Electron-spin-resonance-transient spectroscopy

W. B. Jackson, M. Stutzmann, and C. C. Tsai
Xerox Corporation, Palo Alto Research Center, Palo Alto, California 94304
(Received 10 June 1985; revised manuscript received 28 January 1986)

This paper describes a method for systematically obtaining distributions of annealing and/or creation energies, tunneling distances, and trap depths for specific defect populations from electron-spin-resonance (ESR) transients. The method of analysis, specifically adapted to the constraints imposed by ESR, is presented for obtaining defect distributions from a few transients using time-dependent temperature variations. Specific important cases, numerical examples, and application to actual data are examined. A brief discussion of the experimental methods required to obtain the ESR transients with sufficient sensitivity are described. Finally, issues of resolution and noise, comparison to other transient analysis methods, and possible extensions are presented. A detailed discussion of the application of the method to hydrogenated amorphous silicon is presented in the following paper.

I. INTRODUCTION

In a number of materials, the annealing and creation kinetics of defects are of great scientific and technological interest because the energetics of these processes determines the defect density and the stability of these materials. In particular, materials such as *a*-Si:H, GaAs, and ion-bombarded silicon are important examples where metastable defects can be removed by annealing at increased temperatures. The energies controlling the annealing and creation of these defects are invaluable in elucidating the microscopic defect structures and dynamics. Other materials such as silicon dioxide and silicon nitride contain defects which can be emptied or filled by the variation of temperature and/or the application of fields. The distribution of these states and traps as a function of physical parameters such as distance from the interface or as a function of energy is of interest.

Unfortunately, many measurements which depend on defect concentration, such as carrier lifetime or mobility, do not distinguish between the various defects. Hence, it is extremely valuable to have a method by which the evolution of a specific type of defect can be studied independent of the others. Furthermore, it would be particularly useful to be able to relate the microscopic structure of the defect to its energetics. Since electron spin resonance (ESR) in many cases yields direct information on the microscopic structure of the defects and is capable of distinguishing between various types of defects, one would like to combine ESR with another method to obtain additional information.

One recent method which relates ESR to an energy spectroscopy is spin-dependent deep-level transient spectroscopy (DLTS).[1] In this experiment, the capture and emission rate of defects in a magnetic field is altered by a microwave field. The energy information is derived from the emission and capture rates while the microscopic defect information is derived from the spin-resonance condition. Unfortunately, this measurement is very insensitive, is difficult to interpret, and many defects are not expected to have spin-dependent emission or capture rates. Furthermore, DLTS also requires that the traps be repetitively filled and emptied, and thus it is not appropriate for cases where paramagnetic defects are irreversibly destroyed.

In order to circumvent these problems, we describe an experimental technique which combines ESR with an energy-sensitive spectroscopy. The basis of this method is that the rate of change of a given metastable defect population is observed by ESR as a function of time and temperature. Because ESR is selective to a specific type of defect, the change of defect populations with different spin-resonance transitions can be distinguished. Depending on the cause of the various rates, different types of information about the defect population can be extracted from the observed ESR decays. For example, if the rate distribution occurs because the defects have different microscopic environments, the distribution of defect annealing energies may be determined from the different rates, or if the rate distribution is due to thermal excitation to a band edge, the density of states of the particular defect relative to the band edge can be obtained. Finally, if the transient ESR signal is due to tunneling, the various tunneling distances may be derived from the rate distribution. Consequently, an otherwise uniform spin population can be further resolved into subpopulations with different rate constants.

Although the analysis of the ESR transients in terms of various rates appears straightforward, and periodically, Arrhenius analyses are performed on ESR signals,[2] to the best of our knowledge, it has not been systematically applied to spectroscopically resolve defect distributions. The spectroscopic analysis of electrical transients, while commonly used in capacitance transient spectroscopy,[3] thermally stimulated currents,[4] and luminescence measurements,[5] has only recently been applied to other types of measurements to yield energy distributions. Kostial and Slabeycius have applied a special case of the rate analysis to ultrasonic attenuation measurements in amorphous Se to determine the distribution of energies characterizing structural changes in amorphous Se.[6] The at-

tenuation measurements are not specific to a particular type of defect, however. The reason for the lack of work on ESR transients is due primarily to several experimental considerations particular to ESR. First, the signal-to-noise ratio for ESR is significantly less than luminescence or electrical measurements. Second, ESR spin densities are measured typically by scanning the magnetic field so the time interval between successive spin measurements is often greater than 100 s. In addition, one cannot usually repetitively fill and empty the states. These considerations require that unique experimental methods be used to acquire the ESR decays which will be discussed in this paper and that the data analysis be particularly efficient for obtaining information from a few annealing cycles.

This paper begins with a general discussion of the physical basis and analysis of ESR transient spectroscopy in Sec. II. The general analysis is applied to a number of important special cases. In Sec. III we discuss general experimental methods for obtaining the required ESR transients with sufficient sensitivity. This is followed by a discussion of a number of issues concerning the resolution, signal-to-noise ratio, comparison of the analysis to methods used in other fields, and applicability of the spectroscopy in Sec. IV. This section also proposes possible extensions of the particular analysis developed in this paper. The following paper[7] applies the methods developed in the first paper to the specific case of annealing distributions of metastable dangling bonds in hydrogenated amorphous silicon.

II. ANALYSIS

The framework for extracting energy distributions from ESR transients is presented below. Although in simple cases, the analysis is similar to that used in other areas such as luminescence decays,[5] we have found that a more general method, particularly suited for extracting energy distributions from ESR transients, yields superior results. Finally, some important special cases are discussed.

A. General theory

Consider a solid in which a metastable population of spins has been created either by the formation of new defects or by the population of an existing defect with an unpaired spin. The metastable spins will, in general, have a distribution of some parameter, S. For example, the spins may have a distribution of energies or attempt-to-escape frequencies. Upon annealing and/or passage of time, the spin population will decay according to the relation

$$\frac{dN(S,t)}{dt} = -R(S,t)N(S,t) , \tag{1}$$

where $N(S,t)$ is the density of spins (spins per volume per unit S) at a time t with a rate of decay, $R(S,t)$, characterized by the parameter S. A specific example might be

$$R(S,t) = \nu_0 \exp[-E_a/kT(t)] ,$$

where the distribution parameter S is the activation energy E_a, ν_0 is an attempt frequency, and $kT(t)$ is the time-dependent temperature in energy units. Equation (1) may be solved yielding

$$N(S,t) = N(S,0)\exp\left[-\int_0^t R(S,t')dt'\right] , \tag{2}$$

where $N(S,0)$ is the initial spin distribution that must be determined. The experimentally accessible quantity is the total spin population given by the integral

$$\begin{aligned} N(t) &= \int dS\, N(S,t) \\ &= \int dS\, N(S,0)\exp\left[-\int_0^t R(S,t')dt'\right] . \end{aligned} \tag{3}$$

The purpose of the subsequent analysis is to extract the distribution $N(S,0)$ from the time-varying spin population $N(t)$.

There are several approaches one may take to derive this distribution from $N(t)$. If the rate, $R(S,t)=R(S)$, is independent of time, Eq. (3) can be simplified by a change of variable giving

$$\begin{aligned} N(t) &= \int dS\, N(S,0)\exp[-R(S)t] \\ &= \int_{R(0)}^{R(\infty)} du\, N(u,0)\exp(-ut)/R'(R^{-1}(u)) , \end{aligned} \tag{4}$$

where $u = R(S)$ and $S = R^{-1}(u)$. If the limits extend over a sufficiently large range, this integral can be approximated by the Laplace transform. The solution to Eq. (4) is

$$N(u,0) \simeq R'(R^{-1}(u))L^{-1}(N(t)) , \tag{5}$$

where L^{-1} denotes the inverse Laplace transform. One could therefore obtain the distribution using a numerical inverse Laplace transform. Alternatively, one could, in principle, apply various analysis methods developed for decay of luminescence such as the method of moments[8] or nonlinear least-squares fitting.[5] However, there are a number of problems with these approaches which will be discussed in Sec. III C.

With these difficulties in mind, we derive a iterative method which will extract the distribution from Eq. (3) with very few function evaluations (sometimes as few as two), constraints are easily incorporated, and the tradeoff between noise rejection and resolution is quite readily implemented. Moreover, the method does not assume a distribution *a priori*. More-detailed discussions of the merits of iterative deconvolutions over other methods are discussed in Ref. 9.

The analysis is based on the following observation. All of the time dependence is contained in the exponential factor in Eq. (3). If $\int_0^t R(S,t')$ is less than one, the integrand is approximately $N(S,0)$, while if $\int_0^t R(S,t')dt'$ is greater than one, the integrand is nearly zero. Consequently, Eq. (3) may be approximated by

$$N(t) \simeq \int_{S'(t)}^{\infty} dS\, N(S,0) , \tag{6}$$

where $S'(t)$ is defined implicitly by the equation

$$\int_0^t R(S',t')dt' = 1 . \tag{7}$$

In deriving Eqs. (6) and (7), we have assumed that $\partial R/\partial S' < 0$ and that Eq. (7) yields a single value of S' for each t. If $\partial R/\partial S' > 0$, the limits of integration in Eq. (6)

become 0 and $S'(t)$ and the minus sign in Eq. (9) becomes positive. In all cases of interest, these assumptions are justified. Taking the time derivative of Eq. (6) yields

$$\frac{dN}{dt} \simeq -N(S',0)dS'(t)/dt \tag{8}$$

or

$$\frac{dN}{dS'} \simeq -N(S',0) \ . \tag{9}$$

This important result states that if the spin density is plotted versus $S'(t)$ defined by Eq. (7), the negative slope of this curve approximately gives the distribution of the spins in the parameter S'. Although Eq. (9) is an approximation, it is remarkably accurate and often is sufficiently accurate by itself. The first example in Sec. II D demonstrates that Eq. (9) merely broadens the spectrum by a function of width $2.7kT$.

Equation (9) forms the basis for an iterative method in which one computes a distribution using Eq. (9). Using this distribution, a new transient is computed, and Eq. (9) is applied to the difference between this transient and the observed one. In other words,

$$N_i(S',0)=N_{i-1}(S',0)+\lambda\left\{\frac{d}{dS'}\left[N_{\rm obs}(S')-\int dS\, N_{i-1}(S,0)\exp\left[-\int_{S'(0)}^{S'(t)} R(S,S'')dS''\right]\right]\right\}, \tag{10}$$

where $N_i(S',0)$ is the ith iteration and the initial guess, $N_0(S',0)$ is given by

$$N_0(S',0)=-dN_{\rm obs}/dS' \ . \tag{11}$$

A change of variable from time to S' defined by Eq. (7) was made and $N_{\rm obs}$ is the observed spin decay. The factor λ is a convergence factor which controls the rate of convergence and is typically of order one; large values can increase the convergence rate but may cause instabilities. A more general discussion of the derivation of this iteration equation and its convergence is found in the Appendix.

We now consider several important special cases of Eq. (9) which are of particular interest.

B. Distribution in energy

One of the most common special cases occurs when

$$R(E_a,t)=\nu_0\exp[-E_a/kT(t)] \ , \tag{12}$$

where ν_0 is an attempt frequency, $T(t)$ is the temperature which may be a function of time, and the spin distribution parameter is an activation energy E_a. In this case, Eq. (7) becomes

$$\int_0^t \nu_0\exp[-E_a/kT(t')]dt'=1 \tag{13}$$

and defines E_a uniquely as a function of t. For a given t, Eq. (13) can be solved for E_a using Newton's method. If the temperature is independent of t, Eq. (13) simplifies to

$$E_a(t)=kT\ln(\nu_0 t) \ , \tag{14}$$

and the distribution of spins with respect to the activation energy E_a is given by

$$\frac{dN}{dE_a}=-N_S(E_a(t),0) \ . \tag{15}$$

In other words, Eqs. (14) and (15) indicate that if the spin-density annealing curves are plotted versus $kT\ln(\nu_0 t)$, a universal decay curve should result independent of annealing temperature; the decay curve reflects the initial distribution of activation energies. Equation (15) has two important consequences. First, the monomolecular rate distribution of Eq. (12) caused by a distribution of activation energies is valid if decays for different temperatures align for a single value of the attempt frequency. A distribution of attempt frequencies, for example, would not yield a universal curve. Second, the value of the attempt frequency can be determined by requiring the spin-density annealing curves to be independent of T.

The case of a time-dependent temperature is particularly useful since it allows one to determine the complete distribution from the lowest to the highest activation energies with a single annealing curve as discussed further below. Consequently, one can determine spin distributions in cases where the changes occur over long time periods and are irreversible, e.g., paramagnetic defects created by implant damage or plastic deformation.

C. Distribution of tunneling distances

Another special case of interest occurs when the spin density decays by tunneling from states with a distribution of tunneling distances assuming constant barrier heights. In this case, the decay rate is independent of time and is given by

$$R(r)=\nu_0\exp(-2r/r_0) \ , \tag{16}$$

where r is the tunneling distance (the spin distribution parameter), ν_0 is the attempt frequency, and r_0 is the tunneling decay length. Equation (7) becomes

$$r(t)=(r_0/2)\ln(\nu_0 t) \ , \tag{17}$$

and Eq. (9) becomes

$$\frac{dN}{dr(t)}=-N(r,0) \ . \tag{18}$$

Thus, in this instance, similar to the activation energy case, the slope of the spin density plotted versus $\ln t$ will yield the distribution in tunneling distances. Unlike the case of the activation energy distribution, the prefactor is not temperature dependent but it may be dependent on applied fields in some cases and hence subject to experimental control.

D. Examples

In this section the formulas developed in the preceding sections are applied to specific examples. The first example we consider is the case in which the defects possess a single activation energy. The decay $N(t)$ is given by

$$N(t)=\exp(-Rt)\,, \tag{19}$$

where $R=(10^{10}\ \mathrm{s}^{-1})\exp[(-0.9\ \mathrm{eV})/kT]$ is plotted in Fig. 1(a) versus $kT\ln(\nu_0 t)$ while the derivative is shown Fig. 1(b). Note that the correct position of the peak in the distribution is recovered with some small broadening of the order of $2.7kT$. Hence, at lower temperature the resolution improves. If further improvements in the resolution are required the iterative method using Eq. (11) can be utilized. Increasing the resolution significantly beyond $\sim 2.5kT$ requires a large signal-to-noise ratio which generally is not obtainable for ESR transients.

The second example we consider is the decay of defects with a two-peak distribution for different temperatures (Fig. 2). The decay from this distribution is presented in Fig. 2(a) and shows a plateau region where the lower activation energy defects have been removed while the higher-energy defects remain. Note also that the various temperature annealing curves overlap to yield a universal curve. The derivative is plotted in Fig. 2(b) and accurately reproduces the original distribution. We see that the distribution is quite accurately recovered from the decay without any iteration.

The use of the iterative method is illustrated in Fig. 3. The solid curve represents the actual distribution responsible for the decay. The dashed curves show the results of various stages of the iteration. The initial distribution is given by Eq. (11) and the subsequent iterations are derived from Eq. (10). After only two evaluations of Eq. (10), the results have converged to the correct distribution. The effective number of independent parameters of this spectrum can be estimated by dividing the total energy span by the resolution. Since the resolution of the spectrum in Fig. 3 is ~ 0.08 eV, the spectrum consists of roughly 12 independent parameters.

A more stringent test of the iterative method is shown in Fig. 4. The method is applied to a signal containing 1% noise, and the final resolution of roughly 0.03 eV was obtained after seven iterations of Eq. (10) (dashed curve).

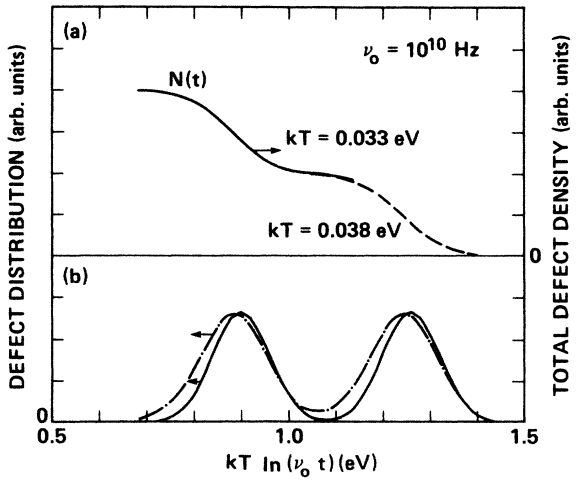

FIG. 2. (a) Defect decay versus $kT\ln(\nu_0/t)$ for the distribution of activation energies given in (b) computed using Eq. (3) for $kT=0.033$ eV (solid line) and $kT=0.038$ eV (dashed line). Note the overlap of the different temperatures. (b) The solid line is the original distribution of activation energies versus energy. The dashed-dotted line indicates the resulting distribution using Eq. (11). Resolution is ~ 95 meV.

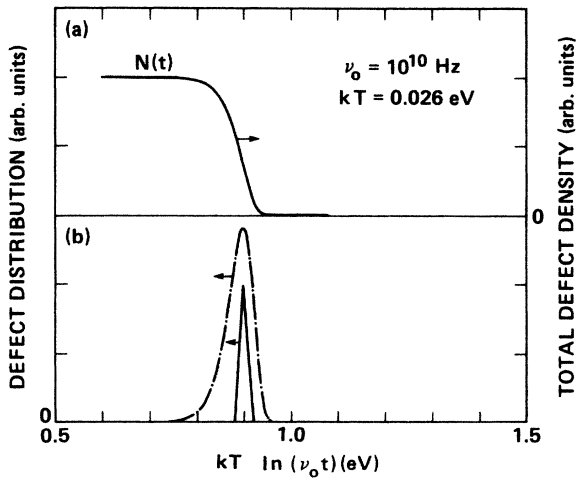

FIG. 1. (a) Defect decay versus $kT\ln(\nu_0 t)$ for the distribution of activation energies given in (b) computed using Eq. (3) for $kT=0.026$ eV. (b) The solid line is the original distribution of activation energies versus energy. The dot-dashed line indicates the resulting distribution using Eq. (9). Resolution is ~ 70 meV.

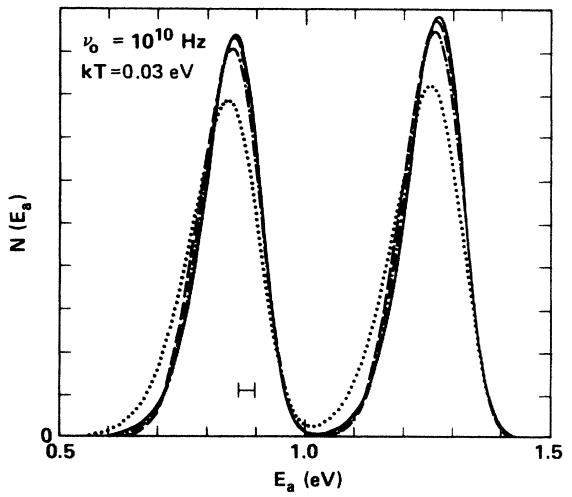

FIG. 3. Solid curve represents the original distribution. From a decay derived using Eq. (3), the starting distribution (dotted) was obtained from Eq. (10), and subsequent distributions using Eq. (9) (dot-dashed and dashed). Resolution is ~ 95 meV.

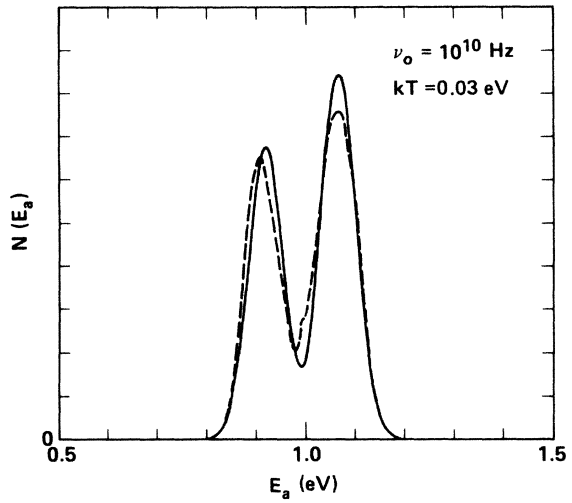

FIG. 4. Distributions obtained from signals with noise. A decay curve was derived from the distribution depicted by the solid line. Relative noise of 1% was added to the decay curve. Equation (10) was used to obtain the dashed curve after seven iterations. The resolution is ~30 meV.

Finally, an example of the case of variable temperature is exhibited in Figs. 5(a)–5(c) on actual data. The curve in Fig. 5(a) represents an experimental decay of metastable silicon dangling bonds in hydrogenated amorphous silicon versus time for the indicated temperature sequence. Further details can be found in the following paper[7] and in Refs. 10 and 11. The activation energy as a function of t was found using Eq. (7) for a distribution of activation energies given by Eq. (11) with $\nu_0 = 10^{10}$ Hz. The equation was solved for E_a for each t using Newton's method for nonlinear equations [Fig. 5(b)]. Figure 5(c) shows the spin distribution plotted versus the energy of Fig. 5(b). Note that the result is a curve without kinks. This same decay curve was also obtained from three decay curves measured at different constant annealing temperatures plotted versus $kT\ln(\nu_0 t)$. This example demonstrates that one or two decay curves are sufficient to extract a distribution, particularly if ν_0 is known.

III. EXPERIMENTAL METHODS

The preceding sections have shown that a primary requirement for the ESR transient analysis is high signal-

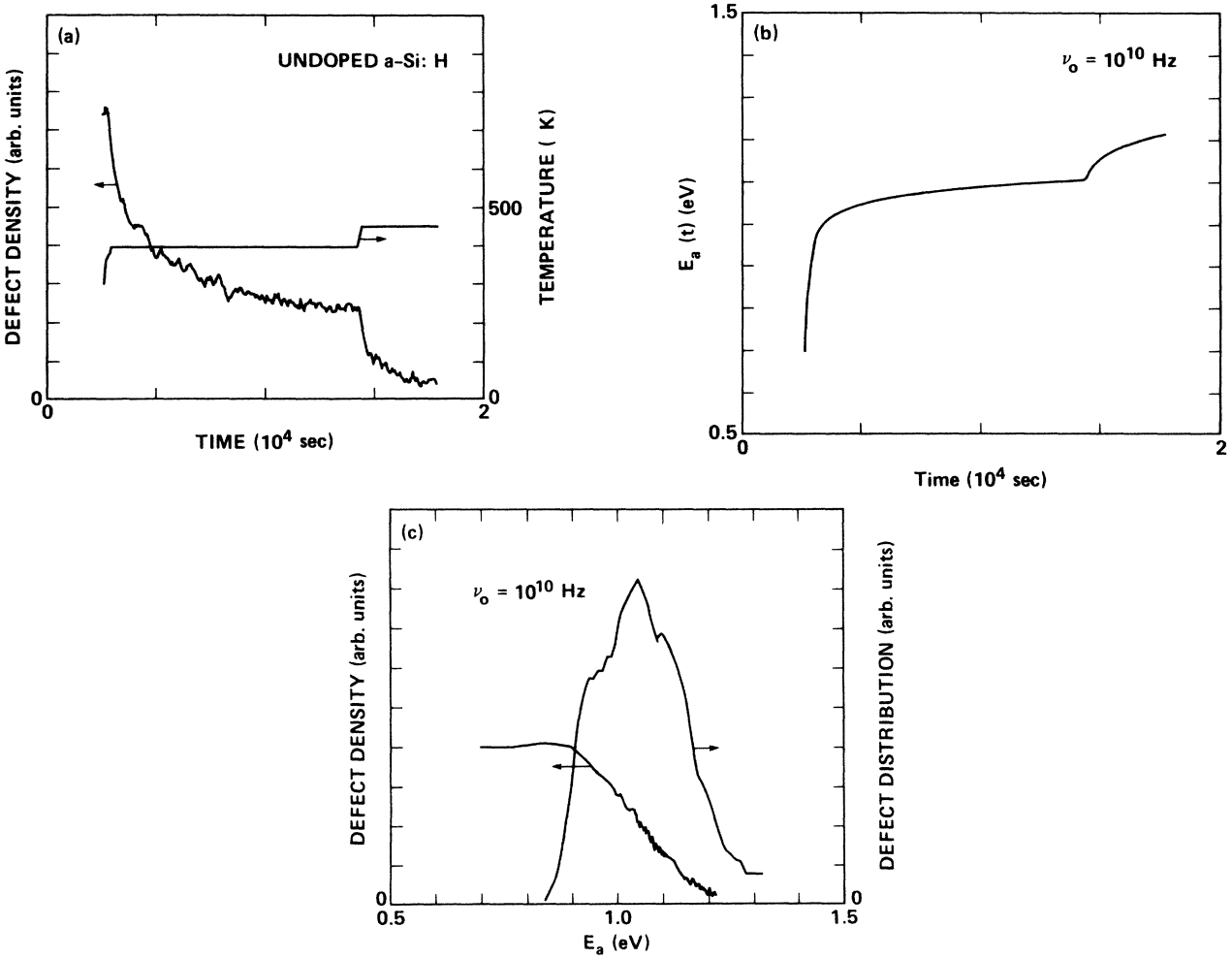

FIG. 5. (a) Total metastable dangling bond defect density in undoped hydrogenated amorphous silicon versus annealing time. The temperature sequence versus time is also presented. (b) Activation energy as a function of time computed by solving Eq. (13) for the temperature sequence shown in (a). (c) Defect decay of (a) plotted versus the energy of (b) and the distribution of the activation energies for the distribution (the derivative). Resolution is ~0.1 eV.

to-noise measurements of spin population transients versus time. Thus, it is important to discuss general methods for obtaining the transients with sufficient time resolution and sensitivity.

Usually, the ESR spin density is obtained by measuring the microwave absorption of a sample in an increasing magnetic field that is modulated with a small high-frequency magnetic field. The peak-to-peak amplitude of the resulting derivative of the microwave absorption line (Fig. 6), is proportional to the spin population. Typically the scan takes a minimum of 100 s, limited by the requirement to keep the speed of the static magnetic field much smaller than the ratio of the ESR linewidth to the time constant used for the synchronous detection. Because in the case of a known constant line shape only the peak-to-peak amplitude is needed to determine the spin density, most of the observation time is wasted providing information on line shape and symmetry repeatedly. Furthermore, the long scan makes the measurements sensitive to spectrometer drifts.

A much more efficient method for determining the spin density is to find it by a double integration of the curve in Fig. 6. This integral, $A(t)$, is related to the spin density by the relation

$$A(t)=CT^{-1}S(H_1,T_1)N(t)\,, \tag{20}$$

where C is a constant, T is the temperature, and the factor $S(H_1,T_1)$ describes the saturation of the spin population as a function of the microwave field H_1, and the spin-lattice relaxation time, T_1. The T^{-1} factor is due to the Curie-law behavior. In general, the temperature dependence of $S(H_1,T_1)$ can be determined by ancillary measurements.

The key to obtaining sufficiently sensitive ESR transients is to obtain a rapid measurement of the area $A(t)$ (up to 100 Hz) using a double modulation method. This can be accomplished by generating a third magnetic field (in addition to the static and high-frequency fields) which varies in time as a triangular or sawtooth waveform of frequency ω between 1 and 100 Hz; the static magnetic field remains constant (Fig. 6). The output of the ESR spectrometer (high-frequency detector) is a waveform that resembles the typical ESR spectrum but repeats at a frequency ω. This waveform can be averaged in a signal averager for periods ranging up to 10 s. At the end of the accumulation period, the double integral of the resulting waveform can be evaluated by a computer and a new average accumulated. The result is that the entire observation time of the spectrometer is now being used to determine the spin density. The time resolution is limited by the speed at which the area can be determined and the sweep rate of the magnetic field. This experimental method works even if the line shape changes as the temperature is altered. Furthermore, one could imagine measuring the decay of two defects simultaneously by applying a magnetic field which switches from one line to the other during a cycle of the third magnetic field waveform. This third magnetic field can be generated by the inclusion of an additional coil within the ESR spectrometer. If the line shape is temperature independent, the magnetic field can be modulated sinusoidally, and the signal averager can be replaced by a lock-in amplifier. Further discussion of experimental details which depend on specific sample characteristics are found in the following paper.[7] In general, we have found that this procedure yields sensitive ESR curves suitable for transient analysis.

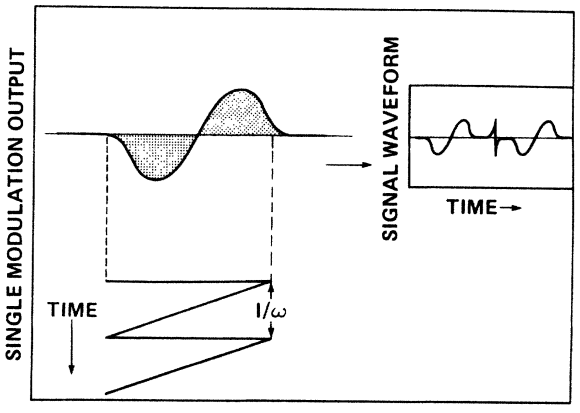

FIG. 6. Procedure for obtaining ESR transients. A third magnetic field with a sawtooth time dependence of frequency ω causes the output of the high-frequency lockin to exhibit the time dependence in the inset. The double integral of the curve in the inset and the shaded region is proportional to the spin density and can be acquired within a $1/\omega$ time period.

IV. DISCUSSION

Finally, we discuss various considerations that are important for the application of ESR transient spectroscopy such as sensitivity and resolution, etc., as well as possible extensions of the spectroscopy. The analysis presented here is compared to those used in other transient analyses.

A. Application of the spectroscopy

One of the most important considerations for applying the present analysis to a new system is to establish the validity of a monomolecular rate since the analysis depends on the fact that that individual defect populations vary independently of each other. If the paramagnetic defects obey different kinetics, the analysis is much more complex. The monomolecular recombination law can be established by observing the decay for different starting defect densities (see following paper[7] for further details). If the result decays are independent of the number of initial defects, the monomolecular relationship is valid. One must ensure that, as the total number of defects is altered, their distribution remains the same.

Another important issue to consider is the position of the Fermi level. Since the spin signal is used to monitor the defect density or the change of occupancy, the Fermi-level position during annealing must be known because the Fermi-level shift can mask actual changes in defect density. Furthermore, the occupation of the existing defect can be altered by the depopulation of another defect. In addition, the temperature dependence of the ESR signal must also be known. For example, the ESR signal can change as a function of temperature due to the Curie temperature dependence of the susceptibility, and the decays

must be corrected for this temperature variation.

A final point to consider is the appropriate equation for the rate $R(S,t)$. If the rate depends on temperature and the decay is plotted versus $S(t)$ defined by Eq. (7), a single universal curve independent of annealing temperature (Fig. 2) should result using the correct value of ν_0. If a single value does not yield a single curve, the spins may exhibit a distribution of attempt frequencies and a more complex extension of the analysis, discussed below, must be invoked.

B. Resolution and noise

In a discussion of a spectroscopic method the noise is intimately related to the resolution. We observed from the single exponential example or from the kernal of Eq. (A6) that the usual resolution of the transient spectroscopy is roughly $2.7kT$. From a physical standpoint, this resolution limit occurs because the change of occupancy due to thermal processes occurs over an energy of kT. For ESR transients, the signal-to-noise considerations generally prohibit significant increase of the resolution beyond this limit. The noise level increases as the resolution is enhanced. Since each iteration increases the resolution in a stepwise fashion, the iterative method is continued beyond the $2.7kT$ limit until the noise becomes significant. Direct inversion of the operator Eq. (A4) or a nonlinear least-squares method does not have this flexibility.

There are several ways to improve the signal-to-noise ratio of the transient analysis. Since the spin density is often plotted versus the logarithm of time, the signal averaging should be logarithmic in time as well. The averaging time for a point in the Nth decade of time should be 10^N times as long as for a point in the first decade. This procedure assures that the smaller signals at later times have the same or better signal-to-noise ratio as the earlier points with larger signals. Another noise-reduction method utilizes filtering to remove the high-frequency noise since the derivative required in Eq. (9) amplifies the high-frequency noise of the measurement. The sliding average methods of Savitzky and Golay[12] allow one to perform the filtering operation simultaneously with the derivative. During each iteration, the derivative in Eq. (10) is computed using the filtering derivative.

Finally, because the derivative amplifies changes in slope, one must be careful when combining decay curves from different temperatures. A small misalignment of the different temperatures can give rise to small artifactual features in the derived distribution. The misalignment can occur for several reasons. First, it takes a longer time period for the high temperatures to stabilize so the defect distribution begins to change during the stabilization period. As a consequence, the zero of the time and the initial defect density are somewhat uncertain. The uncertainty results in some misalignment when curves obtained at different temperatures are combined. This problem is considerably reduced by using the variable temperature analysis and several different temperatures to capture both the low-activation-energy (short time) and the high-energy (long time) components of the decay in a single curve. The initial starting density is thereby ensured to be the same for the different temperatures, and the zero of the time axis is also consistent between the various temperatures. The second case of misalignment arises from slight errors in the attempt frequency and yields curves which systematically are offset as the temperatures increase. The attempt frequency can be adjusted to remove the discontinuities in the slope and the resulting artifacts in the distribution.

C. Comparison to other analyses

The transient analysis presented in Sec. II has a number of advantages over other methods such as nonlinear least-squares fitting, Laplace transforms, Fourier transforms, and the method of moments.[5] Since nonlinear least-squares fitting of the transients from a distribution to the observed transient usually performs as well as these methods,[5] we will only discuss nonlinear least squares. There are a number of problems with this approach.

The first problem is that it is computationally quite expensive. In the general case when the temperature is a function of time (more important for ESR transients), the rate depends on time. One could imagine attempting a nonlinear least-squares fit of the data to Eq. (3) where $N(S,0)$ for various values of S are the fitting parameters. Unfortunately, a reasonable spectrum may consist of up to 20 to 30 points, and since Eq. (3) is computationally quite time consuming, involving a double integral, the optimization problem is formidable.

For example, the distribution obtained in Fig. 3 took two evaluations of Eq. (11) to achieve the correct distribution containing effectively 12 independent parameters. We can estimate the typical number of evaluations that a nonlinear minimization routine might require for a problem of similar complexity. The best of 12 nonlinear minimization routines evaluated in Ref. 13 took an equivalent of 298 function evaluations for a particular ten-parameter model.

There is a good reason for this large discrepancy in the number of function evaluations and corresponding computational effort. The nonlinear fitting routine blindly adjusts all parameters to obtain a good fit, whereas Eq. (10) selects only those variations which are required to obtain a good fit. It is based on the physics of the problem rather than a general mathematical procedure.

The second problem with nonlinear least-squares formulation is more serious, particularly for ESR transients. The resulting least-squares solution oscillates wildly and has unphysical negative regions of the distribution. This phenomenon, well known in various deconvolution problems, occurs because the high-frequency noise of the measurement becomes amplified by the inversion algorithm.[14] Prior information, such as non-negativity and slow variation of the distributions in the parameter S, cause the various $N(S,0)$ values to be interrelated in a complex manner that is difficult to include within a least-squares formulation of the problem.

Finally, the third problem is that the most likely spin distribution is *not* one which minimizes the sum of the squares of the error between theory and experiment as assumed in nonlinear least-squares minimization, but rather

a distribution for which the errors between calculated and experimental transients are randomly distributed with zero mean.[15] Thus, simple nonlinear least-squares fits to models expend a great deal of computation to arrive at a nonphysical distribution.

D. Extensions of the transient analysis

There are a number of ways the transient analysis may be extended to other problems or more complex models. First, one may include rates of defect creation according to the relation

$$\frac{dN(S,t)}{dt}=G(S)-R(S)N(S,t)\ , \tag{21}$$

where $G(S)$ is the generation rate per unit volume per unit S for spins characterized by the parameter S. Equation (21) has a solution given by

$$\begin{aligned}\Delta N(S,t)&=N(S,\infty)-N(S,t)\\&=[N(S,\infty)-N(S,0)]\exp[-R(S)t]\\&=\Delta N(S,\infty)\exp[-R(S)t]\ .\end{aligned} \tag{22}$$

Hence, the final distribution of centers created can be determined using the same analysis as given by Eq. (9).

A second extension consists of the case in which the defects are distributed according to two parameters rather than one.[16] An example of such a case would occur if the defects were distributed in tunneling distance or attempt frequency and in energy. In such a case, Eq. (7) contains a double integration over both parameters. In Fig. 7, the time decay of this population is shown for a distribution in E_a and ν_0. The decay of the population occurs as the line given by the relation $Rt=1$ moves to the left through the distribution. The defects remaining lie to the left of the line. For higher temperatures, the slope of $R(\nu_0,E_a)t=1$ curve increases. In principal, one could perform an analysis similar to that used in tomography to separate the distribution in activation energy and attempt frequency.

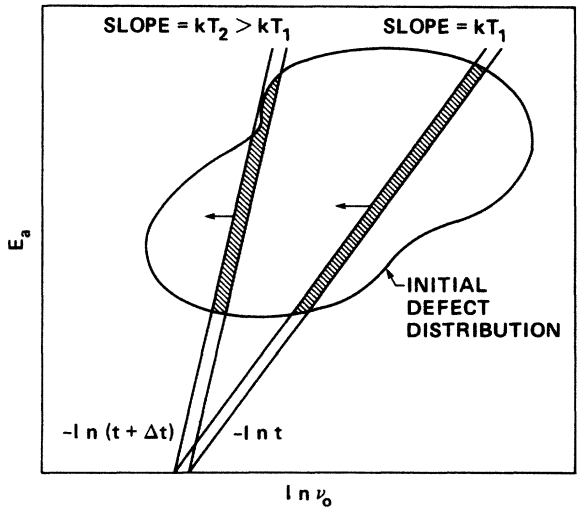

FIG. 7. Decay of a population of defects consisting of a distribution of activation energies and attempt frequencies. The straight lines indicate the condition $R(E_a,\nu_0)t=t\nu_0\exp(E_a/kT)=1$ while the shaded region represents the defects decaying in the interval between t and $t+\Delta t$ for two different temperatures T_1 and T_2. The remaining defects lie to the left of the line while the annealed defects lie to the right of the straight lines.

The third and most important extension is that all the techniques described in this paper can be applied to any measurement which measures monomolecular decays of populations. The analysis may be applied to metastable changes in conductivity, luminescence, current, charge, capacitance, etc. In fact, in the case of weak excitation intensities, luminescence decay curves have been interpreted in terms of a tunneling distribution.[17] Some of the measurements are, of course, equivalent to deep-level transient spectroscopy measurements. The important point is that in some cases, one cannot repetitively populate traps as required by the usual deep-level spectroscopy. A single decay experiment is possible using this analysis.

Thermally stimulated current spectroscopy (TSCP) could also benefit from the transient analysis presented in Sec. II and the Appendix. One of the principal experimental difficulties in performing thermally stimulated current spectroscopy is to achieve a constant rate of temperature increase. The use of Eq. (7) considerably simplifies a TCSP experiment. One merely needs to ramp the temperature in any monotonic increasing manner while the sample temperature is measured. Equation (7) provides the appropriate defect energy contribution to the current.

V. CONCLUSION

In this paper the theoretical basis for ESR transient spectroscopy for studying defect energies has been presented. The spectroscopy enables one to examine the distribution of defect energies for a specific microscopic defect labeled by its characteristic spin signal and hence may be particularly useful in cases where more than one defect is present. The transient analysis may contribute to what is often a major problem in the study of defects in semiconductors, namely, how to relate the microscopic information of the ESR signal of a defect to other properties such as the defect energy.

The basis of the spectroscopy is to analyze the decays of a given spin type versus time, temperature, or other parameters. Using the analysis presented in this paper, the decay curve can be analyzed to obtain the distribution of the spin population in energy or tunneling distances. The iterative method for recovering the defect distributions is substantially less time consuming than nonlinear least-squares minimization methods and more importantly, the distributions are more probable and are physically meaningful.

ACKNOWLEDGMENTS

This work was supported in part by Solar Energy Research Institute (Golden, CO) Contract No. XB-5-04067-1.

APPENDIX

A more general mathematical development of the transient analysis, presented below, justifies the approximations used, the convergence of the iteration scheme, and makes a mathematical connection with the recent developments in modern signal processing.

Starting from Eq. (3), one may take the derivative of the spin density with respect to the time-dependent parameter, $S'(t)$ defined by Eq. (7). The result is

$$\frac{dN}{dS'}=\int dS\, N(S,0)\frac{d}{dS'}\left[\exp\left[-\int_0^t R(S,t')dt'\right]\right]. \quad \text{(A1)}$$

This can be rewritten in the form

$$\frac{dN(S')}{dS'}=\int dS\, N(S,0)H(S,S'), \quad \text{(A2)}$$

where by a change of variable

$$H(S,S')=-R(S,t)\frac{dt}{dS'}\times\left[\exp\left[-\int_{S'(0)}^{S'(t)}R(S,s)\frac{dt}{ds}ds\right]\right]. \quad \text{(A3)}$$

Equation (A2), a typical operator equation, can be written in matrix form

$$\mathbf{dN}=\mathbf{HN}. \quad \text{(A4)}$$

If $\mathbf{G}=(\mathbf{I}-\lambda\mathbf{H})\mathbf{C}$ is a nonexpansive operator where $\mathbf{I}$ is the identity operator, $\mathbf{C}$ is a constraint operator, and λ is a number greater than zero, then it can be shown that the iteration equations

$$\mathbf{N}_0=\mathbf{dN},\qquad \mathbf{N}_i=\mathbf{CN}_{i-1}+\lambda(\mathbf{dN}-\mathbf{HCN})_{i-1},\quad i=2,3,\ldots \quad \text{(A5)}$$

converge to the solution of Eq. (A3).[15] Equation (A5) is identical to Eq. (10). The constraint operator, $\mathbf{C}$, includes any known constraints on the spin distribution $\mathbf{N}$ such as non-negativity or the fact that $N(S')=0$ for large and small values of S'. Using the definition of $\mathbf{H}$ and assuming $\mathbf{C}=\mathbf{I}$, $\mathbf{G}$ can be shown to be a nonexpansive operator for sufficiently small λ and hence, Eqs. (10) and (A5) converge. The approximation used in Eq. (9) is equivalent to replacing $\mathbf{H}$ by $-\mathbf{I}$ in Eq. (A2). In some cases it may be possible to invert Eq. (A4) directly by calculating $\mathbf{H}^{-1}$, but $\mathbf{H}$ is often quite ill-conditioned and $\mathbf{dN}$ contains noise; the iteration equations yield superior results.

For the specific case when $R(S,t)=\nu_0\exp(S/kT)$, $H(S,S')$ has the simple form

$$H(S,S')=-(1/kT)\exp[(S'-S)/kT]\times\exp\{-\exp[(S'-S)/kT]\}. \quad \text{(A6)}$$

If the spacing between points is kT or greater, the matrix representation $\mathbf{H}$ is very close to the identity matrix used to obtain Eq. (9). This explains why Eq. (9) performs well even without iteration.

[1] M. C. Chen and D. V. Lang, Phys. Rev. Lett. **51**, 427 (1983).
[2] G. D. Watkins and J. W. Corbett, Phys. Rev. **121**, 1001 (1961).
[3] H. Okushi and Y. Tokumaru, Jpn. J. Appl. Phys. **20**, 261 (1981); N. M. Johnson, Appl. Phys. Lett. **42**, 981 (1983).
[4] See, for example, M. C. Driver and G. T. Wright, Proc. Phys. Soc. **81**, 141 (1963).
[5] A review and comparison of various methods used to analyze luminescence decays may be found in A. E. McKinnon, A. G. Szabo, and D. R. Miller, J. Phys. Chem. **81**, 1564 (1977).
[6] F. Kostial and J. Slabeycius, Phys. Status Solidi A **87**, K109 (1985).
[7] M. Stutzmann, W. B. Jackson, and C. C. Tsai, Phys. Rev. B **34**, 63 (1986) (following paper).
[8] I. Isenberg, J. Chem. Phys. **59**, 5696 (1973).
[9] R. W. Schafer, R. M. Mersereau, and M. A. Richards, Proc. IEEE **69**, 432 (1981).
[10] M. Stutzmann, W. B. Jackson, and C. C. Tsai, Appl. Phys. Lett. **45**, 1075 (1984).
[11] M. Stutzmann, W. B. Jackson, and C. C. Tsai, Phys. Rev. B **45**, 1075 (1985).
[12] A. Savitzky and M. J. E. Golay, Anal. Chem. **36**, 1627 (1964).
[13] Y. Bard, SIAM J. Numer. Anal. **7**, 157 (1970).
[14] J. B. Abbiss, M. Defrise, C. De Mol, and H. S. Dhadwal, J. Opt. Soc. Am. **73**, 1470 (1983).
[15] H. J. Trussell, IEEE Trans. Acoust. Speech Signal Process. **ASSP-31**, 129 (1983).
[16] Y. A. Gorokhovatskii, A. N. Gubkin, and A. I. Margulev, Phys. Status Solidi A **65**, 485 (1981).
[17] C. Tsang and R. A. Sheet, Phys. Rev. B **19**, 3027 (1979).

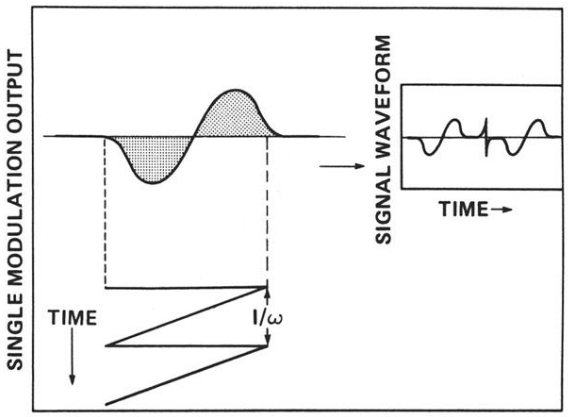

FIG. 6. Procedure for obtaining ESR transients. A third magnetic field with a sawtooth time dependence of frequency ω causes the output of the high-frequency lockin to exhibit the time dependence in the inset. The double integral of the curve in the inset and the shaded region is proportional to the spin density and can be acquired within a $1/\omega$ time period.