PHYSICAL REVIEW B VOLUME 30, NUMBER 10 15 NOVEMBER 1984

*M*Fe center: A configurationally bistable defect in InP:Fe

Mark Levinson, Michael Stavola, and P. Besomi*
AT&T Bell Laboratories, Murray Hill, New Jersey 07974

W. A. Bonner
Bell Communications Research, Murray Hill, New Jersey 07974
(Received 2 July 1984)

We report the observation of a defect in Fe-doped InP which can exist in either of two configurations, for the same charge state. Each configuration exhibits distinct electronic and optical properties. Thermally stimulated capacitance, capacitance transient spectroscopy, and photocapacitance were used to study the properties of the defect in each configuration, and the kinetics of the reversible transformations between configurations. The unique properties of the defect are discussed.

Semiconductor defect-induced deep levels are characterized by localized electronic wave functions which are often strongly coupled to the lattice. Effects of this electron-lattice interaction are evident in such phenomena as carrier capture by multiphonon emission,[1,2] recombination-enhanced defect reactions,[3,4] inverted ordering of energy levels (negative U),[5–7] and most strikingly, metastable effects. These latter include persistent photoconductivity due to *DX* centers in Al-Ga-As (Ref. 8) and Ga-As-P (Ref. 9), and photocapacitance quenching and other anomalous optical properties associated with the EL2 center in GaAs.[10–13] More recently, a remarkable phenomenon involving configurational bistability has been observed for an electron irradiation damage defect in InP, the *M* center.[14–17] This latter center is a multiply charged defect which, for two of its charge states, can exist in either of two configurations, each with distinct electronic and optical properties. Transformations between the configurations are reversible and the configuration which is observed experimentally is dependent upon the thermal, electronic, and optical history of the sample. It appears that the metastability of EL2 may be due to a similar phenomenon,[18] as does that of an *A*-center-related defect in Si.[16,19]

Here we report the observation of a new configurationally bistable defect in Fe-doped InP, which we call the *M*Fe center. Although some of its general features are similar to the *M* center, the configurational transformation kinetics reveal unique and unexpected behavior. The defect is of additional interest due to the widespread use of the Fe impurity to produce semi-insulating InP substrates for device fabrication.

Samples were made from lightly Fe-doped liquid-encapsulated—Czochralski-grown n-type InP. Mesa p^+-n junctions were formed using either Zn- or Cd-doped p^+ liquid-phase epilayers followed by photolithography and etching. Ohmic contacts consisted of (Au,Sn)-Au on the back sides of the wafers and (Au,Zn)-Au on the p^+ sides. The contacts were alloyed at 400 °C for 30 s. The samples were diced, epoxy mounted on TO-18 headers, and wire bonded. C-V profiling revealed a uniform room-temperature free-electron density of 2×10^{15} cm^{-3}.

Thermally stimulated capacitance (TSCAP), photocapacitance, and deep-level capacitance transient spectroscopy (DLTS), were used to determine the electronic and optical properties of the defect in each configuration, and the configurational transformation kinetics. DLTS measurements were made using the boxcar-averager technique. Variable wavelength illumination was obtained with a tungsten source and grating monochromator.

In brief, when occupied by electrons the defect can exist in either of two configurations, which we refer to as *A* and *B*. Configuration *A* is most simply obtained by cooling the sample from above 200 K with no applied bias. Configuration *B* is obtained by cooling with an applied reverse bias. Distinct thermally activated electron emissions are observed for each of the two configurations. They can be detected by one of the capacitance techniques, and they serve as signatures for each configuration. When the system is excited thermally, optically, or electronically (by hole injection) so as to induce a transformation from one configuration to the other, the changing magnitudes of these emission signals can be used to determine the changing fractions of the defects in each configuration. This information allows the transformation rates, and thus the kinetics, to be derived. It is apparent that the emission spectra for both configurations arise from the same defect, because monitoring the transformations by TSCAP shows that an increase in the magnitudes of the capacitance changes of one spectrum is always accompanied by a proportional decrease in those of the other.

The thermally activated electron emissions which are the signatures of each of the two configurations are shown in the TSCAP data of Fig. 1. Curve (a) corresponds to configuration *A*. It was obtained by cooling the sample to ~40 K in darkness at zero bias. Reverse bias was then applied and the capacitance recorded upon rewarming in the normal manner. The only feature is an electron-emission step, labeled ($A1,A2$). This label reflects the apparent two-electron nature of the emission, as will be discussed below. Curve (b) corresponds to configuration *B*. It was obtained by cooling in darkness under reverse bias to ~40 K, followed by a zero bias trap-filling pulse, and rewarming. Two emissions, labeled $B1$ and $B2$, are

 ©1984 The American Physical Society

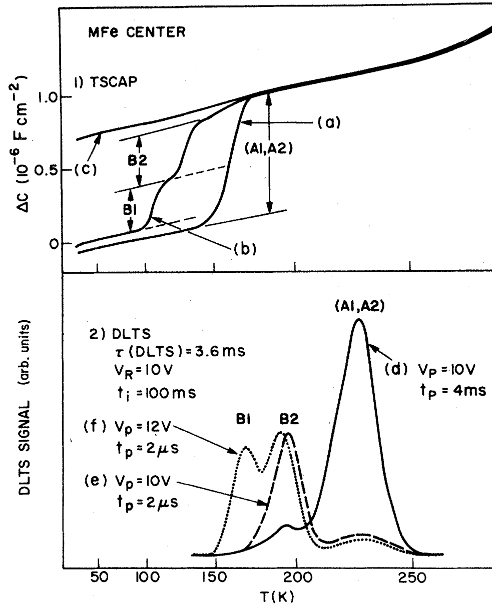

FIG. 1. Thermally activated capacitance (TSCAP) and capacitance transient (DLTS) spectra under various experimental conditions as described in the text. V_R and V_P are the reverse bias and trap-filling pulse voltages, respectively. τ (DLTS) is the DLTS system time constant.

observed. These steps were observed at their maximum magnitudes after the shortest available filling pulse of 50 ns, indicating capture rates at 40 K greater than 2×10^{7} s^{-1}, which imply capture cross sections $\sigma_n \geq 6\times10^{-16}$ cm^2. The residual $(A1,A2)$ step above $B2$, and the separation between curves (a) and (b) below $B1$, appear to be artifacts due to the depletion-layer-edge region. Curve (c) is obtained by cooling with reverse bias and rewarming with no trap-filling zero bias pulse. It represents the temperature-dependent capacitance of the diode in the absence of carrier trapping effects.

The two configurations are more dramatically displayed in the DLTS spectra of Fig. 1. Consistent with the TSCAP data, configuration B gives rise to two peaks $B1$ and $B2$, and configuration A shows one peak $(A1,A2)$ such that the maximum peak heights $B1 \approx B2 \approx \frac{1}{2}(A1,A2)$.

An additional peak was observed at higher temperature, but with a defect concentration a factor of 20 smaller than that of the MFe center. This defect has an electron emission activation energy of 0.66 eV, and an emission section of 8×10^{-14} cm^2. This peak may correspond to the 0.59 eV Fe-related defect in InP reported by Tapster and coworkers.[20]

In contrast to TSCAP methods where only single transients are observed, the DLTS technique relies on repetitive capture and emission processes. As discussed below, some of these processes control the configurational transformations so that the transformations may be occurring during the DLTS measurement. However, if the transformations are occurring on a time scale which is fast compared with the heating rate of the sample (~ 0.1 $K\,s^{-1}$), then a steady state is attained, and with the proper analysis kinetic information can be extracted from the data. During the DLTS measurement, $B \rightarrow A$ takes place during trap-filling pulses (bias off), and $A \rightarrow B$ occurs between pulses (bias on). The peak heights thus reflect the fraction of defects in configuration A, f_A, or configuration B, f_B, at the end of the trap-filling pulse.[21] These fractions are functions of the pulse duration t_p, the time interval between pulses t_i, and the temperature-dependent transformation rates R_{AB} and R_{BA}. It can be shown that at steady state the fractions f_A and f_B are given by[21]

$$f_A = 1 - f_B = (1-e^{-R_{BA}t_p})(1-e^{-(R_{AB}t_i + R_{BA}t_p)})^{-1} . \quad (1)$$

It is seen in Fig. 1(d) that for a constant $t_i = 100$ ms and large t_p, $B2$ almost vanishes and $(A1,A2)$ increases to near its full height. For small t_p [Fig. 1(e)], $B2$ increases to its maximum and $(A1,A2)$ is reduced to a residual peak which is assumed to be due to the depletion-layer-edge region. (Near the depletion-layer edge, the defects are in the Debye tail of the free electrons and thus tend to always be in configuration A.) At the temperature where $B1$ occurs, the configurational transformations are slow compared with the heating rate, and a steady state is not attained. $B1$ is found to appear only after cooling with the bias on, using small t_p and large t_i, and heating quickly.

It is observed that the transformation $A \rightarrow B$ is also promoted by the presence of holes. This is evident in Fig. 1(f) where the trap-filling pulse height is increased so as to forward bias the diode and create hole injection. Both peaks $B1$ and $B2$ are clearly seen. In this case, during the filling pulse holes are captured by defects in configuration A. These defects then transform to B and capture electrons which are subsequently reemitted as $B1$ and $B2$. This hole-induced transformation was also observed at low temperature. A sample was cooled to 35 K at zero bias so as to set configuration A, and an injection pulse applied. A TSCAP measurement then revealed a transformation to B.

Data for the thermally activated emission $(A1,A2)$ are shown in the Ahrrenius plot of Fig. 2(a). The activation energy (corrected for the T^2 dependence of the prefactor) is 0.41 eV with an emission section $\sigma_\infty = 1.9\times10^{-14}$ cm^2.

An experiment was performed which demonstrates that the emission $(A1,A2)$ controls the transformation $A \rightarrow B$: The sample is cooled from a temperature above step $(A1,A2)$ at zero bias to set the defects in configuration A. Reverse bias is applied and the TSCAP monitored while warming until a temperature is reached in the middle of step $(A1,A2)$. The sample is then quickly recooled, and a zero bias trap-filling pulse is applied. A new TSCAP measurement shows that a portion of defects equal to that which had completed the $(A1,A2)$ emission [as shown by the point on the $(A1,A2)$ step where recooling began] now show emissions $B1$ and $B2$. The portion which had not undergone emission $(A1,A2)$, now shows the $(A1,A2)$ step. Thus the emission $(A1,A2)$ results in the transformation to configuration B.

The observed transformation $A \rightarrow B$, due to hole cap-

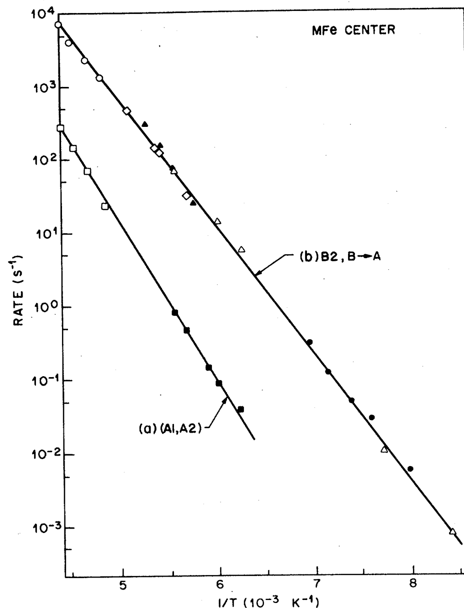

FIG. 2. Ahrrenius plots of thermally activated electron emissions ($A1,A2$) and $B2$, and the configurational transformation $B\to A$. □ and ■ denote ($A1,A2$) emission rates from DLTS and single transient data, respectively. ▲ and ● denote $B2$ emission rates from DLTS and single transient data, respectively. ○ and ◇ denote $B\to A$ rates from DLTS peak heights vs t_p, using ($A1,A2$) and $B2$, respectively. △ denotes $B\to A$ rates from zero bias anneals and TSCAP.

ture in configuration A, is consistent with the above data. It shows that the transformation $A\to B$ is charge state controlled. In other words, for the defect in configuration A, hole capture yields the same result as electron emission, although by a different kinetic process.

Emission-rate data for $B2$ are shown in Fig. 2(b), as is the temperature dependence of the transformation rate $B\to A$. The $B2$ emission was determined in the usual way at higher temperatures from the DLTS peak position. At lower temperatures, single transients were analyzed. The $B\to A$ rate was determined at low temperature by cooling to the measurement temperature under applied bias to set configuration B, and removing the bias for a given time. The bias was then reestablished, followed by a TSCAP scan to observe the fraction of defects which had transformed to configuration A. At higher temperatures, the rate was derived from the peak heights $B2$ and ($A1,A2$) as a function of pulse width t_p, at constant t_i, using Eq. (1) and taking R_{AB} equal to the ($A1,A2$) emission rate at the temperature of the peak.

Ahrrenius behavior was also observed for the emission $B1$. Because of the difficulty in obtaining DLTS signals for this emission, it was examined using single transients.[22] An activation energy of 0.24 eV, with an emission section 3×10^{-17} cm^2, was found.

It is apparent that the kinetics of the transformation $B\to A$ and the electron emission $B2$ are identical within experimental error. The activation energy is 0.35 eV with a preexponential factor of 2.8×10^{11} s^{-1}. This finding is all the more remarkable when it is remembered that these two processes are in principle very different. $B2$ represents an electron emission which takes place when the defects are in the nonequilibrium depletion region. The transformation $B\to A$, on the other hand, appears to involve electron capture and occurs only when the defects are in neutral material with free electrons present.

Photocapacitance measurements[23] were performed to determine the photoionization behavior of emissions $B1$, $B2$, and ($A1,A2$). Photoionization was not observed in configuration A for all photon energies used $\sim0.5<h\nu<1.2$ eV. Likewise, no transformation $A\to B$ was detected. (Above-band-gap light produces holes which promote the transformation $A\to B$, as previously discussed.) That no photoionization occurs for configuration A implies that a very large lattice relaxation occurs here. The optical transition energy is more than three times that of the thermally activated emission, 0.41 eV.

However, photoionizations $B1$ and $B2$ were observed. The optical transition energies for $B1$ and $B2$ are ~0.5 and ~0.6 eV, respectively. Comparison with the corresponding thermal activation energies of 0.24 and 0.35 eV indicates that significant relaxations are also occurring in in configuration B.

Because the transformation $B\to A$ occurs at the same rate as the thermally activated emission $B2$, the effect on the transformation of the optically activated $B2$ emission was examined. The sample was cooled with reverse bias so as to set configuration B. At various temperatures below the thermal $B2$ emission, it was then illuminated with the bias off. TSCAP measurements were then made to reveal any transformation to configuration A. No transformation was observed.

It is thus evident that the transformation $B\to A$ is associated with a thermally activated process. We propose a model in which one configurational barrier provides the rate-limiting process for both the configurational transformation and the thermal emission $B2$. Surmounting this barrier produces an intermediate or "activated" state with no change of charge. Once in the intermediate state, either an electron is emitted so that the defect is unoccupied, or an additional electron is captured, leaving the defect doubly occupied in configuration A. In the absence of free electrons, the emission process occurs. Conversely, with free electrons available, a rapid capture process into configuration A dominates.

These phenomena can be represented by the configuration-coordinate diagram of Fig. 3. The notation is the same as that previously used for the M center.[17] When the defect is unoccupied it is in charge state n and the system is described by the curve labeled C^n, which is common to the two configurations. Upon the capture of one electron the system moves to curve A^{n-1}, into what we have referred to as the intermediate state. The defect then spontaneously undergoes a structural relaxation to a lower energy condition in configuration B, along the curve which is now labeled B^{n-1}. A subsequent electron capture at low temperature leads to a doubly occupied defect in configuration B, as represented by curve B^{n-2}.

Alternatively, from A^{n-1} a second electron capture can bring the system to A^{n-2}, which is the stable condition of

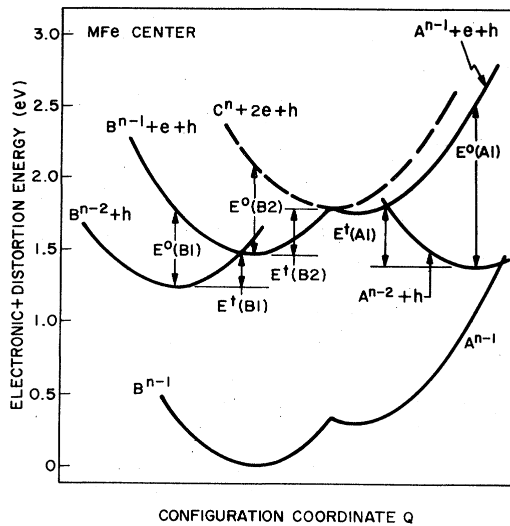

FIG. 3. Configuration coordinate diagram which relates changes in electronic plus lattice distortion energy for each charge state to a configuration coordinate, Q.

configuration A. However, it is assumed that in the temperature range investigated, the relaxation from A^{n-1} to B^{n-1} is faster than capture into A^{n-2}, so the A^{n-1} to B^{n-1} relaxation dominates. The system can only reach A^{n-2} when free electrons are present (bias off) and there is sufficient thermal energy to activate the system from B^{n-1} back to A^{n-1} and then to A^{n-2}. With no free electrons available (bias on), the first electron is simply reemitted and the system moves to curve C^n. Therefore the activation barrier from B^{n-1} to A^{n-1} [(labeled $E^t(B2)$] controls both the $B2$ emission and capture into A^{n-2}, as required by the data.

In this model the emission ($A1, A2$) involves two electrons. The first emission, $A1$, brings the system from A^{n-2} to A^{n-1}. The second electron is either emitted directly to C^n, or is momentarily trapped by the relaxation to B^{n-1}. However, at these temperatures emission from B^{n-1} is immediate because the thermal emission activation energy $E^t(B2)$ is less than that of the first electron to be emitted, $E^t(A1)$. Thus the emission we have called $A2$ is essentially the same as $B2$. Therefore, when the temperature is sufficient to allow emission of the first electron $A1$, the second is emitted immediately thereafter. On the other hand, the ordering of thermal activation energies is normal for configuration B, where the two electron emissions are seen consecutively as $B1$ and $B2$.

A possible microscopic model for the defect involves the kind of electrostatic and lattice-strain-driven defect complex rearrangement mechanism which has been proposed previously for both the M center in InP (Refs. 14 and 15) and EL2 in GaAs.[18] This mechanism can be illustrated by considering a complex consisting of a multiply charged defect C and an ionized shallow donor D, although the mechanism would be equally valid for other types of complexes. In configuration A the two defects are closely associated so that curves A^{n-1} and A^{n-2} correspond to C^0D^+ and C^-D^+, respectively. The transformation to configuration B involves a relaxation which increases the separation of the two species, although they are still bound together. B^{n-1} and B^{n-2} thus correspond to C^0+D^+ and C^-+D^+, while C^n represents C^++D^+. In this picture, the relaxation from A^{n-1} to B^{n-1} would involve a lattice-strain-driven isoelectronic rearrangement $C^0D^+ \rightarrow C^0+D^+$. It should be noted that this representation of the structural relaxation is only schematic. The actual relaxation may involve the motions of several atoms.

The discovery of the MFe center adds to the small but growing list of point defects in covalent semiconductors which exhibit configurational bistability. Like the M center, the MFe center shows clear bistability, where for one or more charge states the defect can remain metastable at low temperature in either configuration. The two configurations can be detected by their distinct optical and electronic properties. The EL2 center in GaAs and the metastable A-center related defect in Si also appear to exhibit configurational bistability, but their particular properties make direct observation of the two configurations more difficult. It is an intriguing question as to whether there is a common mechanism and driving force associated with the configurational instabilities in all these defects, or if some or all cases are unique. It is likely that a knowledge of the chemical and structural identities of the defects will be required to completely understand these phenomena.

---

*Present address: Lasertron, Burlington, MA 01803.

[1]C. H. Henry and D. V. Lang, Phys. Rev. B 15, 989 (1977).
[2]H. Sumi, Phys. Rev. B 27, 2374 (1983), and references therein.
[3]L. C. Kimerling, Solid-State Electron 21, 1391 (1978).
[4]D. V. Lang, Annu. Rev. Mater. Sci. 12, 377 (1982).
[5]P. W. Anderson, Phys. Rev. Lett. 34, 953 (1975).
[6]G. A. Baraff, E. O. Kane, and M. Schlüter, Phys. Rev. Lett. 43, 956 (1979).
[7]G. D. Watkins and J. R. Troxell, Phys. Rev. Lett. 44, 593 (1980).
[8]D. V. Lang, R. A. Logan, and M. Jaros, Phys. Rev. B 19, 1015 (1979).
[9]R. A. Craven and D. Finn, J. Appl. Phys. 50, 6334 (1979).
[10]G. M. Martin, Appl. Phys. Lett. 39, 747 (1981).
[11]G. Vincent, D. Bois, and A. Chantre, J. Appl. Phys. 53, 3643 (1982).
[12]P.Leyral, G. Vincent, A. Nouailhat, and G. Guillot, Solid State Commun. 42, 67 (1982).
[13]E. R. Weber and J. Schneider, Physica (Utrecht) 116B, 398 (1983).
[14]M. Levinson, J. L. Benton, and L. C. Kimerling, Phys. Rev. B 27, 6216 (1983).
[15]M. Levinson, M. Stavola, J. L. Benton, and L. C. Kimerling, Phys. Rev. B 28, 5848 (1983).

[16]J. L. Benton and M. Levinson, in *Defects in Semiconductors II,* edited by S. Mahajan and J. W. Corbett (North-Holland, New York, 1983), p. 95.

[17]M. Stavola, M. Levinson, J. L. Benton, and L. C. Kimerling, Phys. Rev. B **30**, 832 (1984).

[18]M. Levinson, Phys. Rev. B **28**, 3660 (1983).

[19]G. E. Jellison, Jr., J. Appl. Phys. **53**, 5715 (1982).

[20]P. R. Tapster, M. S. Skolnick, R. G. Humphreys, P. J. Dean, B. Cockayne, and W. R. MacEwan, J. Phys. C **14**, 5069 (1981).

[21]M. Levinson (unpublished).

[22]It is noted that the kinetics of Fig. 2 do not explain the observed behavior of DLTS peak $B1$. It must be assumed that as an artifact of the DLTS measurement (but not for TSCAP measurements), in the temperature range where $B1$ appears, there is a second, more rapid mechanism of $B \rightarrow A$ transformation. When the DLTS spectrum is shifted to lower temperature by the use of a longer time constant ($\sim$50 ms), this second $B \rightarrow A$ mechanism also appears to limit the maximum height of peak $B2$.

[23]M. Levinson and M. Stavola, J. Electron. Mater. (to be published).