# A correlation method for semiconductor transient signal measurements

G. L. Miller, J. V. Ramirez, and D. A. H. Robinson

Citation: Journal of Applied Physics **46**, 2638 (1975); doi: 10.1063/1.321896
View online: http://dx.doi.org/10.1063/1.321896
View Table of Contents: http://scitation.aip.org/content/aip/journal/jap/46/6?ver=pdfcov
Published by the AIP Publishing



# A correlation method for semiconductor transient signal measurements

G. L. Miller, J. V. Ramirez, and D. A. H. Robinson

*Bell Laboratories, Murray Hill, New Jersey 07974*
(Received 14 November 1974; in final form 20 December 1974)

A wide variety of experiments in semiconductor physics involve the observation of a particular class of time-dependent transient signals; namely, decaying exponentials. Such signals are characteristic of systems relaxing back to equilibrium following an abrupt change in the populations of carriers occupying the available states. The magnitudes of such signals indicate the density of states involved, while the decay time constants (as a function of temperature) provide information on the energy levels of the states in question. As such, both the amplitudes and the time constants of such signals are of physical interest. The problem of processing such signals has been examined using the theory of linear filtering and as a result a special purpose signal processor has been developed. This processor performs a continuous real-time cross correlation between the experimental signals and an appropriately synchronized locally generated exponential waveform. The resulting system has been used to process signals from transient space-charge measurements on reverse-biased gold-doped silicon junction diodes, providing both correlation spectra that yield energy levels and also trap profiles that expose the spatial distribution of the gold inside the junctions.

PACS numbers: 81.20.M, 85.30.K, 71.55.

## I. INTRODUCTION

In principle, an extremely large number of different kinds of semiconductor transient experiments can be carried out, using optical, electrical, or thermal stimulation and exhibiting relaxation time constants ranging from nanoseconds to hours.[1]

Although the possible variations in such experiments are virtually unlimited,[2] many of them exhibit a common feature; namely, the existence of exponential transients as the physical system approaches a steady state following the application of an abrupt disturbance.

A particularly clear example is illustrated in Fig. 1. This depicts the motion of the depletion layer in a reverse-biased $p^+$-$n$ junction as the bias voltage is switched between two values. At all times the depletion layer thickness $x(t)$ is set by the requirement that the electric field at the back edge of the space-charge region is zero, i.e. at that location the electric field produced by the bias voltage is exactly balanced by the total space charge in the depletion layer. (The crossed circles indicate bound ionized donors, while the square boxes denote electron traps, which may be either filled or empty.)

When the bias voltage is reduced from its quiescent value the traps fill and some of them remain filled when the bias is abruptly returned to its initial value. Now, however, the depletion layer is slightly wider than it was initially since some of the positive charge due to bound ionized donors in the depletion layer is cancelled by the negative charge arising from the trapped electrons.

These trapped electrons have a certain probability of being thermally excited from the traps, in which case they are rapidly swept out of the depletion layer by the electric field, leading to an increase in the net positive charge in the depletion layer and thereby producing a corresponding reduction in the space-charge-region width. This results in an exponential contraction of the space-charge region with a time constant $\tau_D$ given by

$$\tau_D = (g/\gamma N_c)\exp(E_T/kT). \tag{1.1}$$

Here the process is assumed to be due to a single electron trapping level, distance $E_T$ below the conduction band edge, while $k$ is the Boltzmann constant, $T$ the absolute temperature, $g$ the multiplicity of the level, $\gamma$ the capture coefficient, and $N_c$ is the density of states at the conduction band edge.

It is clear in this case that the physical information of interest is contained in the decaying exponential part of $x(t)$ and not in the rectangular part which merely reflects the particular process used to change the trap population. As such, one is led to study the question of how best to filter such decaying exponential signals in order to extract amplitude and time constant information.

It is, of course, understood that while the above particular example provides such exponential signals it is by no means a specialized situation. Similar signals can be expected to arise in general from the relaxation of

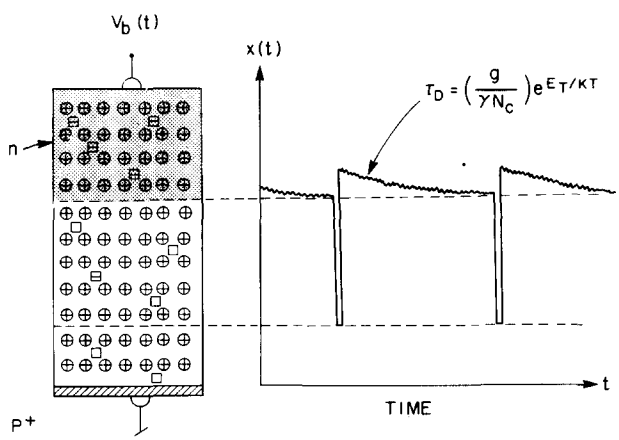



FIG. 1. Depletion layer thickness $x(t)$ in a $p^+$-$n$ junction as a function of time as the reverse-bias voltage is switched between two values.

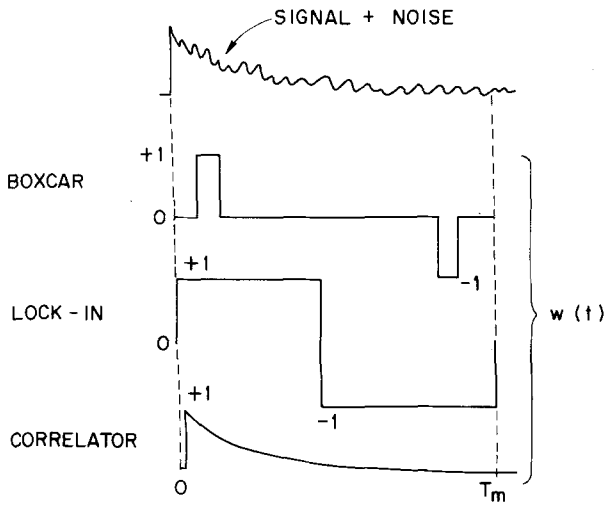



FIG. 2. A representative decaying exponential signal (upper waveform) together with three weighting functions $w(t)$.

any physical system governed by the Boltzmann statistics of a single level.

## II. FILTERING

It is readily established that any linear filtering operation is equivalent to an integral transformation. Specifically, the action of a linear filter acting on an input signal $e_1(t)$ can be characterized by a "weighting function" $w(t)$ together with the relationship

$$e_0(T_m) = \int_0^{T_m} e_1(t) w(t)\, dt. \tag{2.1}$$

Here $e_0(T_m)$ is the output of the filter at the measurement time $T_m$.

Therefore, given a signal of unknown amplitude $A$ but of known shape $s(t)$, in the presence of additive white Gaussian noise $n(t)$, one can seek the optimum form of $w(t)$ in order to obtain the best estimate $E$ of the signal amplitude;

$$E = \int_0^{T_m} [As(t) + n(t)] w(t)\, dt. \tag{2.2}$$

The solution to this problem is known[3] and is given by

$$w(t) \equiv s(t) \tag{2.3}$$

i.e. *the optimum weighting function has the form of the noise-free signal itself, and therefore for the systems of interest here $w(t)$ should be a decaying exponential.*

The significance of this result can be illuminated by considering some other filtering operations that have been applied to such systems, as indicated in Fig. 2. Here the upper waveform represents the typical signal plus noise, while the lower three waveforms comprise different weighting operations. The first of these results from the use of a double boxcar integrator, as introduced by Lang,[4] the second represents the effect of the use of a lock-in amplifier as employed by Kimerling,[5] while the lowest represents the optimum weighting process as prescribed by Eq. (2.3).

It is clear that the utilization of the signal information is quite different in these three cases and it is, therefore, not surprising that the resulting $S/N$ ratio is also different. In fact, calculation shows that the lock-in $S/N$ ratio is theoretically a factor of approximately 2 times worse than that of the exponential correlator, while the boxcar $S/N$ is worse than either (but obviously approaches that of the lock-in monotonically as the width of the boxcar samples approaches $\frac{1}{2}T_m$).

Other performance features are also different; for example, the exact form of the output line shapes and the ability to make a direct determination of the signal decay time constants. Another factor of great practical importance is the behavior of the system during the trap-loading process which generally involves substantial overloading of the rf bridge, or other measuring device, used in the experiment. It is very important that spurious signals generated during this time not be allowed to distort the signal information. Some of these more detailed questions are pursued in Sec. VII.

## III. A CORRELATION SPECTROMETER

There are a number of possible ways of implementing the prescription of Eqs. (2.2) and (2.3). One simple method is indicated in Fig. 3. An RC function generator produces a repetitive decaying exponential, together with an output trigger signal generated at the end of each rundown period. This trigger signal drives the physical system, thereby disturbing the steady state. The resulting experimental exponential signal returns to the correlator and is multiplied by $w(t)$ in an analog multiplier and integrated to form the output signal. Since $w(t)$ has been defined to be unipolar it is convenient to restore the baseline of the input signal through the action of $C_1$, $R_1$, and the synchronized switch $S_1$ which is closed momentarily at the end of each rundown period. This defines the zero level of the signal into the multiplier.

A qualitative view of the signal formation process can be gained by considering input signals of monotonically increasing time constant, 1–6, as indicated in Fig. 4. Each curve is multiplied point by point by $w(t)$ and added. The result is a maximum, for suitably chosen $T_m$, when the input signal time constant $T_s$ equals the reference signal time constant $T_R$. In fact, the normalized correlator output line shape is given by

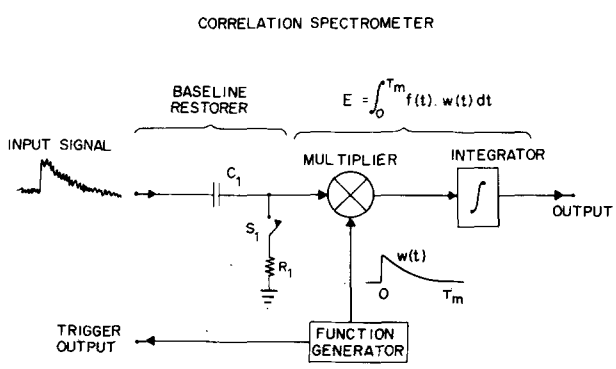



FIG. 3. Block diagram of a correlation spectrometer.

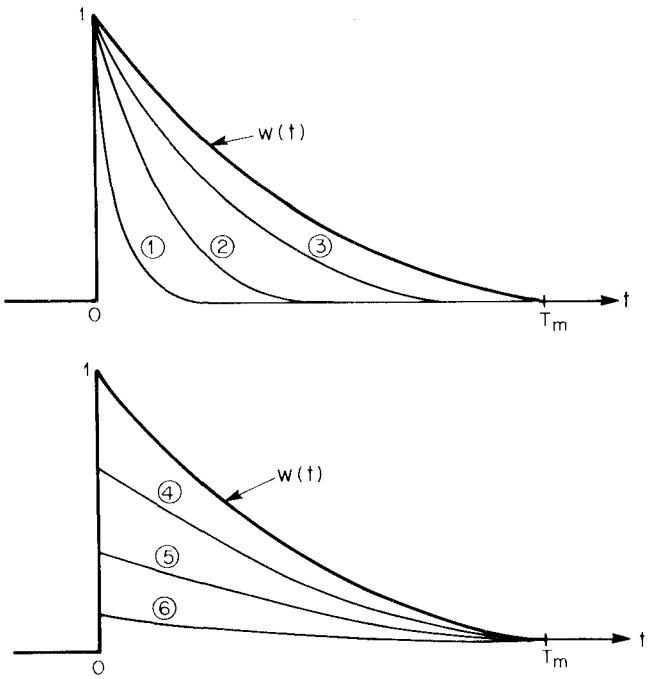



FIG. 4. Multiplier input waveforms (1—6) corresponding to monotonically increasing signal decay time constants. The input signal 1 is shown initially normalized to unit amplitude. The signal zero level is defined by baseline restoration at time $T_m$.

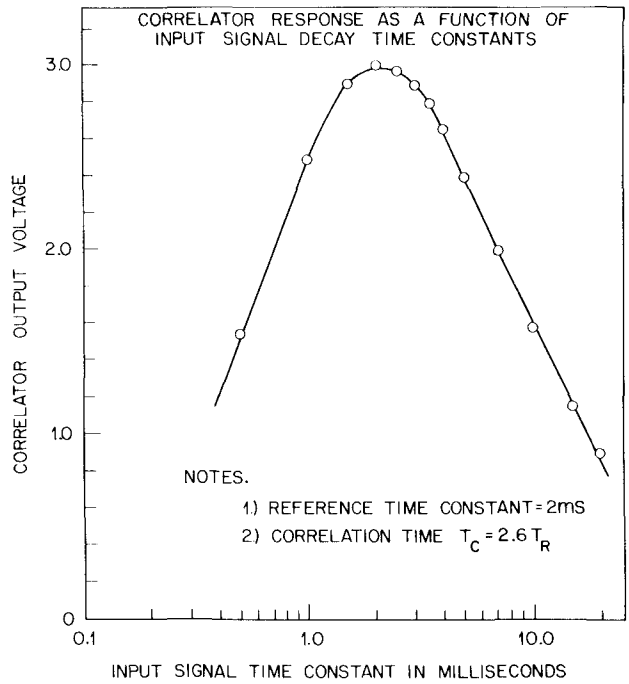



FIG. 5. Correlator response as a function of input signal decay time constants.

$$A = \left\{\left(\frac{T_R^2}{T_M(T_R+T_S)}\right)\exp\left(-\frac{T_M(T_R+T_S)}{T_R T_S}\right)\right\} - \left\{\left(\frac{T_R}{T_M}\right)\exp\left(-\frac{T_M}{T_S}\right)\right\} + \frac{T_R T_S}{T_R(T_R+T_S)}. \qquad (3.1)$$

An experimental plot of the correlator response is shown in Fig. 5. A significant point is that the output is a maximum when the input signal decay time constant equals the reference waveform time constant. This makes the determination of system detrapping time constants particularly simple. Another feature that is noteworthy is that the existence of closed-form expressions such as Eq. (3.1) for the line shape indicates the possibility of investigating *non*exponential decays from line-shape distortion since this precise result holds if and only if the input signal consists of a single

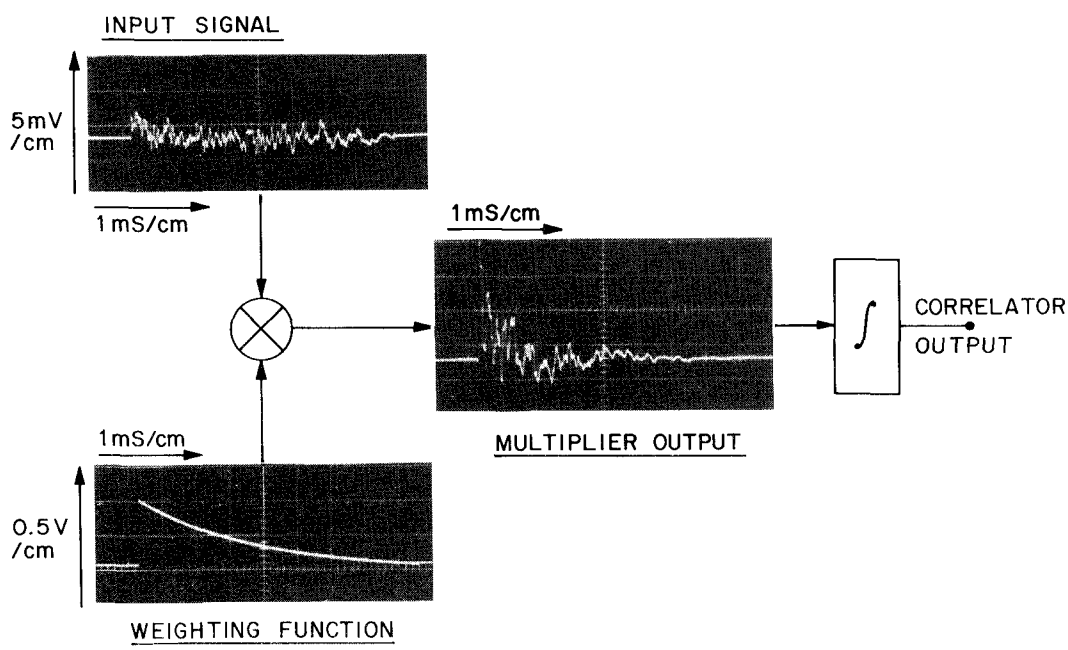



FIG. 6. Representative correlator waveforms which indicate the origin of the $S/N$ ratio enhancement.

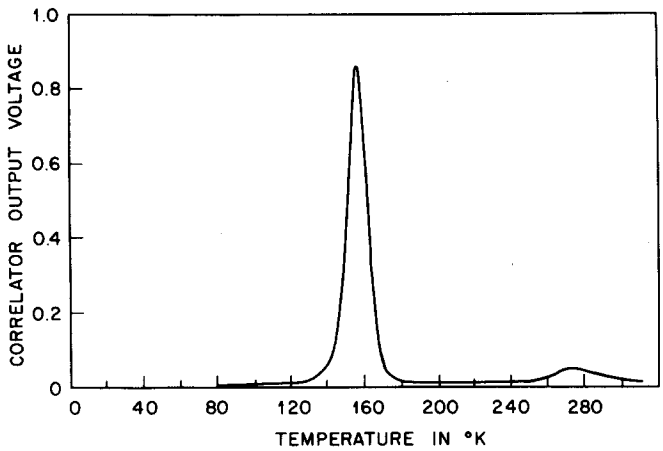



FIG. 7. Correlation spectrum of a gold-doped silicon $p$–$n$ junction. The peaks at 155 and 270 °K correspond to the 0.33- and 0.57-eV levels of gold, respectively.

exponential decay. Any other input signal, for example, one comprising the sum of two exponentials of differing time constant, cannot give rise to the same line shape as that of Eq. (3.1). Consequently, the existence of differing line shapes can be used as an indication of the fact that the signals of interest do not correspond to single exponential processes.

## IV. EXPERIMENTAL RESULTS

The physical meaning of the cross-correlation process can be illuminated with the help of the experimental waveforms shown in Fig. 6. Initially one would be hard put to it to say if there really was a decaying exponential signal buried in the noise in the input signal. However, after multiplication by $w(t)$ it is clear that the majority of the time, during the ~8-ms measurement period, the multiplier output is positive, i.e., a signal is indeed present.

The reason for the improved $S/N$ ratio is simply that one has to pay a noise penalty at each instant only proportional to the magnitude that one expects the signal amplitude to attain at that same time.

As an example of the application of this system, an experiment of precisely the type of Fig. 1 was carried out on a gold-doped silicon $p$-$n$ junction. The depletion layer was cycled through a distance of ~5000 Å while the mean depletion layer depth was held at 17 $\mu$. On feeding the $x(t)$ signal to the correlator, setting the reference time constant to 1.4 ms and sweeping the diode temperature from ~LN to room temperature the plot of Fig. 7 was obtained. The peak at ~150 °K is due to the 0.33-eV level in gold, while the peak at ~270 °K is due to the 0.57-eV level.

An important point in measurements of this type is to ensure that the total depletion layer thickness remains essentially constant during the entire measurement. This is exemplified by the doping profiles shown in Fig. 8 for the diode employed. Each profile shows the range of depletion depths available as the reverse-bias voltage is swept from 0 to 100 V. It is clear that only a distance of $x$~17 $\mu$ can be chosen to stay within the system operating range over the entire temperature span of 77–296 °K.

Determination of trap levels from detrapping time constants is provided by setting the temperature for convenient values of $1/T$ and then slowly sweeping $T_R$ until the correlator output is a maximum. A typical example, for a different type of diode, is shown in Fig. 9. This result was obtained in ~20 min, this time being dominated by the temperature equilibration time at each point.

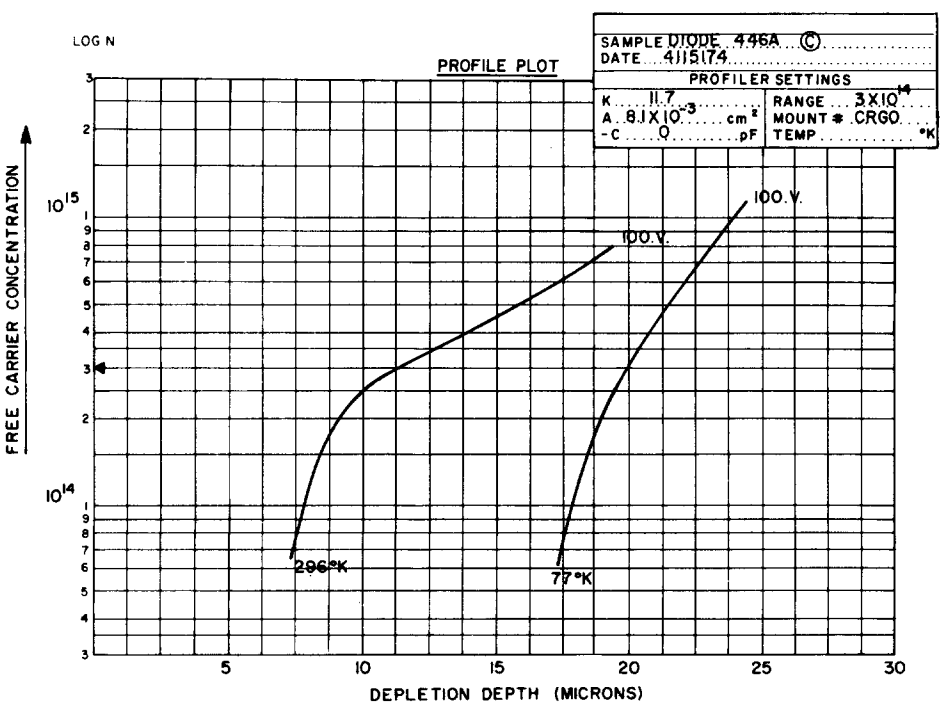



FIG. 8. Profiles of a gold-doped silicon junction taken at two different temperatures. In each case the bias voltage range was 0–100 V.

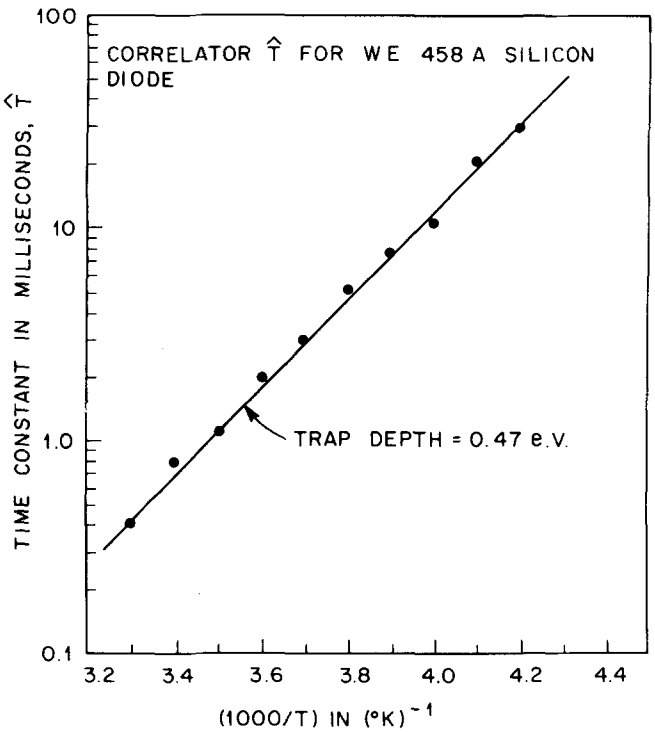



FIG. 9. Representative activation energy plot obtained by adjusting the correlator for maximum output at each temperature.

## V. PROFILING TRAP DISTRIBUTIONS

A variant of the correlator technique involves its operation in synchronism with a feedback profiler.[6] The basis of the feedback profiling method is recapitulated in the upper half of Fig. 10 for a junction exhibiting no trapping phenomena. Electric field changes of constant magnitude (i.e., independent of $x$) are applied to the junction region. These field changes are automatically balanced by the repetitive covering and uncovering of equal quantities of charge $\Delta Q$ at the back edge of the junction depletion layer. Exposing this additional quantity of charge demands a corresponding motion $\Delta x$ in the depletion layer boundary and the reciprocal of the magnitude of this motion then indicates the doping density.

In the presence of traps the situation is modified as shown schematically in the lower half of Fig. 10. This state of affairs is obviously closely related to the regime of Fig. 1, differing in that the motion is relatively smaller and takes place with a 50% duty cycle.

Experimental results, illustrating such effects, are shown in Fig. 11 which depicts the motion of the back edge of the depletion layer as a function of time for four different gold-doped junctions. The horizontal time scale is 1 ms/cm, while the ordinate is ~300 Å/cm for waveforms 1—3 and twice that value for 4. The four $p^+$-$n$ planar oxide masked silicon junctions were initially identical and, subsequently, had gold diffused into them at the temperatures indicated on the profiles. It is apparent that this process resulted in different quantities of gold in solution. This is indicated both by the different free-carrier densities and by the different transient behavior of $\Delta x$.

Qualitatively similar results have been obtained with all the gold-doped silicon junctions so far examined. This has been further investigated with the correlator by synchronizing $w(t)$ with the $\Delta x$ motion imposed by the profiler, and displaying the correlator output as a function of distance on the profile plot as indicated by the result of Fig. 12. Since the trap profile was obtained by operating at the peak of the 1.4-ms correlation spectrum peak at ~270 °K, which corresponds to a gold level, it apparently indicates that the trap profile is itself the gold profile. (The actual interpretation of such results is more complicated, but this is illustrative of the method.)

## VI. NONEXPONENTIAL DECAYS

While exponential correlation is useful for purely exponential signals it is also of interest to consider other possible inputs. An obvious candidate would be a mixed exponential containing two somewhat different time constants. One approach to this problem is that of delaying the turn on of $w(t)$ until the faster signal component has essentially died away. (This is possible with either the boxcar or the correlator method, but is not a possibility with commerical lock-in amplifiers.)

A different, and more radical, attack on the problem is that of adopting an alternative filtering criterion, namely, that $w(t)$ be chosen so that the correlation spectrum yields the *narrowest lines* rather than the best $S/N$ ratio. While this approach is clearly a possibility, it necessarily involves a degradation of the $S/N$ ratio, raising the question of the exact price that must be paid for such line narrowing. This has been investigated by introducing a figure of merit $M$, defined as the $S/N$ ratio at the maximum of a spectral peak divided by the FWHM (full width at half-maximum) of that peak. Examination of a correlation spectrum such as that of Fig. 7 indicates the significance of such a parameter. Obviously, narrower lines should provide correspondingly greater spectral detail, but this would only remain true if the $S/N$ ratio were not itself degraded in the process.

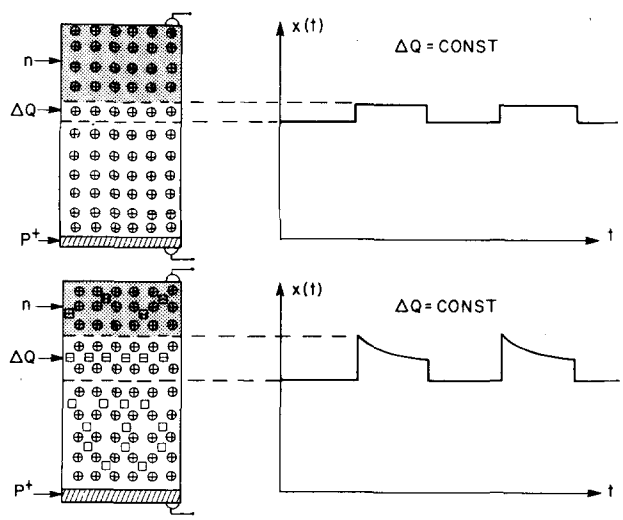



FIG. 10. Schematic representation for feedback profiling action with and without the presence of deep traps.

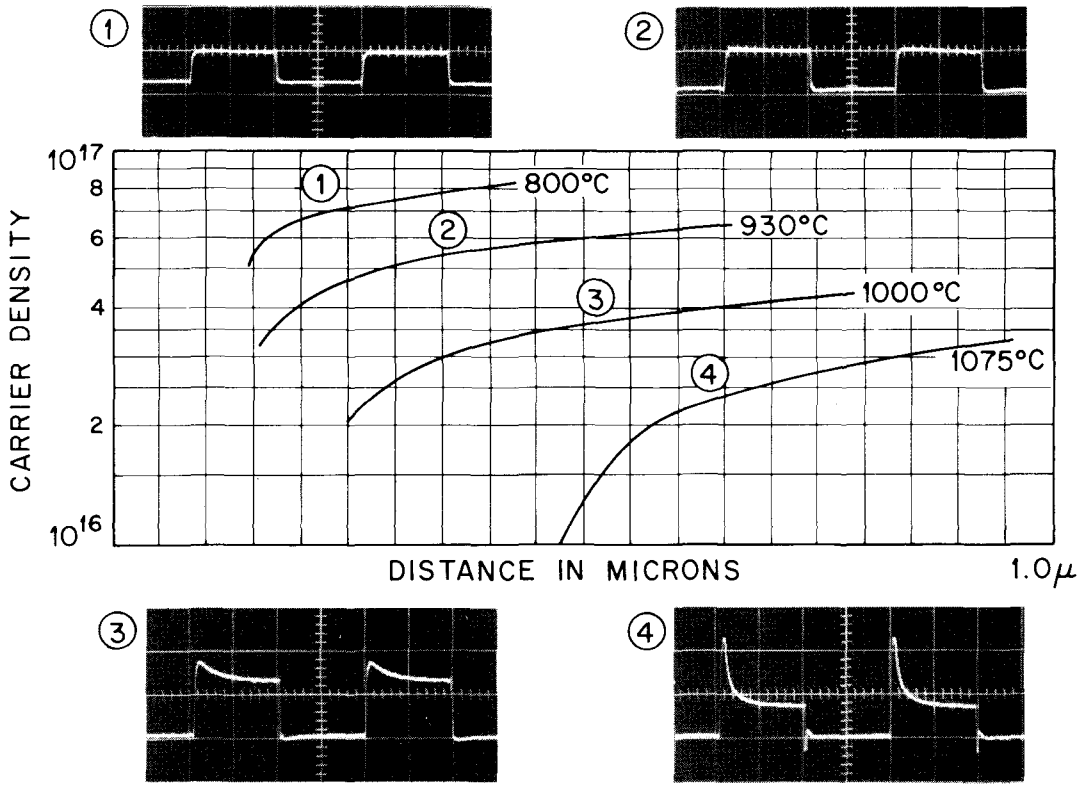



FIG. 11. Profiles and $\Delta v$ waveforms of four different gold- doped silicon junctions.

Use of the parameter $M$ has provided a convenient method for evaluating the potential performance of different weighting functions, as indicated in Sec. VII.

## VII. A COMPARISON OF WEIGHTING OPERATIONS

It has already been pointed out in Eq. (3.1) that the exponential correlator output is a function of the three time constants $T_R$, $T_M$, and $T_S$. Fixing the ratio of $T_R$ to $T_M$, as has been done in Fig. 5, reduces the line shape to a function of two parameters. In addition, for comparison purposes, a convenient choice is to set $T_R = T_M$ and then to introduce the dimensionless ratio $\alpha \equiv T_S/T_M$, as has been done in the first entry in Fig. 13. The function $A(\alpha)$ can then be plotted as a function of $\alpha$ and the amplitude and FWHM of the resulting line determined.

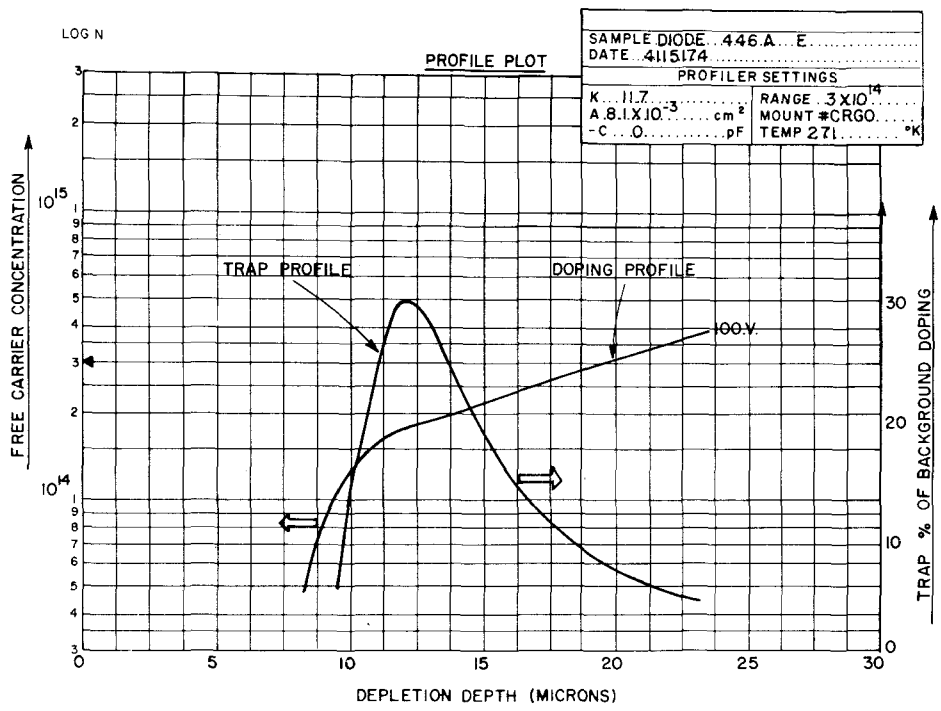



FIG. 12. Superimposed trap and doping profiles in a gold-doped silicon junction.

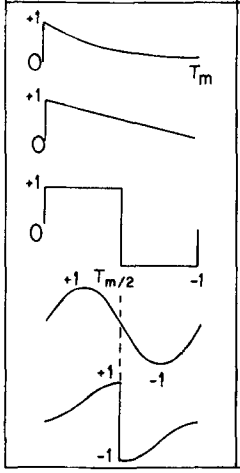
| WEIGHTING FUNCTION | OUTPUT LINE SHAPE | FIGURE OF MERIT |
|---|---|---|
|  | $A(\alpha) = \frac{\alpha+e^{-(\alpha+1)/\alpha}}{\alpha+1} - e^{-1/\alpha}$ | 1.00 |
| | $A(\alpha) = \alpha^2 + \alpha - (\alpha^2+2\alpha+3/2)e^{-1/\alpha}$ | .81 |
| | $A(\alpha) = \alpha(1-e^{-1/2\alpha})^2$ | .55 |
| | $A(\alpha) = \frac{2\pi\alpha^2(1-e^{-1/\alpha})}{4\pi^2\alpha^2+1}$ | .55 |
| | $A(\alpha) = \frac{\alpha}{2}\left[(1-e^{-1/2\alpha})^2 - \frac{(1+e^{-1/2\alpha})^2}{4\pi^2\alpha^2+1}\right]$ | .30 |

FIG. 13. A display of different weighting functions together with their output line shapes and figures of merit.

Similarly, it is a simple matter to integrate the square of the weighting function over the range 0 to $T_M$, thereby determining the corresponding noise contribution. This provides all the information required to compute the figure of merit $M$, and all such figures have been normalized to unity for the exponential weighting operation.

A simple alternative to the exponential would be a linear ramp, and this results in a degradation in performance of ~20% as indicated by the second entry in Fig. 13. The significance of this result is that since sawtooth signal generators are widely commercially available, while exponential generators are not, the sawtooth represents a simple compromise for the function generator of Fig. 3, while only exacting a relatively small performance penality.

Another possibility is Kimerling's method, involving the use of a lock-in amplifier suitably synchronized to the signals generating the disturbance in the sample.[5] This regime corresponds to the bipolar rectangular weighting function on the third line of Fig. 13. This approach is very attractive in the sense that it needs no special equipment since lock-in amplifiers are widely available, although it results in a modest performance penality.

A more serious limitation is that the lock-in is acutely sensitive to the bridge or loading transients (the "rectangular" part of the signal of Fig. 1) since the weighting function has unit magnitude at all times. This sensitivity can be substantially reduced by preceding the weighting by a narror-band filter (as is usually the case in commercial lock-in amplifiers), leading to an effectively sinusoidal over-all weighting function as indicated by the fourth entry in Fig. 13. While the loading transient can be arranged to straddle a zero crossing of this function, the finite slope at such a point still gives rise to a corresponding sensitivity of the output signal to any drifts in the experimental timing conditions.

Ideally one would employ a weighting operation that exhibited both a zero value *and* a zero slope during the loading transient. This is readily achieved both in the case of the exponential correlator and the boxcar method by introducing an adjustable delay between the trigger pulse and the turn on of $w(t)$. While this resource is not available in conventional lock-in amplifiers, a possible modification that has been examined is the split sine-wave function shown at the bottom of Fig. 13. This automatically achieves the desired zero value plus zero slope condition and is quite easy to implement, but it results in a threefold reduction in the over-all system performance.

From this it appears that attempts at line narrowing by the choice of nonexponential weighting operations do not represent attractive possibilities, except in situations in which the $S/N$ ratio is already so large that its substantial degradation can be tolerated without significant loss of spectral detail.

## ACKNOWLEDGMENTS

The authors are indebted to W.L. Brown, L.C. Kimerling, R.A. Boie, and L.D. Yau of Bell Laboratories for useful discussions and particularly to D.V. Lang for detailed intercomparisons of experimental results obtained with the double boxcar and the correlation method.

[1]A.G. Milnes, *Deep Impurities in Semiconductors* (Wiley, New York, 1972), p. 178ff.
[2]C.T. Sah, L.I. Rosier, and A.F. Tasch, Jr., Solid-State Electron. 13, 759 (1970).
[3]D. Middleton, *An Introduction to Statistical Communication Theory* (McGraw Hill, New York, 1960).
[4]D.V. Lang, J. Appl. Phys. 45, 3023 (1974).
[5]L.C. Kimerling (private communication).
[6]G.L. Miller, IEEE Proc. Electron Devices **ED-19**, 1103 (1972).