Equivalent Circuit

See fig for simple eq. circuit. In addition to intrinsics, some parasitic elements: R_G= gate resistance, R_S= source resistance. G_π is a measure of leakage current on the gate; it should be negligible, but if the diode is near turn-on, that may not hold.

Gate-source capacitance $C_{GS} = C_G - C_{DG}$. Total input capacitance on the gate = C_G. Capacitance between gate and drain = C_{DG}. They can be obtained from

$$C_G = \left.\frac{\partial Q_G}{\partial V_{GS}}\right|_{V_{DS}} \text{ and } C_{DG} = -\left.\frac{\partial Q_G}{\partial V_{DS}}\right|_{V_{GS}} \tag{1}$$

C_{DG} contains both intrinsic and parasitic components: intrinsic for the part that is across depletion layer, parasitic that is not.

Intrinsic components: g_{mi}= intrinsic transconductance

$$g_{mi} = \left.\frac{\partial I_{DS}}{\partial V_{GS}}\right|_{V_{DS}} = \left.\frac{\partial I_{DS}}{\partial Q_G}\right|_{V_{DS}} \left.\frac{\partial Q_G}{\partial V_{GS}}\right|_{V_{DS}} = \frac{C_G}{\tau} \tag{2}$$

g_{mi} measures transistor's gain. For max. g_m we need to minimize τ (sounds familiar?). Note: cannot avoid capacitance; it is an intrinsic part of the device.

Further components: output conductance and capacitance.

f_T as usual; for max. frequency of oscillation we have:

$$f_{max} = \frac{f_T}{2\sqrt{G_0(R_G + R_S) + 2\pi f_T C_{DG} R_G}} \tag{3}$$

Basic structure: see fig. Effective gate length definition and channel thickness. Requirement: eff. length should not be significantly different from physical length. Gate design calls for restricted get-finger width. Problem: phase delay along the gate finger. Rule of thumb: Z/2 $< \lambda/16$. 5mm at 1 GHz down to 50μm at 100 GHz.

For microwaves: impedance matching imposes some limits (connecting MESFETs in parallel). Furthermore, temperature must not be allowed to rise above certain level, which reduces the number of cells. Final resort: use circuit combining.

DC characteristics

As for JFETs, pinchoff voltage $V_P = qN_D a^2/(2\epsilon_s)$ = voltage needed to deplete the channel of thickness a.

Threshold voltage: onset of current conduction. $V_T = V_{bi} - V_P$. (all of this for depletion mode devices).

Depletion vs. enhancement mode (normally on and normally off). For enhancement devices, V_T is slightly above 0V (they are depleted in equilibrium).

The max. gate voltage for both cases is limited by forward biasing the gate: cca. 0.7V. Depletion devices have thicker channel regions and/or higher doping. They also provide more currents at highest gate voltage. Breakdown - avalanching between gate and drain.

Saturation region in fig. Saturation current for $V_{GS} = 0$ is called I_{DSS}.

Different model for small devices, based on velocity saturation (fig) . Here the current saturates due to velocity saturation, not as in long channel case. For this to occur we need only 5-10 kV/cm in the channel. Also, dipole formation.

$$I_{DS} = qv_sZN_D(a-h) = qav_sZN_D\left(1-\sqrt{\frac{V_{bi}-V_{GS}}{V_P}}\right) \tag{4}$$

which gives expressions for transconductance:

$$g_{mi} = \frac{v_s\epsilon_s Z}{a}\sqrt{\frac{V_P}{V_{bi}-V_{GS}}} = v_s Z\sqrt{\frac{qN_D\epsilon_s}{2(V_{bi}-V_{GS})}} \tag{5}$$

Other models for JFET/MESFET: field dependent mobility and two-region model (linear and saturated velocity). For small MESFETs ($L \leq 2\mu m$): saturated-velocity model. Full saturation assumed under the gate.

$$I = qv_sZ(a-h)N_D \text{ and for non-uniform doping } I = v_sZ\int_h^a \rho(y)dy. \tag{6}$$

Using the depletion layer width expression for Schottky diode we obtain:

$$I_{DS} = qav_sZN_D\left[1-\sqrt{\frac{V_{bi}-V_{GS}}{V_P}}\right] \tag{7}$$

and for intrinsic transconductance:

$$g_{mi} = \frac{v_s\epsilon_sZ}{a}\sqrt{\frac{V_P}{V_{bi}-V_{GS}}} = v_sZ\sqrt{\frac{qN_D\epsilon_s}{2(V_{bi}-V_{GS})}} = \frac{v_sZ\epsilon_s}{h} \tag{8}$$

In some apps. a constant g_m may be beneficial. This can be accomplished by varying the doping profile, unlike the long gate case. see fig. Equation implies reduction of g_m with application of gate voltage, but g_m/C_{GS} (what is this?) remains constant since $C_{GS} = \epsilon_s/h \rightarrow g_m/C_{GS} = v_sZ = const.$

2-D considerations and dipole layer formation. First look at Si and then GaAs.

Consider a thin Si layer on insulating substrate. For small V_{DS} the region between S and D behaves like a resistor. At larger voltages the velocity does not increase with the same rate as electric field; eventually, velocity saturates. Now add a gate contact.

The width of the opening (depl. layer) depends on the drain voltage applied. The resistance is increased since the channel is constricted. Generally, $I_D = Zqan(x)v(x)b(x)$. Below the critical field, $n = N_D$. Largest reverse bias at drain end. To keep the current continuity the velocity (and el. field) must increase at drain end. This ends once the velocity is saturated (case c))

For even larger drain voltage ($> V_{Dsat}$) the depletion region widens toward the drain, but the point at which electrons reach v_s moves towards source (case d)). This movement towards source results in reduction of voltage at that point so that channel opening increases slightly. This results in increased current coming into the saturation region and positive slope of output I-V curve.

From saturation point to drain potential increases and the depletion layer widens. To maintain current continuity carrier concentration must increase (vel. is saturated) $\rightarrow$ accumulation region is formed up to the point at which depletion layer width is equal to the one at the saturation point.

Since now the channel widens, there is no need for all those accumulated electrons and their concentration drops below the background doping causing positive charge to appear. Fields are still high due to the accumulated electron charge and velocity remain saturated. Any increase in drain voltage above saturation voltage is used up in formation of the dipole layer. Note that it extends beyond the physical gate.

For some applied gate voltage the qualitative picture is the same, but various characteristic points are reached sooner.

GaAs more complex due to different velocity-field curve. The main points are shown in figure. The drift velocity rises to peak at x_1 (max. v point at critical el. field) then falls to the low saturated value at the gate edge. For current continuity, electron concentration increases over the background doping $\rightarrow$ accumulation. Now not only the channel narrowing but also velocity reduction contribute to accumulation. From the peak of el. field to the edge of the gate depl. layer (x_2, x_3), the opposite happens: electrons move faster causing a strong depletion layer. The charges in accumulation and depletion layer are nearly equal and most of the drain voltage drops across this dipole layer.

For very short gate lengths: overshoot effects. Check the Monte Carlo simulation results. For I-V curves: a) electron velocity in the channel is more dependent on drain-source field resulting in larger slope (output conductance), b) no saturation for more negative gate voltages (near pinchoff); result of conduction through substrate. To avoid this leakage a buffer layer is often grown or opposite dopand implanted underneath the gate.

Few rules: choose material with the highest average velocity, make gate length as small as possible. This ensures maximum f_T, f_{max}. With given L, choose L/a around 4–5 for optimum trade-off between f_T, parasitics and short channel effects. Too large L/a $\rightarrow$ reduced current drive, too small $\rightarrow$ high G_o, short channel effects and L_{eff} may be significantly larger than L.

Determine doping (N_D): for digital apps determined by V_T required by logic circuits. For analog apps: V_B, I_D, swing, g_m, etc. For breakdown: $V_{BDG} \approx 5 \cdot 10^{13}/(N_D a)$ ($N_D a < 2.3 \cdot 10^{12} cm^{-2}$). For larger products it reverts to bulk-like breakdown (lower). see fig.

$$V_{BDG} \approx 60 \left(\frac{E_g}{1.1}\right)^{3/2} \left(\frac{10^{16}}{N_D}\right)^{3/4} \tag{9}$$

Due to 2-D nature of field at drain end, V_{BDG} actually higher than in bulk case. For output I-V: $V_B = V_{BDG} + V_T$.

Small-signal Operation

For more complete equivalent circuit see fig. Under normal conditions C_{GS} is much larger than other capacitances. Parasitic resistances R_S, R_D contain contact resistance as well as resistance leading to channel. Some section of the channel is not in velocity saturation, resulting in series resistance R_i that contributes to the charging time (along with R_S and R_G). G_o is a function of velocity saturation and dipole formation; C_{DC} is the capacitance of the dipole domain. C_{DS} is the substrate capacitance. Time delay associated with electrons traversing the channel can be modeled via $g_{mi}e^{-j\omega\tau}$.

Source series resistance reduces transconductance $g_m = g_{mi}/(1 + g_{mi}R_S)$. Series inductance on the source introduces possible stability problems and faster frequency fall-off. Minimization of R_S, L_S very important. High-frequency operation limited by finite transit time and RC constants. Transit time under the gate:

$$\tau = \frac{L}{\mu E_x} \approx \frac{L^2}{\mu V_D} \text{ for const. } \mu; \tau = \frac{L}{v_s} \text{ for vel. saturation} \tag{10}$$

Another constraint: NOISE. see fig for regimes. Performance measured by noise figure (NF or F): ratio of S/N ratio at input and output.

$$NF = 10\log\left(1 + \frac{T_n}{290}\right) \approx 10\log\left(1 + 2\pi f K_f C_{GS}\sqrt{\frac{R_G + R_S}{g_{mi}}}\right) \tag{11}$$

T_n is the equivalent noise temperature, K_f is an empirical factor and depends on geometry and material quality. The last term can be expressed as $(f/f_T)K_f\sqrt{g_{mi}(R_S+R_G}$, indicating that the increase in f_T leads to better noise performance. Same for decrease in gate length. For reduction of R_G, use small gate widths (Z). Also need to know the gain at any given noise figure for cascading purposes.

1/f noise due to various mechanisms, mostly various traps. Not a big problem for linear amps that operate at high frequency, but important for mixers and oscillators. Minimized by reducing traps (buffer layers) and passivation.

Large-signal Operation

Maximize power output or efficiency, or both. In class A, power output is (see fig) $P_m = I_m(V_B - V_{Dsat})/8$. The ratio of rf power output to dc power is drain efficiency:

$$\eta_D = \frac{\eta_{max}(V_B - V_{Dsat})}{V_B + V_{Dsat}} \tag{12}$$

With class A operation and constant $g_m(V_{GS}), \eta_{max}$ =50%.

Power-added efficiency:

$$\eta_{PA} = \frac{P_{out}(rf) - P_{in}(rf)}{P_{in}(dc)} = \eta_D \left(1 - \frac{1}{G}\right) \tag{13}$$

where G is large signal power gain. Important for systems with limited dc power and/or heat dissipation is a problem. Improved by reducing V_{Dsat} $\rightarrow$ parasitic resistances R_S, R_D and using semiconductor with high low-field mobility (intrinsic part).

Scaling: increase power and efficiency by increasing breakdown voltage V_B ?? What about possible trade-offs vs. I_m, f_T, G?

To first order: $V_B \approx V_{BDG}$, which is proportional to max. electric field $\mathcal{E}_{\Updownarrow}$ at breakdown. see fig for various materials. The field between source and drain at breakdown is $\mathcal{E}_B = V_B/L_{eff}$. Therefore, $\mathcal{E}_{\mathcal{B}}$ and $\mathcal{E}_m$ are related since V_B and V_{BDG} are related. The product of f_T and V_B becomes:

$$V_B f_T \approx V_B \left(\frac{v_a}{2\pi L_{eff}}\right) \sim \frac{\mathcal{E}_B v_a}{2\pi} \sim \frac{\mathcal{E}_m v_a}{2\pi} \tag{14}$$

v_a is average velocity under gate. Valid for fixed V_T and $N_D a$ does not shift between two breakdown regions. Fundamental trade-off between V_B and f_T.

Trade-off between V_B and I_m. For $V_B \gg V_{Dsat}$ and $V_{BDG} \gg |V_T|$, the max. rf power is

$$P_{max} = \frac{I_m V_B}{8} \approx \frac{I_m V_{BDG}}{8} \approx \frac{(q v_a Z)(5 \cdot 10^{13})}{8} = K_1 \mathcal{E}_m v_a Z \tag{15}$$

To first order P_{max} and $I_m V_B$ product vary only with material dependent quantities $\mathcal{E}_m, v_a$ and with gate width Z. The quantity without Z is maximum power density (W/mm of gate width). For $f < f_T$ it is constant, but for $f > f_T$ it falls of as 1/f. Explanation: not enough time for transit and the energy gained is only a fraction of what is normally gained in "full" transit. If value of $v_{sat} = 1.2 \cdot 100^7 cm/s$ is used, $P_{d,max} = 1.3W/mm$, close to experiments.

Digital Operation

No CMOS in GaAs due to low hole mobility. Basic inverters: see fig. Active vs. resistive loading. Load characteristic of the active load is a mirror image the I-V curve for depletion mode device. Another possibility: current limiter.

Direct-coupled FET logic: DCFL. Output connected directly to the next stage's gate. For this $V_{DD} < V_{GM}$ should hold (unrealistic), where V_{GM} is forward turn-on voltage. If $V_{DD} > V_{GM}$ the output (high) is clamped at $V_{OH} = V_{GM}$. The load transistor must be able to supply the gate current for the next stage and sustain the voltage $V_{DD} - V_{GM}$.

For voltage transfer function see fig. Saturation due to: a) close spacing of the switching transistor's characteristics in linear region (for high input), and b) compression of g_m near V_T (low input). Noise margins: measure immunity to spurious triggering and crosstalk. For enhancement mode MESFET: V_T high enough to obtain reasonable NM, but low enough not to restrict logic swing. Since $V_{OH} < V_{GM}$ the swing is small and out of necessity V_T must be small, cca. 0.1–0.3V. Consequence: high V_T uniformity is a must.

For switching performance: charging and discharging of capacitances (see fig); the effective node capacitance C_N includes all the components together. Problem: bias dependent. How to calculate various switching times (for definition, see fig):

$$t_F = 0.8(V_{OH} - V_{OL}) < \frac{C_N}{I_T - I_L} >, \quad t_{D,F} = 0.5(V_{OH} - V_{OL}) < \frac{C_N}{I_T - I_L} > \tag{16}$$

$$t_R = 0.8(V_{OH} - V_{OL}) < \frac{C_N}{I_L} >, \quad t_{D,R} = 0.5(V_{OH} - V_{OL}) < \frac{C_N}{I_L} > \tag{17}$$

I_L is the load current, and I_T is switching transistor current (fig). Difficult to find analytic expression due to bias dependence. Rise time t_R is improved by clamping the output voltage (fig). Ring oscillators for gate delay, but that's the best case.

Power dissipation: static+dynamic. $P_{st} \approx V_{DD} I_{LL}/2$ where I_{LL} the load current at output low. $P_{dyn} \approx f C_N V_{DD}^2/2$ for $V_{OH} - V_{OL} \approx V_{DD}$. In this case we get for power-delay product

$$E_{min} = t_D P_D \approx C_N V_{DD}^2/2 \tag{18}$$

FET design: a) increase f_T; transconductance/capacitance is high along the logic swing. b) high g_m; enables large I_{DS} to develop quickly to charge/discharge capacitances. c) for L/a see before. d) adjust the doping for enhancement mode device. e) V_B need not be much larger than V_{DD} (remember the trade-off) f) optimize Z.

Processing of simple MESFET (fig). High frequency performance (figs). Special applications: dual gate FET; used for mixers, high-gain amplifier, amp with gain control. MMICs on GaAs as low-loss substrate and for active devices. Digital process. High-speed characteristics (figs).

JFET: higher forward turn-on voltage (p-n junction), larger logic swing, larger NM, higher yields (V_T related). PBT.