Fig (10-9) represent the temperature corrected transfer curves for a hypothetical FET type. If we want to design an amplifier whose quiescent drain current I_Q will be between points A and B , we draw a line L such that it will cross the transfer curves above A and below B as seen in Fig 10-9. The slope of this line will give us the value of the resistance R_s. This line L may intersect the gate voltage axis at a point in the region of forward bias as in Fig 10,9. In such a case, an external gate voltage which will forward bias the gate must be used. Instead of using a separate dc source to supply the forward bias, we can use a resistive divider which will give us $V_{gg} = aV_{dd}$ where 'a' is the divider ratio. Fig.10-10a shows such a configuration. The value of the resistance R_d must be so chosen that the drain to source voltage will remain within the usable operating area of the output characteristics which is bounded by the voltage saturation characteristic at the low voltages and by the locus of points at which apparent drain to source breakdown occurs at high voltages .(See Fig 10-10b Usually the values of R_s and V_{gg} are increased to take care for temperature variations of I_g (the gate leakage current). Both in Fig 10-10a and Fig I0-8 , the resistance R_s must be bypassed by a capacitor to obtain high ac gain.

The above mentioned methods of stabilization yield a reduction in the spread drain currents, whether the variation is due to difference of characteristics between units or to temperature changes. Other methods may also be used to compensate for

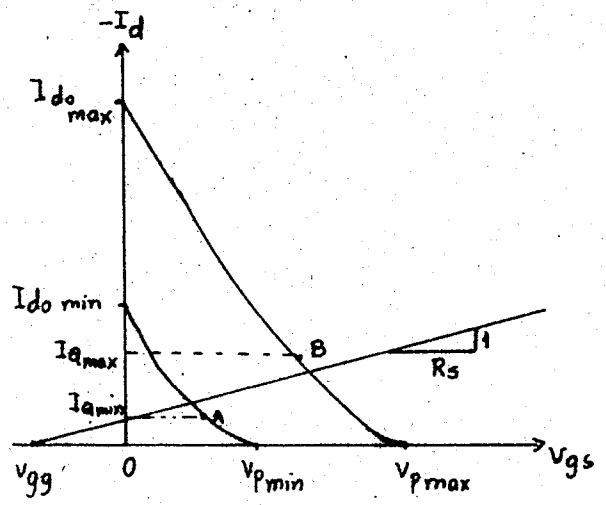


Fig. 10-9 FET transfer curves.

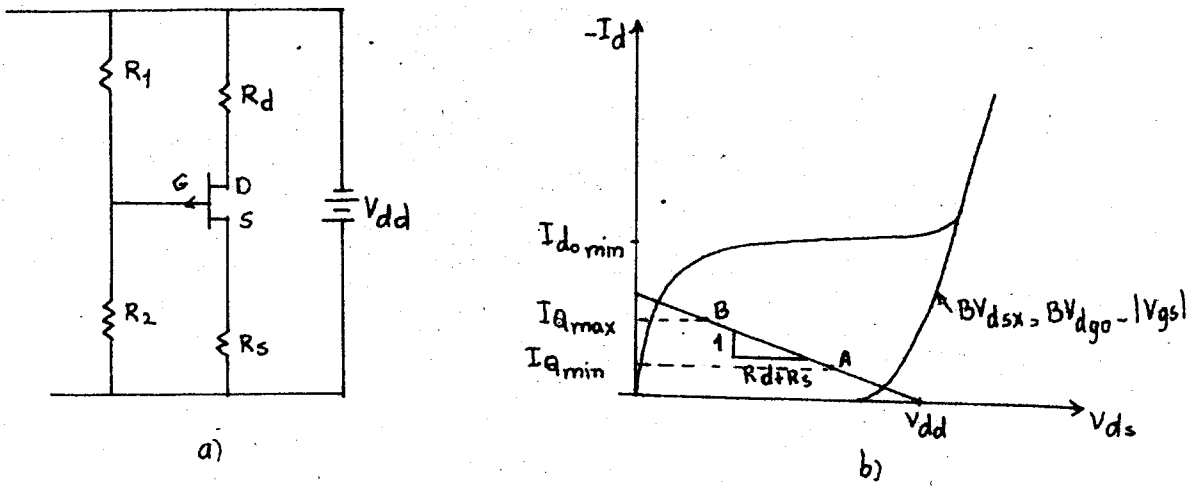


Fig.10-10 a) CS FET amplifier using a voltage divider to supply forward bias to the gate.

b) Output characteristics used to find the value of Rd.

temperature changes.

The best way to compensate for the drift caused by gate leakage current is to use a matched pair of FETs in a direct - coupled difference amplifier often called " The long tailed pair" dc amplifier. See Fig 10-11a. The circuit is symmetrical. Two sources are connected together through small balancing resistors R_s and the output is taken as the voltage difference between the two drains. Fig 10-11b shows a simplified equivalent circuit to be used for analysis. In this figure F_1 and F_2 represent ideal driftless units and current and voltage generators represent drift. The equivalent input drift is $\frac{\Delta V_{in}}{\Delta T}$, and the equivalent output due to this input is $\frac{\Delta V_o}{\Delta dT}$ which must be equal to zero. Writing Kirchoff's voltage rule around the input loop.

$$\frac{\Delta V_{in}}{\Delta T} = \frac{\Delta(V_{gs_1} - V_{gs_2})}{\Delta T} + \frac{\Delta(I_{g_1} - I_{g_2})}{\Delta T}(R_g + R_s) \qquad (10\text{-}50)$$

Here $\Delta(V_{gs_1} - V_{gs_2})/\Delta T$ is equivalent input voltage drift of the FET pair and $\Delta(I_{g1} - I_{g2})/\Delta T$ is the equivalent input current drift of the circuit. Thus we see that both $\Delta V_{gs}/\Delta T$ and $\Delta I_g/\Delta T$ should be matched between the two FETs for minimum drift. $\Delta V_{gs}/\Delta T$ for a given FET is completely specified by I_{do} and V_p. (V_p is pinch off voltage=same as W_o). Matching these two quantities will therefore ensure minimum drift at any bias point. As for the leakage term, the drift can easily be reduced to negligible proportions by matching the gate currents at the highest working temperatures since the current I_g is in all cases exponentially related to temperature.

Another technique of drain current stabilization is the tandem stabilization shown in Fig 10-12. In this set up the source resistor of the normally biased common source FET amplifier is replaced by a bipolar transistor which has its collector current stabilized by any of the conventional means. Since the source and drain current of the FET are equal, the stabilization of the drain current is identical with that of the bipolar transistor. To ensure satisfactory operation, we must avoid bottoming the bipolar transistor at the temperature at which the FET would have

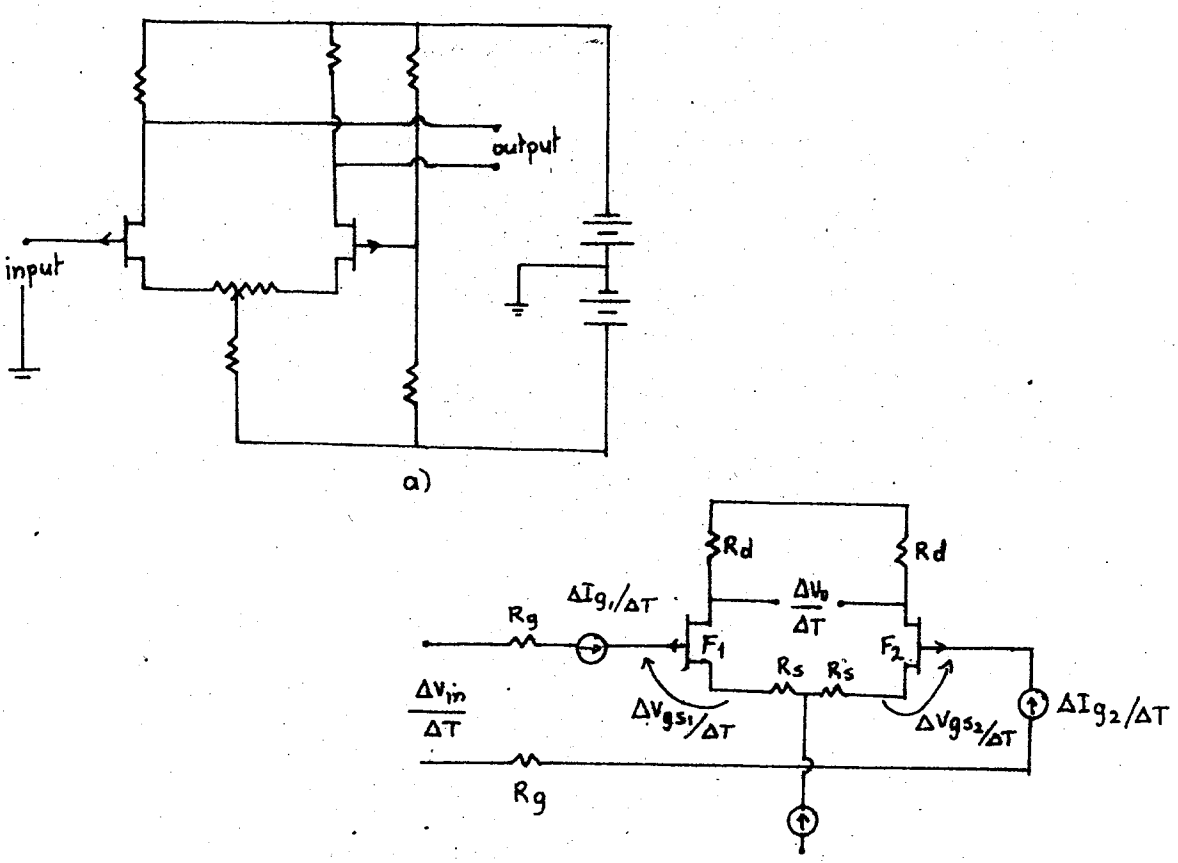


Fig. 10-11. Dc difference amplifier and its drift equivalent.

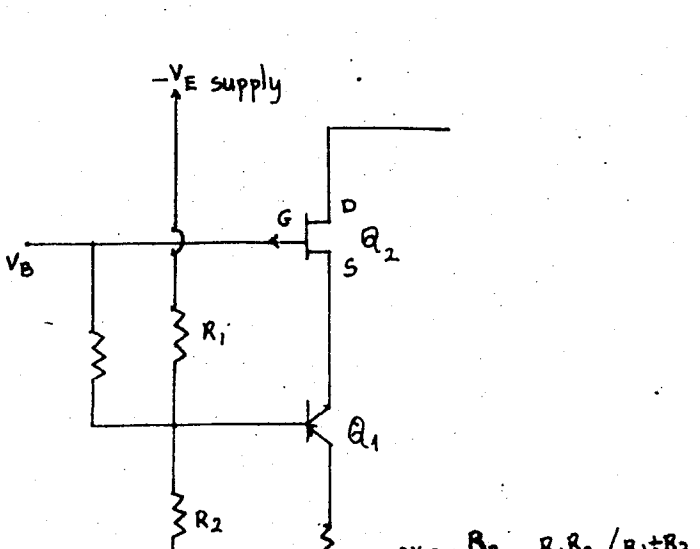


Fig 10-12. Tandem stabilization using a FET and a bipolar transistor.

the lowest drain current (maximum temperature for most units). When the bipolar transistor Q_1 is stabilized

$$\frac{\partial I_c}{\partial T} = S\,\frac{\partial I_{co}}{\partial T} = \frac{R_E + R_B}{R_E + R_B(1-\alpha)} \cdot \frac{I_{co}}{a} = \frac{\partial I_d}{\partial T} \qquad (10\text{-}51)$$

where S = stabilization ratio

R_E = resistor in the emitter of the bipolar transistor ,

$R_B = R_1 R_2 / (R_1 + R_2)$

α = common base current gain of the bipolar transistor ,

I_{co} = bipolar transistor collector cut off current at the temperature at which $\partial I_d/\partial T$ is to be calculated

a = a constant having the dimensions of temperature. 'a' is between 10 and 14°C for currently available transistors

Eq. (10-51) will be correct only if the voltage drop between base and emitter of the bipolar transistor (of the order of 0.33v for Ge and 0.6 v for Si) is negligible compared with the voltage drop V_{CE} across the emitter resistor.

$$V_{CE} = V_{gg} - V_{gs} - R_E I_E \qquad (10\text{-}52)$$

where V_{gg} = standing bias applied to the FET gate

I_E = emitter current

V_{gs} = gate-source voltage for the FET at the required drain current. Since $I_E = I_D$, and $I_D \cong I_{do} - V_{gs} g_m$, we can calculate V_{CE} from the following equation :

$$V_{CE} = V_{gg} - \frac{I_{do} - I_d}{g_m} - R_E I_d \qquad (10\text{-}53)$$

provided that values of I_{do} , and g_m appropriate to the temperature concerned are used. A conservative design criterion for low collector currents is to keep V_{CE} above 1 v .

10-3. Distortion in FET Amplifiers :

The nature of the FET's square law behavior permits both the harmonic and intermodulation distortion of a single stage amplifier to be expressed in a rather simple closed form. Harmonic

distortion in an FET amplifier is determined by many elements in the circuit; bias , operating voltage , load impedance , signal level, device characteristics and variations in input impedance. It is a function of the extent to which the device transfer curve departs from a straight line. By definition, the harmonic distortion produced by an amplifier whose input is a pure sinusoid is

$$D = \frac{r_{ms} \text{ value of all harmonics in the output}}{r_{ms} \text{ value of the fundamental of the output}} \qquad (10\text{-}54)$$

when two or more sine wave signals are summed in a non-linear network, beat frequencies are produced. The distortion produced by these beat frequencies is called intermodulation (IM)distortion

$$IM = \frac{r_{ms} \text{ value of the beat frequencies}}{r_{ms} \text{ value of the complex input}} \qquad (10\text{-}55)$$

10 -3A. Harmonic Distortion:

The single stage amplifier circuit of Fig 10-13 may be used to determine the effect on distortion of variations in e_s , r_{gs} R_L and FET characteristics V_p , g_{ds} , and g_m .

a) Transfer curve and conductances:

As the input signal varies, so do the two conductances g_m and g_{ds} which effect the gain of the stage.

Follow the instantaneous operating point up and left along the load line drawn on the FET output characteristics curves in Fig 10-14. Transconductance g_m increases as V_{gs} approaches zero. This variation in g_m is reflected in the curvature of the transfer curve for the FET. The curvature causes second harmonic distortion. Nonlinearities also exist because g_{ds} also varies with the instantaneous operating output current - g_{ds} increases as V_{gs} approaches zero, and as I_d increases. From the gain equation for the FET:

$$\frac{e_o}{e_s} = \frac{g_m}{1/R_L + g_{ds}} \qquad (10\text{-}56)$$

Eq. (10-56) shows that the effect of a change in g_{ds} is less for large $1/R_L$. As V_{gs} increases , the decrease in g_m is partially offset by a corresponding decrease in g_{ds}.At a certain operating

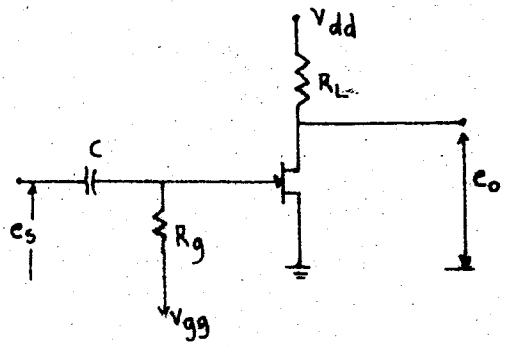


Fig.10-13. Single stage FET amplifier used to determine on distortion of various device and circuit parameters .

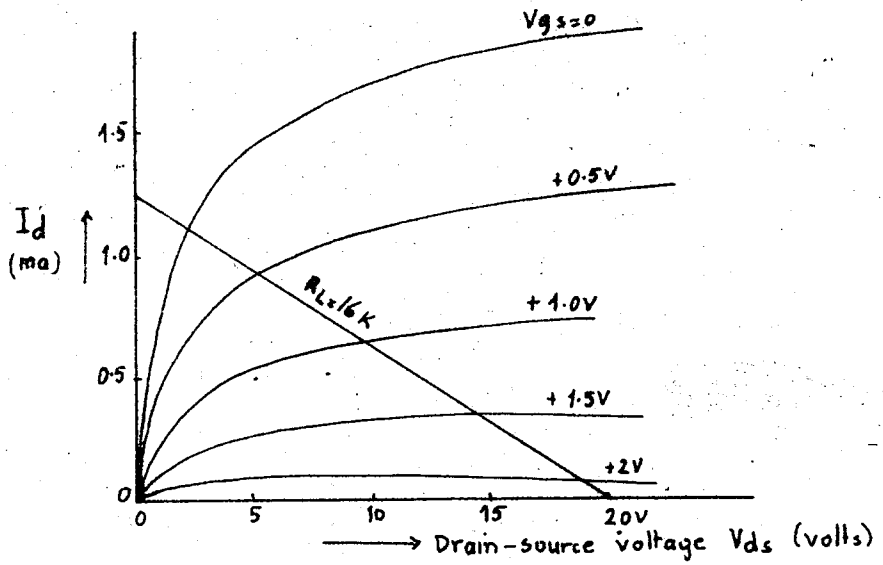


Fig. 10-14. Output characteristics of the FET used in Fig.10-13 , showing the load line and variation of g_{ds} and g_m as one moves along the load line.

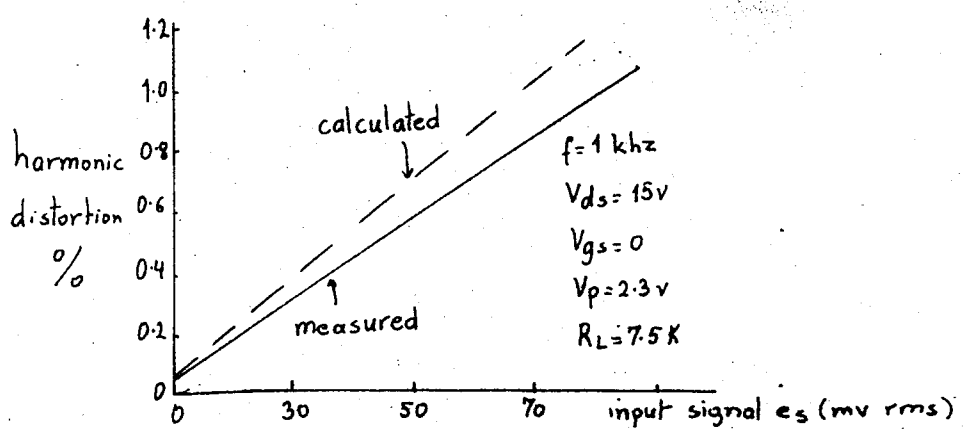


Fig. 10-15. Harmonic distortion vs input signal. Calculated curve is plotted from Eq. 10-60.

point the two distortion sources will nearly cancel to produce a point of minimum distortion.

When a sinusoidal voltage v_g is applied to the input, the total instantaneous gate to source voltage can be written as the sum of the quiescent bias voltage V_{aq} and the instantaneous input signal.

$$V_{gs} = V_{aq} + V_1 \sin \omega t \qquad (10\text{-}57)$$

The total drain current, using the square law approximation is

$$i_d = \frac{I_{do}}{V_p^2} \left(V_{aq} - V_p + V_1 \sin \omega t \right)^2 \qquad (10\text{-}58)$$

$$= \frac{I_{do}}{V_p^2} \left[(V_{aq} - V_p)^2 + \frac{V_1^2}{2} + 2 (V_{aq} - V_p) V_1 \sin \omega t - \frac{V_1^2}{2} \cos 2\omega t \right] \qquad (10\text{-}59)$$

Thus FETs promise to generate only second harmonic distortion and the expression for percent second harmonic distortion is :

$$\% D = \frac{25 V_1}{V_{aq} - V_p} \qquad (10\text{-}60)$$

In a more convenient way, this can be written as

$$\% D = 25 \frac{V_1}{V_p} \sqrt{I_{do}/I_Q} \qquad (10\text{-}61)$$

This expression is valid for small signal distortion due only to the curvature of the transfer characteristic. Required conditions are that the drain voltage saturation region is avoided and that the drain current flows during all portions of the cycle.

b) Gate bias and signal level:

According to Eq 10-60 small signal distortion is directly proportional to the input signal level. However distortion also depends upon the difference between the pinch off voltage and the voltage on the gate(see Eq.10-61). Fig 10-15 shows the total harmonic distortion as a function of input signal. The reason the distortion increases with the input signal is that the gate draws current from the signal source on input peaks. Distortion is reduced by adding the equivalent of grid-leak bias. The gate-return resistor R_g shown in the amplifier circuit of Fig 10-13, develops a gate bias approximately equal to the peak forward signal voltage.

So long as the input capacitor remains charged, the gate ceases to draw current on signal peaks. Thus, distortion becomes a function of frequency. Usually RC is chosen about 1000 times the period of the lowest frequency to be handled. Fig 10-16 shows the harmonic distortion as a function of input signal with RC as a parameter.

Eq.10-60 also indicates that distortion increases as the quiescent operation voltage V_{Qg} approaches V_p . For $V_{Qg}=0$, distortion is inversely proportional to V_p . Fig 10-17a shows a distortion vs V_{Qg} plot of a FET amplifier for two values of input signal. Fig 10-17b plots distortion against V_p for $V_{Qg}=0$. The calculated and measured values agree closely in both plots.

c) Drain voltage:

The effects of V_{ds} and g_{ds} on distortion are illustrated in Fig 10-18a and 10-18b. The V_{ds} effect is in reality a g_{ds} effect as may be seen from the g_{ds} vs V_{ds} curves of Fig 10-18b. Distortion due to g_{ds} becomes large at low V_{ds} due to the very high value of g_{ds} while the distortion at high V_{ds} is due principally to the variation of g_m. At a specific and fairly low V_{ds} the two distortions nearly cancel to produce a minimum in the distortion curve (Fig 10-18a). But operation at the point of minimum distortion is not recommended because its location is uncertain and very near a region of excessive distortion . Operating an FET with the drain voltage near breakdown has been reported as one method for improving bias stability.

d) Long ,short and medium gates.

The physical geometry of the FET also has an effect on distortion. This is due to the effect of the gate length on the value of g_{ds} and the rapidity with which the device output characteristics shift form triode to pentode type as V_{ds} increases. FETs with longer gates show less distortion but have low saturation voltages. Thus a compromise must be made between low distortion and low saturation.

e) Load resistance

Distortion decreases with increasing R_L. The reason for the improvement with increasing R_L is that the g_{ds} induced distortion

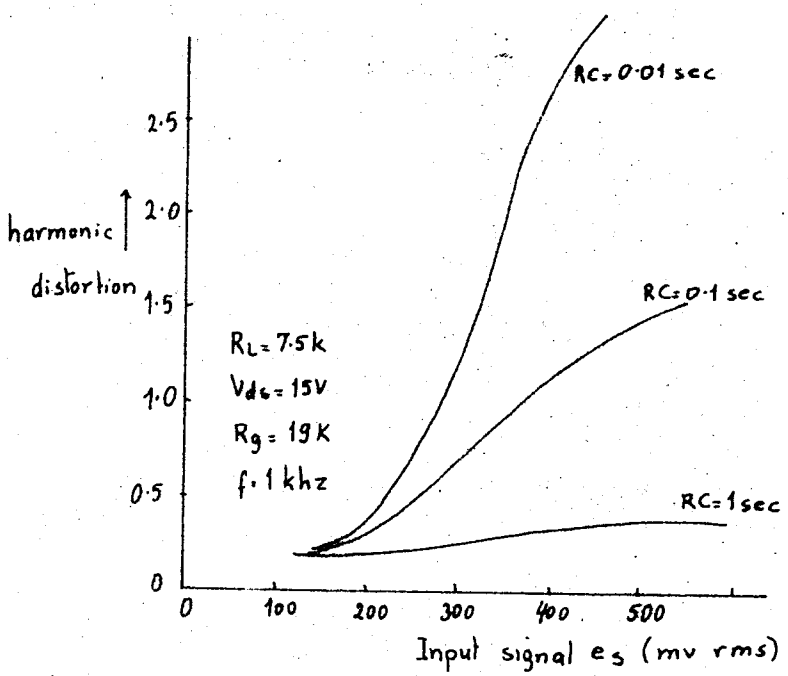


Fig. 10-16. Harmonic distortion vs input signal with RC as a parameter.

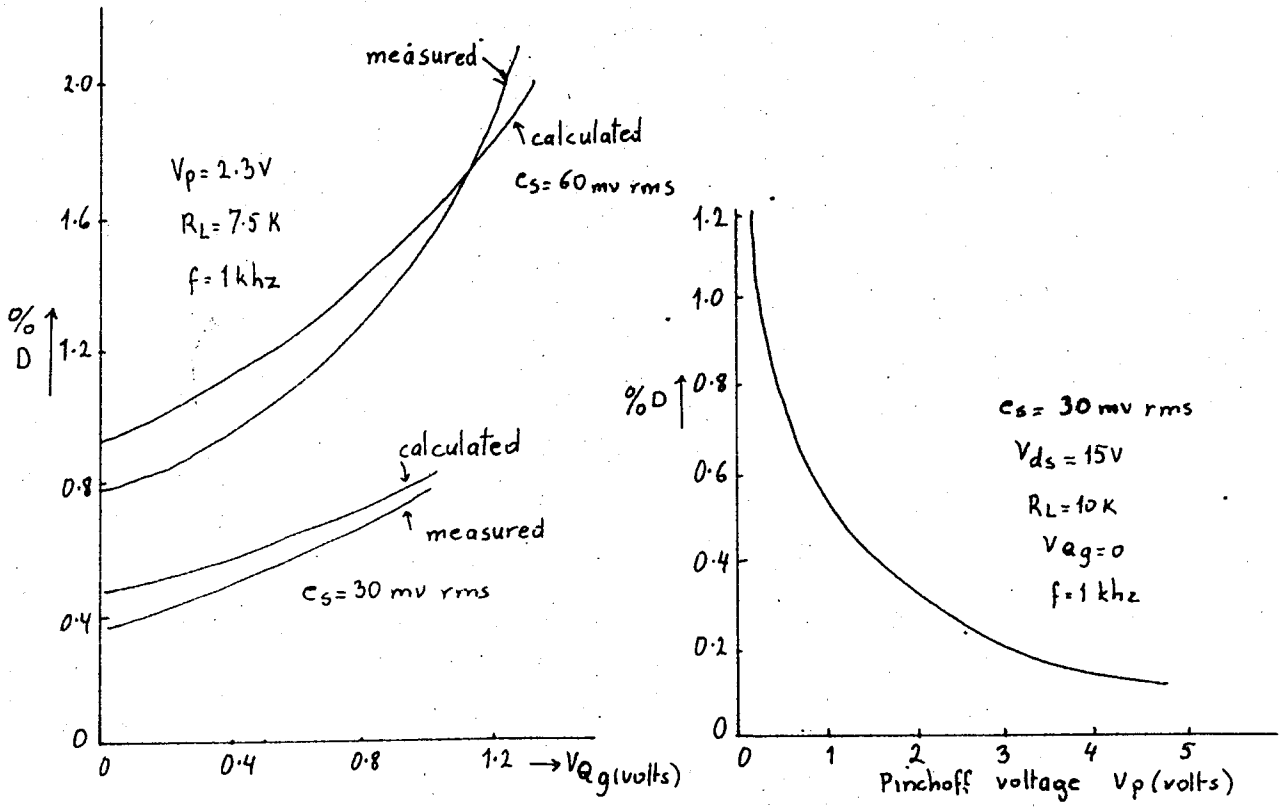


Fig.10-17 a) Harmonic distortion vs quiescent voltage V_{Qg}. (2N2608 FET)

Fig.10-7 b) Harmonic distortion vs V_p with V_{Qg}=0

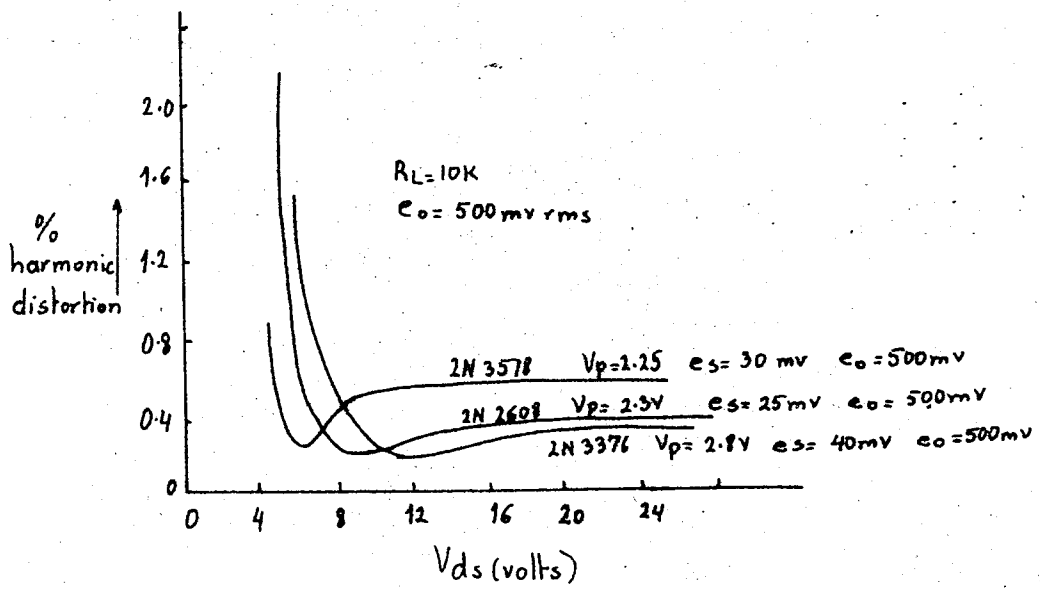


Fig. 18a. Harmonic distortion vs V_{ds}. Drain to source voltage must remain well above the pinch off voltage for distortion to be small .

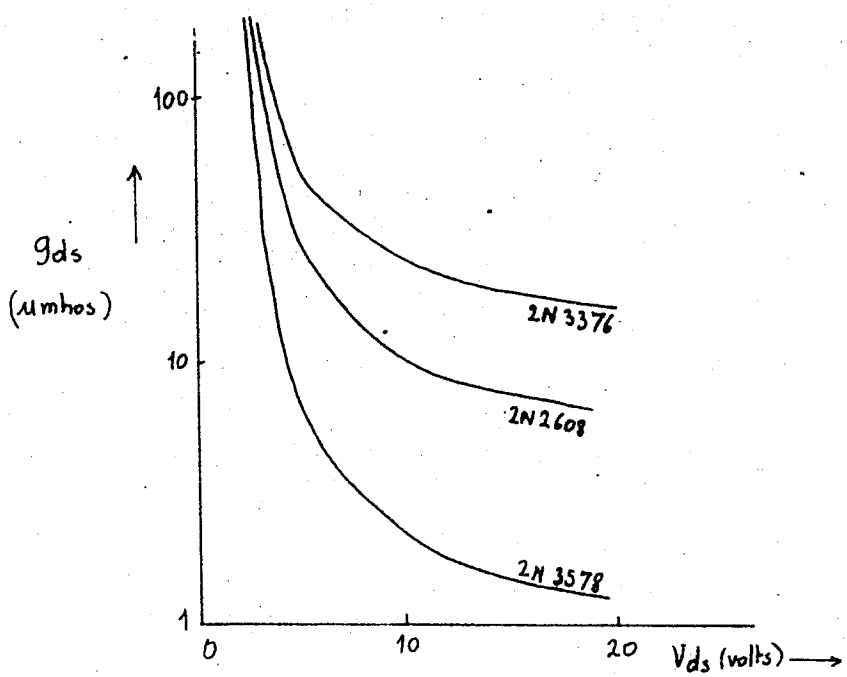


Fig 18b. g_{ds} vs V_{ds} for three FETs with different gate lengths.
2N3578 is a very long gate FET
2N2608 is a medium gate-length device
2N3376 is a short-gate device.

has a greater opportunity to counteract g_m distortion as the load line becomes more nearly horizontal.

10-3B. Intermodulation distortion.

The standard method of measuring I.M distortion of an amplifier consists of summing two sinusoids of known frequency and amplitude ratio at the input and measuring the resulting intermodulation components in the output signal.

$$I_d = \frac{I_{do}}{V_p^2}\left(V_{Qg} - V_p + V_A \sin \omega_1 t + V_B \sin \omega_2 t\right)^2 \qquad (10\text{-}62)$$

when Eq. (10-62) is expanded and only the terms containing the fundamental and intermodulation frequencies are considered, we get

$$i_d = 2\left(V_{Qg} - V_p\right)\left(V_A \sin\omega_1 t + V_B \sin\omega_2 t\right) + V_A V_B\left[\cos(\omega_1 - \omega_2)t - \cos(\omega_1 + \omega_2)t\right] \qquad (10\text{-}63)$$

so that the rms values of the IM term and fundamental are divided:

$$IM = \frac{V_A V_B}{\left(V_{Qg} - V_p\right)\sqrt{2\left(V_A^2 + V_B^2\right)}} \qquad (10\text{-}64)$$

$$= \frac{V_A V_B}{V_p\sqrt{2\left(V_A^2 + V_B^2\right)}}\sqrt{I_{do}/I_Q} \qquad (10\text{-}65)$$

where I_Q = dc drain current at operation point.

$$IM = \frac{g_m V_A V_B}{2 I_Q \sqrt{2\left(V_A^2 + V_B^2\right)}} \qquad (10\text{-}66)$$

10-4 Hybrid Amplifiers:

Amplifiers which use FET's and bipolar transistors alternately can show very high power gain. They can be unilateral and may also have other attractive circuit features. The FETs are usually used in the input stages because they have high input impedances. Thus nine configurations are possible, but of these the ones using CG input stages are not of interest because of their low input impedance and low power gain. The types most widely used are CS-CB , CS-CE, CD-CB, CD-CE, and CS-CC amplifiers.

Fig 10-19 -10-21 show in outline form the hybrid amplifier setups
Note that CS-CB stands for common source -common base
CS-CE stands for common source -common emitter
CD-CB stands for common drain -common base
CD-CE stands for common drain -common emitter
CS-CC stands for common source -common collector.

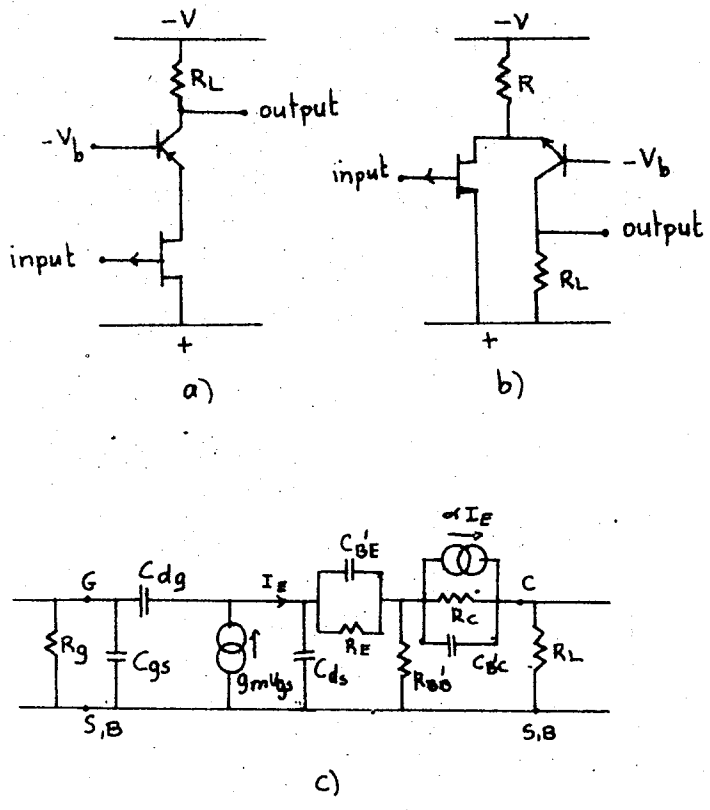


Fig 10-19. CS -CB amplifier
a) cascode CS-CB amplifier
b) the folded cascode circuit
c) the ac equivalent circuit.

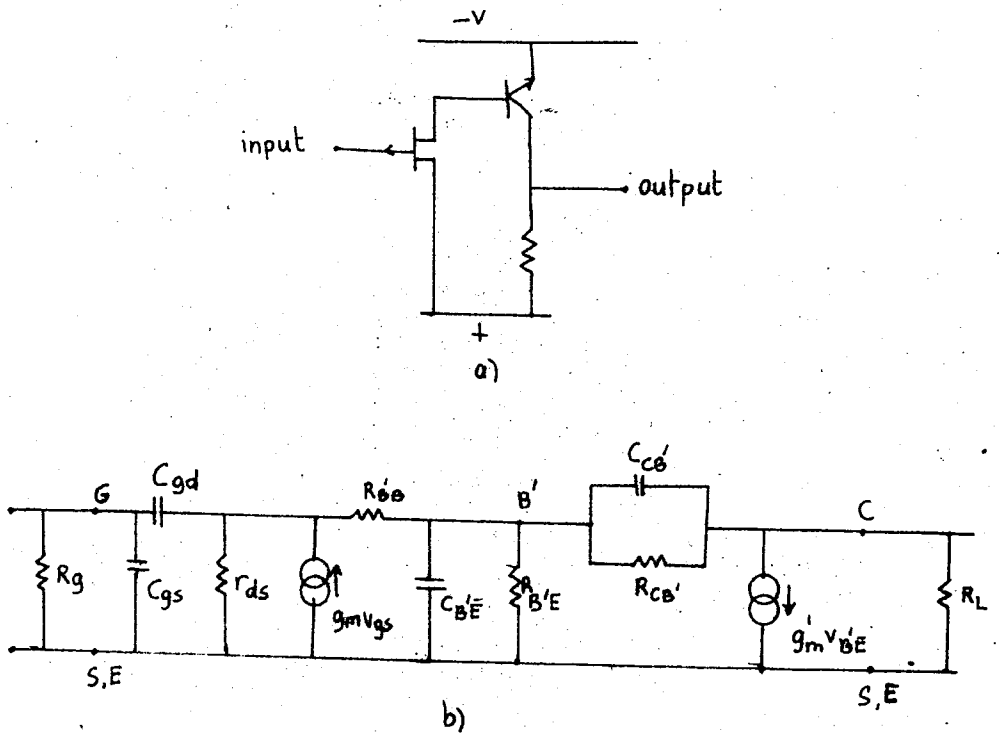


Fig 10-20. CS-CE amplifier

a) cascode folded CS-CE amplifier

b) ac equivalent circuit with the bipolar transistor represented by the hybrid π equivalent

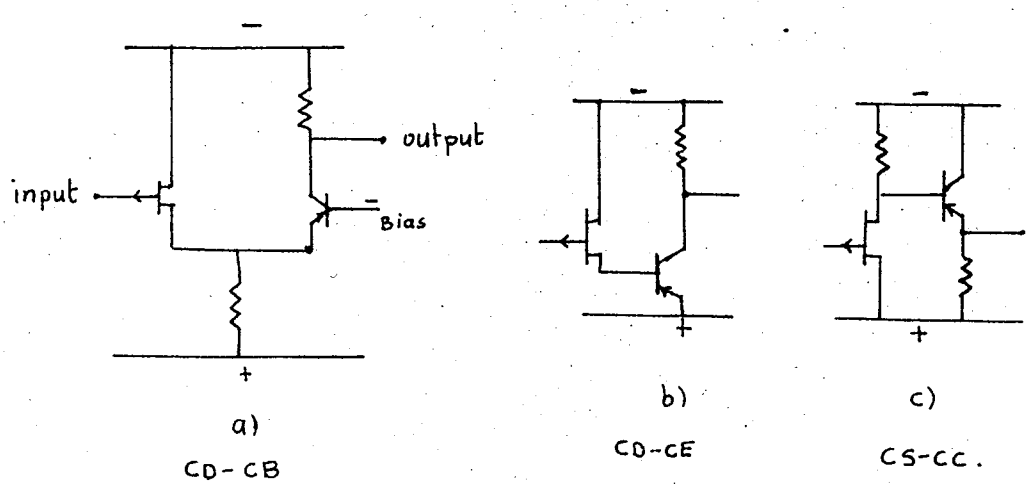


Fig 10-21. Hybrid amplifier configurations

CHAPTER II

LINEAR MOS TRANSISTOR AMPLIFIER CIRCUITS AND FIELD-EFFECT ATTENUATOR CIRCUITS

Depletion mode insulated gate FETs have properties closely similar to those of the junction gate type and can be used in the same circuits. Only two points of difference must be noted. Since the gate is insulated, the gate leakage current is much smaller than that of the junction gate FET, and can be neglected in determining bias conditions. For the present, the MOSFETs are not yet considered low noise device and therefore are not suitable for small signal applications. Enhancement mode units differ from junction gate FETs in the same way as the depletion units do, and in addition require substantially different circuits.

II-I. Single-Stage Amplifiers

There are three basic single-stage amplifier configurations: Common source, common gate, and common drain. In the following discussions the n-channel MOS transistor will be considered unless otherwise noted.

The characteristics of each configuration can be summarized as follows:

a) Common-source amplifier stage

The circuit set up is seen in Fig II-1a. It is characterized by a high input impedance, medium to high output impedance and a voltage gain greater than one.

The voltage gain A_v without feedback is the same as that of the junction gate FET CS configuration.

$$A_v = \frac{g_m r_{ds} R_L}{r_{ds} + R_L} \qquad (II-1)$$

If $R_L \gg r_{ds}$

$$A_v = \mu \qquad (11\text{-}2)$$

If $R_L = r_{ds}$

$$A_v \cong \mu/2 \qquad (11\text{-}3)$$

If $R_L \ll r_{ds}$

$$A_v \cong g_m R_L \qquad (11\text{-}4)$$

When an unbypassed source resistor is introduced into the CS circuit, it produces a negative feedback proportional to the output current. If A_v' is the voltage gain of this circuit,

$$A_v' = \frac{g_m r_{ds} R_L}{r_{ds} + (g_m r_{ds} + 1)R_s + R_L} \qquad (11\text{-}5)$$

where R_s = unbypassed source resistor.
The output impedance of the CS circuit with unbypassed R_s becomes

$$Z_o = r_{ds} + (g_m r_{ds} + 1)R_s \qquad (11\text{-}6)$$

b) Common-drain amplifier

The circuit arrangement is shown in Fig 11-1b. The input impedance is higher than that of the CS. The output impedance is low and there is no phase shift. Voltage gain is less than one and distortion is low. The input circuit capacitance is reduced. The voltage gain A_v is

$$A_v = \frac{R_s}{[(\mu+1)/\mu]R_s + 1/g_m} \qquad (11\text{-}7)$$

If $\mu \gg 1$, Eq.(11-7) reduces to

$$A_v' = \frac{g_m R_s}{1 + g_m R_s} \qquad (11\text{-}8)$$

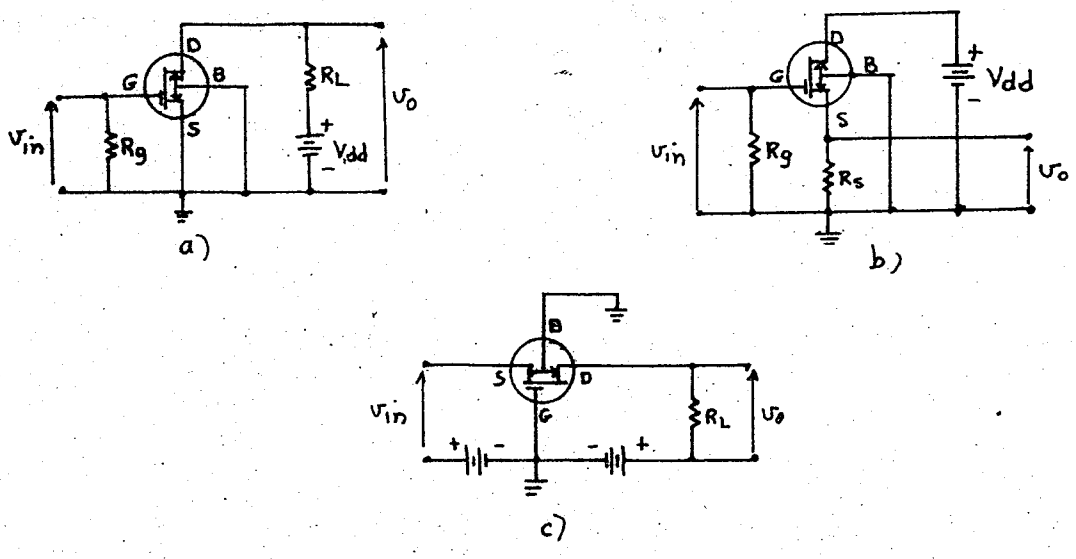


Fig 11-1 Basic circuit configurations : a) Common source ; b) common drain ; c) common gate

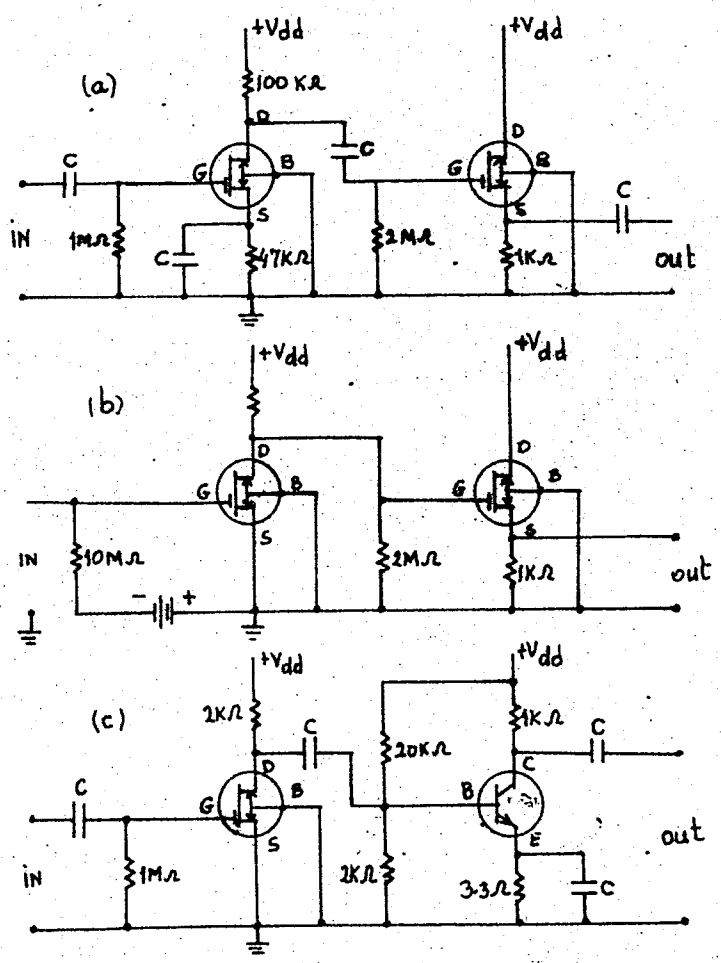


Fig 11-2. Typical cascade circuits using n-channel MOS transistors.

When the resistance R_g is returned to ground, the input resistance R_i is:

$$R_i = R_g \qquad (11\text{-}9)$$

When the resistance R_g is returned to the source, the input resistance R_i' is:

$$R_i' = \frac{R_g}{1 - A_v} \qquad (11\text{-}10)$$

When the load is resistive, the effective input capacitance C_i' of the CD is reduced to

$$C_i' = C_{gd} + (1 - A_v)C_{gs} \qquad (11\text{-}11)$$

The output resistance R_o is

$$R_o' = \frac{r_{ds}R_s}{(g_m r_{ds} + 1)R_s + R_{ds}} \qquad (11\text{-}12)$$

and the output capacitance C_o' is given by

$$C_o' = C_{ds} + C_{gs}\left(\frac{1 - A_v}{A_v}\right) \qquad (11\text{-}13)$$

c) Common-gate amplifier

Fig. 11-1c shows a CG configuration. It is used for impedance matching and also in high frequency circuits. It has a low voltage gain.

$$A_v = \frac{(g_m r_{ds} + 1)\, R_L}{(g_m r_{ds} + 1)R_g + r_{ds} + R_L} \qquad (11\text{-}14)$$

If $g_m r_{ds} \gg 1$ and $r_{ds} \gg R_L$

$$A_v = \frac{g_m R_L}{1 + g_m R_g} \qquad (11\text{-}15)$$

11-2. Cascaded Amplifiers.

MOS transistors may be used in many different cascade arrangements either with other field effect transistors or with bipolar transistors or vacuum tubes. Fig 11-2 gives a few of the more likely combinations. Fig 11-2a has a high input and low output impedance. In Fig 11-2c, the MOS transistor is used as an impedance transformer for the bipolar transistor. With the cascode circuits of Fig 11-3a and b, advantages such as automatic gain control, improved signal to noise ratio and reduced feedback capacitance are obtained. The output resistance of these series cascode pair reaches its height, a constant value, when the drain to source voltage of the driving unit just exceeds the sum of the enhancement voltage on the gate of the driven unit plus the pinch-off voltage of the driven unit. Any further increase in the drain voltage of the driving unit causes the gate of the driven unit to become proportionately more negative than its own source.

11-3. Biasing MOS Transistors

Fig. 11-4 shows the insulated gate transistor with a wide variety of biasing arrangements. The circuits are so designed that there is no power loss due to biasing elements. Circuits a, b, c, and d of Fig 11-4 may be used to obtain a quiescent operating point in the enhancement mode . Of these , Fig 11-4b may be used when the larger drain currents associated with enhancement mode operation can be tolerated -i.e., an inductive load or a low value of resistive load.

The circuit of c provides an increased degree of stability because of the negative feedback. If I_d increases due to a change in temperature, V_d decreases because of the additional voltage drop across R_d and the positive gate voltage, V_g , is decreased. This tends to cancel out the initial drain current increase.

Circuit (d) is a very critical circuit. It has a positive

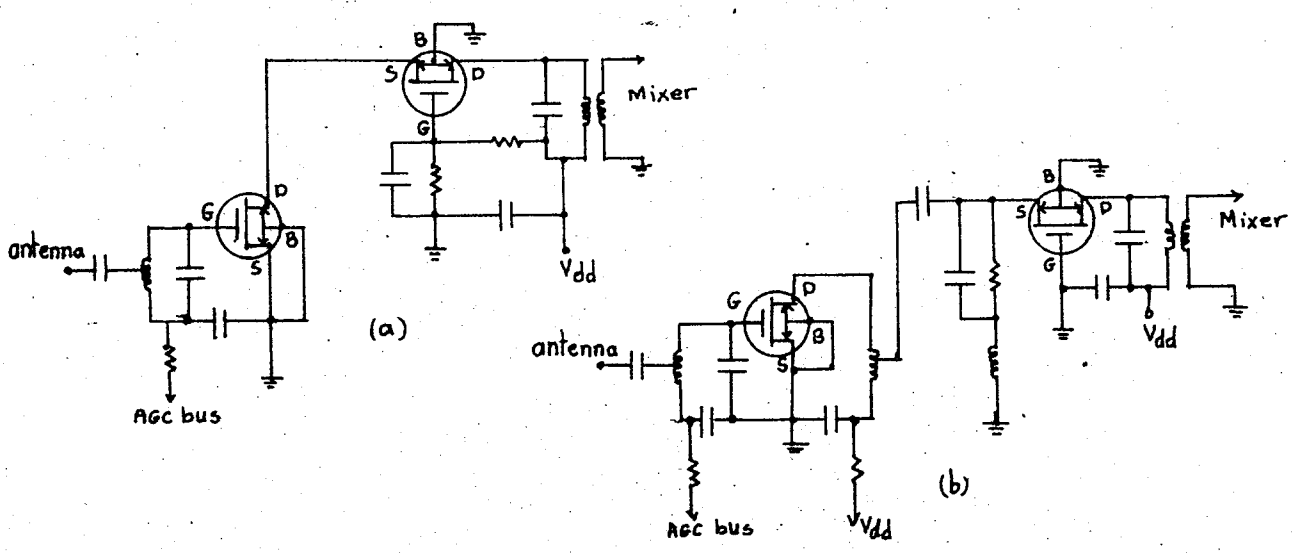


Fig 11-3 Cascode circuits. a) series cascode ; b) parallel cascode.

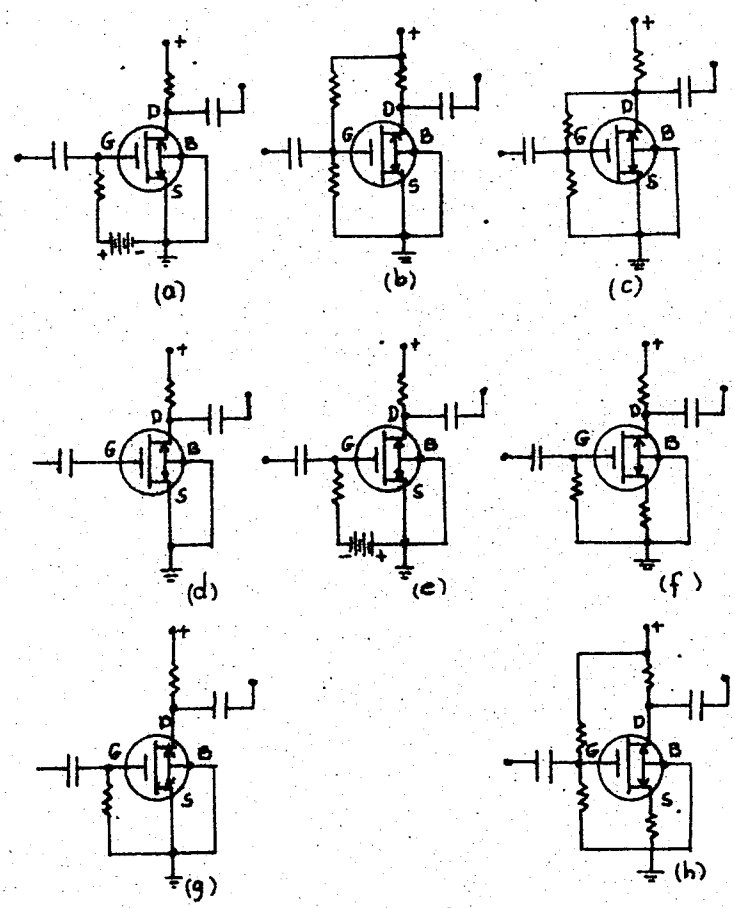


Fig 11-4. Transistor bias circuits for n-channel MOS transistors.

gate bias which is obtained from its internal voltage divider formed by the gate to source and gate to drain leakage resistances. It is used when it is necessary to use the maximum input resistance capabilities of the MOS transistor. For an off set gate depletion type transistor, with no accumulated gate charge the positive bias thus obtained may be about 5% of the drain supply voltage. But in such a case since the gate is floating, it must be protected against damage due to electrostatic voltages. Electrostatic grounding strap which may have an impedance to ground of several megohms may be used to protect the insulation layer from being punctured.

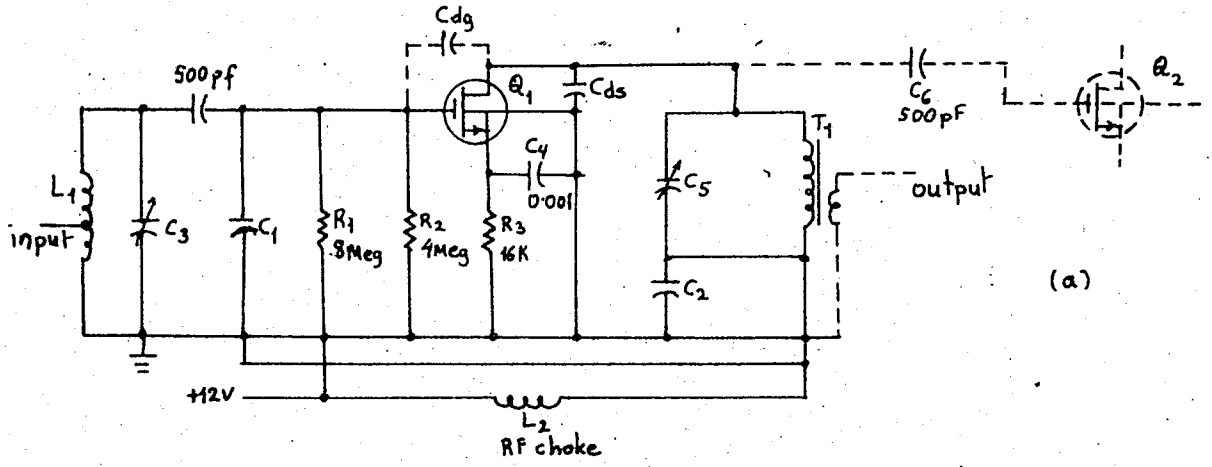
Circuits (e) and (f) are used for operation in the depletion mode. In circuit (f) if I_d increases due to temperature effects, V_d decreases because of the additional voltage drop across both R_d and R_s . The increased voltage drop across R_s (the negative feedback resistor) makes the gate to source voltage more negative. The net result is to restore V_d to its original value. Circuit (f) may be used for low drain current applications in which the load impedance is large and the loss in dc gain due to R_s can be tolerated . In ac amplifier circuits, R_s may be bypassed with a capacitor.

Circuit (g) is used for zero bias operation except for the very large internal gate to source resistance which is of the order of 10^{15} ohms. Due to this internal resistance, the quiescent operating point may be in the enhancement mode as for the circuit of Fig 11-4d.

Circuit (h) is a combination of circuits(b) and (f). A positive voltage is obtained from the divider across the supply voltage which functions to cancel the negative gate bias to a desired degree.

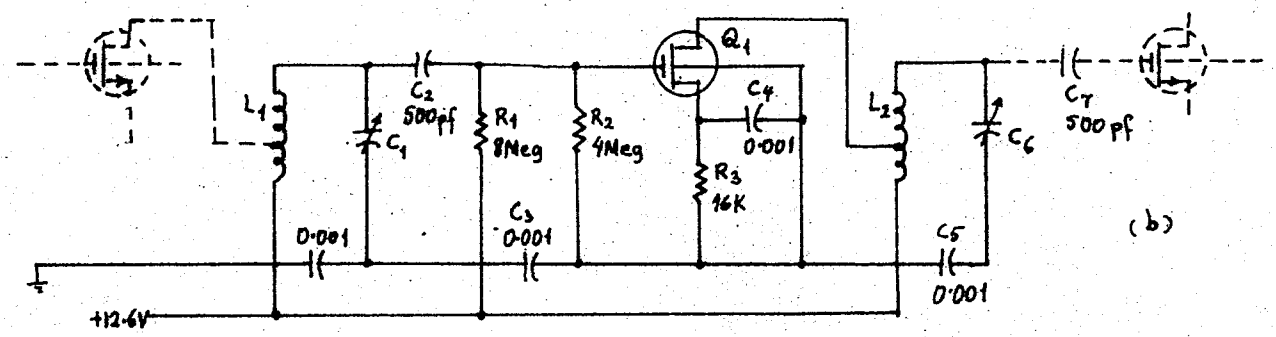
11.4. Some Practical Applications of MOSFETs in Communication:

Fig 11-5 shows some practical amplifier stages . The neutralized common source amplifier of Fig 11-5a delivers the highest



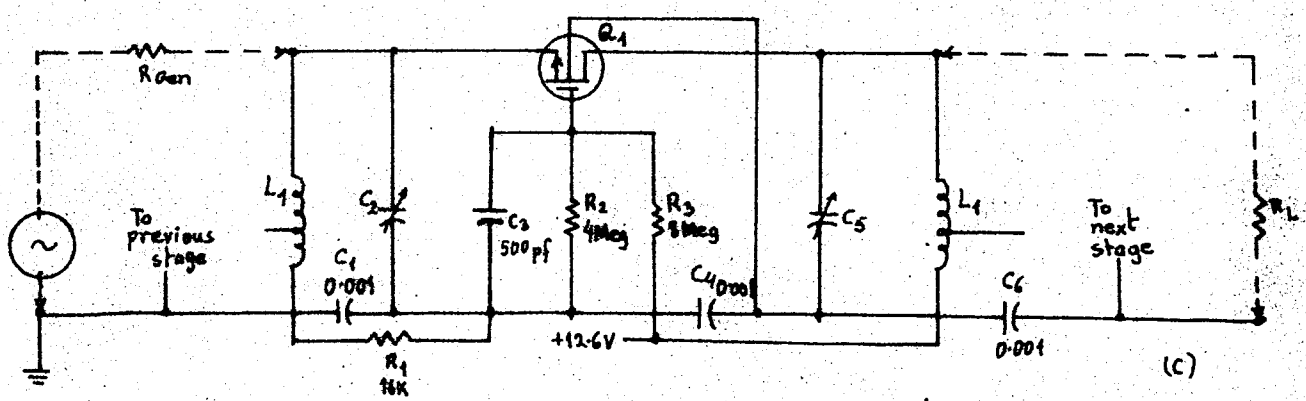
(a)

Neutralized common source amplifier designed with a MOSFET delivers power gain of 20 db at 100 megacycles.



(b)

Unneutralized common source amplifier uses MOSFET. Low Cdg is needed for high power gain.



(c)

MOSFET common-gate amplifier is similar to grounded grid vacuum-tube amplifier. Power gains of 15 db can be obtained at 200 megacycles per sec

Fig 11.5 Some amplifier stages used in communications.

theoretical and practical power gain. Neutralization is accomplished if:

$$\frac{C_{ds}}{C_{dg}} = \frac{C_2}{C_1} \qquad (11\text{-}16)$$

The power gain of a single amplifier stage such as the one of fig 11-5a , is given by

$$P_g = \frac{(g_m^2)}{4\left(\frac{1}{r_d} + \frac{1}{R_T}\right)\left(\frac{1}{R_{in}} + \frac{1}{R_T}\right)} \qquad (11\text{-}17)$$

where R_T = the unloaded impedance of the input and output tank circuits at resonance.

$$R_{in} = R'_o + 1/(wC_{gc})^2 R'_o) \qquad (11\text{-}18)$$

R'_o = loss resistance ($\cong$ 250 ohms for DPT200)
C_{gc} = capacitance between gate metal and channel (approximately 0.8 picofarad)

For high frequency operation, the transconductance measured at low frequency (gm_{LF}) must be converted to its high frequency value gm_{HF} . This is accomplished by

$$gm_{HF} = gm_{LF} / .(1 + jwC_{gc}R'_o) \qquad (11\text{-}19)$$

The high frequency transconductance may also be obtained from relation ship between frequency and transconductance :

$$g_{mHF} = g_{mLF} / \sqrt{1 + (f/fcog_m)^2} \qquad (11\text{-}20)$$

where f = operating freq.
$fcog_m$ = transconductance cut-off frequency
$= 1/2\pi (R'_o)(C_{gc})$.

If several identical amplifiers are cascaded, the power gain per stage reduces to

$$P_g = \frac{(g_m)^2}{\left(\frac{1}{r_d} + \frac{1}{R_T} + \frac{1}{R_{in}}\right)^2} \qquad (11\text{-}21)$$

with high Q circuits a typical power gain of 20 decibels at 100 megacycles per second can be obtained. The theoretical high - frequency noise figure F_{oHF} is

$$F_{oHF} = 1 + \frac{K\,(C_{gc} + C_{dg})\,w}{g_m} \qquad (11\text{-}22)$$

where K = device constant determined by transistor geometry $\cong 2$
Measurements for this circuit indicate a noise figure of 3 to 4 db at 100 megacycles per second.

Fig 11-5b shows an unneutralized common-source amplifier. The resonance resistance R_{TL} for the loaded tank circuit is given by:

$$R_{TL} = 1 \,/\, [\,(\,1/R_{in} + 1/R_T)] \qquad (11\text{-}23)$$

The drain terminal of the MOS transistor is connected to a tap on the output tank to provide a load resistance R_L of

$$R_L = r_d \,/\, [\,w g_m\, C_{dg}\, R_T - 2\,] \qquad (11\text{-}24)$$

The power gain per stage for each of several identical unneutralized amplifiers is :

$$P_g = g_m^2\, R_{TL}\, r_d \,/\, [(g_m\, w\, C_{dg}\, R_{TL}\, r_d) - 2\,] \qquad (11\text{-}25)$$

At high frequencies (2 megacycles and above) , equation (11-25) reduces to

$$P_g \approx g_m \,/ w C_{dg} \qquad (11\text{-}26)$$

The common-gate amplifier of Fig 11-5c is similar to the grounded-grid vacuum tube amplifier. Power gains of 15 db can be obtained at 200 megacycles per second. The maximum theoretical power gain of this stage is given by

$$Pg_{max} = g_m r_d \qquad (11\text{-}27)$$

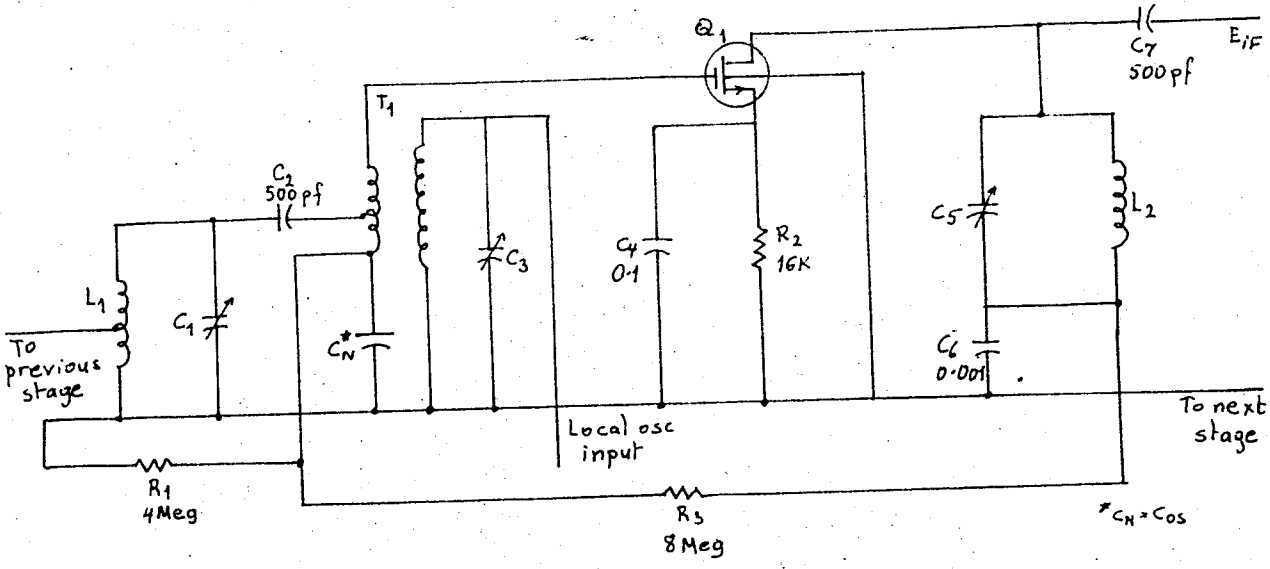


Fig 11-6. Mixer circuit containing MOS FET is superior to conventional transistor mixer circuit in controlling oscillator harmonics.

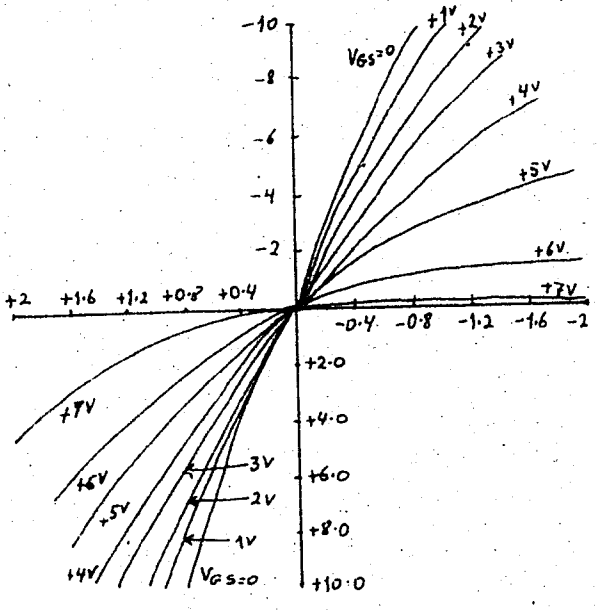


Fig 11-7 Low level output characteristic for the 2N3386, a p-channel junction FET. Drain current is plotted for various positive gate-to-source volts.

Fig 11-6 shows a mixer circuit containing MOSFET. Compared with the convential transistor mixer, with respect to intermodulation and interference from oscillator harmonics, the MOS transistor mixer is superior. Only the second harmonics and frequencies representing Sum and difference components of incoming signals are generated. No intermodulation by-products of the form $(1+n)w_1 - nw_2$ or $(1+n)w_2 - nw_1$, are present as in conventional mixers. The gate voltage excursion, however, should be limited to a range over which the transconductance rises linearly with gate voltage.

The Field-effect transistor operated at 200 Mc is superior to conventional transistors in cross modulation performance. The FET and the conventional high frequency transistor appear to be on par as far as noise figure is concerned but the conventional unit holds an edge in power gain.

Because it is a majority carrier device, the field effect transistor, unlike the conventional transistor can be operated at extremely low temperatures. For example, with 200 Mc operation at the temperature of liquid nitrogen, the FET's power gain increases approximately 4 db and its noise figure drops about 2 db. FET amplifiers that can operate up to 300 Mc have been built. However, selected transistors must be employed in constructing a 300 Mc amplifier.

11-5. Attenuator Circuits

In junction gate and insulated gate field effect transistors, if the drain to source voltage is restricted so that operation is in the non-pinched off region, near the origin, these devices behave like electrically controlled variable ohmic resistors, the gate voltage determining the actual magnitude of the resistance between source and drain. This remains valid even when the polarity of the voltage applied to the drain reverses as shown in Fig 11-7 provided that the gate voltage is large compared with the drain voltage. When the field effect transistor

is operated in this ohmic region, it can be used as a low distortion , voltage controlled attenuator.

The primary advantages of the MOSFET in voltage-controlled attenuator circuits are its low gate power requirements and wide dynamic range.

A simple L-pad arrangement in which the transistor serves as the variable resistance in the low side of the attenuator is shown in Fig 11-8. In this circuit, the maximum attenuation is normally between 60 and 70 db and the minimum signal reduction is 1 or 2 db. Proper performance of this circuit is possible only when the attenuator unit is followed by a high-impedance load such as a common-source MOSFET amplifier.

In another version of the L-pad attenuator, the MOSFET is placed in the series arm. In this case, the shunt arm must have low impedance if the maximum attenuation of 60 to 70 db is desired.

Fig 11-9 and 11-10 illustrate several possible circuit configurations using junction-gate and insulated gate FETs. In Fig 11-9a the voltage attenuation ratio is

$$\frac{Av_{max}}{Av_{min}} = m \left[1 - \left(Av_{max} \right) \right] \qquad (11\text{-}28)$$

where $A_v = \frac{v_{out}}{v_{in}} = \frac{R_{ds}}{R + R_{ds}}$

and $\frac{R_{ds(max)}}{R_{ds(min)}} = m$

In Fig 11-9b, the voltage gain is very nearly

$$A_v = \frac{R_L}{R_{ds}} \qquad (11\text{-}29)$$

Eq. 11-29 can be made equal to or greater than unity.

Fig 11-9c is a bridge circuit in which the minimum signal output can be zero when the bridge balances.

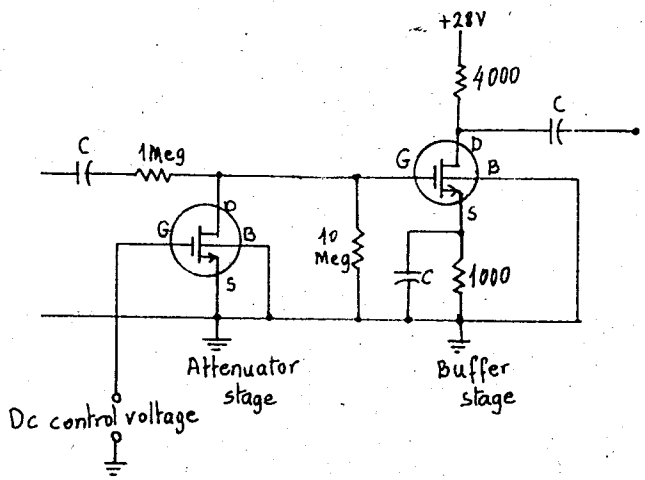
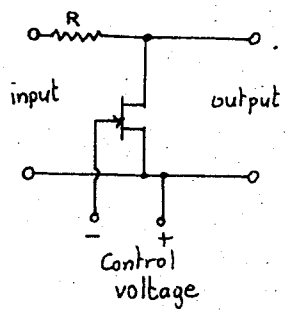
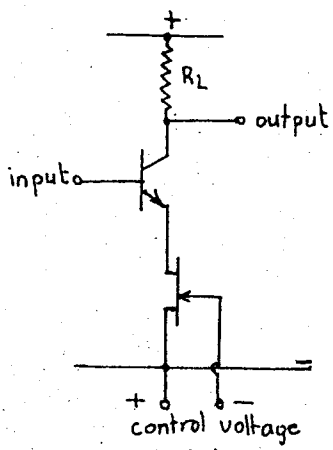


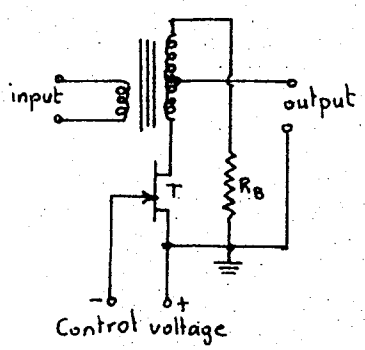
Fig. 11-8 Field effect transistor circuit can attenuate input signals from 60 to 70 db.



a) A voltage controlled attenuator



b) Gain control by variable negative feedback.



c) A bridge gain control circuit

Fig 11-9. Attenuators using FETs.

Fig 11-10 shows how the FET can be used in low pass, high pass and band pass filters. For the LP circuit the design conditions are :

$$R_{in} > 10 \times_{c2} \qquad (11\text{-}30)$$

$$R_{out} < 0.1 \times_{c1} \qquad (11\text{-}31)$$

$$R = R_1 + R_2$$

where R_{in} = input impedance of emitter follower Q_3, and therefore a function of the emitter resistor R_E .

R_{out} = output impedance of Q_3

R_1 and R_2 = source drain resistance of Q_1 and Q_2 respectively.

$w_n = \frac{1}{R\ M\ C_2}$ = resonant or corner frequency

$M^2 = C_1/C_2$

If the damping factor is ρ

$$\rho = \frac{1}{M} + \frac{M}{2}\left(\frac{1}{1 - K^2}\right) \qquad (11\text{-}32)$$

K_2 = voltage gain of Q_3

so that

$$w_n = \frac{1 - K_1 V_{gs}}{R_o\ MC_2} \qquad (11\text{-}33)$$

Thus the frequency becomes a linear function of the FET gate control voltage.

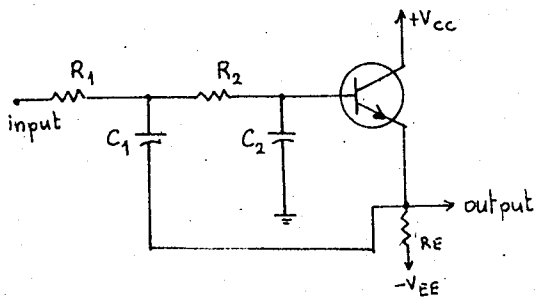
Fig 11-10b shows a basic HP RC filter and its FET subtituted equivalent respectively . In Fig 11-10b, if

$M^2 = C_1/C_2$ $\qquad \lambda^2 = R_2/R_1$ $\qquad K_2$ = gain of Q_3

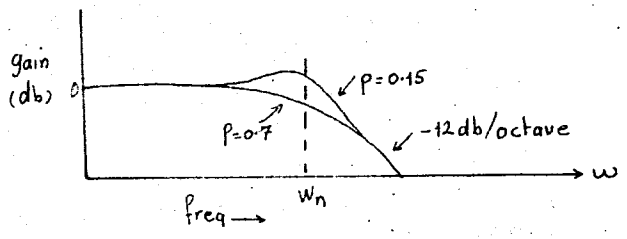
$$\rho = \frac{1 + M^2}{2 + M} + \frac{\lambda}{2M}(1 - K_2) \qquad (11\text{-}34)$$

and if C_1 is made equal to C_2 $\qquad w_n^2 = \frac{1}{R_1 R_2\ C_1 C_2}$

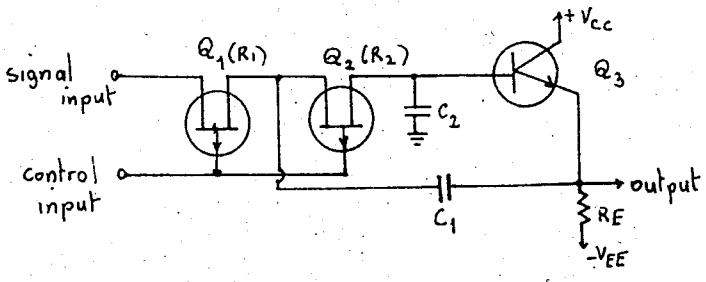
Fig 11-10c shows a band pass filter .



basic RC low pass filter

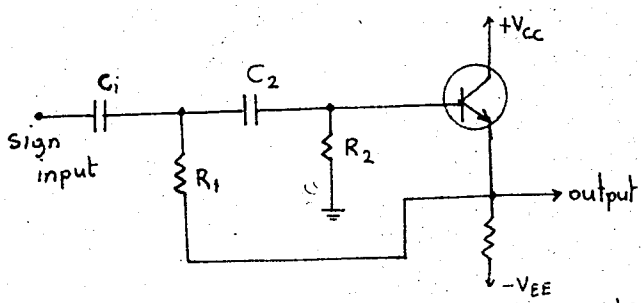


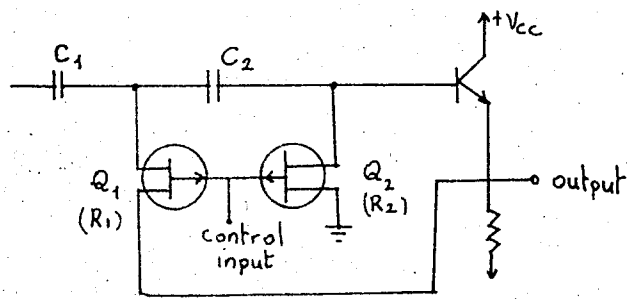
response curve for LP filter



a)

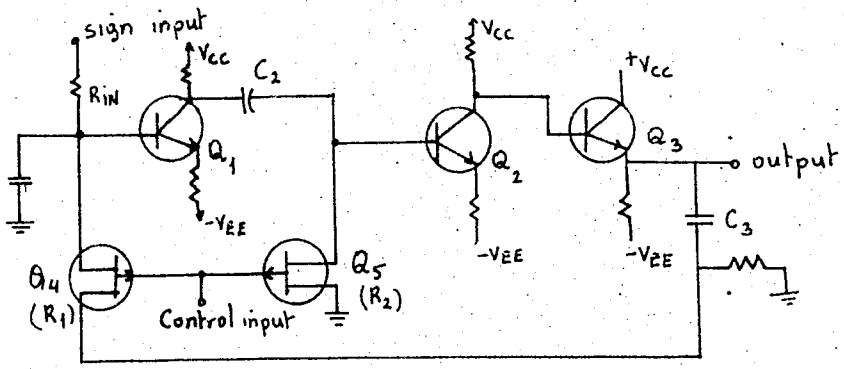
Matched FETs substitute for R_1 and R_2 in the basic LP filter giving active voltage tuned filter.





b)

Basic HP RC filter and its substitute which uses FETs in place of resistors R_1 and R_2.



c) Band pass filter

Fig 11-10 LP, HP, and BP filters using FETs as active resistors.

CHAPTER I2

FIELD EFFECT TRANSISTORS IN NON-LINEAR AND INTERGRATED CIRCUITS

12-1. Analog Switching Circuits

Field effect transistors lack some of the advantages of bipolar transistors in this area when passive load elements are used. The field effect transistors have some inherent limitations such as high "on" or saturation resistance. The saturation voltage i.e., the voltage drop across a transistor biased into conduction, typically is an order of magnitude smaller for bipolar transistor circuits. Thus the ratio of 'off' to 'on' output voltage may be much greater for the bipolar transistor despite a smaller supply voltage. Also the switching speed of the unipolar transistor circuits is slower than that of the bipolar transistor circuits. However, the field effect transistor, especially the insulated gate transistor operating in enhancement mode, contribute some advantages to logic circuits:

a) The FET eliminates the transformer drive that is usually required for analog gate circuits that use conventional chopper transistors.

b) Since FET circuitry does not have input-output off set voltage, transmission at the microvolt-signal level is possible .

c) Because FETs are compatible with integrated-circuit techniques, switches can be made much smaller.

d) Small shifts in the FET's characteristics are not critical, and therefore the switch has long term stability.

e) The high impedance level of the unipolar transistor allows large fan-outs in logic circuits. One stage can drive an indefinite number of stages in parallel because the extremely high gate

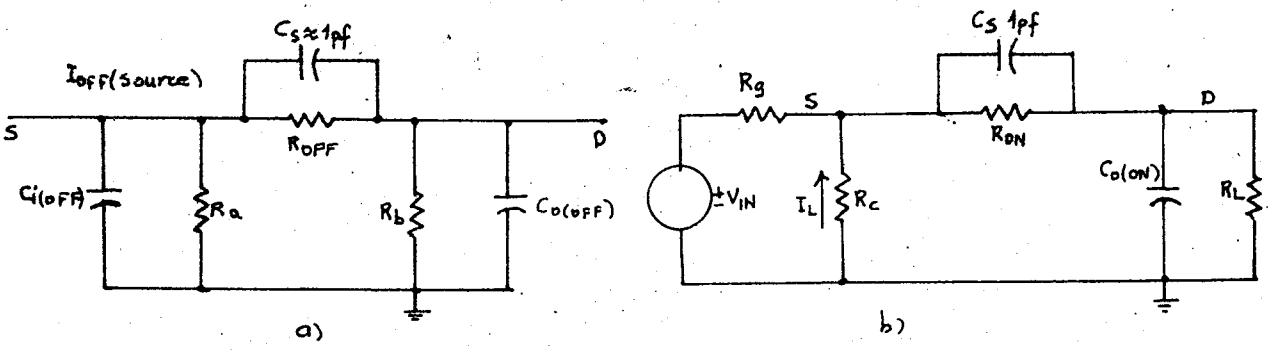


Fig 12-1 ON and OFF equivalent circuit.
a) OFF condition
b) ON condition.

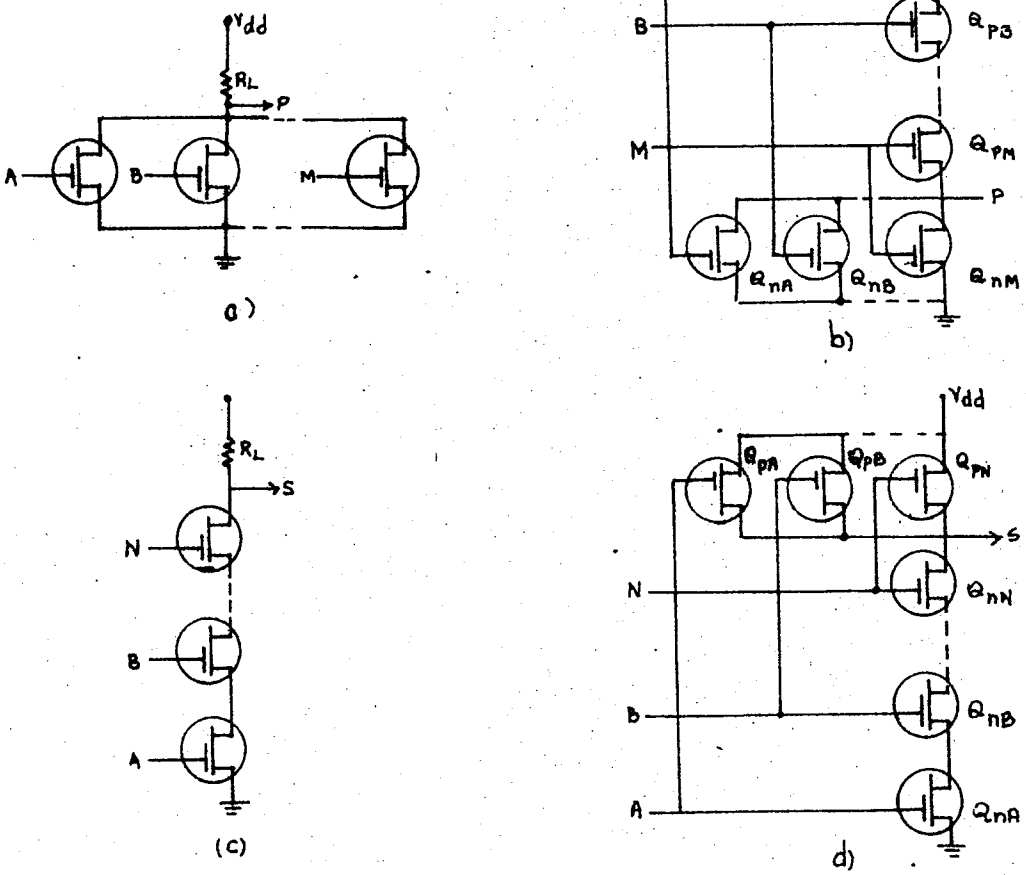


Fig 12-3. Basic logic circuits
a) parallel gate
b) composite complementary circuit analogous to a)
c) series gate
d) composite complementary circuit analogous to c)

input resistance of the driven stages presents essentially no static load to the driving stage. But here a compromise must be made since although dc conditions impose essentially no limitations on fan-out in FET switches, transient requirements restrict fan-out.

f) Complementary field effect transistors have many advantages such as high circuit stability, low power dissipation and symmetry. Because of symmetry almost equal "turn on" and "turn off" switching times are obtained.

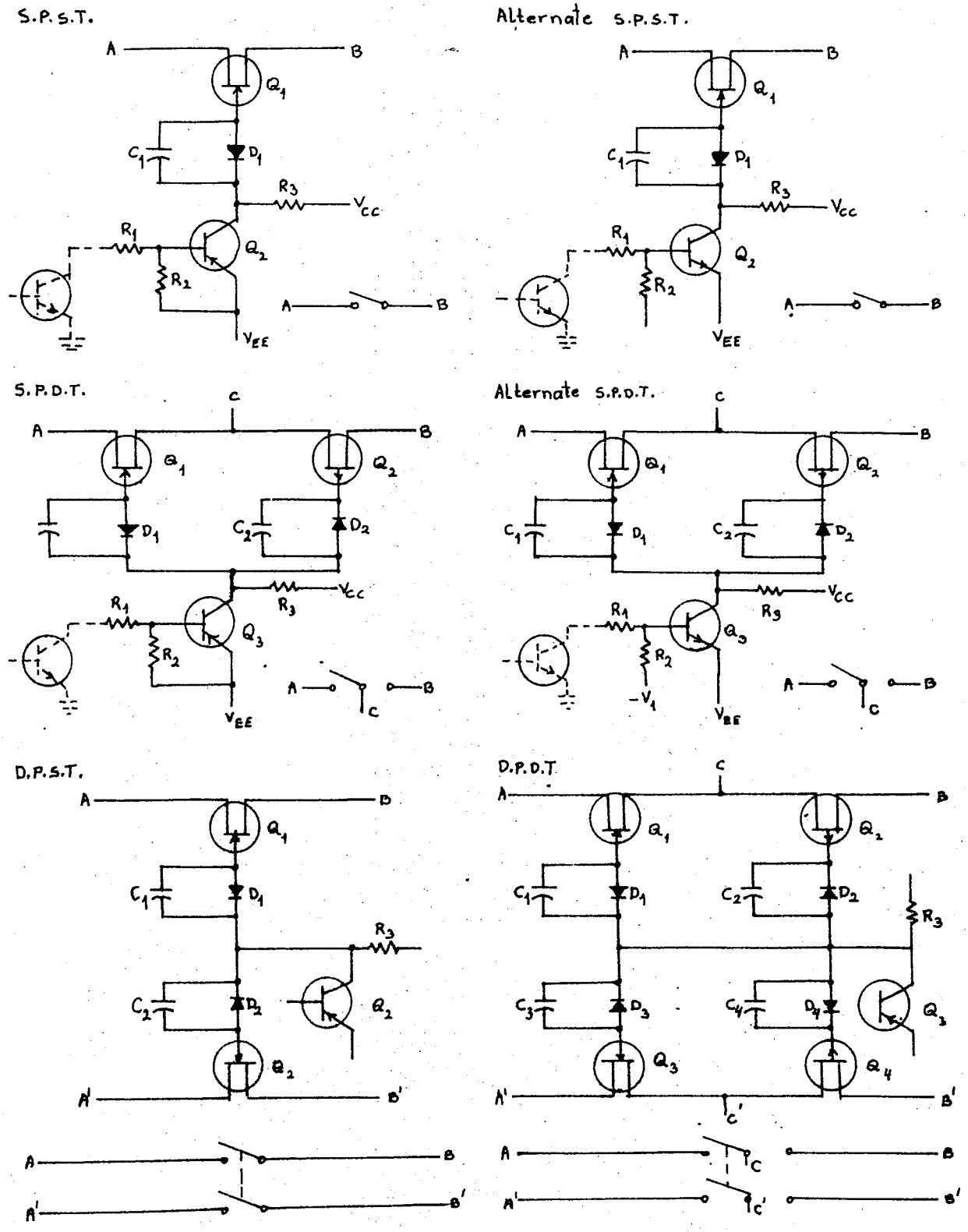
12-2. Equivalent Circuits for Analog Switches.

The off and on conditions may be represented by a set of equivalent circuits as shown on Fig 12-1. The circuit of Fig.12-1a illustrates the off condition. Resistors R_a and R_b represent the source-to-gate and drain-to-gate resistances through which FET leakage current flows because of the reverse biasing of the gate. Capacitor C_s is the stray capacitance. Resistor R_{OFF} is the source-to-drain resistance. The source to gate and drain-to-gate capacitances are $C_{i(OFF)}$ and $C_{o(OFF)}$.

During conduction, the circuit of Fig 12-1b applies. Here RC represents the path for the flow of leakage current I_L from the reverse-biased isolation diode. Its value is a few hundred megohms. Resistor R_{on} is the source to drain resistance during conduction, and R_g is the signal source resistance.

In the design of analog switching circuitry, a choise must be made between the use of an FET with a relatively small value of $C_{i(OFF)}$ and a high value of R_{on} on the one hand, and an FET with a larger $C_{i(OFF)}$ and a smaller R_{on} on the other. The first type of FET provides a switching waveform with a faster rise time than does the second. The second is better for handling very accurate voltages.

FET analog switches may be designed in a wide variety of circuit configurations. Six configurations that can be applied in the design of complex switching functions are shown in Fig. 12-2. Fig 12-3 shows some basic logic circuits and their analogou



S.P.S.T. Single-pole, single throw
S.P.D.T. = single-pole, double throw

D.P.S.T = Double-pole, single-throw
D.P.D.T = Double-throw, double pole

Fig. 12.2. Modular analog switches allow compact design of complex switching functions.

composite complementary circuits respectively.

12-3. Integrated Circuits

Digital logic circuits in satellites and other aerospace equipment must be small and must operate on limited power.

Due to their structural simplicity and the fewer processing steps required and also due to the possibility of direct coupling of digital stages, the insulated gate FETs can be used in integrated circuits. Another advantage of the unipolar transistor over the bipolar transistor in monolithic integrated circuits is that it takes considerably less bias power. Because of the resulting low heat dissipation, high packaging density is possible.

The ladder-like array of MOS transistors was developed as part of the integrated circuit concept. The transistors in this type of array are completely symmetrical. Each diffused region can serve as either source or drain. Sections of such a transistor ladder are isolated by grounding a source drain or gate electrode. Fig 12-4b shows how the circuit of Fig 12-4a can be realized.

12-4. The thin Film Transistor

Weimer has proposed another surface FET which is known as thin-film FET .(TFT). MOS transistors are also surface FETs and the electrical characteristics of these types are quite similar but fabrication techniques are entirely different.

In the TFT the channel is formed of a thin layer of polycrystalline cadmium sulphide. This layer is deposited over metal source and drain electrodes by vacuum evaporation, typically to a thickness of 1μ. The source and drain electrodes may be evaporated gold and are separated by some $5\text{-}50\mu$. The insulating film over the semiconductor is silicon monoxide, and over this a gate is deposited (also a gold film). Thus the transistor consists of four superimposed layers. The operation of this transistor is in the enhancement mode but differs from that of the silicon MOS

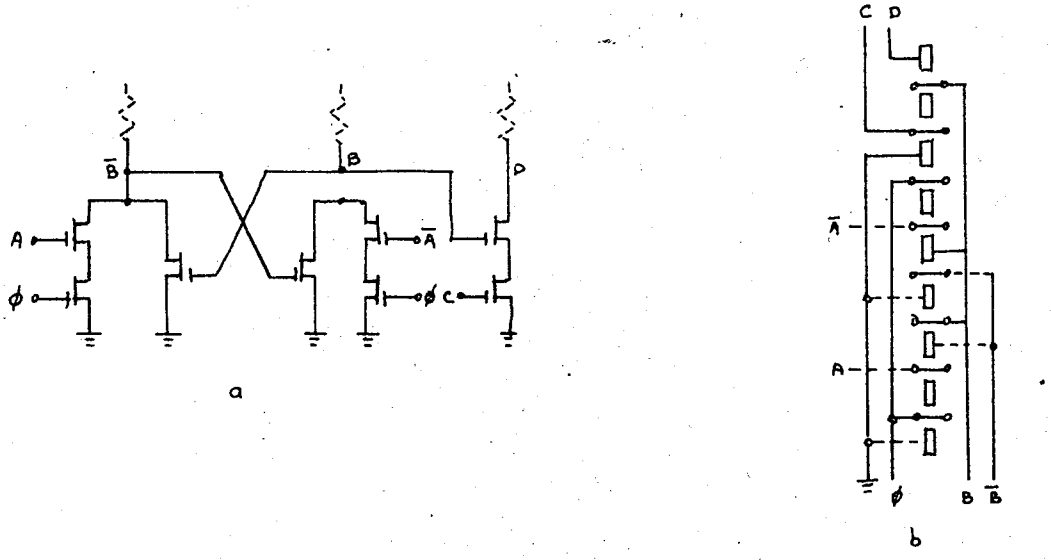


Fig 12-4. a) Typical digital circuit
b) possible inter connection pattern using ladder-type array.

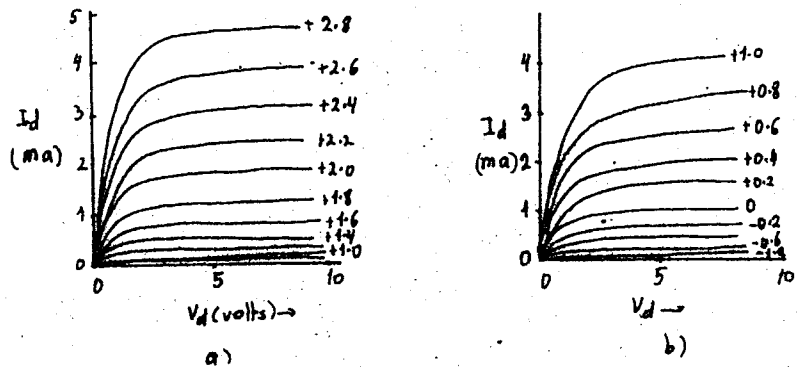


Fig 12-5. Output characteristics of TFT.
a) enhancement type CdS TFT
b) depletion type CdS TFT.

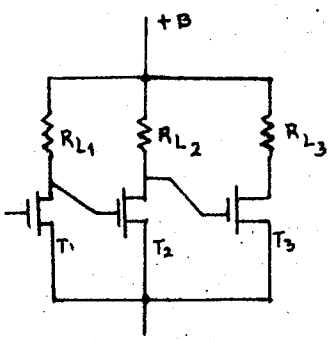


Fig 12-6. Three stage amplifier

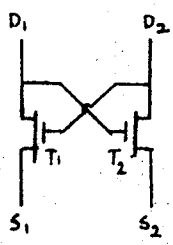


Fig 12-7. TFT flip-flop

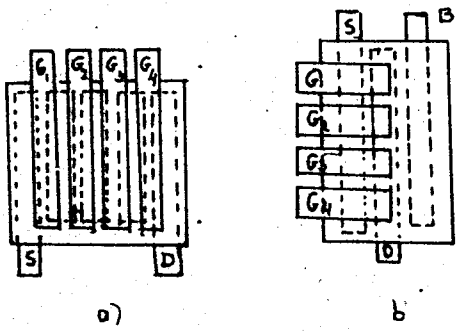


Fig 12.8. a) 'and' gate (TFT)
b) 'Nor' gate (TFT)

device because of the relatively large energy gap of the materials used (such as CdS). Relatively few thermally generated carriers exist and the device is therefore dependent on majority carriers injected from the source electrode. Exactly how the current flows through the polycrystalline semiconductor and is modulated by the gate is not yet clear. The picture is complicated by the fact that the semiconducting layer consists of many small crystallites thus introducing the complications of grain boundaries and surface defects. Many charge carriers injected by the source electrode are held by the surface traps and other immobile sites, and those which are not held contribute to the density of mobile carriers. The initial slow rise of drain current with gate voltage followed by a rapid rise at a large positive bias seen in Fig 12-5 supports the theory that surface traps and states are being filled. But CdS units do not require a large positive gate bias for satisfactory operation.

A TFT circuit can in principle be deposited upon any part of an existing device offering a few sq. milimeters of free space. If the surface is a metal, a preliminary coating of insulator can be applied as substrate for the TFT.

With the developed techniques complete circuits containing hundreds or thousands of active elements deposited in one evaporation sequence is feasible.

Fig 12-6 shows a three-stage direct coupled thin film amplifier. Direct coupling is possible since the insulated gate TFT does not draw appreciable current.

Fig 12-7 shows a direct coupled flip flop based upon the TFT. By inverting one of the triodes in the flip flop the cross connections are simplified.

Fig 12-8 shows 'AND' and 'NOR' gates derived from TFT. The semiconductor itself is used as the load resistor for the NOR gate.

To adequately design, build, and use such integrated devices a new type of engineering skill will have to be developed. Questions such as stability, need for encapsulation and ability to meet design tolerances have yet to be clarified for the TFT before its place in circuit design can fully be assessed.

APPENDIX A

THE D-C CHARACTERISTICS OF FETs WITH DIFFERENT IMPURITY DENSITY PROFILES

The following general equations will be used in deriving the d-c. theory of fets with different profiles. In the following discussions devices with n-type channel and p-type gates will be considered.

A simplifying assumption to be used is that the depletion layer boundary for P-N junction is abrupt at y_p and y_n, where y_p and y_n are the location of the depletion-layer edges in the p-type and n-type regions, respectively.

The potential V of the p-type gate with respect to the n-type channel is obtained from Poisson's equation and is given by ,

$$V(x) = -\frac{q}{\epsilon}\int_{y_p}^{y_n} y\, N(y)\, dy \qquad (A\text{-}1)$$

where

q = magnitude of the electronic charge

ϵ = 8.85×10^{-12} x relative permittivity

$N(y) = N_d - N_a$

y_p, y_n = location of the depletion-layer edges

we know that the net depletion-layer charge is zero, so that

$$\int_{y_p}^{y_n} N(y)\, dy = 0 \qquad (A\text{-}2)$$

Since the gate-channel P-N junctions in an FET are reverse-biased, the gate currents are very small. Therefore we can consider the gates as equipotentials. Making use of Shockley's gradual channel approximation, and also assuming that any variation in the channel width are due to voltage drops caused by current flow in the channel , we can write the drain current I_d expression as :

$$I_d = -Z\frac{dV}{dx}\int_{y_{n_1}}^{y_{n_2}} q\mu_n N(y)\, dy \qquad (A\text{-}3)$$

where Z = channel width

μ_n = electron mobility

Multiplying (A-3) by dx and integrating from source to drain, we get

$$I_d = -\frac{Z}{L} q\mu_n \int_{y_{n_2}}^{y_{n_1}} N(y)\,dy \quad dV$$

$$I_d = -\int_{V_s}^{V_d} g(V)\,dV \qquad (A-4)$$

because from Eq.(1-2) we know that

$$g(V) = \frac{Z}{L} \int_{y_{n_2}}^{y_{n_1}} q\mu_n N(y)\,dy$$

It will be useful to repeat here, for convenience, that

$$V_s = \emptyset + V_{gs}$$

$$V_d = \emptyset + (V_{gs} - V_{ds})$$

V_{gs} = The potential of the gate with respect to the source

V_{ds} = The potential of the drain with respect to the source

$\emptyset$ = Contact potential of the P-N junction($V_{gs} = 0$) This is a negative quantity for an n-type channel and positive quantity for a p-type channel

Note that in the simplified Shockley derivation $\emptyset$ was neglected. If W_o is the gate-drain potential required to produce channel pinch-off when the gradual solution is used, Eq.(A-4) will be valid only for $|V_d| < |W_o|$. For $|V_d| \geq |W_o|$, we can modify Eq.(A-4) as follows

$$I_d = -\int_{V_s}^{W_o} g(V)\,dV \qquad (A-5)$$

only here, in the expression for g(V) the channel length L must be replaced by the effective lengtn Le equal to the value of x at which pinch-off occurs.

The mutual conductance g_m expression can be obtained as follows:

$$g_m = \frac{\partial I_d}{\partial V_{gs}} = \frac{\partial I_d}{\partial V_s} \cdot \frac{\partial V_s}{\partial V_{gs}} + \frac{\partial I_d}{\partial V_d} \cdot \frac{\partial V_d}{\partial V_{gs}} \qquad (A-6)$$

For $|V_d| < |W_o|$, this becomes

$$g_m = g(V_s) - g(V_d) \longrightarrow g_{max} = g(V_s) \qquad (A\text{-}7)$$

For $|V_d| \geq |W_o|$

$$g_m = \frac{Z}{Le} \int_0^{2a} q\mu_n N(y)\, dy \qquad (A\text{-}8)$$

A-1 Linear-Junction Symmetrical Device

Consider Fig 1-3b. The device is symmetrical about $y=a$
From Eq (A-1) and (A-2),

$$y_{n_2}^3 = a^3 \frac{V(x)}{W_o} \qquad (A\text{-}9)$$

where $W_o = -\frac{2qa^2 N_n}{3\epsilon}$

W_o expression is obtained by putting $y_{n2}=a$ in Eq. (A-1). Integrating Eq. (A-4), we get

$$\frac{I_d}{I_{do}} = \frac{5}{2}\left(\frac{V_d}{W_o} - \frac{V_s}{W_o}\right) - \frac{3}{2}\left[\left(\frac{V_d}{W_o}\right)^{5/3} - \left(\frac{V_s}{W_o}\right)^{5/3}\right] \qquad (A\text{-}10)$$

where $I_{do} = -\frac{2}{5}\frac{aZq\mu_n N_n}{L} W_o \qquad (A\text{-}11)$

For $|V_d| \geq |W_o|$ (by replacing $V_d = W_o$ in Eq. (A-10))

$$I_d / I_{do} = 1 - \frac{5}{2}\frac{V_s}{W_o} + \frac{3}{2}\left(\frac{V_s}{W_o}\right)^{5/3} \qquad (A\text{-}12)$$

From Eq. (A-7) and (A-8)

$$g_o = \frac{aZq\mu_n N_n}{L} \qquad (A\text{-}13)$$

$$g_m/G_o = 1 - (V_s/W_o)^{2/3} \qquad (A\text{-}14)$$

A-2 - Abrupt Linear asymmetrical Device

Consider Fig 1-3a.

From Eq.(A-1) and (A-2), the gate to channel potential can be obtained in terms of the space charge widths as follows:

$$V(x) = -\frac{qN_n y_{n2}^3}{3a\epsilon} \qquad (A-15)$$

$$V(x) \cong -\frac{qN_n}{2a\epsilon}\left(\frac{4a^3}{3} + \frac{y_{n1}^3}{3} - a y_{n1}^2\right) \qquad (A-16)$$

At pinch-off $V(x) = W_o$, $y_{n1} = y_{n2}$ so that from (A-15) and (A-16)

$$y_{n_2} = y_{n1} = a \qquad (A-17)$$

$$W_o = -\frac{qN_n a^2}{3\epsilon} \qquad (A-18)$$

Now Eq.(A-15) and (A-16) can be written in terms of W_o.

$$\frac{V(x)}{W_o} = 2 + \frac{1}{2}\left(\frac{y_{n1}}{a}\right)^2 - \frac{3}{2}\left(\frac{y_{n1}}{a}\right)^2 \qquad (A-19)$$

$$g_m = g(V_s) = \frac{G_o}{4}\left[\left(\frac{y_{n1s}}{a}\right)^2 - \left(\frac{V_s}{W_o}\right)^{2/3}\right] \qquad (A-20)$$

where $$G_o = \frac{q\mu_n a Z N_n}{L} \qquad (A-21)$$

y_{n1s} = depletion boundary of gate 1 at the source.

The drain current I_d can be obtained from Eq. (A-4) by changing the variable V to y_n with the aid of the equality $V(x)/W_o = (y_{n2}/a)^3$ and Eq. (A-19). For $|V_d| \geq |W_o|$,

$$I_d = G_o W_o \frac{3}{80}\left[7 - 4\left(\frac{V_s}{W_o}\right)^{5/3} + 2\left(\frac{y_{n1s}}{a}\right)^5 - 5\left(\frac{y_{n1s}}{a}\right)^4\right] \qquad (A-22)$$

Since $I_{do} = -\frac{27}{80} G_o W_o$, the drain current in the pinch-off region may be written as:

$$\frac{I_d}{I_{do}} = \frac{1}{9}\left[-7 + 4\left(\frac{V_s}{W_o}\right)^{5/3} - 2\left(\frac{y_{n1s}}{a}\right)^5 + 5\left(\frac{y_{n1s}}{a}\right)^4\right] \qquad (A-23)$$

y_{n1s} can be obtained from Eq. (A-19), and

$$V_s/W_o = 2 + \frac{1}{2}\left(\frac{y_{n1s}}{a}\right)^3 - \frac{3}{2}\left(\frac{y_{n1s}}{a}\right)^2 \qquad (A-24)$$

APPENDIX B

DETERMINATION OF SOURCE, CHANNEL AND DRAIN RESISTANCES OF AN FET

If a positive voltage is applied to the drain of an n-type FET, while the gate is open circuited, then the gate will float at a potential approximately equal to that at the source end of the channel. The reason for this can be explained as follows: Since the gate is open circuited, the net gate current must be zero. If the gate were to become much more positive than the source end of the channel, a considerable fraction of the length of the gate would be biased in the forward direction and a large current would flow into the gate. If on the other hand, the gate were more negative than the source end of the channel, all of the gate junction would be in the reverse direction and saturation current would flow out of the gate. Therefore the gate must have a potential very close to that of the source end of the channel (in fact the potential is slightly more positive). Thus the open circuit gate potential is closely equal to the potential drop in the source resistance R_s and if the current I_s is measured, the value of R_s can be determined. An electronic voltmeter with a high impedance can be used to measure the open-circuit voltage. To determine the value of R_d, the source and drain connections are interchanged, and the experiment repeated. Then R_o can be determined from the measured value of the sum of R_d, R_o, and R_s. This sum can be determined from the ratio of V_d to I_d for values of V_d sufficiently small compared to W_o of the FETs I_d vs V_d characteristic curve.

APPENDIX C

FIELD EFFECT TRANSISTOR CAPACITANCES

The space charge regions between x and x+dx contain a charge

$$dQ = 2\rho_o (a-b)\, dx \qquad (C\text{-}1)$$
$$= 2\rho_o a \left(\frac{W}{W_o}\right)^{1/2} dx$$

The total charge Q per unit width is then

$$Q = \int_o^L 2\rho_o a \left(\frac{W}{W_o}\right)^{1/2} dx \qquad (C\text{-}2)$$

$$= 2\rho_o a \int_{W_s}^{W_d} \frac{g(W)(W/W_o)^{1/2}}{g(W)(dW/dx)}\, dW \qquad (C\text{-}3)$$

$$= \underbrace{2\rho_o a g_o W_o}_{I} \int_z^y (1-u^2)\, u^{1/2}\, du \qquad (C\text{-}4)$$

where $y = W_d/W_o$, $z = W_s/W_o$, $u = W/W_o$

$$g(W) = g_o\left[1 - \left(\frac{W}{W_o}\right)^{1/2}\right]$$

$$I = \frac{g_o}{L}\left[W_d - W_s - \frac{2}{3}\,\frac{W_d^{3/2} - W_s^{3/2}}{W_o^{1/2}}\right] \qquad (C\text{-}5)$$

$$= \frac{g_o W_o}{L} f_1(y,z)$$

so that we can write Q, as

$$Q = \frac{2\rho_o a L}{f_1(x,z)}\left[\frac{2}{3}\left(y^{3/2} - z^{3/2}\right) - \frac{1}{2}\left(y^2 - z^2\right)\right] \qquad (C\text{-}6)$$

The small signal capacitances per unit width are defined as

$$C_{gg} = \frac{\partial Q}{\partial V_g} = \frac{\partial Q}{\partial y}\cdot\frac{\partial y}{\partial V_g} + \frac{\partial Q}{\partial z}\cdot\frac{\partial z}{\partial V_g} \qquad (C\text{-}7)$$

$$C_{gg} = \frac{1}{W_o}\left(\frac{\partial Q}{\partial y} + \frac{\partial Q}{\partial z}\right)$$

$$C_{gd} = -\frac{\partial Q}{\partial V_d} = -\frac{\partial Q}{\partial y}\frac{\partial y}{\partial V_d} = \frac{1}{W_o}\frac{\partial Q}{\partial y} \qquad (C\text{-}8)$$

and $$C_{gs} = C_{gg} - C_{gd} = \frac{1}{W_o}\frac{\partial Q}{\partial z} \qquad (C\text{-}9)$$

Performing the differentiations, we get

$$C_{gs} = \frac{2p_o aL}{W_o} f_2(y,z) \qquad (C\text{-}10)$$

$$C_{gd} = \frac{2p_o aL}{W_o} f_3(y,z) \qquad (C\text{-}11)$$

where

$$f_2(y,z) = \frac{\left[\frac{2}{3}(y^{3/2} - z^{3/2}) - \frac{1}{2}(y^2 - z^2)\right](1 - z^{1/2}) - f_1(y,z)(z^{1/2} - z)}{[f_1(y,z)]^2} \qquad (C\text{-}12)$$

$$f_3(y,z) = \frac{\left[-\frac{2}{3}(y^{3/2} - z^{3/2}) + \frac{1}{2}(y^2 - z^2)\right](1 - y^{1/2}) + f_1(y,z)(y^{1/2} - y)}{[f_1(y,z)]^2}$$

Note that for saturation $y=1$, so that

$$f_2(1,z) = \frac{3}{2}\frac{(1+z^{1/2})}{(1+2z^{1/2})} \qquad (C\text{-}13)$$

$$f_3(1,z) = 0$$

and when z becomes equal to zero while $y=1$

$$f_2(1,0) = \frac{3}{2} \qquad (C\text{-}14)$$

APPENDIX D

MEASUREMENT TECHNİQUES FOR ADMİTTANCE PARAMETERS FOR FETs

Fig.D-1 shows the bridge configurations for measuring the admittance parameters. Measurements in the frequency range 15 kc/s - 5 Mc/s may be made using a radio frequency bridge.. Measurements for 30 Mc/s and on can be made with transfer function and immittance bridge.

The arrangements for the measurement of the input admittance y_{11} and the output admittance y_{22} are shown in Figs D-1a and D-1b respectively, and the connections for the reverse transfer admittance y_{12} measurement are shown in Fig D-1c . In all these measurements, the value of Y_s read at null , gives us directly the value of the admittances in question. For example , in the case of Fig D-1c when a null condition is obtained at the dedector terminals, The input is effectively short-circuited to ac since the voltage between source and gate is then zero. Under these conditions, the admittance indicated by the bridge is given by

$$y_s = \frac{i_1}{v} = \frac{i_2}{v} = - y_{12}$$

The measurement of the forward-transfer admittance is more difficult. Since

$$y_{21} = -y_2 + g_m$$

and the real part of g_m is greater than the real part of y_2 in the useful frequency range for most of the devices a direct measurement of y_{21} is not possible. Therefore the set up of Fig . D-1d is used : At balance

$$y_s = \frac{i_1}{v} = \frac{i_2}{v} = \frac{-(i_1' + i_2')}{v} = - \frac{i_1'}{v} - \frac{i_2'}{v} = y_{11} + y_{21}$$

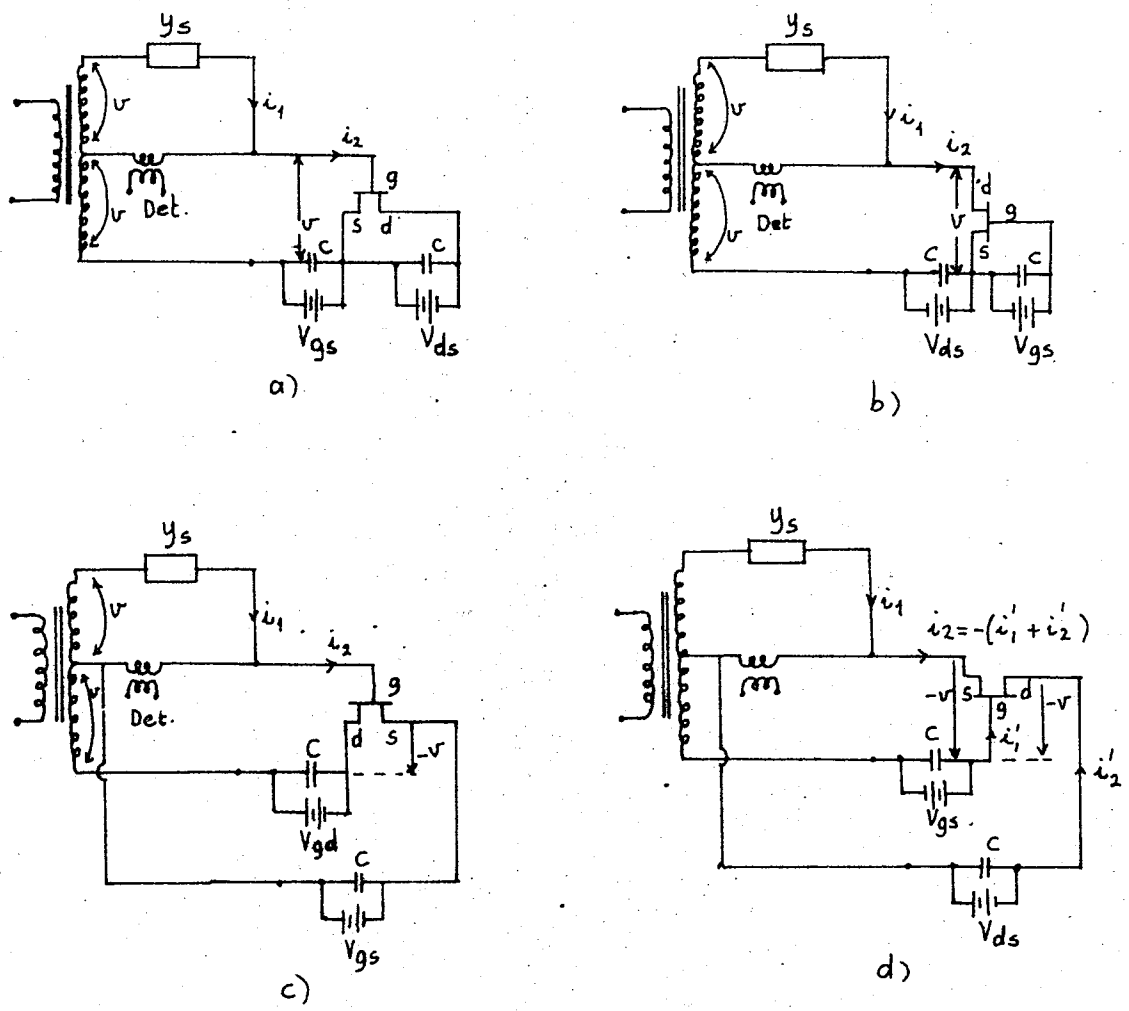


Fig.D-1. Bridge configurations for measuring the admittance parameters

a) Arrangement for measuring y_{11}

b) Arrangement for measuring y_{22}

c) Arrangement for measuring y_{12}

d) Arrangement for measuring $y_{11} + y_{21}$

R E F E R E N C E S

1. Shockley , W. " A Unipolar Field-Effect Transistor," Proc.IRE, Vol. 40 November, 1952, p. 1365.

2. Dacey, G. C. and I. M. Ross, "Unipolar 'Field-Effect' Transistor," Proc. IRE, Vol. 41 August, 1953, p. 970.

3. Dacey, G.C. and I.M.Ross, "The Field-Effect Transistor," Bell Syst. Tech.J., Vol. 34 November, 1955, p.1149.

4. Terman, L.M., "An Investigation of Surface States at a Silicon/Silicon-Oxide Interface Employing MOS Diodes," Tech. Rep. 1655-1, Stanford University, Stanford Electronics Laboratories, Solid-State Electronics Laboratory February 23, 1961.

5. Weimer, P.K.,"The TFT-A New Thin-Film Transistor," Proc. IRE, Vol.50 June,1962, p.1462.

6. van der Ziel, A., "Thermal Noise in Field-Effect Transistors," Proc. IRE, Vol.50 August, 1962, p. 1808.

7. van der Ziel,A., "Shot Noise in Thin Film Transistors" Proc. IRE, Vol.50 September, 1962, p. 1985.

8. Olsen,D.R.,"Equivalent Circuit for a Field-Effect Transistor" Proc. IEEE, Vol.51 January, 1963, p. 254.

9. Bruecke, W.C.,"Noise Measurements in Field-Effect Transistors" Proc. IEEE, Vol. 51 February, 1963, p. 378.

10. Lindholm F.A. and D.C. Latham,"Junction Capacitance of Field-Effect Transistors" Proc.IEEE, Vol. 51 February,1963 p.404

11. van der Ziel,A.,"Gate Noise in Field-Effect Transistors at Moderatly High Frequencies," Proc.IEEE, Vol.51 March, 1963, p.461.

12. Borkan,H.and P.K.Weimer,"An Analysis of the Characteristics of Insulated-Gate Thin-Film Transistors," RCA Rev., Vol. 24 June, 1963, p.153.

13. Hofstein, S.R. and F.P. Heiman,"The Silicon Insulated-Gate Field-Effect Transistor," Proc. IEEE, Vol.51 September,1963, p. 1190.

14. Richer,I. and R.D. Middlebrook,"Power-law Nature of Field - Effect Transistor Experimental Characteristics" Proc. IEEE, Vol.51 August,1963, p.1145. Also Proc. IEEE,Vol.52 March, 1964 p.314.

15. Middelbrook,R.D.,"A Simple Derivation of Field-Effect Transistor Characteristics," Proc.IEEE,Vol.51 August,1963,p.1146.

16. Richer,I.,"Input Capacitance of Field-Effect Transistors ," Proc. IEEE, Vol.51 September,1963 , p. 1249.

17. Halladay,H.E. and W.C. Bruncke,"Excess Noise in Field-Effect Transistors," Proc. IEEE, Vol.51 November,1963, p.1671 .

18. van der Ziel,A.,"Carrier Density Fluctuation Noise in Field-Effect Transistors," Proc.IEEE, Vol.51 November,1963,p.1670

19. Bosenberg.W.,"Transfer Characteristics of Field-Effect Transistors," RCA Rev.,Vol.24 December,1963, p. 688.

20. LeMée, J.,A.V.J. Martin, I.Richer, and R.D. Middelbrook , "Power-Law Nature of Field-Effect Transistor Characteristics" Proc. IEEE, Vol.52 March,1964, p.314.

21. Goldberg,C., S. R. Hofstein, and F. P. Heiman, "Pinch-Off in Insulated-Gate Field-Effect Transistors", Proc. IEEE, Vol.52 April, 1964, p.414

22. van der Ziel,A. and J.W.Ero, "Small-Signal, High-Frequency Theory of Field-Effect Transistor Characteristics" IEEE . Trans, on Electron Devicies,Vol.ED-11 april,1964, p.128.

23. Sah,C.T.,"Theory of Low-Frequency Generation Noise in Junction-Gate Field-Effect Transistors," Proc.IEEE, Vol. 52 July, 1964, p. 795

24. Chopra, A., and L. Vadasz, "Silicon MOS Field-Effect Transistors Using Impurity Redistribution During Oxidatiod ," Proc.IEEE, Vol.52 August, 1964, p.985.

25. Gosling, W., Field-Effect Transistor Applications, London, Temple Press Ltd., 1964.

26. Hayashi, T., and T. Niimi, "Observations of Negative-Resistance Phenomena and Oscillations in the MOS Diode," Proc . IEEE, Vol. 52 August, 1964, p. 986.

27. Cobbold, R.S.C., and F.N. Trofimenkoff, "Field-Effect Transistor with Forward Biased Gate-Channel Junction" Proc. IEEE, Vol.52 September, 1964, p.1073.

28. Zuleeg, R. and V.O. Hinkle,"A Multichannel Field-Effect Transistor", Proc. IEEE, Vol.52 October, 1964, p. 1245.

29. Cobbold, R.S.C. and E.N. Trofimenkoff,"Breakdown Phenomena in Double-Gate Field-Effect Transistors", Proc. IEEE, Vol.5 November, 1964, p.1375.

30. "The Field-Effect Transistor," Part I. II and III, 15 papers in Electronics, Vol.37 November 30, December 14 and December 28, 1964.

31. Josephs, H.C.,"Conditions for Optimizing FET Operation," Proc IEEE, Vol. 53 February, 1965, p.199.

32. White, M.H. and R.C. Gallagher, "Metal-Oxide-Semiconductor (MOS) Small Signal Equivalent Circuit," Proc. IEEE, Vol. 53 March, 1965, p.314.

33. Special issue on Metal-Oxide-Semiconductor (MOS) Devices , IEEE Trans. on Electron Devices, Vol.ED-12 March, 1965.

34. Heiman, F.P. and G. Warfield, "The Effects of Oxide Traps on the MOS Capacitance," IEEE Trans. on Electron Devices, Vol. ED-12 April, 1965, p.167.

35. Leonce J. Sevin, Jr., "Field-Effect Transistors", Mc Graw Hill Book Comp., NewYork ,1965 .

36. Wallmark, J.T., Johnson Harwick,"Field-Effect Transistors, Physics, Technology and Applications," Prentice Hall, Inc., Englewood cliffs, NewJersey, 1966.

37. Tarnay, K., " Temperature Dependence of the Current of the Field-Effect Transistors in the Pinch-off Region", Proc . IEEE, Vol. 53 May,1965, p.485.

38. Zaininger, K.H. and G. Warfield,"Limitations of the MOS Capacitance Method for the Retermination of Semiconductor Surface Properties," IEEE Trans. on Electron Devices, Vol.ED-12 April, 1965, p. 179.