3. Space-Charge Spectroscopy in Semiconductors

By D. V. Lang

With 24 Figures

In this chapter we will discuss the special case of thermally stimulated trap-related phenomena in semiconductors. In particular, we will focus on the effects of trapping and emission of carriers at deep energy levels which are located in the space-charge layer of a pn junction or Schottky barrier. A space-charge layer is essentially depleted of mobile carriers and hence is very much like the bulk insulators discussed in Chap. 2. In addition, however, such layers have certain unique features which both simplify the analysis and make possible a totally different class of experimental techniques. We will discuss the various electrical measurements – capacitance, current, and voltage – which have been used to study deep levels in the space-charge layers associated with semiconductor junctions.

Much of the general groundwork for our discussion has been laid in Chap. 2, especially the rate equations which describe the dynamics of carrier capture and emission at deep energy levels. Therefore, we will begin by pointing out those aspects of thermally stimulated processes which are unique to semiconductor space-charge layers. We will then discuss the transient electrical response of semiconductor junctions containing traps, including such topics as the spatial and temporal variations of sensitivity, trap concentration profiles, and the special problems associated with large trap concentrations. Our treatment of the specific experimental techniques will be divided into two parts. First, we will discuss the so-called single-shot techniques which are the simplest and were historically the first developed. Secondly, we will examine in some detail the more recently developed synchronous-detection methods, which are generally referred to as Deep-Level Transient Spectroscopy (DLTS). The discussion of the implementation and interpretation of the various experimental techniques will be followed by a section on carrier capture measurements. This will include a discussion of the corrections to the measured activation energy which are necessary if the capture cross section or energy level are temperature dependent. Finally, we will conclude the chapter with a brief discussion of the instrumentation useful in the practical application of space-charge spectroscopy.

3.1 Unique Aspects of Thermally Stimulated Processes in Semiconductor Junction Space-Charge Layers

3.1.1 Ideal Space-Charge-Layer Properties

The existence of a space-charge layer at a pn junction or Schottky barrier is a general characteristic of semiconductors. Such a layer is necessary to create the electrostatic potential variation needed to counteract the diffusion potential of the carriers across the junction and hence equalize the Fermi level through the material, i.e., maintain thermal equilibrium. The *ideal* space-charge layer (i.e., without traps) for an asymmetric p^+n step junction is shown in Fig. 3.1. For simplicity we will confine most of our examples in this chapter to the case of p^+n junctions or Schottky barriers on n-type material (for which Fig. 3.1 is also applicable). The extension to n^+p junctions is trivial; the situation for arbitrary junction profiles is usually straightforward but often tedious and beyond the scope of this chapter.

We may think of a semiconductor space-charge layer as somewhat like a variable-width insulator, typical thicknesses are in the range of about 0.1 to 10 μm. Therefore, with reverse bias voltages of 1–100 V, the maximum electric field F_m in the layer is very large, typically of order 10^4–10^5 V cm^{-1}. From a

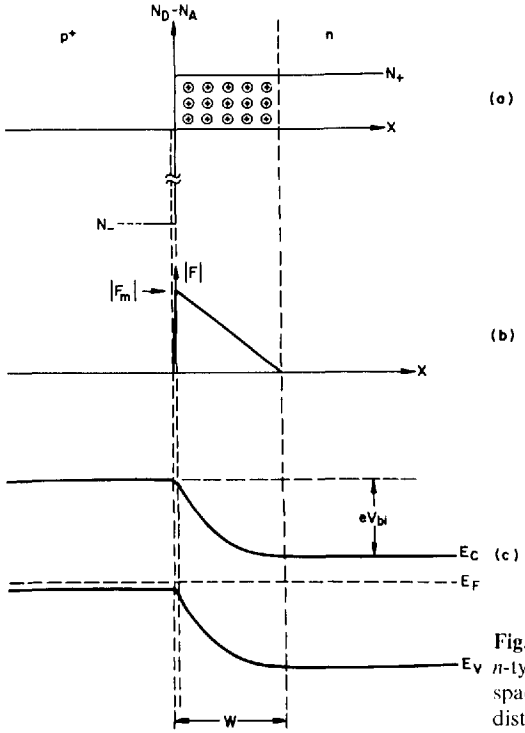

Fig. 3.1a–c. Ideal p^+n step junction (or *n*-type Schottky barrier) at zero bias (a) space-charge distribution, (b) electric field distribution, (c) band-bending diagram

straightforward solution of Poisson's equation for the asymmetric junction in Fig. 3.1 it can be shown [3.1] that

$$|F_m| = \frac{eN_+ W}{\varepsilon} \tag{3.1}$$

where

$$W = \sqrt{\frac{2\varepsilon(V_{bi} + V)}{eN_+}} \tag{3.2}$$

is the width of the space-charge layer, V_{bi} is the built-in potential, V is the applied reverse bias voltage, ε is the dielectric constant of the semiconductor, and e is the electronic unit charge. The capacitance is related to W in exactly the same manner as a parallel plate capacitor, namely

$$C = \frac{\varepsilon A}{W}, \tag{3.3}$$

where A is the area of the junction.

Thus we see that the two principal differences between space-charge layers and bulk insulators are: 1) the width of the layer can be readily adjusted by varying the bias voltage according to (3.2), and 2) a large, spatially varying electric field exists in the space-charge layer even at zero applied bias. The large electric field means that carriers thermally emitted from traps in the space-charge layer are swept out of the layer in a very short time, typically 10^{-10} to 10^{-12} s. Therefore, retrapping effects can be neglected and the analysis of thermal emission transients is considerably simplified as compared to bulk phenomena. The large electric field might also affect the physical properties of the localized defects giving rise to the trap levels. This is both good and bad, for it means on the one hand that it cannot necessarily be assumed that the properties observed are characteristic of the deep level in bulk material at zero field. On the other hand, any electric field effects which exist may be turned to advantage and used to probe the internal symmetry of the localized defect wave function. This latter possibility has not been fully explored but would seem to be a fruitful approach.

The voltage-variable-width feature inherent in semiconductor space-charge layers is basic to nearly all forms of space-charge spectroscopy. Its implications will be discussed in the remainder of this section.

3.1.2 Capacitance and Current Detection of Trap Levels

Trap levels in space-charge layers may be detected either by their effect on the junction *current* or *capacitance*, as illustrated in Fig. 3.2. The current-detection case is almost exactly analogous to the thermally stimulated trap-emission

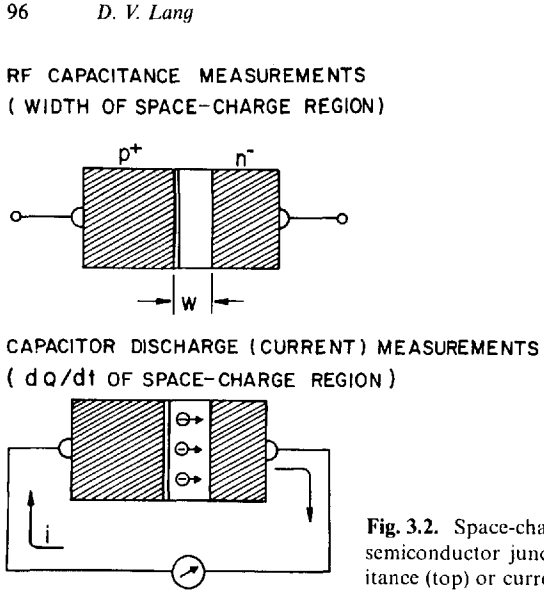

Fig. 3.2. Space-charge spectroscopy of deep levels in semiconductor junctions may be based on either capacitance (top) or current (bottom) measurements

currents which would be seen in a very thin bulk insulator in the absence of retrapping. The capacitance-detection case, which yields the same basic physical information about traps as does current detection, is unique to semiconductor space-charge layers and has no analogy in the bulk insulator case. It is clear from (3.2) that a change in the charge density in the space-charge layer will induce a corresponding change in the width, and hence in the capacitance, of the layer. Thus the space-charge-layer capacitance is a direct measure of the total charge in this layer. Consequently, if the concentration of electrons trapped at deep levels is changed, either by thermal or optical capture or emission, this variation in trapped charge can be readily monitored by observing the corresponding change in the junction capacitance at constant bias voltage. All forms of capacitance spectroscopy are based on this fact. We will discuss the quantitative relationships between the space-charge-layer capacitance and the trapped charge in Sect. 3.2, the purpose of this section is to place semiconductor capacitance measurements into proper perspective relative to the bulk measurements discussed in the other chapters of this volume.

We might view capacitance measurements as utilizing a sort of "charge transducer" effect wherein the semiconductor junction is a combined deep-level sample holder *and* transducer which directly measures the occupation state of the deep levels. Thus the focus in capacitance measurements is *opposite* to that of thermally stimulated conductivity or photoconductivity. In these latter bulk phenomena one measures the *free* carrier concentration and infers from this the properties of traps and recombination centers. In junction capacitance measurements, on the other hand, one directly measures the *trapped* carrier concentration. Thus the trap signals are more-or-less independent from each other and the analysis is considerably simplified. In the case of small trap

concentrations where the trap-induced capacitance change ΔC is much less than the overall junction capacitance C, i.e., when $\Delta C \ll C$, the thermal emission capacitance transients are simple exponential decays which are directly proportional to the rate equations for n, the electron occupation of the trap (see Chap. 2). This is the case of the ideal transducer, where the presence of deep levels does not measurably affect the Δn vs $-\Delta C$ relationship. For *large* trap concentrations, however, the situation is much more complex and the capacitance and current transients are nonexponential (see Sect. 3.2.5). In this chapter, unless otherwise specified, we will be considering the $\Delta C \ll C$ case.

The capacitance measurements which we will be discussing are in the high-frequency limit. In this case the measurement frequency is high enough (typically 1 MHz) so that the oscillating voltage applied to the junction by the capacitance meter or bridge is unable to induce changes in the occupancy of the deep levels. In Sect. 3.4.3 we will discuss the opposite case, where via capture and reemission of carriers the deep-level occupation *can* follow the measurement frequency.

3.1.3 Bias Voltage Pulses

The voltage-variable space-charge-layer width also makes possible another unique feature of space-charge spectroscopy, namely, the so-called majority- and minority-carrier pulses. These voltage pulses superimposed on the steady-state reverse bias make it possible to almost totally decouple the measurements of capture and emission processes at deep levels. This is shown most readily in Fig. 3.3 for a majority-carrier pulse in a p^+n junction. Under steady-state conditions the traps in the upper half of the gap (majority-carrier traps in n-type material) are thermally empty. If the bias voltage is momentarily reduced, as in the middle of Fig. 3.3, part of the region which was formerly within the space-charge layer is now in neutral material so that the traps are below the Fermi level. During the time that the bias is at this lower value the deep levels may capture majority carriers and tend to become filled. Immediately after the pulse, the deep levels are again within the space-charge layer where the capture rate is essentially zero. The capacitance will have changed due to the captured carriers, however, and as these carriers are thermally emitted a capacitance or current transient will be produced. Thus a majority-carrier pulse is essentially a means whereby the majority-carrier concentration can be turned *on* and *off* in a small volume of the sample. Majority-carrier *capture* dominates when the pulse is *on*, whereas majority-carrier *emission* dominates when the pulse is *off*.

A minority-carrier pulse (or injection pulse) corresponds to the diode being forward biased during the pulse so that both majority *and* minority carriers are introduced. This makes possible minority-carrier capture, but one should bear in mind that the steady-state occupation of the level corresponds to a balance between majority- and minority-carrier capture. We will discuss these points in more detail in Section 3.5. An injection pulse could also be generated by using

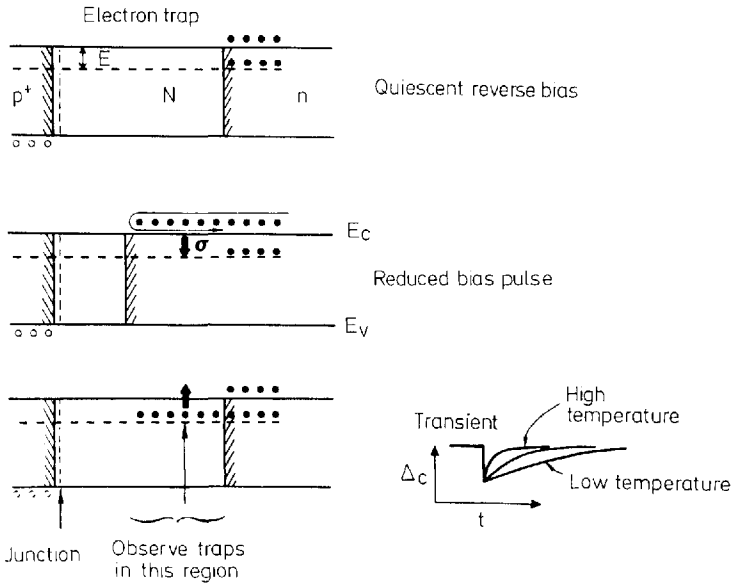

Fig. 3.3. Majority-carrier pulse for p$^+$n junction with band bending omitted for clarity. Trap occupation and space-charge-layer width are indicated corresponding to the conditions before, during, and after the pulse (from top to bottom of the figure, respectively)

an optical pulse or a pulse of high-energy electrons in an electron microscope. Such techniques are useful in Schottky barriers where forward bias injection is not possible.

Finally, let us point out an additional feature of bias voltage variations, namely, the possibility of obtaining a spatial profile of the trap concentration. This can be done by either varying the steady-state bias voltage or by varying the amplitude of the majority-carrier pulse. In either case the spatial region in which the deep levels are observed is varied and the spatial profile $N(x)$ may be obtained as discussed in Sect. 3.2.4.

3.2 Transient Response of a Semiconductor Junction with Traps

3.2.1 Spatial Variations of Sensitivity

The magnitude of a capacitance or current transient corresponding to thermal emission of carriers from a trap depends on the location of the trap within the space-charge layer. Let us calculate the relative capacitance change $[\Delta C/C]_x$ for $n(x)$ trapped electrons in the interval Δx at x, where $0 < x < W$ corresponds to positions within the n-side of the space-charge layer of a p$^+$n junction or n-type Schottky barrier such as shown in Fig. 3.1. From Poisson's equation it is

straightforward to show that the voltage change induced by trapping $n(x)$ electrons at x is

$$\Delta V = \frac{e}{\varepsilon}[N_+ W \Delta W - n(x)\, x \Delta x] \tag{3.4}$$

where N_+ is the positive space-charge concentration at W [note that $N_+(x)$ need not be uniform here, as it was in the example in Fig. 3.1]. Since we are interested in capacitance changes at constant bias, we set $\Delta V = 0$ in (3.4) and have

$$\left[\frac{\Delta C}{C}\right]_x = -\frac{n(x)}{N_+ W^2} x \Delta x . \tag{3.5}$$

This result is the capacitance charge-transducer sensitivity calibration for the $\Delta C \ll C$ case. Note that the sensitivity of the junction capacitance to trapped charge varies linearly from *zero* at the junction to a maximum value at the outer edge of the space-charge layer. This means that for inhomogeneous trap distributions capacitance measurements are very insensitive if the traps are located near the junction (near the *surface* for a Schottky barrier). The relationship in (3.5) also means that *motion* of defects within the space-charge layer will create a capacitance change if the motion produces a net change in the first moment of the defect distribution.

It is also of interest to consider the spatial variations of sensitivity for current transients. We will show that for the case of majority-carrier emission the sensitivity is exactly the opposite of (3.5), i.e., maximum at the junction and zero at the outer edge. Let ΔQ be the charge due to the emission of $\Delta n(x)$ electrons from traps in the interval Δx at x, thus

$$\Delta Q = e \Delta n(x) \Delta x . \tag{3.6}$$

For the case of electron emission in n-type material, some fraction of this charge ΔQ_+ never leaves the space-charge layer since it is needed at the outer edge to produce the accompanying capacitance change in (3.5), i.e.,

$$\Delta Q_+ = e A N_+ \Delta W. \tag{3.7}$$

The net charge which is detected in the external circuit ΔQ_{ext} is the difference between ΔQ and ΔQ_+. With the aid of (3.5), this may be written

$$\Delta Q_{\text{ext}} = e A \Delta n(x) \Delta x \left(1 - \frac{x}{W}\right). \tag{3.8}$$

For *hole* emission, on the other hand, the net charge reaching the external circuit is ΔQ_+. Thus for a thermal emission transient the measured current

density due to deep levels in the interval Δx at x is

$$J(x)=e\Delta x\left\{\alpha n(x)\left(1-\frac{x}{W}\right)+\alpha^*[N-n(x)]\frac{x}{W}\right\}, \tag{3.9}$$

where α and α^* are the thermal emission rates for electrons and holes, respectively. Note that in steady state $\alpha n=\alpha^*(1-n)$ so that the spatially varying terms in (3.9) cancel each other.

The spatial variation of sensitivity for current transients in a p^+n junction in (3.9) is the opposite of the capacitance case for *electron* emission but has the same spatial shape for *hole* emission. The spatially varying terms in (3.9) are actually the displacement current associated with the capacitance change in (3.5). It is therefore not surprising that the sensitivity variations are linear in both cases.

3.2.2 Temporal Variations of Sensitivity

Both capacitance and current transients in the $\Delta C \ll C$ limit are proportional to the trapped electron concentration $n(t)$ for electron emission, or trapped hole concentration $[1-n(t)]$ for hole emission. The time dependence of n is given by

$$\frac{n(t)}{N}=\begin{cases}1 & t\leqq 0\\ \dfrac{\alpha^*}{\alpha+\alpha^*}\left\{1+\dfrac{\alpha}{\alpha^*}\exp[-(\alpha+\alpha^*)t]\right\} & t>0\end{cases} \tag{3.10}$$

for the case where a majority-carrier pulse at $t=0$ completely fills the trap (for the n-type example), and by

$$\frac{n(t)}{N}=\begin{cases}0 & t\leqq 0\\ \dfrac{\alpha^*}{\alpha+\alpha^*}\{1-\exp[-(\alpha+\alpha^*)t]\} & t>0\end{cases} \tag{3.11}$$

for the case where an injection pulse at $t=0$ empties the trap (fills it with holes).

In the context of space-charge spectroscopy, a trap for which $\alpha \gg \alpha^*$ is called a "majority-carrier trap" in n-type material and a "minority-carrier trap" in p-type. The opposite is true for a trap which has $\alpha \ll \alpha^*$. This notation is useful because for capacitance measurements the *sign* of the transient is always *negative* for majority-carrier emission and always *positive* for minority emission, independent of the conductivity type (see also Sect. 3.3.1). Current transients, however, are always of the same sign.

According to (3.5) we see that for $\Delta C \ll C$ the capacitance signal is a direct measure of $n(t)$ and therefore is described by (3.10) or (3.11). The current signal, on the other hand, is proportional to $\alpha n(t)$, for the example of electron emission in n-type material in (3.9). Therefore, the capacitance sensitivity is *independent* of the time constant of the transient whereas the current sensitivity is *directly proportional* to the transient rate. This means that current measurements are most sensitive for fast transients while capacitance is most sensitive for slow transients. We have found that the crossover point between capacitance and current sensitivity for the same trap in the same sample with identical instrumental bandwidths is for time constants roughly of order 1–10 ms.

3.2.3 Transient Magnitudes for Uniformly Distributed Trap Concentrations

In order to evaluate the magnitude of the total signal in a particular sample we must integrate (3.5) or (3.9) over the region of the space-charge layer where emission is actually taking place. To do this we must discuss in more detail the band-diagram of the asymmetric junction case which we are considering. This is shown for zero bias and reverse bias in Fig. 3.4 for the case of uniformly doped *n*-type material, i.e., for a p^+n junction or a Schottky barrier on n-type material.

We must consider two basic regions of the space-charge layer. At the outer edge of the layer there is a transition between the conducting bulk material and the central part of the space-charge layer which is essentially depleted of free carriers. For a particular deep level this gives rise to the so-called *edge region* of thickness λ, defined by the point where the trap level E_T crosses the Fermi level E_F. From Poisson's equation it can readily be shown that

$$\lambda = \sqrt{\frac{2\varepsilon(E_F - E_T)}{e^2 N_+}} \tag{3.12}$$

if N_+ is uniform. Note that in this case λ is independent of bias voltage.

In the edge region the deep level is below the Fermi level and therefore filled with electrons in thermal equilibrium, i.e., there are sufficient free carriers present so that the electron capture rate c_n is much larger than the electron thermal emission rate α. There is also a small additional voltage-dependent edge region λ_V^* near the junction where a sufficient number of free holes have spilled over to make $c_p \gg \alpha^*$ so that the trap is always empty. In the remainder of the space-charge layer the trap level is above the Fermi level and its steady-state occupation is defined by the ratio of thermal emission rates according to (3.10, 11). This inner carrier depletion region is where the capacitance and current transients are generated.

Note that it may seem inconsistent for us to consider here the presence of free carriers within the edge regions of the space-charge layer, but to neglect such carriers in the calculation of W in (3.2). This approach is justified, however,

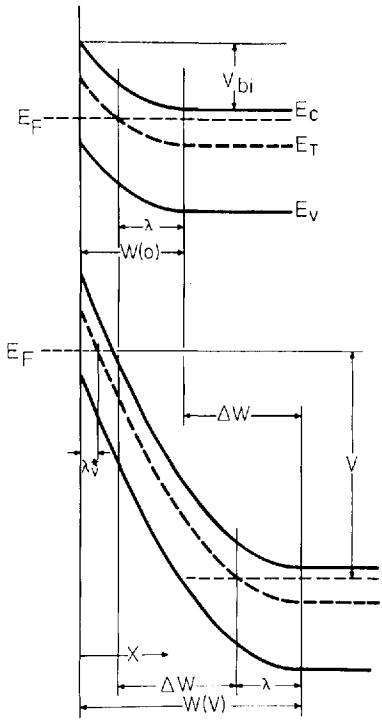

Fig. 3.4. Band-bending diagram for p⁺n step junction (or *n*-type Schottky barrier) at zero bias (top) and at a reverse bias of $V = 3.5\,V_{bi}$ (bottom)

since it has been shown [3.2, 3] that neglect of free carriers in the region $0 < x < W$ introduces negligible errors in calculating the reverse bias capacitance. Thus even though there are sufficient carriers within the edge regions of the space-charge layer to affect the deep-level occupation, there are not enough to significantly change the shape of the junction band bending.

The total capacitance signal produced by a majority-carrier pulse which brings the diode bias to zero for a time long enough to completely fill a trap of concentration N is obtained by integrating (3.5) from $W(0)-\lambda$ to $W(V)-\lambda$, thus

$$\frac{\Delta C}{C} = -\frac{N}{2N_+}\left[1 - 2\frac{\lambda}{W(V)}\left(1 - \frac{C(V)}{C(0)}\right) - \left(\frac{C(V)}{C(0)}\right)^2\right]. \tag{3.13}$$

If one neglected the edge region, the bracketed term in (3.13) would be equal to 1 and the $\Delta C/C$ relationship would reduce to the simple form often given [3.4]. For the example of Fig. 3.4, however, the bracketed term is equal to 0.46, i.e., zero biasing the junction gives slightly less than *half* of the capacitance change that would have been expected if all of the deep levels within the space-charge layer could have been observed. Thus a neglect of the edge region may

introduce substantial error in trap concentration measurements, especially at low bias voltages.

For current transients the integral relationship derived from (3.9) for uniformly distributed traps is most conveniently expressed in terms of the D *parameter* introduced by *Grimmeiss* [3.5], namely

$$J(t)=e(x_2-x_1)\left\{\frac{1}{D}\alpha n(t)+\left(1-\frac{1}{D}\right)\alpha^*[N-n(t)]\right\} \tag{3.14}$$

where

$$\frac{1}{D}=\frac{1}{x_2-x_1}\int_{x_1}^{x_2}\left(1-\frac{x}{W}\right)dx=1-\frac{x_1+x_2}{2W}. \tag{3.15}$$

The positions x_1 and x_2 are the points where the spatially varying hole and electron capture rates, respectively, equal the sum of all relevant emission rates. For thermally stimulated processes $x_1=\lambda_V^*$ and $x_2=W-\lambda$. In most cases D is near to and slightly less than 2, but in general D depends on temperature, trap depth and spatial distribution, and the optical intensity for photocurrent transients.

3.2.4 Measuring Trap Concentration Profiles

There are many ways of using the bias-dependent width of the space-charge layer to obtain both deep- and shallow-level profiles. Obviously, one cannot deduce the spatial profile of deep levels without first knowing the shallow-level profile. Many schemes for obtaining $N_+(x)$ from the $C-V$ behavior of the junction have been discussed [3.6–9].

One of the most convenient ways of measuring the spatial profiles of deep levels, which is especially useful for low concentrations, is to record the magnitudes of a series of capacitance transients produced with majority-carrier pulses of increasing amplitude. This method has the effect of probing the deep-level concentration inward from W as the majority-carrier pulse amplitude is increased from zero. If we denote by $\delta(\Delta C/C)$ the incremental change in the relative capacitance signal due to the traps filled by the small change δV in the majority-carrier pulse of amplitude V, then a convenient relationship is given [3.10] by

$$\delta\left(\frac{\Delta C}{C}\right)=\left(\frac{\varepsilon}{eW^2N_+}\right)\frac{N(x)}{N_+(x)}\delta V \tag{3.16}$$

where x is the width of the space-charge layer during the pulse and C, W, and N_+ are the values corresponding to the steady-state bias between pulses. This relationship is valid for asymmetric junctions or Schottky barriers if $\Delta C \ll C$.

According to (3.16), a linear ΔC vs V plot implies that the trap profile $N(x)$ has the same shape as the shallow-level profile $N_+(x)$. We will show some examples of the use of this relationship in determining deep-level profiles in Sect. 3.4.5.

3.2.5 Capacitance Relationships for Large Trap Concentrations

Capacitance changes in the concentrated-deep-level limit, i.e., $N \gtrsim N_+/3$, are very complex to analyze. A major difficulty is that the capacitance and current transients recorded at constant bias are nonexponential. Since this limit gives rise to the largest signals, it was historically the first to be discovered. Consequently, many treatments of the various possible capacitance relationships have been given in the literature [3.11–21]. Perhaps the most concise and useful is that which relates the concentration N of a uniformly distributed deep-level of energy E_T to the capacitance C_0 immediately following a zero-biasing majority-carrier pulse and the steady-state capacitance C_∞ [3.17, 19, 20], namely,

$$\frac{N}{N_+} = \frac{\left(\frac{C_\infty}{C_0}\right)^2 - 1}{\left[1 - \frac{C_\infty}{C_0}\sqrt{\frac{E_F - E_T}{e(V_{bi} + V)}}\right]^2}. \tag{3.17}$$

It is possible to overcome the problem of nonexponential transients in this regime by utilizing a clever feedback scheme proposed by *Goto* et al. [3.21]. These authors showed that the bias voltage necessary to maintain the diode capacitance at a constant value in a feedback loop is a direct measure of the trapped charge concentration. Thus a measurement of this feedback voltage will produce an exponential transient even for very large trap concentrations. This effect is dramatically demonstrated by Fig. 3.5 which shows the constant-bias and constant-capacitance transients associated with the emission of electrons from the so-called DX center in $Al_xGa_{1-x}As$; in this case $N \cong 8N_+$ [3.22].

3.3 Single-Shot Techniques

3.3.1 Isothermal Transients

The essential physics of all forms of thermally stimulated space-charge spectroscopy – indeed, all of our previous discussion – is based on the thermal-emission transient at some fixed temperature. The more involved techniques which we will introduce later have the benefits of speed and convenience, useful for the study of complex trap spectra, but they cannot produce raw data of any

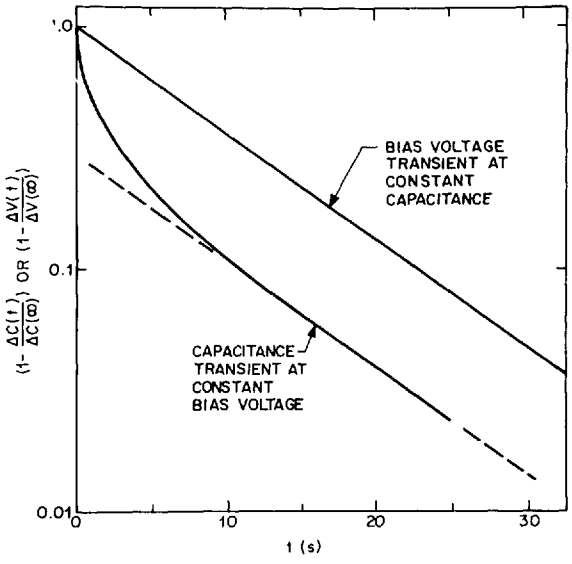

Fig. 3.5. Effect of a large deep-level concentration on the constant-bias and constant-capacitance carrier emission transient

higher quality than can be done with a large number of isothermal transients recorded at many different temperatures.

Isothermal capacitance transients were first proposed as a technique to study traps in semiconductors by *Williams* [3.23], and by *Furukawa* and *Ishibashi* [3.12]. The quantitative expression for a capacitance transient follows from (3.10–12). A capacitance transient due to majority-carrier emission is always *negative*. Figure 3.6 shows such a transient with insets describing the condition of the space-charge layer during the various phases of the transient [3.24]. Similarly, Fig. 3.7 shows a *positive* transient due to minority-carrier emission following an injection pulse [3.24].

Current transients are described by (3.10, 11, 14, 15). Note that the *sign* of current transients is always the same (independent of the nature of the emission process). The current flows in the opposite sense to that in forward bias, thus the voltage drop across the load tends to make the diode bias more positive.

The activation energy $E \equiv E_c - E_T$ for electron emission or $E_A^* \equiv E_T - E_V$ for hole emission can be obtained in the standard manner from the slope of an Arrhenius plot, i.e., a plot of the log of the transient decay rate $\alpha + \alpha^*$ as a function of $1/T$, where T is the absolute temperature at which the transient is recorded. This follows from the relationships

$$\alpha = \nu \exp(-E/kT) \tag{3.18}$$

and

$$\alpha^* = \nu^* \exp(-E_A^*/kT) \tag{3.19}$$

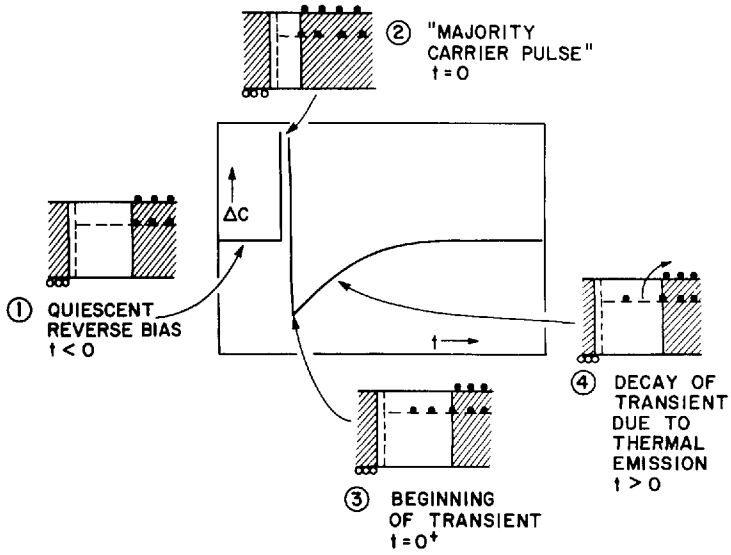

Fig. 3.6. Isothermal capacitance transient for thermal emission from a majority-carrier trap. The insets 1–4 show the conditions of the trap occupation, space-charge-layer width (unshaded), and free carrier concentrations during the various phases of the transient in a p^+n junction

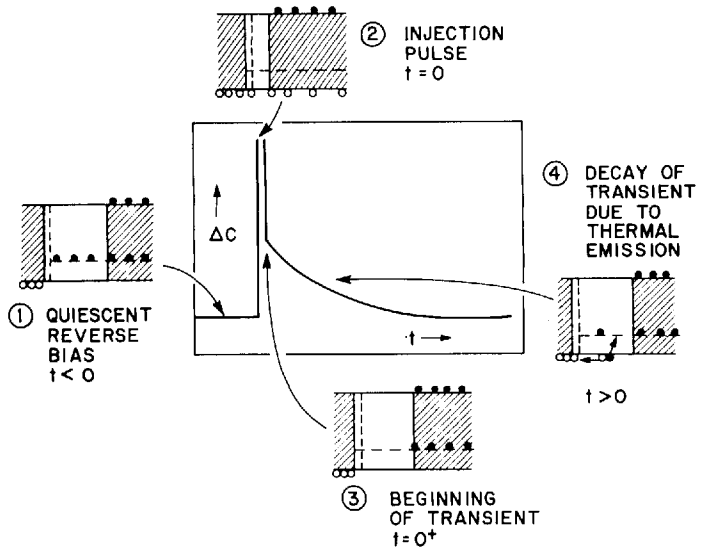

Fig. 3.7. Isothermal capacitance transient for thermal emission from a minority-carrier trap. The insets 1–4 show the conditions of the trap occupation, space-charge-layer width (unshaded), and free carrier concentrations during the various phases of the transient in a p^+n junction

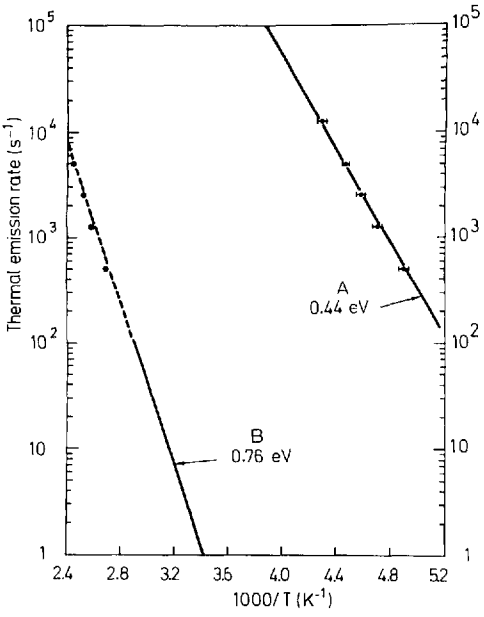

Fig. 3.8. Thermal emission rate α^* versus $1000/T$ for holes trapped on two commonly occurring deep levels of unknown origin in LPE n-GaAs. The data points are DLTS peak positions from Fig. 3.14, the solid lines are isothermal transient data

and the fact that either α or α^* usually dominates the transient rate. An example of such a plot for the hole-emission transients of two common hole traps in GaAs is shown in Fig. 3.8 [3.25]. The activation energy measured in this way is a valid parameter of the trap, but it may not be the true depth of the trap. We will discuss the problems of interpreting activation energies in Sect. 3.5.4 after we have discussed capture processes.

3.3.2 Irreversible Thermal Scans

This class of techniques bears a strong resemblance to many of the thermally stimulated phenomena observed in bulk material. In general, one first prepares the traps in an appropriate initial charge state at some low temperature where thermal emission is negligible. This may be done, for example, by cooling the sample with zero bias (so that all traps are filled with majority carriers) and then applying a reverse bias at low temperature. It is also possible to load the traps with minority carriers by optical excitation or by cooling the sample with forward bias (for the case of a pn junction). Following this initial preparation, the sample is heated at some constant rate q and one observes either current peaks or capacitance steps when the various levels emit their trapped carriers. This is schematically illustrated in Fig. 3.9.

The thermally stimulated current (TSC) method for semiconductor junctions was first proposed by *Driver* and *Wright* [3.26], while the thermally stimulated capacitance (TSCAP) method was first discussed by *Carballes* and

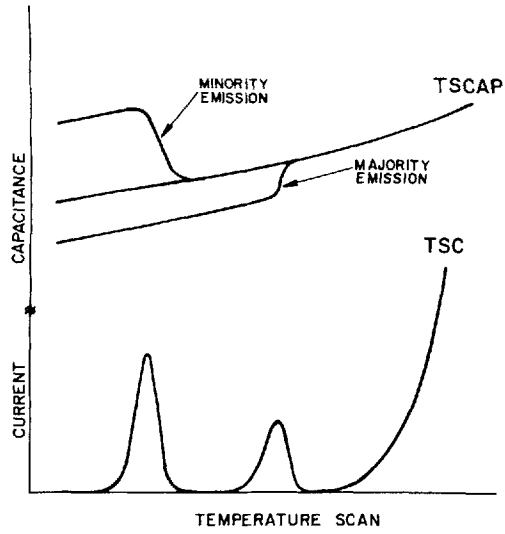

Fig. 3.9. Schematic representation of typical TSCAP and TSC data for a sample containing a majority-carrier trap of concentration N and a shallower minority-carrier trap of concentration $2N$. The increase in current at higher temperatures in the TSC case is the dark current due to steady-state generation and leakage

Lebailly [3.27]. A number of different single-shot thermal techniques have been employed with great success by *Sah* and co-workers to study deep levels in silicon [3.14, 28]. *Buehler* and *Phillips* [3.29] have recently analyzed TSC and TSCAP experiments and have shown that the temperature T_m of the TSC maximum or midpoint of the TSCAP step is related to the activation energy by

$$E = kT_m \ln\left[\frac{\nu k T_m^2}{q(E+2kT_m)}\right] \tag{3.20}$$

for the example of electron emission. Thus if $q = 1\ \mathrm{K\,s^{-1}}$ and $\nu = 10^{12}\ \mathrm{s^{-1}}$ we have $E = 28.9\,kT_m$ for $T_m = 100$ K and $E = 29.5\,kT_m$ for $T_m = 200$ K.

The thermally scanned techniques are less accurate for determining activation energies than is the measurement of the isothermal transient at several temperatures. The scanning techniques do, however, have an important advantage for the case of samples with several different deep levels. Namely, they make it possible to quickly survey nearly the entire range of trap energies and produce a spectrum of the traps in a particular sample. The TSC and TSCAP methods work reasonably well as survey techniques when the trap concentrations are fairly large, e.g., $N \gtrsim 0.1|N_D - N_A|$, and the trap energies are fairly deep, e.g., $E \gtrsim 0.3$ eV. In general, one would expect TSCAP measurements to have a better signal-to-noise ratio than TSC because these scanning techniques are in the low-emission-rate limit of temporal sensitivity variations (see Sect. 3.2.2). Because of this temporal variation, the magnitudes of TSC peaks depend on the heating rate while TSCAP steps do not. In addition, as shown by the current increase at higher temperatures in Fig. 3.9, the TSC technique is subject to considerable problems due to junction leakage or

steady-state generation currents as the temperature is increased; TSCAP, on the other hand, is less sensitive to diode quality. Another advantage of TSCAP over TSC is a characteristic of capacitance vs current measurements in general, namely, the ability to distinguish between majority- and minority-carrier emission by the *sign* of the capacitance change (see Fig. 3.9). However, TSCAP measurements are difficult if the change in the steady-state capacitance with temperature is much greater than the capacitance steps due to thermal emission of carriers; the TSCAP steps are then superimposed on a steeply sloping baseline and are very difficult to resolve. There may be situations, however, in which TSC is preferable over TSCAP, e.g., majority-carrier traps very near the junction (see Sect. 3.2.1) or traps located in the *i* layer of a p-i-n junction (which induce very small changes in C).

Perhaps the major virtue of the single-shot scanning techniques is their relative simplicity. The synchronous detection DLTS techniques to be discussed in the next section are better survey techniques, but they are somewhat more complex to set up. However, in most cases the added complexity is well worth the effort.

3.4 Deep-Level Transient Spectroscopy (DLTS)

3.4.1 The Rate-Window Concept

The basic idea of the DLTS technique is the rate-window concept. This is illustrated in Fig. 3.10. If we consider a train of repetitive bias pulses applied to the sample, we then have a signal which consists of a series of transients with a constant repetition rate. As the temperature is varied, the time constant of the transients varies exponentially with $1/T$, as discussed in Sect. 3.3.1 and shown in Fig. 3.10. Let us assume that there exists a "black box" with the characteristic that it gives no response to this sequence of transients unless the time constant of the transient decay is near some preselected value – the inverse of the "rate window". The output of this black box as the temperature of the sample is scanned is then a series of peaks as shown in Fig. 3.10. Such a plot is called a DLTS *spectrum.*

A DLTS spectrum looks superficially like the signals produced by the single-shot scanning techniques in the last section. However, there are several important differences. First, the DLTS scan is *reversible* and does not depend on the magnitude of the heating rate. Thus one can scan *up* or *down* in temperature at any rate and can even *stop* the scan on a peak to study the capture properties or spatial profile of that particular trap. This latter possibility is a tremendous advantage relative to the irreversible thermal scan methods. Another advantage is that the baseline is always flat, i.e., there is good common-mode rejection of the steady-state capacitance variations. Therefore it is possible to obtain extremely high sensitivity, e.g., $\Delta C/C \sim 10^{-5}$ for $C \sim 100$ pF. This sensitivity may be readily increased by a slower scan with a

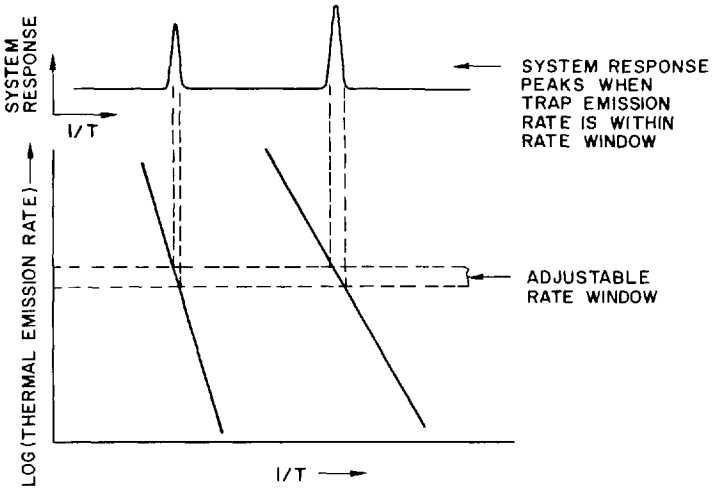

Fig. 3.10. The rate-window concept, which is the basic idea of the DLTS method. This illustrates the peak in the system output response which occurs whenever a component of the decay rate of the input transient signal corresponds to the rate selected by the window

longer final time constant, since the position and shape of the DLTS peaks are independent of the scan rate.

The DLTS method applies equally well to either capacitance or current transients – indeed, to *any* thermally stimulated process which can be repetitively generated by a pulsed excitation. However, there are some differences. First, since the DLTS rate window is usually in the ms to μs regime, current transients are likely to be more sensitive than capacitance transients (see Sect. 3.2.2). This fact is essential to the success of high-resolution DLTS measurements made in a scanning electron microscope, where there are extreme signal-to-noise limitations (see Sect. 3.4.6). However, the price paid for this added sensitivity is loss of the possibility to differentiate between majority- or minority-carrier emission. A second feature of *current* DLTS is the fact that the peak is shifted by somewhat less than half of a linewidth to higher temperature relative to a *capacitance* DLTS peak for the same rate window, i.e., the rate window must be defined slightly differently for current transients. This shift is due to the fact that the magnitude of the current transient is proportional to the thermal emission rate and thus increases very rapidly with temperature, effectively skewing the line-shape toward higher temperature.

3.4.2 Rate-Window Instrumentation

There are a number of ways in which the DLTS rate-window concept may be achieved in practice. The first method proposed in the original DLTS experiments involved the use of a dual-gated integrator (double boxcar) [3.25].

In this method the transient amplitude is sampled at two times t_1 and t_2, after the pulse, as shown in Fig. 3.11. The DLTS signal is the *difference* between the transient amplitude at these two times. Such a difference signal is a standard output feature of a double boxcar. As we can see in Fig. 3.11, there is zero difference between the signal at these two gates for either very slow or very fast transients, corresponding to low or high temperature, respectively. However, when the transient-time constant τ is on the order of the gate separation, a difference signal is generated and the boxcar output passes through a maximum as a function of temperature. This is a DLTS peak. For capacitance measurements the rate window can be expressed in terms of the transient-time constant giving rise to the maximum double-boxcar output [3.25], namely,

$$\tau_{\max} = \frac{t_1 - t_2}{\ln(t_1/t_2)}. \tag{3.21}$$

Equation (3.21) is a very good approximation for the case of relatively wide boxcar gates as well, provided t_1 and t_2 are taken as the *midpoint* of each sampling gate. In fact, the best way to use a boxcar is with relatively wide gates. We have found experimentally that the signal-to-noise ratio is proportional to the square root of the gate width, all else being unchanged. Thus the gate configuration shown in Fig. 3.12a might be a typical operating condition.

An alternate means of implementing the rate window is a lock-in amplifier [3.4, 30, 31, 56]. This is equivalent to having the sinusoidally weighted gates shown in Fig. 3.12b. In order to achieve this condition, however, it is necessary to properly adjust the phase of the lock-in relative to the transient signal. An operational definition of the correct phase relationship is to adjust the lock-in so as to be in quadrature with the bias pulse with the pulse generator triggered by the lock-in reference signal. This corresponds to obtaining a null of the lock-in output for the pulse overload recovery transient under conditions where the trap emission transient is zero.

However, a possible problem with the lock-in DLTS method is that some small contribution from the very large pulse recovery transient cannot be totally avoided. This results in excess noise and in possible baseline shifts during the temperature scan. Note that the duration of the pulse overload recovery depends on the output bandwidth of the capacitance meter and may be considerably longer than the actual pulse. Such a problem does not exist for the boxcar or correlator method, however, since the signal is not sampled during the pulse-overload portion of the waveform. The problem may be reduced, but not totally eliminated, in the lock-in case by utilizing a transmission gate before the input of the lock-in to block the signal during the overload transient [3.31]. Such a gate will cut out the jitter of the overload transient but will not stop the DLTS baseline from shifting as the steady-state capacitance of the sample changes with temperature. To solve this latter problem one needs an active dc baseline restorer such as used in the exponential correlator developed by *Miller* et al. [3.30].

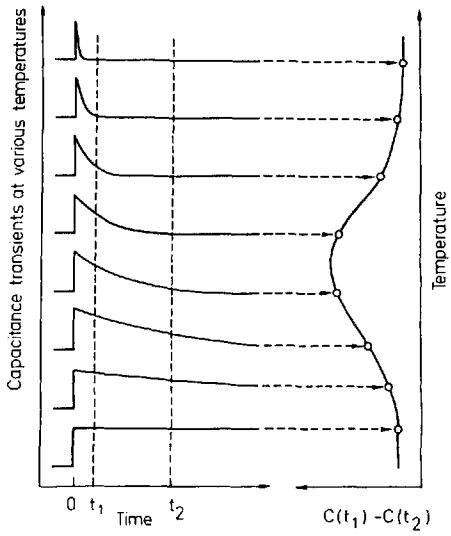

Fig. 3.11. Implementation of a DLTS rate window by means of a double-boxcar integrator with gates set at times t_1 and t_2

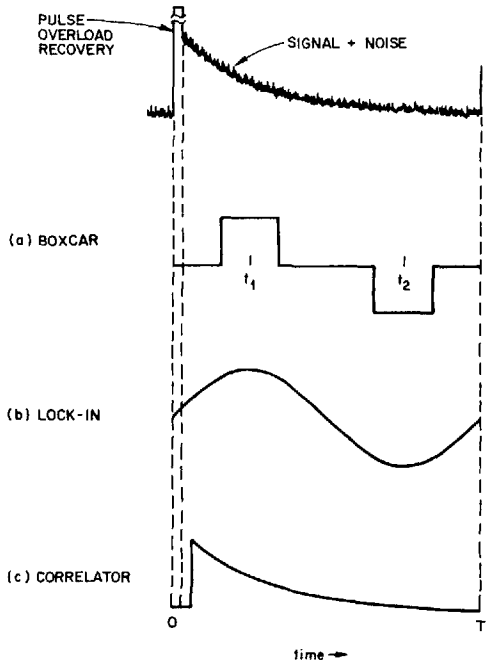

Fig. 3.12a–c. Diagram comparing a typical capacitance transient input signal with three different rate-window operations

The DLTS line shape for the lock-in method has the form [3.4, 30]

$$A(y)=\frac{2\pi(1-e^{+y})}{4\pi^2+y^2}, \tag{3.22}$$

where $y=T/\tau$ with T being the reciprocal of the lock-in frequency and τ the transient time constant. The maximum of this line-shape function occurs for $\tau_{max}=0.424\,T$. This is the lock-in DLTS rate-window relationship for capacitance transients. Note that the maximum of (3.22) differs from the rate window defined by (3.21) for the boxcar gate settings in Fig. 3.12a ($t_1=T/4$, $t_2=3T/4$) by only 7%. This corresponds to a temperature error of $\lesssim 1$ K, which is negligible in most cases. The proper rate window for current transients under these conditions is 2.2 times faster than the capacitance rate window. This is a shift of less than half of a linewidth in (3.22).

In an effort to obtain the highest possible sensitivity in a DLTS measurement, *Miller* et al. [3.30] have analyzed the various techniques using time-domain filter theory. In this analysis the various rate-window techniques are viewed as bipolar weighting functions with which the trap emission signal is multiplied and then integrated to obtain the DLTS signal. Three possible weighting functions are shown in Fig. 3.12. The results show that the exponential weighting function in Fig. 3.12c gives rise to the theoretically *maximum* signal-to-noise ratio for a white noise spectrum. The maximum of the DLTS peak occurs when the time constant of the exponential weighting function is equal to that of the signal.

An instrument which accomplishes this exponential weighting operation is called a *correlator* and has been described by *Miller* et al. [3.4, 30]. An essential feature of this instrument is the active dc baseline restoration circuit. This baseline restorer is essentially a second gate which measures the signal near the end of the transient. Thus the correlator actually has a bipolar weighting function, as it must to achieve the differential measurement mode necessary to produce a DLTS spectrum.

According to the analysis of *Miller* et al. [3.30] the exponential correlator achieves a signal-to-noise ratio which is about two times higher than the lock-in amplifier. The signal-to-noise ratio of the double boxcar is roughly comparable to the lock-in for the gates shown in Fig. 3.12a, but decreases as the square root of the gate width for narrower gates. Thus for very narrow gates the boxcar is quite a bit noisier than the other two techniques.

A *single*-gate-boxcar DLTS scheme has been proposed by *Wessels* [3.32] for the case of current transients. The previously discussed rate-window techniques are, of course, also applicable to current transients, if one takes into account that the peaks are slightly shifted to higher temperature because of the rapid increase in the current transient magnitude with temperature. This rapid increase also means, however, that a *single* gate will generate a DLTS peak. For capacitance transients a single gate will produce a *step* for dc coupling [3.33] and a *peak* for ac coupling between the capacitance meter and boxcar. The

single-gate schemes are not terribly satisfactory, however. The ac-coupled method has the same problem with fluctuations and drifts due to the pulse-overload-recovery transient as does the lock-in method, while the single-gate *current* transient is sensitive to the steady-state leakage current of the sample, which may become significant at high temperatures and cause a baseline shift.

As we mentioned earlier, there is no superior way to gather raw data than to measure the entire transient at many different temperatures. Such a procedure would be prohibitively time consuming if done manually in a sample with many traps. Thus the various survey techniques, such as DLTS, have a distinct advantage in speed at the expense of not recording all of the possible data, such as the actual transient time dependences. However, the use of automated digital data acquisition by a computer makes the analog rate-window techniques unnecessary. With a computer it is possible to rapidly record the actual transients at various temperatures during a thermal scan and to process this data in a variety of different ways. For example, an Arrhenius plot such as in Fig. 3.8 could be produced simultaneously for all of the traps in the sample with just a single scan of temperature. The *same* stored data could also be analyzed using the digital equivalent of one of the weighting functions in Fig. 3.12 to produce a DLTS spectrum.

It is also possible to use a computer to simulate a double boxcar directly, by digitizing the signal sampled during the two gates. This was used by *Lefevre* and *Schulz* [3.34] to obtain the DLTS spectrum corresponding to traps located

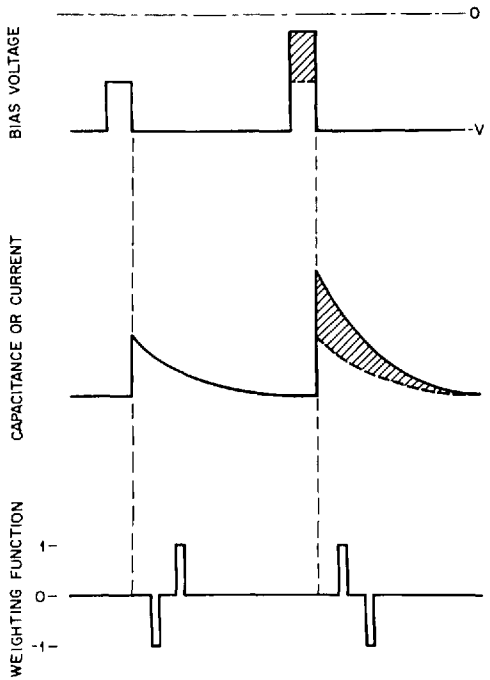

Fig. 3.13. Diagram showing how alternating bias pulses of two different amplitudes produce a DLTS signal corresponding to an internal region of the space-charge layer when analyzed with the alternating weighting function shown. This is called double DLTS

in a narrow region in the *interior* of the space-charge layer. The method, termed *double* DLTS (DDLTS), consists of applying alternating bias pulses of two different amplitudes, as shown in Fig. 3.13. If the digitized gate information is processed as $C(t_1)-C(t_2)$ for one of the pulse amplitudes and as $C(t_2)-C(t_1)$ for the other pulse, then the resultant DLTS signal corresponds to the difference between the transients (shown shaded). The spatial location of the traps giving rise to this "difference transient" corresponds to the range of space-charge-layer widths produced by the shaded region of bias voltages which is the difference between the two bias pulses. The DDLTS method has some advantages if the thermal emission rate of the trapped carriers is electric-field dependent so that the DLTS peak position depends on the location of the trap within the space-charge layer. This method also discriminates against low-frequency noise, i.e., noise of a much lower frequency than the pulse repetition rate. Apparently, some Schottky barrier samples are subject to such noise [3.34].

The shaded difference transient in Fig. 3.13 can also be obtained manually by subtracting one isothermal transient from the other. This was first proposed by *Sah* and *Fu* [3.35] and has been used to study the electric-field dependence of thermal emission rates for deep levels in Si [3.36–38] and GaP [3.10].

3.4.3 Admittance Spectroscopy

If one measures the capacitance of a pn junction or Schottky barrier as a function of temperature at rather low frequencies, e.g., 100 Hz, one finds a trap-related increase in the capacitance at a temperature where the emission and capture rates at the trap are as fast as the measurement frequency [3.4, 39]. In all of our previous discussions we had assumed that the measurement frequency was sufficiently high so that this never occurred. The resulting series of capacitance steps in the low-frequency case look very much like a TSCAP curve (see Fig. 3.9), except that the frequency-response effect does not depend on the direction or rate of the thermal scan. For any such step in the capacitance response corresponding to a frequency-dependent element being able to follow the driving voltage as the temperature is increased, there is necessarily also a *peak* in the ac conductance response of the sample. Measurement of the conductance peaks during a thermal scan was proposed by *Losee* [3.39], and is called admittance spectroscopy.

We have included this technique in the section on DLTS measurements because the conductance peak can actually be viewed as a type of rate-window method for current transients. In this case, however, the bias voltage is not pulsed but is sinusoidally varying. The sinusoidal rate-window weighting function (Fig. 3.12b) is provided by the phase-sensitive detector of the conductance bridge or meter. Indeed, the ac conductance is often measured using a lock-in amplifier with the circuit and phase setting almost identical to that which would be used for the lock-in DLTS method in the case of current

transients. The single exception is the triggered pulse generator which converts the lock-in reference signal into a repetitive series of bias pulses in the DLTS case. For the conductance case the lock-in reference signal is the sinusoidal voltage drive which is applied directly to the sample.

When admittance spectroscopy is viewed in this manner it is easy to see that it has an important limitation relative to other DLTS measurements. The timing of the carrier-capture process cannot be varied independently from the carrier-emission process – both are fixed by sinusoidal waveforms. The more conventional DLTS measurements, on the other hand, separate the carrier-capture and rate-window functions by the use of bias pulses. Thus it is possible to measure the carrier-capture cross sections of the trap, as we will discuss in Sect. 3.5. It is also possible with normal DLTS to study minority-carrier capture and emission; this is not possible with admittance spectroscopy. There are situations, however, for which admittance spectroscopy is better suited than DLTS, for example, in p-i-n junctions where bias pulses do not penetrate into the i layer.

3.4.4 Determination of Activation Energies

The best way to determine the activation energy of a deep level is to construct an Arrhenius plot such as in Fig. 3.8. This can be done simultaneously for nearly all of the traps in a sample by recording several DLTS spectra using different rate windows. An example of such data is shown in Fig. 3.14 for the common, but as yet unidentified, hole traps denoted *A* and *B* which are nearly always present in GaAs grown by liquid-phase epitaxy (LPE) [3.25, 40]. An Arrhenius plot is constructed from these data by plotting the log of the rate window from (3.21) vs the inverse temperature of the DLTS peak T_m^{-1}. The data points in Fig. 3.8 correspond to the rate-window vs peak-temperature data in Fig. 3.14. The relationship of the measured activation energy to the true depth of the deep level is complex, however, and will be discussed in Sect. 3.5.4.

It is also possible to obtain somewhat more limited emission rate data, spanning about one decade, by using the appropriate analytical line-shape function, such as (3.22) for lock-in DLTS. If the FWHM of the lock-in DLTS peak is $T_2 - T_1$, where T_1 and T_2 are the temperatures corresponding to the half-height on the low- and high-temperature side of the peak, respectively, then it can be shown from (3.22) that the activation energy is [3.4]

$$E = \frac{k \ln(15)}{1/T_1 - 1/T_2} - 2kT_m . \tag{3.23}$$

This is a reasonably useful approximate method if the peak is well resolved and the transient is exponential. However, since one does not know a priori from only DLTS data whether or not a particular peak is due to an exponential transient, one must use some care in applying (3.23). Nonexponential transients

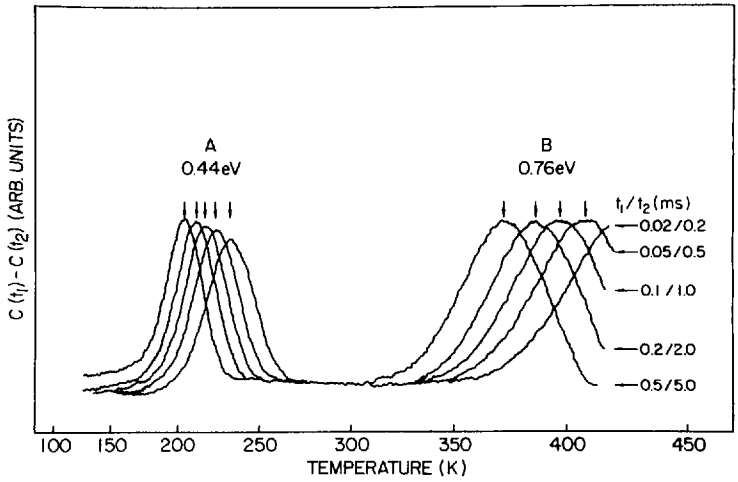

Fig. 3.14. DLTS hole-emission spectra of two hole traps in LPE n-GaAs. The values of t_1 and t_2 corresponding to the various scans are indicated. The arrows indicate the values of T_m used in Fig. 8

typically give rise to DLTS peaks which are broader than the ideal line shape given by (3.22, 23). Even though the ideal line-shape functions are not valid in such cases, the Arrhenius plots constructed from DLTS rate-window vs T_m data for nonexponential transients are still well defined, since the DLTS peak is always composed of a linear superposition of exponential-transient peaks in the $\Delta C \ll C$ case.

A very rough measure of the activation energy of a trap may be estimated directly from the temperature T_m of the maximum of the DLTS peak. This is analogous to the determination of energies from TSC peak positions according to (3.20) and is just about as accurate. If we assume that $\nu = 10^{12}\,s^{-1}$, then with a rate window of $\tau_{max}^{-1} = 50\,s^{-1}$ we find $E \approx 23.7\,kT_m$. Since E depends only logarithmically on ν, this temperature–energy conversion factor is accurate to approximately $\pm 10\%$.

Perhaps one of the most fruitful ways to employ DLTS measurements is to use the spectroscopic nature of the entire DLTS signal vs temperature scan and compare the trap spectra of various samples by always using the same standardized rate window. We have found that $50\,s^{-1}$ is a convenient choice for this routine-survey mode. As shown in Fig. 3.15 for the example of n-GaAs [3.40] it is possible to distinguish characteristic peaks due to various deep levels even for rather complicated spectra. A unique feature of DLTS relative to *optical* spectroscopy is that the various DLTS peak heights are *directly proportional* to the respective deep-level concentrations. Thus *quantitative* analysis of electrically active defects is readily possible with a sensitivity of at least $10^{12}\,cm^{-3}$ in samples with a shallow-level doping of $10^{16}\,cm^{-3}$. With more lightly doped samples the sensitivity is even more impressive, e.g., $< 10^{10}\,cm^{-3}$ traps if $|N_D - N_A| = 10^{14}\,cm^{-3}$.

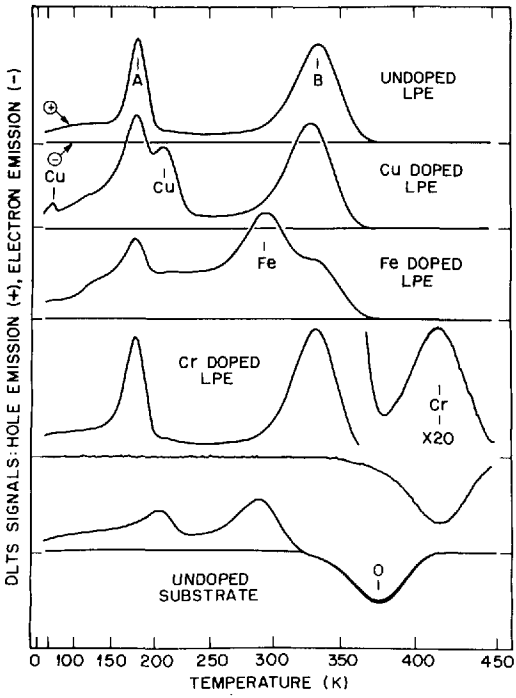

Fig. 3.15. DLTS spectra of five GaAs p^+n junctions with the deep-level doping indicated. The rate window was $50\,s^{-1}$

3.4.5 Trap Concentration Profiles

In Sect. 3.2.4 we discussed a basic relationship useful in determining the spatial profile of trap concentrations, see (3.16). When this is coupled with the DLTS method, a very powerful profiling technique emerges which is capable of simultaneously obtaining the profiles of many different traps of very low concentration in the same sample. Figure 3.16 shows a series of four DLTS spectra of four common hole traps in p-GaAs recorded for different values of reverse bias so that the observed traps are located at various distances x from the n^+p junction [3.41]. By using majority-carrier pulses with an amplitude small compared to the steady-state bias voltage, it was possible to obtain good spatial resolution in these spectra. A series of such spectra taken with various steady-state voltages and pulse amplitudes were analyzed using (3.16) with the shallow-level profile known from *Miller* feedback profiler [3.9] measurements. The trap profiles obtained in this way are shown in Fig. 3.17. Note the good spatial resolution ($<0.1\ \mu m$) and sensitivity (low 10^{13} range). This example was a p-type LPE GaAs layer grown on an n^+-substrate. The data clearly show minute levels of Cu and Fe contamination (0.01–0.03 ppm) which have apparently diffused into the epitaxial layer from the substrate; levels *A* and *B* have apparently been gettered by the n^+-layer.

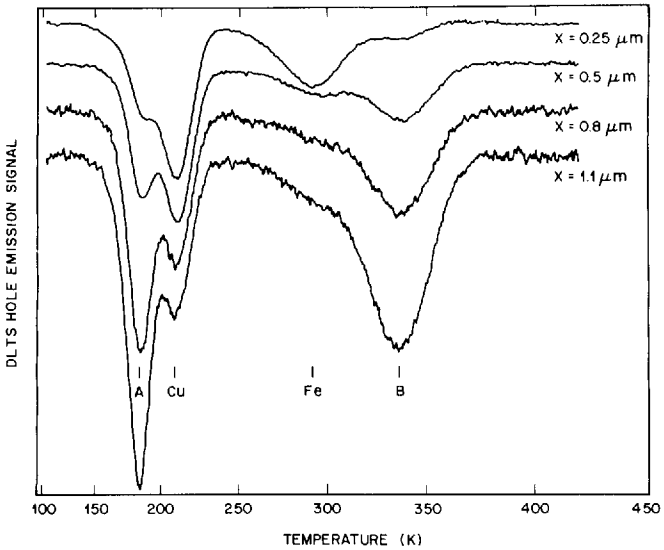

Fig. 3.16. DLTS spectra of an n^+p GaAs junction with spatially varying concentrations of four deep levels. The bias and majority-carrier pulse amplitude were chosen to correspond to the values of x shown. The gain was chosen for each scan to make the concentration scale the same for all four spectra. The rate window was $50\,s^{-1}$

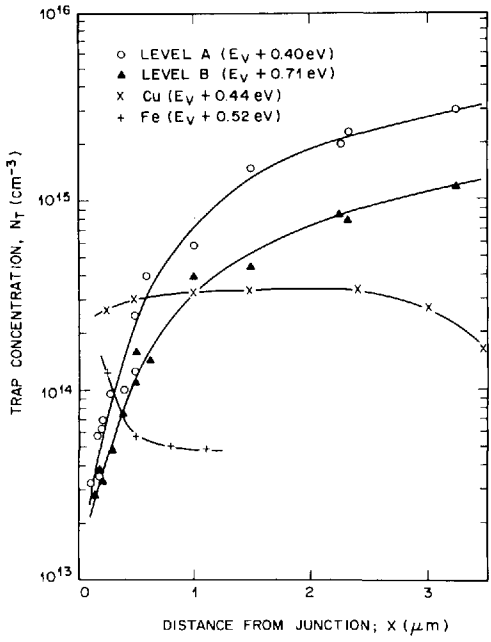

Fig. 3.17. Concentration profiles of the four deep levels shown in Fig. 3.16

3.4.6 Scanning DLTS

The spatial information gained by the profiling method just discussed is only in one dimension, namely, that perpendicular to the junction plane. It is possible to obtain spatial information *within* the junction plane by combining the DLTS technique with the $x-y$ scanning capabilities of a scanning electron microscope (SEM) [3.42]. The idea of such a measurement is quite simple in principle but rather difficult to successfully achieve in practice. In the scanning DLTS (SDLTS) case, the traps are filled by the e–h pairs which are generated by the high-energy electron beam of the microscope. If this beam is pulsed on and off in the same manner as the bias pulses used in normal DLTS measurements, the resulting emission transients can be detected when the beam is off, making it possible to produce a DLTS spectrum. If the temperature of the sample is adjusted to correspond to a DLTS peak, an image of the spatial variation of the trap responsible for that peak can be formed by using the DLTS signal magnitude to modulate the intensity of a CRT display which is rastered in synchronism with the pulsed electron beam.

The major problem with SDLTS measurements is the extremely small thermal-emission signal generated within the small beam-excited region. For example, if the beam were to fill defects in a 5 μm diameter area in a sample 0.5 mm in diameter, then the SDLTS signal would be only 10^{-4} of the normal DLTS signal for a particular deep level in this sample. The SDLTS sensitivity is inversely proportional to the *square* of the spatial resolution. One way to overcome this extreme signal-to-noise problem is to use current transients in the μs region. For reasons discussed in Sect. 3.2.2, such transients are considerably more sensitive than capacitance transients. With this method it has been possible to achieve sensitivities of $\lesssim 10^{15}\,\mathrm{cm}^{-3}$ with fairly low resolution (~ 20 μm) and resolutions as high as ~ 2 μm with lower sensitivity [3.42]. The usefulness of this variation of the DLTS method has only begun to be explored, but it is clear that such spatial resolution capabilities will greatly aid in the study of certain classes of inhomogeneously distributed deep levels in semiconductors.

3.5 Carrier Capture Measurements

3.5.1 Majority-Carrier Pulse

It is clear from Fig. 3.2 that if the duration of the majority-carrier pulse is less than the time necessary to fill the deep levels, the resulting emission transient will be less than its maximum possible value. Thus, for example, DLTS spectra recorded with bias pulses of various lengths may have differing peak heights as shown in Fig. 3.18. It is often more convenient to stop the DLTS scan on a particular peak and record the variation of peak height as a function of pulse width. Such data are shown for the Cr level in LPE n-GaAs in Fig. 3.19, as an

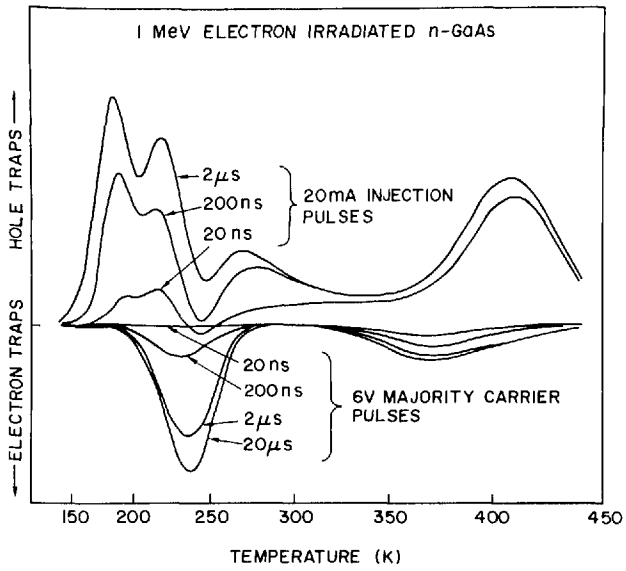

Fig. 3.18. DLTS spectra of deep levels produced by 1 MeV electron irradiation of a p^+n GaAs junction. The effect of incomplete filling of the levels for short pulses is shown. The rate window was $5.1 \times 10^3\ s^{-1}$

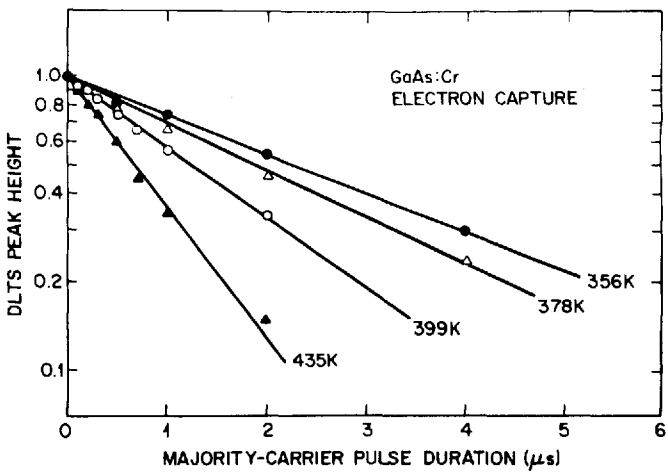

Fig. 3.19. DLTS signal versus majority-carrier pulse durtation for electron capture at Cr levels in LPE n-GaAs at four different temperatures

example. In this case, since we are interested in the majority-carrier capture cross section of a minority-carrier trap, a *double*-pulse sequence has been used. The first pulse is an injection pulse which totally fills the Cr hole trap with holes. The second pulse is a majority-carrier pulse which is applied before the trapped holes are thermally emitted. Thus the electron capture during this second pulse *reduces* the observed DLTS hole-emission signal. For a majority-carrier trap, on the other hand, the signal would *increase* with majority-carrier pulse duration.

Such carrier-capture measurements are typically done at low enough temperatures so that the majority-carrier capture rate (c_n for electrons in n-type material or c_p for holes in p-type material) is much faster than the thermal emission rates α or α^*. Thus, for example, the time dependence of the trapped electron concentration (in n-type material) during a majority-carrier pulse is given by

$$n(t) = N[1 - \exp(-c_n t)] . \tag{3.24}$$

Hole capture is given by an analogous expression. The capture rates are defined in terms of capture cross sections S_n and S_p, average thermal velocities $\langle v_n \rangle$ and $\langle v_p \rangle$, and free carrier concentration n_c and p, as

$$\begin{aligned} c_n &\equiv S_n \langle v_n \rangle n_c \\ c_p &\equiv S_p \langle v_p \rangle p . \end{aligned} \tag{3.25}$$

The experimental data, such as in Fig. 3.19, correspond to the values of $n(t)$ in (3.24) at the *end* of the majority-carrier pulse. Thus a plot of the DLTS or isothermal-transient signal (for the case of minority-carrier traps with the double-pulse sequence) or N *minus* this signal (for the case of majority-carrier traps with a single pulse) versus the pulse duration gives the majority-carrier capture rate directly. From this it is a straightforward matter to obtain the capture cross section from (3.25). Note, however, that $n_c < (N_D - N_A)$ and $p < (N_A - N_D)$ at low temperatures. Capacitance profiling methods [3.6–9] measure $|N_D - N_A|$, not the free-carrier concentration, thus Hall data on the temperature dependence of the free-carrier concentration in similar samples must be known in order to obtain the capture cross sections.

In the edge region of the space-charge layer during the pulse the carrier concentration is much smaller than in neutral material and, in addition, is spatially varying (see Sect. 3.2.3). Therefore, there will be a distribution of majority-carrier capture rates on the edge region which will be much slower than in the remainder of the volume covered by the pulse. This will tend to produce a long, nonexponential tail on the initially-exponential capture data. The presence of such a tail can be seen in Fig. 3.18, where the signal increases rapidly for very short pulses, followed by a small, slower increase for much longer pulses. This effect can be minimized and almost totally eliminated,

however, by using capacitance transients with rather large reverse bias voltages (>10 V); because then the pulse-edge region is a smaller fraction of the total space-charge-layer width and, furthermore, is located in the low-sensitivity region near $x=0$ (see Sect. 3.2.1).

3.5.2 Minority-Carrier (Injection) Pulse

The measurement of minority-carrier capture is operationally very similar to the majority-carrier case just discussed. However, there are several important differences which complicate matters. First, during the injection pulse there are *both* minority and majority carriers present. Thus, for example, the time dependence of the captured electron concentration at an initially empty electron trap in p-type material is given by

$$n(t)=\frac{c_n N}{c_n+c_p}\{1-\exp[-(c_n+c_p)t]\}\,. \tag{3.26}$$

An exactly analogous expression for trapped holes may be written for hole traps in n-type material. Note that the observed capture rate in (3.26) is the *sum* of the electron and hole capture rates and that the steady-state value of n is not necessarily equal to N, but depends on the relative values of c_n and c_p. If the minority-carrier capture rate is much larger than the majority-carrier rate, then we say that the trap is "saturable", i.e., a sufficiently long and intense injection pulse will fill the trap with minority carriers. If, on the other hand, the opposite is true and the majority-carrier capture rate is largest, the trap is "unsaturable" and the minority-carrier occupation will always be near zero. Therefore, an unsaturable minority-carrier trap will not give rise to a transient following an injection pulse and hence will be undetected. It may be possible, however, to detect such deep levels by filling the state optically.

A second difference between minority- and majority-carrier capture is that the minority-carrier concentration needed to obtain the capture cross section from (3.25) depends on the magnitude of the injection-pulse current density J. For the example of electron injection into p-type material, the injected electron concentration n_c is given by [3.43]

$$n_c=\left(\frac{uL_n}{eD_n}\right)J \tag{3.27}$$

where L_n and D_n are the electron diffusion length and diffusion constant, respectively, and where u is the electron injection efficiency of the pn junction, i.e., the fraction of the total forward current corresponding to electron injection

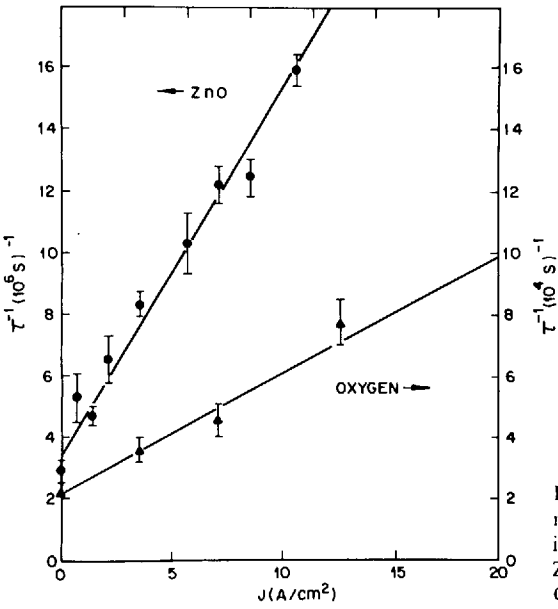

Fig. 3.20. Rate of increase of trapped minority carriers as a function of injection-pulse-current density for the ZnO and oxygen centers in p-type GaP

into the p-side, given by [3.1]

$$u = \left(1 + \frac{D_p L_n p_p}{D_n L_p n_n}\right)^{-1}, \tag{3.28}$$

where n_n and p_p are the free electron and hole concentrations on the n- and p-side of the junction, respectively. Analogous expressions follow readily for hole injection into the n-side of a pn junction.

Thus, according to (3.25–27), the time constant of the increase in the minority-carrier emission signal as a function of the injection-pulse width is a function of the injected current density J during the pulse. An example of the capture-rate variation versus injection-pulse current is shown in Fig. 3.20. This corresponds to injected electron capture at the oxygen and ZnO centers in p-type Ga*P* at 190 K [3.10]. The intercept at zero current is the majority-carrier capture rate while the component of the measured rate which is linear with injected current is the minority-carrier capture rate. This method is considerably more time consuming and somewhat less accurate than the majority-carrier capture measurements of the last section. The inaccuracy stems from the difficulty in knowing the properties of the junction well enough to calculate the minority-carrier concentration from (3.27, 28) with great precision. Therefore it is always advisable to measure capture cross sections by the majority-carrier-pulse method, if this is at all possible.

3.5.3 Optically Generated Carriers

In some cases the carrier capture time is shorter than the duration of the shortest pulse (typically ~ 10 ns). It is then possible to measure the capture cross section by observing the rate of capture of optically generated carriers in the space-charge layer [3.14, 44]. The technique consists of illuminating a semitransparent Schottky barrier with above-band-gap light and measuring the time constant of the capacitance transient due to the capture of photogenerated carriers. If the average drift velocity μF in the space-charge layer is smaller than the thermal velocity $\langle v \rangle$, then the electron density, for example, in the space-charge layer is related to the measured photocurrent J by

$$n_c = J/\mu_n eF \tag{3.29}$$

where the average value of the electric field F must be calculated from the doping profile (see Fig. 3.1) and represents the largest source of error in this measurement. In general, electrons and holes are both present in such an experiment (unless the light is totally absorbed near the surface) so that the capture transient is described by an equation like (3.26).

3.5.4 Temperature-Dependent Capture

In most considerations of thermally stimulated processes it is assumed that the capture cross sections S_n and S_p are independent of temperature. While this may be the case for some deep levels, it is not in general true. In fact, many deep levels in III–V semiconductors, for example, have capture cross sections which are *thermally activated* [3.44, 45], as shown in Fig. 3.21. On the basis of data such as in Fig. 3.21, as well as optical Stokes shifts of various deep levels, it was shown [3.44] that an important, previously neglected, nonradiative capture mechanism for many deep levels is multiphonon emission via lattice relaxation.

It is also possible to have capture cross sections which *decrease* with increasing temperature [3.46]. This is a well-known phenomenon for shallow donors and acceptors at low temperatures (<77 K) and can be explained by capture into excited Coulomb states followed by a cascade of single-phonon emissions into the ground state [3.47]. Similarly, thermal reemission from the excited states of a deep level will also give rise to an effective capture cross section which decreases with temperature [3.44, 48].

The occurrence of temperature-dependent capture cross sections is important for space-charge spectroscopy because it affects the measured activation energy for thermal emission and hence must be taken into account in order to obtain the *true* depth of the level. Thus, for example, if

$$S_n = S_\infty \exp(-E_s/kT), \tag{3.30}$$

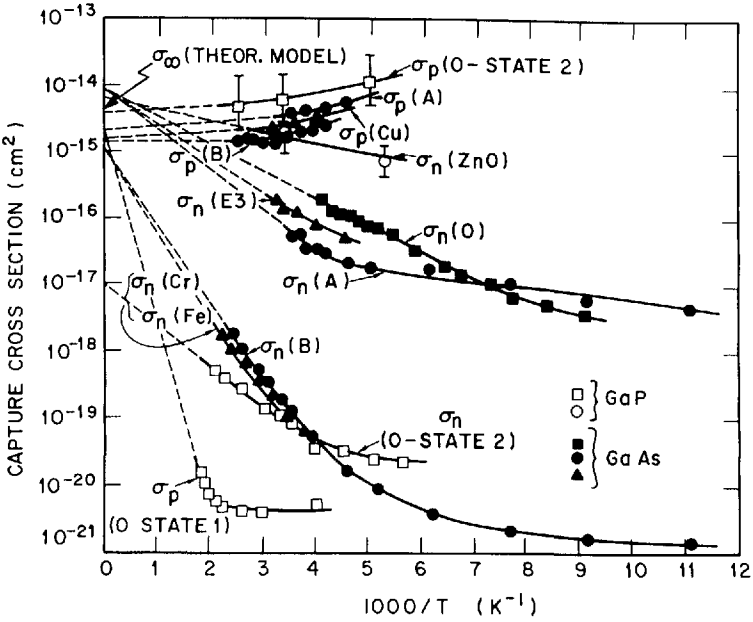

Fig. 3.21. Capture cross section (denoted here by σ) as a function of inverse temperature for seven deep levels in GaAs and two in GaP

such as at higher temperatures in Fig. 3.21, then (3.18) must be modified to become

$$\alpha = \frac{S_\infty \langle v_n \rangle N_c}{g} \exp[-(E+E_s)/kT], \tag{3.31}$$

where g is the degeneracy factor for the deep level [3.46]. Equation (3.31) is a straightforward consequence of the steady-state detailed balance between emission and capture [3.41]. The measured activation energy E_{meas} is related to the true energies of the system by

$$E_{meas} = E + E_s + 2kT, \tag{3.32}$$

where the $2kT$ term arises from the T^2 dependence of $\langle v_n \rangle N_c$. This term is not needed if the Arrhenius plot is constructed as $\log \alpha T^{-2}$ vs $1/T$.

For cases where E_s is negative, i.e., the cross section *decreases* with temperature due to reemission from an excited state of depth E_s, (3.32) is still valid. In this case the measured activation energy E_{meas} is the energy difference between the ground state and excited state [3.48]. This behavior occurs when the thermal emission rate from the excited state is faster than the transition rate from the excited state to the ground state.

In addition to the energy corrections due to the temperature dependence of capture cross sections, there are also corrections to the measured activation energy due to the temperature dependence of the deep-level energy itself. It is helpful to describe the temperature dependence of the deep-level ionization energy in thermodynamic terms. This does not necessarily provide any new physical insight but does make available the powerful machinery of thermodynamic relationships. We first note that the energy E in (3.18) [E_A^* in (3.19)] is the standard Gibbs free energy of the ionization reaction, usually denoted by G. This is the proper free energy to use in a solid, since the pressure P is constant. The slope of an Arrhenius plot is given by

$$\frac{d}{d\frac{1}{T}}\left(\frac{G}{T}\right)_P = G + \frac{1}{T}\left(\frac{dG}{d\frac{1}{T}}\right)_P. \tag{3.33}$$

By making use of the thermodynamic identity

$$\left(\frac{dG}{dT}\right)_P = -S, \tag{3.34}$$

where S is the entropy, we see that the right-hand side of (3.33) is equal to $G+TS$, which is the definition of enthalpy, H. Note that at $T=0$ the enthalpy H and the free energy G are equal. Indeed, it is straightforward to show from the identity in (3.34) that at a given temperature T the value of G obtained by extrapolating the tangent of the true $G(T)$ curve to $T=0$ is identically equal to $H(T)$. Thus H is the proper thermodynamic term for what is often referred to as the "extrapolated $T=0$ energy" of a deep level [3.5, 49, 50]. Since $H=G+TS$, we can see that the parameter which expresses the difference between H and G is the entropy S. Thus an ionization energy which is independent of temperature may also be described by saying that the ionization reaction produces no change in entropy.

By using these proper thermodynamic quantities for the example of electron emission, we may rewrite (3.18) as

$$\alpha = \nu \exp(-G/kT) = \nu \exp(S/k)\exp(-H/kT). \tag{3.35}$$

Thus, since the activation energy as determined from an Arrhenius plot is actually H [after the corrections of (3.32) have been applied], we see that an added entropy term is present in the exponential prefactor. Typical values of the entropy change for ionization reactions in semiconductors are $0 \lesssim S \lesssim 6k$ [3.51, 52]. Therefore, the exponential entropy factor in (3.35) may be as high as 400. Neglect of this factor could thus lead to large errors in determining the capture cross section from (3.31) by using the activation energy as determined from the slope of an Arrhenius plot. Indeed, such discrepancies have been noted [53, 54] and can be explained by the additional factor in (3.35) [3.50].

Care must also be taken in relating energies determined from Arrhenius plots to energies measured optically, since in general the former is H while the latter is G [3.49]. Thus in addition to the problems associated with Stokes shifts and temperature-dependent capture, optical and thermal measurements can only be compared with high accuracy when the temperature dependence of the energy level, i.e., the entropy of ionization, is known.

3.6 Instrumentation

3.6.1 Standard Capacitance Meter with Pulse Transformer

One of the most straightforward ways of making capacitance measurements is to use a standard commercially available 1 – MHz capacitance meter, such as a Boonton model 72B or equivalent. Such a meter works very well for measuring the steady-state capacitance, but has certain shortcomings for measuring capacitance transients. First, its rather slow output time constant (typically ~ 1 ms) gives rise to a rather long pulse-overload-recovery transient, which is a problem for lock-in DLTS measurements and also makes it impossible to measure transients faster than 1 ms. Second, there is no provision made for the introduction of a fast bias pulse. The normal bias input reaches the sample through a heavily filtered dc network which has such a long time constant that it is impossible to use μs pulses.

An alternate scheme, which is adaptable to any type of meter, is the pulse-transformer method proposed by *Henry* et al. [3.43]. The schematic diagram of this pulse-transformer circuit is shown in Fig. 3.22. The transformer can be externally added to the capacitance meter by connecting the sample in series with the secondary winding of the transformer. The voltage of the pulse is measured with a low-capacitance oscilloscope probe connected to the sample lead just as it enters the temperature-control dewar. The current of an injection pulse can be measured with a commerically available current probe.

In practice, the limited bandwidth of most pulse transformers [3.55] makes it impossible to cover with a single transformer the entire ns to ms range often needed for majority-carrier and injection pulses. Thus to cover this range in making carrier capture measurements one typically must use several transformers best suited to different parts of this range. A good choice for routine measurements with a 10 V reverse bias is a 10 V – 10 μs mayority-carrier pulse and a 10 A cm^{-2} – 10 μs injection pulse. A problem that is often encountered with injection pulses is that the pulsed current is driven into the internal circuitry of the capacitance meter and causes instabilities and additional noise at the output. Such effects do not seem to occur with majority-carrier pulses.

A final note for anyone contemplating the use of a capacitance meter in this application is that the signal-to-noise ratio of the meter is proportional to the amplitude of the rf drive level used in measuring the capacitance. This is usually limited to 15 mV_{rms} in order to be well within the small-signal limit for the

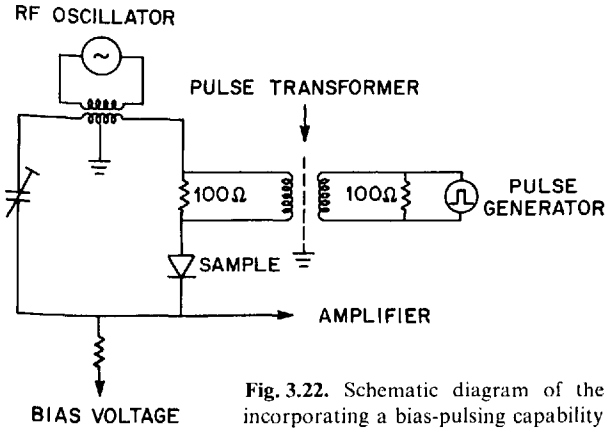

Fig. 3.22. Schematic diagram of the pulse-transformer method of incorporating a bias-pulsing capability into a capacitance meter

measurement of the capacitance of transistors. A much larger drive amplitude can be tolerated – indeed, is *beneficial* – in capacitance spectroscopy measurements. Some manufacturers, e.g., Boonton Electronics, offer a higher drive-voltage option (e.g., 100 mV_{rms}) which is very desirable. The resulting noise level on a DLTS scan at $\tau_{max}^{-1} = 50\,s^{-1}$ is less than 10^{-3} pF.

3.6.2 Fast Capacitance Bridge

The capacitance-meter method of measuring thermal-emission transients has the advantages of ease and simplicity. However, to overcome some of its shortcomings it is necessary to build a faster capacitance bridge from individual components [3.10]. Such a bridge is quite analogous to that which might be used to detect magnetic resonance signals. An updated version of the original bridge [3.10] is shown in Fig. 3.23. This bridge is capable of sub-μs pulse-overload recovery, limited by the recovery time of the amplifiers from overload. In addition, since the bias pulses are directly coupled to the sample, it is possible to span the entire ns–*to*–ms output range of the pulse generator without changing pulse transformers. This is ideal for capture-cross-section measurements.

The frequency at which the bridge is operated is not critical. Obviously, one wants a rather high frequency to obtain the fastest possible overload recovery. However, the frequency must not exceed the limit imposed by the RC time constant of the sample itself. This may be as low as a few MHz for Schottky barriers on thin epitaxial layers grown on insulating substrates. The range of 10–20 MHz seems like a good compromise for most situations.

The purpose of the phase shifter and attenuator which are in parallel with the sample is to provide an impedance of roughly equal magnitude to that of the sample but 180° out of phase. By adjusting this network ("balancing the bridge"), the input to the rf amplifier is kept small enough to be within its linear

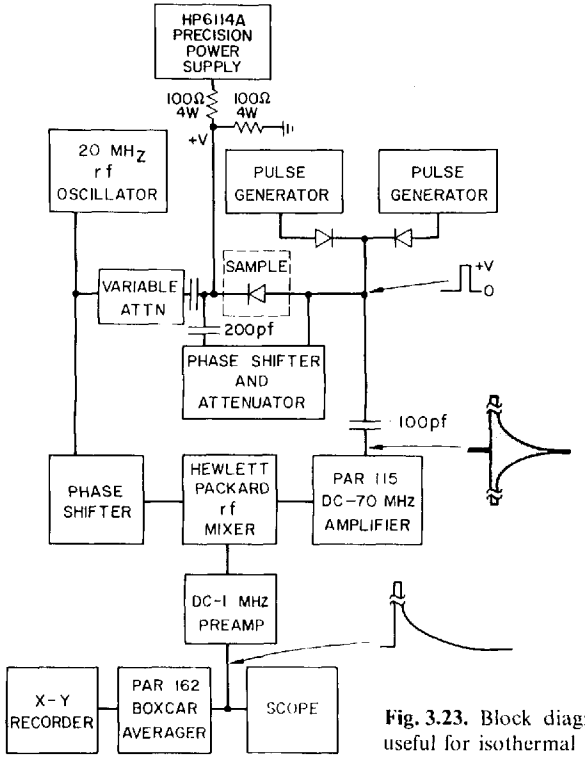

Fig. 3.23. Block diagram of a fast capacitance bridge useful for isothermal transient or DLTS measurements

range. We have found that a passive $50-\Omega$ delay line in series with a tapped variable low-inductance resistor is best. One may also use various transformer-capacitor schemes to accomplish the balancing of the bridge. However, we have found that any circuit containing inductance gives rise to unsatisfactory bias-pulse response, especially for injection pulses.

One of the principal shortcomings of the bridge in Fig. 3.23 is that it is not calibrated, and furthermore its amplitude response depends on the capacitance of the sample. This problem arises because the demands for fast recovery and direct-coupled 50Ω pulse generators dictate attempts at proper 50Ω terminations as much as possible. This, however, gives rise to bridge impedances comparable to that of the sample so that the amplitude response depends on the sample impedance. Thus we have devised a relatively simple means for quickly calibrating the bridge response for each sample *in situ*.

The key to the calibration procedure is a power supply with digital control via thumb-wheel switches so that a small, reproducible voltage step ΔV can be applied to the sample. This voltage step produces a small capacitance change ΔC which can be measured with a calibrated capacitance meter or bridge. By choosing ΔV small enough so that the resulting signal does not overload the bridge in Fig. 3.23, it is possible to measure the bridge output change

corresponding to ΔV and thus calibrate the bridge. Such a calibration is, of course, necessary to obtain deep-level concentrations, but is not needed for energy or capture-rate measurements.

3.6.3 DLTS Spectrometer

This subject has already been partially discussed in Sect. 3.4.2 in connection with the various means of implementing the DLTS rate-window concept. We wish to briefly discuss here the overall experimental system for DLTS. The main elements of a DLTS spectrometer are shown in Fig. 3.24. This system is a combination of the elements needed for isothermal transient measurements (i.e., pulsed bias supply and capacitance meter) with those needed for irreversible thermal scanning techniques such as TSCAP (i.e., temperature scan control and X–Y recorder). In addition, of course, one needs some sort of DLTS rate-window instrumentation (i.e., double boxcar, lock-in, or correlator).

The elements shown in Fig. 3.24 may be implemented in a number of different ways. For example, the bias supply/pulse generator/capacitance meter combination could be either that shown in Fig. 3.22 or the fast bridge shown in Fig. 3.23. The dewar may be of the cold-finger type [3.29] or a heated-gas-flow design. In general the major design rule for the temperature scanning system is to use thermal masses which are as small as possible and to measure the temperature with either a thermocouple or diode placed very close to the

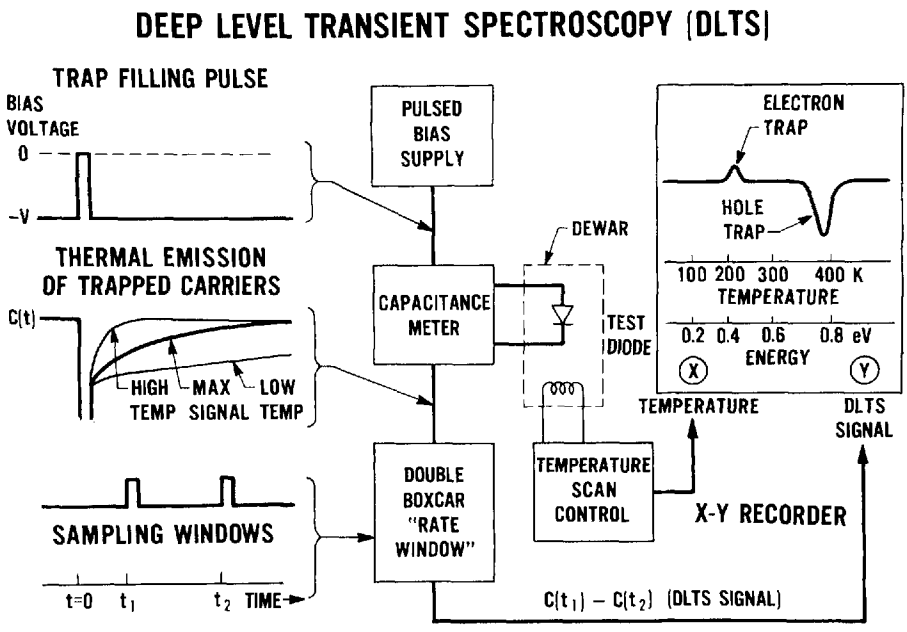

Fig. 3.24. Block diagram showing the essential features of a DLTS spectrometer

sample and in good thermal contact with it. Accurate temperature measurement is perhaps the most difficult experimental problem of thermal scanning techniques. One may test for good thermal coupling between the thermal sensor and the sample by verifying that DLTS peaks occur at the same temperature for both increasing and decreasing thermal scans. Indeed, the thermal emission rate of a particular trap makes a very good thermometer. The electronics for the temperature scan control may be as simple as a heater driven by a constant current or a more complex servo-type temperature controller driven by a voltage ramp. This latter type is advantageous for measuring capture cross sections or concentration profiles since it is possible to stop the scan at a DLTS peak and regulate the temperature at that point.

The DLTS spectrum shown on the $X-Y$ recorder in Fig. 3.24 is meant to be only illustrative but shows a typical temperature scan range (50–450 K) with a rough energy scale corresponding to $\nu = 10^{12}\,\mathrm{s}^{-1}$ and $\tau_{\max}^{-1} = 50\,\mathrm{s}^{-1}$ ($E = 23.7\,kT$). Such a scan can be achieved in about 5–10 min with a temperature accuracy of 1–2 K; this corresponds to $q \sim 1\,\mathrm{K\,s}^{-1}$.

A DLTS spectrometer using current transients would be essentially the same as Fig. 3.24 except that the capacitance meter would be replaced by a current amplifier. For SDLTS measurements in an SEM the DLTS signal at the boxcar output would be applied to a CRT intensity control for image formation and to the $X-Y$ recorder, as shown, for examining spectra from localized regions with the microscope beam held stationary.

We have seen in this chapter that the general DLTS spectrometer in Fig. 3.24 makes possible a wide variety of measurements on deep levels in semiconductor space-charge layers. One can measure the thermal ionization energy, electron and hole capture cross sections, concentration, and spatial distribution of both radiative and nonradiative defects with extremely high sensitivity. This opens a whole new horizon for deep-level studies in semiconductors.

References

3.1 S.M.Sze: *Physics of Semiconductor Devices* (Wiley, New York 1969)
3.2 C.G.B.Garrett, W.H.Brattain: Phys. Rev. **99**, 376–387 (1955)
3.3 G.Lubberts, B.C.Burkey: Solid-State Electron. **18**, 805–809 (1975)
3.4 G.L.Miller, D.V.Lang, L.C.Kimerling: Ann. Rev. Mater. Sci. **7**, 377–448 (1977)
3.5 H.G.Grimmeiss: Ann. Rev. Mater. Sci. **7**, 341–376 (1977)
3.6 J.Hildebrand, R.D.Gold: RCA Rev. **21**, 245 (1960)
3.7 P.J.Baxandall, D.J.Colliver, A.F.Fray: J. Sci. Instrum. **4**, 213 (1971)
3.8 J.A.Copeland: IEEE Trans. ED-**16**, 445 (1969)
3.9 G.L.Miller: IEEE Trans. ED-**19**, 1103 (1972)
3.10 D.V.Lang: J. Appl. Phys. **45**, 3014–3022 (1974)
3.11 A.M.Goodman: J. Appl. Phys. **34**, 329–338 (1963)
3.12 Y.Furukawa, Y.Ishibashi: Jpn. J. Appl. Phys. **5**, 837 (1966); **6**, 503 (1967)
3.13 R.R.Senechal, J.Basinski: J. Appl. Phys. **39**, 3723–3731, 4581–4589 (1968)
3.14 C.T.Sah, L.Forbes, L.L.Rosier, A.F.Tasch, Jr.: Solid-State Electron. **13**, 759–788 (1970)

3.15 Y.Zohta: Appl. Phys. Lett. **17**, 284–286 (1970)
3.16 K.Hesse, H.Strack: Solid-State Electron. **15**, 767–774 (1972)
3.17 G.H.Glover: IEEE Trans. ED-**19**, 138–143 (1972)
3.18 Y.Zohta, Y.Okmura: Appl. Phys. Lett. **21**, 117–119 (1972)
3.19 M.Bleicher, E.Lange: Solid-State Electron. **16**, 375–380 (1973)
3.20 T.Ikoma, B.Jeppson: Jpn. J. Appl. Phys. **12**, 1011–1019 (1973)
3.21 G.Goto, S.Yanagisawa, O.Wada, H.Takanashi: Jpn. J. Appl. Phys. **13**, 1127–1133 (1974)
3.22 D.V.Lang, R.A.Logan, M.Jaros: Phys. Rev. B (to be published)
3.23 R.Williams: J. Appl. Phys. **37**, 3411 (1966)
3.24 D.V.Lang, L.C.Kimerling: Proc. Intern. Conf. Lattice Defects in Semiconductors, Freiberg 1974 (Institute of Physics Conf. Series No. 23, London 1975) p. 581
3.25 D.V.Lang: J. Appl. Phys. **45**, 3022 (1974)
3.26 M.C.Driver, G.T.Wright: Proc. Phys. Soc. (London) **81**, 141 (1963)
3.27 J.C.Carballes, J.Lebailly: Solid State Commun. **6**, 167 (1968)
3.28 C.T.Sah: Solid-State Electron. **19**, 975–990 (1977)
3.29 M.G.Buehler, W.E.Phillips: Solid-State Electron. **19**, 777–788 (1976)
3.30 G.L.Miller, J.V.Ramirez, D.A.H.Robinson: J. Appl. Phys. **46**, 2638–2644 (1975)
3.31 M.D.Miller, D.R.Patterson: Rev. Sci. Instrum. **48**, 237–239 (1977)
3.32 B.W.Wessels: J. Appl. Phys. **47**, 1131–1133 (1976)
3.33 O.Wada, S.Yanagisawa, H.Takanashi: Appl. Phys. **13**, 5–13 (1977)
3.34 H.Lefevre, M.Schulz: Appl. Phys. **7**, 45–53 (1977)
3.35 C.T.Sah, H.S.Fu: Phys. Status Solidi (a) **14**, 59–70 (1972)
3.36 J.M.Herman III, C.T.Sah: Phys. Status Solidi (a) **14**, 405–415 (1972)
3.37 L.D.Yau, W.W.Chan, C.T.Sah: Phys. Status Solidi (a) **14**, 655–662 (1972)
3.38 L.D.Yau, C.T.Sah: Solid-State Electron. **17**, 193–201 (1974)
3.39 D.L.Losee: Appl. Phys. Lett. **21**, 54 (1972); J. Appl. Phys. **46**, 2204–2214 (1975)
3.40 D.V.Lang, R.A.Logan: J. Electron. Mat. **5**, 1053 (1975)
3.41 D.V.Lang, R.A.Logan: J. Appl. Phys. **47**, 1533–1537 (1976)
3.42 P.M.Petroff, D.V.Lang: Appl. Phys. Lett. **31**, 60–62 (1977)
3.43 C.H.Henry, H.Kukimoto, G.L.Miller, F.R.Merritt: Phys. Rev. B**7**, 2499–2507 (1973)
3.44 C.H.Henry, D.V.Lang: Phys. Rev. B**15**, 989–1016 (1977)
3.45 D.V.Lang, C.H.Henry: Phys. Rev. Lett. **35**, 1525 (1975)
3.46 A.G.Milnes: *Deep Impurities in Semiconductors* (Wiley, New York 1973)
3.47 M.Lax: Phys. Rev. **119**, 1502 (1960)
3.48 R.M.Gibb, G.J.Rees, B.W.Thomas, B.L.H.Wilson, B.Hamilton, D.R.Wight, N.F.Mott: Philos. Mag. **36**, 1021–1034 (1977)
3.49 C.M.Penchina, J.S.Moore: Phys. Rev. B**9**, 5217–5221 (1974)
3.50 A.Mircea, A.Mitonneau, J.Vannimenus: J. Phys. (Paris) Lett. **38**, L41–L43 (1977)
3.51 C.D.Thurmond: J. Electrochem. Soc. **122**, 1133 (1975)
3.52 J.A.VanVechten, C.D.Thurmond: Phys. Rev. B**14**, 3539–3550 (1976)
3.53 A.Mircea, A.Mitonneau: Appl. Phys. **8**, 15 (1975)
3.54 D.V.Lang, A.Y.Cho, A.C.Gossard, M.Ilegems, W.Wiegmann: J. Appl. Phys. **47**, 2558 (1976)
3.55 Pulse Engineering Inc., San Diego, California 92112 USA
3.56 L.C.Kimerling: IEEE Trans. NS-**23**, 1497–1505 (1976)