131. W. D. Eades, J. D. Shott, and R. M. Swanson, "Refinements in the Measurement of Depleted Generation Lifetime," *IEEE Trans. Electron Dev.* **ED-30**, 1274–1277, Oct. 1983.
132. S. Venkatesan, R. F. Pierret, and G. W. Neudeck, "A New Lifetime Sweep Technique to Measure Generation Lifetimes in Thin-Film SOI MOSFET's," *IEEE Trans. Electron Dev.* **41**, 567–574, April 1994.
133. D. K. Schroder, M. S. Fung, R. L. Verkuil, S. Pandey, W. C. Howland, and M. Kleefstra, "Corona-Oxide-Semiconductor Device Characterization,'' *Solid-State Electron.* **42**, 505–512, April 1998.
134. G. Duggan and G. B. Scott, "The Efficiency of Photoluminescence of Thin Epitaxial Semiconductors," *J. Appl. Phys.* **52**, 407–411, Jan. 1981.
135. H. J. Hovel, "Solar Cells," in *Semiconductors and Semimetals* (R. K. Willardson and A. C. Beer, eds.) **11**, Academic Press, New York, 1975, 17–20.
136. H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, Oxford University Press, Oxford, 1959.
137. Y. I. Ogita, "Bulk Lifetime and Surface Recombination Velocity Measurement Method in Semiconductor Wafers," *J. Appl. Phys.* **79**, 6954–6960, May 1996.
138. E. Gaubas and J. Vanhellemont, "A Simple Technique for the Separation of Bulk and Surface Recombination Parameters in Silicon," *J. Appl. Phys.* **80**, 6293–6297, Dec. 1996.
139. J. S. Blakemore, *Semiconductor Statistics*, Pergamon Press, New York, 1962.

Schroder, D.K.
Semiconductor Material + Device
Characterization 2/e. Wiley Interscience 1998

CHAPTER 8

MOBILITY

8.1 INTRODUCTION

The carrier mobility influences the device behavior through its frequer response or time response in two ways. First, the carrier velocity is prop tional to the mobility for low electric fields. Hence a higher mobility mate is likely to have a higher frequency response, because carriers take less ti to travel through the device. Second, the device current depends on mobility, and higher mobility materials have higher current. Higher curre charge capacitances more rapidly, resulting in a higher frequency respor

There are several mobilities in use. The fundamental mobility is *microscopic mobility*, calculated from basic concepts. It describes the mobi of the carriers in their respective band. The *conductivity mobility* is deri from the conductivity or the resistivity of a semiconducting material. *Hall mobility* is determined from the Hall effect and differs from conductivity mobility by a factor dependent on the scattering mechanis The *drift mobility* is the mobility measured when minority carriers drift i electric field. It is a device-oriented mobility and therefore very useful. B is not as easy to measure as the Hall mobility, for example, and is not use extensively for that reason.

The geometry has a major influence on the mobility in some devices. example, surface scattering has a major influence in reducing the mobili MOS field effect transistors. The resulting mobility, determined from device current-voltage characteristic, is termed the *effective mobility*. In tion there are considerations that cause further division between *ma* mobility and *minority carrier mobility*. Momentum considerations that electron-electron or hole-hole scattering has no first-order effect on the mobility. However, electron-hole scattering does reduce the mobility, since electrons and holes have opposite average drift velocities. Hence minority carriers experience ionized impurity and electron-hole scattering, while majority carriers experience ionized impurity scattering. We address measurement techniques for the most commonly used mobilities in this chapter.

8.2 CONDUCTIVITY MOBILITY

The conductivity σ of a semiconductor is given by

$$\sigma = q(\mu_n n + \mu_p p) \tag{8.1}$$

For reasonably extrinsic p-type semiconductors $p \gg n$, and the hole or *conductivity mobility* from Eq. (8.1) is

$$\mu_p = \frac{\sigma}{qp} = \frac{1}{q\rho p} \tag{8.2}$$

Measurement of the conductivity and carrier density was one of the first means of determining the semiconductor mobility, namely, the conductivity mobility.[1, 2] The main reasons for its use are ease of measurement and the fact that the Hall scattering coefficient, to be discussed in the following chapter, need not be known. To determine the conductivity mobility, it suffices to measure the majority carrier density and either the conductivity or the resistivity of the sample independently. The method is rarely used today.

8.3 HALL EFFECT AND MOBILITY

8.3.1 Basic Equations for Uniform Layers or Wafers

The Hall effect was discovered by Hall in 1879 when he investigated the nature of the force acting on a conductor carrying a current in a magnetic field.[3] In particular, he measured the transverse voltage on gold foils. Suspecting the magnet may tend to deflect the current, he wrote "... that in this case there would exist a state of stress in the conductor, the electricity pressing, as it were, toward one side of the wire.... I thought it necessary to test for a difference of potential between points on opposite sides of the conductor." A nice discussion of the discovery of the Hall effect including excerpts from Hall's unpublished notebook is given by Sopka.[4]

Discussions [illegible] the Hall effect can be found in many solid-state and semiconductor books. A comprehensive treatment is given by Putley.[5] Th

terization of semiconductor materials because it gives the *resistivity*, the *carrier density*, and the *mobility*. The use of the Hall effect for resistivity measurements is discussed in Chapter 1, and its use in carrier density characterization is discussed in Chapter 2. In this chapter we give a more detailed discussion of the Hall effect and its application to mobility measurements.

Hall found that a magnetic field applied to a conductor perpendicular to the current flow direction produces an electric field perpendicular to both the magnetic field and the current. Consider the p-type semiconductor sample shown in Fig. 8.1. A current I flows in the x-direction, indicated by the holes flowing to the right and a magnetic field B is applied in the z-direction. The current is given by

$j = n|e|v$

$I = n|e|Av$

$$I = qApv_x = qwdpv_x \tag{8.3}$$

The voltage along the x-direction, indicated by V_ρ, is

$$V_\rho = \frac{\rho s I}{wd} \tag{8.4}$$

Ohm's law.

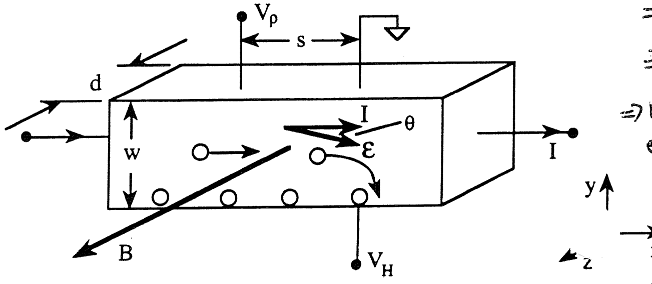
from which the resistivity is derived as

$$\rho = \frac{wd}{s}\frac{V_\rho}{I} \tag{8.5}$$

Consider now the motion of holes in a uniform magnetic field strength B. The force on the holes is given by the vector expression

$$\mathbf{F} = q(\mathscr{E} + \mathbf{v} \times \mathbf{B}) \tag{8.6}$$

The magnetic field in conjunction with the current forces some holes to be deflected to the bottom of the sample, as indicated in Fig. 8.1. For an n-type



holes: $\vec{F} = q\vec{v} \times \vec{B}$

$= qv\hat{x} \times B\hat{z}$

$= qvB\hat{y}$

$\Rightarrow$ holes go down

~~[illegible]~~.

$e^-: \vec{F} =$

$-|e|v(-\hat{x}) \times \hat{B}$

$= |e|vB\hat{y}$

e^- go down

sample, the electrons are also deflected to the bottom of the sample for the same current direction as that in Fig. 8.1, because they flow in the opposite direction to holes and have opposite charge. The holes accumulate at the bottom surface. In the y-direction there is no net force on the holes since no current can flow in that direction and therefore $\mathscr{E}_y = 0$. Combining Eqs. (8.6) and (8.3) allows us to write the vertical electric field as

$$\mathscr{E}_y = Bv_x = \frac{BI}{qwdp} \tag{8.7}$$

The electric field in the y-direction produces the Hall voltage V_H,

$$\Delta V = -\int \vec{E} \cdot d\vec{r}$$

$$\int_0^{V_H} dV = V_H = -\int_w^0 \mathscr{E}_y \, dy = -\int_w^0 \frac{BI}{qwdp} dy = \frac{BI}{qt\overset{\leftarrow d}{p}} \tag{8.8}$$

The Hall coefficient R_H is defined as

$$R_H = \frac{dV_H}{BI} \quad = \frac{1}{qp} \tag{8.9}$$

The angle θ between the current and the net electric field is the Hall angle, given by

$$\tan(\theta) = \frac{\mathscr{E}_y}{\mathscr{E}_x} = B\mu_p \tag{8.10}$$

using Eq. (8.7) and $I = qp\mu_p \mathscr{E}_x wd$. from 8.1 and 8.3

EXERCISE 8.1

Problem

How is R_H converted from the mks to the cgs system?

Solution

For the mks system, the units of R_H are m^3/C for d in m, V_H in V, B in T (1 T = 1 Tesla = 1 Weber/m^2 = 1 V $\cdot$ s/m^2), and I in A. What are the cgs units? One way to determine this is to use Eq. (8.9), i.e.,

$$V_H = \frac{R_H BI}{d} = \frac{R_H(\text{cm}^3/\text{C}) \times 10^{-6}(\text{m}^3/\text{cm}^3) B(G) \times 10^{-4}(T/G) \times I(\text{A})}{d(\text{cm}) \times 10^{-2}(\text{m}/\text{cm})}$$

$$= 10^{-8} \frac{R_H BI}{}$$

or $R_H = 10^8(dV_H/BI)$ for R_H in cm^3/C, d in cm, V_H in V, B in G (Gauss; 10,000 G = 1 T), and I in A. For $B = 5000$ G, $I = 0.1$ mA, and $p = 10^{15}$ cm^{-3}, we find $V_H = 3.1/d$. For a wafer of thickness $d = 5 \times 10^{-2}$ cm, this gives a Hall voltage $V_H \approx 6$ mV and $R_H \approx 60{,}000$ cm^3/C.

Combining Eqs. (8.8) and (8.9) gives

$$p = \frac{1}{qR_H} \tag{8.11}$$

A similar derivation for n-type samples gives

$$n = -\frac{1}{qR_H} \tag{8.12}$$

When both holes and electrons are present, the Hall coefficient becomes[6]

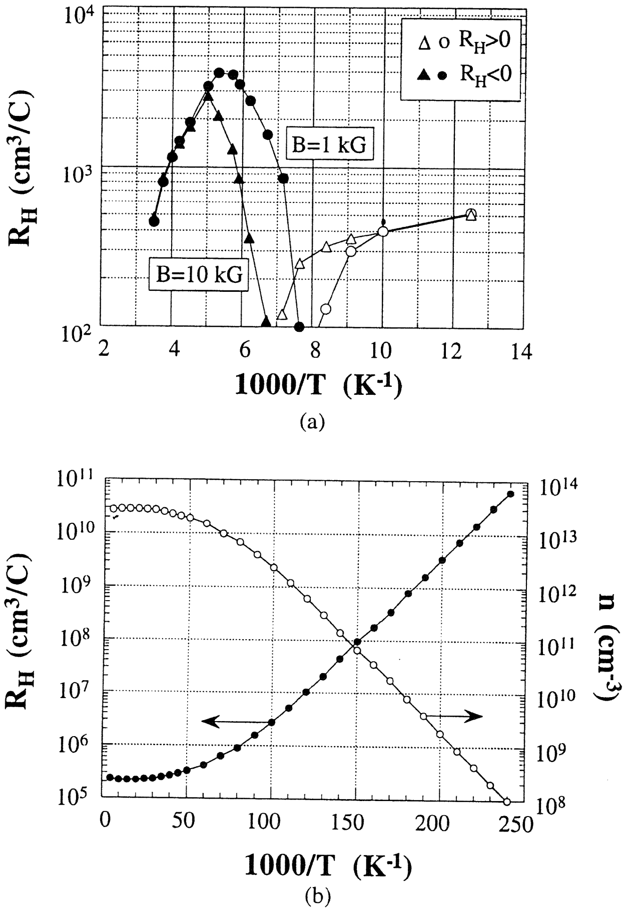
$$\boxed{R_H = \frac{(p - b^2 n) + (\mu_n B)^2 (p - n)}{q\left[(p + bn)^2 + (\mu_n B)^2 (p - n)^2\right]}} \tag{8.13}$$

This expression is relatively complex and depends sensitively on the mobility ratio $b = \mu_n/\mu_p$ and on the magnetic field strength B. In the limit of low and high magnetic field strength, the Hall coefficient becomes

$$B \Rightarrow 0: R_H = \frac{(p - b^2 n)}{q(p + bn)^2}; \qquad B \Rightarrow \infty: R_H = \frac{1}{q(p - n)} \tag{8.14}$$

For Eq. (8.14) to hold in the low field limit, $B \ll 1/\mu_n$ for $p \gg n$ and $B \ll 1/\mu_p$ for $p \ll n$. For a mobility of 1000 cm^2/V $\cdot$ s = 0.1 m^2/V $\cdot$ s this requires $B \ll 10$ T. For mobilities of 10^5 cm^2/V $\cdot$ s, this requirement becomes more severe, with $B \ll 0.1$ T. The high-field limit requires $B \gg 1/\mu_n$ for $p \gg n$ and $B \gg 1/\mu_p$ for $p \ll n$. Hence magnetic fields much larger than 10 T or 0.1 T, respectively, are necessary in this example.

For semiconductors with modest mobilities in the 100 to 1000 cm^2/V $\cdot$ s range and with mobility ratios of $b \approx 3$ to 10, the Hall coefficient is generally found to vary little with magnetic field and Eq. (8.14) with $B \Rightarrow \infty$ or Eqs. (8.11) to (8.12) are used. However, for those semiconductors with high mobilities and high b, the Hall coefficient is found to vary with magnetic field. In addition, the Hall coefficient changes sign as a function of temperature. Such behavior is found in semiconductors like HgCdTe. An example of such a Hall coefficient variation is shown in Fig. 8.2(a) for a n-type HgCdTe



(a)

(b)

Fig. 8.2 (a) Temperature and magnetic field dependent Hall coefficient for HgCdTe showing typical mixed conduction behavior. Reprinted with permission after Zemel et al., Ref. 7. (b) Hall coefficient and electron density for GaAs adapted from Stillman and Wolfe, Ref. 8.

to 300 K, with $n = n_i^2/p \gg p$, because n_i^2 is high for narrow band gap materials. $R_H = -1/qn$ in this temperature range, and it is independent of B. For $T \approx 100$ to 200 K, holes begin to participate and mixed conduction causes R_H to decrease and be magnetic field dependent. Hole conduction dominates at lower temperatures. The Hall coefficient becomes positive and is magnetic field independent. This figure exhibits the temperature and magnetic field dependent behavior of mixed conduction very nicely. Figure 8.2(b) shows the Hall coefficient for GaAs, with neither magnetic field dependence or mixed conduction.[8] Also shown is the electron density derived from the Hall coefficient using Eq. (8.12).

Equations (8.11) to (8.14) are derived under simplifying assumptions of energy-independent scattering mechanisms. With this assumption relaxed, the expressions for the hole and electron densities become[5, 6]

$$p = \frac{r}{qR_H}; \qquad n = -\frac{r}{qR_H} \tag{8.15}$$

where r is the Hall scattering factor, defined by $r = \langle\tau^2\rangle/\langle\tau\rangle^2$, with τ the mean time between carrier collisions. The scattering factor depends on the type of scattering mechanism in the semiconductor and generally lies between 1 and 2. For lattice scattering, $r = 3\pi/8 = 1.18$; for impurity scattering $r = 315\pi/512 = 1.93$, and for neutral impurity scattering $r = 1$.[6, 10] The scattering factor is also a function of magnetic field and temperature; r can be determined by measuring R_H in the high magnetic field limit; that is, $r = R_H(B)/R_H(B = \infty)$. In the high field limit $r \to 1$. The scattering factor has been measured in n-type GaAs as a function of magnetic field and was found to vary from 1.17 at $B = 0.01$ T, as expected from lattice scattering, to 1.006 at $B = 83$ kG.[11] The high fields necessary for r to approach unity are difficult to achieve, and $r > 1$ for most Hall measurements. Typical magnetic fields used for Hall measurements lie between 0.05 and 1 T.

The *Hall mobility* μ_H is defined by

$$\mu_H = \frac{|R_H|}{\rho} = |R_H|\sigma \tag{8.16}$$

The Hall mobility is not identical to the conductivity mobility. Substituting Eq. (8.1) into Eq. (8.16) gives

$$\mu_H = r\mu_p; \qquad \mu_H = r\mu_n \tag{8.17}$$

for extrinsic p- and n-type semiconductors, respectively. Hall mobilities can differ significantly from conductivity mobilities since r is generally larger than unity. For most Hall-determined mobilities, r is taken as unity, but this assumption should be carefully specified.

The schematic Hall sample of Fig. 8.1 has a variety of practical implementations. One of these is the geometry of Fig. 8.3(a). It is, in principle, identical to Fig. 8.1, but has four "legs" for making the voltage contacts and is known as a bridge-type Hall bar. The current flows into 1 and out of 4, the Hall voltage is measured between 2 and 6 or between 3 and 5 in the presence of a magnetic field. The resistivity is determined in the absence of the magnetic field by measuring the voltage between 2 and 3 or between 6 and 5. The equations developed above apply for this geometry. Hall bars are frequently cut out of a wafer with ultrasonic cutting too

A more general geometry is the irregularly shaped sample in Fig. 8.3(b)

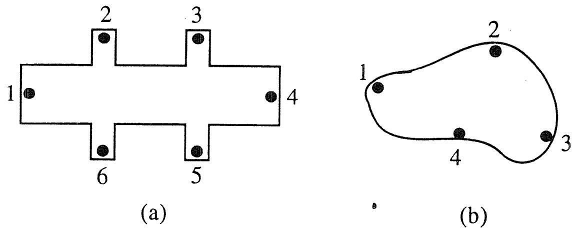


Fig. 8.3 (a) Bridge-type Hall sample, (b) lamella-type van der Pauw Hall sample.

shaped samples is based on conformal mapping developed by van der Pauw.[12, 13] He showed how the resistivity, carrier density, and mobility of a flat sample of arbitrary shape can be determined without knowing the current pattern if the following conditions are met: (1) the contacts are at the circumference of the sample, (2) the contacts are sufficiently small, (3) the sample is uniformly thick, and (4) the sample surface is singly connected, i.e., the sample does not contain isolated holes.

For the sample of Fig. 8.3(b) the resistivity is given by[12]

$$\boxed{\rho = \frac{\pi t}{\ln(2)} \frac{(R_{12,34} + R_{23,41})}{2} F} \tag{8.18}$$

where $R_{12,34} = V_{34}/I$. The current I enters the sample through contact 1 and leaves through contact 2 and $V_{34} = V_4 - V_3$ is the voltage between contacts 4 and 3. $R_{23,41}$ is similarly defined. Current enters the sample through two adjacent terminals and the voltage is measured across the other two adjacent terminals. F is a function of the ratio $R_r = R_{12,34}/R_{23,41}$ only, satisfying the relation

$$\frac{R_r - 1}{R_r + 1} = \frac{F}{\ln(2)} \operatorname{arc\,cosh}\left(\frac{\exp(\ln(2)/F)}{2}\right) \tag{8.19}$$

and is plotted in Fig. 8.4. For symmetric samples (circles or squares) $F = 1$. Most van der Pauw samples are symmetric.

The van der Pauw Hall mobility is determined by measuring the resistance $R_{24,13}$ with and without a magnetic field. $R_{24,13}$ is measured by forcing the current into one and out of the opposite terminal, e.g., terminals 2 and 4 in Fig. 8.3, with the voltage measured across terminals 1 and 3. The Hall mobility is then given by

$$\mu_H = \frac{d \Delta R_{24,13}}{B \rho} \tag{8.20}$$

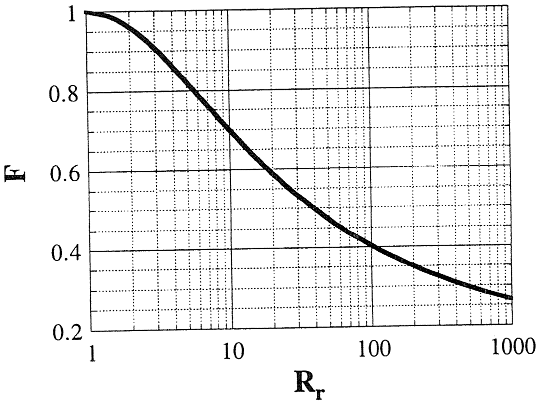


Fig. 8.4 The van der Pauw F factor plotted against R_r.

These equations are for carrier densities per unit volume and for resistivity ρ (ohm · cm). Occasionally it is useful to determine carrier densities per unit area and sheet resistance ρ_s (ohms/square). For uniformly doped samples of thickness d, the *sheet Hall coefficient* R_{HS} is defined as

$$R_{HS} = \frac{R_H}{d} \tag{8.21}$$

and

$$\mu_H = \frac{|R_{HS}|}{\rho_s} \tag{8.22}$$

where $\rho_s = \rho/d$.

The thickness is well defined for bulk samples. For thin layers on substrates of opposite conductivity or on semi-insulating substrates, the active film thickness is not necessarily the total film thickness. If depletion effects caused by Fermi level pinned band bending or surface charges and by band bending at the layer-substrate interface are not considered, the Hall coefficient can be in error as will those semiconductor parameters derived from it.[14, 15] For sufficiently lightly doped films it is possible for the surface-induced space-charge region to deplete the entire film. Hall effect measurements then indicate a semi-insulating film. For semiconducting films on insulating substrates, the mobility is frequently observed to decrease toward the substrate. Surface depletion forces the current to flow in the low-mobility portion of the film, giving apparent mobilities that are lower than actual mobilities.[15] Even the temperature dependence of the surface and interface

8.3.2 Nonuniform Layers

Hall effect measurements are simple to interpret for uniformly doped samples. Nonuniformly doped layer measurements are more difficult to interpret. If the doping density varies with film thickness, then its resistivity and mobility also vary with thickness. A Hall effect measurement gives the *average* resistivity, carrier density, and mobility. For spatially varying mobility $\mu_p(x)$ and carrier density $p(x)$, the Hall sheet coefficient R_{Hs}, the sheet resistance ρ_s, and the average Hall mobility $\langle \mu_H \rangle$ for a p-type film of thickness d are given by[17, 18]

$$R_{HS} = \frac{\int_0^d p(x)\mu_p^2(x)\,dx}{q\left(\int_0^d p(x)\mu_p(x)\,dx\right)^2} \tag{8.23}$$

$$\rho_s = \frac{1}{q\int_0^d p(x)\mu_p(x)\,dx} \tag{8.24}$$

$$\langle \mu_H \rangle = \frac{\int_0^d p(x)\mu_p^2(x)\,dx}{\int_0^d p(x)\mu_p(x)\,dx} \tag{8.25}$$

assuming $r = 1$. x specifies the distance into the sample.

To determine resistivity and mobility *profiles*, Hall measurements must be made as a function of film thickness. The film thickness is varied by two major techniques: removing thin portions of the film by etching and measuring the Hall coefficient repeatedly, and making portions of the film electrically inactive by a reverse-biased space-charge region.

In principle, one can use chemical etching to remove thin layers of the film to be profiled. In practice, it is difficult to remove thin layers reproducibly by chemical etching. The electrochemical profiler, discussed in more detail in Section 2.2.4, has been successfully used to remove thin layers by electrolytic etching of GaAs in Tiron (1,2dihydroxybenzene-3,5disulphonic acid, disodium salt in an aqueous solution).[19] Hall effect measurements are made after each etch. A more common method for reliable layer removal is anodic oxidation and subsequent oxide etch.[17, 18, 20–24] Anodic oxidation consumes a fraction of the semiconductor during oxidation. When the oxide is subsequently etched, that portion of the semiconductor consumed during oxidation is also removed. This method provides for very reproducible semiconductor removal and, w[illegible]oom temperature oxidation, it does not alter the doping profile. A more detailed discussion of anodic oxidation is given in

A second method utilizes a junction formed on the upper surface of the film to be profiled. The film must be sufficiently thin for the reverse-biased space-charge region to be able to deplete it completely, or at least deplete most of its total thickness. This implies that the layer to be profiled must be bounded at its lower surface by an insulator or a junction. The upper junction may be a pn junction, a Schottky barrier junction, or an MOS capacitor. An example, shown in Fig. 8.5, consists of a p-layer on an insulator. The layer is provided with a Schottky gate. The zero-biased metal-semiconductor junction induces a space-charge region of width W under the metal. The insulator could be replaced by a semi-insulating substrate or by an n-type substrate. The square sample is laterally isolated by etching but could be isolated by surrounding it with an n-type film. Four contacts provide for current and voltage probes. When ion-implanted samples are to be profiled, they are usually formed by defining the appropriate Hall sample shape photolithographicly. A p-type implant is then automatically isolated from the n-type substrate by the resulting pn junction.

Van der Pauw measurements provide information on the undepleted film of thickness $d - W$, where d is the total film thickness. A single measurement gives the mobility, the resistivity, and the carrier density averaged over $d - W$. When the Schottky barrier junction is reverse biased, its space-charge region extends into the film, reducing the width of the neutral portion of the film. By measuring the Hall effect as a function of reverse-bias voltage, one can determine mobility, resistivity, and carrier density profiles of the underlying layer. This method has been implemented with MOSFETs for thin Si films on sapphire,[25–27] for Si-on-insulator with the insulator formed by an implanted oxide layer,[28] and with Schottky diodes for GaAs on semi-insulating substrates.[29, 30] A comparison of the destructive "anodize-etch-measure" with the "gated" technique has shown the "gated" method to give more reliable mobilities and to have higher spatial resolution.[31]

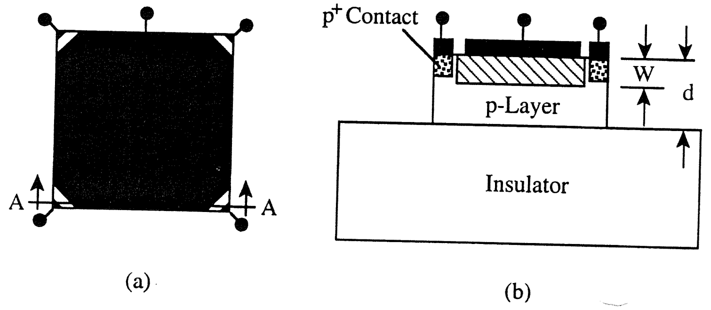


Fig. 8.5 Schottky-gated thin film van der Pauw sample

The spatially varying Hall mobility is determined from the spatially varying sheet Hall coefficient and sheet conductance $\sigma_s = 1/\rho_s$ by the relationship[17, 18, 32]

$$\mu_H = \frac{d(R_{HS}\sigma_s^2)/dx}{d\sigma_s/dx} \tag{8.26}$$

and the spatially varying carrier density is

$$p(x) = \frac{r}{q}\frac{(d\sigma_s/dx)^2}{d(R_{HS}\sigma_s^2)/dx} \tag{8.27}$$

Equations (8.26) and (8.27) are useful when R_{HS} versus x and σ_s versus x curves have been generated.

The mobility and carrier density profiles can also be determined after each layer removal step by making Hall measurements and using the Hall measured values of adjacent layers in the calculations. This is the differential Hall effect (DHE) discussed in more detail in Chapter 1. The average values of mobility and carrier density may differ from the true values if there are large inhomogeneities in the sample. To reduce this effect, it is necessary to make Δx_i, where Δx_i is the thickness of the ith layer, small to approximate the nonuniform film by a uniform film. For ion-implanted and fully annealed samples with no mobility anomalies, the error between the measured and real values in mobility and carrier density is less than 1% if $\Delta x_i < 0.5\Delta R_p$, where ΔR_p is the standard deviation of the implanted profile.[33] A density profile of a boron layer implanted into Si is shown in Fig. 8.6 where the Hall measured profile is compared with the profile determined by secondary ion mass spectrometry and spreading resistance profiling.[34]

Difficulties can arise when there are large mobility variations through the film. Consider a film consisting of two layers of equal thickness. The upper layer has a carrier density of P_1 holes/cm^2 with mobility μ_1 and the lower one has P_2 holes/cm^2 and μ_2.[35] The total hole density is $P_1 + P_2$. The Hall effect measures weighted averages given by[18]

$$P = \frac{(P_1\mu_1 + P_2\mu_2)^2}{P_1\mu_1^2 + P_2\mu_2^2} \tag{8.28}$$

$$\mu_H = \frac{P_1\mu_1^2 + P_2\mu_2^2}{P_1\mu_1 + P_2\mu_2} \tag{8.29}$$

Here, P will be significantly less than $(P_1 + P_2)$ and μ_H will lie between μ_1 and μ_2 for $P_1 > P_2$ and $P_1\mu_1^2 < P_2\mu_2^2$. For example, for $P_1 = 10P_2$ and

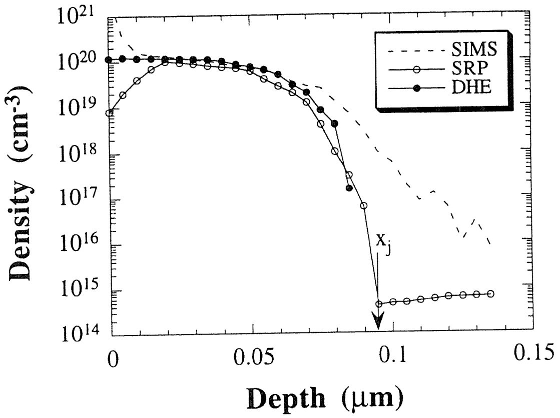


Fig. 8.6 Dopant density profiles determined by DHE, spreading resistance profiling and secondary ion mass spectrometry. Data after Ref. 34. Reprinted from the Januar 1993 edition of *Solid State Technology*. Copyright 1993 by Penn Well Publishin Company.

it is possible for the mobility to be higher than the expected bulk mobility One cause of abnormally high mobilities is the inclusion of metallic precipi tates in the crystal. A thorough discussion of this effect has been given b Wolfe and Stillman.[36]

8.3.3 Multilayers

The discussion in the previous section dealt with the measurement nonuniform films on an "inert" substrate. By "inert" we mean a substrat that does not contribute to the measurement, e.g., an insulating substrate. semi-insulating substrate also approximates this situation and for most pract cal purposes can be considered insulating. A p-film on an n-substrate or a n-film on a p-substrate might be thought to be in the same category, with th space-charge region (scr) between two semiconductors of opposite conducti ity considered an insulating boundary. But this is a more precarious situatio For example, a leaky junction can no longer be considered an insulator. Eve if the insulating properties of the scr are sufficiently good, there may leakage paths along the surface. Or, even worse, the heavily doped contac may be diffused into the substrate, providing a leakage path. Film character zation is then no longer unique to the film, and the substrate properties a reflected in the measurements.

This problem was originally addressed by Nedoluha and Koch[37] and Petritz.[38] Petritz considered a substrate whose surface is inverted by surfa charges. For example, an n-type inversion layer may be formed on a p-ty two-layer interacting configuration was later extended.[39, 40] For a simple two-layer structure with an upper layer having thickness d_1 and conductivity σ_1 and a substrate of thickness d_2 and conductivity σ_2, the Hall constant is given by[37]

$$R_H = \frac{d\left[\left(R_{H1}\sigma_1^2 d_1 + R_{H2}\sigma_2^2 d_2\right) + R_{H1}\sigma_1^2 R_{H2}\sigma_2^2\left(R_{H1}d_2 + R_{H2}d_1\right)B^2\right]}{\left(\sigma_1 d_1 + \sigma_2 d_2\right)^2 + \sigma_1^2\sigma_2^2\left(R_{H1}d_2 + R_{H2}d_1\right)^2 B^2} \tag{8.30}$$

which becomes[38, 40]

$$R_H = \frac{d\left(R_{H1}\sigma_1^2 d_1 + R_{H2}\sigma_2^2 d_2\right)}{\left(\sigma_1 d_1 + \sigma_2 d_2\right)^2} = R_{H1}\frac{d_1}{d}\left(\frac{\sigma_1}{\sigma}\right)^2 + R_{H2}\frac{d_2}{d}\left(\frac{\sigma_2}{\sigma}\right)^2 \tag{8.31}$$

in the low magnetic field limit, and

$$R_H = \frac{R_{H1}R_{H2}d}{R_{H1}d_2 + R_{H2}d_1} \tag{8.32}$$

in the high magnetic field limit. In these equations R_{H1} is the Hall constant of layer 1, R_{H2} is the Hall constant of substrate 2, $d = d_1 + d_2$ and σ is given by

$$\sigma = \frac{d_1}{d}\sigma_1 + \frac{d_2}{d}\sigma_2 \tag{8.33}$$

The magnetic field dependence of Eq. (8.30) can be used to advantage to gain additional information by measuring the Hall coefficient as a function of magnetic field. This is illustrated in Fig. 8.7 for a sample consisting of a p-substrate and an n-layer where $R_{H1} = -1/qn_1$ and $R_{H2} = 1/qp_2$. The Hall coefficients are of opposite sign, making it possible for the measured Hall coefficient to reverse its sign with magnetic field. The Hall coefficient is plotted against the $n_1 d_1$ product. For low $n_1 d_1$ the Hall coefficient is dominated by the p-substrate and is magnetic field independent. Both p_2 and μ_2 can be determined from R_H. For intermediate values of $n_1 d_1$, the Hall coefficient becomes field dependent. Conduction is initially dominated by holes, and then by electrons as the Hall coefficient changes its sign. The carrier density and mobility of both the n-layer and the p-substrate can be deduced from an analysis of the field-dependent R_H using the two-layer model of Eq. (8.30). For high $n_1 d_1$ values, the Hall coefficient is negative, conduction is dominated by the n-layer and R_H becomes again magnetic field independent. A good discussion can be found in Zemel et al.[7]

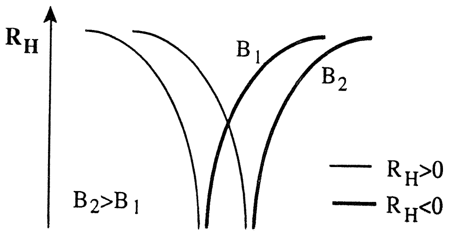


Fig. 8.7 Hall coefficient of a p-type

For $d_1 = 0$, we have $d = d_2$, $\sigma = \sigma_2$ and $R_H = R_{H2}$ with the substrate being characterized. If the upper layer is more heavily doped than the substrate or is formed by inversion through surface states, for example, and the carriers in the substrate freeze out at low temperatures making σ_2 very small, then

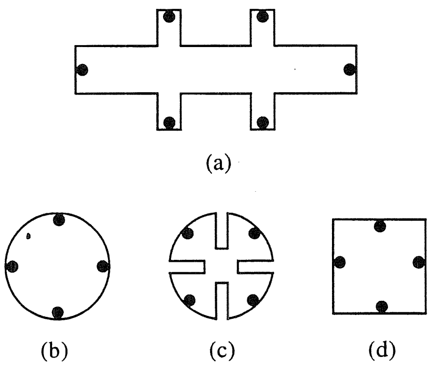
$$\sigma \approx \frac{d_1}{d}\sigma_1; \qquad R \approx R_{H1}\frac{d}{d_1} \tag{8.34}$$

and the Hall measurement characterizes the surface layer. This problem can be especially serious if the existence of the upper layer is not suspected, and it is believed that the substrate is being characterized. Since both conductivity and the Hall coefficient can be erroneous, the resulting Hall mobility will be in error. Examples of an n-type skin on a p-type bulk, an n-type film on p-type bulk, and an n-type skin on n-type bulk are given for HgCdTe and InSb.[40, 41]

8.3.4 Sample Shapes and Measurement Circuits

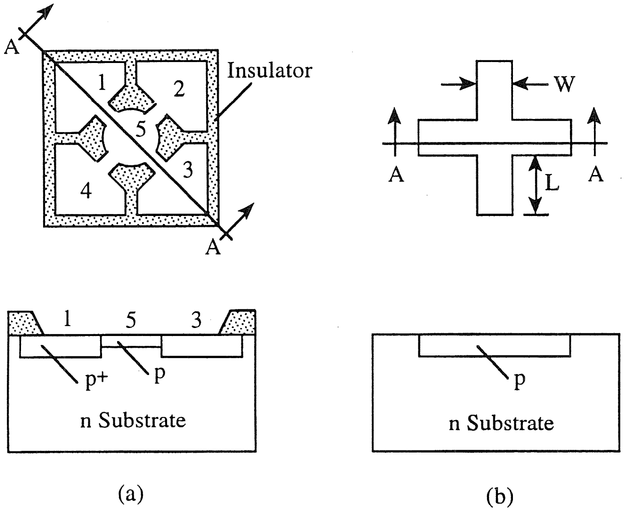
Hall samples come in two basic geometries: bridge type and lamella type. The parallelepiped sample shape of Fig. 8.1 is not recommended because contacts have to be directly soldered to the sample. To ease the contact problem, the Hall bridge has extended arms as shown in Fig. 8.8(a).[42] Both six- and eight-arm geometries can be used. The dimensions are given in ASTM Standard F76.[42] The lamellar specimen may be of arbitrary shape, but a symmetrical configuration is preferred. The sample must be free of geometrical holes; typical shapes are shown in Fig. 8.8(b) to (d). Whereas the bridge-type specimen have extensions for contact placement, the lamella-type specimen may have no such provision. It is important for the contacts to be small and to be placed as close to the periphery as possible.

A few of the common lamella or van der Pauw shapes are shown in Fig. 8.9. The shapes beyond simple circles or squares are usually fabricated by photolithographic methods where it becomes possible provide contact extensions. During the early days of ion implantation development, implant


(a)

(b) (c) (d)

Fig. 8.8 (a) Bridge-type Hall configuration, (b)–(d) lamella-type Hall configuration.



(a) (b)

Fig. 8.9 van der Pauw Hall sample shapes.

contact diffusions 1 to 4. The implant can then be done over the entire wafer, provided the oxide is sufficiently thick to mask against the implant. The area to be measured is region 5. Any implant into the *p*-diffused regions is of no consequence. A transfer length contact resistance test pattern has also been used for Hall measurements. In addition to the contact resistance, specific contact resistance, and sheet resistance, the mobility in the implanted layer and under the contacts, as well as the sheet carrier density, were extracted by applying a magnetic field.[43]

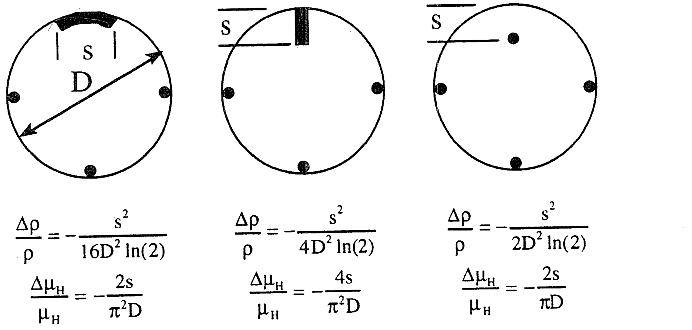
The size and placement of the contacts is important. For van der Pauw samples, the contacts should be point contacts located symmetrically on the periphery. This is not achievable in practice, and some error is introduced

---

samples with contacts spaced at 90° intervals. The contacts are equipotenti areas; three cases are shown in Fig. 8.10. In each case there are three ide contacts, with the fourth being nonideal. The fourth contact is either length s and larger than a point contact or is a point contact displaced distance s from the periphery. Also indicated for each geometry is t relative error in resistivity $\Delta\rho/\rho$ and in mobility $\Delta\mu_H/\mu_H$ introduced by t nonideal contact, valid for small s/D and low $\mu_H B$. The errors are additi to first order if more than one contact is nonideal. Nonideal contact effe are reduced by removing the contacts from the "active" area. One impleme tation is the use of some form of cloverleaf geometry, shown in Fig. 8.8 and Fig. 8.9. The errors due to displaced contacts on square specimen a discussed in Refs. 44 and 45. The placement of the contacts on squa samples is better at the midpoint of the sides than at the corners.[44] T Greek Cross in Fig. 8.9(b) makes use of this type of geometry, where $L \geq 1.02W$ less than 0.1% error is introduced.[45] For square samples w sides of length L having square and triangular contacts of contact length δ the four corners, less than 10% error was introduced for Hall measureme as long as $\delta/L < 0.1$.[46]

The contacts need not be exactly opposite one another, since the magne field reversal routinely made during Hall measurements tends to cancel unbalanced voltage. But for an unbalanced voltage higher than the H voltage, the Hall voltage is the difference of two large numbers, and err are likely to be introduced.

Some samples use a geometry close to that of a conventional semicond tor device. For example, a MOSFET fabricated in a thin film on an insulat substrate has the general shape of the Hall sample in Fig. 8.11, where the regions 1 and 2 are the source and the drain and 7 is the gate. The con regions 3–6 are added for Hall measurements. The Hall voltage is develo between contacts 3 and 4 and 5 and 6. However, the sample is shorted at



$$\frac{\Delta\rho}{\rho} = -\frac{s^2}{16D^2\ln(2)} \qquad \frac{\Delta\mu_H}{\mu_H} = -\frac{2s}{\pi^2 D}$$

$$\frac{\Delta\rho}{\rho} = -\frac{s^2}{4D^2\ln(2)} \qquad \frac{\Delta\mu_H}{\mu_H} = -\frac{4s}{\pi^2 D}$$

$$\frac{\Delta\rho}{\rho} = -\frac{s^2}{2D^2\ln(2)} \qquad \frac{\Delta\mu_H}{\mu_H} = -\frac{2s}{\pi D}$$

Fig. 8.10 Effect of nonideal contact length or contact placement on the resis and mobility for van der Pauw samples. Reprinted with permission from var

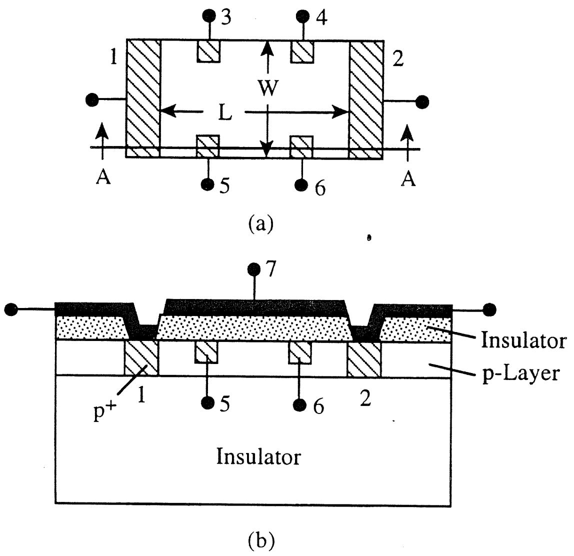


Fig. 8.11 Hall sample with electrically shorted regions at the ends; (a) top view with the gate not shown, (b) cross section along cut A–A.

ends by the source and the drain. This has a significant influence on the interpretation of the measured Hall voltage V_{Hm}. For $L \leq 3W$, V_{Hm} is less than the Hall voltage for samples with $L > 3W$. The Hall voltage V_H for sample dimensions of $L \gg W$ used in the earlier equations in this chapter is related to the measured Hall voltage for short samples by $V_H = V_{Hm}/G$, where G is shown in Fig. 8.12(b).[47] The curves in Fig. 8.12(b) are calculated for the Hall voltage measured across the sample at $x = L/2$. Note that for sample lengths $L \geq 3W$, the shorting effect is negligible, and the measured voltage is the usual Hall voltage.

For a detailed discussion of the measurement procedure and for measurement precautions see ASTM Standard F76.[42] The current and the magnetic field are reversed and the readings averaged for more accurate measurements. Many Hall systems are now computerized, but the principle of the measurement is no different. Special precautions are necessary when the specimen resistance is very high to eliminate current leakage paths and sample loading by the voltmeter. The *guarded* approach utilizes high input impedance unity gain amplifiers between each probe on the sample and the external circuitry.[48] The unity gain outputs drive the shields on the leads between the amplifier and the sample to reduce leakage currents and system time constant by effectively eliminating the stray capacitance in the leads. Measurements [illegible] resistances up to 10^{12} ohms have been made with such a system, and the *guarded* approach has also been automated.[49] Measurements

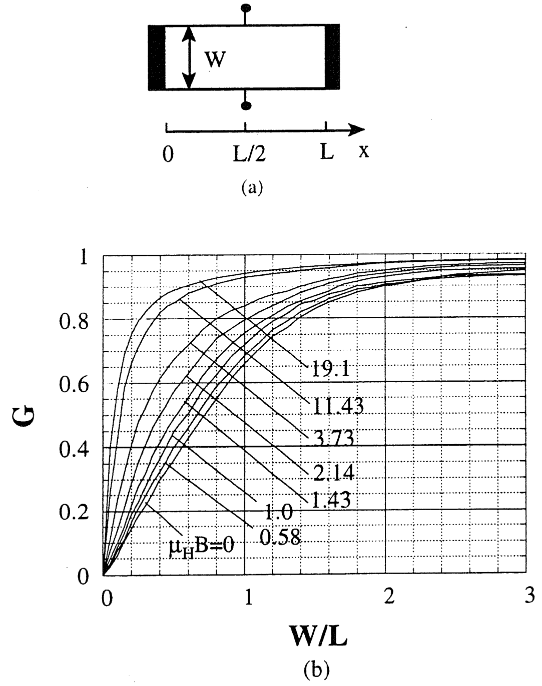


Fig. 8.12 (a) Hall sample with electrically shorted end regions, (b) ratio of measured voltage V_{Hm} to Hall voltages V_H. $G = V_{Hm}/V_H$. Reprinted with permission after Lippmann and Kuhrt, Ref. 47.

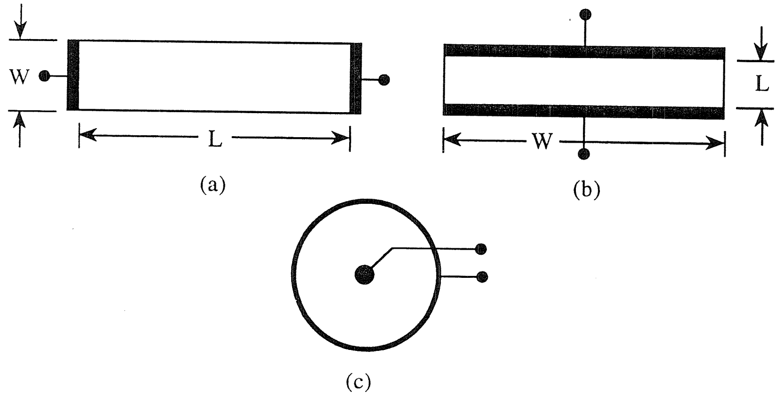
"dark" wafer and introducing a dark spot within the illuminated slit.[50, 51] A resistance measurement along the slit determines essentially the resistance of the small dark spot since the dark spot resistance is much larger than the resistance of the illuminated strip. A resistance map can be obtained by moving the dark spot.

Hall effect profiling measurements have other possible errors. For example, the bottom *pn* junction may be leaky, causing smaller Hall voltages than would be measured for perfect isolation of the film from the substrate. The upper junction in a Schottky contact configuration may also be leaky. Junction leakage currents can be reduced by sample cooling[21] but that may not always be possible. If the upper junction is forward biased to reduce the space-charge region width in order to be able to profile closer to the surface, considerable error is introduced due to the high forward-biased junction current.[29] Although the effect of injected gate current can be corrected,[52] the correction is large and the accuracy of the corrected result may be questionable. Instead of conventional dc measurement circuits, [illegible] circuits can be employed.[29, 30] The device is driven with an ac current at one frequency and a component of a frequency different from the current. The appropriate ac voltages are measured with a lock-in amplifier without interference from the dc leakage current. In one implementation the magnetic field and the current frequencies were 60 Hz and 200 Hz, respectively.[53] The Hall voltage is detected by a lock-in amplifier at the sum frequency of 260 Hz. This eliminates most thermoelectric and thermomagnetic errors associated with dc measurements allowing Hall voltages as low as 10 μV to be measured.

8.4 MAGNETORESISTANCE MOBILITY

Typical Hall-effect structures are either long or of the van der Pauw variety. They require four or more contacts. A long Hall bar is shown schematically in Fig. 8.13(a) with $L \gg W$. Certain semiconductor devices, such as field-effect transistors (FETs) are short with $L \ll W$, shown in Fig. 8.13(b). The Hall electric field, resulting from an applied magnetic field, is nearly shorted by the long contacts and FET structures do not lend themselves well to Hall measurements. The extreme of this short geometry is when one contact is in the center of a circular sample and the other contact is at the periphery, shown in Fig. 8.13(c). The Hall electric field in this Corbino disk[54] is shorted, and no Hall voltage exists. The geometries of Fig. 8.13(b) and (c), however, lend themselves well to *magnetoresistance effect* measurements.

The *resistivity* of a semiconductor generally increases when the sample is placed in a magnetic field. This is known as the *physical magnetoresistance effect* (PMR). It occurs if the conduction is anisotropic, if conduction involves more than one type of carrier, and if carrier scattering is energy dependent. The *resistance* of a semiconductor is also influenced by magnetic fields.[55] The magnetic field causes the path of the charge carriers to deviate from a straight line, raising the sample resistance. This depends on the sample



geometry and is known as the *geometrical magnetoresistance* (GMR). T resistance change as a result of the magnetic field is due to resistivity chang of the semiconductor as well as geometrical effects and is larger the high the sample mobility is. Geometric effects usually dominate. For example, GaAs at room temperature and in a magnetic field of 1 T, the PMR is abo 2%, whereas the GMR is about 50%. The geometric magnetoresistan mobility μ_{GMR} is related to the Hall mobility μ_H by

$$\mu_{GMR} = \xi \mu_H \qquad (8.3$$

where ξ is the magnetoresistance scattering factor given by ξ $(\langle\tau^3\rangle\langle\tau\rangle/\langle\tau^2\rangle^2)^2$.[10] For τ independent of energy, the mean time betwee collisions becomes isotropic, $\xi = 1$ and $\mu_{GMR} = \mu_H$. The physical magn toresistivity change ratio $\Delta\rho_{PMR} = (\rho_B - \rho_0)/\rho_0$ becomes zero under tho conditions, where ρ_B is the resistivity in the presence and ρ_0 in the absen of a magnetic field.

The dependence of the resistance ratio R_B/R_0 is shown in Fig. 8.14 as function of $\mu_{GMR}B$ for rectangular samples of varying L/W ratios.[56] He R_B is the resistance with $B \neq 0$ and R_0 is the resistance with $B = 0$. F long rectangular samples with contacts at the ends of the long sample as Fig. 8.13(a), the ratio is near unity and the magnetoresistance effect is ve small. The ratio is higher for short, wide samples. The highest ratio obtained for the Corbino disk with $L/W = 0$. Figure 8.14 shows the magn toresistance and the Hall effect to be complementary. When one decreas the other increases. For example, we showed in Fig. 8.12 a Hall volta reduction for short, wide samples. But those same sample shapes produ

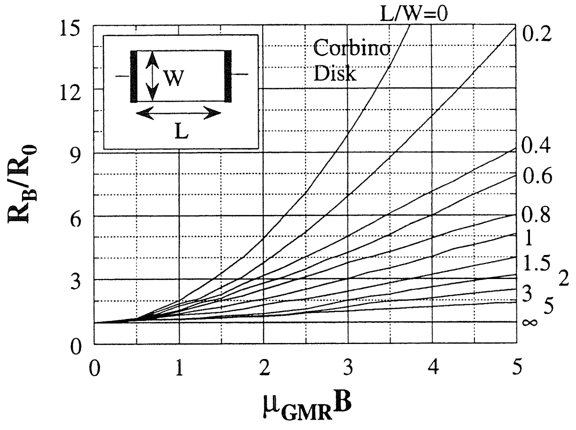


Fig. 8.14 Geometric magnetoresistance ratio of rectangular samples versus μ_{GM} as a function of the length-width ratio. Reprinted with permission after Lippma

maximum magnetoresistance. Magnetoresistance measurements are suitable for field-effect transistors that are short and wide. The current flow in a Corbino disk is radial from the center to the periphery for $B = 0$. With a magnetic field perpendicular to the sample, the current streamlines become logarithmic spirals and the resistance ratio becomes

$$\frac{R_B}{R_0} = \frac{\rho_B}{\rho_0}\left[1 + (\mu_{GMR} B)^2\right] \tag{8.36}$$

Equation (8.36) represents the Corbino disk's curve in Fig. 8.14. Generally, the magnetoresistance scattering factor ξ is taken as unity just as the Hall scattering factor is generally taken to be unity. This is done for simplicity and because the scattering mechanisms are not known precisely. Measurements of μ_{GMR} on a modified Corbino disk geometry and of μ_H on Hall samples from the Corbino disk showed ξ to be unity for GaAs within experimental error.[57, 58] The measurements were performed for magnetic fields up to 0.7 T and temperatures from 77 to 400 K.[58] Under those conditions $\rho_B \approx \rho_0$ and $\mu_{GMR} \approx \mu_H$. Making the additional assumption of $\mu_H \approx \mu_p$, the mobility is given by

$$\mu_p \approx \frac{1}{B}\sqrt{\frac{R_B}{R_0} - 1} \tag{8.37}$$

The mobility is obtained from the slope of a plot of $(R_B/R_0 - 1)^{1/2}$ versus B. The mobility can be profiled by using a Corbino disk with a Schottky gate and measuring the resistance as a function of the gate voltage.[59]

The use of Corbino disks is inconvenient because of its special geometrical configuration. However, as is evident from Fig. 8.14, rectangular sample shapes with low L/W ratios are equally suitable for magnetoresistance measurements.[60] For rectangular samples with low L/W ratios and $\mu_{GMR} B < 1$, Eq. (8.36) is replaced by[56, 57]

$$\frac{R_B}{R_0} = \frac{\rho_B}{\rho_0}\left[1 + (\mu_{GMR} B)^2(1 - 0.54L/W)\right] \tag{8.38}$$

If the error in the determination of μ_{GMR} is to be less than 10%, then the aspect ratio L/W must be less than 0.4. For typical FET structures with $L/W \ll 1$, Eq. (8.38) is a close approximation to Eq. (8.36), and it is for that reason that Eq. (8.36) is generally used in GMR measurements. Magnetoresistance measurements were first used for GaAs Gunn effect devices.[57, 61] It is a rapid technique that can be used for functional devices, requiring no special test structures. Instead of measuring the resistance as a function of the FET gate voltage, it is also possible to determine the mobility from

The magnetoresistance mobility measurement method has been applied
metal-semiconductor FETs (MESFETs) as well as to modulation-dop
FETs (MODFETs). By using the magnetic field dependence of the GM
effect, it is possible to extract the mobilities of the various conducting regio
and subbands in MODFETs.[63] The method has been used to determine t
mobility dependence on gate electric field.[64] Effects of gate currents f
Schottky-gate devices and series resistance effects must be corrected.[52, 62]
Gate current corrections are particularly important when the gate becom
forward biased. Contact resistance, which is of only secondary importance f
Hall measurements, is very important for GMR measurements because
adds to the measured resistance and contact resistance is relatively indepe
dent of magnetic field. When the mobility is measured as a function of ga
bias, the average mobility is measured for each value of gate voltage. Bo
the *average* and the *differential mobilities* can be determined from transco
ductance measurements.[66]

The GMR effect is not universally applicable the way the Hall effect
shown by Eq. (8.36). Assuming that $\rho_B/\rho_0 \approx 1$, which is a reasona
assumption, we find that in order to observe a resistance change, $\Delta R/R$
$(R_B - R_0)/R_0$ of, say, 10%, the condition $\mu_{GMR} \geq 0.3/B$ must be met.
typical magnetic fields of 0.1 to 1 T, this requires $\mu_{GMR} \geq 30{,}000$ to 3
$cm^2/V \cdot s$. These are the kinds of mobilities found in MESFETs and MO
FETs made in III-V materials, especially at low temperatures. These are
very materials that have been successfully characterized by GMR. For hig
magnetic fields, as obtained in superconducting magnets, lower mobilities
be determined. Silicon, whose mobility lies in the 500–1300 $cm^2/V \cdot s$ ra
is unsuitable for magnetoresistance measurements because its GMR is ne
gibly small for typical laboratory magnet fields.

8.5 TIME-OF-FLIGHT DRIFT MOBILITY

The *time-of-flight* method to determine the mobility of minority carriers
first demonstrated in the well-known *Haynes-Shockley experiment*.[67–69]
first comprehensive mobility measurements for Ge and Si were made
this technique by Prince.[70] The principle of the method is demonstrated
the *p*-type semiconductor bar in Fig. 8.15(a). A drift voltage $-V_{dr}$ prod
an electric field $\mathscr{E} = V_{dr}/L$ along the bar. Minority electrons are injecte
negative polarity pulses at the *n*-emitter. The injected electron packet
from the emitter to the collector in the applied electric field to be colle
by the collector.

The electrons are injected as a narrow pulse at $t = 0$. For $t > 0$
diffuse and recombine with majority holes as they drift along the
Consequently, the minority carrier pulse broadens due to diffusion an
area decreases by recombination. The pulse shape is given as a functio

space and time by the expression[71]

$$\Delta n(x,t) = \Delta n(x,0)\exp\left(-\frac{(x-vt)^2}{4D_n t} - \frac{t}{\tau_n}\right)$$

$$= \frac{N}{\sqrt{4\pi D_n t}}\exp\left(-\frac{(x-vt)^2}{4D_n t} - \frac{t}{\tau_n}\right) \quad (8.39)$$

where N is the electron density (electrons/cm^2) in the packet at $t = 0$ at the point of injection. The first term in the exponent describes diffusion and drift, and the second term describes recombination.

The time for the electron packet to drift from emitter to collector is given by $t_d = d/v$, where d is the spacing between contacts shown in Fig. 8.15(a), and $v = \mu_n \varepsilon$ the electron packet velocity. The normalized output voltage waveform according to Eq. (8.39) is shown in Fig. 8.15(b) along with data points from Ref. 72. Calculated output voltages are shown in Figs. 8.15(c) and (d) as a function of spacing d and as a function of lifetime τ_n. Note the area decrease and pulse broadening with time in (c) and the pulse amplitude dependence on lifetime in (d).

The delay time t_d is determined by measuring the output pulses versus time for varying amplitude input pulses and extrapolating to zero injection. Alternately, the injection pulse amplitude can be reduced until the peak position of the output pulse no longer shifts in time. This ensures low-level injection with the injected carrier density well below the majority carrier equilibrium density, eliminating any local disturbance of the electric field by the minority carrier pulse.

With the velocity given by $v = \mu_n \mathscr{E}$, the drift mobility is determined from the relationship

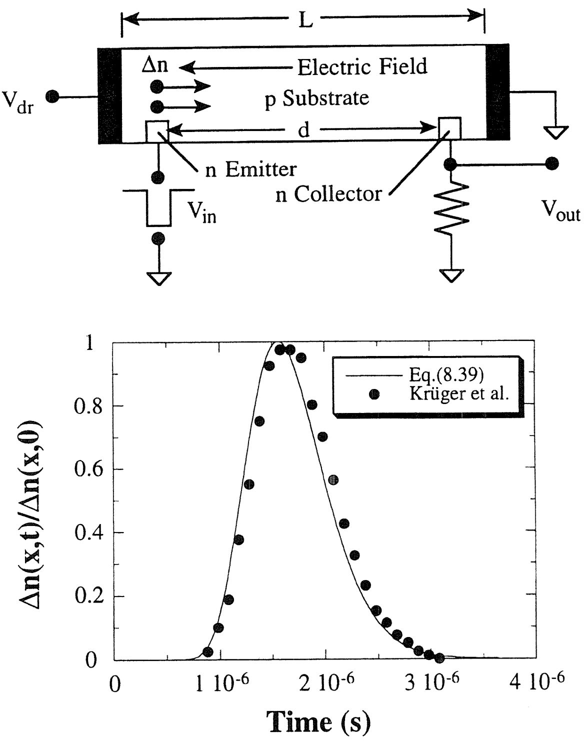
$$\mu_n = \frac{d}{t_d \mathscr{E}} \quad (8.40)$$

The time-of-flight method actually measures the *minority carrier velocity* or the *minority carrier mobility*. It is therefore useful for the determination of the carrier velocity-electric field behavior. This relationship is difficult to determine with other mobility measurement techniques.

To determine the diffusion constant D_n, the collected pulse width is measured at half its maximum amplitude. It can be shown (see Problem 8.9) that D_n is given by

$$D_n = \frac{(d\Delta t)^2}{16\ln(2)t_d^3} \quad (8.41)$$

where Δt is the pulse width



(a)

Fig. 8.15 (a) Drift mobility measurement arrangement and normalized output vol
age pulse (μ_p = 180 cm^2/V · s, τ_n = 0.67 μs, T = 423 K, $\mathscr{E}$ = 60 V/cm).

The lifetime is determined by measuring the collected electron packe
pulse at times t_{d1} and t_{d2}, corresponding to the two drift voltages, V_{dr1} an
V_{dr2}. In the ideal case with no minority carrier trapping, the collected puls
has the predicted Gaussian shape and the lifetime is obtained by comparin
the corresponding output pulse amplitudes V_{01} and V_{02}. The electron lifetim
is then[72]

$$\tau_n = \frac{t_{d2} - t_{d1}}{\ln(V_{01}/V_{02}) - 0.5\ln(t_{d2}/t_{d1})} \quad (8.4$$

If carrier trapping is present, $\log(A_p)$ should be plotted against the del
time t_d, where A_p is the pulse area. The slope of such a plot is $-1/\tau_n$. O
can also fit theory to experiment and determine μ_n, D_n and τ_n.[72]

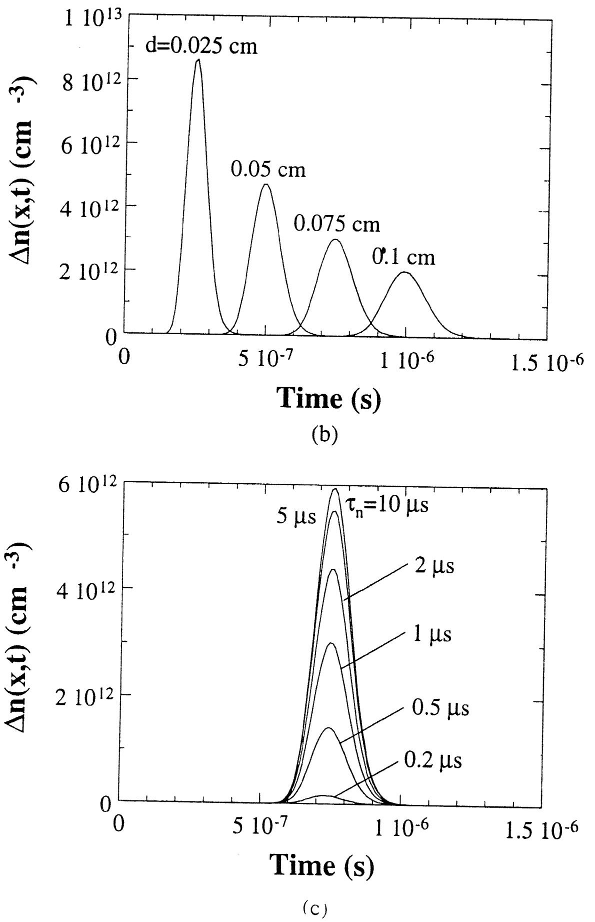


Fig. 8.15 (b) output voltage pulses (μ_n = 1000 cm^2/V·s, τ_n = 1 μs, T = 300 K, $\mathcal{E}$ = 100 V/cm, N = [illegible] cm^{-2}), (c) output voltage pulses (μ_n = 1000 cm^2/V·s, d = 0.075 cm, T = 300 K, $\mathcal{E}$ = 100 V/cm, N = [illegible] cm^{-2}).

[illegible] injection with the [illegible] is followed by electrical detection. [illegible] diffuse and are [illegible]

The ehps can also be created by a pulsed electron beam[76] or by placing the sample into a microwave circuit. In the latter case the electron beam is deflected at microwave frequencies across the sample, and the resulting microwave current is detected.[77, 78] The drift velocity is determined from the amplitude and phase of the microwave current.[77, 79] The region between emitter and collector can be oxidized and provided with a gate, if there is concern that carriers recombine at the surface.[72] An appropriate gate voltage biases the surface into accumulation, effectively reducing surface recombination. Surface recombination has been discussed in Chapter 7.

Mobility or carrier velocity measurements as a function of electric field are not difficult with the circuit of Fig. 8.15(a) at *low* electric fields. Measurements at *high* electric fields are more difficult due to sample heating when high voltages, applied to the semiconductor sample, generate very high currents. To overcome this limitation, a configuration is employed in which the electric field is developed in the scr of a reverse-biased junction. The reverse-biased junction ensures low current at high electric field. The drift of carriers through such a high electric field region is more suitable for determining the drift velocity than the mobility as a function of the electric field, however.

We demonstrate the principle of the method in Fig. 8.16(a), consisting of two parallel plates. Voltage $-V_1$ is applied to the cathode. Electrons, liberated at the cathode by UV light, for example, drift with velocity v_n from the cathode to the anode in the electric field generated by V_1. The electron charge $Q_N = qN$ C/cm^2 induces charges Q_C and Q_A in the cathode and anode, respectively, with $Q_N = Q_C + Q_A$. The arrows represent electric field lines from Q_C and Q_A terminating on Q_N. The electric field lines due to the applied voltage are not shown.

The charge on both plates redistributes itself continuously as the charge between the plates drifts from the cathode to the anode. The anode charge is $Q_A = 0$ at $t = 0$ and $Q_A = Q_N$ at $t = t_t$, where t_t is the transit time defined by

$$t_t = \frac{W}{v_n} \qquad (8.43)$$

[illegible] changes from zero to Q_N, the charge flows through the external [illegible] current $I(t)$, given by

$$[illegible] = \frac{Q_N A}{t_t} = \frac{Q_N A v_n}{W}, \qquad 0 \le t \le t_t \qquad (8.44a)$$

$$[illegible] \qquad (8.44b)$$

[illegible] only during [illegible] transit time

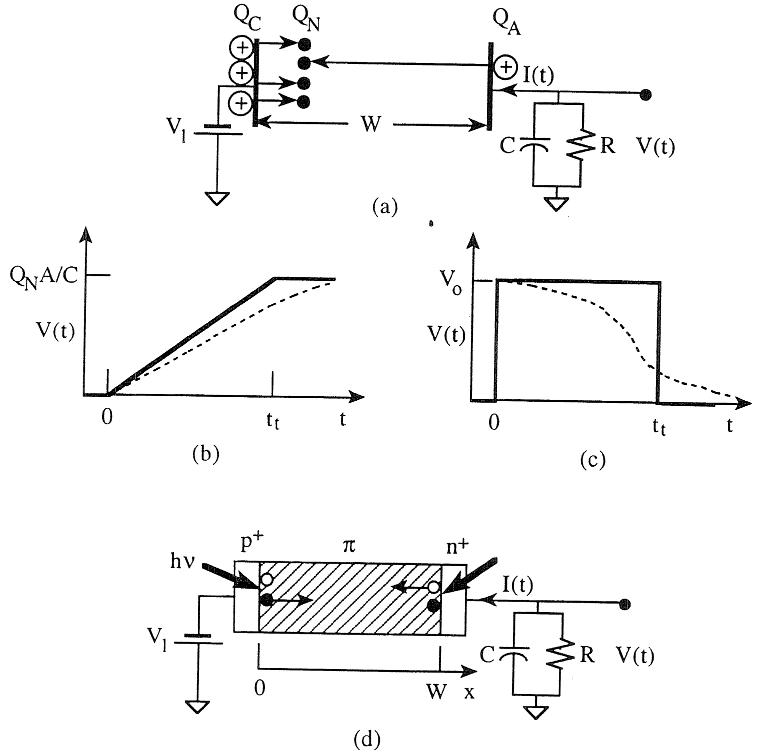


Fig. 8.16 (a) Time-of-flight measurement schematic, (b) output voltage for $t_t \ll RC$, (c) output voltage for $t_t \gg RC$, (d) implementation with a $p^+\pi n^+$ diode.

The sample, connecting leads, and input to the voltage-sensing circuit all contain capacitances. These are all lumped into C. R is the load resistance in Fig. 8.16(a). The output voltage is

$$V(t) = \frac{Q_N A v_n R}{W}(1 - e^{-t/RC}) = V_0(1 - e^{-t/RC}) \qquad (8.45)$$

EXERCISE 8.2

Problem

Derive Eq. (8.45).

Solution

In the frequency domain we have

Taking the Laplace transform gives

$$V(s) = Z(s)I(s) = \frac{R}{1 + sRC}I(s) = \frac{R}{s(1 + sRC)}\frac{Q_N A v_n}{W}$$

using a step current of $I(s) = I/s = (Q_N A v_n/W)(1/s)$, where "$s$" is the Laplace operator. Taking the inverse Laplace transform gives

$$V(t) = \frac{Q_N A v_n R}{W}(1 - e^{-t/RC}) = V_0(1 - e^{-t/RC})$$

Equation (8.45) has two limits that are of interest for transit time measurements.

1. For $t_t \ll RC$, the voltage becomes

$$V(t) \approx \frac{V_0 t}{RC} = \frac{Q_N A v_n t}{WC}, \qquad 0 \le t \le t_t \qquad (8.46a)$$

$$V(t) = \frac{Q_N A}{C}, \qquad t > t_t \qquad (8.46b)$$

In this approximation, the RC circuit acts as an integrator, and the resulting voltage is shown in Fig. 8.16(b) by the solid line.

2. For $t_t \gg RC$, the voltage becomes

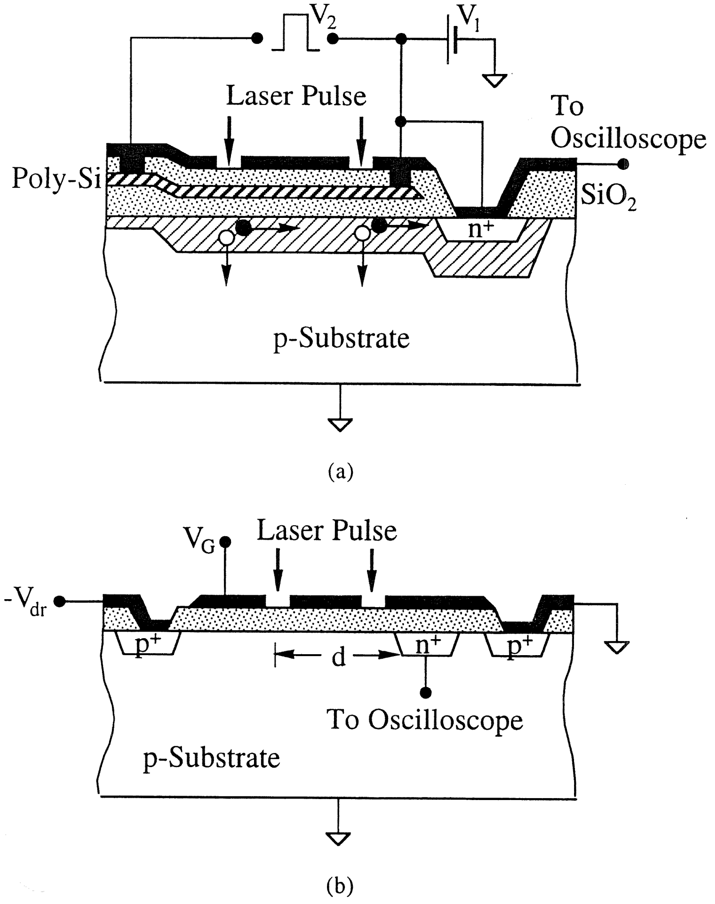
$$V(t) \approx V_0 = \frac{Q_N A v_n R}{W}, \qquad 0 \le t \le t_t \qquad (8.47a)$$

$$V(t) = 0, \qquad t > t_t \qquad (8.47b)$$

The RC time constant in this approximation is so small that the capacitor never charges and $V(t) \approx RI(t)$. The resulting voltage is shown in Fig. 8.16(c) by the solid line. The transit time can be determined for either case and the carrier velocity is extracted from t_t.

This time-of-flight method can be implemented with the $p^+\pi n^+$ junction in Fig. 8.16(d). The π region is a lightly doped p-region in this figure. Bias voltage V_1 depletes the π region entirely. Shallow penetration excitation (high energy light or an electron beam) from the *left* creates ehps near $x = 0$. The holes flow into the p^+ contact layer and the electrons drift to $x = W$, allowing the *electron velocity* to be determined. With excitation from the *right*, holes drift to the left, and the *hole velocity* is measured. This test structure can be used for both kinds of carriers. There is, of course, a dc dark generated, time-varying current I_1. Obviously, I_1 must be higher than I_{dk} to be able to detect the output voltage.

Two slightly different implementations of time-of-flight measurement geometries are shown in Fig. 8.17. Both use *pn* diodes combined with MOS structures. Figure 8.17(a) shows a gate-controlled diode with both diode and gate biased to V_1.[82, 83] This ensures deep depletion under the gate so that an inversion layer cannot form. The gate of the gate-controlled diode consists of a high-resistivity polysilicon film whose sheet resistance is around 10 kohms/square. The voltage pulse V_2 with 200 ns pulse length and 10 kHz repetition rate creates a periodic voltage along the gate as well as along the semiconductor. This voltage in turn generates a lateral electric field in the semiconductor. Optical pulses, from a mode-locked Nd:YAG laser, are directed to two openings defined in a metal gate. The absorbed photons



(a)

(b)

create electron-hole pairs in the semiconductor. The holes drift into th substrate and the electrons drift along the surface to the collecting diode t produce a current pulse in the output circuit. By injecting minority carrie into two locations, defined by optical apertures, the difference in arriv times is used to determine the drift velocity. To obtain the field dependenc of the mobility, the *lateral* or *tangential* electric field is varied by changing V To determine the gate voltage dependence of the mobility, the *normal* (*vertical* electric field is varied by adjusting V_1. A detailed description of th technique is given in Cooper and Nelson.[82]

The electric field in the semiconductor in Fig. 8.17(b) is obtained not fro a voltage drop along a polysilicon gate but from a voltage applied betwee two p^+ contacts in the semiconductor itself.[84] The Al gate is used to set th surface potential, but the lateral electric field is independent of the vertic electric field since the lateral field does not originate from a gate voltag The continuous gate is also a light shield with two slits for the laser pulses generate ehps. Optical pulses with 70 ps pulse widths from a mode-locke Nd-YAG laser have been used. The minority carrier packets are collected b the n^+ collector and displayed on a sampling oscilloscope. The circuits Figs. 8.15 and 8.17 are, in principle, very similar. The chief difference lies the method of minority carrier injection. In Fig. 8.15 minority carriers a injected electrically, and in Fig. 8.17 they are injected optically. In all of the techniques, it is important for carrier trapping to be either eliminated accounted for in the data analysis.[75, 85] The dashed lines in Fig. 8.16(b) ar (c) indicate the effects of trapping.[81]

The saturation velocity can also be determined from MOSFET curren voltage data. For short channel MOSFETs ($L_m \leq 1$ μm—L_m is the mas defined gate length), the drain current under saturation conditions in th presence of source resistance R_S, can be written as[86]

$$I_{D,\text{sat}} = \frac{W_{\text{eff}} v_{\text{sat}} \mu_{\text{eff}} C_{\text{ox}} (V_{GS} - V_T - I_{D,\text{sat}} R_S)^2}{2 v_{\text{sat}} L_{\text{eff}} + \mu_{\text{eff}} (V_{GS} - V_T - I_{D,\text{sat}} R_S)} \qquad (8.4$$

Solving for $I_{D,\text{sat}}$ and dropping higher order $I_{D,\text{sat}}$ terms, allows Eq. (8.48) be written as

$$\frac{1}{I_{D,\text{sat}}} = \frac{2R_S W_{\text{eff}} v_{\text{sat}} C_{\text{ox}} + 1}{W_{\text{eff}} v_{\text{sat}} C_{\text{ox}} (V_{GS} - V_T)} + \frac{2(L_m - \Delta L)}{W_{\text{eff}} \mu_{\text{eff}} C_{\text{ox}} (V_{GS} - V_T)^2} \qquad (8.4$$

A plot of $1/I_{D,\text{sat}}$ versus L_m has intercepts $(1/I_{D,\text{sat}})_{\text{int}}$ and $L_{m,\text{int}}$ given

$$\left(\frac{1}{I_{D,\text{sat}}}\right)_{\text{int}} = \frac{2R_S W_{\text{eff}} v_{\text{sat}} C_{\text{ox}} + 1}{W_{\text{eff}} v_{\text{sat}} C_{\text{ox}} (V_{GS} - V_T)} - \frac{2\Delta L}{W_{\text{eff}} \mu_{\text{eff}} C_{\text{ox}} (V_{GS} - V_T)^2} \qquad (8.5$$

$$L_{m,\text{int}} = \Delta L - \frac{\mu_{\text{eff}} (V_{GS} - V_T)(2R_S W_{\text{eff}} v_{\text{sat}} C \quad 1)}{2} \qquad (8.5$$

Substituting Eq. (8.50) into (8.51) gives

$$L_{\mathrm{m,int}} = \Delta L + \frac{2R_S W_{\mathrm{eff}} \upsilon_{\mathrm{sat}} C_{\mathrm{ox}} + 1}{W_{\mathrm{eff}} \upsilon_{\mathrm{sat}} C_{\mathrm{ox}} (V_{\mathrm{GS}} - V_T)} \frac{L_{\mathrm{m,int}}}{(1/I_{\mathrm{D,sat}})_{\mathrm{int}}} = \Delta L + A \frac{L_{\mathrm{m,int}}}{(1/I_{\mathrm{D,sat}})_{\mathrm{int}}} \tag{8.52}$$

Note that Eq. (8.52) no longer contains μ_{eff}. Plotting $L_{\mathrm{m,int}}$ versus $L_{\mathrm{m,int}}/(1/I_{\mathrm{D,sat}})_{\mathrm{int}}$ has the slope A. Plotting A versus $1/(V_{\mathrm{GS}} - V_T)$ gives a line with slope S, which leads to υ_{sat} through the expression

$$\upsilon_{\mathrm{sat}} = \frac{1}{W_{\mathrm{eff}} C_{\mathrm{ox}} (S - 2R_S)} \tag{8.53}$$

8.6 MOSFET MOBILITY

The conductivity, Hall, and magnetoresistance mobilities are *bulk* mobilities. Surfaces play a relatively minor role in their determination. The carriers are free to move throughout the sample and a mobility, averaged over the sample thickness, is measured. The main scattering events determining the mobility are *lattice* or *phonon scattering* and *ionized impurity scattering*. *Neutral impurity scattering* becomes important at low temperatures, where ionized impurities become neutral due to carrier freeze-out. For some semiconductors there is additionally *piezoelectric scattering*. Each scattering mechanism is associated with a mobility. According to Mathiessen's rule, the net mobility μ depends on the various mobilities as[87]

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \cdots \tag{8.54}$$

and the lowest mobility dominates.

In this section we are concerned with additional scattering mechanisms that occur when the current carriers are confined within a narrow region as in an inversion layer in a MOSFET. The location of the carriers at the oxide-semiconductor interface introduces additional scattering mechanisms like Coulomb scattering from oxide charges and interface states, as well as surface roughness scattering. These additional scattering sources reduce the MOSFET mobility below the bulk mobility.[88] Quantization of carriers in inversion layers further reduces the mobility.[89, 91]

8.6.1 Effective Mobility

The MOSFET drain current is due to drift and diffusion of the carriers in the

The considerations for p-channel devices are similar. The drain current I_D can be written as

$$I_D = \frac{W \mu_{\mathrm{eff}} Q_n V_{\mathrm{DS}}}{L} - W \mu_{\mathrm{eff}} \frac{kT}{q} \frac{dQ_n}{dx} \tag{8.55}$$

where Q_n is the mobile channel charge density (C/cm^2), and μ_{eff} the effective mobility. The effective mobility is measured at low drain voltage typically 50–100 mV. However a lower V_{DS} is better for μ_{eff} determination. For low V_{DS}, one can assume the channel charge to be fairly uniform from source to drain, allowing the diffusive second term in Eq. (8.55) to be dropped. Solving for μ_{eff} then gives

$$\boxed{\mu_{\mathrm{eff}} = \frac{g_d L}{W Q_n}} \tag{8.56}$$

where the drain conductance g_d is defined as

$$g_d = \left. \frac{\partial I_D}{\partial V_{\mathrm{DS}}} \right|_{V_{\mathrm{GS}} = \mathrm{constant}} \tag{8.57}$$

How is Q_n to be determined? Two approaches are commonly used. In the first, the mobile channel charge density is approximated by

$$Q_n = C_{\mathrm{ox}} (V_{\mathrm{GS}} - V_T) \tag{8.58}$$

When μ_{eff} is determined with Eqs. (8.56) and (8.58), one usually observes significant drop of mobility near $V_{\mathrm{GS}} = V_T$. The reason for this is twofold First, Eq. (8.58) is only an approximation to the true value of Q_n, and second the threshold voltage is not well known. As discussed in Chapter 4, the threshold voltage is not uniquely defined. To use Eq. (8.58), for $t_{\mathrm{ox}} \approx 10$ nm V_G must exceed V_T by about 0.5 V for an error of less than 10%. For thinne oxides, the error increases for a given V_G above V_T.

The approach giving better results is based on a direct measure of Q from capacitance measurements, with the mobile channel density determine from the gate-to-channel capacitance/unit area C_{GC} according to

$$Q_n = \int_{-\infty}^{V_{\mathrm{GS}}} C_{\mathrm{GC}}\, dV_{\mathrm{GS}} \tag{8.59}$$

Then C_{GC} is measured using the connection of Fig. 8.18(a). The capacitanc meter is connected between the gate and the source-drain connected to

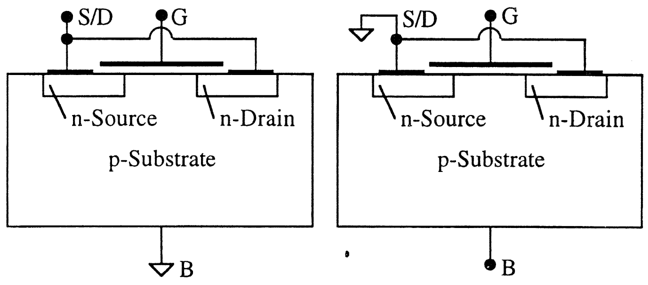


Fig. 8.18 Configuration for (a) gate-to-channel, (b) gate-to-substrate capacitance measurements.

Appendix 6.1. What is measured in this connection is shown in more detail in Fig. 8.19. For negative gate voltage [Fig. 8.19(a)], the channel region is accumulated and the overlap capacitances $2C_{ov}$ are measured. For $V_{GS} > V_T$ [Fig. 8.19(b)], the surface is inverted and all three capacitances, $2C_{ov} + C_{ch}$, are measured. A typical $C_{GC} - V_{GS}$ curve is shown in Fig. 8.20(a). Subtracting $2C_{ov}$ from this curve and integrating gives the $Q_n - V_{GS}$ curve of Fig. 8.20(a). Figure 8.20(b) gives the drain output characteristics. These curves give the drain conductance g_d from the slope at low V_{DS}. Extracting the mobility from Fig. 8.20 through Eq. (8.56) gives the mobility shown in Fig. 8.21.

Even if the mobility is determined with Eqs. (8.56) and (8.59), there are still some sources of error that are usually ignored. Nevertheless, we will briefly mention them. C_{GC} is most commonly measured as shown in Fig. 8.18(a). Clearly in this configuration, $V_{DS} = 0$, but the drain current is obviously measured with $V_{DS} \neq 0$. It is very common to use $V_{DS} = 100$ mV for I_D measurements. A much better choice is to use as small a drain voltage as possible. However, if V_{DS} is too low, the measurement becomes noisy, but $V_{DS} \approx 20$–50 mV is reasonable. $V_{DS} \neq 0$ introduces an error in $\mu_{eff} - V_{GS}$ data, primarily near $V_{GS} = V_T$, because Q_n reduces as V_{DS} is increased for a given V_{GS}.[92-94] Modifying the measurement circuit slightly allows a drain bias to be applied during the C_{GC} measurement, with the capacitance measured

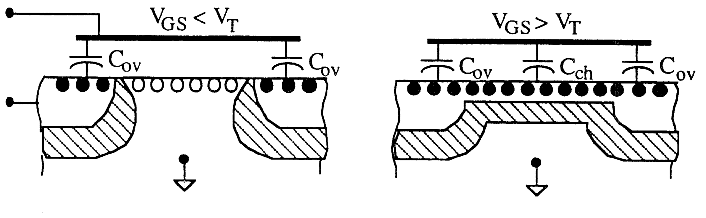


Fig. 8.19 Surface conditions for gate-to-channel capacitance measurements for (a)

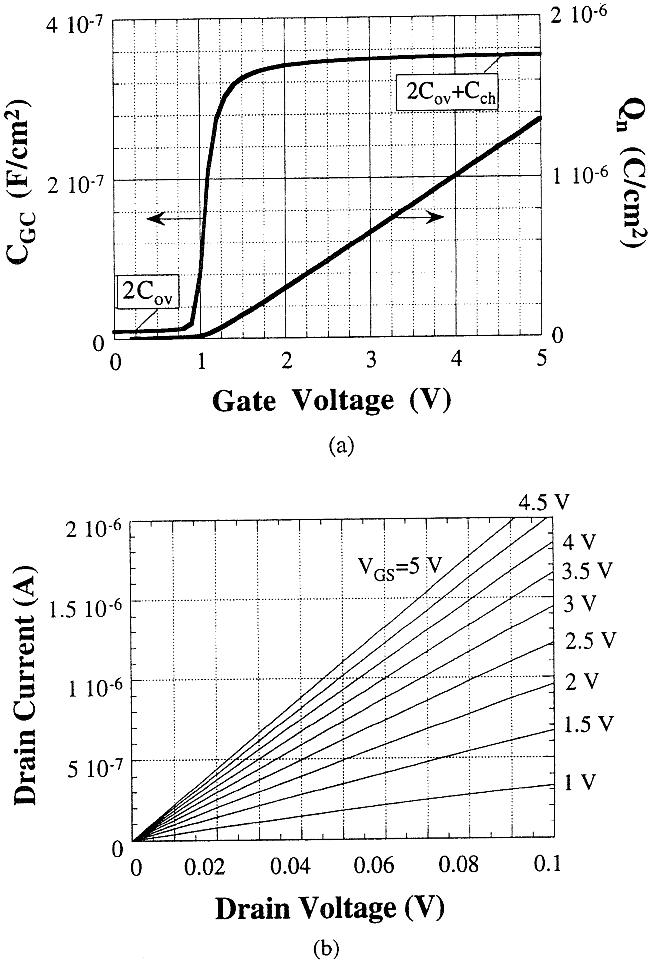


Fig. 8.20 (a) C_{GC} and Q_n versus V_{GS}; (b) I_D versus V_{DS}. $W/L = 10\ \mu m/$ μm, $t_{ox} = 10$ nm, $N_A = 1.6 \times 10^{17}$ cm^{-3}.

between G and S (C_{GS}) with the drain reverse biased.[94] Then the G to capacitance (C_{GD}) is measured. C_{GC} is $C_{GS} + C_{GD}$. Another error is tl neglect of the overlap capacitances C_{ov} in Fig. 8.19, although it may l permissible to neglect these capacitances for large MOSFETs with ga lengths of 100 μm or so. Nevertheless some error is introduced if the capacitances are not considered in the analysis.

A further error is introduced by assuming the drain current to be dr current only, i.e., neglecting the diffusion term dQ_n/dx in Eq. (8.55). Whi this may be a good approximation for operation ab threshold, for V near V_T, diffusion current begins to be important. In fact, as is well know