

Metal/Semiconductor Ohmic Contacts

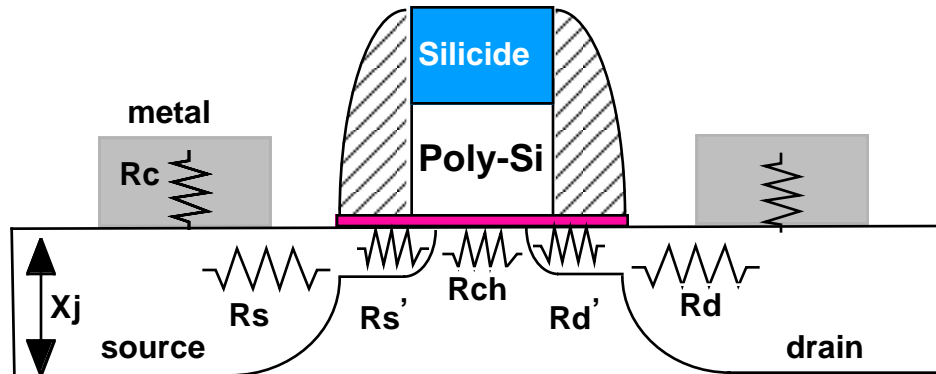


Fig. 1 components of the resistance associated with the S/D junctions of a MOS transistor.

| Year | 1997 | 1999 | 2003 | 2006 | 2009 | 2012 |
|-----------------------|------------|------------|------------|------------|------------|------------|
| Min Feature Size | 0.25 μ | 0.18 μ | 0.13 μ | 0.10 μ | 0.07 μ | 0.05 μ |
| Contact x_j (nm) | 100-200 | 70-140 | 50-100 | 40-80 | 15-30 | 10-20 |
| x_j at Channel (nm) | 50-100 | 36-72 | 26-52 | 20-40 | 15-30 | 10-20 |

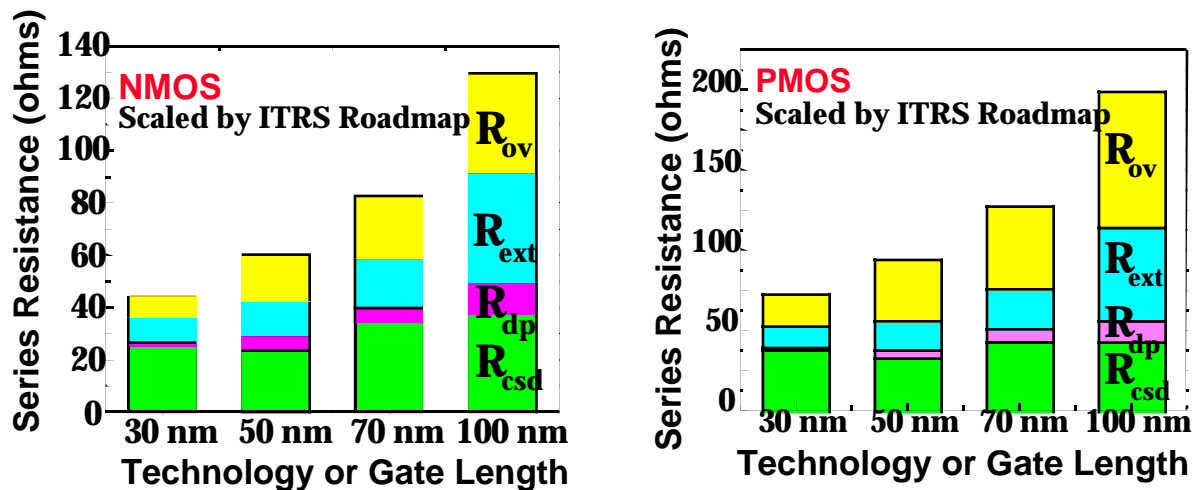


Fig. 2. Various components of the resistance associated with the shallow junctions of NMOS and PMOS transistors for different technology nodes. (Source: Jason Woo, UCLA)

Conduction Mechanisms for Metal/Semiconductor Contacts

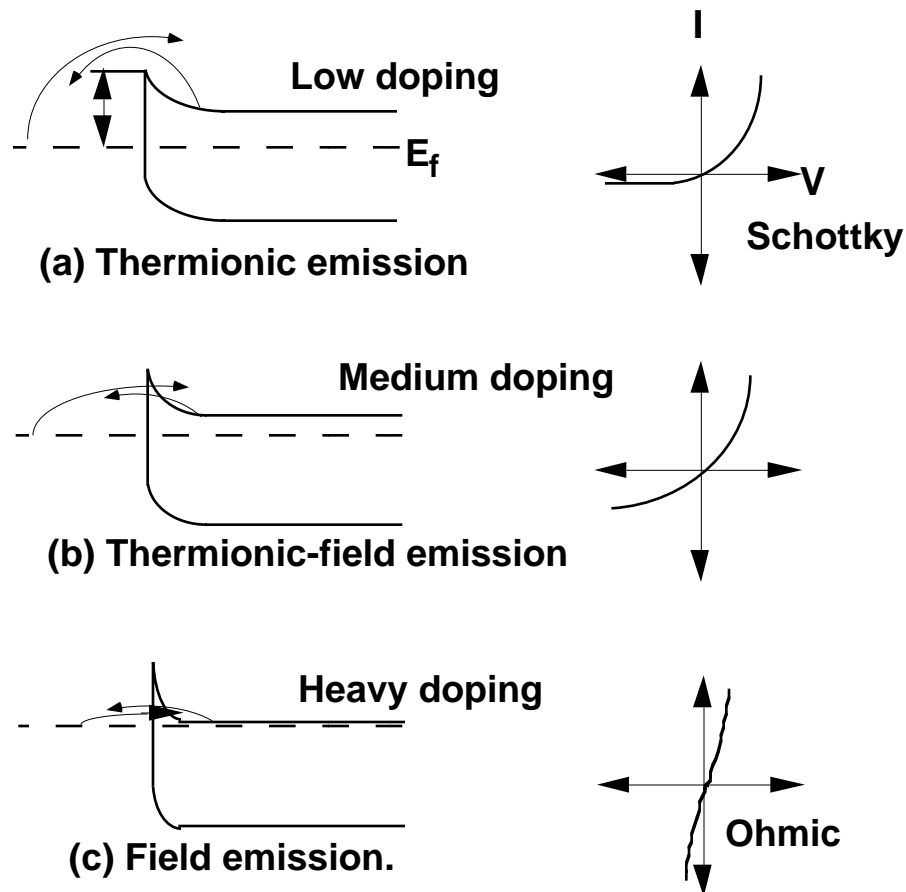


Fig. 3. Conduction mechanisms for metal/n-semiconductor contacts as a function of the barrier height and width. (a) Thermionic emission; (b) thermionic-field emission; (c) field emission.

(1) Thermionic emission (TE), occurring in the case of a depletion region so wide that the only way for electrons to jump the potential barrier is by emission over its maximum (Fig. 3a). The barrier height is reduced from its original value as a result of image force barrier lowering.

(2) Field emission (FE), consisting in carrier tunneling through the potential barrier. This mechanism, which is the preferred transport mode in ohmic contacts, takes place when the depletion layer is sufficiently narrow, as a consequence of the high doping concentration in the semiconductor (Fig. 3c).

Contact Resistance and Specific Contact Resistivity (ρ_c)

Contact resistance is a measure of the ease with which current can flow across a metal-semiconductor interface. In an ohmic interface, the total current density J entering the interface is a function of the difference in the equilibrium Fermi levels on the two sides.

The band diagram in the Fig. 4 may be used as an aid in describing the majority current flow in the block of uniformly heavily doped semiconductor material of length l with ohmic contacts at each end. The applied voltage V drives a spatially uniform current I through the semiconductor bulk and ohmic contacts of cross sectional area A . Then, under the low-current

assumption that the voltage drop across both metal-semiconductor contacts is identical, the I-V relation becomes:

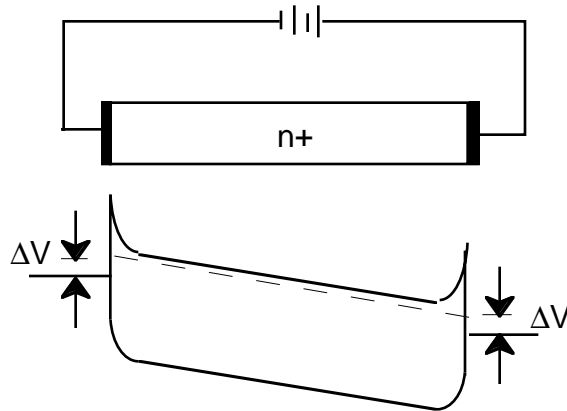


Figure 4: Ideal contacts to a heavily doped semiconductor with uniform current density.

$$V = V_{bulk} + 2V_{contact} = (R_{bulk} + 2R_{contact})I = \quad (1)$$

$$R_{bulk} = \frac{dV_{bulk}}{dI} = \frac{\rho l}{A} \quad (2)$$

$$R_{contact} = \frac{dV_{contact}}{dI} = \frac{\rho_c}{A}$$

where ρ is the bulk resistivity and ρ_c specific contact resistivity that can be defined through the component resistances.

Since the voltage required to drive current through a good ohmic contact is small we restrict the ρ_c definition to zero applied voltage.

$$\rho_c = \lim_{V \rightarrow 0} \left(\frac{dV_{contact}}{dJ} \right) \Omega \text{ cm}^2 \quad (3)$$

where J is the current density I/A. Alternatively (3) can be defined as

$$J = \frac{v_{metal} - v_{semicond}}{\rho_c} \quad (3a)$$

Thermionic Emission - Schottky Contact

For a Schottky contact the current governed by thermionic emission over the barrier is given by

$$J_S = A^* T^2 \exp\left(\frac{-2\phi_B}{kT}\right) \left(e^{qV/kT} - 1 \right) \quad (4)$$

where A^* is Richardson's constant. The specific contact resistivity as calculated by Eq. (3) is

$$\rho_c = \frac{k}{qA^*T} \exp\left(\frac{2\phi_B}{kT}\right) = \frac{kT}{qJ_s} \quad (5)$$

Tunneling - Ohmic Contacts

An ohmic contact is defined as one in which there is an unimpeded transfer of majority carriers from one material to another, i.e., the contacts do not limit the current. The way to achieve such a contact is by doping the semiconductor heavily enough that tunneling is possible. It is usual to heavily dope the Si regions N^+ or P^+ so that an ohmic contact is insured. Suppose N_d (or N_a) in the semiconductor is very large. Then the depletion region width at the metal - semiconductor interface

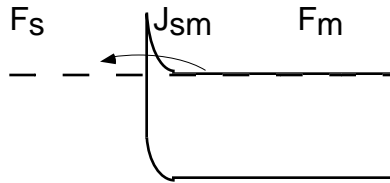
$$X_d = \sqrt{\frac{2 K \epsilon_o \phi_i}{q N_d}}$$

becomes very small. When $X_d < \approx 2.5\text{--}5\text{nm}$, electrons can “tunnel” through the barrier. This process occurs in both directions $M \rightarrow S$ and $S \rightarrow M$ so the contact shows very little resistance and becomes ohmic.

To calculate an approximate value for the required doping,

$$N_{d_{\min}} \approx \frac{2 K \epsilon_o \phi_i}{q X_d^2} \approx 6.2 \times 10^{19} \text{ cm}^{-3} \quad \text{for } X_d = 2.5 \text{ nm}$$

This is a relatively easy value to achieve in practice and is normally how ohmic contacts are made in integrated circuits.



For a tunneling contact the net semiconductor to metal current is given by

$$J_{sm} = \frac{A^* T}{k} \int F_s P(E) (1 - F_m) dE \quad (6)$$

where F_s and F_m are Fermi-Dirac distribution functions in metal and semiconductor respectively, and $P(E)$ is the tunneling probability given by

$$P(E) \sim \exp\left(-\frac{2\Phi_B}{h} \sqrt{\frac{\epsilon_s m^*}{N}}\right) \quad (7)$$

Where m^* is the effective mass of the tunneling carrier and h is the Plank's constant. The analysis to calculate current is more is somewhat more complicated, resulting in

$$J_{sm} \propto \exp\left[-2x_d \sqrt{2m^* (q\phi_B - qV)/h^2}\right]$$

Specific contact resistivity can be calculated using equations described above and is of the form

$$\rho_c = \rho_{co} \exp\left(\frac{2\phi_B}{h} \sqrt{\frac{\epsilon_s m^*}{N}}\right) \text{ ohm-cm}^2 \tag{8}$$

Where ρ_{co} is a constant dependent upon metal and the semiconductor. Specific contact resistivity, ρ_c primarily depends upon

- the metal-semiconductor work function, ϕ_B ,
- doping density, N , in the semiconductor and
- the effective mass of the carrier, m^* .

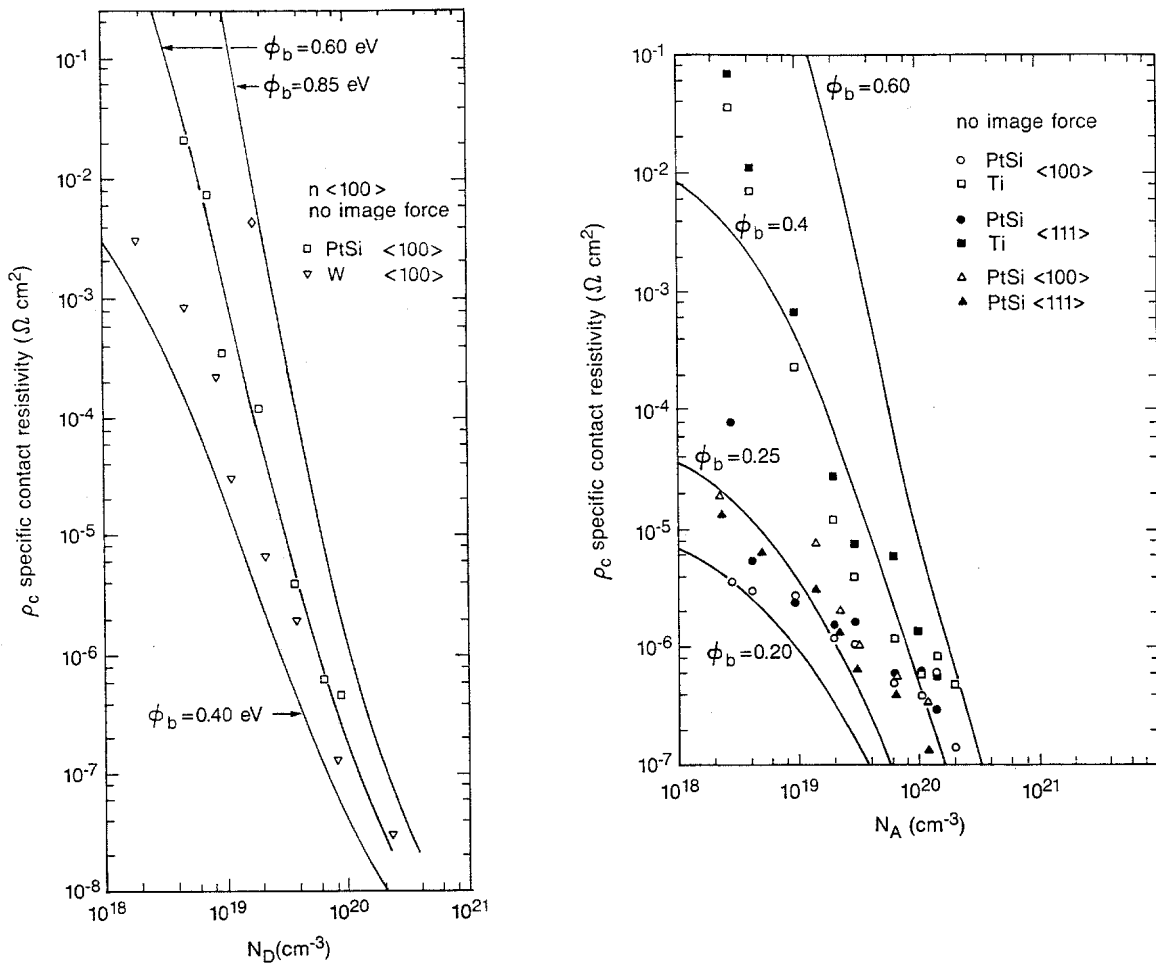


Fig. 5. Specific contact resistivity of metal contacts to n-type and p-type Si. Solid lines are calculated from the model. (Ref: S. Swirhun, Electrochem. Soc., Oct. 1988)

Observations

1. Specific contact resistivity, $\rho_c \downarrow$ as barrier height \downarrow
3. For a given doping density contact resistance is higher for n-type Si than p-type. This can be attributed to the barrier height
2. Specific contact resistivity, $\rho_c \downarrow$ as doping density \uparrow
 - Doping density can't be scaled beyond solid solubility.
 - N type dopants have higher solid solubility than P type dopants

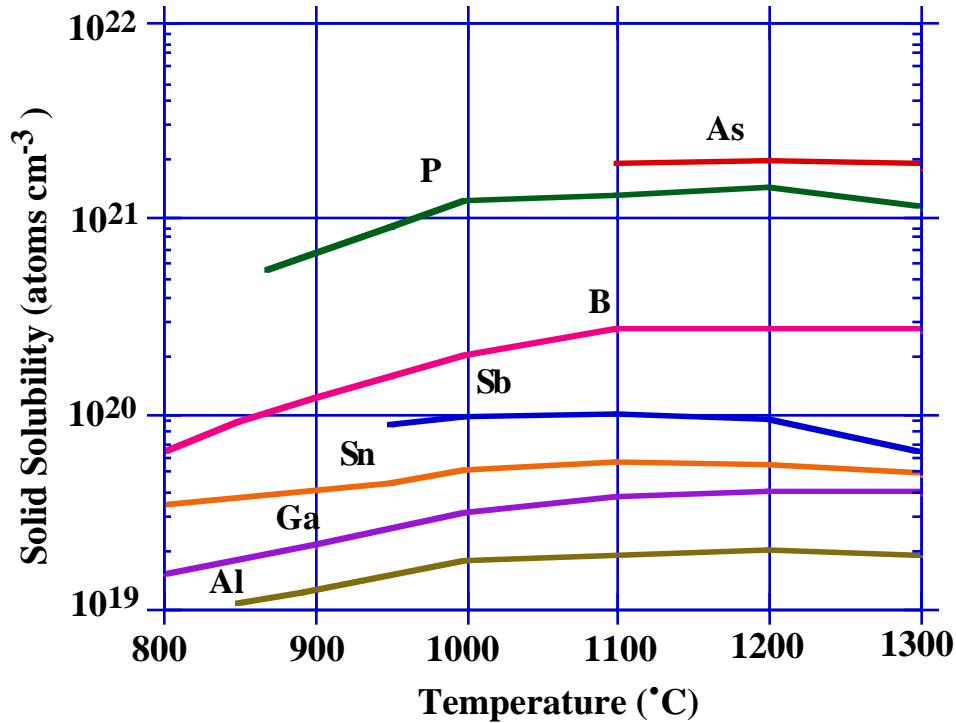


Fig. 8. Solid solubility of dopants in Si (Ref: Plummer & Griffin, Proc. IEEE, April 2001)

Barrier Height

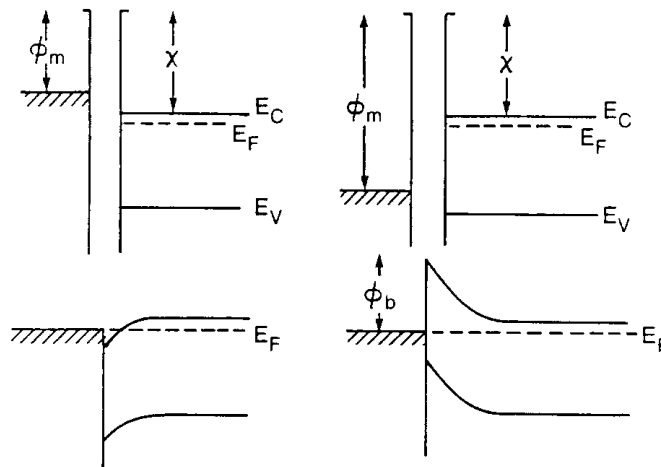


Figure: 9 Accumulation and depletion type contacts.

ρ_c is the physical parameter that describes the transport of majority carriers across heavily doped Si-metal interfaces. However, experiment and modeling of ohmic conduction is still crude. An ohmic contact is generally modeled as a heavily doped Schottky (diode) contact.

The Schottky model predicts that upon bringing in contact Si with electron affinity χ , and a metal of work function ϕ_m , a barrier of height $\phi_b = (\phi_m - \chi)$ which is independent of semiconductor doping will be formed. Since measured ϕ_m values for a variety of metals range from about 2.0 to 5.5 eV, and $\chi_{Si} \approx 4.15$ eV, this model should predict both accumulation and depletion (Fig. 9) metal-semiconductor contacts. This is generally not seen with Si; there is little evidence for the existence of any accumulation type metal to heavily doped Si contact. The reason is poorly understood but related to the restructuring of the metal-silicon surface. All practical n and p type ohmic contacts to Si are depletion type.

The barrier heights that are used in modeling ohmic contact to Si are empirical values, usually measured by capacitance-voltage, current-voltage or photoemission techniques. Metal and silicide barrier heights to both n and p type Si as a function of metal work function are illustrated in Fig. 10. The thin vertical lines connect data points for the same metal. The stronger ϕ_b dependence of metallic suicides on ϕ_m has led to the postulation that some interface cleanliness or the presence of an interfacial layer affects barrier height. Silicides are known to make more intimate contact to Si.

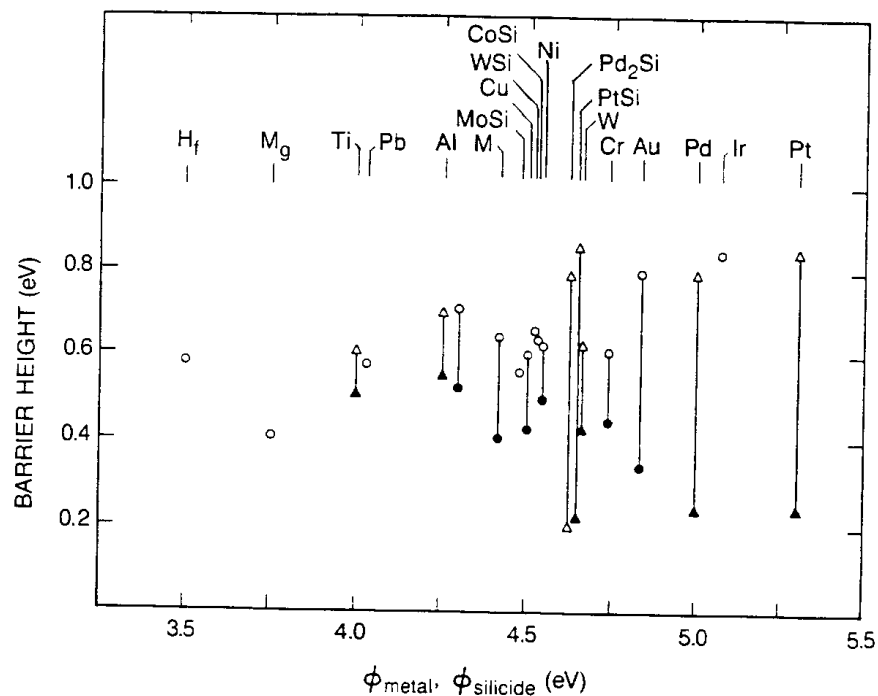


Figure 10: Metal-semiconductor barrier height to n- and p-type Si (ϕ_{bn} - hollow symbols and ϕ_{bp} solid symbols) vs. metal work function. (Ref: S. Swirhun, PhD Thesis, Stanford Univ. 1987)

It can be noticed that the Fermi level pinning is roughly at the same energy within the forbidden gap for both n and p type Si (i.e. the sum of ϕ_{bn} and ϕ_{bp} , is approximately E_g suggesting that interface and structural factors pin the Fermi level because of a very high density of interface states (Fig. 11). Note that for ohmic contacts we never need worry about the occupancy of these states changing, because of very small potential drop across the contact.

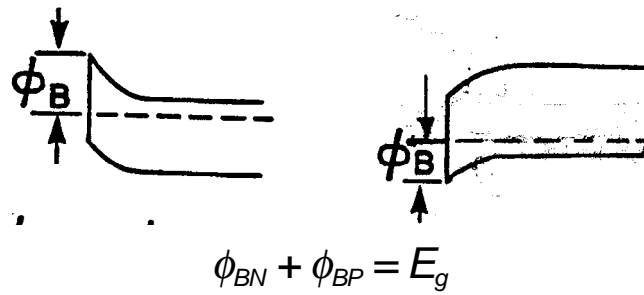


Figure. 11 Metal-semiconductor barrier height to n-type and p-type Si

Accurate Modeling of Contact Resistance

In practice it is difficult to construct a practical sized contact that passes a uniform current density over its area so this definition is usually considered in the limit as the elemental contact area approaches zero.

For a uniform current density ρ_c can be defined as contact resistance per unit area. However the situation becomes complicated in real device structures as the current distribution is non-uniform. Fig. 5 illustrates the current crowding to the front edge of a planar metal to semiconductor resistor contact. In such situation we can't use Eq. (2) to calculate contact resistance.

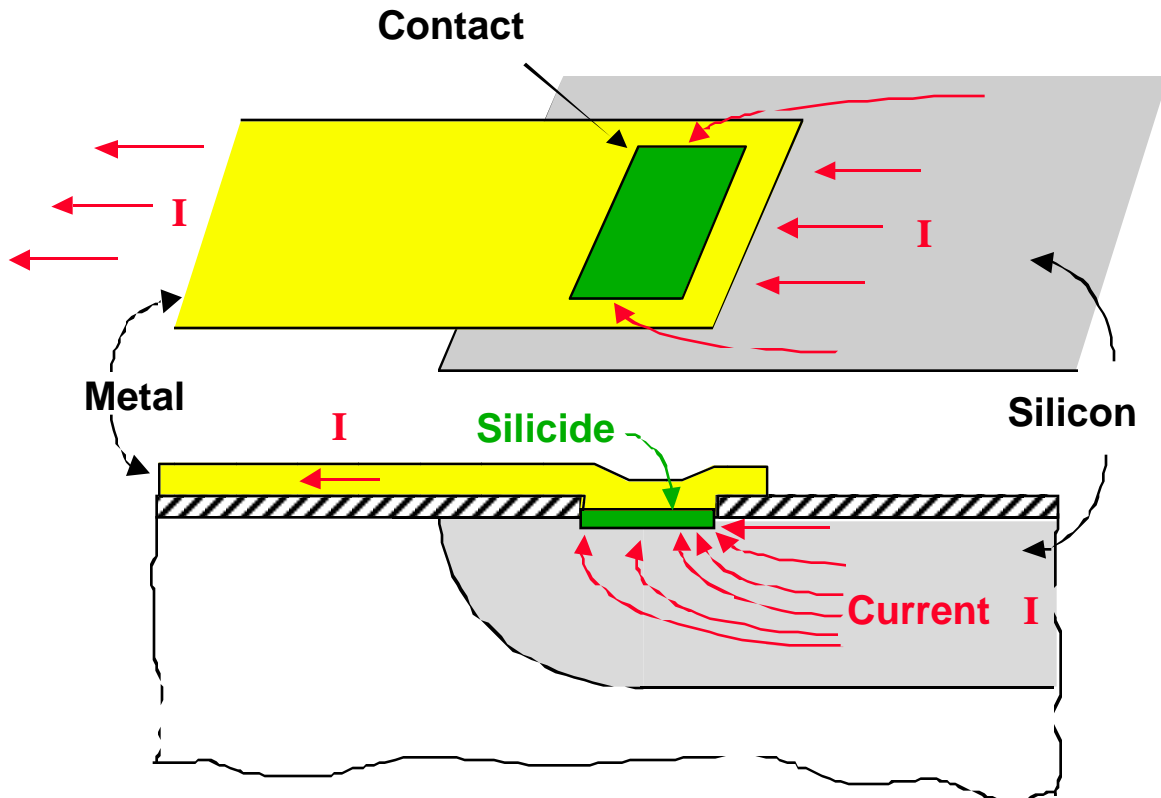


Figure : Non-uniform current distribution in a contact.

Generalized Contact Model

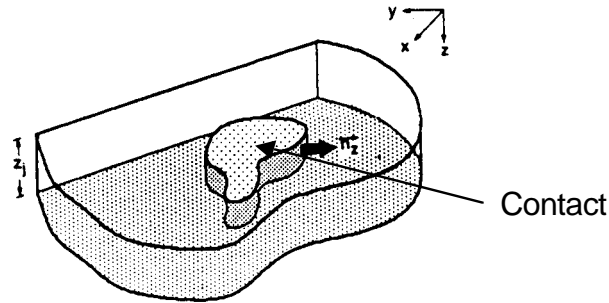


Fig. 12. General topology of the contact system. The contact surface is located at $Z = 0$; Z_j is the effective thickness of the semiconductor layer.

In general, the contact system can only be adequately described by the three basic transport equations, namely the Poisson and the two carrier continuity equations in 3-D. Under most circumstances, the equations can be simplified, and 2-D and 1-D models might be sufficient.

A. 3-Dimensional Model

The three-dimensional contact system has no restriction in the topology. Both metal potential v_m and semiconductor potential v are functions of the spatial coordinate x, y, z . In the heavily doped semiconductor region normally used in VLSI contacts, the following approximations can be made:

- (1) The effect of minority carriers is neglected. This assumption is equivalent to neglecting the depletion depth or band bending in the semiconductor region at the contact interface with respect to the depth of the semiconductor layer. (The depletion region is where minority-carrier effects such as recombination become significant.) The total current density J is then approximately the same as the majority carrier current density because the metal-semiconductor interfaces inject far more majority carriers than minority carriers.
- (2) By quasi-neutrality, the majority carrier concentration is equal to the active dopant density. Therefore, only the majority-carrier continuity equation requires solving in the semiconductor region beneath the contact.

The majority carrier continuity equation outside the contact becomes

$$\nabla \cdot \mathbf{J} = \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = 0 \quad (9)$$

The current density \mathbf{J} in the semiconductor is given by

$$\mathbf{J} = -\sigma \mathbf{E} = \sigma \nabla v \quad (10)$$

where v is the potential at coordinate (x, y, z) . By combining these two equations we obtain an equation similar to Ohms law

$$\nabla \cdot \sigma \nabla v = 0 \quad (11)$$

This formulation also applies to the metal region with a similar set of expressions. If the metal conductivity is much larger than of the semiconductor, which is generally the case, v_m becomes constant over the entire interface. The entire contact system will then be

governed by the semiconductor potential v , the only variable that needs to be determined for a given metal-semiconductor system. The total current can be evaluated over any sectioned surface A by

$$I_{tot} = -\int \mathbf{J} \cdot d\mathbf{A} \quad (12)$$

Solution of the above equations with appropriate boundary conditions will give the necessary information about contact resistance. The 3-D model is simple in concept, but difficult in its computation and generalization. Therefore it is advantageous to simplify the equations and boundary conditions to 2-D, which are much more tractable and still produce useful insights.

B. 2-dimensional Model

The transformation from 3-D to 2-D involves a few simplifications. First, the contact interface is regarded as a 2-D surface perpendicular to the z axis as illustrated in Fig. 12. The semiconductor or diffusion layer is located below the contact surface with an effective thickness of z_j . The conductivity is assumed to be independent of the x and y spatial variables, i.e., $\sigma = \sigma(z)$. This simplification is valid for most of the modern VLSI technologies with planar diffusion layers. The aim of the 2-D model is to lump all the effects of the z -axis into just one single parameter R_s , the sheet resistance of the diffusion layer is given by

$$R_s = \left(\int \sigma(z) dz \right)^{-1} \quad (13)$$

The metal plane potential v_m , seen by the contact will be essentially constant because the metal layer is usually much more conductive than the semiconductor layer. If this constant metal potential is set at zero then 3D equations in the contact region can be simplified to the Helmholtz equation (see the paper by Loh et al. for details)

$$\nabla^2 V = \frac{R_s V}{\rho_c} = \frac{V}{l_t^2} \quad (14)$$

with l_t as the transfer length defined as $l_t = \sqrt{\rho_c / R_s}$. In the other bulk region where there is no contact surface on top, the Laplace equation describes the potential by

$$\nabla^2 V = 0 \quad (15)$$

A solution of these equations gives the I-V relationship at the contact interface. By comparing the experimental data with the 2-D model an accurate value of ρ_c can be extracted. This accurate value can then be used for further calculation of the contact resistance for the appropriate structure.

C One-dimensional Model

One more spatial variable can be eliminated if the potential changes only slightly, and not affecting other potentials along the variable axis. The contact system is oriented such that the y -axis variation is neglected. The Helmholtz equation becomes

$$\frac{\nabla^2 V(x)}{\partial^2 x} = \frac{V(x)}{l_t^2} \quad (16)$$

The Laplace equation also reduces to Ohm's law. All the boundary conditions become trivial: $V = V_i$ at the contact leading edge ($x = 0$) and $\partial V/\partial x=0$ at the contact trailing edge ($x = l$).

The potential can be shown as

$$V(x) = V_i \frac{\cosh\left(\frac{l-x}{l_t}\right)}{\cosh\left(\frac{l}{l_t}\right)} \quad (17)$$

and total current is simply

$$I_{tot} = \frac{W}{R_s} \frac{\partial V}{\partial x} \Big|_{x=0} \quad (18)$$

Historically, this is the approach taken to model the distributive effect of current entering a contact window. A solution of Eq. (17) and (18) will give contact resistance.

D. Zero-Dimensional Model

Under very special circumstances, such as in large contact windows, an extremely high value of ρ_c or when very small contacts are encountered, the I-D model can be degenerated into the O-D model or the one-lump model. This is the simplest of all existing models. Although its validity is scarce, it is the most common model used as a first pass to estimate the upper bound of ρ_c . In other words, its accuracy is poor, but it offers a very intuitive "feel" of the contact resistance. This model states that the potential is constant in the semiconductor layer and the current density entering the contact window is uniform. From (3a) the macroscopic "definition" of ρ_c appears as

$$R_c = \frac{\rho_c}{A} \quad (19)$$

Intuitively, the contact resistance R_c of a contact will approach ρ_c/A as the contact sizes decreases below the transfer length l_t .

Measurement of Contact Resistance and Specific Contact Resistivity (ρ_c)

The contact resistance R_c is a measured V/I ratio of a structure, that is controlled by the contact size, structure layout, semiconductor doping density, and specific contact resistivity ρ_c . Whereas ρ_c is a fundamental quantity governed by the interface, R_c accounts for the layout dependent non-uniform current flow pattern. A more accurate 2D or 3D analysis is required for accurate calculations of R_c .

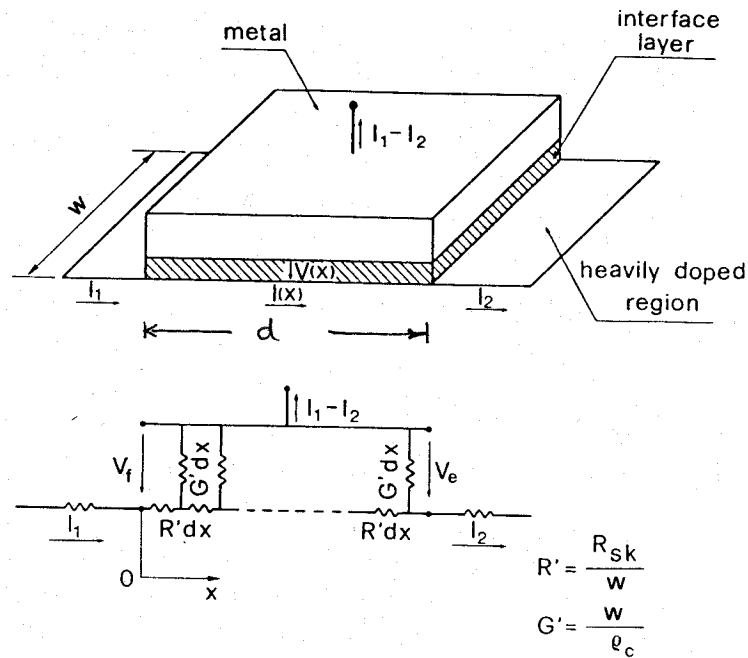


Fig. 13. 1D transmission line contact model.

The simple 1D model assumes a semiconductor modeled by a distributed sheet resistance R_s and no vertical extent. The metal sheet resistance is assumed negligible, i.e., uniform v_m . Assuming a long transmission line the contact resistance is given by the characteristic impedance of the line which can be obtained by a solution of Eq. (17) and (18). The transmission line like model gives current density that decreases roughly exponentially from the leading edge of the contact to the trailing edge.

$$I(x) = I_1 \exp\left(-\frac{x}{\sqrt{\rho_c/R_s}}\right) = I_1 \exp(-x/l_t) \tag{20}$$

The characteristic length of the transmission line $l_t = \sqrt{\rho_c/R_s}$ is the distance at which 63% of the current has transferred into the metal. This model is valid only for an electrically long contact ($d \gg l_t$).

Contact resistance test structures are usually fabricated with other conventional test devices on the same die or wafer to monitor a particular process. Therefore, the most commonly used contact test structures for extraction of ρ_c , are planar devices: the cross bridge Kelvin resistor (CBKR), the contact end resistor (CER), and the transmission line tap resistor (TLTR). In all of these structures, a current is sourced from the diffusion level up into the metal level via the contact window. A voltage is measured between the two levels using two other terminals. The contact resistance for each structure is simply this voltage divided by the source current. It is important to realize that each device measures the voltage at a different position along the contact as shown in Fig. 14; hence the resistance values measured are different, and must be clearly defined and distinguished from one another.

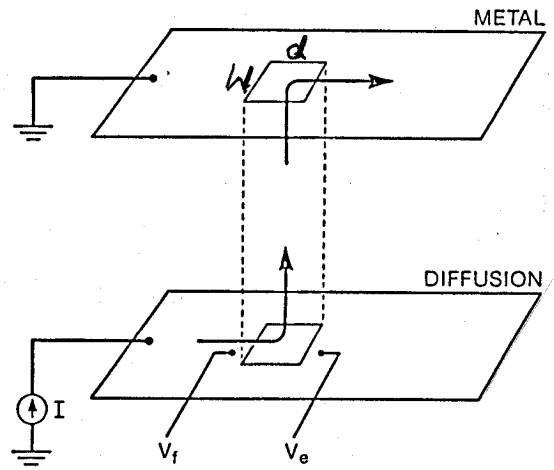


Fig. 14

The *front* resistance R_f is defined as the ratio of the voltage drop V_f across the interfacial layer at the front edge of the contact, where the current density is the highest, to the total current I_1 flowing into the contact. For a bounded structure ($l_2 = 0$) it can be shown by solving of Eq. (17) and (18) that

$$R_f = V_f / I_1 = \frac{\sqrt{R_s \rho_c}}{w} \coth(d / l_t) \tag{21}$$

For a very large value of l_t or for $d \ll l_t$

$$R_f \approx \frac{\rho_c}{wd} \tag{22}$$

which is the equation for uniform current density. A convenient structure to make the measurements is the *transmission line tap resistor* in which several contacts are made to a long diffused line.

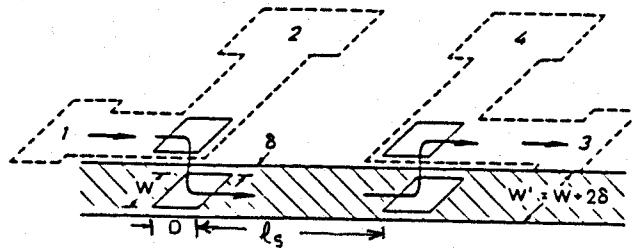


Fig. 15 Transmission line tap resistor for R_f measurement

$$V_{24} = V_f + IR_{Si} + V_f$$

$$R_t = \frac{V_{24}}{I} = 2R_f + R_s l_s / w \tag{23}$$

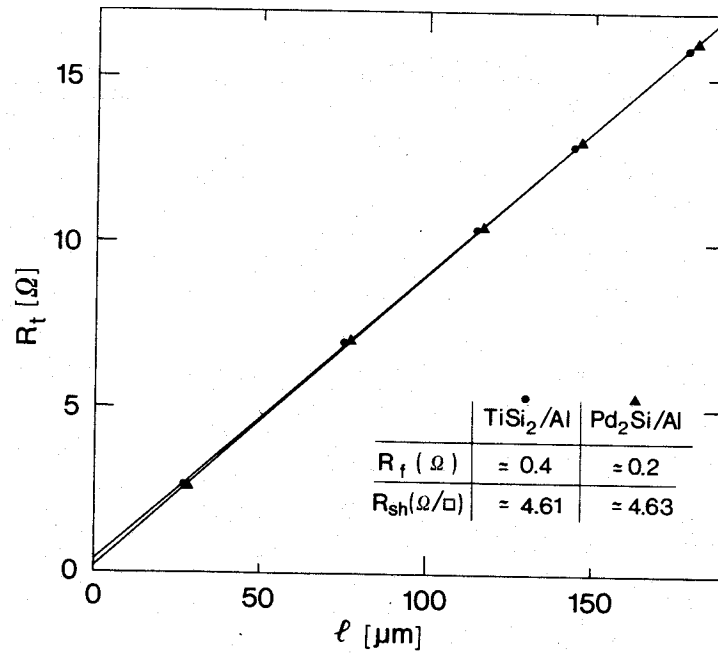


Fig. 16. Characteristic plot of the total resistance as a function of the contact separation in a TLTR structure. The nominal dimensions were $l_1 = 30, l_2 = 80, l_3 = 120, l_4 = 150, l_5 = 180, w = 40, W = 50 \mu\text{m}$.

By varying the value of l_5 the value of R_f can be determined as shown in the figure above. However, since R_f is a smaller number in comparison to the R_{Sj} the relative accuracy of this method is depends upon the measurement accuracy. Slight error in the measurement of R_f can result in a large error in the calculated value of R_f .

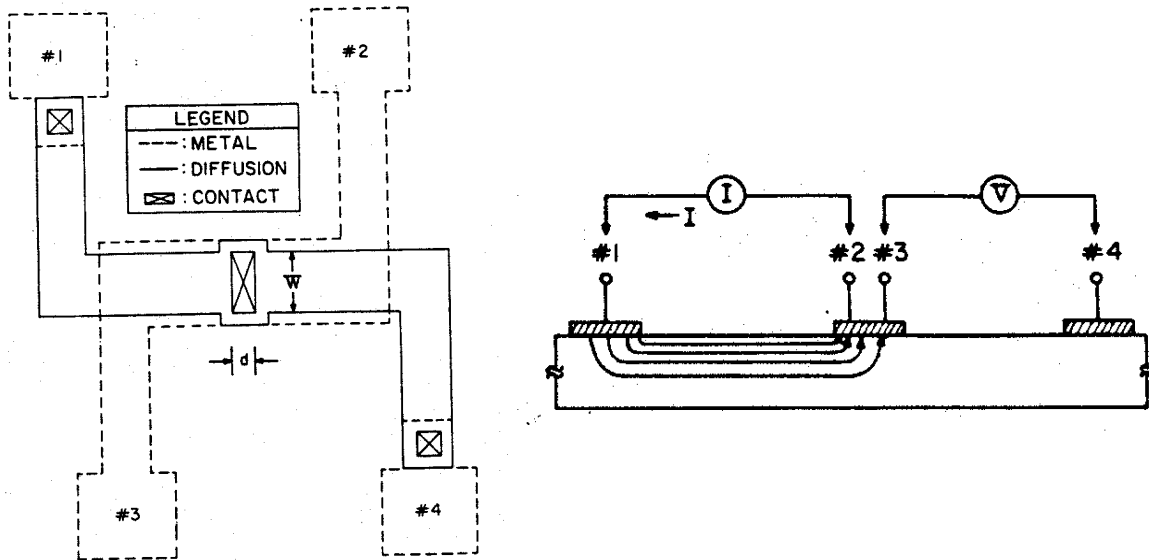


Fig 17. (a) Test pattern for R_e contact end resistance measurement. Where current I is forced through pads 1 and 2 and voltage V across pads 3 and 4 is sensed. $R_E = V/I$. (b) Cross-sectional view of (a)

Similarly the **end resistance** R_e is defined as the ratio of the voltage drop V_e across the interfacial layer at the back edge of the contact, where the current density is the lowest, to the total current I flowing into the contact and is given by

$$R_e = V_e / I = \frac{V_m - V_4}{I} = \frac{\sqrt{R_s \rho_c}}{w \sinh(d/l_t)} \tag{24}$$

The most popular structure is the *cross bridge Kelvin resistor*. It assumes a uniform current density and uses Eq (2) for calculations.

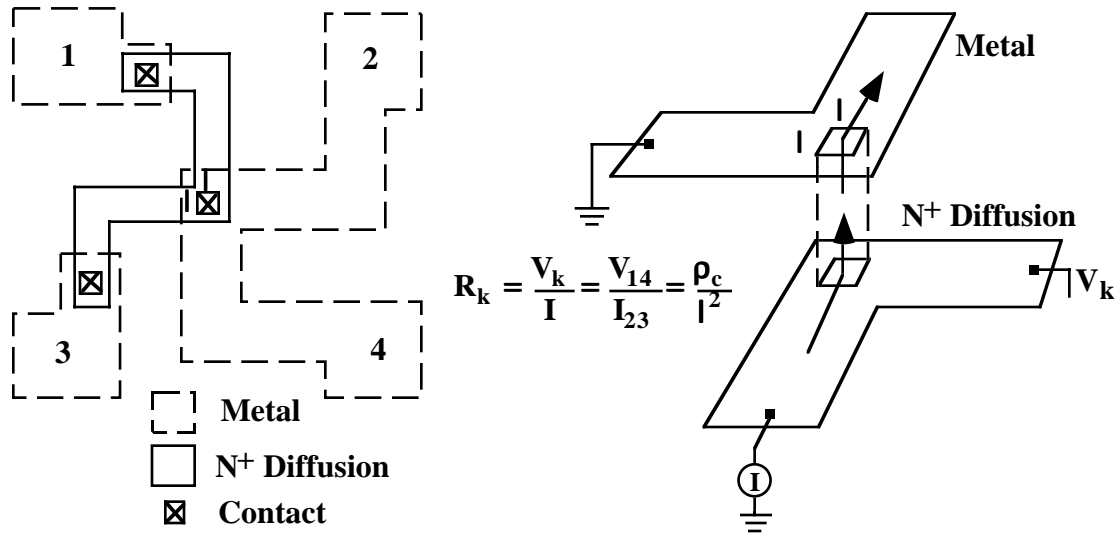


Fig. 18 cross bridge Kelvin resistor

In all of these structures we have assumed that current flows only in one direction. In reality the current flow is highly 2D.

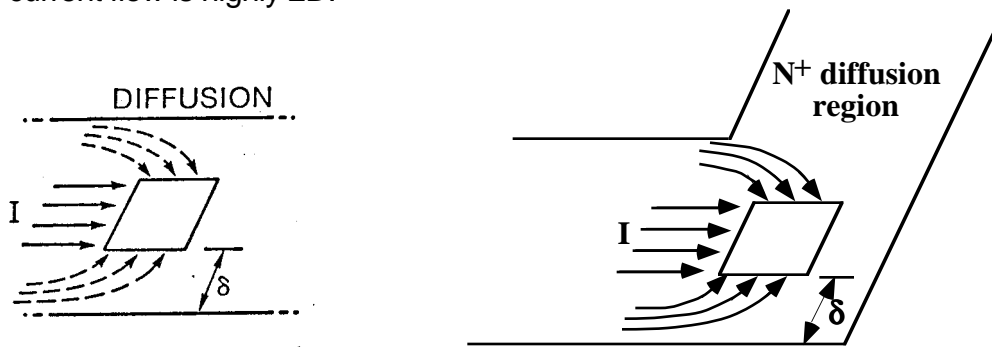


Fig. 19 Schematic of 2D current flow in the contacts.

This leads to overestimation of ρ_c . For example in a Kelvin structure the measured resistance does not scale with the contact area (Fig. 20).

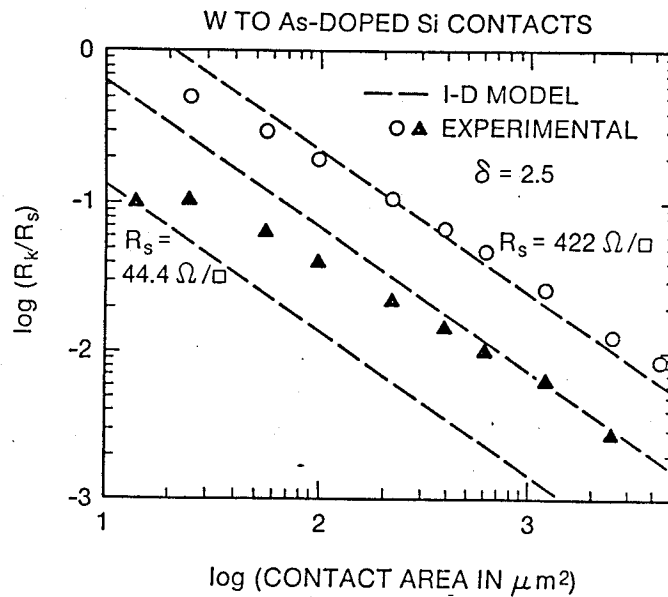


Fig. 20 Contact resistance as a function of contact area for Kelvin structure

A detailed 3D or at least 2D simulation should be used to determine the correct values of specific contact resistivity, ρ_c .

By comparing the experimental data with the 2-D model an accurate value of ρ_c can be extracted. This is illustrated in Fig. 21 for the case of Kelvin structure. This accurate value can then be used for further calculation of the contact resistance for the appropriate structure.

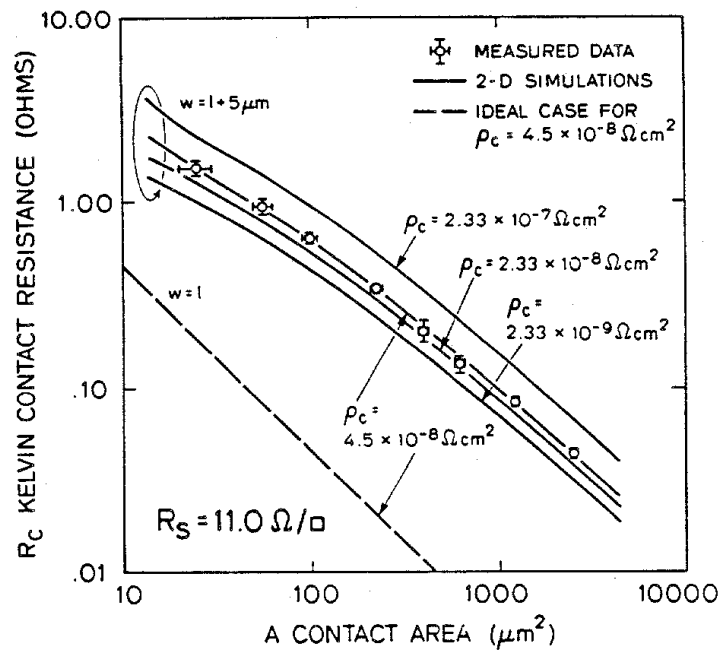


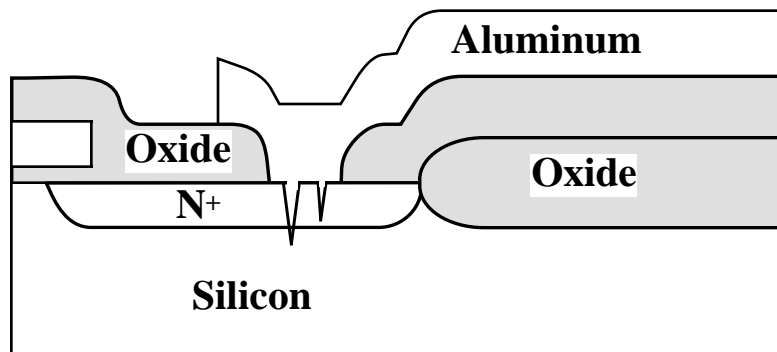
Fig. 21. Kelvin resistance vs contact area. Diffusion width(w) is larger than contact window size (l) by $5 \mu\text{m}$ in the top set of curves. The ideal case is $w = l$. Sheet resistance of the diffusion is $11 \Omega/\text{sq}$, the simulation parameter ρ_c is varied from 2.33×10^{-7} to $2.33 \times 10^{-9} \Omega\text{cm}^2$.

Requirements of Ohmic Contacts

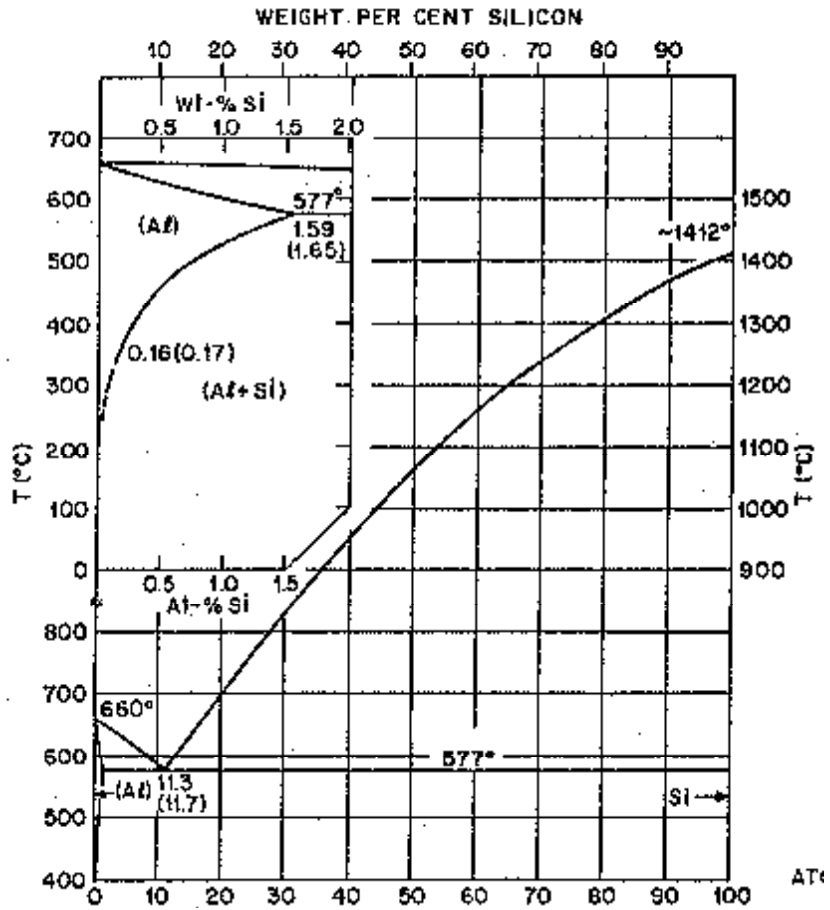
1. Low contact resistance to both N^+ and P^+ regions
2. Ease of formation (deposition, etching)
3. Compatibility with Si processing (cleaning etc.)
4. No diffusion of the contact metal in Si or SiO_2
5. No unwanted reaction with Si or SiO_2 and other materials used in backend technology.
6. No impact on the electrical characteristics of the shallow junction
7. Long term stability

Aluminum Contacts to Si

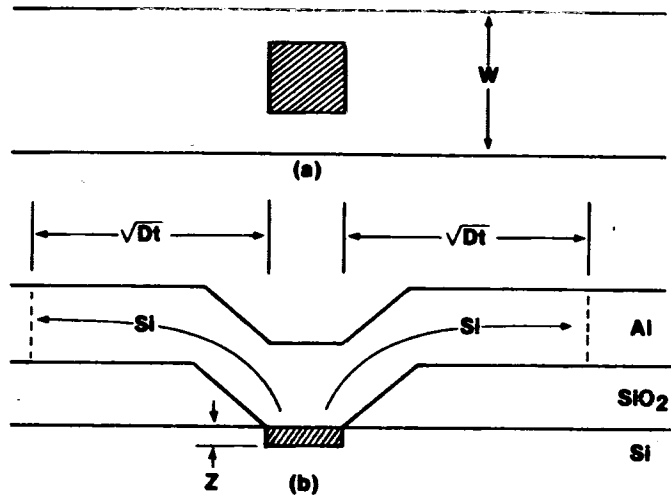
Al has been used for a long time to make contacts to Si because it meets many of the above requirements.



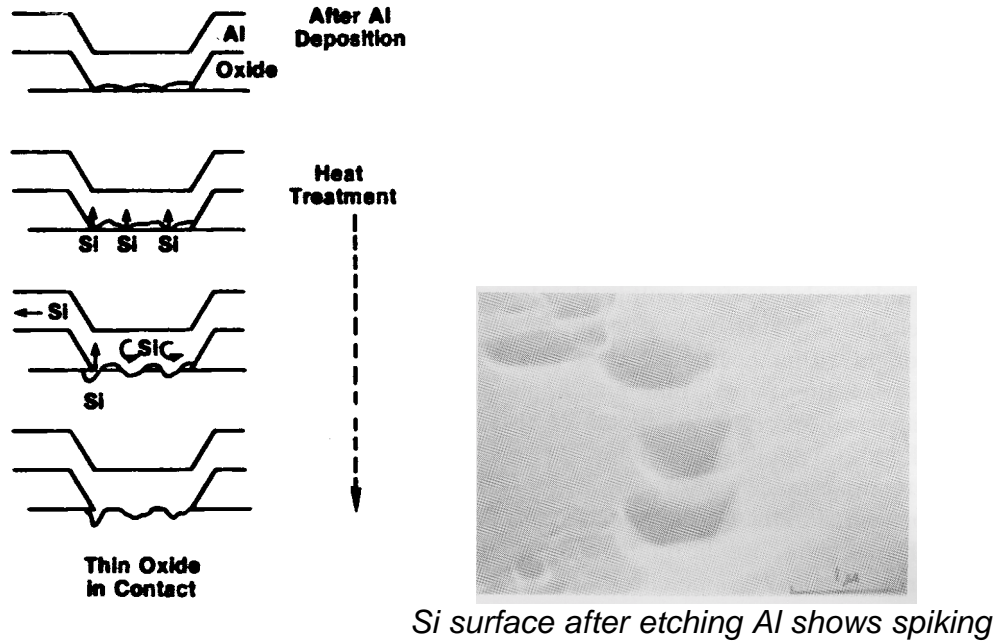
For shallow junctions it suffers with the problem of *junction spiking*.



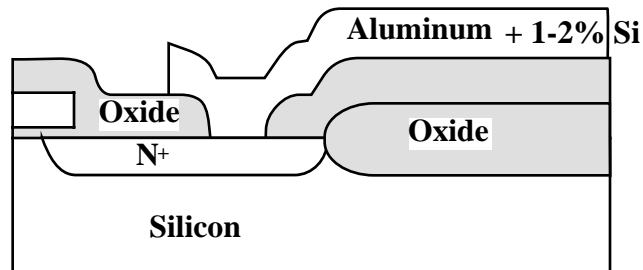
- Silicon has high solubility in Al ~ 0.5% at 450°C
- Silicon has high diffusivity in Al. At 450°C $D = 10^{-8} \text{ cm}^2/\text{sec}$



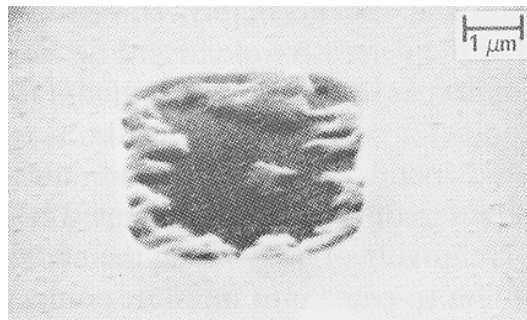
- Si diffuses into Al. At 450°C $\sqrt{Dt} \approx 40\mu\text{m}$



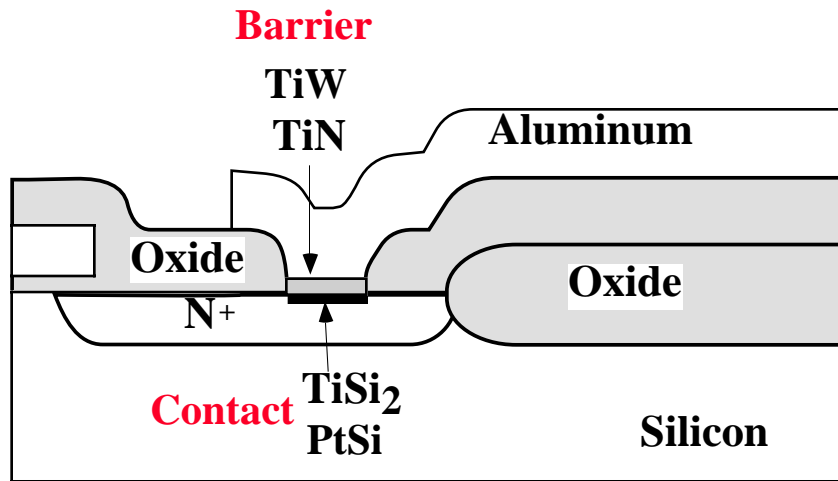
- Voids form in Si which fill with Al: “Spiking” occurs.
- Pure Al can’t be used for junctions < 2-3 μm
- By adding 1-2% Si in Al to satisfy solubility requirement junction spiking is minimized



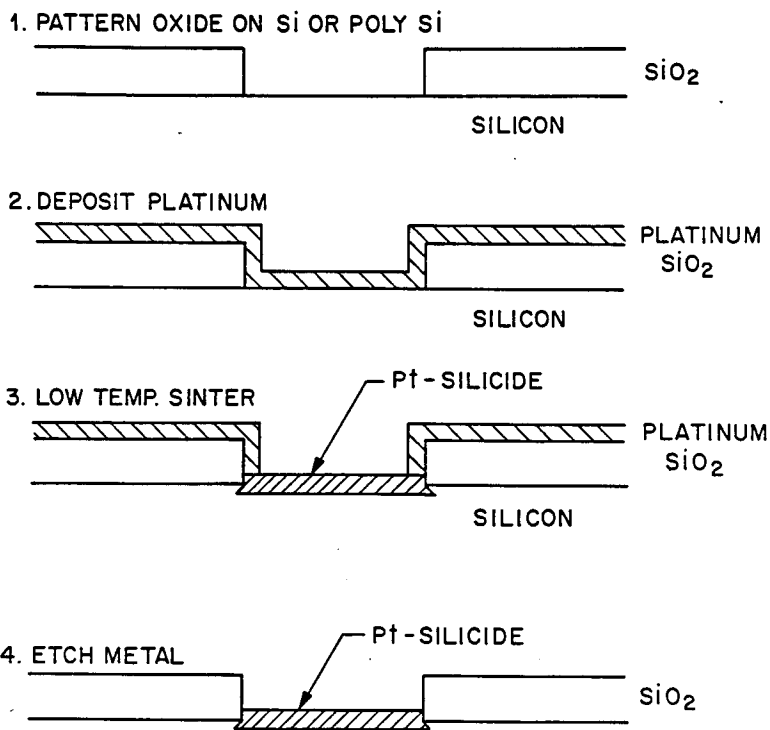
- But Si precipitation can occur when cool down to room temperature \Rightarrow bad contacts to N⁺ Si as the Si precipitates are saturated with Al which is a p type dopant



Silicide Contacts



- Silicides like PtSi, TiSi₂ make excellent contacts to Si

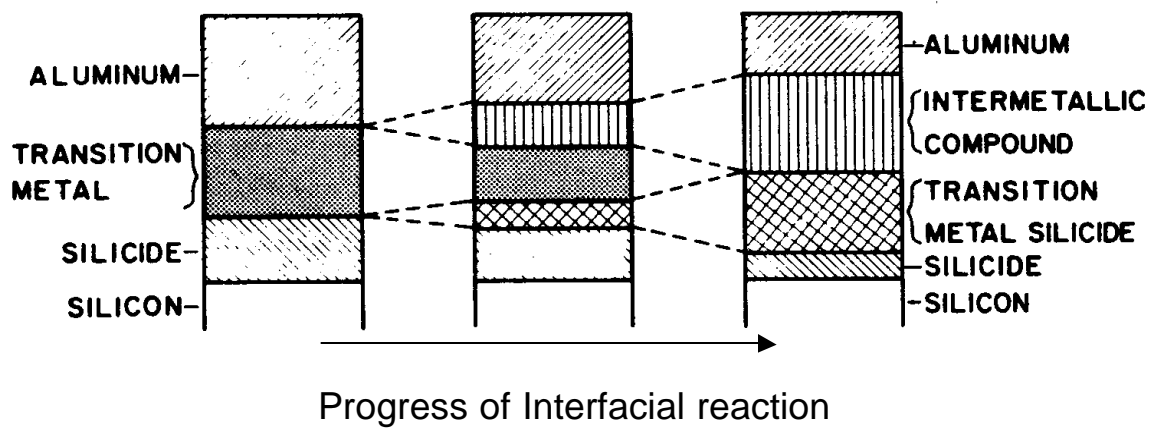


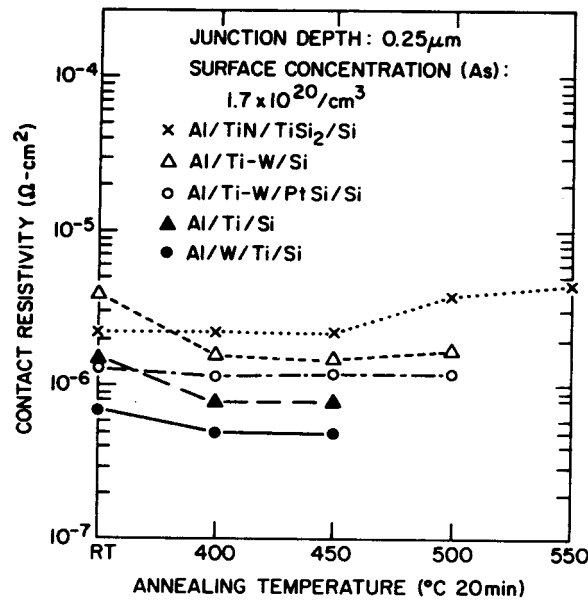
- However, they react with Al
- A barrier like TiN or TiW prevents this reaction

Barriers

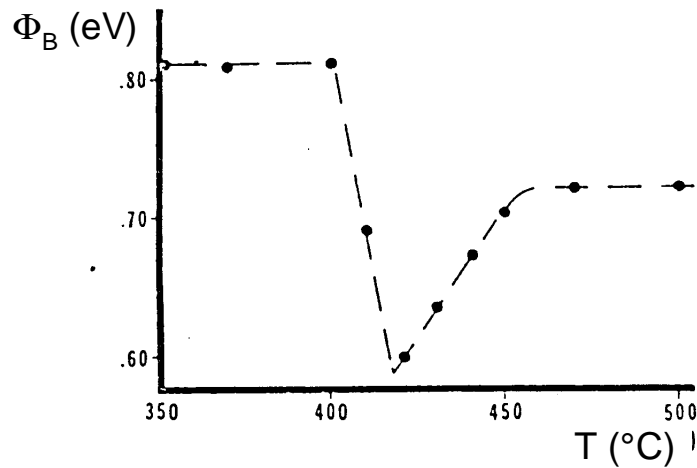
| Structure | Failure Temperature (°C) | Failure Mechanism (Reaction products) |
|--|--------------------------|--|
| Al/PtSi ₂ /Si | 350 | Compound formation (Al ₂ Pt, Si) |
| Al/TiSi ₂ /Si | 400 | Diffusion (Al ₅ Ti ₇ Si ₁₂ , Si at 550°C) |
| Al/NiSi ₂ /Si | 400 | Compound formation (Al ₃ Ni, Si) |
| Al/CoSi ₂ /Si | 400 | Compound formation (Al ₉ Co ₂ , Si) |
| Al/Ti/PtSi ₂ /Si | 450 | Compound formation (Al ₃ Ti) |
| Al/Ti ₃₀ W ₇₀ /PtSi ₂ /Si | 500 | Diffusion (Al ₂ Pt, Al ₁₂ W at 500°C) |
| Al/TiN/TiSi ₂ /Si | 550 | Compound formation (AlN, Al ₃ Ti) |

- Silicides react with Al at T < 400°C
- A barrier like TiN or TiW prevents this reaction upto 500°C





Contact resistivity vs anneal temperature for a variety of metallization systems indicating the effectiveness of various barriers (Ref: Wittmer, JVST 1984)



Without the barrier the contacts will be severely degraded as shown by change in barrier height in an Al/PtSi Schottky contact *without the barrier*. At higher temperature Al reacts with PtSi reducing the barrier height. If there was a shallow junction being contacted, it would have been damaged.