K. E. Drangeid
R. Sommerhalder

# Dynamic Performance of Schottky-barrier Field-effect Transistors*

**Abstract:** The dynamic performance of Schottky-barrier field-effect transistors is discussed, with the aim of finding in a most simple way the physical parameters on which the dynamic properties of a FET depend, how strong they influence the dynamic qualities of FET's, and what recommendations can be given as to proper choice of material or structure for FET's with good high-frequency performance.

## Symbols

| | |
|---|---|
| $b$ | gate width |
| $C_0$ | channel-to-gate capacitance per unit length |
| $C_{sg}$ | source-gate capacitance |
| $C_{sd}$ | source-drain capacitance |
| $C_{dg}$ | drain-gate capacitance |
| $d$ | transistor thickness |
| $E$ | electric field |
| $E_c$ | critical electric field |
| $f_0$ | frequency limit |
| $i$ | channel current |
| $L$ | gate length |
| $l$ | channel thickness |
| $L_{eff}$ | effective length of saturated channel section |
| $N$ | density of donor centers |
| $Q$ | charge of free carrier per unit |
| $R_d$ | drain resistance |
| $R_0$ | resistance per unit length |
| $R_s$ | source resistance |
| $v$ | drift velocity |
| $v_s$ | saturation drift velocity |
| $V_{sd}$ | source-drain bias voltage |
| $V_{gc}$ | gate-channel voltage |
| $V_{sg}$ | source-gate bias voltage |
| $V_D$ | diffusion voltage |
| $z$ | impedance |
| $z_{sd}$ | source-drain impedance |
| $z_{sg}$ | source-gate impedance |
| $z_{gd}$ | gate-drain impedance |
| $z_r$ | right-end channel impedance |
| $\phi$ | phase of impedance |
| $\tau$ | time constant |
| $\mu$ | mobility |

## I. Introduction

The theory presented here is the result of a search conducted with the aim of finding in a most quick and simple manner—at the expense of numerical accuracy—the physical parameters on which the dynamic properties of the field-effect transistor (FET) depend, how strong they influence the dynamic qualities of FET's, and what recommendations can be given as to proper choice of material or structure for FET's with good high-frequency performance.

We shall first derive an equivalent network for a FET which permits approximate calculation of the dynamic response of a field-effect transistor to small, periodically applied gate signals, if the static situation in respect to the operating point is already known. Static solutions on which the dynamics are based will be introduced, with special emphasis placed on the case of a small gate electrode length (i.e., a gate length comparable to the thickness of the transistor), and on the effects due to carrier velocity saturation. Plots of source-to-gate impedance, gate-to-drain impedance, source-to-drain impedance and transconductance are shown for the cases where there was available information on the dc conditions at the operating point.

The main result will be that lumped FET's need not be slow compared to bipolar transistors. Additional high-frequency potentialities may be found if sophisticated devices such as distributed FET's are studied, but this is not done here. Ratios of $g_m/C$ in the 10 to 30 GHz range may be expected if the transistors are made small enough

The authors are located at the IBM Zurich Research Laboratory, 8803 Rüschlikon-ZH, Switzerland.

* This work has been supported partly by Office of Aerospace Research, U. S. Air Force contract F19628-68-C-0343.

and from a highly doped material with a high value of saturated drift velocity. Typical data may be: (1) carrier concentration $N \approx 10^{17}/\text{cm}^3$ (or higher, but electric breakdown at the gate will probably soon set an upper limit); (2) saturated drift velocity $v_s \approx 1$ to $2 \times 10^7$cm/sec; and (3) gate length and interspace to source and drain $\delta \approx 1$ micrometer. In addition, the transistors have to be thin enough (a fraction of one micrometer thick) for reduction of their dc losses to tolerable limits.

## II. Equivalent network of a FET

Let us consider a two-dimensional field-effect transistor structure, as shown in Fig. 1, with width $b$ perpendicular to the cross-section.

Let us suppose further that the *static* properties of the FET are known, in particular the resistance per unit length $R_0$, of the conducting channel at position $x$, the capacitance per unit length $C_0$, of the conducting channel against the gate electrode at position $x$, where the coordinate $x$ measures the length of the boundary line separating the depleted and conducting regions, with its origin at the end between source and gate electrodes.

If we now apply a *small and periodic* gate voltage at angular frequency $\omega$, a small ac current $\Delta i(x, \omega)$ will flow through the conducting channel.

The current $i$ is partly due to a change in the number of free carriers per unit length, $Q$, of the conducting channel, and partly due to a change of their drift velocity, $v$:

$$\Delta i = \Delta Q \cdot v + Q \cdot \Delta v = \Delta(Qv). \tag{1}$$

In addition, $\Delta i$ and $\Delta Q$ are related by the equation of continuity

$$\frac{\partial \Delta i}{\partial x} + \frac{\partial \Delta Q}{\partial t} = 0$$

$$\frac{\partial \Delta i}{\partial x} = -\frac{\partial \Delta Q}{\partial t} = -j\omega \Delta Q. \tag{2}$$

Combining Eqs. (1) and (2) yields

$$\Delta i = -\frac{v}{j\omega}\frac{\partial \Delta i}{\partial x} + Q \Delta v. \tag{3}$$

So far, no explicit dependence of the drift velocity, $v$, of the free carriers on the electric field, $E$, in the channel has been introduced. It is done now, using the simplified relation shown in Fig. 2, which takes into account the typical low field (constant mobility, $\mu$) and high field (constant drift velocity, $v_s$) ranges properly but assumes a sharp transition at $E = E_c$ from one to the other region instead of a smeared-out one.

Then, for $E < E_c$ (for constant mobility there is no space charge in the channel) a change $\Delta Q$ of free carriers in the conducting channel is associated with a change of the channel-to-gate voltage, $\Delta V(x, \omega)$, given by

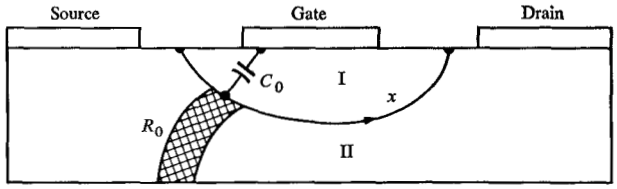
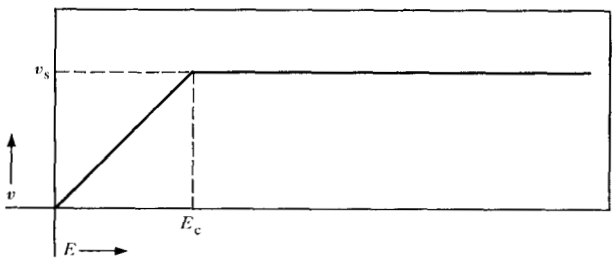



**Figure 1** Field-effect transistor structure and notations: I = Depleted region; II = Conducting channel region; $R_0$ = Resistance per unit width of conducting channel; $C_0$ = Channel-to-gate capacitance per unit width of conducting channel.

**Figure 2** Drift velocity versus electric-field relationship. $v_s$ = Saturation drift velocity; $E_c$ = critical electric field.




$$\Delta V = \frac{1}{C_0}\Delta Q = -\frac{1}{j\omega C_0}\frac{\partial \Delta i}{\partial x}. \tag{4}$$

In addition, we have

$$\left.\begin{aligned} \Delta v = \mu \Delta E &= -\mu \frac{\partial \Delta V}{\partial x} \\ \mu Q &= \frac{1}{R_0} \end{aligned}\right\} \tag{5}$$

so that we need to find from Eqs. (3), (5) and (4),

$$\Delta i = -\frac{v(x)}{j\omega}\frac{\partial \Delta i}{\partial x} + \frac{1}{R_0(x)}\frac{\partial}{\partial x}\left(\frac{1}{j\omega C_0(x)}\frac{\partial \Delta i}{\partial x}\right). \tag{6}$$

Equation (6) is a differential equation for the ac current. It can be solved provided that $v(x)$, $C_0(x)$ and $R_0(x)$ are known.

For $E > E_c$ we have $\Delta v = 0$. Therefore, Eq. (3) simplifies directly into

$$\Delta i = -\frac{v_s}{j\omega}\frac{\partial \Delta i}{\partial x}. \tag{6'}$$

The characteristic channel impedance $z(x, \omega)$ becomes with Eqs. (4) and (6),

$$\frac{1}{z} = v(x)C_0(x) - \frac{1}{R_0(x)}\frac{C_0(x)}{\dfrac{\partial \Delta i}{\partial x}}\frac{\partial}{\partial x}\left(\frac{1}{C_0(x)}\frac{\partial \Delta i}{\partial x}\right) \tag{7}$$

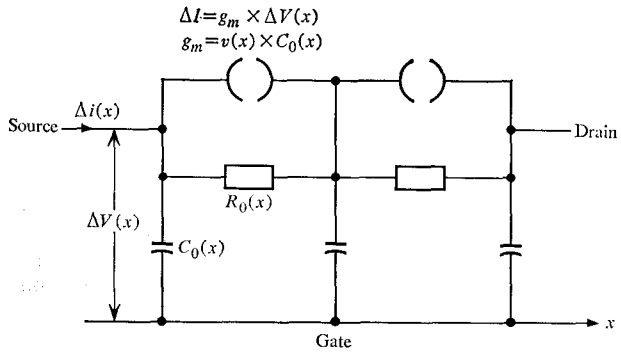
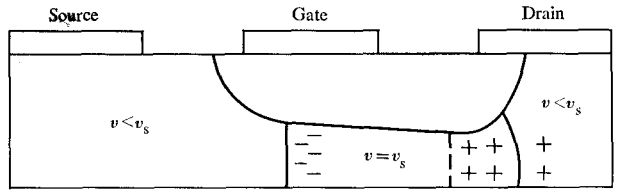



**Figure 3** Equivalent electric network of a field-effect transistor.

**Figure 4** Depletion region for a FET with a drift-velocity saturated conducting channel (schematic).




for $E < E_c$, and reaches, at position $x = x_0$ for which $E(x_0) = E_c$, a value

$$1/z = v_s C_0(x_0). \tag{7'}$$

An equivalent electric network corresponding to Eqs. (6) and (6') is a set of FET's forming the transmission line for wave propagation as shown in Fig. 3. For the drift velocity saturated channel section, $R_0 \rightarrow \infty$ because the mobility used here is $\mu = \partial v/\partial E$ and for $E > E_c$, $\mu = 0$. Sections with $C_0 = 0$ may exist at the source or drain directed ends of the channel; for these sections $g_m = 0$, so that nothing more than their total resistance $R_s$ (source resistance) or $R_d$ (drain resistance) must be taken into account.

A simple qualitative discussion may convince us that the methods of wave propagation are well suited to get an understanding for the solution of Eqs. (6) and (7).

For $E > E_c$, Eq. (6') shows after immediate integration that

$$\Delta i \approx \exp\left(-\frac{j\omega}{v_s} x\right), \tag{8}$$

which is an unattenuated current wave propagating through the channel with saturation drift velocity $v_s$. A very similar delay-line type solution

$$\left(\Delta i \approx \exp -\int \frac{j\omega}{v(x)}\, dx\right) \tag{9}$$

also exists in the case $E < E_c$, provided that the second term on the right-hand side of Eq. (6) is negligibly small compared to the first one. If, on the other hand, the second right-hand term of Eq. (6) predominates over the first, we are merely left with the differential equation for an *RC*-transmission line.

Note that the delay-line solutions are unidirectional (source $\rightarrow$ drain) waves, while the *RC*-line solutions are two-directional (source $\leftrightarrows$ drain) current waves.

## III. Some remarks on static performance, mainly transconductance

Unfortunately there is not ample theoretical information available on such static data as requested by Eq. (6). Shockley's[1] well-known gradual-case theory applies only to the case of a very-long-gate electrode (gate length $\gg$ transistor thickness) and, further, neglects drift velocity saturation of the carriers running through the conducting channel. The authors[2] handled, by means of an analog model, the case of FET's with a gate electrode length as small as the thickness of the transistor, but likewise ignored drift-velocity saturation in the conducting channel. Inclusion of these latter effects seems to call for a sophisticated numerical analysis, but sufficient results are not yet available to us. In this paper we suggest that for an operating point in the saturation region of the drain-current versus drain-voltage characteristics, a drift-velocity saturated channel section exists, as indicated schematically in Fig. 4, whose length increases with increasing drain voltage and may become predominating over the unsaturated section. The accumulation of charges is treated by Kennedy.[5] Before going into Eq. (6), we discuss with this diagram some static aspects involved with respect to the transconductance, $g_m$. If $g_m$ is mainly limited by the drift velocity saturation, we can identify the transconductance of the transistor with that of the saturated channel part of Fig. 4. In order to calculate the latter expression, we distinguish by subscripts l and r, values referring to the left- and right-hand ends of the saturated channel, respectively, and find with Eqs. (7) and (4) in a straightforward manner that the transconductance is equal to the reciprocal left-hand end channel impedance:

$$g_m = \frac{\Delta i_r}{\Delta V_l} = \frac{\Delta i_l}{\Delta V_l} = \frac{1}{z_l} = v_s \cdot C_{0l}. \tag{10}$$

The result expressed by Eq. (10) differs remarkably from the well-known formula:

$$g_m = \frac{\sigma b d}{L} \tag{11}$$

of the Shockley theory, where $\sigma$ = conductivity, $b$ = device width, $d$ = device thickness and $L$ = gate length. This difference is not surprising, because Eq. (11) was

derived for a different geometry (a large gate electrode length, $L$, compared to the transistor thickness, $d$) and because drift velocity saturation was ignored. Either the difference in geometry or the neglectance of drift-velocity saturation might account for the discrepancy in the results obtained for $g_m$. In order to analyze this question further, we now consider a transistor with a long gate electrode ($L \gg d$), but take drift-velocity saturation into account.

In this case, we have made the assumption that the drift velocity saturation is confined to an extremely thin region at the right-hand end of the conducting channel. This assumption is based on the fact that the output conductance is low for higher drain voltages, and therefore the drain current is almost independent of drain voltage. If the drain current is almost constant, this means that the geometry of the unsaturated part of the depletion-layer will change very little when drain voltage is varied. We assume, then, constant mobility in the whole channel. Thus it is only the *boundary condition* at the right-hand end of the channel which becomes different, according to whether or not drift-velocity saturation is taken into account. Shockley assumed $v_r \to \infty$, which corresponds to a right-end channel impedance $z_r = 0$ (instead of $z_r = 1/v_s C_{0r}$), and as a consequence of this he obtained drain-current saturation for complete pinch-off instead of for a finite width, $l_r$, of the conducting channel (see Fig. 5).

In order to derive the formal expression for the transconductance, we have to combine the Shockley equation[3]†

$$I_{sd} = \frac{\sigma b d}{L}\left[\left[1 - \frac{2}{3}\sqrt{\frac{V_r}{\frac{Ned^2}{2\varepsilon\varepsilon_0}}}\right]V_r + \left[1 - \frac{2}{3}\sqrt{\frac{V_{sg}}{\frac{Ned^2}{2\varepsilon\varepsilon_0}}}\right]V_{sg}\right] \tag{12}$$

where ($\sigma$ = conductivity, $N$ = free carrier concentration, $\varepsilon$ = dielectric constant, $b$ = width of the transistor, $L$ = gate electrode length, $d$ = thickness of the transistor) relating the drain current, $I_{sd}$, to the left- and right-hand end channel-to-gate voltages $V_{sg}$ (source-to-gate voltage) and $V_r$, respectively, with the expression for the saturated current

$$I_r = (Ne\, v_s) b\, l_r = I_{sd}, \tag{13}$$

($v_s$ = saturation-drift velocity) where according to the Shockley theory the width of the channel is

$$l_r = d\left[1 - \sqrt{\frac{V_r}{\frac{Ned^2}{2\varepsilon\varepsilon_0}}}\right], \tag{14}$$

and then find by straightforward calculation for small changes of current, $\Delta I_{sd}$, and voltage, $\Delta V_{sg}$,

† $I_A \to I_{sd}/2$, $a \to d$, $(U_A - U_B) \to V_r$ have to be substituted to adopt the notations, and correct for a slightly different geometry.

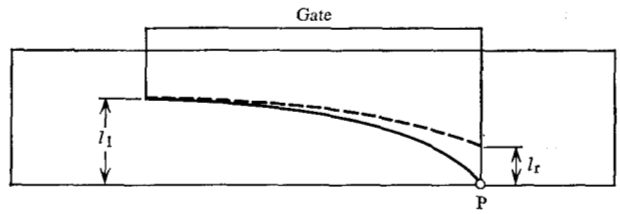



**Figure 5** Schematic plots of depletion layer according to Shockley's theory (—), and modified owing to drift-velocity saturation (– –). G = Gate electrode; P = Pinch-off point; $l_1$, $l_r$ = Left and right-hand widths of conducting channel.

$$g_m = \frac{\Delta I_{sd}}{\Delta V_{sg}} = \frac{g_{m(\text{Shockley theory})}}{1 + \left(\frac{1}{v_s \cdot C_{0r}}\right)\cdot\frac{l_r}{l_1}\cdot g_{m(\text{Shockley theory})}}, \tag{15}$$

where

$$g_{m(\text{Shockley theory})} = \frac{\sigma b d}{L}\left[1 - \sqrt{\frac{V_{sg}}{\frac{Ned^2}{2\varepsilon\varepsilon_0}}}\right], \tag{16}$$

$$C_{0_r} = \varepsilon\varepsilon_0 b/(d - l_r) \text{ and} \tag{17}$$

$$l_1 = d\left[1 - \sqrt{\frac{V_{sg}}{\frac{Ned^2}{2\varepsilon\varepsilon_0}}}\right]. \tag{18}$$

Since in this calculation we have not explicitly made use of the right-hand end channel impedance, it is instructive to verify Eq. (10) for this special case. Indeed we easily find from Eqs. (13) and (14):

$$z_r = \frac{\Delta V_r}{\Delta I_{sd}} = \frac{1}{C_{0_r}\cdot v_s}. \tag{19}$$

Figure 6 shows an example of $g_m$, $g_{m\,(\text{Shockley theory})}$ and $1/v_s\, C_{0r}$ versus free carrier concentration, $N$, for the data $b = 1$ mm, $L = 4\mu$, $d = 1\mu$, $\mu = 2500\ \text{cm}^2/\text{V}\cdot\text{sec}$, $V_{sg} = 0$ volt. The essence of the plots is that the Shockley theory formula is essentially correct as long as $g_{m\,(\text{Shockley theory})} \ll C_{0r}\cdot v_s$, but fails for $g_{m\,(\text{Shockley theory})} \gg C_{0r}\cdot v_s$. If the last-mentioned condition holds, $g_m$ is rather close to $C_{0r}\cdot v_s$, although, in principle, a larger transconductance than this can be achieved. Equation (10) is valid for the saturated part of the channel with either a long ($L \gg d$) or short ($L = d$) gate electrode, and $C_{0r}\cdot v_s$ is a "figure of merit" for the largest transconductance achievable.

Although the purpose of this paper is to discuss the performance of a lumped device, it is appropriate to make some remarks with regard to a distributed FET. By this we mean an FET with a width $b$ so large that source, gate and drain electrodes have to be considered as transmission lines for wave propagation. The lumped device may be considered as a cross-section of the distributed device.

Let us first see what transconductance values are required in order to obtain voltage, current or power amplifi-

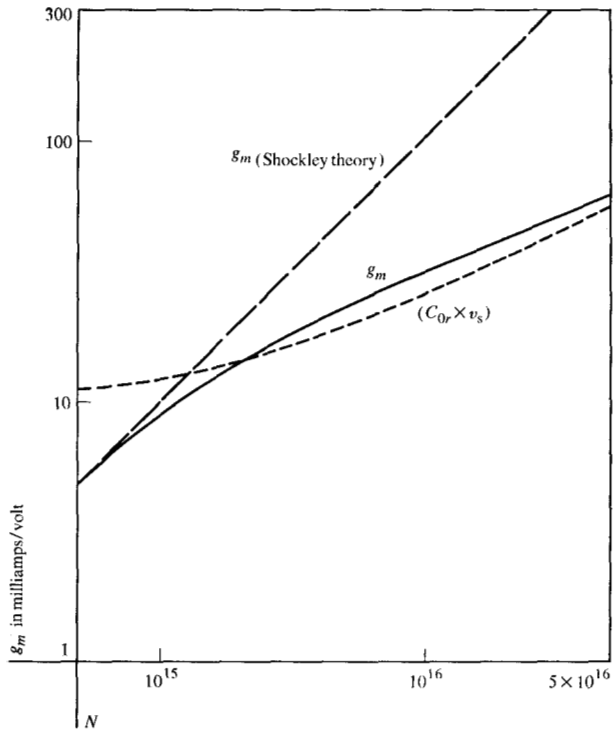
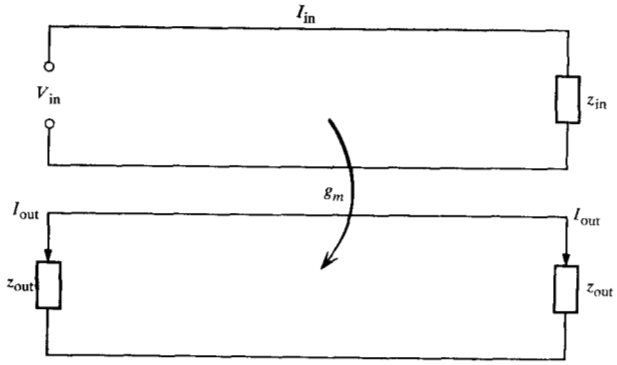



**Figure 6** Gradual case transconductance $g_m$ as a function of free carrier concentration, $N$. Values of $g_{m(\text{Shockley Theory})}$ and $(C_{0r} \cdot v_s)$ are plotted for comparison. $L = 4\mu$, $d = 1\ \mu$, $b = 1$ mm, $\mu = 2500\ \text{cm}^2/\text{V}\cdot\text{sec}$, $V_{sg} = 0$.

**Figure 7** Distributed FET. Schematic plot of source-to-gate and source-to-drain transmission lines.




cation in a distributed FET. Figure 7 shows a simplified schematic plot of the distributed device. Source-to-gate and source-to-drain lines are supposed to have no losses. The source-to-gate line is terminated by an impedance $z_{in}$ and carries a current $I_{in} = V_{in}/z_{in}$, if $V_{in}$ stands for the applied voltage. The source-to-drain line is terminated at both ends by an impedance $z_{out}$ and carries a current $I_{out} = g_m \cdot V_{in}$.

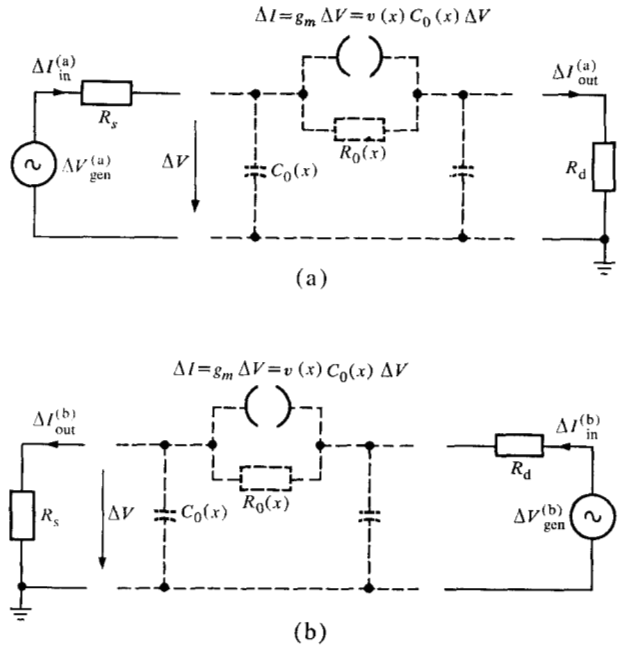



**Figure 8** Equivalent network relating $\Delta I_{in}^{(a,b)}$ and $\Delta I_{out}^{(a,b)}$ to $\Delta V_{gen}^{(a,b)}$.

The conditions for current, voltage and power amplification, respectively, read:

$$\frac{I_{out}}{I_{in}} = \frac{g_m z_{in}}{2} > 1, \tag{20}$$

$$\frac{V_{out}}{V_{in}} = \frac{g_m z_{out}}{2} > 1, \tag{21}$$

$$\frac{(IV)_{out}}{(IV)_{in}} = \frac{g_m^2 z_{in} z_{out}}{4} > 1. \tag{22}$$

If for further simplicity $z_{in} = z_{out} = z$, Eqs. (20), (21) and (22) read simultaneously:

$$g_m > 2/z. \tag{23}$$

Thus for $z = 50\Omega$ we need a transconductance larger than $g_m = 40$ mA/V.

Such high transconductance values may be achieved by means of a large capacity per unit length, $C_0$, which implies that the depleted layer has to be thin or that the transistor has to be wide enough. Unfortunately the width of the transistor is limited because the metal contacts, which are the conductors of the transmission line, are very thin, and the resistance of these conductors gives rise to attenuation. As a consequence, considerable dc losses occur in the transistor unless a sufficiently thin wafer plate has been chosen.

To be specific, we apply the gradual-case theory to the following data: $L = 1\mu$, $d = \frac{1}{3}\mu$, $b = 1$ mm, $N = 5 \times$

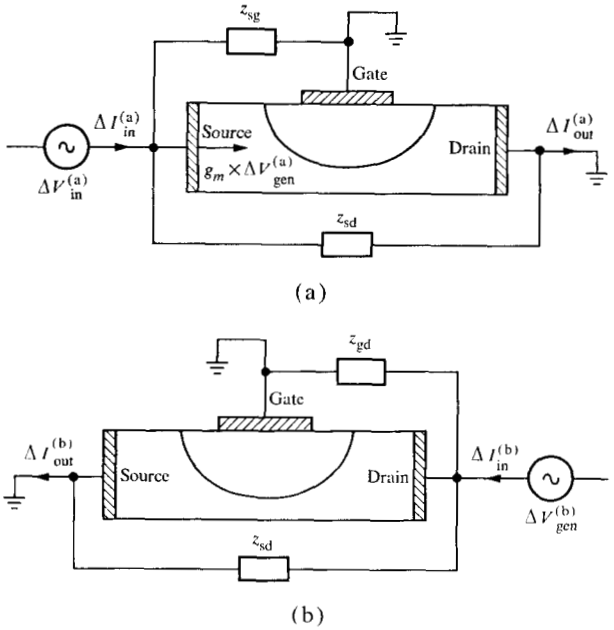



**Figure 9** "Vierpol parameter" of the field-effect transistor. $g_m$ = transconductance; $z_{sd}$ = source-to-drain impedance; $z_{sg}$ = source-to-gate impedance; $z_{gd}$ = gate-to-drain impedance.

$10^{16}/cm^3$, $\mu = 500$ cm$^2$/Volt·sec†, $v_s = 10^7$ cm/sec, $\varepsilon\varepsilon_0 = 10^{-12}$ Coulomb/Volt·cm, $V_{sg} = 0$, $V_{sd} = 1.4$ Volt. We then need to find a drain current $I_{sd} = 120$ mA, thus a dc power of $P = 170$ mW, and a transconductance $g_m = 64$ mA/V. The thickness of the depletion layer is zero at the source-directed end of the gate electrode, and 1870 Å at the drain-directed end of the gate, if we assume that no carrier accumulation takes place in the channel under the gate.

If, however, such thin depleted layers are required in order to achieve the transconductance value needed for distribution, one may question whether the assumption of a completely depleted barrier region is good enough. Although little experimental information referring to this aspect is available, it is encouraging. Silicon transistors have been prepared on layers as thin as 0.2 micrometer. The measured dc transconductance was comparable with that calculated for a completely depleted barrier region.

## IV. Dynamics of FET's: transconductance, impedances between electrodes and their frequency dependence

Let us apply a generating voltage $\Delta V_{gen}$ to either the source or drain electrode, while the other two electrodes (gate and either drain or source electrode) are kept at an ac "ground potential." $\Delta V_{gen}$ will produce an input (either source or drain) current $\Delta I_{in}$ and an output (either drain or source) current $\Delta I_{out}$, which can be calculated according to Figs. 8a and 8b, where currents and voltages corresponding to the two different cases are distinguished by an appropriate index letter "a" or "b". The reason for doing this calculation is that the FET's source-to-gate impedance, $z_{sg}$, gate-to-drain impedance, $z_{gd}$, source-to-drain impedance, $z_{sd}$, and transconductance, $g_m$, have a simple representation in terms of

$$\frac{\Delta I_{in}^{(a)}}{\Delta V_{gen}^{(a)}}, \frac{\Delta I_{out}^{(a)}}{\Delta V_{gen}^{(a)}}, \frac{\Delta I_{in}^{(b)}}{\Delta V_{gen}^{(b)}}, \frac{\Delta I_{out}^{(b)}}{\Delta V_{gen}^{(b)}}.$$

In fact we may easily verify by means of Figs. 9a and 9b that:

$$\left.\begin{aligned} \frac{\Delta I_{in}^{(a)}}{\Delta V_{gen}^{(a)}} &= g_m + \frac{1}{z_{sd}} + \frac{1}{z_{sg}} \\ \frac{\Delta I_{out}^{(a)}}{\Delta V_{gen}^{(a)}} &= g_m + \frac{1}{z_{sd}} \end{aligned}\right\} \tag{24}$$

$$\left.\begin{aligned} \frac{\Delta I_{in}^{(b)}}{\Delta V_{gen}^{(b)}} &= \frac{1}{z_{sd}} \\ \frac{\Delta I_{out}^{(b)}}{\Delta V_{gen}^{(b)}} &= \frac{1}{z_{sd}} + \frac{1}{z_{gd}}. \end{aligned}\right\} \tag{25}$$

We have numerically calculated $z_{sd}$, $z_{gd}$, $z_{sd}$ and $g_m$ as a function of frequency, $f$, for Examples 2 and 3 listed in Table 1.

• *Example 1*

The static information was obtained from the gradual case theory developed in Section III, which takes drift velocity saturation into account. The source-to-drain voltages $V_{sd}$ listed in Table 1 are the smallest required to achieve drain-current saturation with the prescribed source-to-gate voltage $V_{sg} = -1$ V. Note that these drain voltages depend on the ratio of gate electrode length to transistor thickness, which would not be true for the Shockley theory. The dynamic performance would not have been different, if a larger drain-voltage had been chosen.

Figure 10 shows the shape of the depleted regions. We recall that drift-velocity saturation is reached at the drain-directed end of the channel only, and that there is no "pinch-off" as would be according to the Shockley theory.

The 3-dB-point frequencies, $f_{3\,dB}$, with corresponding phase shifts, $\phi_{3dB}$, for $g_m$ are listed in Table 2. We have also listed for comparison the respective values appropriate to $v_s \to \infty$ (Shockley theory) and $V_{sd} \geq 3$V, without going into details of the calculation.

Since an algebraic expression for the input data of the dynamic equation was used here, there should be no difference between the dynamically calculated low-

† This order of magnitude for $N$ and $\mu$ was measured by Th. Mohr of this laboratory in epitaxially grown Si layers as thin as 0.2 micrometer. Further, these numbers are compatible with data collected by Motorola.[4]

**Table 1** Calculations for Examples 1, 2 and 3.

| *Example no.* | *Transistor geometry device width* $= 10^{-1}$ cm: *Gate length,* $L$ | *Transistor thickness,* $d$ | *Static operation point:* *Drain voltage,* $V_{sd}$ | *Gate voltage,* $V_{sg}$ | *Material data* $\varepsilon\varepsilon_0 = 10^{-12}\frac{C}{V\cdot cm}$; *Free carrier conc.* $N = 5 \times 10^{15}/cm^3$; *Drift velocity* $v$; *Mobility* $\mu$ | *Method for obtaining dc-information* |
|---|---|---|---|---|---|---|
| 1 | $3 \times 10^{-4}$ cm<br>$4 \times 10^{-4}$ cm<br>$10 \times 10^{-4}$ cm | $10^{-4}$ cm | 1.37 V<br>1.63 V<br>2.24 V | −1 V | $\mu = 1250\frac{cm}{V\cdot sec}$;<br>$v_s = 9 \times 10^6$ cm/sec<br>$v = \mu E$ for $v < v_s$<br>$v = v_s$ for $E \geq 7 \times 10^3$ V/cm | Gradual Case Theory according to Sect. II (incomplete pinch-off) |
| 2 | $10^{-4}$ cm | $10^{-4}$ cm | 1.8 V | −2 V | $\mu = 1250\frac{cm^2}{V\cdot sec}$;<br>$v_s = 9 \times 10^6$ cm/sec<br>$v = \mu \cdot E$ for $v < v_s$<br>$v = v_s$ for $E \geq 7 \times 10^3$ V/cm | Analog model |
| 3 | $10^{-4}$ cm | $10^{-4}$ cm | 13 V | −2 V | $\mu = 1250\frac{cm^2}{V\cdot sec}$;<br>$v_s = 9 \times 10^6$ cm/sec<br>$v = \mu E$ for $v < v_s$<br>$v = v_s$ for $E \geq 7 \times 10^3$ V/cm | Analog model for unsaturated channel sections. Drift saturated section treated as a delay line |

**Figure 10** Depleted regions for Example 1.

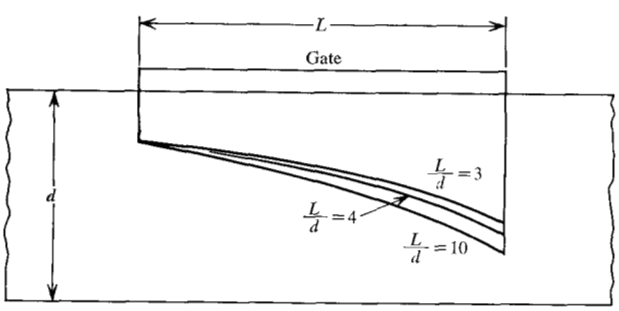



**Table 2** Three-dB point frequencies, with phase shifts.

| $L/d$ | $f_3$ dB GHz | $\phi_3$ dB Degrees | $f_3$ dB GHz | $\phi_3$ dB Degrees |
|---|---|---|---|---|
| | *Drift velocity limited* | | *Drift velocity not limited* | |
| 3 | 18.0 | −66 | 33.0 | −80 |
| 4 | 11.5 | −70 | 19.0 | −80 |
| 10 | 2.5 | −77 | 3.0 | −80 |

frequency transconductance and the static value according to Eq. (15). This in fact is true, and is put into evidence in Table 3.

For frequencies up to the 3-dB point, $g_m$ may be presented by

$$\frac{g_{m(f)}}{g_{m(0)}} = \frac{e^{-j2\pi\nu\tau_1}}{1 + j2\pi\nu\tau_2}, \tag{26}$$

where the time constants $\tau_1$ and $\tau_2$ are to be taken from Table 4.

The frequencies $f_{45}$ corresponding to a phase shift of 45° for $z_{sg}$ are listed in Table 5 together, for comparison, with the respective values appropriate to $v_s \to \infty$ (Shockley theory) and $V_{sd} \geq 3$V (drain current saturation).

At low frequencies, $z_{sg}$ is roughly

$$z_{sg(f)} = \frac{1}{2\pi fC} \tag{27}$$

where

$$C = \frac{\varepsilon\varepsilon_0 b}{d}\int_0^L \frac{dx}{1 - l(x)} \tag{28}$$

is the capacity of the depletion layer. In principle, it is not quite correct to calculate $C$ as if no drain current were

**Table 3** Calculations of low-frequency transconductance.

| | $L/d$ | | |
|---|---|---|---|
| $g_m$ *according to:* | 3<br>mA/V | 4<br>mA/V | 10<br>mA/V |
| Static Eq. (16) | 10.0 | 8.8 | 4.5 |
| Dynamic equation $f \to 0$ | 10.1 | 8.8 | 4.6 |

**Table 4** Time constant.

| $L/d$ | $\tau_1$ in psec | $\tau_2$ in psec | $\tau_3$ in psec |
|---|---|---|---|
| 3 | 3.4 | 8.8 | 3.1 |
| 4 | 6.0 | 14.0 | 5.3 |
| 10 | 3.6 | 64.0 | 27.0 |

**Table 5** Frequencies corresponding to 45° phase shift.

| $L/d$ | $f_{45°}$ GHz | |
|---|---|---|
| | *Drift velocity limited* | *Drift velocity not limited* |
| 3 | 52.0 | 60.0 |
| 4 | 30.0 | 34.0 |
| 10 | 5.8 | 5.3 |

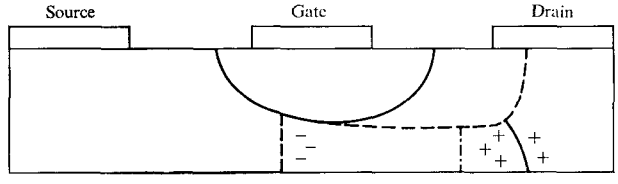
**Table 6** Calculations for $C$ allowing for drain current flow.

| $L/d$ | $\left(\frac{1}{2\pi f z_{sg}}\right)_{f\to 0}$ in pF | $C$ in pF |
|---|---|---|
| 3 | 0.45 | 0.49 |
| 4 | 0.57 | 0.64 |
| 10 | 1.20 | 1.50 |

**Table 7** Gain-bandwidth products compared with those from Shockley theory.

| $L/d$ | $f_0$ in GHz | |
|---|---|---|
| | *Drift velocity limited* | *Drift velocity not limited* |
| 3 | 3.2 | 7.8 |
| 4 | 2.2 | 4.4 |
| 10 | 0.5 | 0.7 |

**Figure 11** Depletion layers for Examples 2 (—) and 3 (– – –). The drift-velocity saturated section is put into evidence by its space charge. –·–·– marks the neutral cross-section.




flowing, and therefore we need to find some discrepancy, which is put into evidence in Table 6.

For frequencies corresponding to phase shifts up to 45°, $z_{sg}$ may be approximated by

$$z_{sg(p)} = \frac{1}{2\pi f C}\,\frac{1}{1 + j\cdot 2\pi f \tau_3}, \qquad (29)$$

where the time constant $\tau_3$ has to be taken from Table 4.

The gain-bandwidth product $f_0 = (1/2\pi)(g_m/C)$ is listed in Table 7, together for comparison with the corresponding values given by the Shockley theory. The data listed demonstrate the increasing importance of taking drift velocity saturation into account with decreasing $L/d$.

So far, the source resistance, $R_s$, was supposed to be zero. $R_s$ has mainly the effect of increasing $\tau_3$.

We conclude by recalling that for all Examples 1, no longitudinal electric fields have been taken into account in the depletion layer. Although this gradual case approximation procedure is generally used, it might not be justified, even for $L \gg d$, in the neighborhood of the drain-directed end of the gate electrode. Therefore, emphasis will be placed on the succeeding examples to use a better static background in this region, and to point out the consequences on the dynamic performance of field-effect transistors.

• *Examples 2 and 3*

Examples 2 and 3 have been chosen so as to demonstrate the delay effects involved with an extended channel section where drift-velocity saturation is established. The static pictures for 2 and the assumed picture for 3 are plotted in Fig. 11. In Example 2, the source-to-drain voltage is so small that the carrier drift-velocity is everywhere below its saturation value except for the most narrow cross-section of the conducting channel; in Example 3 the drain voltage is so large that drift-velocity saturation is reached

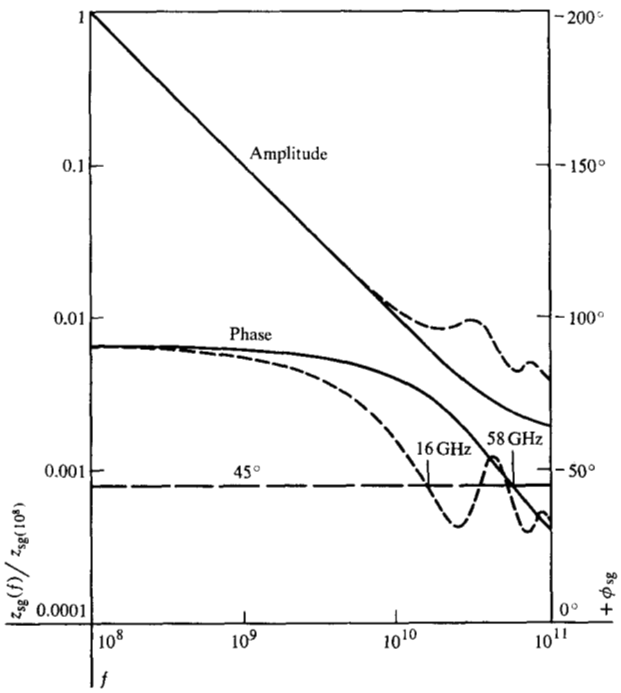



**Figure 12** Absolute value $|z_{sg}|$ and phase $\phi_{gd}$ of source-to-gate impedance for Examples 2 (—) and 3 (– – –).

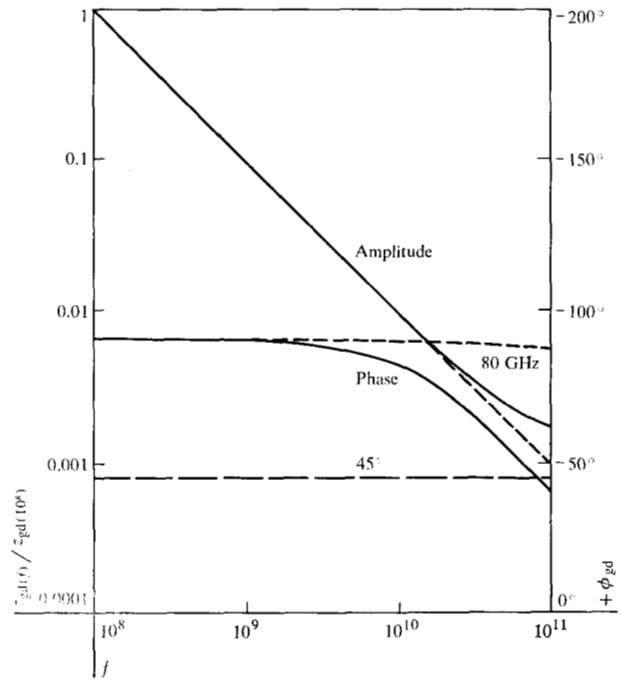



**Figure 13** Absolute value $|z_{gd}|$ and phase $\phi_{gd}$ of gate-to-drain impedance for Examples 2 (—) and 3 (– – –) versus frequency, $f$.

in an appreciable section of the channel. The saturated part of the channel contains negative space charge at the source-directed end and positive charge at the drain-directed end. The positive charge extends into the region where $v < v_s$. In our treatment, however, the space charge outside the drift-velocity saturated section was ignored.

The frequency dependence of $z_{sg}$, $z_{gd}$ and $g_m$ is plotted in Figs. 12, 13, and 14, ($z_{sd} \rightarrow \infty$) and the values of $R_{sd}$, $C_{sg}$, $C_{gd}$ and $g_m$ as determined from the respective zero-frequency impedances are listed in Table 8. The $g_m$ values listed above compare well with those predicted earlier by the authors from a less-detailed discussion of drift-velocity saturation effects (Fig. 7 of Ref. 2).

A striking result evident from Fig. 12 is that oscillations of $z_{sg}$ appear at frequency intervals of roughly $f = 5 \times 10^{10}$ Hz for Example 3. The origin of these can be understood. Suppose that a voltage pulse (Fig. 15a) has been applied to the gate electrode. It will produce at the left end of the drift-velocity saturated channel (see Fig. 15b for a schematic plot) a charge pulse which propagates with velocity $v_s$. After a time $\tau = L_{eff}/v_s$ ($L_{eff}$ = effective length of saturated channel section) it will leave the channel at its right end. We may suggest that thereby a charge pulse of opposite sign is produced in order to bring the channel back to its originally neutral state. If we produce a new pulse at this moment and continue to do so at a rate of $f = 1/\tau$, we ensure that no gate current is set up

**Figure 14** Absolute value $|g_m|$ and phase $\phi_{g_m}$ of transconductance for Examples 2 (—) and 3 (– – –) versus frequency.

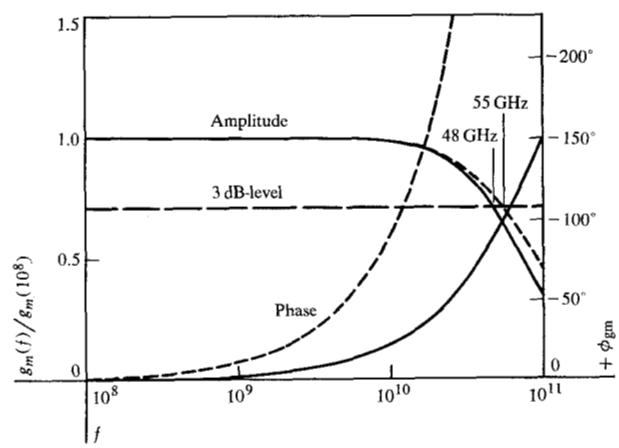



**Table 8** Values of $R_{sd}$, $C_{sg}$, $C_{gd}$ and $g_m$ determined from zero-frequency impedances.

| *Example* | $R_{sd}$ | $C_{sg}$ | $C_{gd}$ | $g_m$ |
|---|---|---|---|---|
| 2 | $\infty$ | 0.25 pF | 0.16 pF | 11.5 mA/V |
| 3 | $\infty$ | 0.49 pF | 0.03 pF | 12 mA/V |

in the transistor, which means that the input impedance is infinitely large (losses have been neglected). Appropriate data are roughly $L_{eff} = 2 \times 10^{-4}$ cm and $v_s = 10^7$ cm/sec, which yield $\nu_0 = 5 \times 10^{10}$ Hz, thus the right order of magnitude. In the frequency range below that for oscillator effects, $z_{sg}$ may be represented by

$$z_{sg(\nu)} = \frac{1}{2\pi f C_{sg}} \frac{1}{1 + j2\pi f \tau_3}. \tag{30}$$

The appropriate time constants $\tau_3$ are listed in Table 9. Evidently there is a pronounced difference in favor of Example 2, on which we should like to comment.

The source-to-gate capacitance is larger for Example 3 than for Example 2 (Table 8). If the difference $\Delta C_{sg}$ is due to the capacitance associated with the saturated channel section of Example 3, it should be roughly $\Delta C_{sg} = C_{0r} \cdot L_{eff}$. Since we have $C_{0r} = 10^{-9}$ $F$/cm and $L_{eff} = 2 \times 10^{-4}$ cm we need to find $\Delta C_{sg} = 0.2$ pF, which is the right order of magnitude. An increase in source-to-gate capacitance is unfavorable, because it gives rise, in connection with the source resistance $R_s$, to additional phase shift between input voltage and current, and additional damping. This effect may be minimized either by reducing $R_s$ by means of technological tricks, or by choosing the smallest possible source-to-drain voltage required (for a given source-to-gate voltage) to achieve drift-current saturation.

The situation is just the reverse with respect to the gate-to-drain impedance $z_{gd}$ (see Table 8 and Fig. 13). The Miller capacitance $C_{gd}$ is smaller for Example 3 than for Example 2. The gate-to-drain impedance may be represented by

$$z_{gd(f)} = \frac{1}{2\pi f C_{gd}} \frac{1}{1 + j2\pi f \tau_4} \tag{31}$$

with time constant $\tau_4$ as listed in Table 9. Thus Example 3 is dynamically more favorable than Example 2.

The transconductance, $g_m$, is represented below its 3 dB level by

$$\frac{g_{m(f)}}{g_{m(0)}} = \frac{e^{-j2\pi f \tau_1}}{1 + j2\pi f \tau_2}, \tag{32}$$

where the relaxation times $\tau_1$ and $\tau_2$ are contained in Table 9. We see from Fig. 14, that for Example 2 the 3 dB point is located at 48 GHz with a corresponding phase shift of $\approx 85°$; for Example 3 corresponding data are 55 GHz and $\approx 450°$. The difference in phase shifts must roughly be equal to the delay $fL_{eff}/v_s \times 360°$ [see Eq. (8)] associated with the waves propagating along the saturated channel section in Example 3, and in fact we find good agreement for $f = 55$ GHz, $L_{eff} = 2 \times 10^{-4}$ cm and $v_s = 10^7$ cm/sec. Pronounced delay effects may cause serious instability problems in a distributed device; therefore the cross-sectional elements in a distributed device

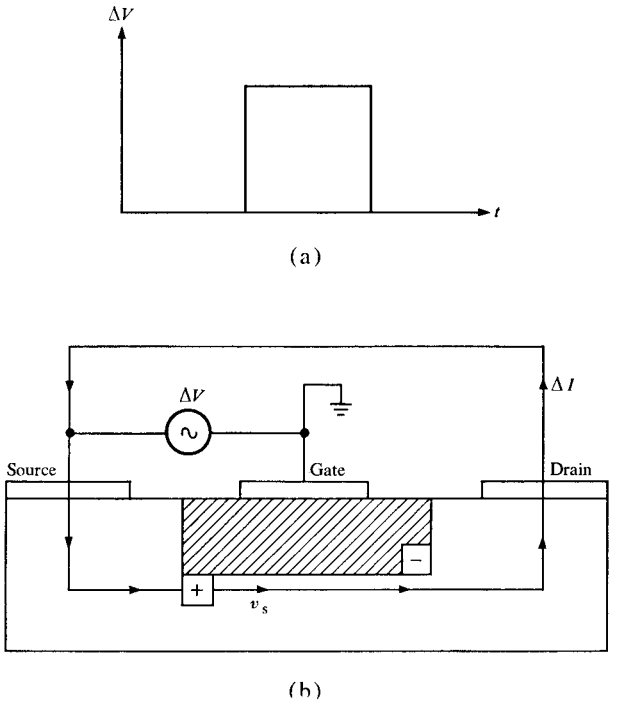



**Figure 15** Applied source-to-gate voltage pulse a) and schematic plot of current pulse product b). Hatched area is the depleted region.

**Table 9** Time constants for Examples 2 and 3, in psec.

| *Example* | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_4$ |
|---|---|---|---|---|
| 2 | 2.3 | 3.3 | 2.8 | 2.0 |
| 3 | 20.0 | 2.9 | 10.0 | 0.0 |

should be operated with the smallest possible source-to-drain voltage required for a given source-to-gate voltage to obtain drift-current saturation.

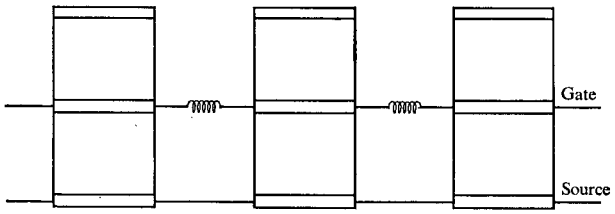
If we remember that a saturated channel permits current propagation without attenuation, we may also understand why the 3-dB point of the transconductance is located at a higher frequency for Example 3 than for Example 2.

The gain bandwidth product $f_0$ for amplification is not very much different for Examples 2 and 3; the average value becomes

$$f_0 = \frac{1}{2\pi} \frac{g_m}{C_{sg} + C_{gd}} \approx 4 \text{ GHz}. \tag{33}$$

This is a drastically smaller value than the 15 GHz attained with neglection of drift-velocity saturation.

The result for $f_0$ obtained here completes Table 7 for the geometry $L/d = 1$. We recognize that $f_0$ increases with decreasing $L/d$ and saturates around $L/d \approx 3$.




(a)

(b)

**Figure 16** a) Source-to-date channel of distributed FET, b) input signal and distortion of transmitted signal.

**Figure 17** Gain-bandwidth product.

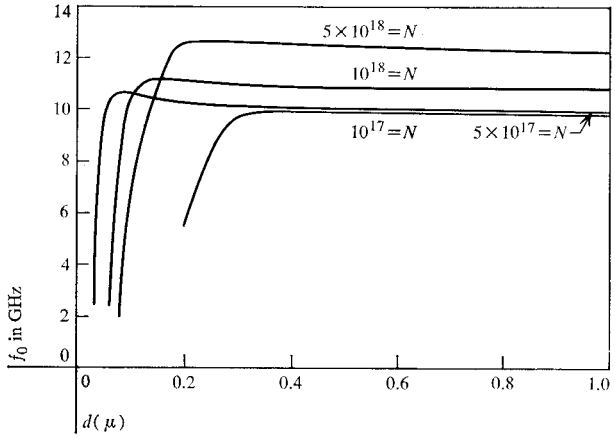



**Figure 18** Saturated drain current.

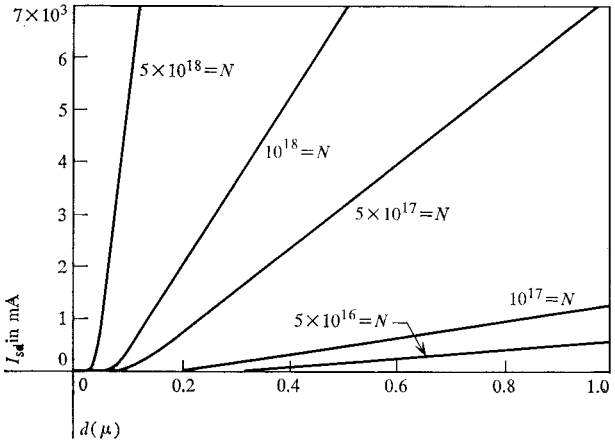



**Table 10** Relation between mobility and doping level.

| $\mu$, $\frac{cm^2}{V\cdot sec}$ | 640 | 520 | 350 | 290 | 160 |
|---|---|---|---|---|---|
| $N$, $cm^{-3}$ | $5\times10^{16}$ | $10^{17}$ | $5\times10^{17}$ | $10^{18}$ | $5\times10^{18}$ |

The $g_m/C$ ratio is certainly a "figure of merit" for the lumped device. We want to point out that, in addition, it also becomes figure of merit for the distributed device, provided that the cross-sectional element has a conducting channel which consists predominantly of the drift-velocity saturated section, so that:

$$g_m \approx C_{0r} \cdot v_s \tag{34}$$

and

$$C_{sg} \approx C \approx C_{0r} L. \tag{35}$$

In order to develop the argument, let us recall that the basic idea of distribution is to get rid of the frequency-limiting effects due to capacitance in a lumped device by connecting a large number of lumped devices in parallel, with inductive coupling elements in between (Fig. 16b). The parameter $C_s$ then becomes a cross-sectional element in an $LC$-transmission line (the source-to-gate channel), on which signals can be transmitted without distortion up to the highest frequencies, if losses are neglected.

An implicit assumption made in this argument is that a cross-sectional element has a source-to-gate impedance which is a capacitance up to the highest frequencies used for transmission. But it is clear that at frequencies $f \gg 1/2\pi\tau$, where $\tau$ stands for the flight-time of the free carriers along the gate, this is no longer the case: at these frequencies, the time needed to charge or discharge the depleted region becomes so large that signals propagating along the source-to-gate channel become distorted (Fig. 16a), so that distribution no longer presents the expected advantage. If we calculate $f_0$ by means of Eqs. (34) and (35), it turns out that the upper frequency limit for undistorted transmission on the source-to-gate channel of a distributed device is equal to the frequency limit for amplification with subsequent stages of a lumped device:

$$f_0 = \frac{1}{2\pi} = \frac{g_m}{C} = \frac{1}{2\pi}\frac{L}{v_s} = \frac{1}{2\pi\tau}. \tag{36}$$

In this consideration the parasitic capacitances and the capacitances at both edges of the gate have not been taken into account. These capacitances can be compensated for, and therefore the upper limit for undistorted transmission will be higher than the gain bandwidth product $f_0$.

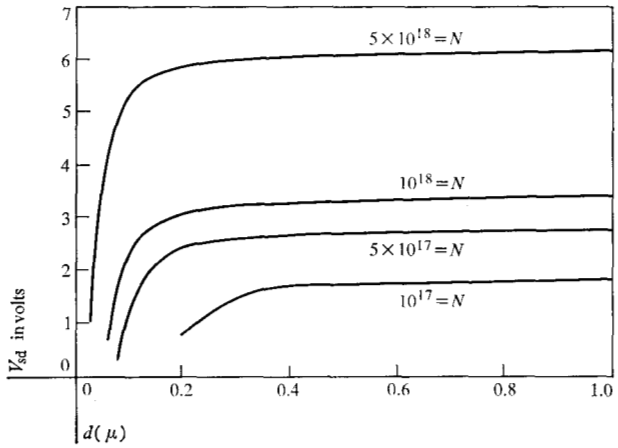



**Figure 19** Minimum drain voltage for current saturation.

**Figure 20** Dissipated dc power.

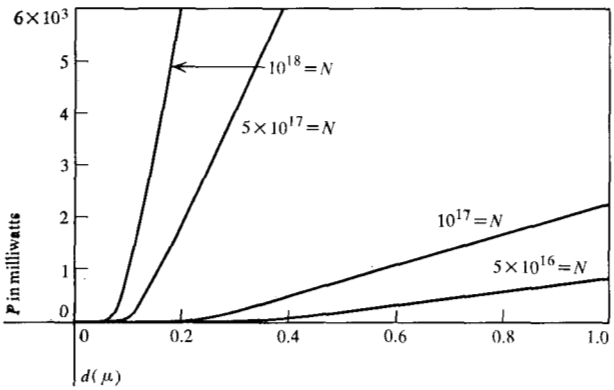
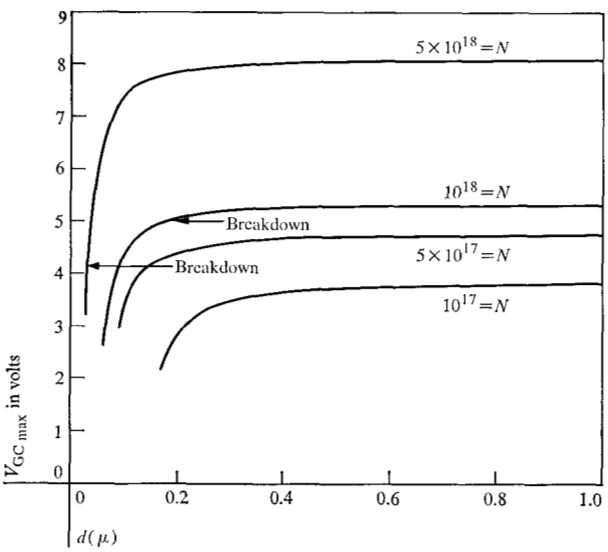



**Figure 21** Highest gate-to-channel voltage.



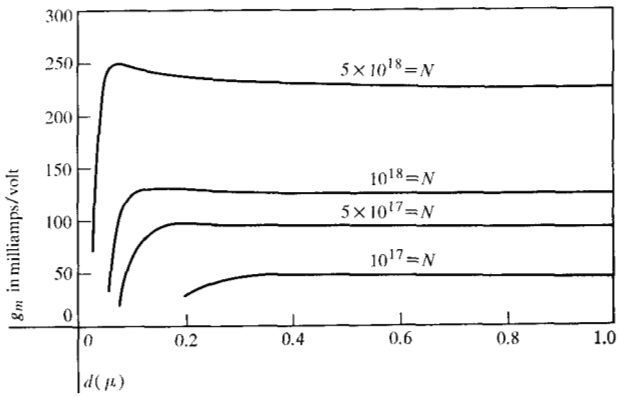




**Figure 22** Transconductance.

## V. Discussion of results

The frequency limit for a given FET in a circuit depends on the way in which the transistor is being used. In wideband lumped devices the gain-bandwidth product

$$f_0 = \frac{g_m}{2\pi C}$$

is certainly a "figure of merit."† We find within the framework of the present study that $f_0 \approx 10$ GHz can be obtained with Si Schottky-barrier FET's. The transistors have to be made from highly doped material (free carrier concentration $N \approx 10^{17}/\text{cm}^3$ or higher, until electric breakdown at the gate junction sets an upper limit) with a gate length of 1 $\mu$m (and equal distances to source and drain). In addition they have to be thin, no thicker than a few tenths of a $\mu$m, in order to keep the dc losses within tolerable limits. $f_0$ should rise to the 30 GHz range if GaAs is used instead of Si, due to the larger saturation drift velocity in the former material, if not, instabilities cause disturbing effects.

The value $f_0$ for silicon FET's is displayed in Fig. 17 as a function of transistor thickness, $d$, and for various carrier concentrations, $N$. The plots refer to $b = 1$ mm, $L = 1$ $\mu$m (distances to source and drain are equal), $V_{sg} = -2$V, $R_{sg} = 0$ and a mobility versus doping level relationship as listed in Table 10.

Some additional data: the saturated drain current, $I_{sd}$, and the smallest source-to-drain voltage, $V_{sd}$, needed for current saturation are depicted in Figs. 18 and 19. For the so-defined operating points we then plotted the dissipated dc power, $P$ (Fig. 20), the highest gate-to-channel voltage, $V_{gc_{max}}$ (Fig. 21), and the transconductance $g_m$ (Fig. 22).

† If FET's are used in a distributed device, such as for example, in a chain amplifier, then

$$f_{max} = \frac{f_0}{2}\sqrt{\frac{R_{sd}}{R_{sg}}}, \tag{37}$$

which is the highest frequency for power gain with matched input and output, is the appropriate "figure of merit." Here $R_{sd}$ and $R_{sg}$ stand for source-to-drain resistance and source-to-gate resistance (including the resistance along the gate contact), respectively.

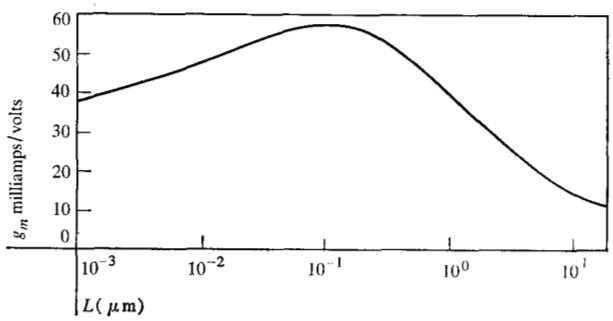



**Figure 23** Optimized transconductance as a function of gate width.

Figure 23 shows finally how the highest achievable transconductance (optimized by choosing the transistor thickness appropriately at a constant doping level $N = 0.8 \times 10^{17}/cm^3$ and a gate voltage $V_{sg} = I_{sd} \cdot R_{sg} + V_d$, where $V_d$ denotes a diffusion voltage $V_d = 0.5$ V) depends on the size of the transistor structure. The transconductance would admit a maximum value for a gate length equal to some tenths of a micrometer, which cannot be realized for technological reasons at present. However, since the maximum is flat, a one-micrometer structure comes quite close to the best conditions.

## VI. Acknowledgments

The programming and computing work was carried out by H. Keller. His contribution is gratefully acknowledged. We should also like to thank Prof. G. Kohn, who first initiated the idea of a Schottky-barrier FET in our laboratory.

## References

1. W. Shockley, *Proc. IRE* **40,** 1365 (1952).
2. K. E. Drangeid and R. Sommerhalder, *Helv. Phys. Acta* **41,** 766 (1968).
3. H. Salow et al., *Der Transistor,* Springer-Verlag, 1963, p. 395.
4. Motorola, *Integrated Circuit Design,* 1964.
5. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. Develop.,* **14,** 95 (1970, this issue).

*Received September 25, 1969*