Chapter 4
Temperature-Dependent Electrical Characteristics of Semiconductor Devices

Gábor Farkas

Thermal transient testing is a true interdisciplinary segment of engineering. It has its roots in thermodynamics, but it also requires firm knowledge of electronics to ensure stable powering for the devices under test and to count with secondary effects caused by internal tester circuits and cabling.

The electrical characteristics of semiconductor devices, that is, what currents flow in various material regions at certain applied voltages and how these quantities depend on the temperature in these regions, are determined by the laws of solid-state physics.

In the typical academic curricula of engineers, the subjects are presented with focus on some narrow topics and wide gaps among them. For example, a student encounters electron shells and orbits in chemistry, then possibly learns about diode and transistor characteristics in electronics, jumping to logic gates in digital design, and suddenly switches to operating systems in computer science.

In this chapter, we do not want to convey a deep and continuous knowledge connecting all above. We also want to avoid just giving a reference to thick tomes of solid-state physics with many unrelated topics. Instead, in a short introduction (Sect. 4.1), we sketch the train of thought which leads from the physical background to the temperature-dependent electrical characteristics of semiconductor devices, and we refer to literature with exhaustive information for those who need a deeper insight. Those who focus on resulting device characteristics and their temperature dependence can directly refer to all subsequent sections following the introduction.

G. Farkas (✉)
Siemens Digital Industry Software STS, Budapest, Hungary
e-mail: Gabor.Farkas@siemens.com

M. Rencz et al. (eds.), *Theory and Practice of Thermal Transient Testing of Electronic Components*, https://doi.org/10.1007/978-3-030-86174-2_4

4.1 Basic Laws of Solid-State Physics

Solid-state physics is built on a few simple principles, but the elaboration to the level where the temperature-dependent equations governing the behavior of semiconductor devices can be obtained needs a difficult mathematical apparatus.

It is a common practice in this branch of physics to use theoretical models that yield analytic equations for physical processes and then to use empirical correction factors to fit the formulae to measured properties. In this section we use simple models that describe the basic physical effects, especially those related to the temperature dependence of semiconductor devices.

Those who need deeper knowledge of the physical background may refer to several classic books. General *semiconductor physics* is treated in depth in [10, 13]. A more focused work on the *physics of LED devices* is [11]; in an easier approach, also the first chapters of [7] can be useful.

There are also useful *online resources* on the topic in a well-structured linked arrangement; these are generally less bulky than printed books. Instead of reading them from begin to end, generic search for related keywords on the Internet often directly results in links to appropriate chapters of [12, 14].

Short and useful online summaries on the derivation of the equations which determine the characteristics of diodes or other devices built of pn junctions are given in [98–100].

A useful *Appendix to the solid state physics of heaters and sensors* is available online at [15] which embraces most of the solid-state physics background in moderate length.

4.1.1 Band Structure of Semiconductors, Electrons, and Holes

In atoms, the electrons can have only discretized energy levels assigned to electron shells. In crystals with a high number of atoms, the electron shells disperse into broader bands of allowed energy levels.

In semiconductors and insulators, the bands are separated with bandgaps, electrons cannot possess energies belonging to these forbidden energy ranges. In these materials there exists a last energy band which is nearly fully populated by electrons; this is called valence band, because the interaction of its electrons establishes the bonding between crystal atoms. The next higher energy band, called conduction band, is nearly empty and is separated from the valence band by a bandgap of W_g energy distance.

The distribution of electrons in a space with potential energy is governed by the *Fermi-Dirac* (F-D) function. It specifies the probability that an electron occupies an existing state at energy level W:

$$f(W) = \frac{1}{e^{(W - W_F)/kT} + 1} \tag{4.1}$$

The average energy of all electrons is the W_F Fermi energy. In general, particles of average energy do not necessarily exist.

When the particle concentration is low, the Fermi-Dirac distribution does not significantly differ from the *Maxwell-Boltzmann* (M-B) distribution which says that at higher energy, the concentration of particles diminishes exponentially:

$$f(W) \approx e^{-(W - W_F)/kT} \tag{4.2}$$

The F-D and M-B distributions are temperature dependent; at higher temperatures, more electrons appear around the same energy level.

Those materials in which at room temperature (or at the temperature of use) a small but observable number of electrons, corresponding to the F-D distribution, have a high enough energy to escape from the valence band into the conduction band are called *semiconductors*. In a pure (intrinsic) semiconductor, the number of empty energy levels left in the valence band is necessarily equal to the number of electrons elevated into the conduction band.

As it is rather hard to track the behavior of the large number of electrons in the valence band, the concept of *holes*, quasiparticles corresponding to the missing electrons, is introduced. Instead of following the position of empty places in the abundance of electrons, holes are handled as particles of positive charge, residing at the highest energies of the valence band. The elevation of an electron into the conduction band is interpreted as the generation of an electron-hole pair; the return to the valence band is handled as a recombination of two particles of different energy.

The equal number of electrons in the conduction band of the intrinsic semiconductor and the holes in the valence band inflicts that the average W_F energy is near to the middle of the bandgap. This energy level is called the W_i intrinsic Fermi energy.

The temperature dependence of the F-D and M-B distributions determines the $n_i = p_i$ concentration of electrons and holes in intrinsic semiconductors, as proved in [15]

$$\begin{aligned} n_i &= G \cdot T^{\frac{3}{2}} \cdot e^{\frac{-W_g}{2kT}} \\ n_i^2 &= G^2 \cdot T^3 \cdot e^{\frac{-W_g}{kT}}, \end{aligned} \tag{4.3}$$

where T is the absolute temperature, k is the Boltzmann constant connecting thermal energy to temperature and the G factor cumulates structural parameters of the crystal lattice with no significant temperature dependence.

When an external E electric field is applied on the crystal, all mobile electrons start moving in it against the direction of the electric field, due to their $-q$ elementary negative charge. This movement is known as electric current. Again, it is easier to

describe the flow of the high number of electrons in the valence band as the movement of a few holes of positive charge toward the direction of the external field.

In free space (vacuum), an electron can "fly" freely. It may have any energy and momentum, both of them determined by its m mass and v velocity. The momentum is proportional the velocity, $p_w = m \cdot v$; the kinetic energy is quadratically dependent as $W_{kin} = ½\, m \cdot v^2$. A deeper interpretation of p_w is presented in [15].

At absolute zero temperature, the lattice atoms, which are of positive charge, form a perfect periodic potential in space, repeated in all directions by the lattice constant of the crystal. In this undisturbed periodic potential, braking and accelerating forces influence the movement of an electron. As a consequence, electrons and holes act as having an m_n* and m_p* effective mass, not equal to the mass of the free electron. As a further consequence, the laws of quantum mechanics imply that in a periodic potential space due to the braking and accelerating forces, the minimum energy of the moving electron does not necessarily coincide with the minimum momentum. (Momentum has some more generalized meaning in quantum mechanics; a deeper definition is presented in [15].) In the perfect periodic potential space, no obstacle perturbs the movement of electrons; they can accelerate freely as determined by their m* effective mass.

Those semiconductors in which the minimum energy of electrons and the maximum energy of holes belongs to the same (typically zero) momentum are called semiconductors of *direct bandgap*. In other semiconductors, as silicon or silicon carbide, electrons of lowest energy have nonzero momentum; these are materials with *indirect bandgap*.

4.1.2 The Concentration of Charge Carriers and Their Motion in Semiconductors

At nonzero temperatures, the lattice atoms oscillate around their position in the crystal lattice. Moving electrons interact with the vibrating atoms, "collide" in a broader sense, and exchange energy with them. These interactions ensure that the temperature of the electrons remains the same as the temperature of the crystal atoms until no extreme external energy is applied on the material. The electrons at T temperature move stochastically at v_{th} average velocity; as a consequence of the M-B distribution, their kinetic energy is

$$W_{kin} = \frac{1}{2} m^* \cdot {v_{th}}^2 = \frac{3}{2} kT \tag{4.4}$$

Equation (4.4) is the definition of the T absolute temperature, and it presents how the k Boltzmann constant connects the concept of kinetic energy to temperature.

In real crystals also other, less temperature-dependent imperfections of the lattice arise. These crystal defects can be atoms of other elements occupying regular

positions in the lattice, or in interstitial position, and dislocations. A large imperfection is the surface of the bulk semiconductor itself.

Applying E electric field on the semiconductor, the persistent energy exchange with the lattice prevents unlimited acceleration of the charge carriers. The braking effect of the interactions with lattice atoms and crystal defects can be expressed in the empirical quantity of the μ mobility. The electric field causes a drift of carriers of v velocity: $v_n = \mu_n \cdot E$ for electrons and $v_p = \mu_p \cdot E$ for holes. The movement of charge carriers provides the electric current:

$$J_{ndrift} = q \cdot \mu_n \cdot E, \quad J_{pdrift} = q \cdot \mu_p \cdot E, \tag{4.5}$$

J is the current density and q is the elementary charge. The electrons have negative charge and move opposite to the E field; both current components have the same direction. Equation (4.5) is the differential form of the Ohm law, $J = \sigma \cdot E$, where σ is the electrical conductivity of the material and J is the sum of the currents of electrons and holes. The thermal power generated in unit volume is $P_V = J \cdot E$.

In electronics generally current is imagined as a continuous flow of charge carriers, and a small perturbation, *noise*, is superposed on this continuous flow. A signal-to-noise ratio is an eminent descriptor of a measurement. In reality, the electrons at T temperature stochastically move at high v_{th} thermal velocity, and a small v drift velocity is superposed on their random movement. Better to say, because their temperature does not change independently from the temperature of the crystal, the stochastic movement is turned slightly in the direction of the E field by a v drift velocity.

The v_{th} thermal velocity in bulk semiconductor is about 10^7 cm/s. For comparison, the μ_n electron mobility in intrinsic silicon is about 1400 cm^2/Vs, and the μ_p mobility of holes is around 480 cm^2/Vs at room temperature; and both quickly diminish at higher temperature. The field strength is limited by the presence of mobile carriers; the charge carriers' movement is really best described as a small superposed drift.

The density of electrons and holes in a semiconductor can be influenced by doping, by substituting, for example, silicon atoms in the lattice with trivalent atoms (acceptors) or pentavalent atoms (donors). The substituent atoms can be considered dopants, if the 3/2 kT thermal energy at room temperature (or at the temperature of actual use) is sufficient to ionize almost all the dopant atoms. In a neutral material section of doped semiconductor, $n \approx N_D$ in an *n-type* material of N_D donor concentration and $p \approx N_A$ in a *p-type* material of N_A acceptor concentration.

In an *n-type* semiconductor, for example, the large N_D concentration of freely moving electrons enhances the probability of the recombination of holes, the amount of holes will drastically sink in donor-doped regions. In such a material, the electrons are considered majority charge carriers and the holes are minority charge carriers.

In most cases even the concentration of the majority charge carriers remains low in the sense that only a low number of available energy levels of the conduction band are occupied. The majority and minority carriers are in dynamic balance through

generation and recombination. It is a general law of reaction kinetics that as long as the concentrations of two reagents are "thin," the elevation of the concentration of one component proportionally lowers the concentration of the other. This is the mass action law, in the actual case resulting in $np = {n_\mathrm{i}}^2$.

The mass action law is maintained through different recombination mechanisms. In direct bandgap materials, electrons and holes can directly recombine in a single recombination event, emitting a photon of zero mass and approximately W_g energy. Such materials are intermetallic compounds like GaAs or GaN. The rate of recombination depends on both n and p concentrations, such it is related to ${n_\mathrm{i}}^2$ and on its temperature dependence described in (4.3). In the literature these recombination mechanisms related to the concentration of two particles are called *bimolecular recombination*, a term borrowed from the description of chemical reactions.

In materials with indirect bandgap, the carriers have to lose their excess momentum during the travel from the conduction band to the valence band; such recombination has to occur through interaction with particles of high mass. These can be primarily vibrating lattice atoms. In order to formalize this interaction, the thermal energy of the vibration is attributed to virtual particles called phonons.

The energy and momentum can also be dispersed in an interaction with dopant atoms and crystal defects. Moreover, special intermediate energy levels, "traps" can be established in the bandgap with intentional placement of appropriate impurity atoms of other elements into the crystal lattice. With offering multiple successive jumps through the intermediate levels, the presence of traps increases the recombination probability. This trap-assisted recombination mechanism is generally known as Shockley-Read-Hall (SRH) recombination.

The concentration of majority carriers is equal to the dopant concentration, a temperature-independent parameter permanently determined by the semiconductor technology. Similarly, the number of crystal defects and traps is less dependent of the temperature. This way the rate of indirect recombination predominantly depends on the concentration of minority carriers, which is governed through the mass action law on the temperature dependence of the n_i intrinsic carrier concentration. These recombination mechanisms are frequently referred to as *monomolecular recombination*.

SRH recombination can be seen as being in a halfway between different mechanisms; the number of traps which are used as intermediate levels for the travel of electrons is present in unvarying number, but at the end electrons recombine holes.

The high number of majority carriers and the reduced number of minority carriers in doped semiconductors also shift the average W_F Fermi level from the W_i halfway position of the bandgap toward the band edges where majority carriers reside. From the M-B distribution approximation

$$n = n_\mathrm{i} e^{\frac{W_\mathrm{F} - W_\mathrm{i}}{kT}}, \qquad p = n_\mathrm{i} e^{\frac{W_\mathrm{i} - W_\mathrm{F}}{kT}}, \tag{4.6}$$

and n or p are tied to the N_A or N_D concentration of dopants.

The particle concentration of charge carriers can be kept at steady high n_1 or p_1 level with an appropriate charge transfer mechanism into a region of the semiconductor. If other neighboring regions are of lower carrier concentration, the random thermal movement of particles results in a net particle diffusion toward those regions. The movement of charged particles represents a J_{diff} diffusion current.

Formerly it was observed that the E electric field represents a force driving the charged particles, while their motion is hindered by "collisions," interaction with obstacles in the lattice. The obstacles were condensed into a μ mobility (or rather into $1/\mu$).

It is less obvious but still true that concentration difference of randomly moving particles on two sides of a surface corresponds to a force.[1] The velocity of the particles is v_{th}, and the obstacles against their movement are the same in the drift and diffusion current mechanisms.

Introducing now an empirical D diffusion constant to consider the obstacles against the particle stream induced by the concentration gradient, we get that

$$J_{\text{ndiff}} = q \cdot D_{\text{n}} \cdot dn/dx, \qquad J_{\text{pdiff}} = q \cdot D_{\text{p}} \cdot dp/dx, \tag{4.7}$$

again the obstacles rather correspond to $1/D$. In (4.7) it is exploited that the particle flow streams toward lower concentrations, but the electron has $-q$ charge.

The relation between μ and D can be easily constructed in a simple "thought experiment" in which a dn/dx concentration gradient is caused by an E field which is maintained until the drift and diffusion current components equalize. The result is the well-known Einstein relation, $D_{\text{n}} = \mu_{\text{n}} \cdot V_{\text{T}}$, $D_{\text{p}} = \mu_{\text{n}} \cdot V_{\text{T}}$.

The $V_{\text{T}} = kT/q$ temperature-dependent factor is known as thermal voltage. This quantity is of principal importance in the temperature dependence of the characteristics of semiconductor devices. Around room temperature (300 K), the thermal voltage can be calculated, with $k = 1.38 \cdot 10^{-23}$ J/K and $q = 1.6 \cdot 10^{-19}$ As, one gets $V_{\text{T}} = 26$ mV.

4.1.3 The pn Junction

When in a single crystal a p-type and an n-type semiconductor are brought together, the large difference in electron concentration of the two sides causes a diffusion current of electrons from the n-type material across the metallurgical interface into the p-type material. Similarly, the difference in hole concentration causes a diffusion current of holes from the p-type to the n-type material. Due to this diffusion process, the region at the interface becomes almost completely depleted of mobile charge carriers. The gradual depletion of the charge carriers gives rise to a space charge created by the charge of the ionized donor and acceptor atoms that is not

[1] This force drives, for example, the steam engine.

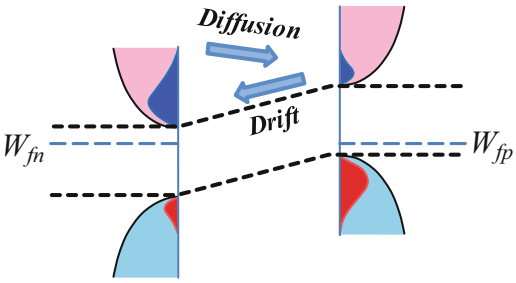


Fig. 4.1 A pn junction in equilibrium. The diffusion and drift components of the current are in balance in the depletion region. The V_D diffusion voltage corresponds to the energy difference either in the conduction or the valence band

compensated by the mobile charges any more. This region of the space charge is called the *space-charge region* or *depleted region* and is schematically illustrated in Fig. 4.1. Regions outside the depletion region, in which the charge neutrality is conserved, are denoted as the quasi-neutral regions.

The space charge in the depleted region results in the formation of an internal electric field which forces the charge carriers to drift in the opposite direction than the concentration gradient. The diffusion currents continue to flow until the forces acting on the charge carriers, namely, the concentration gradient and the internal electrical field, compensate each other. The driving force for the charge transport does not exist any longer, and no net current flows through the pn junction.

Integrating the electric field along the depleted region yields the *contact potential* between the two material types, which is called at semiconductors, referring to the process which established it, the V_D *diffusion potential* (also called built-in potential). The integral can be calculated only knowing the actual doping profile, the concentration of donors and acceptors around the metallurgical interface. However, V_D can be also established directly from the energy levels in equilibrium.

Each electron which stepped from the n region into the p region caused a sinking in the energy on the p side and an elevation of it on the n side; the process ended when the average electron energy, the W_F Fermi level, equalized over the whole structure. From the Maxwell-Boltzmann approximation of the Fermi-Dirac particle distribution, it directly follows through (4.6) that

$$V_D = [(W_F - W_i)|_{n\ side} - (W_F - W_i)|_{p\ side}] = \frac{kT}{q} \cdot \ln\left(\frac{N_A \cdot N_D}{n_i^2}\right) \tag{4.8}$$

Equation (4.8) yields the temperature-dependent diffusion voltage of the junction calculated from the N_A and N_D doping concentrations and the temperature-dependent intrinsic carrier concentration.

In Fig. 4.1 the charge carrier densities on the two sides of the depletion region and the band edges are shown, in a pn junction in equilibrium. The densities of available states for electrons and holes are shown as light red and light blue shapes; the dark blue and red "bubbles" correspond to the actual particle densities at a given energy

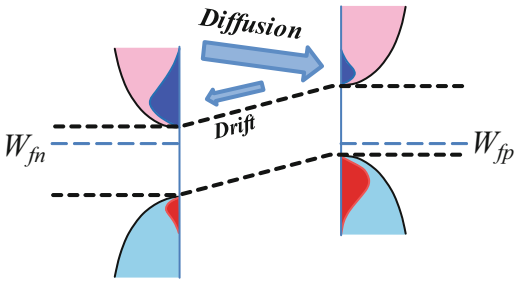


Fig. 4.2 A forward-biased pn junction

level. The n and p total concentrations can be calculated integrating the "bubbles" along the conduction and valence band. The V_D diffusion voltage corresponds to the energy difference either in the conduction or the valence band.

When an external V_F voltage is applied to a pn junction, the potential difference between the n-type and p-type regions will change, and the electrostatic potential across the space-charge region will become $V_D - V_F$ (Fig. 4.2).

The height of the energy barrier which hinders the flow of the majority carriers of one side toward the other side is diminished by V_F; the tail of higher-energy electrons (high end of the dark blue bubble) extends over the barrier with exponentially growing number of electrons. Assuming a few conditions, such as the number of injected carriers remains significantly below the native majority carriers at both sides and no significant recombination occurs in the short depletion layer, the I–V characteristics of a pn junction can be easily derived.

This is a train of thought of simple elementary steps, but because of its length, we have to refer to [15]. The calculation yields the expected exponential growth of the current in a forward-biased *pn* junction, the ideal Shockley equation:

$$I_F = I_0 \cdot \left(e^{\frac{V_F}{V_T}} - 1 \right). \tag{4.9}$$

4.2 Resistive Heaters and Sensors, Active Devices

The previous section focused on the temperature-dependent quantities in semiconductor materials and on the characteristics of the pn junction composed of semiconductor regions of different doping. Still, this overview already contributes to understanding the operation of many actual devices used as heaters and sensors in thermal transient testing, such as the following:

Resistive Sensors

Equation (4.5) describes the drift current in semiconductors. It has to be noted that metals do not differ from highly doped semiconductors in that aspect that a high number of electrons ("majority carriers") are available in them next to an abundance

of available empty states which can be occupied. The equation reveals the origin of the Ohm law in both metals and semiconductors and indicates that the probability of interaction of the charge carriers with lattice atoms grows at higher temperature. This explains the positive thermal coefficient of series resistance in diodes and MOS devices, resistive heaters, and resistor sensors.

The temperature dependence of $1/\mu$ can be well approximated by an exponential function. For example, the popular PT100 platinum sensor has 100 Ω resistance at 0 °C, and it grows by 385 ppm/K with the change of temperature.

Resistive Heaters

Equation (4.5) also indicates that in a unit volume in which an E electric field causes a flow of J current density, a thermal power of $P_V = E \cdot J$ is generated. This results in the whole volume of a resistive material region $P = V \cdot I = I^2 \cdot R$ heat generation (V is the voltage drop across the region and I is the current flowing through it, $R = V/I$.)

It has to be noted that recently novel *active devices* have been introduced based on sudden drastic, thermally influenced change of the electric conductivity. For example, vanadium dioxide (VO_2) transits from semiconductor phase into metal phase at around 67 °C, changing its electrical conductivity from $\sigma = 10^{-1}$–10^{-2} S/cm to $\sigma = 10^3$–10^4 S/cm [102].

Thermistors with Negative Thermal Coefficient

These temperature sensors are simply a piece of semiconductor (sintered metal oxide) in which the number of charge carriers quickly grows with the temperature, as indicated by (4.3).

Thermocouples and Thermoelectric Cooler (TEC, Peltier) Devices

In (4.8), we introduced the V_D contact potential (diffusion potential) as the difference of the Fermi levels in materials when they are in contact and equilibrium. The formula is equally valid for two metals in contact; the contact potential is the difference of the Fermi levels; just the "depleted region" is infinitesimally thin, or not that thin at all when their connection is oxidized or contaminated.

This contact potential cannot be measured by a voltage meter, or used as voltage source in a circuit in which all components are at the same temperature, because the other components are also constructed of metals and semiconductors and the sum of V_D values in a closed circuit equals to zero. However, as (4.8) indicates, keeping different components at diverse temperatures, the different $V_D(T)$ temperature-dependent values in series result in a net temperature-dependent voltage (Seebeck effect).

In a reverse effect, a flow of current generates temperature difference at the junction of two materials (Peltier effect). In TEC devices the net resulting temperature can be calculated from a superposition of the Ohm, Seebeck, and Peltier effects.

Active Devices

The previous section expounded that when there is a potential barrier in the way of the flow of charge carriers (current), then a tiny shift in the barrier height causes a

major alteration in the current. Beyond the barrier height, the current also depends on the presence of charge carriers, and such depends heavily on the temperature.

In the last 120 years of electronics, the operation of all active devices has been based on the shifting of a W_b barrier by an action needing low energy. The shift regulates the stream of (charged) particles, resulting in high energy change.

This operation is analogous to the function of a valve which can modify a flow of a substance with minor energy investment. This analogy is reflected in the names of active devices from the beginning, for example, vacuum tubes are referred to as *valve*, and analogous solid-state active devices are *TRAN*-sfer re-*SISTORS*, such as BJT, FET, and HEMT.

4.3 Diodes

As a consequence of the previous assumptions in Sect. 4.1, the forward current (I_F) – forward voltage (V_{Fpn}) characteristics of a pn junction follows the Shockley equation:

$$I_F = I_0 \cdot \left(e^{\frac{V_{Fpn}}{mV_T}} - 1 \right). \tag{4.10}$$

The V_T and I_0 parameters are temperature dependent:

$$V_T = kT/q, \tag{4.11}$$

and

$$\begin{aligned} I_{01} &\sim n_i = G \cdot T^{3/2} \cdot e^{\frac{-W_g}{2kT}} \\ I_{02} &\sim n_i^2 = G^2 \cdot T^3 \cdot e^{\frac{-W_g}{kT}}, \end{aligned} \tag{4.12}$$

where I_{01} and I_{02} are the saturation current constituents maintained by monomolecular and bimolecular recombination mechanisms, respectively.

The two recombination components are typically reflected, *more or less justified* in a single m parameter:

$$I_0 = G_{sum} \cdot T^{3/m} \cdot e^{\frac{-W_g}{mkT}}. \tag{4.13}$$

In practical measurements, as justified in Sects. 5.4 and 6.1, diodes are typically driven by a controlled I_F forward current, for electric and thermal stability reasons. The V_{Fpn} forward voltage at controlled I_F forward current will be

$$V_{\mathrm{Fpn}} = m \cdot V_{\mathrm{T}} \cdot \ln \frac{I_{\mathrm{F}}}{I_0} \tag{4.14}$$

Real diodes have characteristics similar to (4.14) at lower current. At higher current, additional recombination and diffusion mechanisms modify the basic effect expressed by the Shockley equation.

Moreover, the ohmic regions in the semiconductor at the two sides of the junction and the wiring in the package add an additional voltage drop to the total V_{F} forward voltage on the device. Cumulating these secondary effects into an R_{S} internal electrical series resistance (and neglecting the very small – I_0 term), we get

$$V_{\mathrm{F}} = V_{\mathrm{Fpn}} + V_{\mathrm{FRs}} = m \cdot V_{\mathrm{T}} \cdot \ln \frac{I_{\mathrm{F}}}{I_0} + I_{\mathrm{F}} \cdot R_{\mathrm{S}}, \tag{4.15}$$

m is the device specific constant from (4.13) called *ideality factor*. An ideal diode is supposed to have $m = 1$.

Real silicon diodes may have various m values. Signal diodes are heavily doped with impurities establishing trap levels in order to promote SRH recombination; their ideality factor is nearer to $m = 2$. The shorter lifetime of charge carrier results in faster switching operation. In power devices SRH recombination enhances current leakage at high reverse voltages and reduces injected current density, it is advantageous to keep the trap density low, and their ideality factor is nearer to $m = 1$.

The number of available energy states at the edge of the bandgap in semiconductors differs from the number of states in the partially filled conduction band of metals. For this reason the $3/m$ exponent in (4.13) changes to $2/m$ in Schottky diodes, where the potential barrier is formed at the interface of a metal and a semiconductor.

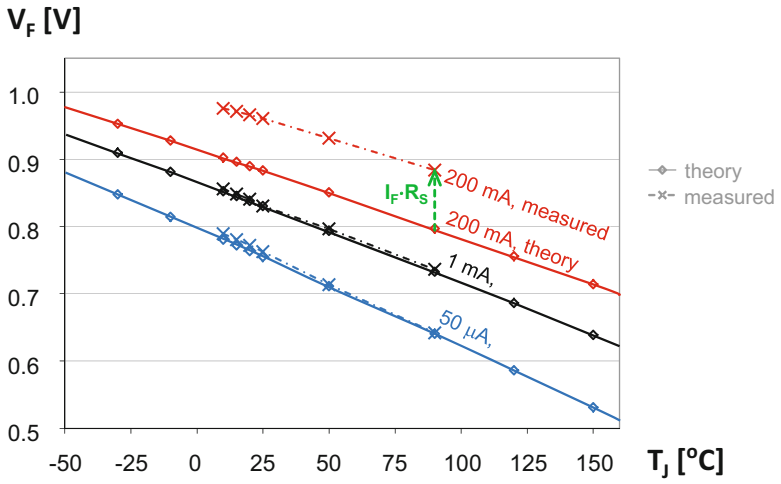


Fig. 4.3 Temperature dependence of the forward voltage, diode at different bias currents

Equations (4.13) and (4.15) offer a good analytical approach for the nonlinear diode characteristics in a broad current and temperature range.

Figure 4.3 shows the V_F forward voltage of a silicon diode at different bias currents (50 µA, 1 mA, 200 mA). The **x**-shaped marks in the plot show measured V_F values of the actual diode at various temperatures from 10 °C to 90 °C. Using the measured values at 1 mA bias, we calculated the I_0 and m values and produced analytic curves (dash-dotted lines) for other bias than 1 mA and in an extrapolated temperature range (−50 °C to 150 °C).

The curves, corresponding to the analytic equations above, are obviously nonlinear, but show only very small nonlinearity over a broad temperature range. The R_S series resistance can be ascertained from the difference of the modeled and measured values at higher currents, as shown by the green $I_F \cdot R_S$ vector in the chart.

4.3.1 Differential Properties of the Diode Characteristics

In many cases we study the change of the *I–V* characteristics at small current or temperature changes. The quantities depicting this partial variations around an operating point are the $R_D(I_F, T_J) = dV_F/dI_F$ *differential (electric) resistance* and the $S_{VF}(I_F, T_J) = dV_F/dT$ *temperature sensitivity*. Sometimes in engineering the so-called K factor is used, $K = 1/S_{VF} = dT/dV_F$.

Composing the dV_F/dI_F derivative of (4.15) one obtains that the differential resistance of a diode is approximately inversely proportional to the I_F forward current:

$$R_D = m \cdot V_T/I_F + R_S \tag{4.16}$$

In most cases R_D is rather low, 26 Ω at 1 mA forward current, or 26 mΩ at 1 A forward current for an "ideal" diode ($m = 1$) at room temperature. This ensures high noise tolerance of thermal measurements as it will be shown in Sect. 5.7.

In a few steps, compiling the temperature derivative of (4.14), at a proper point inserting the dI_0/dT derivative from (4.13), it can be found that

$$S_{VFpn} = \frac{dV_{Fpn}}{dT} = \frac{V_F - 3V_T - W_g/q}{T} \tag{4.17}$$

It is important that m which was present in both (4.13) and (4.14) has not disappeared; it is just incorporated into the V_F term.

A first estimation on the value S_{VF} can be done easily. Around room temperature (300 K), as a consequence of (4.8) and (4.15), the forward voltage of realistic *silicon* diodes in their typical operating points is around $V_F = 500$–800 mV. Inserting this V_F into (4.17), Table 4.1 lists a few calculated S_{VF} values. For the homogeneity of the measures, the bandgap is given as the $V_g = W_g/q$ bandgap voltage; its change with the temperature is estimated based on the empirical Varshni formula.

Table 4.1 Calculated S_{VF} sensitivity values at some temperatures

T [K]	V_{Fpn} [mV]	V_T [mV]	$3V_T$ [mV]	V_g [mV]	S_{VF} [mV/K]
300	500	26	78	1120	−2.33
300	800	26	78	1120	−1.33
400	500	34.5	103.5	1097	−1.75
400	800	34.5	103.5	1097	−1.01

The table can be easily extended to other temperature and forward voltage ranges, and other semiconductor materials of different W_g bandgap. With the actual values it is proven that silicon diodes expose around room temperature and in a broad current range an S_{VF} temperature sensitivity factor between −1 and −2.5 mV/K, a negative temperature coefficient. In wide bandgap materials, V_F and S_{VF} can be significantly higher.

One can express the dependence of S_{VF} on the current density through the Shockley equation. It is generally known that at *higher current density*, the S_{VF} *sensitivity diminishes*, but only modestly through many orders of magnitude as the logarithmic nature of (4.15) indicates.

From a more detailed calculation:

$$\Delta S_{VF} = \frac{V_{F2} - V_{F1}}{T} = \frac{m \cdot V_T(\ln \frac{I_{F2}}{I_{F1}})}{T} = \frac{m \cdot kT(\ln \frac{I_{F2}}{I_{F1}})}{q \cdot T} \tag{4.18}$$

Surprisingly, ΔS_{VF} does not seem to depend on temperature. If still it does, it is just because the Shockley equation is merely an approximation of physical reality. I_F is typically the forced quantity; in this way, it is an external parameter not dependent on temperature. Thus

$$\Delta S_{VF} = \frac{m \cdot k}{q} \cdot \ln \frac{I_{F2}}{I_{F1}} \tag{4.19}$$

The k/q constant is 86 μV/K. According to this; and considering the logarithm of the current ratio, 10% decrease in I_F current will result in $8.22 \cdot 10^{-3}$ mV/K growth in the temperature dependence (with $m = 1$ ideality factor, and any temperature).

It has to be noted that S_{VF} is a negative number, and so is ΔS_{VF} because $W_g/q > V_F$. If I_F grows a full decade, ΔS_{VF} changes by $k/q \cdot \ln(10) = -0.2$ mV/K. At a growth by 5 decades, $\Delta S_{VF} = k/q \cdot \ln(100000) = -0.99$ mV/K.

So far we have proven that V_F is of negative S_{VF} temperature coefficient which slightly diminishes at high current densities. However, the R_S component can also have negative temperature coefficient until secondary diffusion effects around the depleted region and included into the series resistance dominate. R_S turns to positive temperature coefficient at high currents due to the ohmic nature of the semiconductor regions adjacent to the pn junction.

Table 4.2 Calculation of the Z temperature-induced relative growth of the forward current at various constant V_F forward voltages, expressed in percentage

V_F [V]	$Z_{@T\,=\,300\ K}$ (%)	$Z_{@T\,=\,400\ K}$ (%)
0.4	10.27	5.80
0.5	8.98	5.07
0.6	7.70	4.35
0.7	6.41	3.63
0.8	5.12	2.90

The change of V_{Fpn} can also be calculated analytically for larger I_F current changes. Equation (4.15) yields for large current differences in an ideal diode:

$$V_{F1} - V_{F2} = V_T \cdot \ln(I_1/I_2) + R_S \cdot (I_2 - I_1) \tag{4.20}$$

For calculating the change of the voltage caused by the first, ideal term one can use the approximate values $V_T = 26$ mV (at room temperature) and $\ln(10) \approx 2.3$. At ten times higher current, the V_{Fpn} forward voltage on the junction grows by approximately $V_T \cdot \ln(10) = 60$ mV. Similarly, at hundred times higher current, the forward voltage grows by 120 mV, and at ten thousand times higher current, it grows by 240 mV.

When for some reason a pn junction is driven by steady V_{Fpn} forward voltage, the current grows extremely fast with increasing temperature. In order to obtain a manageable expression for this change, it is reasonable to relate the current change to the current itself. From the Shockley equation, the Z ratio of the temperature-induced current change and the forward current itself can be expressed as

$$Z = \frac{1}{I}\frac{dI}{dT} = \frac{1}{mT}\left(3 + \frac{V_g - V_F}{V_T}\right) \tag{4.21}$$

The Z ratio of the current growth at fixed V_F forward voltage is calculated in Table 4.2 with V_T and V_g values taken from Table 4.1, at 300 and 400 K absolute temperature and $m = 1$ ideality factor.

A further consequence of the high dependence of the forward current on temperature at fixed forward voltage is the *current crowding* effect. This may undermine the accuracy of the forward voltage to temperature calibration and of the thermal transient measurement process because it causes an uncertainty in the interpretation of the results.

A diode can be imagined as many elementary pn structures in parallel. As it was presented in Sect. 3.5.1 when uniform power is applied on the chip surface, then a bell-shaped nonuniform temperature distribution develops on it. At high currents uniform powering can be achieved if the series R_S resistance has positive thermal coefficient. Switching to a low measurement current, the elementary junctions have the same V_F forward voltage as they are connected in the chip. This causes the current threads to cumulate into the hotter elementary structures.

Electrothermal simulation tools can properly model the temperature and current distribution on the chip surface, but they yield the result only for one certain

Table 4.3 Calculation of the current crowding of a low I_F forward current at temperature inhomogeneity on the diode surface

ΔT [°C]	$K_{@Z = 4\%}$	$K_{@Z = 5\%}$	$K_{@Z = 6\%}$
1	1.04	1.05	1.06
2	1.08	1.10	1.12
3	1.12	1.16	1.19
4	1.17	1.22	1.26
6	1.27	1.34	1.42
8	2.36	2.46	2.56
10	3.45	3.58	3.72

geometry and one powering. An estimation on the current crowding can be done simply, supposing that the temperature difference on the surface is typically a few centigrade between the hottest and coldest elementary diodes.

Table 4.2 indicates that 1 °C temperature difference causes a few percent change in the forward current.

Table 4.3 presents the typical K ratio of the current density between the coldest elementary junction and an other junction in parallel, which is hotter by ΔT. Supposing 4–6% relative Z growth of the current density, $K = (1+Z)$ for 1 °C difference. More centigrade difference multiplies this change; the K current ratio between elementary diodes follows a "compound interest" scheme.

The table proves that a temperature gradient of a few degrees makes the current cumulate into the hottest surface region of a diode.

4.3.2 Negative and Positive Thermal Coefficients in the Diode Characteristics

In the next example, the sections of negative and positive thermal coefficient in the diode characteristics are analyzed in a simulation experiment.

Example 4.1: Negative and Positive Thermal Coefficients in the Diode Characteristics

In the example we use the basic diode model of the popular LTSpice software for modeling the internal pn junction. Realistic devices have positive thermal coefficient at high currents, which can be attributed to the decrease of mobility in bulk semiconductor and to the wiring within the package. These effects are modeled by external resistors of 0.15 Ω resistance and a positive thermal coefficient of 0.2%/K (shown as 0.15 tc = 2 m in the figure).

Separate diodes and resistors are used to represent the devices kept at $T_J = -25$ °C, 25 °C, 75 °C, and 125 °C temperatures (Fig. 4.4).

The resulting temperature-dependent diode characteristics are shown in Fig. 4.5; the negative and positive thermal coefficient regions and the zero thermal coefficient point ZTC can be observed.

(continued)

Example 4.1 (continued)

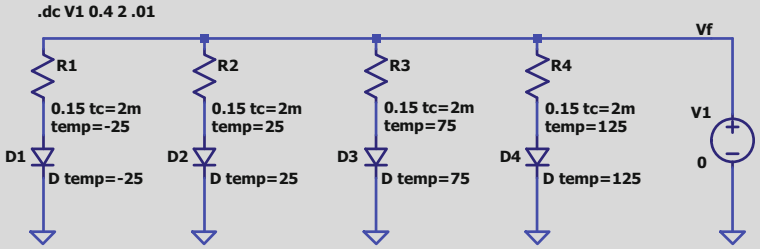


Fig. 4.4 Circuit scheme for analyzing the temperature and current dependence of the diode characteristics in an LTSpice simulation. The resistor added to the standard diode model represents the external portion of the series resistance

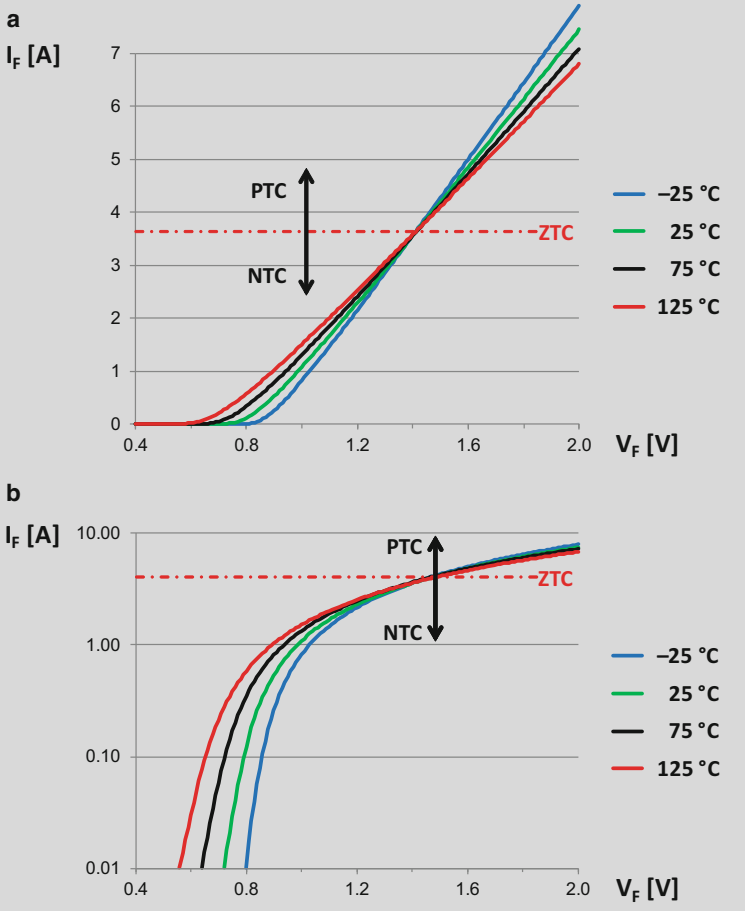


Fig. 4.5 Simulated temperature-dependent diode characteristics from the scheme of Fig. 4.4, plotted with linear (**a**) and logarithmic (**b**) current scale

We can identify an NTC region at lower currents dominated by the junction, a PTC region at higher currents dominated by the series resistance and a ZTC zero temperature coefficient point at ~1.4 V, 3.7A.

4.3.3 *Electrothermal Model of a Diode*

In Fig. 4.6 an LTSpice implementation of the electrical characteristics of a diode is shown. The Eqs. (4.10)–(4.13) are realized as controlled sources. In order to simplify the formulae in the source definitions, the conversion between centigrade and kelvin is realized as a constant "voltage" source and V_T is provided by a "temperature" controlled voltage source. R_S is constant in this simple model, a more complex representation can be a current controlled voltage source, with, e.g, a V = I (IF) * ({R0}+{dR}*V(Tj)) formula.

The temperature and forward voltage response of a diode in the thermal environment represented by the simple compact thermal model of Fig. 4.7, can be monitored, as shown in Fig. 4.8.

First, a heating can be followed. The current forced through the diode grows linearly from 10 mA to 1 A in the first 1 µs. (The current is multiplied by 10 in order to make it visible in the chart. The curvature corresponds to linear growth in log-lin scale.)

In the next second the growth of the temperature and the decay of the forward voltage can be observed. The "bumps" of the curve correspond to the two thermal time constants in Fig. 4.7, τ_1=1 ms, τ_2=100 ms.

The forward voltage on the diode corresponds to the V_g bandgap voltage parameter of the I_0 source, which was set to 1.1 V.

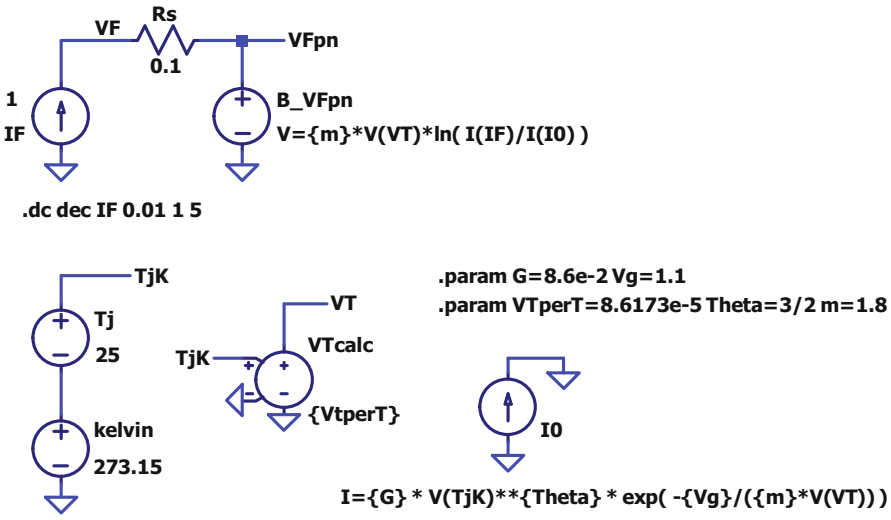


Fig. 4.6 LTSpice implementation of the electrical characteristics of a diode with the help of controlled sources

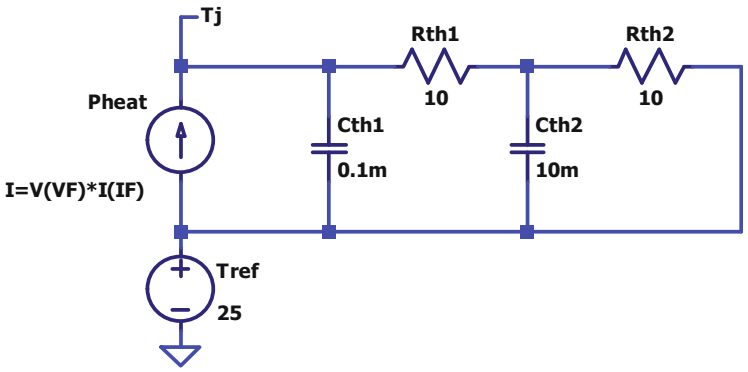


Fig. 4.7 Thermal subcircuit of the diode; heat source and a simplified equivalent Foster chain

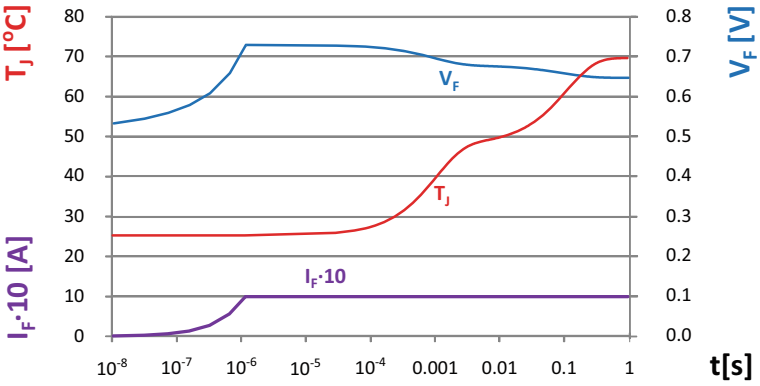


Fig. 4.8 Transient simulation of the equivalent LTSpice circuit. The diode is switched on to $I_{heat} = 1$ A heating current. Values of I_F, T_J and V_F are shown

4.4 MOS Transistors

A correct analysis of thermal effects in field effect devices can be carried out using simulator programs following the flow of particles in the crystal, charge carriers, and phonons representing the thermal behavior of the lattice atoms. An example of such tools is presented in [101].

The characteristic quantity of insulated gate devices is the V_{th} threshold voltage, above which the device current grows in a quadratic way. Because of the quick growth, the definition of the threshold shall be not very sharp. Practical guides define the presence of an I_D drain current of 1 μA or 1 mA or similar for an actual device as current limit, and all these yield very similar V_{th} threshold voltages.

A more physical approach can be based on the operation of the device. Figure 4.9 presents the cross section of a (lateral) MOS structure. The n^+ doping profiles at the source and drain diffusion (S and D in the figure) induce the appearance of depleted regions around the metallurgical interface of p- and n-type semiconductor (Fig. 4.9b).

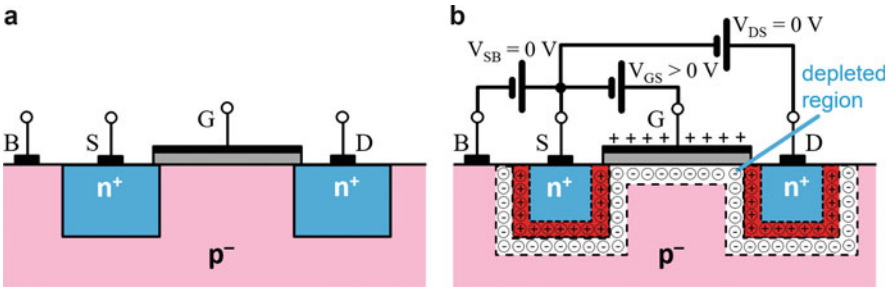


Fig. 4.9 Schematic drawing of a MOSFET device, (**a**) cross section with doping profiles (**b**) depleted area around pn junctions of the source and gate diffusion and under the gate

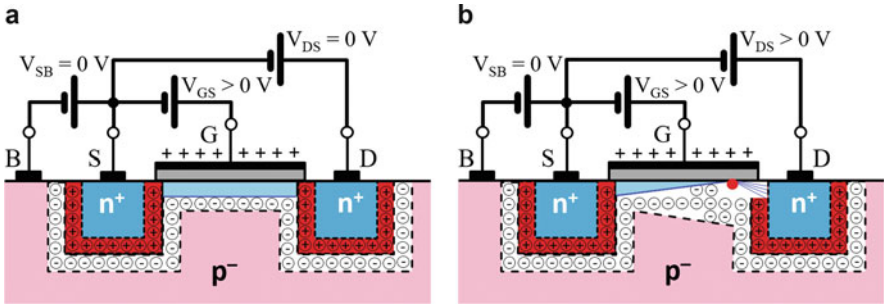


Fig. 4.10 Schematic drawing of a MOSFET device, (**a**) inverted channel under the gate biased by higher V_{GS} voltage (**b**) increased depleted area around the drain diffusion and pinch-off at elevated V_{DS}

An external voltage applied *between* the gate (G in the figure) and the bulk semiconductor, which is most frequently connected to the source, bends the band structure on the semiconductor-oxide interface. At a high enough V_{GS} voltage, the bending of the potential causes the W_i intrinsic level to turn to the opposite side of the W_F Fermi level, as explained more in detail in [15]. The surface will be *inverted*; in the actual case, it starts behaving as p-type semiconductor instead of the bulk n type (Fig. 4.10a).

The accepted interpretation of inversion is that a semiconductor is *fully inverted* when the n concentration below the gate reaches the original p concentration of the bulk material; the conducting *channel* below the oxide reaches the electric conductivity of the bulk.

For composing the device characteristics, it is easier to describe the equations as potentials, rather than energies. Using the $\Phi = q \cdot W$ notation for potentials, the threshold voltage of an n-channel device can be written as

$$V_{th} = \Phi_{GC} - 2\Phi_{FS} - \frac{Q_{B0}}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \tag{4.22}$$

the inversion channel appears at $2\Phi_{FS}$ bending of the Fermi level, that is, the electron concentration in the channel is the same as the hole concentration of the p-body. Φ_{GC} is the contact potential between the gate material and the p-body, $\Phi_{GC} = \Phi_{FS} - \Phi_{FG}$, with a $\Phi_F = q \cdot W_F$ notation.

Φ_{FS} is dependent on T temperature. From Sect. 4.1, at N_A acceptor dopant concentration of the body:

$$\Phi_{FS} = (kT/q) \cdot \ln(n_i/N_A), \quad \text{and} \quad n_i^2 = G \cdot T^{3/2} \cdot e^{-W_g/kT} \tag{4.23}$$

(as before, W_g is the bandgap in the semiconductor, k is the Boltzmann constant and q is the elementary charge).

In (4.22) C_{ox} is the gate oxide capacitance per unit area, and Q_{ox} is the charge per unit area on the surface due to oxide traps, interface states, etc. At least these two latter are not strongly temperature dependent. However, the charge of ionized acceptors in the depleted layer under the channel is, as calculated from the length of the depletion region at V_F voltage, from [15]:

$$-\frac{Q_{B0}}{C_{ox}} = \frac{\sqrt{2qN_A\varepsilon_{Si} \cdot |-2\Phi_{FS}|}}{C_{ox}} \tag{4.24}$$

where ε_{Si} is the permittivity of the silicon. All these temperature-dependent factors result in an approximately −4 mV/K temperature coefficient of V_{th}.

For calculating R_{DSON} (the resistance of the channel when the device is switched on) we introduce the $V_{OV} = V_{GS} - V_{th}$ "overdrive" voltage, which produces electrons in the channel. If C is the capacitance of the gate/body structure,

$$C = (\varepsilon_{ox}/d_{ox}) \cdot W \cdot L = C_{ox} \cdot W \cdot L \tag{4.25}$$

the total charge in the channel of W width and L length is $Q = C \cdot V_{OV}$. Applying a small V_{DS} voltage to the device; an electric field of $E = V_{DS}/L$ appears along the channel. Electrons in the channel move from the source to drain with the drift velocity:

$$v_{drift} = \mu_n E = \mu_n (V_{DS}/L) \tag{4.26}$$

In a typical power MOS, we can count with the following numeric data when it is switched "on": $\mu_n = 1000$ cm^2/Vs, $V_{DS} = 100$ mV, $L = 1$ µm $= 10^{-4}$ cm.

With these data from (4.26), we get $E = 1$ kV/cm and $v_{drift} = 10^6$ cm/s. This superposed v_{drift} is low; the speed of the random thermal movement of electrons at room temperature from the $½\ m^* \cdot v^2 = 3/2\ kT$ formula is over 10^7 cm/s.

In case of a saturated MOSFET, most dissipation occurs between the pinch region and drain (right side of Fig. 4.10b), and here a few "hot" electrons, that is, electrons of high speed, carry the same current which is forwarded by plenty of carriers of slow v_{drift} in the open channel.

From the above equations, the so-called modified Shichman-Hodges equation for the MOS characteristics can be derived [12, 14]. It distinguishes between *subthreshold* mode, *triode* mode, and *saturation* mode.

The separation curve between the triode and saturation range is $V_{DS\_sat} = V_{GS} - V_{th}$. At this gate voltage, the channel thickness goes to zero; this is the pinch-off situation.

The characteristics in different operation ranges can be formulated as:

- $V_{GS} \leq V_{th}$, $I_F = 0$, cutoff/subthreshold operation range with low current;
- $V_{GS} > V_{th}$, $V_{DS} \leq V_{DS\_sat}$, linear/ohmic/triode operation range:

$$I_D = \mu C_{ox} \cdot \frac{W}{L} \cdot \left[(V_{GS} - V_{th}) \cdot V_{DS} - \frac{V_{DS}^2}{2} \right] \tag{4.27}$$

- $V_{GS} > V_{th}$, $V_{DS} > V_{GS} - V_{th}$, saturation operation range:

$$I_D = \mu C_{ox} \cdot \frac{W}{2L} \cdot \left[(V_{GS} - V_{th})^2 \right] \cdot \left[1 + \lambda_L \cdot (V_{DS} - V_{DS\_sat}) \right] \tag{4.28}$$

Condensing $\mu \cdot C_{ox}$ into a K_p parameter, and denoting $V_{GS} - V_{th}$ as V_{OV}, one gets the more convenient form:

for $V_{OV} > 0$, $V_{DS} \leq V_{DS\_sat}$, linear/ohmic/triode operation range

$$I_D = K_p \cdot \left[V_{OV} \cdot V_{DS} - \frac{V_{DS}^2}{2} \right] \tag{4.29}$$

for $V_{OV} > 0$, $V_{DS} > V_{DS\_sat}$, saturation operation range

$$I_D = \frac{K_p}{2} \cdot V_{OV}^2 \cdot [1 + \lambda_L \cdot (V_{DS} - V_{DS\_sat})] \tag{4.30}$$

The λ_L empirical parameter describes the shortening of the MOS channel due to the voltage drop between the pinch and the drain.

Figure 4.11 presents an LTSpice simulation with a generic MOSFET model. The main parameters are set to $K_p = 60$ A/V^2, $V_{th} = 2$ V with a β negative thermal coefficient of $S = 6$ mV/K, and channel-length modulation parameter of $\lambda = 0.1$ 1/V.

Figure 4.12 is the output characteristics of the MOSFET device when simulated in the circuit scheme of Fig. 4.11.

Each curve in the plot shows the I_D drain current value when V_{DS} drain to source voltage is applied. Applying higher V_{GS} voltage on the control pin which is now the gate, one will experience higher current at the same drain-source voltage.

Figure 4.13 is the logarithmic version of the plot.

Focusing now our attention on the conducting open channel, we get for the initial section of Figs. 4.12 and 4.13 where V_{DS} is low:

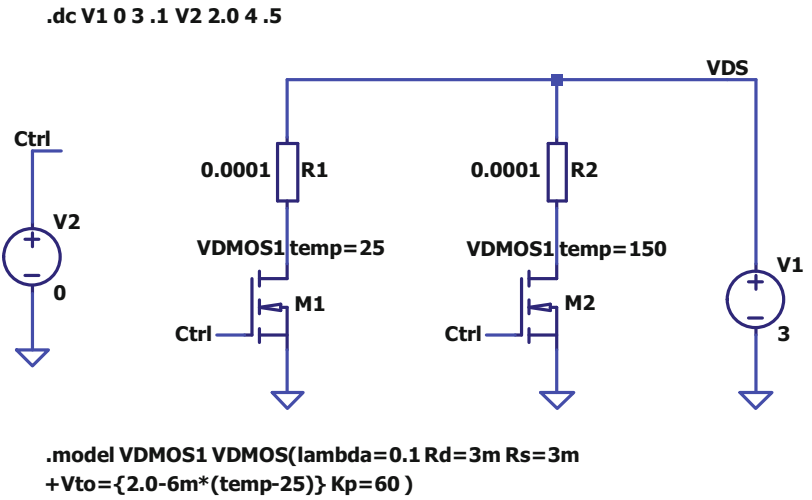


Fig. 4.11 LTSpice analysis of two identical MOS transistors at 25 and 150 °C device temperature. The equation of the MOS model is shown. The λ channel shortening parameter is set to 0.1, the threshold voltage V_{th} to 2 V, its S thermal coefficient to −6 mV/K

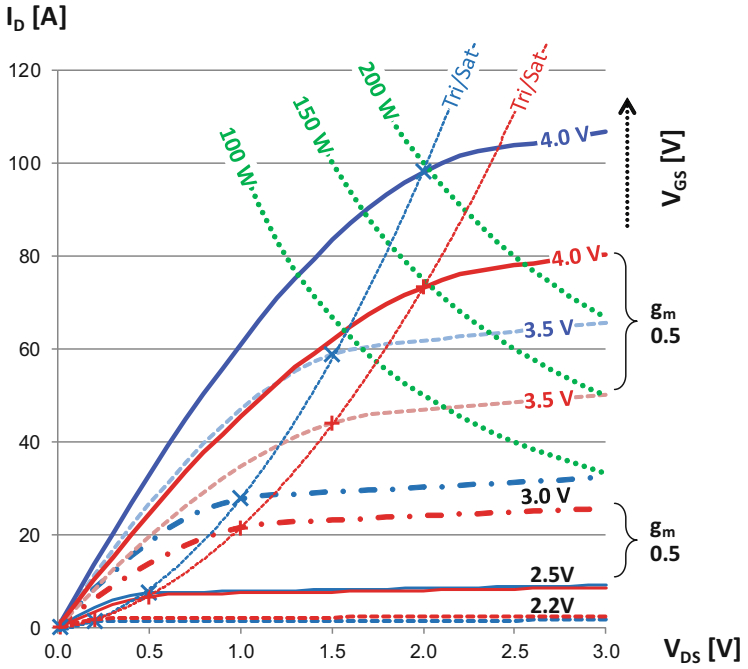


Fig. 4.12 Output characteristics of a MOSFET, cold (blue) hot (red), lin-lin scale. At fixed high current V_{DS} is of positive thermal coefficient, at low current of negative thermal coefficient. The zero thermal coefficient point is around $V_{GS} = 2.5$ V. Triode/ohmic and saturation regions are shown

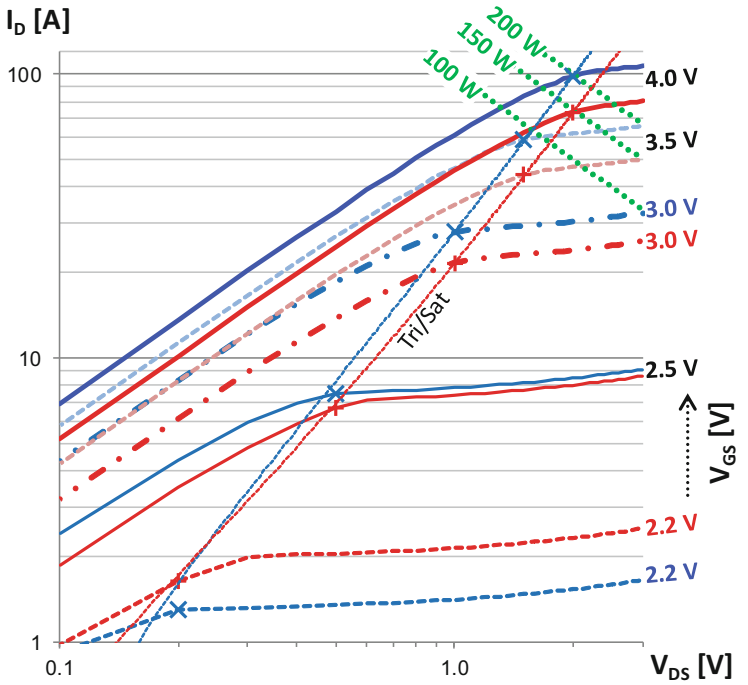


Fig. 4.13 Output characteristics of a MOSFET, cold (blue) hot (red), log-log scale. At fixed high current, V_{DS} is of positive thermal coefficient, at low current of negative thermal coefficient. The zero thermal coefficient point is around $V_{GS} = 2.5$ V. Triode/ohmic and saturation regions are shown

$$I_D = \mu_n C_{ox}(W/L) \cdot V_{OV} \cdot V_{DS}$$
$$G_{DSON} = \mu_n C_{ox}(W/L) \cdot V_{OV} = \mu_n K \cdot V_{OV} \tag{4.31}$$

the device acts as a resistor; its G_{DS} conductance is proportional to V_{OV}. In a smaller range, we can assume a linear dependence of V_{th} and μ_n on temperature with a gradient of α and β, respectively. Fixing the gate overdrive at $V_{OV} = V_{OV0}$, we get:

$$G_{DSON} = \mu_{n0}(1 - \alpha \cdot dT) \cdot K \cdot [V_{GS0} - V_{th0}(1 - \beta dT)];$$

$$\begin{aligned} G_{DSON} &= \mu_{n0}(1 - \alpha \cdot dT) \cdot K \cdot (V_{GS0} - V_{th0} + V_{th0}\beta dT) \\ &= K \cdot (\mu_{n0} - \mu_{n0}\alpha \cdot dT) \cdot (V_{OV0} + V_{th0}\beta dT) \\ &= K \cdot (\mu_{n0}V_{OV0} - \mu_{n0}V_{OV0}\alpha dT + \mu_{n0}V_{th0}\beta dT - D\alpha\beta dT^2 \end{aligned} \tag{4.32}$$

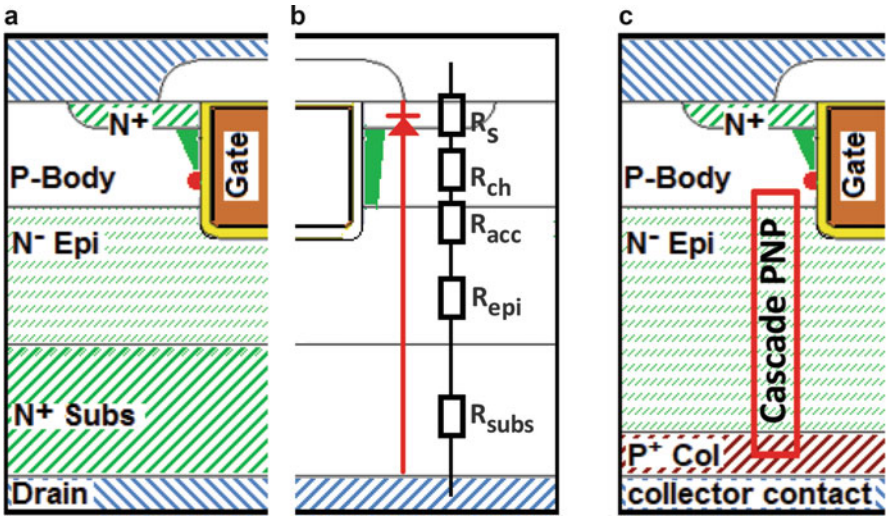


Fig. 4.14 (**a**) Elementary section of a vertical TrenchMOS device, (**b**) constituents of the R_{DSON} resistance (**c**) elementary section of a vertical IGBT device with the cascade pnp bipolar junction transistor shown

Hence, α and β are small; the quadratic term at the end of (4.32) can be neglected; G_{DSON} is determined by a mix of positive and negative temperature coefficients. This can be well observed in Figs. 4.12 and 4.13.

A typical construction of power MOSFET devices is the TrenchMOS structure. A dense net of elementary vertical MOSFETs ensures equal current distribution on the surface. Figure 4.14a, b shows an elementary section of a device.

One can identify the reverse body diode (in red) between the p-body and N^+ source. For illustration only, on the left side, a saturated channel with a pinch point is shown; on the right side, there is a continuous channel corresponding to low V_{DS}.

The figure indicates that R_{DSON} is influenced by the following components:

- R_s: source resistance, low due to high N^+ doping.
- R_{ch}: channel resistance, determined by gate voltage.
- R_{acc}: resistance from the accumulation region.
- R_{epi}: resistance from the epitaxial top layer of silicon; this layer controls the blocking voltage of the MOSFET.
- R_{subs}: resistance from the silicon substrate.

The technological parameters of different regions contributing to the R_{DSON} resistance can be adjusted so as to minimize the channel resistance. However, their setting compromises other device parameters, especially the blocking (breakdown) voltage. In low-voltage transistors for switching purposes, the p-type body can be doped more heavily in such a way substantially decreasing the channel resistance.

The critical technological parameter is the thickness and donor doping of the N^- doped epitaxial layer. Thin epitaxy of higher doping results in low blocking voltage

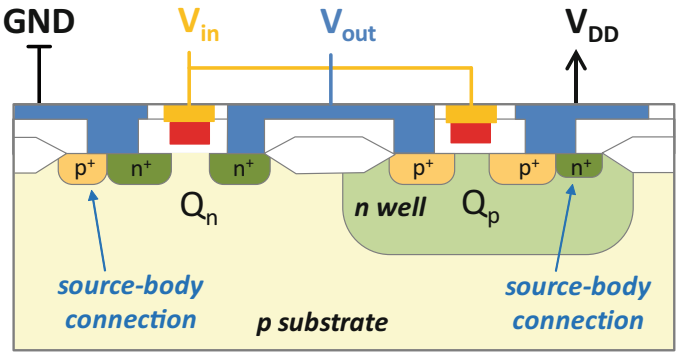
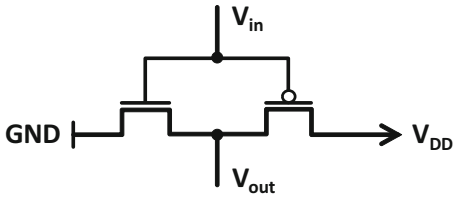


Fig. 4.15 CMOS inverter, cross section. A lateral NMOS transistor in the p substrate and a lateral PMOS transistor in the n well are shown

Fig. 4.16 CMOS inverter, schematic



and low associated R_{epi} series resistance. In high-voltage transistors, this layer is thick and lightly doped; the associated resistance increases R_{DSON} critically.

Depending on gate material and oxide construction, the MOSFET can be of depletion or enhancement type, that is, it does or does not possess a conducting channel at $V_{\text{GS}} = 0$.

The construction of MOS transistors has also a primary role in the thermal transient testing of integrated circuits. Nowadays the majority of these circuits is manufactured in some version of the CMOS technology. Figure 4.15 shows the cross section of an inverter circuit. This is the basic building block for most functionalities in the device, such as logical gates, amplifiers, and output stages. Figure 4.16 presents the two complementary transistors in the equivalent circuit scheme of the inverter.

As the figures illustrate, there is a chain of p and n regions in series between the VDD and GND rails of the circuit. This large diode covers the active area of the whole chip. Reverse biasing this large diode, we get a "dull" but robust and nondestructive way for powering and sensing during a thermal transient test.

4.5 Insulated Gate Bipolar Transistor (IGBT) Devices

In the previous subsection, we stated that high-voltage MOS devices perform poorly because of the long and lightly doped epitaxial layer which is the drain of the transistor. The voltage drop at high current in high-voltage insulated gate devices can be significantly lowered if a large number of charge carriers is injected into the epitaxial layer externally, for example, from an additional pn junction.

IGBTs have many construction variants alike MOS transistors. A frequent form is presented in Fig. 4.14c. This structure is identical to the vertical MOSFET of Fig. 4.14a, except that the N^+ drain region responsible for lowering the series resistance is replaced with a P^+ collector layer, thus forming a vertical pnp bipolar junction transistor. The additional P^+ region forms a cascade scheme of the pnp transistor and the surface n-channel MOSFET (Fig. 4.17).

IGBTs combine the simple control of power MOSFETs due to their insulated gate with the high-current and low-saturation-voltage capability of bipolar transistors.

At high currents IGBTs feature a low V_F forward voltage drop on the base-emitter junction of the pnp device, compared to high-voltage MOSFETs, while those are more advantageous at lower current where a proportional voltage drop occurs just across R_{DSON}.

In IGBTs the injection of minority carriers (holes) from the collector P^+ region into the N^- drift region considerably reduces the resistance of this latter. Moreover, at higher current this injection increases, resulting in the known merely logarithmic increase of V_F.

However, this resultant reduction in on-state voltage drop brings about several disadvantages:

- The series pn junction blocks reverse current flow; IGBTs do not conduct in the reverse direction. In switching applications where reverse current flow is needed, a discrete freewheeling diode is to be placed antiparallel with the IGBT to ensure conduction in the opposite direction.
- The reverse bias rating of the N^- drift region to collector P^+ diode is low, only tens of volts. If in an application a reverse voltage appears on the device, an additional series diode must be used.

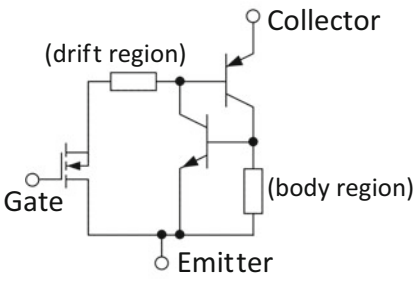


Fig. 4.17 Equivalent circuit scheme of an IGBT device

- The minority carriers injected into the N^- drift region take time to enter and exit or recombine at turn-on and turn-off. This results in longer switching times, and consequently in higher switching loss compared to MOSFETs.

All these drawbacks can be mitigated using MOSFET devices manufactured on wide bandgap materials, where high blocking voltage is granted as a result of the high W_g value.

4.6 Semiconductor Devices on Wide Bandgap Materials

During the previous decades, the devices of power electronics were mostly realized on silicon. This offered a very robust technology working fine even at 150 °C operating temperature. For this reason, the thermal testing standards (e.g., of JEDEC or AQG) have taken for given such features of silicon like

- pn junctions available for powering and sensing
- Mostly linear temperature dependence of parameters
- Availability of normally-off devices for power electronics
- No slowly moving surface charges in MOS structures
- Rather long lifetime of charge carriers in depleted regions

Newly, compound semiconductor devices have been introduced into power applications where their excellent properties can justify their higher cost. They work well at extremes like at 77 K in liquid nitrogen for minimum noise, or at high temperatures up to 350 °C. They show low channel resistance due the high carrier mobility and high blocking voltages, thanks to their wide bandgap.

Silicon carbide (SiC) is a material of a wide and indirect bandgap; for this reason, it has always been intended for power and high-temperature devices.

Gallium nitride (GaN) has a direct bandgap; minority carriers are not present in it as they would recombine after an extremely short lifetime, they have no chance to reach an opposite material layer (i.e., collector). Accordingly, only unipolar transistors can be made of it. On the other hand, on an AlGaN/GaN interface a

Table 4.4 Thermal testing methods for device types realized in certain materials

Material	Si	SiC	GaN	Note
Device				
BJT	St	E	×	
MiSFET	St	Td	Td	Also as MIS gate HEMT
JFET	St	E	Td	Also as Schottky gate HEMT
HEMT	×	×	Td	
IGBT	St	Td	×	
pn diode	St	E	×	
Schottky diode	St	Td	E	

two-dimensional (2DEG) electron gas forms easily and offers low sheet resistances in a high electron mobility transistor (HEMT) device.

Optoelectronics used to be the main market for the direct III–V-semiconductor GaN. In this role it became better and a cheaper and entered the power market.

Table 4.4 lists the existing device types and also hints on the existing knowledge base available for their thermal testing. (In the table the abbreviations mean: St: standard testing method, E: the device exists, ×: the device does not exist, Td: thermal transient is defined and explained in this book.)

The preeminent material properties of these materials are given in [17, 103]. In most parameters, GaN is slightly superior to SiC and provides a three to five times better ON resistance in insulated gate devices. The thermal conductivity of GaN is low that would impede its use in power applications; for this, it is always used as a thin epitaxial layer on a cheap substrate material.

Generally, Eqs. (4.22)–(4.32) also apply for MOSFET devices built on wide bandgap material, resulting in device characteristics similar to Figs. 4.12 and 4.13.

Equation (4.32) predicts that the $R_{\mathrm{DSON}}(T)$ channel resistance has a temperature dependence with sections of positive and negative thermal coefficients, governed by the shrinking of the V_{T} threshold voltage at lower and the increase of the μ mobility at higher temperatures. At practical current levels, the negative thermal coefficient can be observed only at extreme low temperature in silicon devices. A simple test proving this effect in the −196 °C to 25 °C range is presented in [83].

In SiC devices, the turning point between negative and positive temperature coefficient sections can be around room temperature; its actual position depends on the operation point, that is, the V_{DS} drain-source voltage and I_{D} drain current.

Figure 4.18 investigates the temperature dependence of the channel resistance in two MOSFET devices. R_{DSON} is interpreted in a "static" manner, that is, as $V_{\mathrm{DS}}/I_{\mathrm{DS}}$ at low drain-source voltage (see also Chap. 6, Sect. 6.2.1). The chart compares the

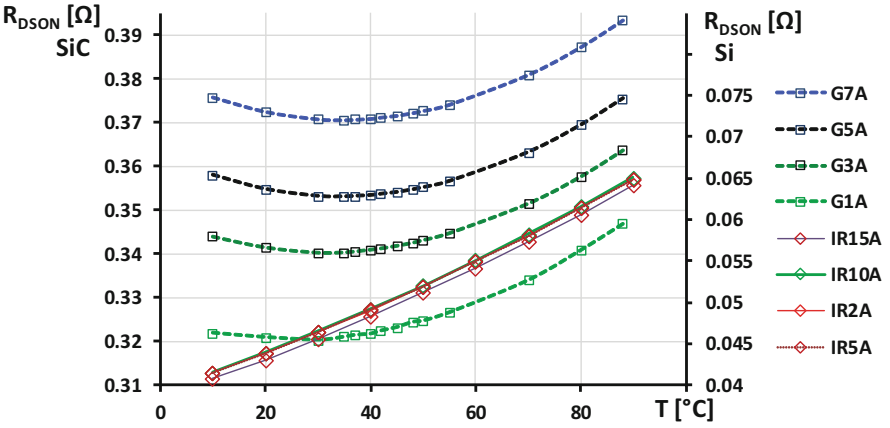


Fig. 4.18 Measured temperature-dependent "static" $R_{\mathrm{DSON}} = V_{\mathrm{DS}}/I_{\mathrm{DS}}$ channel resistance of a SiC MOSFET and a Si MOSFET. The nonmonotonous temperature dependence of R_{DSON} in the SiC device can be observed

popular IRFP260N silicon MOSFET with 40 mΩ nominal channel resistance to the commercially available silicon carbide G3R350MT12D MOSFET, with a claimed channel resistance of 350 mΩ at room temperature and 499 mΩ at 175 °C in its datasheet.

The chart presents the change of the channel resistance of the SiC MOSFET at $I_D = 1$ A, 3 A, 5 A, and 7 A current values (key G1A to G7A, curves scaled on the left vertical axis). The channel resistance of the larger Si MOSFET, which has a ten times lower channel resistance, is shown at $I_D = 2$ A, 5 A, 10 A, and 15 A (key IR2A to IR15A, curves scaled on the right vertical axis).

The higher current dependence and the nonmonotonous temperature dependence of R_{DSON} in the SiC device can be well observed. The turning point from negative to positive temperature coefficient is around 30 °C and shifts toward higher temperatures at higher current.

Conversely, for the Si device, the temperature coefficient of R_{DSON} is positive in the examined 10 °C–90 °C range, and so it remains from deep sub-zero to the maximum rated temperature of 175 °C. The temperature dependence follows an $R_{DSON}(T) = R_{DSON0} \cdot \exp(\alpha_\lambda(T - T_0))$ formula, as a consequence of (4.29) and (2.53).

The R_{DSON} channel resistance is one of the most often used temperature sensitive parameters (TSPs) in thermal transient measurements, described in detail in Chaps. 5 and 6, Sect. 6.2. The nonmonotonous nature of the channel resistance in SiC MOSFETs limits its use as TSP for this device category.

4.7 High Electron Mobility Transistor (HEMT) Devices

HEMT is a type of field effect transistor, where the channel is formed at the interface of two layers (usually AlGaN and GaN) of different bandgap (heterojunction). The layers are grown upon each other on top of a carrier substrate. As the consequence of the abrupt change of the different electric fields at the material interface (polarization fields), a two-dimensional electron gas (2DEG) layer is formed at this which acts as the channel of the transistor. In this 2DEG layer, the electron mobility is much higher than in the bulk material, which makes the device very efficient in high-frequency applications [104, 105]. The substrate material has no direct effect on the electric behavior; however, it affects significantly the price, yield, and the thermal performance. Often used materials are silicon, sapphire, SiC, or GaN. As the channel forms naturally without external biasing, the classic HEMT devices have a depletion mode (normally ON) characteristics; a negative gate voltage needs to be applied to turn the device off. The classic GaN HEMTs have a Schottky barrier gate contact.

HEMT devices are widely used as millimeter wave RF amplifiers; newly they are conquering high-frequency switching mode power conversion applications as well. In order to meet industry requirements, for safer operation and simpler gate drive circuits, there is extensive development of modified structures to achieve optimal performance with enhancement mode (normally normally off) operation [129].

Using band engineering techniques like “recessed gate” [130], “fluorine gate” [131], and thin AlGaN barrier layer [132], the threshold voltage can be increased above 0 V realizing the enhancement mode operation. Recent advancements combining these techniques with additional gate dielectric layer (MIS HEMT) can reduce gate leakage current and decrease switching and conduction losses [129, 133].

In GaN devices, the R_{DSON} channel resistance, related to the sheet resistance of the 2DEG electron gas, has positive thermal coefficient in the temperature range of interest for power devices. However, many constructions suffer from a time-variant behavior, due to the varying amount of trapped charge at off-state at the AlGaN barrier surface (current collapse). Recent constructions aim at reducing this effect which restricts the use of GaN devices in all power switching applications; this also improves the usability of channel resistance as TSP in thermal transient measurements.