RF INTEGRATED CIRCUITS: NOTES

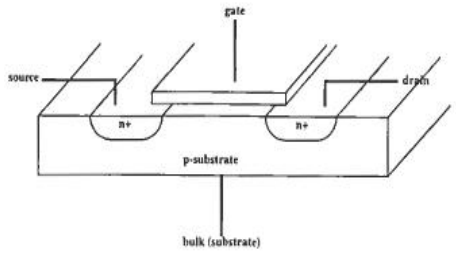
UNIT-2

1. REVIEW OF MOS DEVICE PHYSICS - MOS DEVICE REVIEW:-

1.1. Introduction:- This chapter focuses attention on those aspects of transistor behavior that are of immediate relevance to the RF circuit designer. Separation of first-order from higher-order phenomena is emphasized. so there are many instances when crude approximations are presented in the interest of developing insight. As a consequence this review is intended as a supplement to - rather than a replacement for - traditional rigorous treatments of the subject. In particular, we must acknowledge that today's deep submicron MOSFET is so complex a device that simple equations cannot possibly provide anything other than first-order (maybe even zeroth-order) approximations to the truth. The philosophy underlying this chapter is to convey a simple story that will enable first-pass designs which are then verified by simulators using much more sophisticated models. Qualitative insights developed with the aid of the zeroth-order models enable the designer to react appropriately to bad news from the simulator. We design with a simpler set of models than those used for verification.

1.2. MOSFET PHYSICS: THE LONG-CHANNELAPPROXIMATION:- As you well know, a basic n-channel MOSFET as shown in figure 1.1 consists of two heavily doped n-type regions, the source and drain, which constitute the main terminals of the device. The gale was made of metal in early incarnations but is now made of heavily doped polysilicon, whereas the bulk of the device is p-type and is typically rather lightly doped. In much of what follows we will assume that the substrate (bulk) terminal is at the same potential as the source. However it is extremely important to keep in mind that the substrate constitutes a fourth terminal whose influence cannot always be ignored.

As an increasing positive voltage is applied to the gate holes are progressively repelled away from the surface of the substrate. At some particular value of gate voltage (the threshold voltage *V,),* the surface becomes completely depleted of charge. Further increases in gate voltage induce an *inversion layer*, composed of electrons supplied by the source (or drain), that constitutes a conductive path ("channel") between source and drain. When the gate-source voltage is several *kT/q* above *V"* the device is said to be in strong inversion



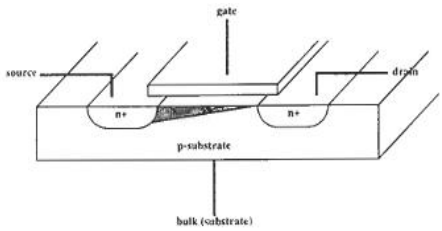


Figure 1.1.n-channel MOSFET Figure 1.2.n-channel MOSFET shown of boundary between triode and saturation

The foregoing discussion implicitly assumes that the potential across the semiconductor surface is a constant (i.e .. that there is zero drain-to-source voltage). With this assumption, the induced inversion charge is proportional to the gate voltage above the threshold, and the induced charge density is constant along the channel. However if we do apply a positive drain voltage *V.* then the channel potential must increase in some manner from zero at the source end to *V* at the drain end. TIle net voltage available to induce an inversion layer therefore decreases as one approach the drain end of the channel. Hence, we expect the induced channel charge density to vary from a maximum at the source (where *Vs'* minus the channel potential is largest) to a minimum at the drain end of the channel (where *Vg.,* minus the channel potential is smallest), as shown by the shaded region representing charge density in Figure 1.2

Specifically, the channel charge density has the following form is given by

$$Q_n(y) = -C_{ox}\{[V_{gs} - V(y)] - V_t\},$$

where *Qn (.v)* is the charge density at position y, *Cox.* is ϵox/Iox" and *V(y)* is the channel potential at position *y.* Note that we follow the convention of defining the y-direction as along the channel, Note also that Co., is a capacitance *per unit area,* The negative sign simply reflects that the charge is made up of electrons in this NMOS example above Equation is all we really need in order to derive the most important equations governing the terminal characteristics.

1.2.1. DRAIN CURRENT IN THE LINEAR (TRIODE) REGION:- The linear or triode region of operation is defined as one in which *V,s* is large enough (or *Vds* small enough) to guarantee the formation of an inversion layer for the whole distance from source to drain. From our expression for the channel charge density we see that it has a zero value when

$$[V_{gs} - V(y)] - V_t = 0.$$

The charge density thus first becomes zero at the drain end at some particular voltage, Hence, the boundary for the triode region is defined by

$$[V_{gs} - V_{ds}] - V_t = 0 \implies V_{ds} = V_{gs} - V_t = V_{dsat}.$$

As long as *Vd, < Vdsat* the device will be in the linear region of operation. Having derived an expression for the channel charge and having defined the linear region of operation, we are now in a position to derive an expression for the device current in terms of the terminal variables, Current is proportional to charge times velocity is

$$I_D = -WQ_n(y)v(y).$$

The velocity at low fields (remember, this is the "long channel" approximation) is simply the product of mobility and electric field. Hence,

$$I_D = -WQ_n(y)\mu_n E.$$

Where W is the width of the device, Substituting now for the channel charge density, we obtain

$$I_D = -WC_{ox}[V_{gs} - V(y) - V_t]\mu_n E.$$

Next we note that the (v-directed) electric field *E* is simply (minus) the gradient of the voltage along the channel. Therefore

$$I_D = \mu_n C_{ox} W[V_{gs} - V(y) - V_t]\frac{dV}{dy}.$$

$$I_D\,dy = \mu_n C_{ox} W[V_{gs} - V(y) - V_t]\,dV.$$

Now integrating on both side solve the Id is

$$\int_0^L I_D\,dy = I_D L = \int_0^{V_{ds}} \mu_n C_{ox} W[V_{gs} - V(y) - V_t]\,dV.$$

At last we have the following expression for the drain current in the triode region is given by

$$I_D = \mu_n C_{ox}\frac{W}{L}\left[(V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right].$$

Note that the relationship between drain current and drain-to-source voltage is nearly linear for small *Vds.* Thus, a MOSFET in the triode region behaves as a voltage controlled resistor. The strong

sensitivity of drain current to drain voltage is qualitatively similar to the behavior of vacuum tube triodes, which lend their name to this region of operation.

1.2.2.DRAIN CURRENT IN SATURATION:- When *Vds* is high enough so that the inversion layer does not extend all the way from source to drain, the channel is said to be "pinched off." In this case, the field felt by the channel charge ceases to increase, causing the total current to remain constant despite increases in *Vds*. Calculating the value of this current is easy; all we need do is substitute Vdsat, for *Vds* in our expression for current is

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{gs} - V_t)V_{dsat} - \frac{V_{dsat}^2}{2}\right].$$

This simplifies to

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{gs} - V_t)^2.$$

hence, in saturation, the drain current has a square-law dependence on the gate to source voltage and is (ideally) independent of drain voltage. Because vacuum tube pentodes exhibit a similar insensitivity of plate current to plate voltage. this regime is occasionally called the pentode region of operation.

The transconductance of such a device in saturation is easily found from differentiating our expression for drain current is

$$g_m = \mu_n C_{ox}\frac{W}{L}(V_{gs} - V_t).$$

This may be expressed as

$$g_m = \sqrt{2\mu_n C_{ox}\frac{W}{L}I_D}.$$

Thus, in contrast with bipolar devices. a long-channel MOSFET's transconductance depends only on the square root of the bias current.

1.2.3. CHANNEL· LENGTH MODULATION:- So far, we've assumed that the drain current is independent of drain-source voltage in saturation. However, measurements on real devices always show a disappointing lack of such independence. The primary mechanism responsible for a nonzero

output conductance in long-channel devices is channel-length modulation (CLM).Since the drain region forms a junction with the substrate, there is a depletion region surrounding the drain whose extent depends on the drain voltage. As the drain *v*oltage increases, the depletion zone's width increases as well effectively shortening the channel. Since the effective length thus decreases the drain current increases.

To account for this effect, the drain current equations for both triode and saturation are modified as follows the

$$I_D = (1 + \lambda V_{ds})I_{DO}.$$

where *Ido* is the drain current when channel-length modulation is ignored and where the parameter λ is a semi-empirical constant whose dimensions are those of inverse voltage. The reciprocal of λ is often given the symbol *VA* and called the Early voltage, after the fellow who first explained nonzero output conductance in bipolar transistors (where it is caused by an analogous modulation of base width with collector voltage).*The* graphical significance of the Early voltage is that it is the common extrapolated zero current intercept of the *Vds-Id* device curves. Measurements on real devices almost never permit extrapolation to a single point but engineers are rarely bothered by such trivial details.

1.2.4. DYNAMIC ELEMENTS:- So far, we've considered only DC parameters. Let's now take a look at the various capacitances associated with MOSFETs. These capacitances limit the high-frequency performance of circuits. so we need to understand where they come from and how big they are First, since the source and drain regions from reverse-biased junctions with the substrate. One expects the standard junction capacitance from each of those regions to the substrate, these capacitances are denoted *Cjsb* and *Cjdb*. as shown in Figure 1.3.where the extent of the depletion region has been greatly exaggerated.

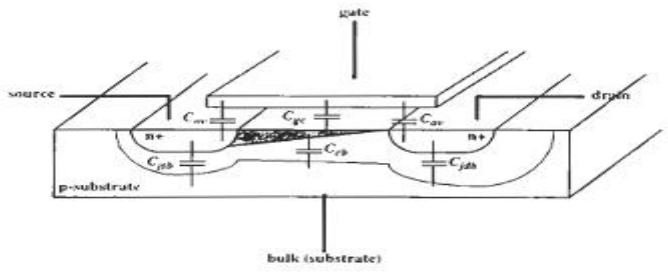


Figure 1.3.Mosfet with capacitances

There are also various parallel plate capacitance terms in addition to the junction capacitances. The capacitors shown as *Cov* in Figure 1.3 represent gate-source and gate-drain *overlap* capacitances; these are highly undesirable but unavoidable. During manufacture, the source and drain regions may diffuse laterally by an amount similar to the depth that they diffuse. Hence, they bloat out a bit during processing and extend underneath the gate electrode by some amount. As a crude approximation. Once may take the amount of overlap, *LD*• as 2/3 to 3/4 of the depth of the source-drain diffusions, hence the equation is given by

$$C_{ov} \approx \frac{\varepsilon_{ox}}{t_{ox}} W L_D = 0.7 C_{ox} W x_j.$$

Where *Xj* is the depth of the source-drain diffusions. ϵox. is the oxide's dielectric constant (about *3.9*ϵo*)* and tox is the oxide thickness. The parallel plate overlap terms arc augmented by fringing and thus the "overlap “capacitance would be nonzero even in the absence of physical overlap. In this context, once should keep in mind that (in modern devices) the gate electrode is actually considerably thicker than the channel is long. so the relative dimensions of Figure *1.3* are misleading. Think of a practical gate electrode as a tall oak tree instead of thin plate. In addition the interconnecting wires to the source and drain are hardly of negligible dimensions. See Figure 1.4. Because the thickness of the gate electrode now scales little (if at all), the "overlap" capacitance now changes somewhat slowly from generation to generation.

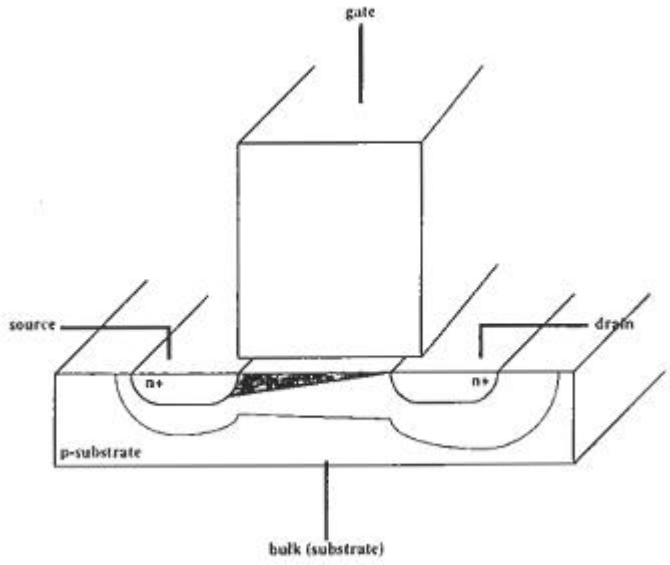


figure *1.4.* More accurate depiction of relative gate electrode size.

Another parallel plate capacitance is the gate-to-channel capacitance *Cgc.* Since both the source and drain regions extend into the region underneath the gate the effective channel length decreases by twice the bloat. *Ld,* Hence, the total value of *Cgc* is

$$C_{gc} = C_{ox}W(L - 2L_D).$$

In strong inversion the charge carriers at the surface and those in the bulk are of the opposite type. In between there is a depletion region. As a result, there is also a capacitance between the channel and the bulk, *Ccb.* that behaves as a junction capacitance. Its value is approximately

$$C_{cb} \approx \frac{\varepsilon_{Si}}{x_d} W(L - 2L_D),$$

where *x d* is the depth of the depletion layer whose value is given by

$$x_d = \sqrt{\frac{2\varepsilon_{Si}}{qN_{sub}}|\phi_s - \phi_F|}.$$

The quantity within absolute-value bars is the difference between the surface potential and the Fermi level in the substrate. In strong inversion (for both triode and saturation regions), this quantity has a magnitude of twice the Fermi level.

1.2.5. HIGH-FREQUENCY FIGURES OF MERIT:- It is perhaps natural to attempt to characterize multidimensional quantities with a single number: laziness is universal after all. In the specific case of high-frequency performance two figures of merit are particularly popular. These are *WT* and Wmax, which are the frequencies at which the current and power gains (respectively) are extrapolated to fall to unity. It is worthwhile to review briefly their derivation since many engineers forget the origins and precise meanings of these quantities. Often drawing incorrect inferences as a result.

The most common expression for *WT* assumes that the drain is terminated in an incremental short circuit while the gate is driven by an ideal current source. As a consequence of the shorted termination. *WT* does not include information about drain-bulk capacitance. The current-source drive implies that series gate resistance similarly has no influence on *WT* Clearly. both *rg* and *Cjdh* can have a strong effect on high-frequency performance but *WT* simply ignores this reality. Furthermore.

the gate-to-drain capacitance is considered only in the computation of the input impedance: its feed forward contribution to output current is neglected. With these assumptions the ratio of drain current to gate current is

$$\left|\frac{i_d}{i_{in}}\right| \approx \frac{g_m}{\omega(C_{gs}+C_{gd})},$$

which has a unity value at a frequency

$$\omega_T = \frac{g_m}{C_{gs}+C_{gd}}.$$

Now the frequency at which the (extrapolated) current gain goes to unity really has no fundamental importance; it is simply easy to compute. Perhaps more relevant is the frequency at which the maximum power gain is extrapolated to fall to unity. To compute will in general is quite difficult. however. so we will invoke several simplifying assumptions to make an approximate derivation possible. Specifically we compute the input impedance with an incrementally shorted drain and ignore the feed forward current through *Cgd* just as in the computation of *WT* .We do consider the *feedback* from drain to gate through *Cgd* in computing the output impedance. However which is important because computation of the maximum power gain requires termination in a conjugate match.

With these assumptions we can calculate the power delivered 10 the input by the current source drive as simply

$$P_{in} = \frac{i_{in}^2 r_g}{2}.$$

Where *rg the* series gate resistance is the only dissipative element in the input circuit. The magnitude of the short-circuit current gain at high frequencies is approximately given by the same expression used in the computation of *WT is*

$$\left|\frac{i_D}{i_{in}}\right| \approx \frac{\omega_T}{\omega}.$$

It is also straightforward to show that the resistive pan of the output impedance is roughly

$$g_{out} \approx g_m \cdot \frac{C_{gd}}{C_{gd} + C_{gs}} = \omega_T \cdot C_{gd}.$$

If the conjugate termination has a conductance of this value then the power gain will be maximized. with half of the R", generator's current going into the conductance of the termination and the balance into the device itself. The total maximum power gain is therefore

$$\frac{P_L}{P_{in}} \approx \frac{\frac{1}{2}\left(\frac{\omega_T}{\omega} \cdot i_{in} \cdot \frac{1}{2}\right)^2 \frac{1}{(\omega_T \cdot C_{gd})}}{\frac{i_{in}^2 r_g}{2}} \approx \frac{\omega_T}{\omega^2 4 r_g C_{gd}}.$$

which has a unity value at a frequency given by

$$\omega_{max} \approx \frac{1}{2}\sqrt{\frac{\omega_T}{r_g C_{gd}}}.$$

II is clear that Wn,"' depends on the gale resistance, so it is more comprehensive in this regard than *WT*· Because judicious layout can reduce gale resistance to small values, Wmax can be considerably larger than *WT* for many MOSFETs. The output capacitance has no effect Wmax. Because it can be tuned out with a pure inductance and therefore does not limit the amount of power that may be delivered to a load.

Measurements of both Wmax and *WT* are carried out by increasing the frequency until a noticeable drop occurs in maximum power gain or current gain. A simple *ex*trapolation to unity value then yields Wmax and *WT*. Because these arc extrapolated values. it is not necessarily a given that one may actually construct practical circuits operating at say Wmax These figures of merit should instead be taken as rough indications of high-frequency performance capability.

1.2.6. TECHNOLOGY SCALING IN THE LONG-CHANNEL LIMIT :- Now that we have examined both the static and dynamic behavior of MOSFETs. We can derive an approximate expression for *WT* in terms of operating point. process parameters and device geometry. We have already derived an expression for *gm.* so all we need is an expression for the requisite capacitances. To simplify the derivation

let us assume that *Cgs* dominates the input capacitance and is itself dominated by the parallel plate capacitance. Then in saturation

$$\omega_T \approx \frac{g_m}{C_{gs}} \approx \frac{\mu_n C_{ox}(W/L)(V_{gs} - V_t)}{\frac{2}{3}WLC_{ox}} = \frac{3}{2}\frac{\mu_n(V_{gs} - V_t)}{L^2}.$$

Hence *WT* depends on the inverse square of the length and increases with increasing gate-source voltage. Remember, though that this equation holds only in the long channel regime.

1.3. OPERATION IN WEAK INVERSION (SUBTHRESHOLD):- In simple MOSFET models (such as the one we've presented so far), the device conducts no current until an inversion layer forms. However mobile carriers don't abruptly disappear the moment the gale voltage drops below *Vt.* In fact exercising a little imagination. one can discern a structure reminiscent of an n-p-n bipolar transis- tor when the device is in the subthreshold regime with the source and drain regions functioning as emitter and collector respectively, and the (non inverted) bulk behaving a bit like a base.

As *Vgs* drops below threshold the current decreases in an exponential fashion much like a bipolar transistor. Rather than dropping at the 60 mV/decade rate of such a bipolar, however the current in all real MOSFETs drops more slowly (e.g.100mV/decade) owing to the capacitive voltage division between gate and channel and between channel and source or bulk.

1.4. MOS DEVICE PHYSICS IN THE SHORT-CHANNEL REGIME:- The continuing drive to shrink device geometries has resulted in devices so small that various high-field effects become prominent at moderate voltages. The primary high-field effect is that of velocity saturation. Because of scattering by high-energy ("optical") phonons, carrier velocities eventually cease to increase with increasing electric field. As the electric field approaches about 10^6 V/m in silicon, the electron drift velocity displays a progressively weakening dependence on the field strength and eventually saturates at a value of about 10^5 *m/s.*

In deriving equations for long-channel devices, the saturation drain current is assumed to correspond to the value of current at which the channel pinches off. In short-channel devices, the current saturates when the carrier velocity does.

To accommodate velocity saturation, begin with the long-channel equation for drain current in saturation is

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{gs} - V_t)^2.$$

which may be rewritten as

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{gs} - V_t)V_{dsat}.$$

Here the long-channel *Vdsat* is denoted *Vdsatl* and is equal to *(Vgs - Vt).*

As stated earlier. the drain current saturates when the velocity does, and the velocity saturates at smaller voltages as the device gets shorter. Hence, we expect *Vdsat* to diminish with channel length.

It can be shown that *Vdsat may* be expressed more generally by the following approximation is

$$V_{dsat} \approx (V_{gs} - V_t) \parallel (LE_{sat}) = \frac{(V_{gs} - V_t)(LE_{sat})}{(V_{gs} - V_t) + (LE_{sat})}.$$

$$I_D = \frac{\mu_n C_{ox}}{2} \frac{W}{L} (V_{gs} - V_t)[(V_{gs} - V_t) \parallel (LE_{sat})].$$

where *Esat,* is the field strength at which the carrier velocity has dropped to half the value extrapolated from low-field mobility.

It should be clear from the preceding equations that the prominence of "short channel" effects depends on the ratio of (Vgs - *Vt)/ L* to *Esat* If this ratio is small, then the device still behaves as a long device; the actual channel length is irrelevant. All that happens as the device shortens is that less gate overdrive (Vgs – *Vt)* is needed to induce the onset of these effects.

Given our definition of *E*sat the drain current may be rewritten as

$$I_D = WC_{ox}(V_{gs} - V_t)v_{sat}\left[1 + \frac{LE_{sat}}{V_{gs} - V_t}\right]^{-1}.$$

A typical value for *Esat* is about 4 x 10^6 V/m. Although *Esat* is somewhat process- dependent. we will treat it as constant in all that follows.

For values of (Vgs - *Vt)/ L* that are large compared with *Esat* the drain current approaches the following limit is

$$I_D = \frac{\mu_n C_{ox}}{2} W(V_{gs} - V_t)E_{sat}.$$

That is, the drain current eventually *ceases to depend on the channel length.* Furthermore, the relationship between drain current and gate-source voltage becomes incrementally *linear* rather than square-law.

1.4.1. EFFECTOF VELOCITY SATURATION ON TRANSISTOR DYNAMICS :- In view of the first-order effect of velocity saturation on the drain current, we ought to revisit the expression for *WT* to see how device scaling affects high-frequency performance in the short-channel regime.

First, let's compute the limiting transconductance of a short-channel MOS device in saturation

$$g_m = \frac{\partial I_D}{\partial V_{gs}} = \frac{\mu_n C_{ox}}{2} W E_{sat}.$$

Using the same numbers as for the limiting saturation current, we find that the transconductance should be roughly 1mS per micron of gate width (easy numbers to remember: everything is roughly 1 something per *lim.* Note that the only practical control over this value at the disposal of a device designer is through the choice of *tox,* to adjust Cox unless a different dielectric material is used.

To simplify calculation of *WT,* assume (as before) that *Cg* dominates the input capacitance. Further assume that short-channel effects do not appreciably influence charge sharing, so that *C"* still behaves approximately as in the long-channel limit is

$$C_{gs} \approx \tfrac{2}{3} WLC_{ox}.$$

Taking the ratio of *gm* to *Cgs* then

$$\omega_T \approx \frac{g_m}{C_{gs}} \approx \frac{(\mu_n C_{ox}/2) W E_{sat}}{\frac{2}{3} WLC_{ox}} = \frac{3}{4} \frac{\mu_n E_{sat}}{L}.$$

We see that the *WT* of a short-channel device thus depends on 1/ *L* instead of on $1/L^2$ additionally, note that it does not depend on bias conditions (but keep in mind that this independence holds only in saturation) nor on oxide thickness or composition.

1.4.2. THRESHOLD REDUCTION :- We've already seen that higher drain voltages cause channel shortening, resulting in a nonzero output conductance. When the channel length is small, the electric field associated with the drain voltage may extend enough toward the source that the effective threshold diminishes. *This drain-indllced barrier lowering* (DIBL, pronounced "dibble") can cause dramatic increases in subthreshold current (keep in mind the exponential sensitivity of the subthreshold current). Additionally. it results in a degradation in output conductance beyond that associated with simple channel-length modulation.

A plot of threshold voltage as a function of channel length shows a monotonic decrease in threshold as length decreases. At the *0.5-Jum* level, the threshold reduction can be 100-200 mV over the value in the long-channel limit, corresponding to potential increases in subthreshold current by factors of 10 to 1000.To reduce the peak channel field and thereby mitigate high-field effects, a lightly doped drain (LDD) structure is almost always used in modern devices. In such a transistor, the doping in the drain region is arranged to have a spatial variation, progressing from relatively heavy near the drain contact to lighter somewhere in the channel. In some cases, the doping profile results in overcompensation in

the sense that higher drain voltages actually *increase* the threshold over some range of drain voltages before ultimately decreasing the threshold. Not all devices exhibit this *reverse short-channel effect,* since its existence depends on the detailed nature of the doping profile. Additionally, PMOS devices do not exhibit high-field effects as readily as do NMOS transistors since the field strengths necessary to cause hole velocity to saturate are considerably higher than those that cause electron velocity saturation.

1.4.3. SUBSTRATE CURRENT: - The electric field near the drain can reach extraordinarily large values with moderate voltages in short-channel devices. As a consequence, carriers can acquire enough energy between scattering events to cause impact ionization upon their next collision. Impact ionization by these "hot" carriers creates hole-electron pairs and, in an NMOS device, the holes are collected by the substrate while the electrons flow to the drain (as usual). The resulting substrate current is a sensitive function of the drain voltage, and this current represents an additional conductance term shunting the drain to ground. This effect is of greatest concern when one is seeking the minimum output conductance at high drain-source voltages.

1.4.3. GATE CURRENT :- The same hot electrons" responsible for substrate current can actually cause *gate current,* The charge comprising this gate current can become trapped in the oxide, causing upward threshold shifts in NMOS devices and threshold reductions in PMOS devices. Although this effect is useful if one is trying to make nonvolatile memories, it is most objectionable in ordinary circuits because it degrades long-term reliability,

2. DISTRIBUTED SYSTEMS:- MAXWELL AND KIRCHHOFF

$$\nabla \cdot \mu_0 \mathbf{H} = 0.$$

$$\nabla \cdot \varepsilon_0 \mathbf{E} = \rho.$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t}.$$

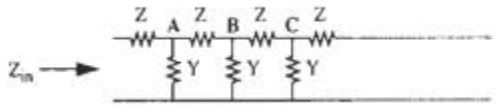
$$V = \oint \mathbf{E} \cdot d\mathbf{l} = \oint (-\nabla \phi) \cdot d\mathbf{l} = 0.$$

$$\nabla \cdot \mathbf{J} = \nabla \cdot (\nabla \times \mathbf{H}) = 0.$$

$$c = 1/\sqrt{\mu_0 \varepsilon_0}.$$

$$l \ll \lambda.$$

2.1DRIVING-POINT IMPEDANCE OF ITERATED STRUCTURES:- Specifically, consider the infinite ladder network shown in Figure 2.1. Even though resistor symbols are used here, they represent arbitrary impedances. To find the driving-point impedance of this network without summing an infinite series, note that the impedance to the right of node C is the same as that to the right of B, and the same as that to the right of A



$$Z_{in} = Z + [(1/Y) \parallel Z_{in}].$$

Which expands to

$$Z_{in} = Z + \frac{Z_{in}/Y}{1/Y + Z_{in}} \Longrightarrow (Z_{in} - Z)\left(\frac{1}{Y} + Z_{in}\right) = \frac{Z_{in}}{Y}.$$

Solving for Zin

$$Z_{in} = \frac{Z \pm \sqrt{Z^2 + 4(Z/Y)}}{2} = \frac{Z}{2}\left[1 \pm \sqrt{1 + \frac{4}{ZY}}\right].$$

Where the special case Z=1/Y=R

$$Z_{in} = \left(\frac{1+\sqrt{5}}{2}\right)R \approx 1.618R.$$

This ratio of Zin to *R* is known as the *golden ratio* (or *golden section)* .

IDEAL TRANSMISSION LINE AS INFINITE LADDER NETWORK

Let's now consider the more general case of the input impedance in the limit where *modZY<<1*. and where we continue to disallow negative values of *Zin* In that case we can simplify the result to

Zin=√Z/Y

We sec that if *Z/Y* happens to be frequency-independent. then the input impedance will also be frequency-independent." One important example of a network of this type is the model for an ideal transmission line. In the case of a lossless line. Z =*sl* and *Y* = sC, where Land C represent differential (in the mathematical sense) circuit elements. The input impedance (called the *characteristic* impedance Zo) for an ideal loss less infinite transmission line is therefore

$$Z_{in} \approx \sqrt{Z/Y} = \sqrt{sL/sC} = \sqrt{L/C}.$$

The resolution is that the energy doesn't end up as heat if the line is truly infinite. The energy just keeps traveling down the line forever and so is lost to the external world just as if it had heated up a resistor and its environs; the line acts like a black hole for energy.

3. TRANSMISSION LINES IN MORE DETAIL:-

3.1 LUMPED MODEL FOR LOSSY TRANSMISSION LINE:- To derive the relevant parameters of a lossy line consider an infinitesimally short piece of line. of length *dz;* as shown in Figure 3.1. Here the elements *L, R,* C and G are all quantities *per unit length* and simply represent a specific example of the more general case considered earlier. The inductance accounts for the energy stored in the magnetic field around the line. While the series resistance accounts for the inevitable energy loss (such as due to skin effect) that all ordinary conductors exhibit. The shunt capacitance models the energy stored in the electric field surrounding the line and the shunt conductance accounts for the loss due to mechanisms such as ordinary ohmic leakage as well as loss in the line's dielectric material.

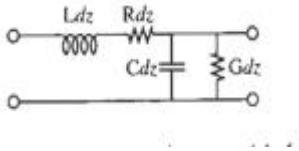


Figure3.1 Lumped *RLC* model of infinitesimal transmission-line segment

3.2. CHARACTERISTIC IMPEDANCE OF A LOSSY TRANSMISSION LINE:- To compute the impedance of a lossy line.

$$Z_0 = Z\,dz + [(1/Y\,dz) \parallel Z_0] = Z\,dz + \frac{Z_0}{1 + (Y\,dz)Z_0}.$$

We will consider the limiting behavior of this expression as *dz* approaches zero. So we may use the first-order binomial expansion of 1/(1 +.r)

$$\begin{aligned} Z_0 &= Z\,dz + \frac{Z_0}{1 + (Y\,dz)Z_0} \\ &\approx Z\,dz + Z_0[1 - (Y\,dz)Z_0] = Z_0 + dz(Z - YZ_0^2). \end{aligned}$$

Cancelling Zo from both sides. we see that the final term in parentheses must equal zero. The characteristic impedance is thus

$$Z_0 = \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}}.$$

If the resistive terms are negligible, the equation for Zo= Zin=√L/C, because the impedance approaches=√L/C at sufficiently high frequency, independent of R and G, it is sometimes called transient or pulse impedance.

4. BEHAVIOR OF FINITE-LENGTH TRANSMISSION LINES:

4.1. TRANSMISSION LINE WITH MATCHED TERMINATION:-

The driving-point impedance of an infinitely long line is simply Zo. Suppose we cut the line somewhere, discard the infinitely long remainder, and replace it with a single lumped impedance of value Zo .The driving-point impedance must remain Zo there's no way for the measurement apparatus to distinguish the lumped impedance from the line it replaces. Hence, a signal applied to the line simply travels down the finite segment of line, eventually gets to the resistor, heats it up, and contributes to global warming.

4.2. TRANSMISSION LINE WITH ARBITRARY TERMINATION:- In general, a transmission line will not be terminated in precisely its characteristic impedance. A signal traveling down the line maintains a ratio of voltage to current that is equal (of course) to Zo until it encounters the load impedance. The termination impedance imposes its own particular ratio of voltage to current. however and the only way to reconcile the conflict is for some of the signal to reflect back toward the source.

To distinguish forward (incident) quantities from the reflected ones we will use the subscripts i and *r*. respectively, If *E,* and *I,* are the incident voltage and current. then it's clear that

$$Z_0 = \frac{E_i}{I_i}.$$

At the load end of things, the mismatch in impedances gives rise to a reflected voltage and current. We still have a linear system, so the total voltage at any point on the system is the superposition of the incident and reflected voltages. Similarly the net current is also the superposition of the incident and reflected currents, because the current components travel in opposite directions, the superposition here results in a subtraction. Thus, we have

$$Z_L = \frac{E_i + E_r}{I_i - I_r}.$$

We may rewrite this last equation to show an explicit proportionality to Zo as follows

$$Z_L = \frac{E_i + E_r}{I_i - I_r} = \frac{E_i}{I_i}\left[\frac{1 + E_r/E_i}{1 - I_r/I_i}\right] = Z_0\left[\frac{1 + E_r/E_i}{1 - I_r/I_i}\right].$$

The ratio of reflected to incident quantities at load end of the ΓL line is called ΓL and will generally be complex. Using ΓL the expression for *Z.* becomes

$$Z_L = Z_0\left[\frac{1 + E_r/E_i}{1 - I_r/I_i}\right] = Z_0\left[\frac{1 + \Gamma_L}{1 - \Gamma_L}\right].$$

Solving for ΓL and is given by

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}.$$

If the load impedance equals the characteristic impedance of the line, then the reflection coefficient will be zero. If a line is terminated in either a short or an open, then the reflection coefficient will have a magnitude of unity: this value is the maximum magnitude it can have (for a purely passive system such as this one, anyway).

We may generalize the concept of the reflection coefficient so that it is the ratio of the reflected and incident quantities at any arbitrary point along the line is

$$\Gamma(z) = \frac{E_r e^{\gamma z}}{E_i e^{-\gamma z}} = \frac{E_r}{E_i} e^{2\gamma z} = \Gamma_L e^{2\gamma z}.$$

Here we follow the convention of defining z = 0 at the load end of the line and to eating the driving source at z = -*l* with this convention the voltage and current at any point z along the line may be expressed as

$$V(z) = V_i e^{-\gamma z} + V_r e^{\gamma z}.$$
$$I(z) = I_i e^{-\gamma z} - I_r e^{\gamma z}.$$

As always the impedance at any point *t.* is simply the ratio of voltage to current is

$$Z(z) = \frac{V_i e^{-\gamma z} + V_r e^{\gamma z}}{I_i e^{-\gamma z} - I_r e^{\gamma z}} = Z_0\left[\frac{1 + \Gamma_L e^{2\gamma z}}{1 - \Gamma_L e^{2\gamma z}}\right].$$

Substituting for ΓL and doing a whole heck of a lot of crunching yields

$$\frac{Z(z)}{Z_0} = \frac{\frac{Z_L}{Z_0}(e^{-\gamma z} + e^{\gamma z}) + (e^{-\gamma z} - e^{\gamma z})}{\frac{Z_L}{Z_0}(e^{-\gamma z} - e^{\gamma z}) + (e^{-\gamma z} + e^{\gamma z})}.$$

Writing this expression in a more compact form, we have

$$\frac{Z(z)}{Z_0} = \frac{\frac{Z_L}{Z_0} - \tanh \gamma z}{1 - \frac{Z_L}{Z_0}\tanh \gamma z}.$$

In the special case where the attenuation is negligible (as is commonly assumed to permit tractable analysis), a considerable simplification results is

$$\frac{Z(z)}{Z_0} = \frac{\frac{Z_L}{Z_0} - j\tan\beta z}{1 - j\frac{Z_L}{Z_0}\tan\beta z} = \frac{Z_L\cos\beta z - jZ_0\sin\beta z}{Z_0\cos\beta z - jZ_L\sin\beta z}.$$

Here, z is the actual coordinate value and will always be zero or negative. As a final comment notes that this expression is periodic. Such behavior is strictly observed only in lossless lines, of course, but practical lines will behave similarly as long as the loss is negligible. Periodicity implies that one need consider the impedance behavior only over some finite section (specifically. a half-wavelength) of line.

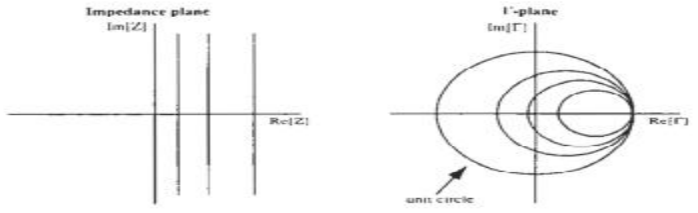
5. THE SMITH CHART-PLOTTING GAMMA:-

The expression for the reflection coefficient in terms of the normalized load impedance is

$$\Gamma = \frac{\frac{Z_L}{Z_0} - 1}{\frac{Z_L}{Z_0} + 1} = \frac{Z_{nL} - 1}{Z_{nL} + 1}. \qquad (1)$$

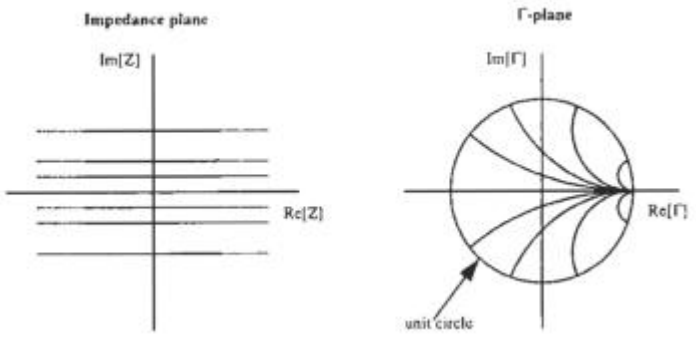
The relationship between the normalized load impedance and Γ is bi-unique: knowing one is equivalent to knowing the other. This observation is important because the familiar curves of the Smith chart are simply a plotting in the Γ-plane, of contours of constant resistance and reactance. The relationship between impedance and r given in Equation 1 may be considered a mapping of one complex number into another. In this case it is a special type of mapping known as a *bilinear transformation* because it is a ratio of two linear functions. Among the various properties of bilinear transformations a particularly relevant one is that circles remain circles when mapped. In this context a line is considered a circle of infinite radius. Hence circles and lines map into either circles or lines.

With the aid of Eqn. 1 it is straightforward to show that the imaginary axis of the Z plane maps into the unit circle in the Γ -plane, while other lines of constant resistance in the Z -plane map into circles of varying diameter that are all tangent at the point Γ =1 as shown in Figure 5.1.



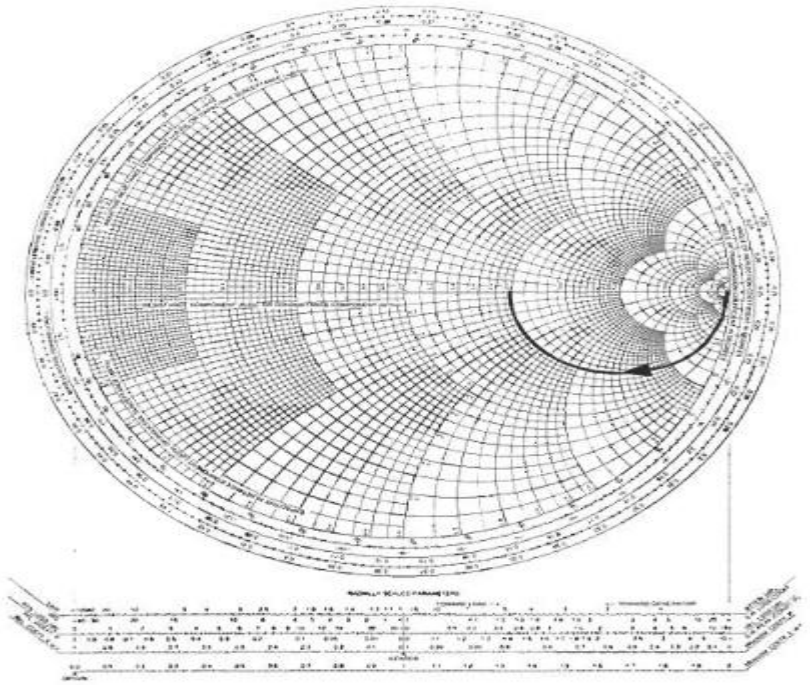
Mapping of constant-resistance lines in Z-plane to circles in Γ-plane.

Lines of constant reactance are orthogonal to lines of constant resistance in the Z-plane, and this orthogonality is preserved in the mapping. Since lines map to lines or circles we expect constant-reactance lines to transform to the circular arcs shown in Figure 5.2. The Smith chart is just the plotting of both constant-resistance and constant-reactance contours in the r-plane without the explicit presence of the Γ -plane axes.



Mapping of constant-reactance lines in Z-plane to contours in Γ-plane.

The bottom half of the Z plane maps into the bottom half of the unit circle in the r-plane, and thus capacitive impedances are always found there Similarly The top half of the Z-plane corresponds to the top half of the unit circle and inductive impedances. Progressively smaller circles of constant resistance correspond to progressively larger resistance values. The point Γ = -1 corresponds to zero resistance (or reactance), and the point Γ = 1 corresponds to infinite resistance (or reactance).



The Smith chart with series *RC* example.

it is always negative in sign the locus must be just the bottom half of that constant resistance circle. traversed clockwise from Γ = 1 as frequency increases as shown above figure.

6. BANDWIDTH ESTIMATION USING OPEN-CIRCUIT TIME CONSTANTS:-

Finding the -3-dB bandwidth of an arbitrary linear network can be a difficult problem in general, Consider, for example. the standard recipe for computing bandwidth

(1) derive the input-output transfer function (using node equations, for example);

(2) set *s = jw*

(3) find the magnitude of the resulting expression;

(4) set the magnitude = *1/√2* of the "midband" value; and

(5) solve for *w.*

The method of open-circuit time constants (OC_{TS}), also known as "zero value" time constants, was developed in the mid-1960s at MIT. As we shall see, this powerful technique allows us to estimate the bandwidth of a system almost by inspection and sometimes with surprisingly good accuracy. More important, and unlike typical circuit simulation programs, OC_{TS} *identify* w*hich elements are responsible for band- width limitations.*The great value of this property in the design of amplifiers hardly needs expression.

To begin development of this method, let us consider all-pole transfer functions only. Such a system function may be written as follows as is given by

$$\frac{V_o(s)}{V_i(s)} = \frac{a_0}{(\tau_1 s + 1)(\tau_2 s + 1)\cdots(\tau_n s + 1)},$$

where the various time constants mayor may not be real Multiplying out the terms in the denominator leads to a polynomial that we shall express as

$$b_n s^n + b_{n-1} s^{n-1} + \cdots + b_1 s + 1,$$

here the coefficient *bn* is simply the product of all the time constants and *b1* is the sum of all the time constants. (In general, the coefficient of the $\mathbf{s^j}$ term is computed by forming all unique products of the n time constants taken *j* at a time and summing all n!/j! (n - *j)* ! such products.).

We now assert that, near the - 3-dB frequency, the first-order term typically dominates over the higher-order terms so that (perhaps) to a reasonable approximation we have

$$\frac{V_o(s)}{V_i(s)} \approx \frac{a_0}{b_1 s + 1} = \frac{a_0}{(\sum_{i=1}^{n} \tau_i)s + 1}.$$

The bandwidth of our original system in radian frequency as estimated by this first- order approximation is then simply the reciprocal of the effective time constant:

$$\omega_h \approx \frac{1}{b_1} = \frac{1}{\sum_{i=1}^{n} \tau_i} = \omega_{h,est}.$$

Before proceeding further, we should consider the conditions under which our neglect of the higher-order terms is justified. Let us examine the denominator of the transfer function near our estimate of *Wh.* For the sake of simplicity, we start with a second-order polynomial with purely real roots.

Now, at our estimated -3-dB frequency, the original denominator polynomial is

$$-\tau_1\tau_2\omega_{h,est}^2 + j(\tau_1 + \tau_2)\omega_{h,est} + 1.$$

Furthermore, the bandwidth estimate is usually conservative in the sense that *the actual bandwidth will almost always be at least as high as estimated by this method.*

Fortunately, it is possible to relate the desired time-constant sum, *b1* to (more or less) easily computed network quantities. The new recipe is thus as follows. Consider an arbitrary linear network comprising only resistors, sources (dependent or independent), and m capacitors. Then:

(a) compute the effective resistance *Rj0* facing each *jth* capacitor with all of the

other capacitors removed (open-circuited, hence the method's name);

(b) form the product $\tau jw = Rj0Cj$ (the subscript *0* refers to the open-circuit condition) for each capacitor;

(c) sum all *m* such "open-circuit" time constants.

the sum of open-circuit time constants formed in step (c) is in fact precisely equal to the sum b1 of the pole time constants.

6.1. OBSERVATIONS AND INTERPRETATIONS:-

The origin of this properly of OCτs may be regarded intuitively as follows. *The reciprocal of each* jth *open-circuit time is the bandwidth that the circuit would exhibit if that jth capacitor were the only capacitor in the systems.* Thus each time constant represents a *local bandwidth degradation term.* The method of OCτs then states that the linear combination of these individual local limitations yields an estimate of the total bandwidth. The value of OCτs derives directly from the identification and approximate quantification of the local bandwidth bottlenecks.

6.2. ACCURACY OF OCτs :- The nature and magnitude of the problem are best illustrated with an example. Consider the simplest possible case. a two-pole transfer function is

$$H(s) = \left[\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega_n} + 1\right]^{-1}.$$

The $OC_{\tau s}$ bandwidth estimate is found from the coefficient of the *s* term

$$\omega_h \approx \frac{\omega_n}{2\zeta};$$

may be shown that the actual bandwidth is

$$\omega_h = \omega_n\left[1 - 2\zeta^2 + \sqrt{2 - 4\zeta^2 + 4\zeta^4}\right]^{1/2}.$$

n this particular case, we see that the OC_τ estimate predicts monotonically increasing bandwidth as the damping ratio zeta will approaches zero, while the actual bandwidth asymptotically approaches about *1.55wn.* Thus it is possible for OC_τ estimates to be *optimistic* - in this case, wildly so. At a zeta of about 0.35, OC_τ estimates are correct; for any higher damping ratio. $OC_{\tau s}$ are pessimistic. Fortunately, the poles of amplifiers are usually designed to have relatively high damping ratios (to control overshoot and ringing in the step response, and to minimize peaking in thc frequency response) and so. for most practical situations, OC_τ estimates are pessimistic.

6.3. USEFUL FORMULAS :-

(a) Common-Source Input Resistance:- At frequencies where we may neglect quasistatic effects, the input resistance is effectively infinite (not a function of *gmb* or CLM /DIBL). In this era of super thin gate oxides leakage is more than measurable and the input resistance is consequently not truly infinite. However we will pretend that for most purposes it still remains high enough that infinity is a useful approximation.

(b) Common-Source Output Resistance :- Use a test current source, it. Since *vgs= vbs.* we can simply sum the two transconductances into a single one (let's call it *gm out).* Compute the resistance looking into the drain. and call it "rout". The net resistance at the drain will therefore be "rout" in parallel with any external resistance, *RL.* Thus we have

$$v_{gs} = -i_t R_S,$$

$$v_{r0} = (i_t - g_{m,\text{tot}} v_{gs}) r_0 = i_t(1 + g_{m,\text{tot}} R_S) r_0,$$

$$V_{\text{test}} = i_t R_S + i_t(1 + g_{m,\text{tot}} R_S) r_0 = i_t[R_S + r_0 + g_{m,\text{tot}} r_0 R_S].$$

So, at last, the resistance looking into the drain (and so *not* including R_L) is

$$r_{\text{out}} = R_S + r_0 + g_{m,\text{tot}} r_0 R_S = R_S + r_0 + (g_m + g_{mb}) r_0 R_S.$$

(c) common-Source Voltage Gain: Multiplication of the total resistance at the drain by the effective transconductance gives us the gain. To find the effective transconductance, short the drain to ground (incrementally speaking) and then measure the ratio of short-circuit output current to input voltage is

$$v_{gs} = v_{in} - i_{out} R_S,$$
$$v_{bs} = -i_{out} R_S;$$
$$i_{out} = g_m v_{gs} + g_{mb} v_{bs} - i_{out} \frac{R_S}{r_0}$$
$$= g_m(v_{in} - i_{out} R_S) + g_{mb}(-i_{out} R_S) - i_{out} \frac{R_S}{r_0};$$
$$g_{m,eff} = \frac{i_{out}}{v_{in}} = \frac{g_m}{1 + (g_m + g_{mb}) R_S + R_S/r_0} = \frac{g_m}{1 + (g_m + g_{mb} + g_0) R_S},$$

where we have used $g_0 = 1/r_0$.

The voltage gain is therefore

$$A_v = -g_{m,eff}(r_{out} \parallel R_L)$$
$$= -\frac{g_m}{1 + (g_m + g_{mb} + g_0) R_S} ([R_S + r_0 + (g_m + g_{mb}) r_0 R_S] \parallel R_L)$$

which simplifies initially to

$$A_v = -\frac{g_m R_L}{1 + (g_m + g_{mb} + g_0) R_S} \cdot \frac{R_S + r_0 + (g_m + g_{mb}) r_0 R_S}{R_S + r_0 + (g_m + g_{mb}) r_0 R_S + R_L}.$$

After a little more work, the equation's complexity reduces further, leading to reasonable expression:

$$A_v = \frac{g_m R_L}{1 + (g_m + g_{mb}) R_S + g_0(R_L + R_S)}.$$

(d) Common-Gate Input Resistance :- Use a test voltage source. Note that once again *vgs=vbs* and so we can merge the two transconductances into a single *gm, tot* Compute the resistance looking into the source first; then accommodate *Rs* at the end if desired or necessary. You may use superposition to simplify the derivation (treat transconductances as independent current sources in this computation, since the control voltage is fixed *throughout this particular experiment).*

We start with

Vgs=-vtest

With transconductance *gm,tot* disabled, we compute one contribution to the test current

$$i_{t1} = \frac{v_{test}}{r_0 + R_L}.$$

Next, disable (short out) the test voltage source, taking care not to zero out *gm,tot* and compute the other contributions to the test current. The current from the transconductance is added to the current flowing through *ro* ,Then

$$i_{t2} = -g_{m,tot} v_{gs} + \left(-i_{t2} \frac{R_L}{r_0}\right) = g_{m,tot} v_{test} - i_{t2} \frac{R_L}{r_0}$$

$$\Longrightarrow \quad i_{t2} = \frac{g_{m,tot} v_{test}}{1 + R_L/r_0} = \frac{g_{m,tot} r_0 v_{test}}{r_0 + R_L};$$

$$i_t = i_{t1} + i_{t2} = \frac{v_{test}}{r_0 + R_L} + \frac{g_{m,tot} r_0 v_{test}}{r_0 + R_L} = \frac{v_{test}}{r_0 + R_L}(1 + g_{m,tot} r_0).$$

Hence, the resistance looking into the source is

$$r_{in} = \frac{v_{test}}{i_t} = \frac{r_0 + R_L}{1 + g_{m,tot} r_0} = \frac{r_0 + R_L}{1 + (g_m + g_{mb}) r_0}.$$

Note that, if *RL* <<r0 and if the transistor's intrinsic voltage gain *gmro>>1*. then the resistance looking into the source is approximately

$$r_{in} \approx \frac{1}{g_m + g_{mb}}.$$

If you desire the total resistance between the source node and ground, merely compure the parallel combination of rin and *Rs.*

(e) Common-Gate Output Resistance :- This resistance is precisely the same as the output resistance of the degenerated common-source amplifier is

$$r_{out} = R_S + r_0 + g_{m,tot} r_0 R_S = R_S + r_0 + (g_m + g_{mb}) r_0 R_S.$$

If we drive the source terminal directly with a voltage source then *Rs* = 0 and the equation for the resistance looking into the drain collapses to simply ro. That remains a reasonably good approximation if *Rs* is small compared with *ro* and if *(gm+gmb)Rs* is small compared with unity.

(f) Common-Gate Voltage Gain:- As before, let's first compute the effective transconductance of the amplifier, defined here as the ratio of short-circuit drain current to source voltage is

$$v_{gs} = -v_{in},$$

$$v_{bs} = v_{gs};$$

$$i_{out} = -g_{m,tot} v_{gs} + \frac{V_{in}}{r_0} = v_{in}(g_{m,tot} + g_0)$$

$$\Longrightarrow \quad g_{m,eff} = g_{m,tot} + g_0 = (g_m + g_{mb} + g_0).$$

The voltage gain from source to drain is therefore

$$A_{v0} = g_{m,eff}(r_{out} \parallel R_L) = (g_m + g_{mb} + g_0)(r_0 \parallel R_L).$$

where we have used the fact that, in this first computation, we are driving the source terminal directly with a voltage source. Consequently, the total drain resistance is just ro parallel with RL, so the gain in above Eqn. is that from the source to drain, as stated.

The voltage divider formed at the input must be taken into account in order to complete our calculation of the overall gain is

$$A_v = A_{v0} \cdot \frac{r_{in}}{r_{in} + R_S},$$

$$A_v = (g_m + g_{mb} + g_0)(r_0 \parallel R_L) \cdot \frac{\dfrac{r_0 + R_L}{1 + (g_m + g_{mb})r_0}}{\dfrac{r_0 + R_L}{1 + (g_m + g_{mb})r_0} + R_S},$$

which simplifies a bit to

$$A_v = (g_m + g_{mb} + g_0)(r_0 \parallel R_L) \cdot \frac{r_0 + R_L}{r_0 + R_L + R_S + (g_m + g_{mb})r_0 R_S}$$

and still further to

$$A_v = \frac{(g_m + g_{mb} + g_0)R_L}{1 + g_0(R_L + R_S) + (g_m + g_{mb})R_S}.$$

As expected, the gain expression collapses to more familiar forms when body effect is neglected, thc source resistance *Rs* is zero, the transistor's output conductance is zero.

(g) Source-Follower Resistance :- Again. we may treat the source follower's *input* resistance as infinite if we neglect gate leakage and nonquasistatic effects. This result is independent of body effect and CLM/DIBL.

The source follower's *output* resistance is the same as the input resistance of a common-gate stage is

$$r_{out} = \frac{r_0 + R_L}{1 + g_{m,tot} r_0} = \frac{r_0 + R_L}{1 + (g_m + g_{mb}) r_0}.$$

In most source followers *RL* is chosen very small compared with *r0.* In such cases, the output resistance simplifies to

$$r_{out} \approx \frac{r_0}{1 + (g_m + g_{mb}) r_0}.$$

If, in addition, the transistor's intrinsic voltage gain is high, then the unity factor in the denominator can be neglected, simplifying the output resistance expression even further is

$$r_{out} \approx \frac{1}{g_m + g_{mb}}.$$

(h) Source-Follower Voltage Gain:-Again, compute an effective transconductance and then multiply by the output resistance. That output resistance needs to include any external loading that might appear in parallel with the resistance seen when looking into the transistor's source. Now

$$v_{gs} = v_{in},$$
$$v_{bs} = v_{gs}$$

For output resistance only; *vbs*=0 only for effective transconductance calculations,therefore

$$i_{out} = g_m v_{in} - i_{out} \frac{R_L}{r_0},$$
$$g_{m,eff} = \frac{i_{out}}{v_{in}} = \frac{g_m}{1 + R_L / r_0}.$$

Overall voltage gain is thus

$$A_{v0} = g_{m,\text{eff}}(r_{\text{out}} \parallel R_S) = \frac{g_m}{1 + R_L/r_0} \cdot \left[\left(\frac{r_0 + R_L}{1 + (g_m + g_{mb})r_0} \right) \parallel R_S \right].$$

which, after some effort, simplifies to

$$A_{v0} = \frac{g_m r_0 R_S}{(g_m + g_{mb})r_0 R_S + (r_0 + R_L + R_S)}.$$

Alternatively, we may write

$$A_{v0} = \frac{g_m R_S}{(g_m + g_{mb})R_S + 1 + g_0(R_L + R_S)} = \frac{1}{1 + \dfrac{g_{mb} + g_0}{g_m} + \dfrac{1 + g_0 R_L}{g_m R_S}}.$$

from which it is perhaps somewhat easier to see that the voltage gain can only *approach* unity. as we would expect from a source follower. As usual the reader should verify that the exact equation simplifies to the numerous known approximations in the case where body effect is zero. transistor *go* may be neglected. *RL* is zero. and so forth,

(i) Resistances for OC Time Constant Calculations:- Using a current source in a hybrid model, we find that the effective resistance may be stated in the simple. universal. mnemonic form we've already cited

$$r_{eq} = r_{\text{left}} + r_{\text{right}} + g_{m,\text{eff}} r_{\text{left}} r_{\text{right}},$$

where *rleft* is the resistance seen between the capacitor's left terminal and ground, *rrght* is that seen between the right terminal and ground, and *gm,eff* is the effective transconductance - defined as the ratio of short-circuit output current to input voltage, This equation may also be expressed as

$$r_{eq} = r_{\text{left}} + r_{\text{right}} - A_{vf} r_{\text{left}},$$

where we have recognized the product of effective transconductance and *rright* as(minus) the voltage gain between the two terminals of the capacitor.

(j) Resistance facing *cgd* : For the drain-gate capacitance the effective resistance is

$$R_G + (r_{out} \parallel R_L) + \frac{g_m}{1 + (g_m + g_{mb} + g_0)R_S}(r_{out} \parallel R_L)R_G,$$

$$r_{out} \parallel R_L = [R_S + r_0 + (g_m + g_{mb})r_0 R_S] \parallel R_L.$$

The expression for the effective resistance simplifies considerably in the case of a common-gate connection. Where *Rc* is often zero (or very small).

n a source-follower connection, where *RL* is typically zero (or very small), an even greater simplification results, In that case, *RG* is perhaps a good approximation to the resistance facing *Cgd.* In other words, the expression for this resistance is somewhat complicated only for the case of a common-source amplifier.

(j) Resistance facing *cgs* :- Here, the same basic two-port model applies. but now the model parameters must be chosen to reflect the fact that the capacitor in question connects between the gate and source terminals. That is the amplifier under consideration has its input at the gate and provides an output at the source. If we use *Rc* to denote the resistance to the left, then resistance to the right is simply the total output resistance of a source follower (i.e., the resistance looking into the source terminal in parallel with any external *Rs).* Similarly, the effective transconductance is also merely that of a source follower. Thus we have actually derived all of the pieces already. We just need to put them together

$$r_{eq} = R_G + \left[\left(\frac{r_0 + R_L}{1 + (g_m + g_{mb})r_0}\right) \parallel R_S\right] - \left[\frac{g_m}{1 + R_L/r_0}\right]R_G\left[\left(\frac{r_0 + r_L}{1 + (g_m + g_{mb})r_0}\right) \parallel R_S\right],$$

which we try to simplify as follows:

$$R_{eq} = R_G + \left[\frac{R_S}{r_0 + R_L + R_S + (g_m + g_{mb})r_0 R_S}\right][(r_0 + R_L) - g_m r_0 R_G].$$

After some crunching, we finally get something that does look simpler:

$$r_{eq} = \frac{R_G(r_0 + R_L + R_S + g_{mb}r_0 R_S) + R_S(r_0 + R_L)}{r_0 + R_L + R_S + (g_m + g_{mb})r_0 R_S}.$$

Collecting terms in a slightly different way may also prove useful:

$$r_{eq} = \frac{R_G + R_S + g_{mb}R_G R_S + g_0(R_G R_L + R_G R_S + R_S R_L)}{1 + (g_m + g_{mb})R_S + g_0(R_L + R_S)}.$$

Note that, in the limit of no body effect and infinite *ro* the equivalent resistance facing *cgs* indeed simplifies to a result we've seen before is

$$r_{eq} \approx \frac{R_G + R_S}{1 + g_m R_S}.$$

(k) Resistance facing *cdb:-* In UIC general case. this is just the resistance looking into the drain of a transistor in parallel with any additional resistance connected to that drain (call that extra resistance RL to be consistent with our notation so far). We've already found the resistance looking into the drain is

$$r_{out} = R_S + r_0 + g_{m,tot} r_0 R_S = R_S + r_0 + (g_m + g_{mb}) r_0 R_S.$$

So, just place that value in parallel with any *RL* that happens to be present.

(l) Resistance facing *csb:-* Similarly, this resistance is JUS! that looking into the source of a transistor in parallel with any other resistance connected to it (we've been calling it *Rs).* Thus, the source-follower output resistance equation is what we need here is

$$r_{out} = \frac{r_0 + R_L}{1 + (g_m + g_{mb}) r_0}.$$

Just compute the parallel combination of *rout* and *Rs* to get the final answer.

7. BANDWIDTH ESTIMATION USING SHORT-CIRCUITS TIME CONSTANTS:-

In the case of short-circuit time constants. we instead assume that all of the zeros are at the *origin* and that there are as many poles as zeros. Thus, the corresponding system function may be written as follows

$$\frac{V_o(s)}{V_i(s)} = \frac{ks^n}{(s + s_1)(s + s_2) \cdots (s + s_n)}.$$

where the various pole frequencies mayor may not be real and where *k* is simply a constant to fix up the scale factor, Multiplying out the terms in the denominator leads to a polynomial that we shall

express as

$$s^n + b_1 s^{n-1} + \cdots + b_{n-1} s + b_n,$$

here the coefficient hi is the sum of all of the pole frequencies. and *bn* is the product of all of the pole frequencies. (In general. the coefficient of the *s!* term is computed by forming all unique products of the n frequencies taken *j* at a time and summing alln!/j!(n - *j)!* such products.)

We now assert that. near the low-frequency -3-dB breakpoint. the higher-order terms dominate the denominator. so that we obtain

$$\frac{V_o(s)}{V_i(s)} \approx \frac{ks^n}{s^n + b_i s^{n-1}} = \frac{ks}{s + \sum_{i=1}^{n} s_i}.$$

he low-frequency -3-dB point of our original system in radian frequency as estimated by this first-order approximation is then simply the sum of the pole frequencies is

$$\omega_l \approx b_1 = \sum_{i=1}^{n} s_i = \omega_{l,\text{est}}.$$

Before proceeding further, we should consider the conditions under which our neglect of the lower-order terms is justified. Let us examine the denominator of the transfer function near our estimate of *Wl.* For the sake of simplicity we consider a second-order polynomial with purely real roots, S1 and *S2.*

Now, at our estimated -3-dB frequency, the original denominator polynomial is

$$-\omega_{l,\text{est}}^2 + j\omega_{l,\text{est}}(s_1 + s_2) + s_1 s_2.$$

Substituting our expression for the estimated - 3-dB point, we obtain

$$-[s_1^2 + s_2^2 + 2s_1 s_2] + j[s_1^2 + s_2^2 + 2s_1 s_2] + s_1 s_2.$$

Furthermore. the low-frequency cutoff estimate is conservative in the sense that *the actual cutoff frequency will almost always be as low as or lower than estimated by this method.*

Consider an arbitrary linear network comprising only resistors. sources (dependent or independent), and m capacitors. Then

(a) Compute the effective resistance *Rjs* facing each jth capacitor with all of the other capacitors short-circuited (the subscript *s* refers to the short-circuit condition for each capacitor);

(b) Compute the "short-circuit frequency *1/(RjsCj);*

(c) Sum all m such short-circuits frequencies.

The sum of the reciprocal short-circuits time constants formed in step (c) turns out to be precisely equal to the sum *b1* of the pole frequencies. Thus at last, we have

$$\omega_{1,\text{est}} = \sum_{j=1}^{m} \frac{1}{R_{js} C_j}.$$

7.1. ACCURACY OF SC$_{Ts}$

s with OC $_{Ts}$ truncation of the denominator polynomial means that one must be careful not to place too much faith in the ability of SC $_{Ts}$ to provide accurate estimates of *Wl* in all cases. This caveat notwithstanding. it should be clear that an SC $_{Ts}$ estimate is in fact exact for a first-order network since *no* truncation of terms is involved there. Not surprisingly then the SC $_{Ts}$ estimate will be quite accurate if a network of higher order happens to be dominated by one pole (here. that means that one pole is much *higher* in frequency than all the other poles).

8. RISETIME, DELAY, AND BANDWIDTH

8.1. DELAY OF SYSTEMS IN CASCADE:- Defining delay in terms of *moments of the impulse response.* As seen in Figure 8.1 one delay measure is the time it takes for the impulse response to reach its "center of mass," that is, the normalized value of its first moment

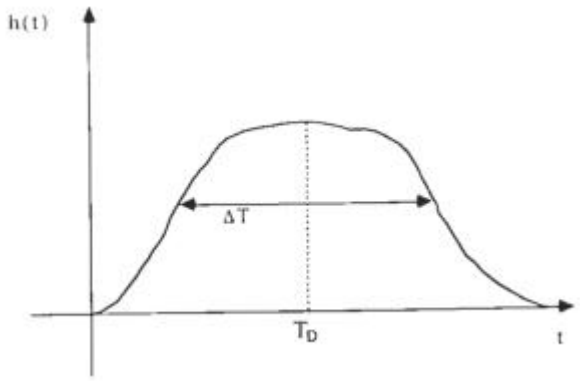


Figure8.1 impulse response

$$T_D \equiv \frac{\int_{-\infty}^{\infty} th(t)\,dt}{\int_{-\infty}^{\infty} h(t)\,dt}.$$

This quantity is also known as the Elmore delay in some of the literature, after the fellow who first used this moment-based approach. This particular measure of time delay derives much of its utility from the fact that it is readily related to a number of Fourier transform identities allowing us to exploit the full power of linear system theory. Specifically the first moment is

$$\int_{-\infty}^{\infty} th(t)\,dt = -\frac{1}{j2\pi}\frac{d}{df}H(f)\bigg|_{f=0}.$$

The normalization factor is simply the DC gain

$$\int_{-\infty}^{\infty} h(t)\,dt = H(0).$$

So that

$$T_D \equiv \frac{\int_{-\infty}^{\infty} th(t)\,dt}{\int_{-\infty}^{\infty} h(t)\,dt} = -\frac{1}{j2\pi H(0)}\frac{d}{df}H(f)\bigg|_{f=0}.$$

Using this definition. one finds that the Elmore delay for a single-pole low-pass system is just the pole time constant τ.

Now that Monsieur Fourier has graciously helped us out by providing a definition of time delay purely in terms of his transforms, the derivation becomes straightforward. Specifically consider two systems with impulse responses *h* 1(t) and *h* 2(t).

with corresponding Fourier transforms *HI(f)* and *H2(f)·* From basic linear system theory, we know that the Fourier transform of these two systems in cascade is just the product of the individual transforms. so that Htot = *H1.H2* The overall time delay is therefore

$$T_{D,\text{tot}} = -\frac{1}{j2\pi H_1(0)H_2(0)}\frac{d}{df}H_1H_2\Big|_{f=0}.$$

which we may expand to obtain

$$T_{D,\text{tot}} = -\frac{1}{j2\pi H_1(0)H_2(0)}\left[H_2(0)\frac{dH_1}{df}\Big|_{f=0} + H_1(0)\frac{dH_2}{df}\Big|_{f=0}\right].$$

From this we immediately (well okay, maybe not quite *immediately*) note that

$$T_{D,\text{tot}} = T_{D1} + T_{D2}.$$

which was to be shown. We see that use of this particular definition of lime delay has led us to the intuitively satisfying result that the overall delay of a cascade of systems is simply the sum of the individual delays.

8.1. RISETIME OF SYSTEMS IN CASCADE:-

Just as we employed the first moment of the impulse response in defining the time delay, we find the second moment useful in defining the risetirne. Referring again to Figure 8.1, note that the quantity *deltaT* is a measure of the duration of the impulse response and hence also a measure of the risetime of the step response (since the step response is the integral of the impulse response). Specifically, delta*T* is twice the "radius of gyration" about the "center of mass" *(TD)* of *h* (t).

$$\left(\frac{\Delta T}{2}\right)^2 \equiv \left[\frac{\int_{-\infty}^{\infty} t^2 h(t)\,dt}{\int_{-\infty}^{\infty} h(t)\,dt} - (T_D)^2\right].$$

Again this definition allows the use of Fourier transform identities. In particular,

$$\int_{-\infty}^{\infty} t^2 h(t)\, dt = -\frac{1}{(2\pi)^2}\frac{d^2}{df^2}H(f)\bigg|_{f=0}.$$

So that

$$t_{\text{rise}}^2 = (\Delta T)^2$$
$$= 4\left[\frac{-\frac{1}{(2\pi)^2}\frac{d^2}{df^2}H(f)\Big|_{f=0}}{H(0)} - \left(-\frac{1}{j2\pi H(0)}\frac{d}{df}H(f)\bigg|_{f=0}\right)^2\right].$$

Here we have made use of the equation for delay,

$$t_{\text{rise}}^2 = \frac{4}{(2\pi)^2 H(0)}\left[-\frac{d^2}{df^2}H(f)\bigg|_{f=0} - \frac{1}{H(0)}\left(\frac{d}{df}H(f)\bigg|_{f=0}\right)^2\right].$$

The Elmore risetime for a single-pole low-pass system is 2τ. we now consider two systems. each with its own risetime. Then

$$t_{\text{rise,tot}}^2 = \frac{4}{(2\pi)^2 H_1(0) H_2(0)}\left[-\frac{d^2}{df^2}H_1 H_2\bigg|_{f=0} - \frac{1}{H_1(0)H_2(0)}\left(\frac{d}{df}H_1 H_2\bigg|_{f=0}\right)^2\right].$$

After a small algebraic miracle. this leads to the desired result at last is

$$t_{\text{rise,tot}}^2 = t_{\text{rise1}}^2 + t_{\text{rise2}}^2.$$

Thus we see that the *squares* of the individual risetimes add linearly to yield the square of the overall risetime. Stated alternatively, the individual risetirnes add in root-sum-squared (rss) fashion to yield the overall risetime is

$$t_{\text{rise,tot}} = \sqrt{t_{\text{rise1}}^2 + t_{\text{rise2}}^2}.$$

If the individual systems satisfy this requirement then the relationships derived here will hold well. The greater the departure from the step-response monotonicity condition the less appropriate the use of these formulas.

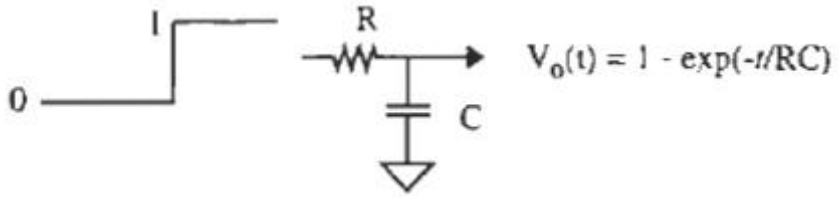
8.2. BANDWIDTH-RISETIME RELATIONS:-

The simple *RC* low-pass filter displayed in Figure 8.2. Given the equation for the response to a unit voltagestep. it is straightforward to compute the 10-90% rise time is

$$t_{\text{rise}} = RC \ln\left(\frac{0.9}{0.1}\right) \approx 2.2RC.$$

$$\omega_{-3\,\text{dB}} t_{\text{rise}} \approx 2.2,$$

here W-3dB is the -3-d8 bandwidth in radians per second and lri'" is the 10-90% risetime in response to a step.



Note that this value is about 10% higher than the Elmore risetirne computed earlier. In addition to the risetime, we already know that the -3-d8 bandwidth (in radians per second) is simply *IIRC.* Hence, the bandwidth-risetirne product is in fact about 2.2, as the rule states.

Now we can Consider for example the step response of a two-pole system

$$V_o(t) = 1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \sin\left(\sqrt{1-\zeta^2}\,\omega_n t + \Phi\right).$$

$$\Phi = \tan^{-1}\left[\frac{\sqrt{1-\zeta^2}}{\zeta}\right].$$

The -3-dB bandwidth of this system is given by

$$\omega_h = \omega_n\left(1 - 2\zeta^2 + \sqrt{2 - 4\zeta^2 + 4\zeta^4}\right)^{1/2}.$$

Let's use these formulas to explore what happens as we change zeta value. for the extreme case of a damping ratio of zero the risetirne and bandwidth are

$$t_r\big|_{\zeta=0} = \frac{1}{\omega_n}[\sin^{-1} 0.9 - \sin^{-1} 0.1] \approx \frac{1.02}{\omega_n},$$

$$\omega_h\big|_{\zeta=0} \approx 1.55\omega_n.$$

so that the corresponding bandwidth-risetirne product is

$$\omega_h t_r\big|_{\zeta=0} \approx 1.6,$$

or about *72%* of the value obtained for the first-order case.For a reasonably well-damped system. we might expect closer agreement with the first-order result. As a specific example. if we set zeta= 1/ √*2* then the risetirne and bandwidth are

$$t_r\big|_{\zeta=1/\sqrt{2}} \approx \frac{2.14}{\omega_n}$$

$$\omega_h\big|_{\zeta=1/\sqrt{2}} = \omega_n.$$

$$\omega_h t_r\big|_{\zeta=1/\sqrt{2}} \approx 2.14,$$

or a value within a small percentage of the first-order result. Note that the product of bandwidth and Elmore risetime is 2.0 for a single-pole system.

In general. the bandwidth-risetime product will range bel ween 2 and 2,2 if the system is well damped (or more precisely. if the impulse response is unipolar so that the step response is monotonic. for the same reasons that prevailed in the moment-based expressions for risetime and time delay); the product will decrease if the system is not very well damped. However even in the case of no damping at all. we have seen that the bandwidth-risetime product still does not deviate that much. Because

most systems of practical interest are generally well damped. we can expect their bandwidth-risetime product to be about 2.2. Therefore measurement of the step response risetirne is often an expedient way to obtain a reasonably accurate estimate of the bandwidth: only one experiment has to be performed and step excitations are often easier to generate than sine waves.

8.3. OPEN CIRCUIT TIME CONSTANTS, RISETIMEADDITION, AND BANDWIDTH SHRINKAGE:-

Consider a cascade of *N* identical amplifiers. each of which is single-pole with a time constant r. Combining the risetime addition rule with the bandwidth-risetirne relationship yields

$$\mathrm{BW} \approx \frac{1}{\sqrt{\sum_1^N \tau^2}} = \frac{1}{\tau\sqrt{N}}.$$

Compare that approximate result with the more exact (but still approximate) relationship

$$\mathrm{BW} \approx \frac{\sqrt{\ln 2}}{\tau\sqrt{N}} \approx \frac{0.833}{\tau\sqrt{N}}$$

Note that the method of open-circuit lime constants would predict quite a different result. Since the effective time constant is found there by summing all the individual time constants, the OCr-estimaled bandwidth would be

$$\mathrm{BW} \approx \frac{1}{\tau N}.$$

The difference is significant. and underscores yet again how the use of open-circuit time constants can lead to extremely pessimistic estimates of bandwidth if a single pole does not dominate the transfer function.

9. ZEROS TO ENHANCE BANDWIDTH:-

9. 1.THE SHUNT-PEAKED AMPLIFIER:- A technique that satisfied this requirement of large bandwidth at low cost is known as *shunt peaking, a* shunt-peaked amplifier is sketched in Figure 9, 1. This amplifier is a standard common-source configuration with the addition of the inductance. If we

assume that the transistor is ideal, then the only elements that control the bandwidth are R, L, and C. The capacitance C may be taken to represent all the loading on the output node, including that of a subsequent stage (perhaps arising from the input capacitance of another transistor, for example). The resistance R is the effective load resistance at that node and the inductor provides the bandwidth enhancement, as we now proceed to demonstrate.

Given our assumptions, we may model the amplifier for small signals as shown in Figure 9.2, it's clear from the model that the transfer function Vout/Iin is just the impedance of the RLC network, so it should be straightforward to analyze. Before launching into a detailed derivation, though, let's think about why adding an inductor this way should give us a bandwidth extension.

First, we know that the gain of a purely resistively loaded common-source amplifier is proportional to *g m RL*. We also know that when a capacitive load is added, the gain eventually falls off as frequency increases because the capacitor's impedance diminishes, The addition of an inductance in series with the load resistor provides an impedance component that increases with frequency (i.e., it introduces a zero), which helps offset the decreasing impedance of the capacitance leaving a net impedance that remains roughly constant over a broader frequency range than that of the original *RC* network.

An equivalent time-domain interpretation may be provided by considering the step response. The inductor delays current *Row* through the branch containing the resistor, making more current available for charging the capacitor, reducing the risetime.

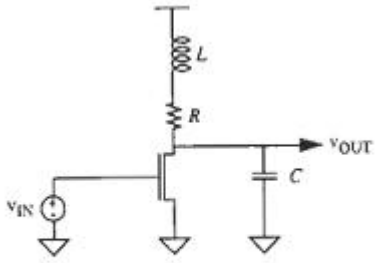


FIGURE 9.1. Shunt-peaked amplifier.

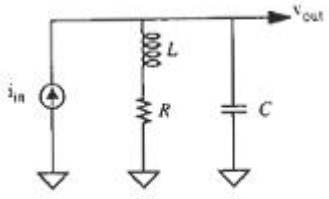


FIGURE 9.2. Model of shunt-peaked amplifier.

To the extent that a faster risetirne implies a greater bandwidth an appropriate choice of inductor therefore increases the bandwidth. Formally the impedance of the *R LC* network may be written as

$$Z(s) = (sL + R) \parallel \frac{1}{sC} = \frac{R[s(L/R) + 1]}{s^2LC + sRC + 1}.$$

In addition to a zero there are two poles (possibly complex). definitely violating the conditions for open-circuit time constants, Since the gain of the amplifier is the product of *g,"* and the magnitude of *Z(5).*let's now compute the latter as a function of frequency is

$$|Z(j\omega)| = R\sqrt{\frac{(\omega L/R)^2 + 1}{(1 - \omega^2 LC)^2 + (\omega RC)^2}}.$$

with the simple *RC* case there is a term in the numerator (from the zero) that increases with increasing frequency, Furthermore. the 1- w^2LC term in the denominator contributes to an increase in MOD Z for frequencies below the *LC* resonance as well. Both of these terms extend bandwidth.

To facilitate subsequent derivations we introduce a factor m defined the ratio of the *RC* and *L/R* time constants is

$$m = \frac{RC}{L/R}.$$

Then our transfer function becomes

$$Z(s) = (sL + R) \parallel \frac{1}{sC} = \frac{R(\tau s + 1)}{s^2\tau^2 m + s\tau m + 1}.$$

Where time constant =L/R.

The magnitude of the impedance normalized to the DC value (= *R)* as a function of frequency is then

$$\frac{|Z(j\omega)|}{R} = \sqrt{\frac{(\omega\tau)^2 + 1}{(1 - \omega^2\tau^2 m)^2 + (\omega\tau m)^2}},$$

$$\frac{\omega}{\omega_1} = \sqrt{\left(-\frac{m^2}{2} + m + 1\right) + \sqrt{\left(-\frac{m^2}{2} + m + 1\right)^2 + m^2}},$$

where W1 is the uncompensated -3-dB frequency (= *1/RC).*

the problem then is to choose a value of m that leads to some desired behavior. Maximizing the bandwidth is one obvious possibility. After a certain amount of effort one finds that this maximum occurs at a value of

$$m = \sqrt{2} \approx 1.41,$$

which extends the bandwidth to a value about 1.85 times as large as the uncompensated bandwidth, Anyone who has labored to meet a tough bandwidth specification can well appreciate the value of nearly doubling bandwidth through the addition of a single inductance at no increase in power.

$$\omega = \omega_1\sqrt{1 + \sqrt{5}} \approx 1.8\omega_1.$$

Hence the bandwidth in this case is still quite close to the maximum, Further calculation shows that the peaking is substantially reduced to about 3%.

Carrying out this method manually is frequently labor-intensive, but in this particular example, a straightforward calculation reveals that the magic value of m is

$$m = 1 + \sqrt{2} \approx 2.41,$$

which leads to a bandwidth that is about 1.72 times as large as the unpeaked case .Hence, at least for the shunt-peaked amplifier, both a maximally flat response and a substantial bandwidth extension can be obtained simultaneously.

10. THE SHUNT-SERIES AMPLIFIER:- One particularly useful broadband circuit that employs negative feedback is the shunt-series amplifier. Its name derives from the use of a combination of shunt and series feedback and its utility derives from the relative constancy of input and output impedances over a broad frequency range (which makes cascading much less complicated) as well as from its normally associated with negative feedback – namely a reduced dependency on device parameters improved distortion broader bandwidth and a nicer complexion.

the shunt-series amplifier is depicted in Figure 9.1 where *Rs* now denotes the resistance of the input source and *RL* is the load resistance. Thus the amplifier core consists of just *RF*. R1 and the transistor. To understand how this amplifier works initially assume that *RI* is large enough (relative to the reciprocal of the transistor's *gill)* that it degenerates the overall transconductance to approximately

I/R1• Since *RI* is in series with the input and output loops. the degeneration by RI is the "series" contribution to the name of this amplifier.

To continue the analysis assumes also that *RF* is large enough that its loading on the output node may be neglected. With these assumptions the voltage gain of the amplifier from the gate to the drain is approximately - *R L/ R*1·

Although we have assumed that *RF* has but a minor effect on gain. it has a controlling influence on the input and output resistance, Specifically. it reduces both quantities through the (shunt) feedback it provides. Additionally the reduction of input and output resistances helps to increase the bandwidth still further by reducing the open-circuit time-constant sum.

To compute the input resistance Rin we use the fact that the gain from gate to drain is approximately *–RL/R1* If as seems reasonable we may neglect gate current then the input resistance is due entirely to current flowing through *RF*· Applying a test voltage source at the gate terminal allows us to compute the effective resistance in the usual way. Just as in the classic Miller effect connecting impedance across two nodes that have an inverting gain between them results in a reduction of impedance. Formally Rin is given by

$$R_{in} = \frac{R_F}{1 - A_V} \approx \frac{R_F}{1 + R_L/R_1}.$$

where *Av* is the voltage gain from gate to drain

Now to compute the output resistance apply a test voltage source to the drain node and again take the ratio of ***v test to I test***

$$R_{out} = \frac{R_F + R_S}{1 + R_S/R_1} \approx \frac{R_F}{1 + R_S/R_1}.$$

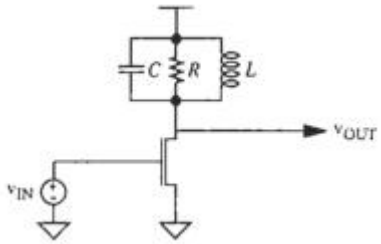
If the source and load resistances are equal (a particularly common situation in discrete realizations). then the denominators of Eqn. 18 and Eqn. 19 are approximately equal. Since the numerators are also approximately equal it follows that Rin and Rout, arc themselves nearly equal. If *Rs = RL = R* then we may write is

$$R_{out} \approx R_{in} \approx \frac{R_F}{1 + R/R_1} \approx \frac{R_F}{1 - A_V}.$$

11. TUNED AMPLIFIERS:-

11.1:- COMMON-SOURCE AMPLIFIER WITH SINGLE TUNED LOAD

the gain-bandwidth product should be roughly independent of center frequency, consider the amplifier shown in Figure 9.15 (biasing details have been omitted). If we drive from a zero-impedance source (as shown) and if we can neglect series gate resistance, then the drain-gate capacitance *Cgd* may be absorbed into the capacitance C. In that case, we can model the circuit as an ideal transconductor driving a parallel *RLC* tank. At low frequencies the inductor is a short and the incremental gain is zero, whereas at high frequencies the gain goes to zero because the capacitor acts as a short. At the resonant frequency of the tank, the gain becomes simply *8mR* since the inductor and capacitor cancel.



For this circuit the total -3-dB bandwidth is, as usual. simply *1/RC* Hence, the product of gain (measured at resonance) and bandwidth is just

$$G \cdot \mathrm{BW} = g_m R \cdot \frac{1}{RC} = \frac{g_m}{C}.$$

The fundamental difference between these two approaches is, of course, due to the cancellation of the load capacitance by the inductor in the tuned amplifier. As long as we have direct access to the terminals of any parasitic capacitance (and can make them appear across the tank), we can resonate out this capacitance with an appropriate choice of inductance and obtain a constant gain-bandwidth product at any arbitrary center frequency.

11.2. DETAILED ANALYSIS OF THE TUNED AMPLIFIER:- In particular, the choice of a zero source resistance and zero gate resistance allowed us to absorb the drain-gate capacitance into the tank network, permitting the inductance to offset its effects, Since *Cgd* might have a more serious effect if it were no longer possible to absorb it directly into the tank let's now allow for nonzero source resistance and nonzero series gate resistance, as shown in Figure 11.1. The corresponding incremental model is depicted in Figure 11.2

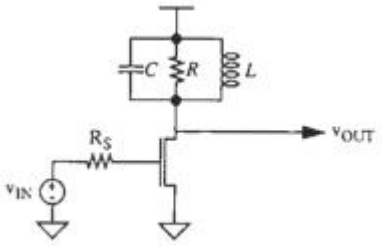


Figure 11.1 Amplifier with single tuned load.

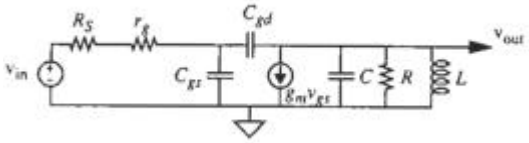


figure11.2. Incremental model for circuit.

Using this model, we can compute two important impedances (actually admittances, to be precise), First, we'll find the equivalent admittance seen to the left of the *RLC* tank; then we'll find the admittance seen to the right of the source resistance *Rs.*

In carrying out this analysis, it is better to apply a test voltage source across the tank to find the equivalent admittance seen to its left. Remember, you'll get the same answer whether you use a test voltage or a test current (assuming you make no errors, or at least the same errors), but a test voltage is more convenient here because it most directly fixes the value of *vgs* the voltage that determines the value of the controlled source.

The end of the result is that the admittance seen by the tank consists of an equivalent resistance (which we'll ignore for now) in parallel with an equivalent capacitance. This capacitance is given by

$$C_{eq} = C_{gd}[1 + g_m R_{eq}] = C_{gd}[1 + g_m(R_S + r_g)].$$

Notice that Ceq can be fairly large. This is actually an alternative manifestation of the Miller effect, now viewed from the output port. Some fraction of the voltage applied to the drain appears across *ligr'* where it excites the *g",* generator. The resulting current adds to that through the capacitors and must be supplied by the test source, so the source sees a lower impedance. One component of that current is due to a simple capacitive voltage divider and is thus in phase with the applied voltage. It therefore represents a resistive load on the tank, causing a gain reduction. Another component of the current leads the applied voltage and therefore represents an additional capacitive load on the tank,

The additional capacitive loading by Ceq shifts downward the resonant frequency of the output tank. Although this shift can be compensated by a suitable adjustment of the inductance, it is generally inadvisable to operate in a regime where the resonant frequency depends critically on poorly

controlled, poorly characterized, and potentially unstable transistor parasitics. It is therefore desirable to select C relatively large compared with the expected variation in parameters, so that the total tank capacitance remains fairly independent of process and operating point. The unfortunate tradeoff is a reduction in the gain-bandwidth product for a given transconductance.

A more serious effect of *Cgd* becomes apparent when we consider the input impedance (or, more directly, the input admittance).

$$y_{in} = \frac{y_L y_F}{y_L + y_F} + \frac{g_m y_F}{y_L + y_F}.$$

where Yin is the admittance seen to the right of *Cgs* "Yf is the admittance of *Cgd,* and *YI.* is the admittance of the *RLC* tank." If, as is often the case, the magnitude of the feedback admittance *Y'F* is small compared to that of *YL,* then we may write

$$y_{in} \approx y_F + \frac{g_m(j\omega C_{gd})}{y_L}.$$

The significance of this result becomes apparent when you observe that *yt.* has a net negative imaginary part at frequencies where the tank looks inductive (i.e., below resonance), so that the second term on the right-hand side of the equation (and therefore Yin) can have a *negative real* part; that is, the input of the circuit can act as if a negative resistor were connected to it. Having negative resistances around can encourage oscillation (which is just fine if this is your intent, but more typically is not).

We certainly have all of the necessary ingredients: inductance, capacitance, and negative resistance. If there were no *Cgd* there would be no such problem .The difficulty with *Cgd,* then, is that it couples the input and output circuits in potentially deleterious ways. It loads the output tank and decreases gain, detunes the output tank, and can cause instability. This latter problem is particularly severe if one attempts to add a tuned circuit to the input as well. Furthermore, even before true instability sets in, the interaction of tuned circuits at both ports may make it extremely challenging to achieve proper tuning.

12. CASCADED AMPLIFIERS:-

12.1. BANDWIDTH SHRINKAG:- Let's suppose that each amplifier stage has a unit DC gain (LO simplify the math marginally) and a single pole. The amplifier's transfer function is then

$$H(s) = \frac{1}{\tau s + 1}.$$

A cascade of *n* such amplifiers will therefore have an overall transfer function of

$$A(s) = \left(\frac{1}{\tau s + 1}\right)^n.$$

We find the bandwidth in the standard way by computing the magnitude of the transfer function and solving for the -3-dB roll off frequency is

$$|A(j\omega)| = \left|\left(\frac{1}{j\omega\tau + 1}\right)\right|^n = \frac{1}{\sqrt{2}},$$

$$\left(\frac{1}{\sqrt{(\omega\tau)^2 + 1}}\right)^n = \frac{1}{\sqrt{2}},$$

$$[(\omega\tau)^2 + 1]^n = 2,$$

and solving for the bandwidth at last gives us

$$\omega = \frac{1}{\tau}\sqrt{2^{1/n} - 1}.$$

That is, the bandwidth of the overall amplifier is the bandwidth of each stage, multiplied by some funny factor. As n approaches infinity, the overall bandwidth tends toward zero.

The precise form of the bandwidth shrinkage is perhaps a little hard to see from this formula. For large *n* though, we can simplify the term under the radical sign to make the relationship substantially clearer. Mathematicians would suggest using a series expansion of $2^{1/n}$ and then using only the first couple of terms. We begin by recognizing that

$$2^{1/n} = \exp\{\ln(2^{1/n})\} = \exp\left\{\frac{1}{n}\ln 2\right\}.$$

Then, for large *n* we can write

$$\exp\left\{\frac{1}{n}\ln 2\right\} \approx 1 + \frac{1}{n}\ln 2.$$

We thus derive the interesting result that the bandwidth behaves approximately as follows as

$$\omega = \frac{1}{\tau}\sqrt{2^{1/n} - 1} \approx \frac{1}{\tau}\sqrt{\frac{1}{n}\ln 2} \approx \frac{0.833}{\tau\sqrt{n}}.$$

That is, the bandwidth shrinks as the inverse square root of the number of stages, at least in the limit of large *n.*

12.2. OPTIMUM GAIN PERSTAGE:-

With the bandwidth shrinkage formula, we're in a position to identify the optimum strategy to maximize bandwidth in a cascaded amplifier, given a stated gain requirement and technology constraints.

Again, we'll assume that all the stages are identical (because if one were slower than any other, it would represent the bandwidth bottleneck for the whole amplifier),each with a single pole *whose frequency depends inversely on the stage gain.* That is, each stage has a constant gain-bandwidth product, so that stage gain and bandwidth trade off linearly. Our goal is to find the number of stages that, for a given overall gain requirement, maximizes the bandwidth (and hence the overall gain-bandwidth product).

Assume that the overall gain is to be *G,* so that each amplifier stage must have a gain of $G^{1/n}$ If each stage has the same gain-bandwidth product *WT,* then the single stage bandwidth will be

$$\mathrm{BW}_{ss} = \frac{\omega_T}{G^{1/n}}.$$

From the approximate bandwidth shrinkage formula, we may write the following expression for the bandwidth of the total amplifier is

$$\mathrm{BW}_{tot} \approx \frac{\omega_T}{G^{1/n}} \cdot \frac{\sqrt{\ln 2}}{\sqrt{n}}.$$

The reciprocal of the bandwidth (somewhat handier for what we'll do shortly) is thus

$$\frac{1}{\mathrm{BW}_{\mathrm{tot}}} \approx \left(\frac{1}{\omega_T\sqrt{\ln 2}} \cdot \sqrt{n}\right) G^{1/n}.$$

We'll now maximize the total bandwidth by minimizing its reciprocal is

$$\frac{d}{dn}\left(\sqrt{n}G^{1/n}\right) = 0.$$

Taking the derivative, cancelling terms and solving yields

$$\ln(G^{1/n}) = \tfrac{1}{2} \implies G^{1/n} = e^{1/2}.$$

According to this analysis, the gain per stage should therefore be chosen as the square root of *e* if we want to maximize the overall bandwidth .The number of stages corresponding to this optimum is

$$n = 2\ln G,$$

and the overall bandwidth corresponding to this condition is

$$\mathrm{BW}_{\mathrm{tot}} = \omega_T \cdot \sqrt{\frac{\ln 2}{2e \cdot \ln G}} \approx \frac{0.357\omega_T}{\sqrt{\ln G}}.$$

From this last expression, we can see that the overall bandwidth is relatively insensitive to the value of overall gain when this optimum is chosen. In fact, the product of bandwidth and the *square root of the log* of gain is constant.

12.3. THE SUPERREGENERATIVE AMPLIFIER:-

A greatly simplified superregenerative amplifier is shown in Figure 9.22. Notice that there is a *negative* resistor in this system. We can always synthesize the equivalent of a negative resistance by using active devices. so invoking its existence here is perfectly realistic.

As a consequence of the negative resistance. the *RC* time constant has a negative value and the pole is in the right-half v-planc, We therefore obtain an exponential *growth* in the capacitor voltage

whenever the sampling switch is open. The longer we wait before closing the switch again. the greater the gain. by an exponentially growing factor.

A quantitative analysis of this amplifier is straightforward. Assume that the switch closes for an infinitesimally short time and that the input source is capable of instantly charging up the capacitor. It is not necessary even to approach these conditions in reality. but accepting these assumptions simplifies the analysis without introducing any fundamental errors. When the switch opens, the capacitor voltage ramps up exponentially from the initial voltage (which is Vin) is

$$v_{out}(t) = v_{in}e^{t/RC}.$$

his exponential growth is allowed to continue for a period *T.* and the exponentially growing signal is averaged by an ideal low-pass filter. Hence

$$\overline{v_{out}(t)} = \frac{1}{T}\int_0^T v_{in}e^{t/RC}\,dt = \frac{RC}{T}(e^{T/RC} - 1)v_{in}.$$

If the time constant *RC* is short compared with the sampling period *T.* we get a gain factor that is exponentially related to that ratio. To involve bandwidth explicitly, note that sine we have a sampled system we must satisfy the Nyquist sampling criterion. Hence, we must choose 1/ *T* higher than twice the highest frequency component of the input signal. That is, we must have

$$\mathrm{BW} < \frac{1}{2T}.$$

Therefore the product of bandwidth and the log of the gain for this type of amplifier are

$$\mathrm{BW} \cdot \ln G = \frac{1}{2RC} + \frac{1}{2T}\ln\left(\frac{RC}{T}\right).$$

Thus, we see that the superregenerative amplifier trades off bandwidth for the log of gain. to a reasonable approximation. As in the case of a cascade of amplifiers the implication is that the bandwidth changes little as one varies the gain over a large range. In fact, if one is willing to endure a sufficiently long regeneration interval. The gain can be made enormously large and so allow the overall gain-bandwidth product to exceed that of the active device(s) involved. Another difference between the superregenerative amplifier and a cascade of con- ventional singlepole amplifiers is that

it accomplishes this gain-bandwidth tradeoff with just one *RC*. The periodically time-varying nature of this amplifier endows it with some of the properties of a high-order system with only one energy-storage element.