Deep-level transient spectroscopy: A new method to characterize traps in semiconductors

Cite as: Journal of Applied Physics **45**, 3023 (1974); https://doi.org/10.1063/1.1663719
Submitted: 08 January 1974 . Published Online: 06 October 2003

D. V. Lang

Deep-level transient spectroscopy: A new method to characterize traps in semiconductors

D. V. Lang

Bell Laboratories, Murray Hill, New Jersey 07974
(Received 8 January 1974)

A new technique, deep-level transient spectroscopy (DLTS), is introduced. This is a high-frequency capacitance transient thermal scanning method useful for observing a wide variety of traps in semiconductors. The technique is capable of displaying the spectrum of traps in a crystal as positive and negative peaks on a flat baseline as a function of temperature. It is sensitive, rapid, and easy to analyze. The sign of the peak indicates whether the trap is near the conduction or valence band, the height of the peak is proportional to the trap concentration, and the position, in temperature, of the peak is uniquely determined by the thermal emission properties of the trap. In addition, one can measure the activation energy, concentration profile, and electron- and hole-capture cross sections for each trap. The technique is presented with a simple theoretical analysis for the case of exponential capacitance transients. Various traps in GaAs are used as examples to illustrate certain features of the DLTS technique. Finally, a critical comparison is made with other recent capacitance techniques.

I. INTRODUCTION

In an effort to characterize the traps present in a semiconductor, one would prefer a technique that is sensitive, rapid, and straightforward to analyze. Such a technique should be able to distinguish between majority- and minority-carrier traps and should provide information about the concentrations, energy levels, and capture rates of these traps. In addition, one would prefer such a technique to be spectroscopic in the sense that the signals due to different traps be resolved from one another and to be reproducible in position when plotted against a single variable. In order to be useful as a survey technique, it is also important that any possible method should be capable of measuring traps over a wide range of depths, including both radiative and nonradiative centers.

Luminescence[1] has been widely used with great success in the study of shallow centers. The main features resposible for this success are the sensitivity, speed, and spectroscopic nature of the technique. These characteristics are thus of primary importance in evaluating any new technique aimed at studying those traps missed by luminescence, namely, the deeper nonradiative centers about which so little is known at present.

The most promising of the new techniques aimed at studying these nonradiative centers use the capacitance of a p-n junction or Schottky barrier as a probe to monitor the changes in the charge state of the centers. However, the types of capacitance techniques used to date[2-10] have lacked either the sensitivity, the speed, the range of observable trap depths, or the spectroscopic nature which would have been necessary to make them a practical means of doing spectroscopy on nonradiative centers in a large number of samples.

In this paper, we present a capacitance transient thermal scanning technique which has all of the above features, and thus is a highly attractive method either for studying unknown traps in a relatively small set of samples or for use in routine monitoring of known traps on a large scale. This technique, which we call deep-level transient spectroscopy (DLTS), is a high-frequency (MHz range) junction capacitance technique, and as such has advantages over thermally stimulated current (TSC) methods in view of its better immunity to noise and surface channel leakage current and its ability to distinguish between majority- and minority-carrier traps.[2] It has advantages over the best previous survey method, thermally stimulated capacitance (TSCAP),[2,3] in that the ultimate sensitivity and range of observable trap depths is much greater. It is more versatile than the admittance spectroscopy technique[4] since it is not limited to majority-carrier traps.

Briefly, the DLTS measurement system consists of a sensitive capacitance measurement apparatus with good transient response, one or two pulse generators to make rapid changes in the diode bias, a dual-gated signal integrator, and X-Y recorder, and a variable temperature cryostat. The presence of each trap is indicated by a positive or negative peak on a flat baseline plotted as a function of temperature. The heights of these peaks are proportional to their respective trap concentrations, the sign of each peak indicates whether it is due to a majority- or minority-carrier trap, and the positions of the peaks are simply and uniquely determined by the integrator gate settings and the thermal emission properties of the respective traps. By the proper choice of experimental parameters it is possible to measure the thermal emission rate, activation energy, concentration profile, and capture rate of each trap.

In Sec. II we review the basic concepts of pulsed bias capacitance transients. An introduction to DLTS and the expressions needed to analyze the spectra are presented in Sec. III. In Sec. IV we outline the various DLTS modes of operation and give representative examples of data on several GaAs samples. Section V is a comparison of this method with other capacitance and conductance survey techniques. Finally, in Sec. VI we summarize the contents of the paper.

II. PULSED BIAS CAPACITANCE TRANSIENTS

In order to explain DLTS we must first consider the more basic problem of capacitance transients. The use of capacitance transients for trap studies in semiconductors is well known.[5-8] This technique is used to obtain information about an impurity level in the depletion region of a Schottky barrier or p-n junction by observing the capacitance transient associated with the return to

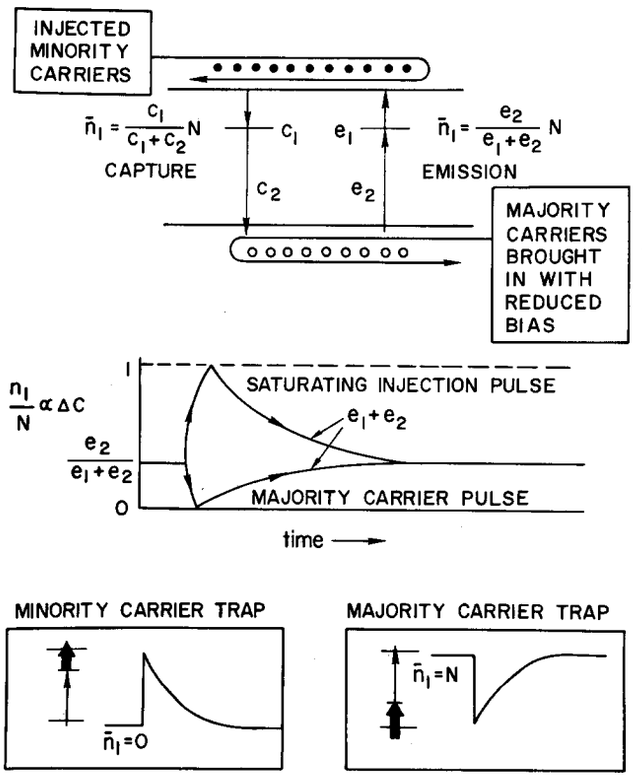

FIG. 1. Schematic summary of the emission and capture processes which describe a particular trap in p-type material and give rise to the characteristic capacitance transient of that trap.

thermal equilibrium of the occupation of the level following an initial nonequilibrium condition. One can measure the time constant of this transient as a function of temperature and obtain the activation energy for the level. The initial magnitude of the transient is related to the concentration of the trap. The form of this technique used in DLTS makes use of one or more voltage pulses applied to the sample in order to define the initial conditions. The concepts and notation of the pulsed bias method have been described previously.[7,9] For completeness we will outline the method here as well.

For simplicity, we will describe the situation in p-type material in an asymmetric n^+p diode. The results are, of course, completely general with only trivial changes of notation. In more symmetric junctions the main difference is that traps on *both* sides of the junction will give comparable signals, whereas for the highly asymmetric case the depletion region is primarily in the low-doped material, and consequently one is much more sensitive to traps only on that side of the junction. Thus, in our discussion we will consider *only* those traps in the low-doped p-side of the n^+p diode.

Figure 1 is a schematic summary of the emission and capture processes which characterize a particular trap. The capture and thermal emission rates for minority carriers (electrons in this example) are c_1 and e_1, respectively. The capture and thermal emission rates for majority carriers (holes) are c_2 and e_2, respectively. In the quiescent state of the system the diode is reverse biased and the observable traps are within the depletion region. Thus, the capture rates are zero and the occupation of the level is determined by the thermal emission rates e_1 and e_2. As indicated in Fig. 1, the steady-state electron occupation of a level is

$$\bar{n}_1 = [e_2/(e_1+e_2)]N, \tag{1}$$

where N is the concentration of the trap. We define an electron (minority-carrier) trap as one which tends to be empty ($\bar{n}_1=0$) of electrons (minority carriers), and thus capable of capturing them. Likewise, a hole (majority-carrier) trap is one which tends to be full ($\bar{n}_1=N$) of electrons (minority carriers), and thus capable of having a trapped electron recombine with a hole, i.e., capture a hole. Then according to Eq. (1), an electron trap has to have $e_1 \gg e_2$ and a hole trap $e_2 \gg e_1$. This is indicated by the heavy and light arrows at the bottom of Fig. 1. The emission rates are proportional to a Boltzmann factor, and thus depend exponentially on the energy difference between the trap level and the conduction band (electron emission) and the trap level and the valence band (hole emission). Because of this, electron traps tend to be in the upper half of the gap and hole traps in the lower half.

As shown in Fig. 1, a capacitance change is caused by using a bias pulse to introduce carriers, and thus changes the electron occupation of a trap from the steady-state value in Eq. (1). As this population returns to equilibrium, the capacitance returns to its quiescent value. For the simple linear rate equations illustrated in Fig. 1, the transient is an exponential function of time with a rate constant equal to e_1+e_2. One of these rates usually dominates, thus the transient rate for a majority-carrier trap is e_2 and for a minority-carrier trap is e_1. The sign of the capacitance change depends on whether the electron occupation of the trap had been increased or decreased by the pulse. An increase in trapped minority carriers causes an increase in the junction capacitance. As indicated in the bottom of Fig. 1, the capacitance transient due to a

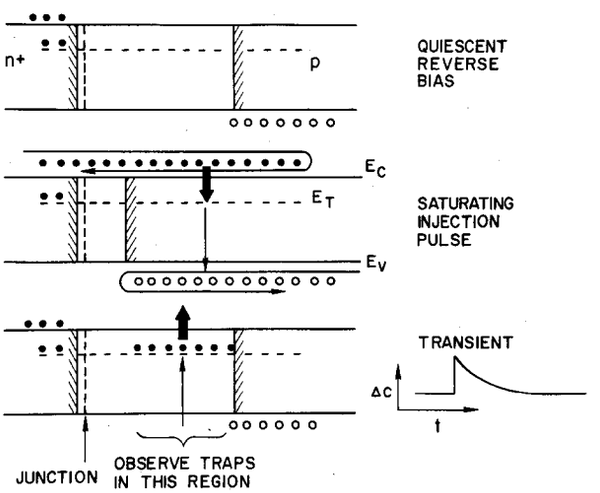

FIG. 2. Injection pulse sequence which is used to produce a capacitance transient for a minority-carrier trap. The energy-vs-distance diagrams (with band bending omitted for simplicity) show the p^+n junction depletion region (edges denoted by shaded lines) as well as the capture and emission processes and trap occupation before, during, and after an injection pulse.

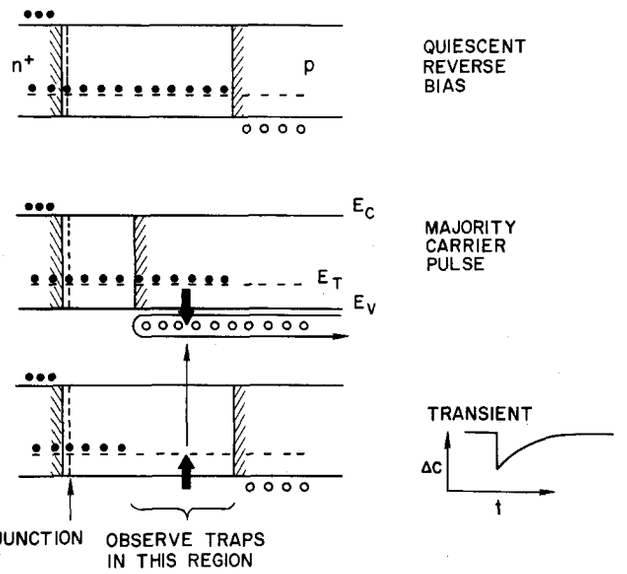

FIG. 3. Majority-carrier pulse sequence which is used to produce a capacitance transient for a majority-carrier trap. The energy-vs-distance diagrams (with band bending omitted for simplicity) show the p^+n junction depletion region (edges denoted by shaded lines) as well as the capture and emission processes and trap occupation before, during, and after a majority- carrier pulse.

minority-carrier trap is always positive and is induced only by injected minority carriers, whereas the transient due to a majority-carrier trap is always negative and is induced only by majority carriers.

There are two main types of bias pulses, namely, an injection pulse which momentarily drives the diode into forward bias and injects minority carriers into the region of observation shown in Fig. 2, and a majority-carrier pulse[11] which momentarily reduces the diode bias and introduces only majority carriers into the region of observation shown in Fig. 3. As indicated in Fig. 1, the steady-state electron occupation during a bias pulse is

$$\bar{n}_1 = [c_1/(c_1+c_2)]N, \tag{2}$$

where c_1 is the minority-carrier (electron) capture rate proportional to the concentration of injected minority carriers and c_2 is the majority-carrier (hole) recombination rate proportional to the majority-carrier concentration. In all situations we will consider, the capture rates are much larger than the emission rates, which can be neglected during the bias pulse. An injection pulse which introduces a large enough number of electrons so as to make $c_1 \gg c_2$ and overwhelm the trap emptying process will completely fill the trap with electrons; such a pulse is called a saturating injection pulse. A majority-carrier pulse, on the other hand, introduces only holes, and thus tends to empty all traps of electrons, i.e., fill them with holes.

Figure 2 is an injection pulse sequence which is used to produce a capacitance transient for the case of a minority-carrier (electron) trap, while Fig. 3 shows the analogous majority-carrier pulse sequence for a majority-carrier (hole) trap. Both Figs. 2 and 3 are schematic energy-vs- distance diagrams of the n^+p junction which we have been describing with the band bending due to the junction electric field omitted for simplicity. The edges of the depletion region are indicated by shaded lines. Figures 2 and 3 show, from top to bottom, the depletion region, capture and emission processes, and trap electron occupation before, during, and after the appropriate bias pulse. Figure 4 is a schematic illustration of the time dependences of the various experimental parameters associated with a majority-carrier trap pulse sequence (top) and a minority-carrier trap injection pulse sequence (bottom).

The concentration of a trap can be obtained directly from the capacitance change corresponding to completely filling the trap with a saturating injection pulse (in the case of minority-carrier trap) or the largest possible majority-carrier pulse (in the case of a majority-carrier trap). The relationship for an electron trap in an n^+p step junction is simply

$$N = 2(\Delta C/C)(N_A - N_D), \tag{3}$$

where N is the trap concentration, ΔC is the capacitance change at $t=0$ due to a saturating injection pulse, C is the capacitance of the diode under quiescent reverse-biased conditions, and $N_A - N_D$ is the net acceptor concentration on the p side of the junction where the trap is observed. A more precise determination of trap concentration under more general doping conditions may be made by concentration profiling as discussed in Sec. IV.

III. DEEP-LEVEL TRANSIENT SPECTROSCOPY: THEORY

The physics of DLTS is in the capacitance transients

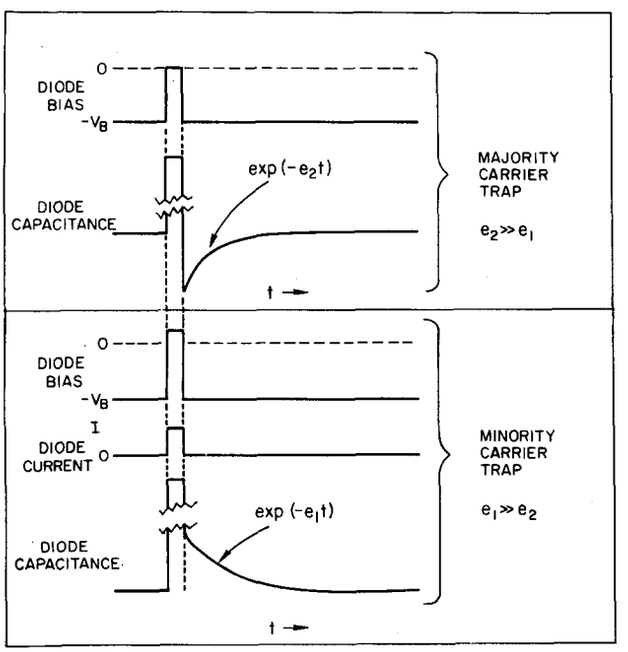

FIG. 4. Typical time dependences involved in pulsed bias capacitance transients for majority- and minority-carrier traps. The upper half is a majority-carrier pulse sequence while the lower half is an injection-pulse sequence.

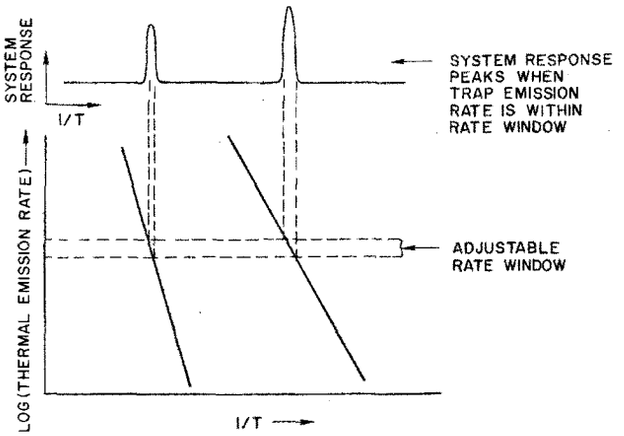

FIG. 5. Schematic illustration of the basic idea of DLTS method, namely, the rate window concept. The lower part of the figure is a typical activation energy plot for the case of two traps. The upper part of the figure shows the resulting response of a capacitance transient measurement apparatus equipped with a rate window.

outlined in Sec. II. The primary contribution of the DLTS scheme as presented here lies in the convenience and speed with which the information inherent in a series of capacitance transient experiments may be obtained. That is to say, the information which may be obtained by several 10-min DLTS thermal scans could also be obtained from a more tedious and time consuming point-by-point observation of the capacitance transients at many different fixed temperatures. This latter alternative may be helpful in the detailed study of an unknown trap since it can be more precise, but in the initial characterization of many traps in an unknown sample or in high-volume monitoring work the DLTS scheme has distinct advantages.

The essential feature of DLTS is the ability to set an emission rate window such that the measurement apparatus only responds when it sees a transient with a rate within this window. Thus, if the emission rate of a trap is varied by varying the sample temperature, the instrument will show a response peak at the temperature where the trap emission rate is within the window. These emission rates are thermally activated and by the principle of detailed balance can be given as

$$e_1 = (\sigma_1 \langle v_1 \rangle N_{D1}/g_1) \exp(-\Delta E/kT), \tag{4}$$

where σ_1 is the minority-carrier capture cross section, $\langle v_1 \rangle$ is the mean thermal velocity of minority carriers, N_{D1} is the effective density of states in the minority-carrier band, g_1 is the degeneracy of the trap level, and ΔE is the energy separation between the trap level and the minority-carrier band. An exactly similar equation holds for e_2 with all subscripts changed from 1 to 2 and with all quantities referring to majority carriers. A standard means of characterizing the depth of a trap is to construct a plot of $\log e_1$ or $\log e_2$ vs $1000/T$ and to report the slope of the resulting straight line as the activation energy of the trap. This quantity is, of course, only equal to ΔE if the prefactors in Eq. (4) are independent of temperature. Since these quantities are, in general, temperature dependent one must use more care to obtain an accurate ΔE value.

Figure 5 illustrates a typical activation energy plot and shows how a measurement system which only responds to transients within a selected rate window can resolve the signals from different traps as a function of temperature. It is the strong temperature dependence of the thermal emission rates in Eq. (4) which allows the rate window thermal scan in Fig. 5 to resolve the signals.

The DLTS scheme presented here makes use of a dual-gated signal averager[12] (double boxcar) to precisely determine the emission rate window and to provide signal averaging capability to enhance the signal-to-noise ratio for the detection of low-concentration traps. When coupled with the fast-response capacitance transient measurement apparatus detailed in Ref. 7, the maximum rate window attainable is on the order of 10^5 sec^{-1}. There is in principle no lower limit on rate windows, but in practice it is inconvenient to go below 1 sec^{-1} because the thermal scan time becomes rather long if one needs to signal average. Thus, one can cover a range of at least five orders of magnitude in rate windows with currently available components.

The use of a double boxcar to select the rate window is illustrated in Fig. 6. The capacitance transients are observed on a fast-response capacitance bridge, such as in Ref. 7. (A commercial capacitance meter may

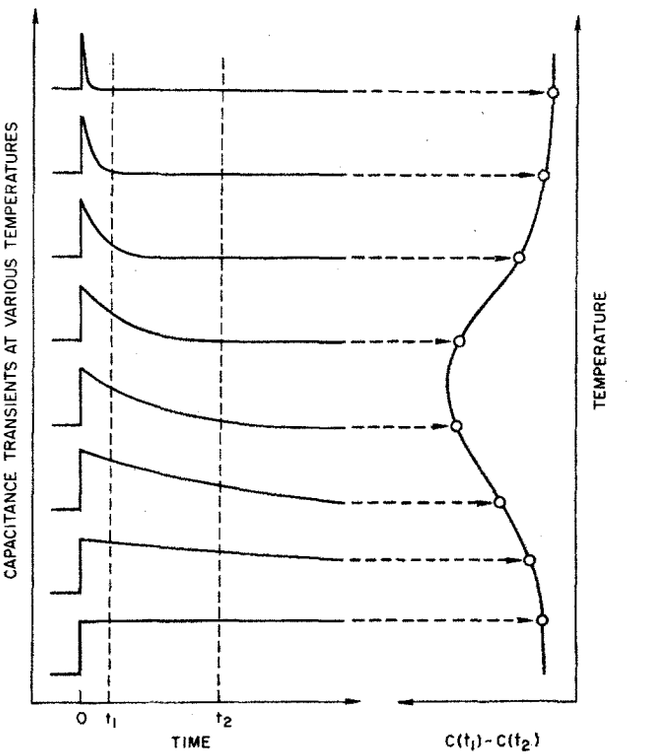

FIG. 6. Illustration of how a double boxcar is used to define the rate window. The left-hand side shows capacitance transients at varius temperatures, while the right-hand side shows the corresponding DLTS signal resulting from using the double boxcar to display the difference between the capacitance at time t_1 and the capacitance at time t_2 as a function of temperature.

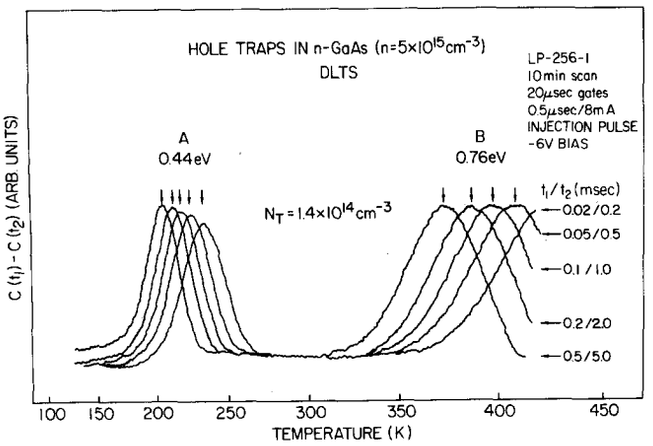

FIG. 7. Typical experimental DLTS spectra for hole traps in n-GaAs. The two traps are labeled A and B and have activation energies measured from the valence band of 0.44 and 0.76 eV, respectively. The trap concentrations are both 1.4×10^{14} cm^{-3}. Five different spectra are shown corresponding to the five rate windows determined by the values of t_1 and t_2.

also be used, but one is then limited to much lower rates due to the slower response times of such instruments.) A series of such capacitance transients for a typical trap at various temperatures is shown schematically on the left-hand side of Fig. 6. According to Eq. (4) and Fig. 5, the emission rate is very small for low temperatures and becomes more rapid as the temperature is increased. This example shows a minority-carrier trap (since ΔC at $t=0$ is positive) which is filled at $t=0$ by a saturating injection pulse. The actual signal is repetitive on a time scale somewhat larger than that shown in Fig. 5. These transient signals are fed into a double boxcar with gates set at t_1 and t_2 as shown in Fig. 6. The signal from the differential output of the boxcar is applied to the y axis of an X-Y recorder. This is simply the capacitance at t_1, $C(t_1)$, minus the capacitance at t_2, $C(t_2)$. It is clear from Fig. 6 that this quantity, $C(t_1)-C(t_2)$, goes through a maximum when τ, the inverse of the transient rate constant, is on the order of t_2-t_1. Thus, the values of t_1 and t_2 determine the rate window for a DLTS thermal scan.

Let us now derive the relationship between τ_{max}, the value of τ at the maximum of $C(t_1)-C(t_2)$ vs T for a particular trap, and the positions of the two gates, t_1 and t_2. The quantity that has been somewhat loosely referred to as the rate window is now defined to be τ_{max}^{-1}. For simplicity we will derive the desired expression assuming infinitesimally narrow gates. It can be shown, however, that our result also holds under the less stringent requirement that the gate width be much less than τ_{max}. The complete solution for arbitrary gate widths is straightforward but somewhat cumbersome, and thus will not be presented here. Let us define the normalized DLTS signal, $S(T)$, shown on the right-hand side of Fig. 6 as

$$S(T)=[C(t_1)-C(t_2)]/\Delta C(0), \tag{5}$$

where $\Delta C(0)$ is the capacitance change due to the pulse at $t=0$. Then for exponential transients we have

$$S(T)=[\exp(-t_1/\tau)]-[\exp(-t_2/\tau)], \tag{6}$$

where the temperature dependence of τ is given by Eq. (4). This can also be written

$$S(T)=\exp(-t_1/\tau)[1-\exp(-\Delta t/\tau)], \tag{7}$$

where $\Delta t=t_2-t_1$. The relationship between τ_{max} and t_1 and t_2 is simply determined by differentiating $S(T)$ with respect to τ and setting the result equal to zero. The desired expression is then

$$\tau_{max}=(t_1-t_2)[\ln(t_1/t_2)]^{-1}. \tag{8}$$

Thus, the emission rate corresponding to the maximum of a trap peak observed in a DLTS thermal scan is a precisely defined quantity and may be used along with the temperature corresponding to the peak maximum in constructing a semilog activation energy plot such as Fig. 5. That is to say, at the maximum of the DLTS signal, one can measure the temperature and calculate τ_{max} from Eq. (8) to get one point of a $\log e_1$- or $\log e_2$-vs-$1000/T$ plot. Other points can similarly be obtained from other scans made with different gate settings, and thus different values of τ_{max} and different trap peak positions. The magnitude of the peak maximum can be related to $\Delta C(0)$ via Eqs. (7) and (8), and thus trap concentrations can be derived via Eq. (3).

It is clear from the preceeding analysis that the rate at which the temperature is varied plays no role in determining the shape or position of the DLTS peaks. However, this statement is only true within reasonable bounds. That is to say, if the temperature measurement device is reading the true diode temperature and if the boxcar time constant is short enough so as not to distort the signal, then the rate or direction of the thermal scan does not affect the data. This means that one could observe the cool-down spectrum with one rate window and the warm-up spectrum with a different rate window to get a rough measure of the various activation energies in a single temperature cycle. In fact, the reproducibility of an up and down scan with the same rate window is a good way to verify that the measured temperature is the true temperature and that the signal shape is not being distorted by the electronics.

IV. DLTS MODES OF OPERATION

In making a series of thermal scans to determine activation energies one must vary t_1 and t_2 to obtain different rate windows for each scan according to Eq. (8). Three schemes for systematically varying these parameters come to mind: namely, (i) t_1 fixed, vary t_2; (ii) t_2 fixed, vary t_1; and (iii) t_1/t_2 fixed, vary both t_1 and t_2. Of these three, the last alternative seems preferable for reasons we will now discuss.

Figure 7 shows a typical series of DLTS spectra with five different rate windows. The two minority-carrier traps (hole traps in this case) in the n-GaAs sample are labeled A and B for convenience. The activation energies measured by careful point-by-point observation of the capacitance transients at fixed temperatures are 0.44 and 0.76 eV from the valence band, respectively. The concentrations of both of these traps is about 1.4×10^{14} cm^{-3} in an undoped sample with $n\approx5\times10^{15}$ cm^{-3}

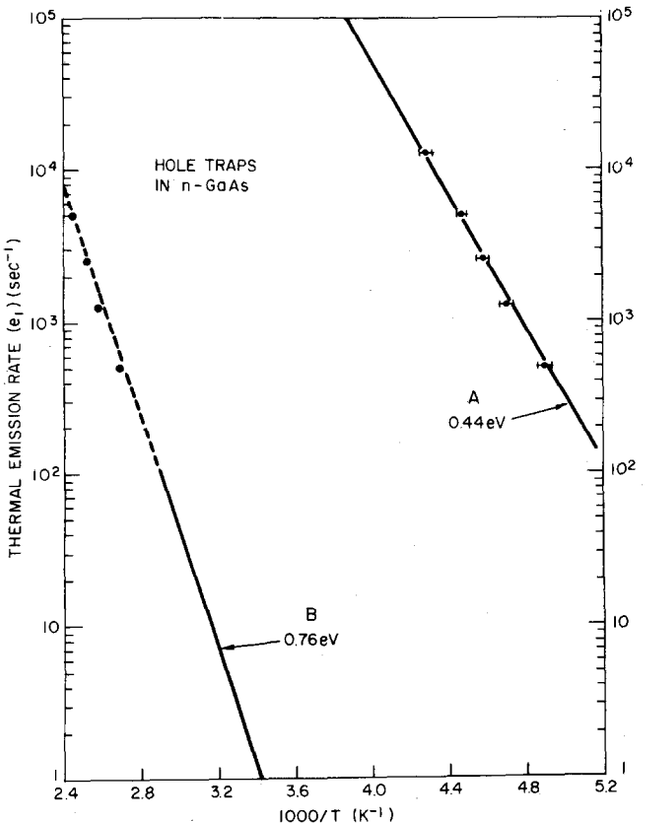

FIG. 8. Thermal emission rates vs $1000/T$ determined from the DLTS spectra in Fig. 7. The solid lines are determined from careful fixed-temperature measurements of the capacitance transients; the dotted lines are extrapolations of this data. The error bars on the DLTS data represent the uncertainty in locating the peaks in the spectra.

grown by liquid-phase epitaxy. The variation of t_1 and t_2 in Fig. 7 is according to (iii) above. The important feature to note is that the peaks shift more or less rigidly without changing their shape significantly as one varies the rate window. If one uses method (i) and fixes t_1 while varying t_2, the peaks change considerably in both size and shape with the low-temperature side moving as t_2 is varied, while the high-temperature side moves very little at all. Method (ii) produces exactly the opposite result with the high-temperature side moving as might be expected. Equation (8) is of course valid in all three cases and the data is consistent among the three ways of doing things. It is, however, easier to locate the peak as well as aesthetically more pleasing to use the fixed ratio scheme (iii). In addition, data reduction is easier since $\ln(t_1/t_2)$ is then a constant in Eq. (8).

It is of interest to compare the rate-vs-$1/T$ data obtained from Fig. 7 with that obtained from a careful series of measurements at various fixed temperatures. The temperatures corresponding to the maxima of the trap peaks are marked in Fig. 7 with arrows. The corresponding emission rates at these temperatures can be simply calculated from the t_1 and t_2 values via Eq. (8). The resulting data points are shown in Fig. 8 with the horizontal error bars indicating the uncertainty in locating the peak maxima in Fig. 7. The solid lines are obtained by many careful fixed-temperature observations of these capacitance transients. The dashed line is an extrapolation in the case of trap B since the DLTS peak was observed in a different temperature range from the fixed-point data. The agreement between the two methods is really remarkable in view of the vast difference in time and effort involved in taking the data. As a demonstration of the accuracy of the DLTS method, let us again emphasize that what we have done in Fig. 8 is to take only the positions of the peak maxima in Fig. 7 with Eq. (8) and the gate settings and with no adjustable parameters we obtain very good agreement with what is essentially the correct result.

It is also easy to obtain trap concentrations from a DLTS thermal scan. Since τ_{max} is known from the set values of t_1 and t_2, one can use Eq. (6) or (7) in conjunction with Eq. (5) to obtain $\Delta C(0)$ at the trap peak maximum. Then if the injection or majority-carrier pulse is large enough and long enough to completely fill the trap, one may use Eq. (3) to obtain the trap concentration. This last condition may be easily checked by making several scans with increasing larger and longer pulses until the trap peak no longer increases in size. Monitoring the concentrations of traps of known properties in a series of samples is made particularly convenient by the DLTS method. One knows *a priori* where on the temperature scale a given trap will appear with a particular rate window as well as the magnitude and length of majority carrier or injection pulse necessary to completely fill the trap and thus make Eq. (3) valid. A cool-down scan with a slow rate window followed by a warm-up scan with a fast rate window should suffice to bring most types of traps into view in a manageable temperature range. We should point out, however, that the above scheme is strictly valid only in the case of traps with *exponential* transients. The nonexponential case does occur,[7] especially for intermediate and shallow traps, and care must be taken when analyzing this situation. Relative comparisons of the same trap in a series of samples with the same rate window is still valid, however. The only problem is that one cannot use Eqs. (5)—(8) to obtain $\Delta C(0)$ in the nonexponential case, thus absolute determinations of N via Eq. (3) are more difficult.

Let us now discuss the more complex measurements of concentration profiles and capture rates. The procedure is exactly the same as presented in Ref. 7, except that all traps are measured simultaneously on a DLTS thermal scan without the difficulty of separating signals that occur in a fixed-temperature method. We will only, very briefly, outline the technique here; the reader is referred to Ref. 7 for a complete discussion.

Concentration profiling of the traps in a DLTS scan is straightforward, but it requires a number of scans for good resolution. Figures 9 and 10 show typical concentration profile data on the minority-carrier traps A and B in n-GaAs. The essence of the method is exactly the same here as at fixed temperatures in Ref. 7, namely, one looks at the dependence of the signal magnitude on the voltage of the majority-carrier pulse. For majority-carrier traps this is particularly

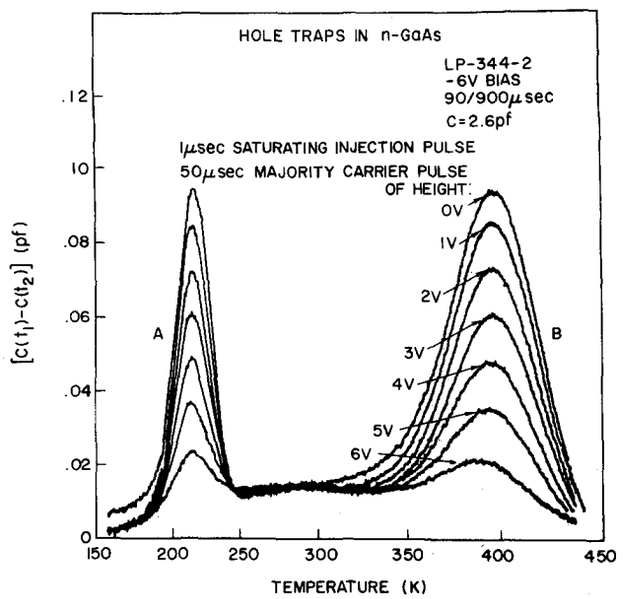

FIG. 9. Example of concentration profiling with DLTS. The seven superimposed spectra show the same hole traps in LPE n-GaAs as were shown in Fig. 7. On all runs the traps are filled with holes by a 1-μ sec saturating injection pulse followed a few μ sec later by a 50-μ sec majority-carrier pulse of variable amplitude. The resulting DLTS signal destruction due to successive scans with increasing amplitude of the second pulse is the basic information used to determine the trap concentration profile.

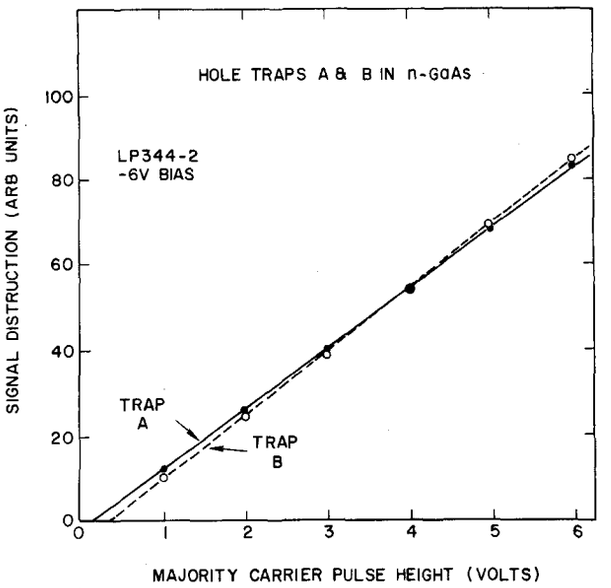

FIG. 10. Signal destruction vs majority-carrier pulse height (in volts) for the DLTS scans in Fig. 9. The linear dependence obtained indicates a uniform concentration for these two traps of 1.4×10^{14} cm^{-3}.

simple: as one varies the majority-carrier pulse from zero to a maximum value just below injection the trap peaks increase in size, corresponding to the larger portion of the quiescent depletion region in which the traps have been filled (see Fig. 3). Here one must use a long enough pulse to completely fill those traps with majority carriers in the region selected by the pulse. One then makes a plot of signal vs majority-carrier pulse voltage similar to Fig. 10. If this is a straight line, then, as shown in Ref. 7, the trap concentration profile is proportional to the carrier profile. Deviations from linearity indicate corresponding deviations from the carrier profile. Minority-carrier traps as shown in Figs. 9 and 10 require two pulse generators for profiling. The first, as in a normal DLTS scan, is a saturating injection pulse which fills all traps with minority carriers in the region of observation (see Fig. 2). The second, as in Ref. 7, is a majority-carrier pulse of adjustable voltage which is long enough to completely empty the traps of minority carriers in the region selected by the voltage of the pulse. As in the case of majority-carrier traps above, one makes a series of scans with progressively larger majority-carrier pulses (Fig. 9), but here one plots the signal *destruction* vs pulse voltage in Fig. 10 to obtain the profile. By using Eq. (A7) of Ref. 7, the data in Fig. 10, and a carrier profile for this sample we find that both traps A and B have uniform concentrations of 1.4×10^{14} cm^{-3}.

Capture rates are determined by measuring the trap peak heights as a function of the injection or majority-carrier pulse *width* as in Ref. 7. Figure 11 shows a typical example taken from radiation damage traps in n-GaAs. Note here also that the sign difference for majority- and minority-carrier traps is graphically illustrated. To measure both capture rates for a given trap, one must use two pulse generators—one to measure the filling rate, the second to measure the emptying rate once the trap is full. Majority-carrier capture rates are straightforward to measure. One simply plots the signal height vs the majority-carrier pulse width to obtain the capture rate directly. Minority capture is

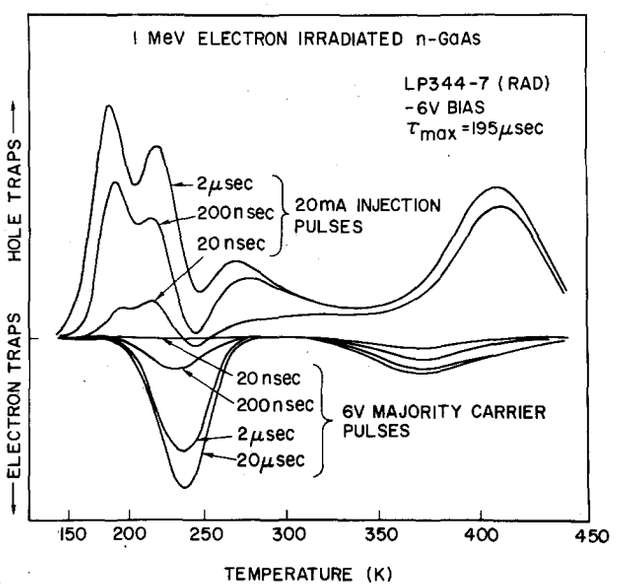

FIG. 11. Typical example of the dependence of trap peak heights on the *width* of the injection or majority-carrier pulse. The traps shown in this figure have been produced by irradiating a p^+n LPE GaAs diode with 1-MeV electrons. Note that the electron traps (majority-carrier traps in this case) give negative signals and are only produced by majority-carrier pulses, while hole traps give positive signals and are only produced by injection pulses.

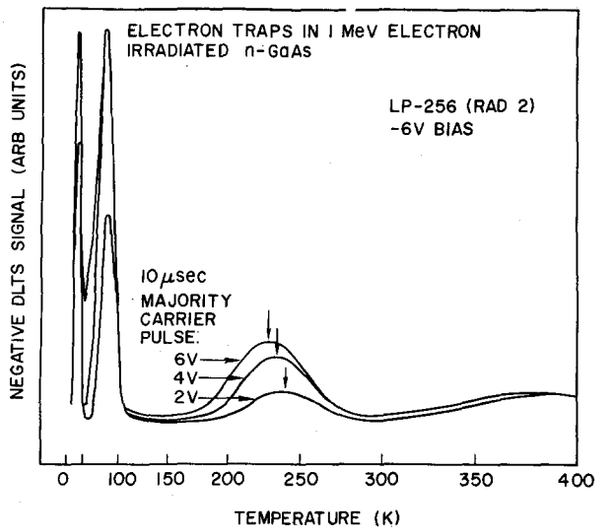

FIG. 12. Example of the trap peak shift exhibited by nonexponential transients which are due to a dependence of the emission rate on the junction electric field. The sign of the DLTS signal has been inverted.

more complex since the rate depends on the concentration of injected minority carriers. By plotting the filling rate for a minority-carrier trap or the emptying rate for a majority-carrier trap as a function of injected current, one may obtain minority-carrier capture rates. Absolute values are hard to obtain in GaAs because of the difficulty in relating the injected minority-carrier concentration to the measured injection pulse current. This problem is less severe in Si and GaP, however.[7]

Perhaps the most important traps are those which act as fast nonradiative recombination centers. Since such traps tend to dominate the recombination kinetics, they are desirable in fast semiconductor switches, for instance, but detrimental to the performance of semiconductor lasers and light-emitting diodes. The DLTS method provides a particularly simple way to bias the spectrum in favor of such fast capturing traps while suppressing the signal from slower traps, namely, to use very short pulses as shown in Fig. 11. Such short pulses will be able to fill a fast trap but unable to appreciably fill a slower trap. Thus, the pulse width is a very effective discriminator in the DLTS method to suppress the signals from slow, possibly less interesting, traps.

Finally, a word should be said about the problem of nonexponential transients raised earlier. We again emphasize that all of the previous analysis is strictly correct only for exponential capacitance transients. The DLTS method still works very well for nonexponential transients but care must be taken in obtaining absolute trap parameters from the data. As mentioned earlier, however, sample-to-sample relative comparisons with the same rate window are perfectly valid regardless of the shape of the transient.

The nonexponential behavior due to spatial variations in the emission rate of a trap caused by the spatial variation in the junction electric field is relatively easy to detect by DLTS directly. Such behavior is typified by the ZnO center in GaP.[7] An example of a spatially inhomogeneous emission rate detected by DLTS is shown in Fig. 12 for radiation-induced traps in n-GaAs. To observed this type of inhomogeneity one follows the procedure for concentration profiling outlined above. If as in Fig. 12 the peak position and shape changes as the majority-carrier pulse voltage is changed, then spatial inhomogeneity is present. This can be understood by considering the depletion region to be divided into n regions in each of which the signal is a perfect exponential, but with different rate constants in each region. Each of these n signals will appear in a different position on a DLTS scan. When all are present together, one observes a broad trap peak in which the individual components are not resolved. When one looks at fewer of these signals, as is done in a concentration profile sequence, this broad peak appears to shift and narrow as some of the n component lines of the original are removed. A homogeneous trap peak due to a pure exponential transient will remain fixed in position and shape during a concentration profile as in Fig. 9.

V. COMPARISON OF DLTS WITH OTHER CAPACITANCE TECHNIQUES

In this section we will briefly discuss the DLTS method as it relates to other capacitance techniques currently employed in studying traps in semiconductors. The other main techniques which we will discuss are thermally stimulated capacitance (TSCAP),[3] edge region TSCAP,[2] admittance spectroscopy,[4] and photocapacitance.[10] Another technique that has been widely used is thermally stimulated current (TSC); however, we will not make detailed comparisons with this technique since its shortcomings relative to high-frequency capacitance techniques have already been pointed out.[2] If there are circumstances where current measurements are desirable, however, it is worth noting that the basic DLTS scheme is applicable to current transients as well as capacitance transients.

Photocapacitance[10] is sensitive but not particularly useful as a survey technique since the data must be taken point by point and a rather involved analysis of optical cross sections is needed to obtain accurate trap parameters. In addition, it is limited to traps deeper than about 0.3 eV. Photocapacitance could be useful, however, if one were interested in the optical properties of a deep trap.

In addition to DLTS, the other technique which may be considered spectroscopic is the so-called admittance spectroscopy method recently reported by Losee.[4] Like DLTS, this method presents resolved peaks corresponding to each trap and is independent of thermal scan rate or direction. It is, however, limited to only majority-carrier traps. Unlike most other capacitance techniques, it is at its best for shallow traps with decreasing sensitivity for the deeper traps.

The previous best capacitance survey method was TSCAP.[2,3] DLTS may be viewed as an improved version of TSCAP which has the following advantages: (i) much greater sensitivity, (ii) greater range of observable

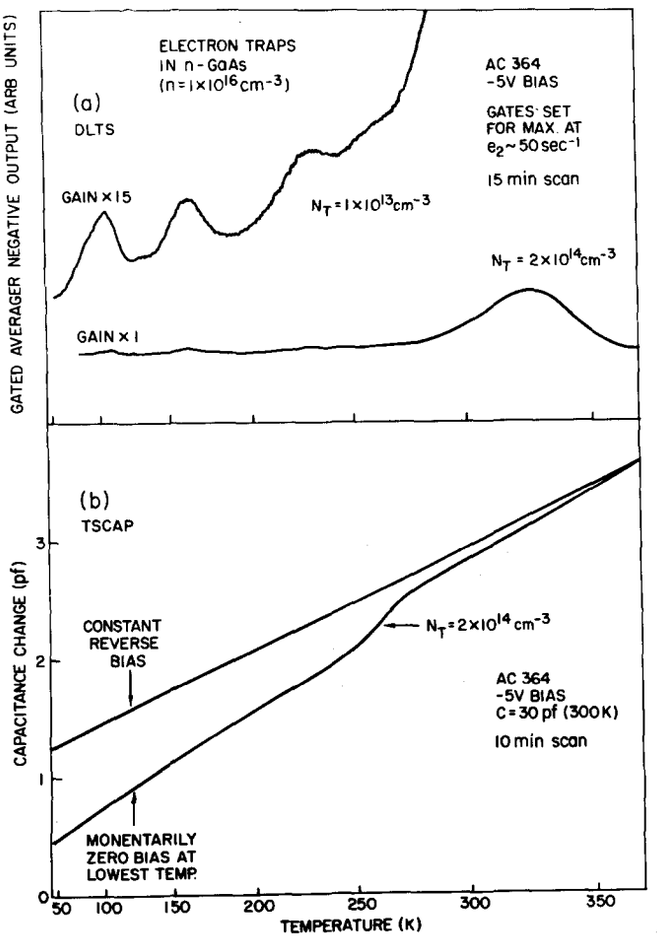

FIG. 13. As an example of the sensitivity of DLTS, this figure shows three electron traps in molecular beam epitaxial n-GaAs with 1×10^{13} cm^{-3} concentration. (a) shows a DLTS spectrum which displays these traps with good S/N. (b) shows a TSCAP run on the same sample which fails to observe these low-concentration traps. Two larger concentration traps are visible in the TSCAP run, only one of these is on scale in the DLTS scan.

trap depths, and (iii) much more convenient to use and to interpret.

The sensitivity of DLTS is graphically illustrated in Fig. 13. This shows electron traps in n-type GaAs ($n=1\times10^{16}$ cm^{-3}) grown by molecular beam epitaxy.[13] Figure 13(a) shows two DLTS thermal scans with different gains. Figure 13(b) shows a TSCAP run on the same sample. Note the three trap levels with about 1×10^{13} cm^{-3} concentration in the upper DLTS scan. We could have detected traps at 10^{12} cm^{-3} and even lower (i.e., $\Delta C/C \lesssim 10^{-4}$) with a longer scan to allow more averaging time. The trap at 2×10^{14} cm^{-3} concentration is visible in both the DLTS and TSCAP runs at about 330 and 260 K, respectively. The three low-concentration traps should be in the TSCAP spectrum somewhere below about 200 K, but are clearly not visible. A differential TSCAP run would have had a better chance of seeing these traps, but it still could not compete with the ultimate sensitivity of DLTS. The reason for the sensitivity of DLTS is the same as for the well-known case of lock-in detection, namely, a signal with a rapid repetion rate can be detected and amplified in a relatively quiet region of the noise spectrum which is well away from the unavoidable low-frequency drifts and other $1/f$ noise in the signal processing equipment. That is to say, with identical scan rates and filter time constants the TSCAP or differential TSCAP signal (which is essentially at dc for scans of several minutes or more) will be subject to more problems of drift, etc., than an equivalent DLTS signal which could be observed with a repetition rate as high as 20 kHz.

Another factor in the increased sensitivity of DLTS is the fact that no baseline subtraction or correction is necessary. This is because the signal is due only to the repetitive transient population changes of traps, and thus there is no need to subtract the baseline due to the temperature dependence of the over-all diode capacitance as is the case in TSCAP. Figure 13(b) shows how the diode capacitance can vary with temperature, even without deep-trap population changes. With high gain this baseline shift can be servere in some samples, even with differential TSCAP.

The wide range of traps observable by DLTS is illustrated in Fig. 14. This shows three majority-carrier traps in electron-irradiated n-GaAs. The rate window for this scan is 10^4 sec^{-1}, which is still a factor of 10 less than the maximum. By starting the thermal scan at about 30 K, one is able to observe traps with activation energies of 0.08, 0.17, and 0.38 eV relative to the conduction band. By contrast, the effective rate window of TSCAP is typically on the order of 0.1 sec^{-1} and therefore, one is limited to traps deeper than approximately 0.25 eV. Edge region TSCAP [2] can observe such shallow and intermediate majority-carrier traps, but it much less sensitive than ordinary TSCAP and thus limited to only large concentration species such as intentionally doped centers.

In terms of convenience of use and interpretation, two of the main features of DLTS which make it an improvement over TSCAP are the independence of

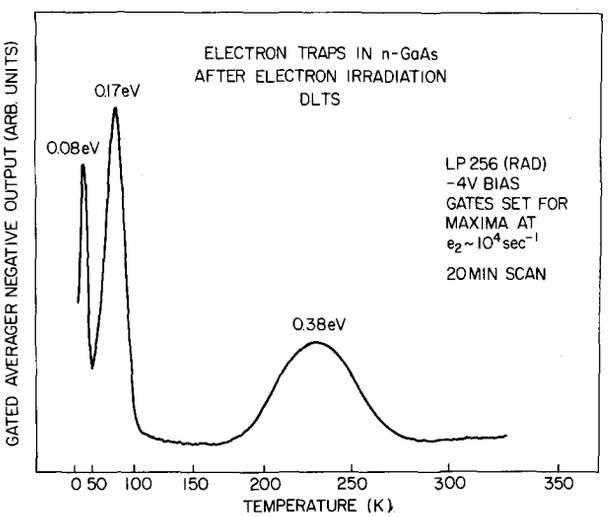

FIG. 14. Example of shallow electron traps in 1-MeV-electron-irradiated n-GaAs as observed by DLTS. The activation energies measured relative to the conduction band are 0.08, 0.17, and 0.38 eV. The sign of the DLTS signal has been inverted.

thermal scan rate and the ease with which the rate window can be changed and measured for activation energy analysis. The characteristics of the thermal scan are particularly important for a differential TSCAP spectrum. Relating the size of the derivative for a particular trap peak to the true value of ΔC needed for concentration measurements is considerably more involved than the direct ΔC values obtained from a DLTS scan. In addition, one must take pains to make the differential TSCAP scan rate reproducible and linear to get even relative concentration values.

One drawback of DLTS is that it might miss minority-carrier traps that cannot be saturated at practical levels of forward current (i.e., c_2 is extremely large). Even if one could saturate such a trap during the injection pulse, it is possible that it might be able to empty during the fall time of the pulse. An example of such a level is the second electron state of the oxygen double donor in GaP.[9,10] Another problem occurs if one must use a Schottky barrier instead of a p-n junction. In this case, one misses the minority-carrier traps, since forward current causes no injection of minority carriers. However, these cases can still be handled by a combination of DLTS and photocapacitance. One uses a majority-carrier pulse while at the same time illuminating the sample with an appropriate wavelength of light from a monochrometer. In this case, a negative transient due to optically filling the trap between pulses is observed for a minority-carrier trap and the rate constant is the sum of the thermal emission rate e_1 and the net optical rate for filling the trap. When e_1 is much larger than the optical rate, the trap can no longer be filled between pulses and the transient disappears. At this temperature one sees a positive step in the DLTS baseline similar to the sort of signal observed in TSCAP. The size of the transient at temperature below the step depends on the concentration of the trap as well as the intensity and wavelength of the light.

From this comparison with the best of the previous capacitance techniques it is clear that DLTS is a significant improvement in terms of sensitivity, trap depth range, and ease of operation and analysis. The only large class of traps which DLTS cannot observe are the shallow centers such as donors, acceptors, and bound excitons. This is by no means a drawback, however, since these shallow centers can easily be detected by luminescence. What DLTS in effect does is to extend trap spectroscopy into the regime of the deep and intermediate traps missed by luminescence, especially the nonradiative centers about which there is almost a complete lack of understanding.

VI. SUMMARY

The DLTS technique is introduced and analyzed. The basic underpinnings of the method lie in the concepts of p-n junction capacitance transients due to the filling and emptying of traps at fixed temperatures. These concepts are briefly discussed along with the pulsed bias method of defining the initial conditions for a transient. The DLTS method is then explained by showing in detail the results of using a dual-gated signal averager to extract information from a repetitive capacitance transient during a temperature scan. Thermal emission rate measurements of traps in GaAs made by DLTS are compared with more accurate and tedious measurements made at a series of fixed temperatures. The various modes of operation are described and illustrated which enables one to use DLTS to determine activation energies, trap concentrations, concentration profiles, and carrier capture cross sections for both majority and minority carriers. The problem of obtaining this information from nonexponential capacitance transients is discussed.

In the final section the DLTS method is compared with other capacitance techniques, namely, thermally stimulated capacitance (TSCAP), edge region TSCAP, admittance spectroscopy, and photocapacitance. The results of this comparison show that DLTS has the most sensitivity, the widest range of trap depths, and is the most convenient to use and analyze of these methods. The development of DLTS has brought the capacitance method to the stage where for deep and intermediate traps it is in the same class as luminescence is for the shallow traps. That is to say, DLTS has the sensitivity, the spectroscopic nature, and the speed and ease of analysis to make it a practical means of doing spectroscopy on a large number of traps in a large number of samples.

ACKNOWLEDGMENTS

The author is indebted to R.G. Logan and H.G. White for providing the liquid-phase epitaxy samples and to A.Y. Cho for the molecular beam epitaxy material. Thanks are due A.J. Williams for diode fabrication and mounting. The electron irradiation was performed in cooperation with L.C. Kimmerling with whom several helpful discussions were also held. Throughout the course of this work, the author has had numerous discussions and helpful comments from C.H. Henry for which he is especially grateful.

[1]A.A. Bergh and P.J. Dean, Proc. IEEE 60, 156 (1972).
[2]C.T. Sah and J.W. Walker, Appl. Phys. Lett. 22, 384 (1973).
[3]J.C. Carballes, J. Varon, and T. Ceva, Solid State Commun. 9, 1627 (1971); C.T. Sah, W.W. Chan, H.S. Fu, and J.W. Walker, Appl. Phys. Lett. 20, 193 (1972); M.G. Buehler, Solid-State Electron. 15, 69 (1972).
[4]D.L. Losee, Appl. Phys. Lett. 21, 54 (1972).
[5]The capacitance transient technique was first used by R. Williams, J. Appl. Phys. 37, 3411 (1966).
[6]A comprehensive treatment of capacitance and current transients is given by C.T. Sah, L. Forbes, L.L. Rosier, and A.F. Tasch, Jr., Solid-State Electron. 13, 759 (1970).
[7]D.V. Lang, preceding paper, J. Appl. Phys. 45, 3014 (1974).
[8]L.D. Yau and C.T. Sah, Phys. Status Solidi A 6, 561 (1971).
[9]C.H. Henry, H. Kukimoto, G.L. Miller, and F.R. Merritt, Phys. Rev. B 7, 2499 (1973).
[10]H. Kukimoto, C.H. Henry, and F.R. Merritt, Phys. Rev. B 7, 2486 (1973).
[11]This is referred to as a clear pulse in Ref. 7.
[12]Princeton Applied Research model 162 with model 164 gate modules.
[13]A.Y. Cho, J. Vac. Sci. Technol. 8, S31 (1971).