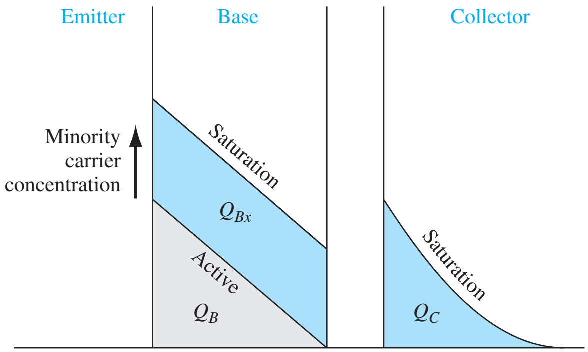

Figure 12.44 | Charge storage in the base and collector at saturation and in the active mode.

change instantaneously. Recall that when the transistor is biased in saturation, both the B–E and B–C junctions are forward biased. The charge Q_{BX} in the base must be removed to reduce the forward-biased B–C voltage to 0 V before the collector current can change. This time delay is called the *storage time* and is denoted by t_s. The storage time is the time between the point at which V_{BB} switches to the time when the collector current is reduced to 90 percent of its maximum saturation value. The storage time is usually the most important parameter in the switching speed of the bipolar transistor.

The final switching delay time is the fall time t_f during which the collector current decreases from the 90 percent to the 10 percent value. During this time, the B–C junction is reverse biased but excess carriers in the base are still being removed, and the B–E junction voltage is decreasing.

The switching-time response of the transistor can be determined by using the Ebers–Moll model. The frequency-dependent gain parameters must be used, and normally the Laplace transform technique is used to obtain the time response. The details of this analysis are quite tedious and are presented here.

12.7.2 The Schottky-Clamped Transistor

One method frequently employed to reduce the storage time and increase the switching speed is the use of a Schottky-clamped transistor. This is a normal npn bipolar device with a Schottky diode connected between base and collector, as shown in Figure 12.45a. The circuit symbol for the Schottky-clamped transistor is shown in Figure 12.45b. When the transistor is biased in the forward-active mode, the B–C junction is reverse biased; hence, the Schottky diode is reverse biased and effectively out of the circuit. The characteristics of the Schottky-clamped transistor—or simply the Schottky transistor—are those of the normal npn bipolar device.

When the transistor is driven into saturation, the B–C junction becomes forward biased; hence, the Schottky diode also becomes forward biased. We may recall from our discussion in Chapter 9 that the effective turn-on voltage of the Schottky diode is approximately half that of the pn junction. The difference in turn-on voltage means that most of the excess base current is shunted through the Schottky diode and away

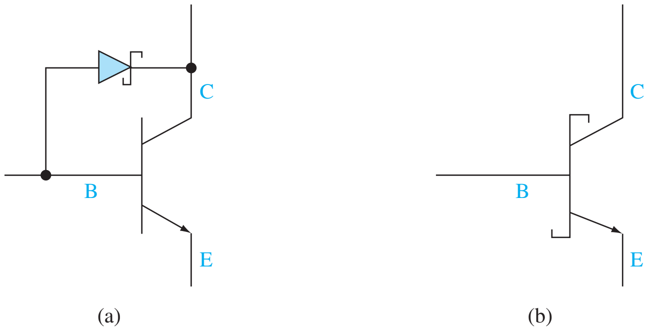

Figure 12.45 | (a) The Schottky-clamped transistor. (b) Circuit symbol of the Schottky-clamped transistor.

from the base so that the amount of excess stored charge in the base and collector is drastically reduced. The excess minority carrier concentration in the base and collector at the B–C junction is an exponential function of V_{BC}. If V_{BC} is reduced from 0.5 to 0.3 V, for example, the excess minority carrier concentration is reduced by over three orders of magnitude. The reduced excess stored charge in the base of the Schottky transistor greatly reduces the storage time—storage times on the order of 1 ns or less are common in Schottky transistors.

*12.8 | OTHER BIPOLAR TRANSISTOR STRUCTURES

This section is intended to briefly introduce three specialized bipolar transistor structures. The first structure is the polysilicon emitter bipolar junction transistor (BJT), the second is the SiGe-base transistor, and the third is the heterojunction bipolar transistor (HBT). The polysilicon emitter BJT is being used in some recent integrated circuits, and the SiGe-base transistor and HBT are intended for high-frequency/high-speed applications.

12.8.1 Polysilicon Emitter BJT

The emitter injection efficiency is degraded by the carriers injected from the base back into the emitter. The emitter width, in general, is thin, which increases speed and reduces parasitic resistance. However, a thin emitter increases the gradient in the minority carrier concentration, as indicated in Figure 12.19. The increase in the gradient increases the B–E junction current, which in turn decreases the emitter injection efficiency and decreases the common-emitter current gain. This effect is also shown in the summary of Table 12.3.

Figure 12.46 shows the idealized cross section of an npn bipolar transistor with a polysilicon emitter. As shown in the figure, there is a very thin n^+ single-crystal silicon region between the p-type base and the n-type polysilicon. As a first approximation to the analysis, we may treat the polysilicon portion of the emitter as low-mobility silicon, which means that the corresponding diffusion coefficient is small.

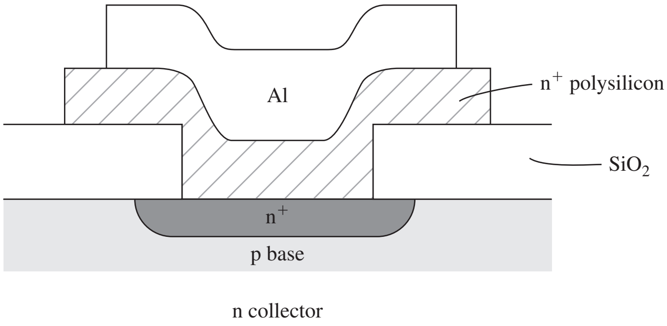

Figure 12.46 | Simplified cross section of an npn polysilicon emitter BJT.

If the neutral widths of both the polysilicon and single-crystal portions of the emitter are much smaller than the respective diffusion lengths, then the minority carrier distribution functions will be linear in each region. Both the minority carrier concentration and diffusion current must be continuous across the polysilicon/silicon interface. We can therefore write

$$eD_{E(\text{poly})}\frac{d(\delta p_{E(\text{poly})})}{dx} = eD_{E(\text{n}^+)}\frac{d(\delta p_{E(\text{n}^+)})}{dx} \tag{12.106a}$$

or

$$\frac{d(\delta p_{E(\text{n}^+)})}{dx} = \frac{D_{E(\text{poly})}}{D_{E(\text{n}^+)}}\cdot\frac{d(\delta p_{E(\text{poly})})}{dx} \tag{12.106b}$$

Since $D_{E(\text{poly})} < D_{E(\text{n}^+)}$, then the gradient of the minority carrier concentration at the emitter edge of the B–E depletion region in the n^+ region is reduced as Figure 12.47 shows. This implies that the current back-injected from the base into the emitter is reduced so that the common-emitter current gain is increased.

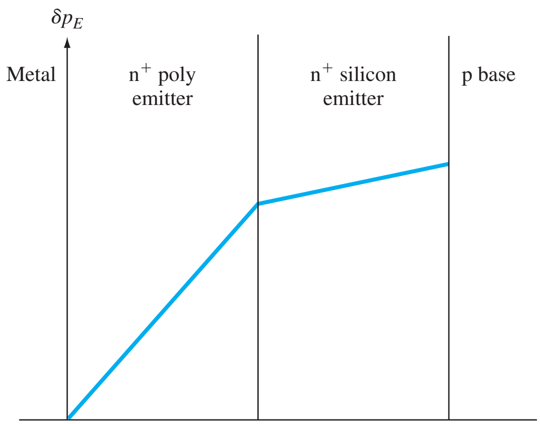

Figure 12.47 | Excess minority carrier hole concentrations in n^+ polysilicon and n^+ silicon emitter.

12.8.2 Silicon–Germanium Base Transistor

The bandgap energy of germanium (Ge) (~0.67 eV) is significantly smaller than the bandgap energy of silicon (Si) (~1.12 eV). By incorporating Ge into Si, the bandgap energy will decrease compared to pure Si. If Ge is incorporated into the base region of a Si bipolar transistor, the decrease in bandgap energy will influence the device characteristics. The desired Ge concentration profile is to have the largest amount of Ge near the base–collector junction and the least amount of Ge near the base–emitter junction. Figure 12.48a shows an ideal uniform boron doping concentration in the p-type base and a linear Ge concentration profile.

The energy bands of a SiGe-base npn transistor compared to a Si-base npn transistor, assuming the boron and Ge concentrations given in Figure 12.48a, are shown in Figure 12.48b. The emitter–base junctions of the two transistors are essentially identical, since the Ge concentration is very small in this region. However, the bandgap energy of the SiGe-base transistor near the base–collector junction is smaller than that of the Si-base transistor. The base current is determined by the base–emitter junction parameters and hence will be essentially the same in the two transistors. This change in bandgap energy will influence the collector current.

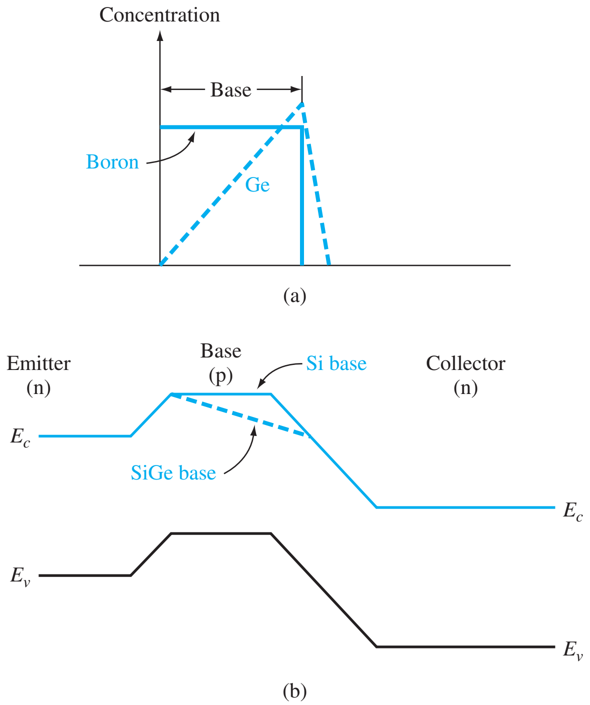

Figure 12.48 | (a) Assumed boron and germanium concentrations in the base of the SiGe-base transistor. (b) Energy-band diagram of the Si- and SiGe-base transistors.

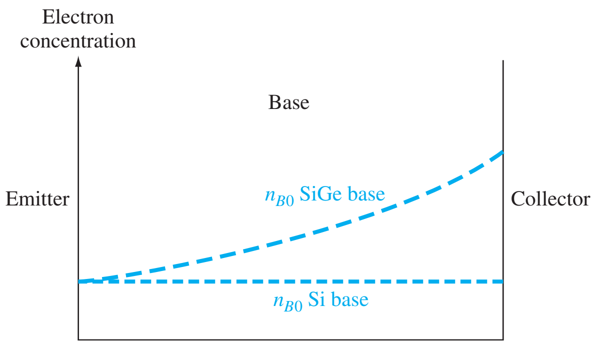

Figure 12.49 | Thermal-equilibrium minority carrier electron concentration through the base of the Si- and SiGe-base transistors.

Collector Current and Current Gain Effects Figure 12.49 shows the thermal-equilibrium minority carrier electron concentration through the base region of the SiGe and Si transistors. This concentration is given by

$$n_{B0} = \frac{n_i^2}{N_B} \tag{12.107}$$

where N_B is assumed to be constant. The intrinsic concentration, however, is a function of the bandgap energy. We may write

$$\frac{n_i^2(\text{SiGe})}{n_i^2(\text{Si})} = \exp\left(\frac{\Delta E_g}{kT}\right) \tag{12.108}$$

where n_i (SiGe) is the intrinsic carrier concentration in the SiGe material, n_i(Si) is the intrinsic carrier concentration in the Si material, and ΔE_g is the change in the bandgap energy of the SiGe material compared to that of Si.

The collector current in a SiGe-base transistor will increase. As a first approximation, we can see this from the previous analysis. The collector current is found from Equation (12.36a), in which the derivative is evaluated at the base–collector junction. This means that the value of n_{B0} in the collector current expression in Equation (12.37) is the value at the base–collector junction. Since this value is larger for the SiGe-base transistor (Figure 12.49), the collector current will be larger compared to the Si-base transistor. Since the base currents are the same in the two transistors, the increase in collector current then implies that the current gain in the SiGe-base transistor is larger. If the bandgap narrowing is 100 meV, then the increase in the collector current and current gain will be approximately a factor of 4.

Early Voltage Effects The Early voltage in a SiGe-base transistor is larger than that of the Si-base transistor. The explanation for this effect is less obvious than the explanation for the increase in collector current and current gain. For a bandgap narrowing of 100 meV, the Early voltage is increased by approximately a factor of 12. Incorporating Ge into the base region can increase the Early voltage by a large factor.

Base Transit Time and Emitter–Base Charging Time Effects The decrease in bandgap energy from the base–emitter junction to the base–collector junction induces an electric field in the base that helps accelerate electrons across the p-type base region. For a bandgap narrowing of 100 meV, the induced electric field can be on the order of 10^3 to 10^4 V/cm. This electric field reduces the base transit time by approximately a factor of 2.5.

The emitter–base junction charging time constant, given by Equation (12.87), is directly proportional to the emitter diffusion resistance r'_e. This parameter is inversely proportional to the emitter current, as seen in Equation (12.88). For a given base current, the emitter current in the SiGe-base transistor is larger, since the current gain is larger. The emitter–base junction charging time is then smaller in a SiGe-base transistor than that in a Si-base transistor.

The reduction in both the base transit time and the emitter–base charging time increases the cutoff frequency of the SiGe-base transistor. The cutoff frequency of these devices can be substantially higher than that of the Si-base device.

12.8.3 Heterojunction Bipolar Transistors

As mentioned previously, one of the basic limitations of the current gain in the bipolar transistor is the emitter injection efficiency. The emitter injection efficiency γ can be increased by reducing the value of the thermal-equilibrium minority carrier concentration p_{E0} in the emitter. However, as the emitter doping increases, the bandgap narrowing effect offsets any improvement in the emitter injection efficiency. One possible solution is to use a wide-bandgap material for the emitter, which will minimize the injection of carriers from the base back into the emitter.

Figure 12.50a shows a discrete aluminum gallium arsenide (AlGaAs)/gallium arsenide (GaAs) heterojunction bipolar transistor, and Figure 12.50b shows the energy-band diagram of the n-AlGaAs emitter to p-GaAs base junction. The large potential barrier V_h limits the number of holes that will be injected back from the base into the emitter.

The intrinsic carrier concentration is a function of bandgap energy as

$$n_i^2 \propto \exp\left(\frac{-E_g}{kT}\right)$$

For a given emitter doping, the number of minority carrier holes injected into the emitter is reduced by a factor of

$$\exp\left(\frac{\Delta E_g}{kT}\right)$$

in changing from a narrow- to wide-bandgap emitter. If $\Delta E_g = 0.30$ eV, for example, n_i^2 would be reduced by approximately 10^5 at $T = 300$ K. The drastic reduction in n_i^2 for the wide-bandgap emitter means that the requirements of a very high emitter doping can be relaxed and a high emitter injection efficiency can still be obtained. A lower emitter doping reduces the bandgap narrowing effect.

The heterojunction GaAs bipolar transistor has the potential of being a very high-frequency device. A lower emitter doping in the wide-bandgap emitter leads to

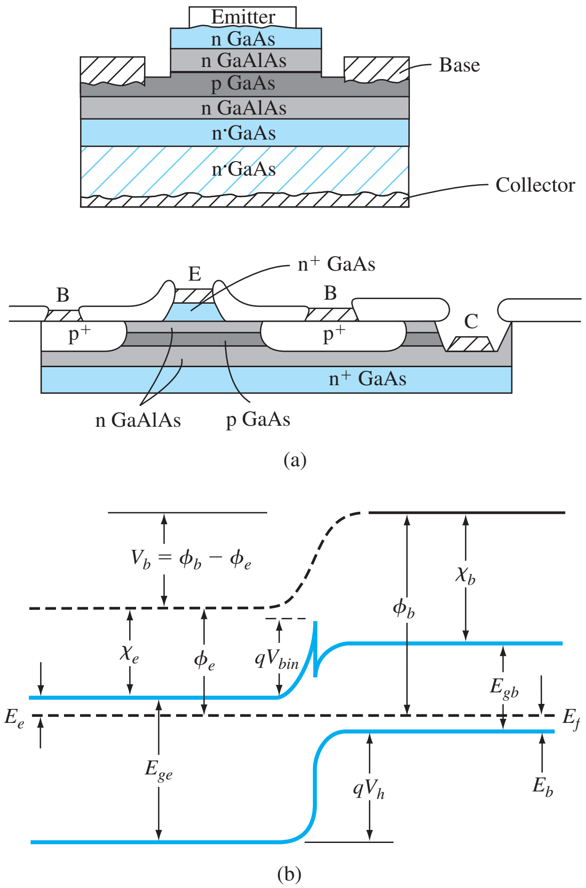

Figure 12.50 | (a) Cross section of AlGaAs/GaAs heterojunction bipolar transistor showing a discrete and integrated structure. (b) Energy-band diagram of the n-AlGaAs emitter and p-GaAs base junction.
(From Tiwari et al. [20].)

a smaller junction capacitance, increasing the speed of the device. Also, for the GaAs npn device, the minority carriers in the base are electrons with a high mobility. The electron mobility in GaAs is approximately five times that in silicon; thus, the base transit time in the GaAs base is very short. Experimental AlGaAs/GaAs heterojunction transistors with base widths on the order of 0.1 μm have shown cutoff frequencies on the order of 40 GHz.

One disadvantage of GaAs is the low minority carrier lifetime. The small lifetime is not a factor in the base of a narrow-base device, but results in a larger B–E recombination current, which decreases the recombination factor and reduces the current gain. A current gain of 150 has been reported.

12.9 | SUMMARY

- There are two complementary bipolar transistors—npn and pnp. Each transistor has three separately doped regions and two pn junctions. The center region (base) is very narrow, so the two pn junctions are said to be interacting junctions.
- In the forward-active mode, the B–E junction is forward biased and the B–C junction is reverse biased. Majority carriers from the emitter are injected into the base where they become minority carriers. These minority carriers diffuse across the base into the B–C space charge region where they are swept into the collector.
- When a transistor is biased in the forward-active mode of operation, the current at one terminal of the transistor (collector current) is controlled by the voltage applied across the other two terminals of the transistor (base–emitter voltage). This is the basic transistor action.
- The minority carrier concentrations are determined in each region of the transistor. The principal currents in the device are determined by the diffusion of these minority carriers.
- The common-base current gain, which leads to the common-emitter current gain, is a function of three factors—emitter injection efficiency, base transport factor, and recombination factor. The emitter injection efficiency takes into account carriers from the base that are injected back into the emitter, the base transport factor takes into account recombination in the base region, and the recombination factor takes into account carriers that recombine within the forward-biased B–E junction.
- Several nonideal effects are considered:

  1. Base width modulation, or Early effect—the change in the neutral base width with a change in B–C voltage, producing a change in collector current with a change in B–C or C–E voltage.
  2. High-injection effects that cause the collector current to increase at a slower rate with base–emitter voltage.
  3. Emitter bandgap narrowing that produces a smaller emitter injection efficiency because of a very large emitter region doping concentration.
  4. Current crowding effects that produce a larger current density at the emitter edge than in the center of the emitter.
  5. A nonuniform base doping concentration that induces an electric field in the base region, which aids the flow of minority carriers across the base.
  6. Two breakdown voltage mechanisms—punch-through and avalanche.

- Three equivalent circuits or mathematical models of the transistor are considered. The Ebers–Moll model and equivalent circuit are applicable in any of the transistor operating modes. The Gummel–Poon model is convenient to use when nonuniform doping exists in the transistor. The small-signal hybrid-pi model applies to transistors operating in the forward-active mode in linear amplifier circuits.
- The cutoff frequency of a transistor, a figure of merit for the transistor, is the frequency at which the magnitude of the common-emitter current gain becomes equal to unity. The frequency response is a function of the emitter–base junction capacitance charging time, the base transit time, the collector depletion region transit time, and the collector capacitance charging time.
- The switching characteristics are closely related to the frequency limitations although switching involves large changes in currents and voltages. An important parameter in switching is the charge storage time, which applies to a transistor switching from saturation to cutoff.

GLOSSARY OF IMPORTANT TERMS

alpha cutoff frequency The frequency at which the magnitude of the common-base current is $1/\sqrt{2}$ of its low-frequency value; also equal to the cutoff frequency.

bandgap narrowing The reduction in the forbidden energy bandgap with high emitter doping concentration.

base transit time The time that it takes a minority carrier to cross the neutral base region.

base transport factor The factor in the common-base current gain that accounts for recombination in the neutral base width.

base width modulation The change in the neutral base width with C–E or C–B voltage.

beta cutoff frequency The frequency at which the magnitude of the common-emitter current gain is $1/\sqrt{2}$ of its low-frequency value.

collector capacitance charging time The time constant that describes the time required for the B–C and collector–substrate space charge widths to change with a change in emitter current.

collector depletion region transit time The time that it takes a carrier to be swept across the B–C space charge region.

common-base current gain The ratio of collector current to emitter current.

common-emitter current gain The ratio of collector current to base current.

current crowding The nonuniform current density across the emitter junction area created by a lateral voltage drop in the base region due to a finite base current and base resistance.

cutoff The bias condition in which zero- or reverse-biased voltages are applied to both transistor junctions, resulting in zero transistor currents.

cutoff frequency The frequency at which the magnitude of the common-emitter current gain is unity.

early effect Another term for base width modulation.

early voltage The value of voltage (magnitude) at the intercept on the voltage axis obtained by extrapolating the I_C versus V_{CE} curves to zero current.

emitter–base junction capacitance charging time The time constant describing the time for the B–E space charge width to change with a change in emitter current.

emitter injection efficiency factor The factor in the common-base current gain that takes into account the injection of carriers from the base into the emitter.

forward active The bias condition in which the B–E junction is forward biased and the B–C junction is reverse biased.

inverse active The bias condition in which the B–E junction is reverse biased and the B–C junction is forward biased.

output conductance The ratio of a differential change in collector current to the corresponding differential change in C–E voltage.

CHECKPOINT

After studying this chapter, the reader should have the ability to:

- Describe the basic operation of the transistor.
- Sketch the energy bands of the transistor in thermal equilibrium and when biased in the various operating modes.
- Calculate, to a good first approximation, the collector current as a function of base–emitter voltage.

- Sketch the minority carrier concentrations throughout the transistor under the various operating modes.
- Define the various diffusion and other current components in the transistor from the minority carrier distribution curves.
- Explain the physical mechanisms of the current gain limiting factors.
- Define the current-limiting factors from the current components in the transistor.
- Describe the physical mechanism of base width modulation and its effect on the current–voltage characteristics of the transistor.
- Describe the voltage breakdown mechanisms in a bipolar transistor.
- Sketch the simplified small-signal hybrid-pi equivalent circuit of the transistor biased in the forward-active mode.
- Describe qualitatively the four time-delay or time-constant components in the frequency response of the bipolar transistor.

REVIEW QUESTIONS

1. Describe the charge flow in an npn bipolar transistor biased in the forward-active mode. Is the current by drift or diffusion?
2. Define the common-emitter current gain and explain why, to a first approximation, the current gain is a constant. What is the relation between the common-emitter and common-base current gains?
3. Explain the conditions of the cutoff, saturation, and inverse-active modes.
4. Sketch the minority carrier concentrations in a pnp bipolar transistor biased in the forward-active mode.
5. Define and describe the three limiting factors in the common-base current gain. Why does the base doping concentration affect the emitter injection efficiency?
6. Describe base width modulation. Sketch an *I–V* curve that shows the base width modulation effect.
7. What is meant by high injection?
8. Explain emitter current crowding.
9. Define I_{CBO} and I_{CEO}, and explain why $I_{CEO} > I_{CBO}$.
10. Sketch a simplified hybrid-pi model for an npn bipolar transistor and explain when this equivalent circuit is used.
11. Describe the time-delay factors in the frequency limitation of the bipolar transistor.
12. What is the cutoff frequency of a bipolar transistor?
13. Describe the response of a bipolar transistor when it is switching between saturation and cutoff.

PROBLEMS

(Note: In the following problems, use the transistor geometry shown in Figure 12.13. Assume $T = 300$ K unless otherwise stated.)

Section 12.1 The Bipolar Transistor Action

12.1 For a uniformly doped $n^{++}p^{+}n$ bipolar transistor in thermal equilibrium, (*a*) sketch the energy-band diagram, (*b*) sketch the electric field through the device, and (*c*) repeat parts (*a*) and (*b*) for the transistor biased in the forward-active region.