

CHAPTER 1

DEVICE PHYSICS PRIMER

Before we begin thinking about analog circuit design we must first understand some basic device physics. Let us first review some electrostatics and transport concepts.

1.1 BASIC ELECTROSTATICS

The fundamental law of electrostatics is Coulomb's law given by

$$F = \frac{q_1 q_2}{4\pi\epsilon r^2} \quad (1.1)$$

which describes the force between two charges q_1 and q_2 separated by a distance r . The inverse square law relationship is a consequence of the homogeneity of space i.e there is no preferred direction for the laws of physics. Therefore, the force should be equal in all direction. In two dimensions this would be circle and in 3-D it describes a sphere. It can be intuitively understood by examining the form of the denominator. It can be seen that it is the surface area of a sphere of radius r . Assume that q_1 is equally distributed on the surface of this sphere the effective charge at any point is then given by $\frac{q_1}{4\pi r^2}$. It can be seen then that the force felt by q_2 is the interaction of q_2 with the effective charge of q_1 at point r divided by ϵ . The effective charge $\frac{q_1}{4\pi r^2}$ divided by ϵ is the effective force per unit charge at the point r and is called the field \vec{E} due to charge q_1 . The direction of the field, \vec{E} pointing normal to the surface of the sphere.



Figure 1.1 Gauss Law Surface

To be able to perform useful calculations with the basic laws we need to be able to compute the effect of multiple charges. Since the system is linear the law of superposition (vector addition) applies. So the field due to multiple charges on a charge at point j due to charges at i is given by $E_{ij} = \sum \frac{q_i}{4\pi\epsilon(r_i-r_j)^2}$. If we do not consider the underlying discrete nature of charge we can make a continuum approximation and replace the summation with the integral giving the following equation for the field $E = \frac{1}{4\pi\epsilon} \int \frac{\rho dV}{r^2}$ where ρ is the volume charge density and r is the distance. Unfortunately unless the charge density is uniform the computation gets difficult. Luckily there is a quantity, namely flux, that can considerably ease the calculation. Flux as the name suggests is the total number of *particles* that are flowing through an area normal to the plane per unit time. The reason why the normal direction is important is that any other direction can be decomposed into a normal and parallel component, while the particle's normal component passes through the surface and hence contributes to the flux, the parallel component does not. However, the term *particles* is a bit too strict and evokes an idea of physical particles in the mind's eye. Though this mental picture is correct and can aid in understanding, it is a limited description. Any *vector field* can have an associated flux. It is, of course a vector field because the direction i.e. normal to the plane is important. An easy way to only account for the normal component of the field is to take the dot product of the field under consideration with the oriented surface vector under consideration. The perceptive reader will be quick to note that this will only work if the surface is a plane. However, this is a minor impediment since we can always approximate any surface as a number of small planes oriented normal to the patch under consideration. The flux is, of course, a scalar. Let us now consider the flux of the electric field through this surface due to some charge enclosed within it. To perform this calculation we find the Electric field \vec{E}_j at patch j and take the dot product with the oriented surface vector \vec{n}_j . By repeating this operation throughout the entire surface we can compute the flux $\phi = \sum \vec{E}_j \cdot \vec{n}_j$. If we make the surface infinitesimally small then we can replace the summation with the integral $\phi = \int_{Surface} \vec{E} \cdot d\vec{n}$.

So how does this formulation help us? Let us consider a simple surface of a sphere of radius r that encloses a charge Q ¹. The electric field in this case is the same at every point and is given by $\frac{Q}{4\pi\epsilon r^2}$. The integral is simply the surface area of the sphere which yields $\phi = \frac{Q}{4\pi\epsilon r^2} \cdot 4\pi r^2 = \frac{Q}{\epsilon}$. This is a remarkable result that shows that the flux going through a sphere does not depend on the surface area of the sphere. This result is entirely a consequence of the inverse square law behavior of the electric field. Moreover, this result does not just hold for spheres it holds for any arbitrary surface as we will soon show.

¹This charge need not be a single charge. It can be a distributed charge then we need to ensure that the sphere radius, r , is many times the radius of a sphere tightly enclosing this charge. The radius r is measured from the center of mass of the charge distribution.

1.2 TRANSPORT

There are two fundamental modes of transport drift and diffusion. In recent years due to ever decreasing device sizes ballistic transport has become important as well. Let us start with diffusion, Fick's 1st law in one dimension, which we will later derive from a molecular perspective, is

$$\phi(x, t) = -D\nabla C(x, t) \quad (1.2)$$

Where D is the diffusion constant, the negative sign arises from the fact the flux direction is opposite to the direction of the gradient. $C(x,t)$ is the concentration which is a function of both space and time. The units of Flux are of course $\frac{\text{particle}}{m^2s}$, therefore the diffusion constant must have units of $\frac{m^2}{s}$.

Let us derive Fick's first law from a molecular perspective, to do this we need to define a few key parameters. The first is τ_f which is the mean time between collisions in a system in thermal equilibrium. By this we mean that the molecules are jiggling about due to temperature and they randomly collide with each other. In between collisions they travel with a mean velocity \bar{v} and travel a distance l . To make matters simple we will consider a fictitious box in one dimension with a large number of particles in thermal equilibrium at temperature T. Imagine now that we are looking at a small box of length $2l$, defined as in figure ???. Now in the interval of a mean time between collision half the particles will travel from the right ($-l/2$) of the box to the left half, while half the particles will travel from the left ($l/2$) to the right. The reason half the particles travel is because every particle has a equal probability that it will travelling to the right or left. Now the let the flux in the $+x$ direction be ϕ_+ and the flux in the $-x$ direction be ϕ_- . The net flux in $+x$ direction is $\phi = \phi_+ - \phi_-$. Now the flux is given by the number of particles/unit area/unit time. The number of particles = concentration *volume = concentration*area* l . So

$$\phi_+ = \frac{1}{2} \frac{C(x - \frac{l}{2}, t)Al}{A\tau_f} \quad (1.3)$$

and similarly

$$\phi_- = \frac{1}{2} \frac{C(x + \frac{l}{2}, t)Al}{A\tau_f}. \quad (1.4)$$

So

$$\phi = \frac{l}{2\tau_f} C(x - \frac{l}{2}, t) - C(x + \frac{l}{2}, t). \quad (1.5)$$

From a macroscopic point of view $x \gg l$ so then we can expand the concentration using a Taylor's series around the point x. i.e.

$$C(x - \frac{l}{2}, t) = C(x, t) - \frac{l}{2} \frac{dC(x, t)}{dx} + \text{higher order terms} \quad (1.6)$$

$$C(x + \frac{l}{2}, t) = C(x, t) + \frac{l}{2} \frac{dC(x, t)}{dx} + \text{higher order terms} \quad (1.7)$$

Ignoring higher order terms since $\frac{l}{2}$ is small and substituting into the equation for flux we get,

$$\phi = \frac{l}{2\tau_f} \left((C(x, t) - \frac{l}{2} \frac{dC(x, t)}{dx}) - (C(x, t) + \frac{l}{2} \frac{dC(x, t)}{dx}) \right) \quad (1.8)$$

$$\phi = -\frac{l^2}{2\tau_f} \frac{dC(x, t)}{dx} \quad (1.9)$$

Comparing this with Fick's first law, we see that

$$D = \frac{l^2}{2\tau_f} \quad (1.10)$$

From the equipartition theorem we have

$$\frac{1}{2} m \bar{v}^2 = \frac{1}{2} \kappa T \quad (1.11)$$

$$\bar{v}^2 = m \left(\frac{l}{\tau_f} \right)^2 = m \frac{D}{\tau_f} \quad (1.12)$$

$$D = \tau_f \frac{kT}{m} \quad (1.13)$$

For semiconductor devices we are generally interested in charged particles and to be consistent with other standard textbooks we will use $n(x, t)$ to denote the concentration of charged carriers. So the diffusive current is then given by

$$J_e^{diffusion} = qD_e \frac{dn(x, t)}{dx} \quad (1.14)$$

where we assumed that the charged particle was an electron with charge $-q$ and concentration $n(x, t)$. If instead we were interested in positive charge carriers, holes, with charge q and concentration $p(x, t)$,

$$J_h^{diffusion} = -qD_h \frac{dp(x, t)}{dx} \quad (1.15)$$

The negative and positive sign ensures that the current direction is consistent. We define the positive current direction to be the flow of positive charges. In this convention an electron current and hole current with the same concentration profile will oppose each other.

We looked at transport due to random thermal motion namely diffusive transport.

How about transport due to a constant driving force F , this is named drift.

Now we know from $\vec{F} = m \frac{d\vec{v}}{dt}$. Separating the variables and integrating we get $\vec{v}^{final} - \vec{v}^{initial} = \frac{\vec{F}}{m} (t^{final} - t^{initial})$. As we know from our discussion on diffusion, particles collide with each other and the mean time between collisions is τ . The particle travels with an average velocity \bar{v} between these collisions. As long as the force is not large enough to change the mean time between collisions we can substitute these parameters to yield $\vec{v} = \frac{\vec{F}\tau}{m}$.

From the definition of flux we have $\vec{J} = C(x, t)\vec{v}$, where \vec{J} is the flux. Substituting for \vec{v} from the above equation we get, $\vec{J} = C(x, t) \frac{\vec{F}\tau}{m}$

Again we can extend this to charged particles by multiplying the concentration by the appropriate charge and multiplying and dividing by q and noting that $\frac{\vec{F}}{q} = \vec{E}$, for electrons we get $J_e^{\vec{drift}} = -q^2 n(x, t) \frac{\vec{E}\tau}{m_e^*}$ and $J_h^{\vec{drift}} = q^2 p(x, t) \frac{\vec{E}\tau}{m_h^*}$ for holes, where we have replaced classical mass with quantum mechanically corrected effective masses due to the interaction of the charged carrier with the ionic potentials. Now we can define the electron and hole mobilities as $\mu_e = \frac{q\tau}{m_e^*}$ and $\mu_h = \frac{q\tau}{m_h^*}$ respectively. So we get $J_e^{\vec{drift}} = -q\mu_e n(x, t)\vec{E}$ and $J_h^{\vec{drift}} = q\mu_h p(x, t)\vec{E}$. We can also relate the velocity to the field as $\vec{v}_e = \mu_e \vec{E}$ and $\vec{v}_h = \mu_h \vec{E}$. We can one further simplification by noting that $q\mu_e n(x, t)$ has the units of conductivity, σ_e . So we can write $J_e^{\vec{drift}} = -\sigma_e \vec{E}$ and similarly $J_h^{\vec{drift}} = \sigma_h \vec{E}$ which is, of course, Ohm's Law.

In our above discussion it was implicitly assumed that the collision time was unaffected by the presence of a field.

1.3 QUANTUM STATISTICS

To Be Written

1.4 METAL-SEMICONDUCTOR JUNCTION

To Be Written

1.5 P-N JUNCTION

To Be Written

1.6 TWO TERMINAL MOS STRUCTURE

Consider a metal-oxide-semiconductor(p-type) sandwich brought into intimate contact. If there is zero potential across of it i.e. we connect a wire between the metal and semiconductor, then electrons will flow from the metal into the semiconductor if the work function of the metal is above the work function of the semiconductor. Now there is a layer of positive charge in the metal oxide interface and there is layer of negative charge in the semiconductor. The 'built-in' potential is given by the difference in work functions of the metal and semiconductor.

$$\phi_{bi} = W_s - W_M \tag{1.16}$$

One can also divide the built-in potential into the drop across the oxide and the drop across the semiconductor. For the sake of convenience we will measure all potential will respect to the bulk i.e. $\phi_{bulk} = 0$. SO we have $\phi_{bi} = \phi_{ox} + \phi_s$

In a p-type semiconductor the total charge in the semiconductor due to a charge on the gate is given by

$$\rho_s = q(p - n - N_A) \tag{1.17}$$

Since there is no current flow in a MOS structure we can use equilibrium physics to compute the relationship between the charge and the potential given by the Boltzmann relation.

$$n = n_{bulk} \exp\left(\frac{q\phi}{kT}\right) \quad (1.18)$$

$$p = p_{bulk} \exp\left(\frac{-q\phi}{kT}\right) \quad (1.19)$$

$$(1.20)$$

We also have the following relationship at equilibrium, far away in the bulk charge neutrality has to hold since the charge in the gate has been compensated. So at the bulk

$$q(p_{bulk} - n_{bulk} - N_A) = 0 \quad (1.21)$$

Recall that a semiconductor in equilibrium maintains an equilibrium concentration between electron and hole concentration given by

$$n_{bulk} p_{bulk} = n_i^2 \quad (1.22)$$

But we also know that the number of holes is approximately equal to the number of dopant atoms since the holes created to the dopant atoms are much greater than the intrinsic hole concentration. so $p_{bulk} \approx N_A$ using this relation in equation 1.22 we get $n = \frac{n_i^2}{N_A}$. Substituting this in the Boltzmann relation we get

$$n = \frac{n_i^2}{N_A} \exp\left(\frac{q\phi}{kT}\right) \quad (1.23)$$

$$p = N_A \exp\left(\frac{-q\phi}{kT}\right) \quad (1.24)$$

$$(1.25)$$

We need to ensure that the charge neutrality is preserved in the bulk where the potential is zero. Using this boundary condition in the above equation gives

$$n = \frac{n_i^2}{N_A} \quad (1.26)$$

$$p = N_A \quad (1.27)$$

$$(1.28)$$

However when substituted into the charge neutrality equation this does not hold. So we need to add a compensating charge of $\frac{n_i^2}{N_A}$ to the overall charge equation.

Making this addition and collecting common terms we get the following relationship

$$\rho_s = qN_A \left(\exp\left(\frac{-q\phi}{kT}\right) - 1 \right) - \frac{n_i^2}{N_A^2} \left(\exp\left(\frac{q\phi}{kT}\right) - 1 \right) \quad (1.29)$$

Poisson's equation from electrostatics is given by $\frac{d^2\phi}{dx^2} = -\frac{\rho_s}{\epsilon_s}$ Using this we get

$$\frac{d^2\phi}{dx^2} = \frac{qN_A}{\epsilon_s} \left(1 - \exp\left(\frac{-q\phi}{kT}\right) \right) + \frac{n_i^2}{N_A^2 \epsilon_s} \left(\exp\left(\frac{q\phi}{kT}\right) - 1 \right) \quad (1.30)$$

Now this is rather nasty equation at first sight but we can perform one simplification using a trick from calculus which is $\frac{d}{dx} \left(\frac{d\phi_s}{dx} \right)^2 = 2 \frac{d\phi_s}{dx} \frac{d^2\phi_s}{dx^2}$ so we get

$$\frac{d}{dx} \left(\frac{d\phi}{dx} \right)^2 = 2 \frac{qN_A}{\epsilon_s} \left(1 - \exp \frac{-q\phi}{kT} \right) + \frac{n_i^2}{N_A^2} \left(\exp \frac{q\phi}{kT} - 1 \right) \frac{d\phi}{dx} \quad (1.31)$$

Integrating on both sides with respect to x we get

$$\left(\frac{d\phi}{dx} \right)^2 = \frac{2qN_A}{\epsilon_s} \left(\phi + \frac{kT}{q} \left(\exp \frac{-q\phi}{kT} - 1 \right) \right) + \frac{n_i^2}{N_A^2} \left(\frac{kT}{q} \left(\exp \frac{q\phi}{kT} - 1 \right) - \phi \right) \quad (1.32)$$

Taking the square root on both sides we get

$$\frac{d\phi}{dx} = \pm \sqrt{\frac{2qN_A}{\epsilon_s} \left[\left(\phi + \frac{kT}{q} \left(\exp \frac{-q\phi}{kT} - 1 \right) \right) - \frac{n_i^2}{N_A^2} \left(\frac{kT}{q} \left(\exp \frac{q\phi}{kT} - 1 \right) - \phi \right) \right]^{\frac{1}{2}}} \quad (1.33)$$

Now we could try and perform an integration by separating the variables but unfortunately a closed form expression cannot be obtained and we have to resort to numerical methods due to the rather complicated function in ϕ . We can however gain some insight into the nature of this function. The first question we have to ask ourselves is which root is valid. We turn to the intuitive physical picture to answer this. We know that $-\frac{d\phi}{dx}$ represents the Electric field. Now the electric field is pointing in the $+x$ direction when the charge in the semiconductor layer is negative or the MOS structure is in depletion or inversion, so this means the negative root is valid. In accumulation we have the positive root valid since the field points in the $-x$ direction, we should keep this in mind when we try making simplifications. Now if we want to write the function without having to worry about this we note that the sign of ϕ gives the negative of the sign to use, of course we need to normalize so we divide by $\|\phi\|$. So we can write the function of ϕ as

$$\mathcal{F}(\phi) = -\frac{\phi}{\|\phi\|} \left[\left(\phi + \frac{kT}{q} \left(\exp \frac{-q\phi_s}{kT} - 1 \right) \right) + \frac{n_i^2}{N_A^2} \left(\frac{kT}{q} \left(\exp \frac{q\phi_s}{kT} - 1 \right) - \phi \right) \right]^{\frac{1}{2}} \quad (1.34)$$

So we can rewrite equation 1.33 as

$$\frac{d\phi}{dx} = \mathcal{F}(\phi) \sqrt{\frac{2qN_A}{\epsilon_s}} \quad (1.35)$$

Now we can, at least symbolically, integrate both sides

$$\int \frac{d\phi}{\mathcal{F}(\phi)} = \int_0^x \sqrt{\frac{2qN_A}{\epsilon_s}} dx \quad (1.36)$$

$$\int \frac{d\phi}{\mathcal{F}(\phi)} = \sqrt{\frac{2qN_A}{\epsilon_s}} x \quad (1.37)$$

$$(1.38)$$

Now we can perform some simplifications based on our intuitive picture by examining the different regions of operation.

1.6.1 Accumulation

In accumulation we know that the concentration of electrons is very small. If we look at the second half of our function $\mathcal{F}(\phi)$ we see that it is multiplied by a pre-factor which is the electron concentration $\frac{n_i^2}{N_A}$ divided by the dopant concentration N_A , this is a very small number in accumulation and can be ignored. Now the value surface potential in accumulation is sufficiently larger than kT/q . Then we can further simplify the function so we have

$$\mathcal{F}(\phi) = \sqrt{\frac{kT}{q}} \exp \frac{-q\phi_s}{2kT} \quad (1.39)$$

So now the integral becomes

$$\int \frac{d\phi}{\exp \frac{-q\phi}{2kT}} = \sqrt{\frac{kT}{q} \frac{2qN_A}{\epsilon_s}} x \quad (1.40)$$

Integrating this function from the surface to the bulk we get

$$\exp \frac{q\phi}{2kT} - \exp \frac{q\phi_s}{2kT} = \sqrt{\frac{q^2 N_A}{2kT \epsilon_s}} x \quad (1.41)$$

The square root of the ratio between the thermal potential to divergence of the electric field is the Debye length², L_D . So we can rewrite the above equation as

$$\exp \frac{q\phi}{2kT} - \exp \frac{q\phi_s}{2kT} = \frac{x}{\sqrt{2}L_D} \quad (1.42)$$

At the end of the accumulation the potential goes to zero. Therefore we have the thickness of the accumulation layer x_{acc} is given by

$$x_{acc} = \sqrt{2}L_D \left(1 - \exp \frac{q\phi_s}{2kT} \right) \quad (1.43)$$

Since the surface potential is sufficiently larger than kT/q and negative we get

$$x_{acc} \approx \sqrt{2}L_D \quad (1.44)$$

From this we can compute the surface charge in the accumulation layer (due to holes) The excess hole concentration (since the background concentration cannot give rise to charge since it is everywhere balanced by the dopant charge) due to the charge on the gate is given by

$$Q_h = \int_0^{x_{acc}} qN_A \left(\exp \frac{-q\phi}{kT} - 1 \right) dx \quad (1.45)$$

²The Debye length is a molecular expression of the drift-diffusion balance. Explicitly is the characteristic length over which the charge must be distributed since we cannot have concentrated sheet of charge.

We can change the limits of integration from x to ϕ and noting that $\frac{d\phi}{dx} = \mathcal{F}(\phi)\sqrt{\frac{2qN_A}{\epsilon_s}}$ and \mathcal{F} is given by equation 1.34 Making this substitution we get

$$Q_h = \int_{\phi_s}^0 qN_A \sqrt{\frac{\epsilon_s}{2qN_A}} \sqrt{\frac{q}{kT}} \exp \frac{q\phi}{2kT} \left(\exp \frac{-q\phi}{kT} - 1 \right) d\phi \quad (1.46)$$

$$Q_h = \frac{\epsilon_s L_D}{\sqrt{2}} \int_{\phi_s}^0 \left(\exp \frac{-q\phi}{2kT} - \exp \frac{q\phi}{2kT} \right) d\phi \quad (1.47)$$

The second term is small so we can neglect it and perform the integration yielding

$$Q_h = \frac{\epsilon_s}{\sqrt{2}L_D} \frac{2kT}{q} \left(\exp \frac{-q\phi_s}{2kT} - 1 \right) \quad (1.48)$$

If we examine this closely we see that the surface charge is given by the molecular capacitance/per unit area times the thermal potential with a multiplying factor.

1.6.2 Depletion

In depletion again we can neglect the term due to electrons since it is small. However the potential is positive and larger than a few kT/q so we have

$$\mathcal{F}(\phi) = -\sqrt{\phi} \quad (1.49)$$

Now performing the integration on equation 1.38 we get

$$\sqrt{\phi_s} - \sqrt{\phi} = \sqrt{\frac{qN_A}{2\epsilon_s}} x \quad (1.50)$$

We can computer the maximum width of the depletion layer as

$$x_{depmax} = \sqrt{\frac{2\phi_{sth}\epsilon_s}{qN_A}} \quad (1.51)$$

where we have denoted the surface potential at maximum depletion as ϕ_{sth}

The total depletion charge is given by

$$Q_s = -qN_A x_{depmax} = -\sqrt{2qN_A\phi_{sth}\epsilon_s} \quad (1.52)$$

1.6.3 Inversion

Inversion is a bit tricky to deal with since it is really composed of two regions an inverted region near the surface and a depleted region away from the surface. Now we need a delineating condition in terms of potential so that we can treat these two regions. If the total electron concentration in the region due the potential at that region is less than the dopant concentration then it should be in depletion and is above then it should be in inversion. At the onset of inversion the total electron concentration must be equal to the dopant concentration Therefore

$$n = \frac{n_i^2}{N_A} \exp q\phi_{sth}kT = N_A \quad (1.53)$$

$$\phi_{sth} = 2 \frac{kT}{q} \ln \frac{N_A}{n_i} = 2\phi_f \quad (1.54)$$

Since the characteristics change in the semiconductor as a function of surface potential we will use a piecewise approximation to make our lives easier. When $\phi > 2\phi_f$ the function

$$\mathcal{F}(\phi) = -\sqrt{\frac{n_i^2}{N_A^2} \left(\frac{kT}{q} (\exp \frac{q\phi - 2\phi_f}{kT} - 1) + 2\phi_f \right)} \quad (1.55)$$

$\phi \leq 2\phi_f$ then we can use the depletion approximation derived earlier.

Now we will follow a slightly different path to avoid confusing the main issue with too much math. Now equation 1.35 is the Electric field in the semiconductor i.e.

$$\vec{E}_s = \mathcal{F}(\phi) \sqrt{\frac{2qN_A}{\epsilon_s}} \quad (1.56)$$

Now the divergence of the field gives the total charged enclosed in the volume divided by the permittivity of the semiconductor. The surface charge density is, however, given by the field at the surface times the permittivity so we get

$$Q_s = \mathcal{F}(\phi_s) \epsilon_s \sqrt{\frac{2qN_A}{\epsilon_s}} \quad (1.57)$$

Substituting the expression for $\mathcal{F}(\phi_s)$ we get

$$Q_s = -\epsilon_s \sqrt{\frac{qN_A}{\epsilon_s}} \sqrt{\left(\frac{n_i^2}{N_A^2} \frac{kT}{q} (\exp \frac{q\phi_s - 2\phi_f}{kT} - 1) + 2\phi_f \right)} \quad (1.58)$$

We can get the inversion layer charge by subtracting the depletion layer charge from the total charge.

$$Q_i = -\sqrt{2qN_A\epsilon_s} \left[\sqrt{\left(2\phi_f + \frac{n_i^2}{N_A^2} \frac{kT}{q} (\exp \frac{q\phi_s - 2\phi_f}{kT} - 1) \right)} - \sqrt{2\phi_f} \right] \quad (1.59)$$

We now would like to relate the inversion charge in the channel to the applied potential. For this we turn to Kirchoff's voltage law which is an expression of energy conservation i.e.

$$V + \phi_{bi} = \phi_{ox} + \phi_s \quad (1.60)$$

Now we can bring the semiconductor into a "flat-band" condition by applying an appropriate potential on the gate w.r.t the semiconductor bulk. Of course the value of this potential must be exactly equal to the negative built-in potential. Expressing this mathematically and substituting into equation 1.60. we get

$$V_{FB} = -\phi_{bi} \quad (1.61)$$

$$V = V_{FB} + \phi_{ox} + \phi_s \quad (1.62)$$

Since there is no charge in the oxide the field in the oxide is a constant, so the potential is linearly related to the field via the oxide thickness i.e.

$$\phi_{ox} = -\vec{E}_{ox} x_{ox} \quad (1.63)$$

At the semiconductor oxide interface the displacement field must be equal so we get

$$\vec{E}_{ox} = \frac{\epsilon_s}{\epsilon_{ox}} \vec{E}_s \quad (1.64)$$

$$\vec{E}_s = \frac{Q_s}{\epsilon_s} \quad (1.65)$$

$$\vec{E}_{ox} = \frac{Q_s}{\epsilon_{ox}} \quad (1.66)$$

$$\phi_{ox} = -\frac{Q_s x_{ox}}{\epsilon_{ox}} \quad (1.67)$$

The capacitance per unit area of the oxide is given by

$$C_{ox} = \frac{\epsilon_{ox}}{x_{ox}} \quad (1.68)$$

Using this relationship we can express the potential drop across the oxide as

$$\phi_{ox} = -\frac{Q_s}{C_{ox}} \quad (1.69)$$

The total charge in the semiconductor can be expressed as sum of the inversion and depletion layer charges $Q_s = Q_i + Q_{dep}$. Using this relationship we get

$$\phi_{ox} = -\frac{Q_i + Q_{dep}}{C_{ox}} \quad (1.70)$$

$$V = V_{FB} + \phi_s - \frac{Q_i + Q_{dep}}{C_{ox}} \quad (1.71)$$

$$Q_i = -C_{ox} \left(V - (V_{FB} + \phi_s + \frac{Q_{dep}}{C_{ox}}) \right) \quad (1.72)$$

The expression for Q_i depends on the linear value of ϕ_s in equation 1.72 but also the exponential value of ϕ_s in equation 1.59. However, since the surface potential does not change much (due to the exponential nature of charge dependence in the inversion layer on the surface potential) we can approximate $\phi_s = 2\phi_f$ for small values of ϕ_s . So we get

$$Q_i = -C_{ox} \left(V - (V_{FB} + 2\phi_f + \frac{\sqrt{q\epsilon_s N_A 2\phi_f}}{C_{ox}}) \right) \quad (1.73)$$

We can make a couple of simplifications, we can define the potential required to bring on the onset of inversion as the threshold voltage and also define a body factor that relates the control of the bulk on the electrostatics of the channel to that of the gate.

$$\gamma = \frac{\sqrt{2q\epsilon_s N_A}}{C_{ox}} \quad (1.74)$$

$$V = V_{th} \mapsto Q_i = 0 \quad (1.75)$$

$$V_{th} = V_{FB} + 2\phi_f + \gamma\sqrt{2\phi_f} \quad (1.76)$$

$$Q_i = -C_{ox}(V - V_{th}) \quad (1.77)$$

1.6.4 Subthreshold

In the depletion domain of the MOS structure there is finite electron concentration in the depletion region. The reason for this is that the potential applied on the gate makes the surface attractive enough for the thermally generated electrons. We can calculate this electron concentration by simply using the Boltzmann relation i.e.

$$Q_e = -q \frac{n_i^2}{N_A} \int_0^\infty \exp \frac{q\phi}{kT} dx \quad (1.78)$$

However, we notice that the electron concentration is mainly at the surface of the MOS structure. We will denote the thickness of this layer as t and the surface potential is approximately constant around t . Performing the integration we get.

$$Q_e = -qt \frac{n_i^2}{N_A} \exp \frac{q\phi_s}{kT} \quad (1.79)$$

Now we need to relate ϕ_s to the applied potential V . If we look at the structure closely we see that the potential at the surface is the result of capacitive charge division between the gate capacitor and the depletion capacitor. So we get

$$\frac{d\phi_s}{dV} = \frac{C_{ox}}{C_{ox} + C_{dep}} \quad (1.80)$$

$$\phi_s = \frac{V_{gb}}{n} + K \quad (1.81)$$

$$K = -\frac{V_{th}}{n} + 2\phi_f \quad (1.82)$$

$$\phi_s = \frac{(V_{gb} - V_{th})}{n} + 2\phi_f \quad (1.83)$$

$$(1.84)$$

Substituting this and collecting terms we get

$$Q_e = -qt \frac{n_i^2}{N_A} \exp \frac{2q\phi_f}{kT} \exp \frac{q(V_{gb} - V_{th})}{nkT} \quad (1.85)$$

1.7 MOSFET LONG-CHANNEL

1.7.1 Current flow in Strong Inversion

Let us derive the current that flows in a mosfet in strong inversion. Current = Charge \times velocity Now the channel charge = Width of Channel \times Inversion Layer Charge \times mobility \times Electric field. Inversion layer charge = Channel Capacitance

\times potential in the channel. Consider a point y in the channel. The charge at this point

$$Q(y) = -C_{ox}(V_g - V_{tn} - V(y)) \quad (1.86)$$

$$J = qNv(y) \quad (1.87)$$

$$N = \frac{Q(y)}{t} \quad (1.88)$$

$$I = JWt \quad (1.89)$$

$$I = WQ(y)\mu_n E(y) \quad (1.90)$$

$$E(y) = -\frac{dV(y)}{d(y)} \quad (1.91)$$

$$I = WC_{ox}Q(y)\mu_n \frac{dV(y)}{d(y)} \quad (1.92)$$

$$\int_0^L Idy = \int_{V_s}^{V_d} W\mu_n C_{ox}(V_g - V_{tn} - V(y)) dV(y) \quad (1.93)$$

$$IL = W\mu_n C_{ox} \left((V_g - V_{tn})V(y) \Big|_{V_s}^{V_d} - \frac{V(y)^2}{2} \Big|_{V_s}^{V_d} \right) \quad (1.94)$$

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_{tn})(V_d - V_s) - \frac{1}{2}(V_d^2 - V_s^2) \right) \quad (1.95)$$

We can rewrite this to refer to V_s as follows:

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_s + V_s - V_{tn})(V_d - V_s) - \frac{1}{2}(V_d^2 - V_s^2) \right) \quad (1.96)$$

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_s) - V_{tn} \right) (V_d - V_s) + V_s(V_d - V_s) - \frac{1}{2}(V_d^2 - V_s^2) \quad (1.97)$$

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_s - V_{tn})(V_d - V_s) + V_s V_d - V_s^2 - \frac{V_d^2}{2} + \frac{V_s^2}{2} \right) \quad (1.98)$$

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_s) - V_{tn} \right) (V_d - V_s) + V_s V_d - \frac{V_s^2}{2} - \frac{V_d^2}{2} \quad (1.99)$$

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_s) - V_{tn} \right) (V_d - V_s) - \frac{1}{2}(V_d^2 + V_s^2 - 2V_d V_s) \quad (1.100)$$

$$I = \mu_n C_{ox} \frac{W}{L} \left((V_g - V_s) - V_{tn} \right) (V_d - V_s) - \frac{1}{2}(V_d - V_s)^2 \quad (1.101)$$

We can write $V_g - V_s$ as V_{gs} , meaning the gate voltage w.r.t the source and $V_d - V_s$ as V_{ds} , the drain voltage w.r.t the source.

$$I = \mu_n C_{ox} \frac{W}{L} \left(V_{gs} - V_{tn} \right) V_{ds} - \frac{1}{2} V_{ds}^2 \quad (1.102)$$

In Saturation the potential across the inversion layer is $V_{DS} \cong V_{GS} - V_{th}$ using this relation we get

$$I = \mu_n C_{ox} \frac{W}{2L} (V_{gs} - V_{tn})^2 \quad (1.103)$$

This is only approximately right.

1.7.2 The body effect

We have implicitly assumed till now that the body and the source at the same potential so the reference can be interchanged without any change in the equation. However, when a finite source body potential exists we need to modify the equations. This is fairly straight forward since all the equations that we use depend on the surface potential referred to the bulk. We can take the finite source bulk potential into account by substituting $\phi_s^{bulk} \mapsto \phi_s^{source} + V_{sb}$ and that $\phi_s^{source} \cong 2\phi_f$ in strong inversion. When we use this substitution in the threshold voltage we get

$$V_{th} = V_{FB} + 2\phi_f + V_{sb} + \gamma\sqrt{2\phi_f + V_{sb}} \quad (1.104)$$

$$V_{th} = V_{FB} + 2\phi_f + \gamma\sqrt{2\phi_f} + \gamma\sqrt{2\phi_f + V_{sb}} - \gamma\sqrt{2\phi_f} \quad (1.105)$$

$$V_{th} = V_{th0} + \gamma\sqrt{2\phi_f + V_{sb}} - \gamma\sqrt{2\phi_f} \quad (1.106)$$

This is especially important in computing the current in the channel. Since the threshold changes along the channel due to the potential between the channel and the bulk. We can easily model this effect by modifying equation ??,1.77 to give

$$Q_i = -C_{ox} \left(V_{gs} - V(y) - (V_{th0} + \gamma\sqrt{V(y) + 2\phi_f} - \gamma\sqrt{2\phi_f}) \right) \quad (1.107)$$

Substituting this in equation 1.95 we get

$$I = WC_{ox}Q(y)\mu_n \frac{dV(y)}{d(y)} \quad (1.108)$$

$$I = WC_{ox}\mu_n \left(V_{gs} - V(y) - (V_{th0} + \gamma\sqrt{V(y) + 2\phi_f} - \gamma\sqrt{2\phi_f}) \right) \frac{dV(y)}{d(y)} \quad (1.109)$$

Integrating this equation we get

$$\int_0^L Idy = \int_0^{V_{ds}} WC_{ox}\mu_n \left(V_{gs} - V(y) - (V_{th0} + \gamma\sqrt{V(y) + 2\phi_f} - \gamma\sqrt{2\phi_f}) \right) dV(y) \quad (1.110)$$

Substituting the limits and simplifying we get

$$IL = WC_{ox}\mu_n \left[(V_{gs} - V_{th0})V_{ds} - \frac{V_{ds}^2}{2} \frac{2}{3} \gamma((V_{ds} + 2\phi_f)^{3/2} - (2\phi_f)^{3/2}) + \gamma\sqrt{2\phi_f}V_{ds} \right] \quad (1.111)$$

$$I = \frac{W}{L} \mu_n C_{ox} \left[\left(V_{gs} - V_{th0} + \gamma\sqrt{2\phi_f} - \frac{V_{ds}}{2} \right) V_{ds} - \frac{2}{3} \gamma((V_{ds} + 2\phi_f)^{3/2} - (2\phi_f)^{3/2}) \right] \quad (1.112)$$

We can also use this in the saturation equation by substituting V_{ds} with $V_{ds} = V_{gs} - V_{th}$ when the inversion charge at the drain end is 0.

1.7.3 Subthreshold Current

To derive the subthreshold equation we turn to diffusive flow since the electron concentration is very small and hence drift does not play a major role. So we have

$$I_{ds}^{sub} = -qWD_n \frac{dn}{dx} \quad (1.113)$$

Since there is no recombination in the channel (assuming the channel is smaller than the recombination length) the differential can be replaced by the linear equivalent and we get

$$I_{ds}^{sub} = qWD_n \frac{Q_s - Q_d}{L} \quad (1.114)$$

The drain and source electron concentration can be found using equation 1.79 and changing the reference for the surface potential making this substitution we get

$$Q_s = -qt \frac{n_i^2}{N_A} \exp \frac{2q\phi_f}{kT} \exp \frac{q\phi_s^{source}}{kT} \quad (1.115)$$

$$\phi_s^{source} = \phi_s^{bulk} - V_{sb} \quad (1.116)$$

$$\phi_s^{source} = \frac{(V_{gb} - V_{th})}{n} - V_{sb} \quad (1.117)$$

$$\phi_s^{source} = \frac{(V_{gb} - V_{sb} + V_{sb} - V_{th})}{n} - V_{sb} \quad (1.118)$$

$$\phi_s^{source} = \frac{(V_{gs} - V_{th})}{n} - V_{sb} \frac{n-1}{n} \quad (1.119)$$

$$Q_s = -qt \frac{n_i^2}{N_A} \exp \frac{2q\phi_f}{kT} \exp \frac{(V_{gs} - V_{th})}{n} - qV_{sb} \frac{n-1}{n} \quad (1.120)$$

Similarly we can compute the electron concentration at the drain side

$$Q_d = -qt \frac{n_i^2}{N_A} \exp \frac{2q\phi_f}{kT} \exp \frac{q(V_{gb} - V_{th})}{n} - qV_{db} \quad (1.121)$$

$$Q_d = -qt \frac{n_i^2}{N_A} \exp \frac{2q\phi_f}{kT} \exp \frac{q(V_{gs} - V_{th})}{n} + q \frac{V_{sb}}{n} - qV_{db} \quad (1.122)$$

$$(1.123)$$

Then I_{ds} is given by

$$I_{ds}^{sub} = \frac{qWtD_n n_i^2}{LN_A} \exp \frac{2q\phi_f}{kT} \exp \frac{q(V_{gs} - V_{th})}{nkT} \exp \frac{-qV_{sb}(n-1)}{nkT} \left(1 - \exp \frac{-qV_{ds}}{kT} \right) \quad (1.124)$$

If $V_{ds} \gg \frac{kT}{q}$, the current becomes independent of the drain-source voltage and the transistor is in saturation and the current is given by,

$$I_{dssat}^{sub} = \frac{qWtD_n n_i^2}{LN_A} \exp \frac{2q\phi_f}{kT} \exp \frac{q(V_{gs} - V_{th})}{nkT} \exp \frac{-qV_{sb}(n-1)}{nkT} \quad (1.125)$$

1.8 MOSFET SHORT-CHANNEL

As devices scale, effects that were ignored in the earlier derivation need to be taken into account for an accurate prediction of the drain current we will examine these effects and understand the physical reasons for them.

1.8.1 Mobility Degradation

The vertical field imposed by the gate causes the charge carriers to frequently collide at the semiconductor-insulator interface. The quality of the $Si - SiO_2$ interface is usually of very quality so we can assume that most of the collision energy is carried by phonons. Under this simplifying assumption we can relate the mobility to the applied field. In a low field condition the mobility is set by bulk scattering mechanisms, namely ionized impurity scattering. In strong inversion the inversion charge effectively screens the bulk charge and the mobility is dominated by the surface scattering effects. Mobility is inversely proportional to the normal field so we can express the mobility as

$$\mu_0 \propto \frac{1}{|E_0(x)|} \quad (1.126)$$

$$\mu_{av} \propto \frac{1}{|E_{av}(x)|} \quad (1.127)$$

$E_0(x)$ is the field that gives rise to the low field mobility μ_0 under bulk scattering dominated conditions and E_{av} is the average field that gives rise to the surface scattering dominated mobility μ_{av} .

In general both mechanisms are present and we can use 'Mathieson's rule' to get an effective mobility,³ μ_{eff} .

$$\frac{1}{\mu_{eff}} = \frac{1}{\mu_0} + \frac{1}{\mu_{av}} \quad (1.128)$$

We can substitute equation 1.127 and simplify to get

$$\mu_{eff} = \frac{\mu_0}{1 + \left| \frac{E_{av}}{E_0} \right|} \quad (1.129)$$

It should be noted that experimentally it has been found that the effective field dependent mobility is given by

$$\mu_{eff} = \frac{\mu_0}{\left(1 + \frac{E_{av}}{E_0} \right)^\nu} \quad (1.130)$$

The value of ν that gives the best fit for electron mobility is 2 and for hole mobility it is 1. However, to get a tractable closed form equation ν is often assumed to be 1.

The average field is given by the expectation of the field with the probability given by the charge at the point in a vertical slice of the MOS structure, this of course needs to be normalized. Mathematically

$$E_{av} = \frac{\int E(x)qn(x)dx}{\int qn(x)dx} \quad (1.131)$$

³This can be intuitively derived from the superposition of fields and the inverse relationship between the applied field and the mobility

We can change the integration variable from $x \mapsto E_i(x) + E_{dep}$ by using Gauss's law $\frac{dE(x)}{dx} = \frac{-qn(x)}{\epsilon_s}$ we get

$$\vec{E}_{av} = \frac{\int (E_i(x) + E_{dep}) dE_i(x)}{\int dE_i(x)} \quad (1.132)$$

Performing the integration we get

$$\vec{E}_{av} = \frac{1}{2} E_i(x) + E_{dep} \quad (1.133)$$

The field is given by $\frac{Q}{\epsilon_s}$ making this substitution we get

$$\vec{E}_{av} = \frac{1}{\epsilon_s} \left(\frac{Q_i}{2} + Q_{dep} \right) \quad (1.134)$$

To get a sense for how device scaling affects the average field we can relate the inversion layer charge and depletion charge to the applied voltage i.e

$$Q_i = -C_{ox} (V_{gs} - V_{th}) \quad (1.135)$$

$$Q_{dep} = -C_{ox} (V_{th} - V_{FB} - 2\phi_f) \quad (1.136)$$

Using these relationships we get

$$\vec{E}_{av} = \frac{-\epsilon_{ox} \left(\frac{V_{gs} + V_{th}}{2} - V_a \right)}{\epsilon_s x_{ox}} \quad (1.137)$$

Where we have made use of the $C_{ox} = \frac{\epsilon_{ox}}{x_{ox}}$ and $V_a = V_{FB} + 2\phi_f$.

1.8.2 Velocity Saturation

If the field is high enough the carrier temperature is no longer in equilibrium with the lattice and hot carrier transport must be considered. The effective carrier thermal velocity now becomes the bottleneck since the phonon collision does not effectively randomize the particle's velocity. Experimentally the velocity dependence on the field can be modelled as

$$v(E) = \frac{\mu_{eff} \vec{E}}{\left(1 + \frac{E}{E_{sat}} \right)^n} \quad (1.138)$$

We can compute the channel current under velocity saturation conditions as

$$\vec{J}_{ds} = v(E) \frac{Q_i}{t} I = JWt = Wv(E)Q_i \quad (1.139)$$

Substituting the relation for the inversion layer charge in equation ?? we get

$$I_{ds} = -\mu_{eff} \frac{\vec{E}}{\left(1 + \left| \frac{E}{E_{sat}} \right| \right)} C_{ox} W (V_{gs} - V_{th} - V(y)) \quad (1.140)$$

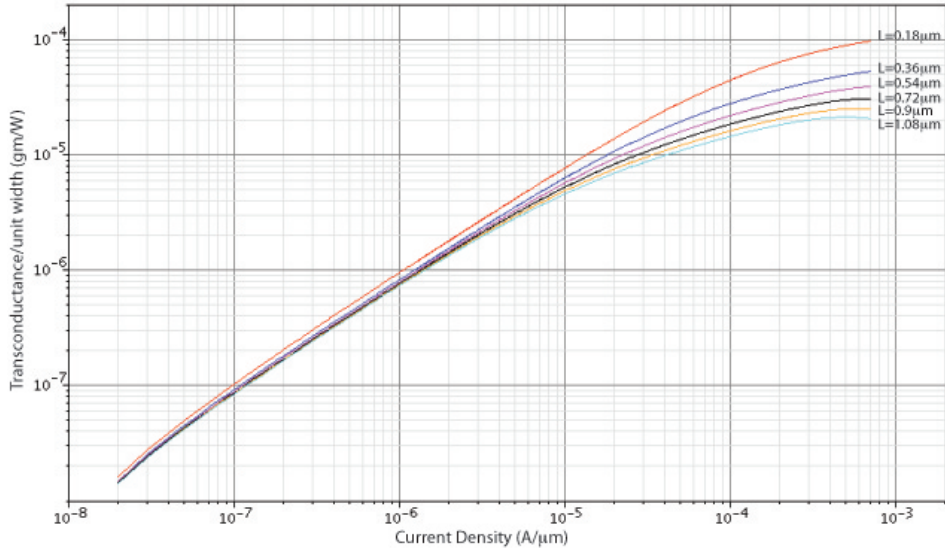


Figure 1.2 gm vs current density in a 0.18 μ m technology

Using the relation $E(y) = -\frac{dV(y)}{dy}$ we get

$$\int_0^L I_{ds} \left(1 + \left| \frac{dV(y)/dy}{E_{sat}} \right| \right) dy = \mu_{eff} C_{ox} W \int_0^{V_{ds}} (V_{gs} - V_{th} - V(y)) dV(y) \quad (1.141)$$

Performing the integration and simplifying we get

$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L} \left(V_{gs} - V_{th} - \frac{V_{ds}}{2} \right) \frac{V_{ds}}{1 + \left| \frac{V_{ds}}{LE_{sat}} \right|} \quad (1.142)$$

In saturation $V_{ds} \mapsto V_{gs} - V_{th}$, substituting this in equation 1.142 we get

$$I_{dssat} = \frac{\mu_{eff}}{2 \left(1 + \left| \frac{(V_{gs} - V_{th})}{LE_{sat}} \right| \right)} C_{ox} \frac{W}{L} (V_{gs} - V_{th})^2 \quad (1.143)$$

1.9 PROBLEMS

1.10 FURTHER READING

REFERENCES

1. Edward M. Purcell, *Electricity and Magnetism*, McGraw-Hill, New York, 1985.
2. Jesus Del Alamo, *Integrated Microelectronic Devices: Physics and Modelling*, Prentice-Hall, New York, 2002.
3. Robert F. Pierret, *Semiconductor Device Fundamentals*, Addison Wesley, 2nd edition, 1996