JPL Publication 00-06

An Introduction to Space Radiation Effects on Microelectronics

L. D. Edmonds
C. E. Barnes
L. Z. Scheick
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

May 2000

The work described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration, Code AE. This work was funded by the NASA Electronic Parts and Packaging Program (NEPP).

ABSTRACT

The use of microelectronic devices in spacecraft requires that these devices preserve their functionality in the radiation environment found in space. This introductory course presents a summary of the current understanding of the effects of radiation on microelectronic devices. Topics include an overview of the radiation environment and various types of effects (cumulative ionization, displacement damage, and single event effects) that the environment may have on each of the common device technologies. The interpretations and limitations of data obtained from standard test methods are also discussed. This course is intended for non-experts who are seeking explanations of terminology and fundamental concepts. The reader is expected to be acquainted with semiconductor devices on a level treated by introductory textbooks but is not required to have any prior knowledge of radiation effects.

Contents

Contents (Continued)

Figures

Tables

I. INTRODUCTION

The use of microelectronic devices in spacecraft requires that these devices preserve their functionality in the radiation environment found in space. This introductory course presents a summary of the current understanding of the effects of radiation on microelectronic devices. Topics include an overview of the radiation environment and various types of effects (cumulative ionization, displacement damage, and single event effects) that the environment may have on each of the common device technologies. The interpretations and limitations of data obtained from standard test methods are also discussed. This course is intended for non-experts who are seeking explanations of terminology and fundamental concepts. The reader is expected to be acquainted with semiconductor devices on a level treated by introductory textbooks, such as the book by Streetman [1.1], but is not required to have any prior knowledge of radiation effects.

The objective of this course is to provide the reader with an awareness and general knowledge of the degradation or loss of functionality that can occur in microelectronic devices exposed to the space radiation environment. Upon completion of this course, the reader should have an introductory-level understanding of the following:

- The radiation environments and basic effects (cumulative ionization, etc.)
- The meanings and effects of oxide trapped charge and interface states
- Enhanced low dose rate effects in bipolar linear circuits
- Multiple total dose failure modes in mixed signal devices
- Displacement damage effects
- Single event upsets in digital circuits
- Single event transients in analog circuits
- Single event latchups in various circuits
- Single event gate rupture in power MOSFETs and other devices
- Single event burnout in power MOSFETs and power bipolar devices
- Single event functionality interrupt in DRAMs and other devices
- Stuck bits in DRAMs
- Single event dielectric rupture in FPGAs
- Test methods (and limitations) for determining device radiation responses

References

[1.1] B.G. Streetman, *Solid State Electronic Devices*, second edition, Prentice-Hall, 1980.

2. RADIATION SOURCES

Because this course is intended primarily for NASA users, the focus will be on the natural space radiation environment encountered by the majority of NASA missions. This includes galactic cosmic rays (GCRs), particles emitted by solar events, and particles trapped in the Earth's radiation belts. A smaller number of missions may experience radiation from onboard radioisotopes and/or particles magnetically trapped around other planets (e.g., Jupiter). Apart from an occasional and brief mention, the latter environments are not discussed in this course. The neutron environment relevant to high-altitude avionics is also excluded from this discussion.

Describing the radiation environment in terms of flux or fluence versus energy for each type of particle is the most versatile (not always the most convenient) description because all radiation related quantities can be derived from this information. Flux through a surface is the rate that particles cross the surface divided by the area of the surface. An integral (in energy) flux refers to the set of particles having energies exceeding the specified value. A differential (in energy) flux is the negative (to produce a positive quantity) of the energy derivative of the integral flux. Conversely, the integral flux evaluated at a selected energy is obtained by integrating the differential flux from the selected energy to all higher energies. The selected energy is the lower integration limit, so we integrate the differential flux (a positive quantity) instead of the energy derivative of the integral flux (a negative quantity).

Directional flux is the flux, through a surface that is perpendicular to the specified direction, from trajectories within a small solid angle about that direction, divided by the solid angle. The omnidirectional flux is the directional flux integrated in solid angle over 4π steradians. If the flux is isotropic, the omnidirectional flux is simply the directional flux multiplied by 4π steradians. Whether the flux is isotropic or not, the rate that particles enter a spherical region (counting only entries and not exits, and assuming that the region does not contain radiation sources which contribute to the flux by supplying particles that exit without first entering) is the omnidirectional flux times the cross sectional area of the sphere. The cross sectional area is the area of a circle, which is one-fourth the surface area of the sphere. If the flux is isotropic, the rate that particles enter an arbitrary convex region (counting only entries and not exits and assuming no sources within) is the omnidirectional flux multiplied by one-fourth the surface area of the region. The latter statement can be derived from mathematical arguments showing that the directional average projected area, projected in each direction, is one-fourth the surface area. It can also be derived from a simpler physical argument which notes that (for an isotropic flux), the rate per area that particles cross a surface element (from one half-space to the other but not including both directions) is one-fourth the omnidirectional flux, and this quantity is multiplied by the surface area to obtain the rate that particles enter the region.

For example, suppose a particle detector produces a count when a particle enters some region. Suppose the region is a cube and each face has area A. The surface area is 6A, so the count rate in an isotropic flux is the omnidirectional flux multiplied by 3A/2.

Directional and omnidirectional fluxes apply to an arbitrary directional distribution, including a unidirectional particle beam. For a unidirectional beam, the directional flux contains a Dirac delta function describing the angular dependence. The flux is not isotropic, so earlier statements regarding arbitrary convex regions do not apply, but the earlier statement regarding spheres does apply. The omnidirectional flux is the rate that particles hit a sphere divided by the cross-sectional area. This is the same as the conventional flux measure for a unidirectional beam; the rate per area that particles cross a surface element perpendicular to the beam. The omnidirectional flux for a unidirectional beam is therefore the same as the conventional flux measure. The rate that particles hit a sphere is the same for any directional distribution having the same omnidirectional flux. Therefore, the rate that particles hit a sphere is the same for an isotropic environment as for a unidirectional beam, if the beam flux (rate per area that particles cross a surface element perpendicular to the beam) is equal to the omnidirectional flux for the isotropic environment.

Fluence associated with a specified time period is the time integral of flux, i.e., a number of particles (instead of rate) divided by area. Unless otherwise stated, flux and fluence are assumed in this course to be omnidirectional. A spectral shape can refer to either a differential or integral flux, but in either case it refers to the shape (independent of the normalization or flux units) of the plot of flux versus energy. One spectrum is said to be "harder" than another if it is more skewed towards the higher energies. A hard spectrum of either electrons, protons, or heavy ions (any ion with atomic number greater than one), is able to travel farther through matter than a soft spectrum.

Environments can be compared using flux or fluence terminology, but comparisons are sometimes also made in terms of total ionizing dose (TID), displacement damage, or single event upset (SEU) rates (from direct ionization or indirect ionization). These topics are discussed in Chapter 3. The present chapter assumes that the reader either already has some understanding of these topics, or will gain an understanding later after reading Chapter 3.

2.1 Trapped Particles

One component of the radiation environment important to Earth-orbiting systems consists of electrons and protons trapped in the Earth's magnetic field. There are also some trapped heavy ions, but their energies are low enough to be stopped by typical spacecraft shielding [2.1]. In contrast, the electrons and protons can penetrate spacecraft shielding and are usually the only trapped particles that are a concern for Earth orbits (although trapped heavy ions are very important to some Jupiter orbits).

The regions of trapped electrons and protons, sometimes called radiation belts and sometimes called the Van Allen belts, are most significant between the altitudes of

approximately 1,000 km to 32,000 km. Their extent is greater than this, but the particle flux is smaller outside this altitude range.

A low-order approximation for Earth's magnetic field is that from a slightly tilted (from Earth's axis of rotation) and offset (from Earth's center) magnetic dipole. This approximation completely fails in the outer regions of the magnetic field, which are influenced by interactions with the solar wind, but the approximation is adequate for some qualitative concepts. Charged particle motion in the magnetic field can be visualized as a spiral motion around a magnetic field line superimposed with a longitudinal motion that follows the field line. This longitudinal motion follows a magnetic field line until the particle reaches a certain "mirror point", at which time the longitudinal motion reverses and the particle comes back along the field line in the opposite direction; unless the mirror point is below the top of the atmosphere (with "top" loosely defined here), in which case the particle will be removed from the radiation belt by the atmosphere. Each particle having mirror points outside the atmosphere bounces back and forth between a northern and southern mirror point. The mirror points depend on particle trajectory characteristics that have a statistical distribution, so the mirror points themselves have a statistical distribution for a collection of many particles. This influences the variation of particle flux when comparing different locations on the same magnetic field line. However, there is a more dramatic variation in particle flux when comparing locations on different magnetic field lines having different altitudes at the magnetic equator. The physical explanation is complicated and not yet completely understood, but the end result is that a spacecraft environment is strongly dependent on altitude, even for equatorial orbits.

The altitude distribution is significantly different for protons than for electrons. The distribution of electrons, whose energies reach a maximum of about 7 MeV, shows two altitude peaks at about 4,000 km and 24,000 km, giving rise to two belts with a relatively unpopulated slot between them. In the case of protons, their location is restricted primarily to one belt, but their energies can be much higher - in excess of several hundred MeV. Depending on altitude, the flux of all electrons having energies greater than 100 KeV can exceed $10^8/cm^2$-s [2.1], and the flux of all protons having energies greater than 500 KeV can also exceed $10^8/cm^2$-s [2.1].

As previously stated, the dipole best representing Earth's magnetic field is tilted and offset from Earth's center. The result is a special geographic location, off the coast of Brazil and called the South Atlantic Anomaly (SAA), in which an intense radiation belt is encountered at a lower altitude than would be the case without an offset. This intense region reaches down to the upper part of the atmosphere. This is not an issue for orbits that would be in the intense radiation belts anyway (a very harsh environment), but it is an issue for spacecraft attempting to avoid harsh environments. A low-altitude orbit that would otherwise (without the offset) be very benign, by avoiding the radiation belts, is less benign because of periodic passes through the SAA. Most of the radiation effects seen by such orbits is from these passes through the SAA.

The radiation belts respond to solar activity. In addition to supplying particles that can populate the belts, solar activity can also affect the belts by producing changes in the geomagnetic field (magnetic storms), and by producing an atmospheric effect discussed below. Solar activity can create two types of time variations in the radiation belts. One type is random, in the sense of originating from events (e.g., solar particle events or magnetic storms) that cannot be predicted in advance. Some random variations are sporadic, and some others persist for months or longer. The other type of time variation is a systematic change between solar minimum and solar maximum. The outer electron belt shows large (order of magnitude) sporadic time variations [2.2]; large enough to mask any systematic variations that might otherwise be observed. Therefore the standard environmental models refer only to long-term (many-year) averages for these particles. The proton belt and inner electron belt also exhibit random variations, so standard environmental models refer only to time averages. However, the random variations for these particles are small enough for the systematic variations to be observable, so the standard models provide one time average for solar maximum conditions and another time average for solar minimum conditions. The systematic trend for the inner electron belt is to be more intense during solar maximum. The systematic trend for protons is significant only at the low altitudes, and the trend is to be less intense during solar maximum. This is because the increased solar energy output increases the atmospheric density at altitudes of 200 to 1000 km, which makes the atmosphere more effective at removing protons [2.2].

The trapped particle flux at a particular location is non-isotropic, particularly in low-altitude orbits. Measured data in the standard environmental models refer only to time averages of omnidirectional flux, but it is possible to predict the directionality of the flux by supplementing these data with physical analysis. The directionality is relevant if the spacecraft shielding is non-isotropic and the spacecraft maintains a fixed orientation relative to the environment. However, this level of detail involves complex calculations that are useful only if the spacecraft mass distribution and the orientation of the spacecraft relative to the environment are known. Such calculations are occasionally done, but the most common engineering practice is to assume that the environment is isotropic.

The energies of the trapped particles are low enough so that mass shielding thicknesses practical for spacecraft use can provide significant protection (particularly for the electrons). However, localized or "spot" shielding may be needed to obtain adequate protection without adding excessive weight to the spacecraft. Also, mass shielding produces "diminishing returns". While the first 100 mils of aluminum (for example), will substantially reduce the environment from the unshielded level, an additional 100 mils produces a much smaller additional reduction.

2.2 GCR

Galactic cosmic rays (GCRs) originate from outside the solar system and consist of about 85% protons, approximately 14% alpha particles, and about 1% heavier nuclei

[2.2]. GCRs contain the most energetic heavy ions and protons, reaching energies in excess of 10 GeV/nucleon. Plots or tabulations produced by standard models start at 10 MeV/nucleon and go up in energy from there. Although the GCRs are mostly protons, the proton component is rarely important to spacecraft electronics, excluding proton detectors. In terms of TID and displacement damage (discussed later in Sections 3.2 and 3.4), the effect of GCR protons is much less than that from solar event protons in interplanetary space, or trapped protons in an Earth orbit. In terms of SEU produced by indirect ionization (discussed later in Section 3.3), the relative abundance of GCR protons is not large enough to compensate for the small cross section for producing such an event, so GCR heavy ions dominate the GCR induced SEU rate. GCR protons are only important to devices that respond to the very lightly ionizing high-energy protons by direct ionization (e.g., a proton detector). In contrast, the heavy ions are an important source of SEU. In fact, the primary impact of GCR on spacecraft electronics is through heavy-ion induced single event effects.

The GCR environment in interplanetary space changes with the phase of the solar cycle. This is because solar activity creates magnetic disturbances in the heliosphere, and these disturbances deflect some GCR particles so that they do not enter the solar system. The result is that the GCR environment is milder during solar maximum than during solar minimum. This "solar modulation" removes the lower energy particles more effectively than the higher energy particles, so the environments are nearly the same at the higher energies (equal to or greater than a few GeV per nucleon), but the solar maximum environment is milder at the lower energies. This makes the solar maximum GCR environment the hardest (in terms of energy distribution) of all interplanetary space heavy ion environments. It is so hard that the mass shielding required to reduce this environment by a factor of two is about six inches of aluminum or equivalent. The solar minimum GCR environment is reduced by a larger factor, but this factor is bounded by the requirement that the solar minimum environment is at least as severe as the solar maximum environment.

The solar maximum and solar minimum GCR environments can be compared by comparing the SEU rates they produce in a device. The spectral shapes are different, so an exact comparison depends on the device, but typical ratios of rates are between three and five, with the smaller rate applying to solar maximum.

In addition to solar modulation, magnetic shielding is also provided by Earth's magnetic field for Earth orbiting spacecraft. This shielding deflects some particles so that they do not reach the orbit. As with the solar modulation, this shielding is more effective at removing the lower energy particles. The only heavy-ion environment harder than interplanetary solar maximum GCR is the GCR environment at a location heavily protected by a planetary magnetic field. It should be emphasized that "hard" in this context means that the spectrum is skewed towards high energies, which is relevant to the additional protection provided by mass shielding. This should not be interpreted to mean that a magnetically shielded environment is severe from the point of view of causing SEU. The highest energy particles are actually less able to produce SEUs. This is because there is a worst energy for a given ion moving through a given medium; the energy at

which the interaction with the medium is greatest (a measure of the interaction, called LET, is discussed later in Section 3.1). The energies of nearly all GCR ions are larger than this worst value, so the interaction decreases with increasing ion energy. Because the highest energy particles are less able to produce SEUs, magnetic shielding that removes the lower energy particles produces a comparatively benign environment. In fact, strong protection from a planetary magnetic field is probably the only case in which mass shielding can be detrimental. An extreme thickness capable of slowing down some of the GCR particles can make these particles more capable of causing an SEU.

In terms of orbit average flux, polar orbits are less protected by magnetic shielding than a low-inclination orbit at the same altitude. The orbit-average GCR environment for a low altitude (e.g., 1000 km) *polar* orbit can be compared to the interplanetary GCR environment by comparing the SEU rates they produce in a device. The spectral shapes are somewhat different, so an exact comparison depends on the device, but typical ratios of rates are between three and five, with the smaller rate applying to the orbit average. The spectral shape for a low-altitude *low-inclination* GCR environment is so radically different than that for interplanetary space that there are no "typical" factors that compare heavy-ion induced SEU rates. Rate ratios can be anywhere from roughly an order of magnitude (applicable to very susceptible, or soft, parts that will also upset by trapped protons not included in this discussion) to many orders of magnitude (applicable to harder parts).

The interplanetary GCR environment is not exactly homogeneous throughout the solar system, but the common engineering practice is to use the same models for all locations in interplanetary space. The community-accepted model is in a set of computer codes called CREME (Cosmic Ray Effects on MicroElectronics), with the most recent version called CREME96 [2.3]. This code can also include modulation by Earth's magnetic field. The GCR environment is isotropic in interplanetary space. It is not exactly isotropic in an Earth orbit, but the common engineering practice is to assume that it is.

2.3 The Anomalous Component

Another source of particles, composed of helium and some heavier ions with energies less than ~50 MeV/nucleon, was named the "anomalous component" because sometimes it is present and sometimes not. The anomalous component is believed to originate as neutral interstellar gas that diffuses into the heliosphere, becomes singly ionized by solar radiation or by charge exchange, and is then convected by the solar wind to the outer heliosphere [2.4]. The ions are then accelerated, probably at the solar wind termination shock, and propagate to the vicinity of the Earth [2.4]. These particles are seen at 1 AU only during solar minimum, but the details differ from one solar minimum to another [2.4].

The anomalous component is not a very severe environment compared to GCR in interplanetary space behind moderate or heavy spacecraft shielding, because shielding is effective against this relatively low-energy environment. However, the anomalous

component can be significant compared to GCR for spacecraft heavily protected by magnetic shielding (a low-altitude, low-inclination Earth orbit) but only lightly protected by mass shielding. The reason is that the anomalous component is only singly ionized and this tends to make it able to penetrate magnetic shielding. If mass shielding does not remove it, while magnetic shielding does remove most of the GCR environment, the anomalous component is a significant part of the surviving heavy ions. However, this is not an issue during the solar maximum time period when the anomalous component is not present.

2.4 Solar Event Particles

The sun occasionally emits bursts of energetic charged particles into interplanetary space. Physicists studying the mechanisms of particle acceleration are careful to distinguish solar flares from coronal mass ejections, but the semantics are not important to spacecraft designers that only care about the end result, so the phrase "solar flare" is frequently found in the literature discussing either kind of event. To avoid irritating some of the physicists, this course uses the phrase "solar event" for any burst of protons or heavy ions from the sun. Electrons are also emitted, but they are not important to any of the radiation effects discussed in this course, so the focus is on protons and heavy ions.

Large solar events are correlated with the eleven-year solar cycle, and are much more probable or frequent during solar maximum than during solar minimum. These events are not predictable, except on a statistical basis. The time profiles of solar event particles generally show a rapid rise (a few hours or less), followed by a slow decay (several hours to several days). The flux is approximately isotropic during the decay phase. Different events have different shapes (spectral shapes) and sizes, but the very large event that occurred in August 1972 is often used as a reference [2.5]. A more recent event in October 1989 produced a similar proton spectrum [2.1], which encourages us to continue to use the August 1972 event as a reference. The measured proton spectrum shows that the fluence consisting of all protons with energies exceeding 10 MeV, accumulated over the duration of the event, is about two hundred times the fluence of such particles accumulated over one year of solar minimum GCR. Stated another way, two hundred years of solar minimum GCR would be needed to accumulate the same fluence of such particles as produced by one of these events.

Unlike GCR protons, solar event protons can be important to ionizing dose, displacement damage, and single event effects. Given that an event occurs (even if it is not particularly large), the solar event protons are the primary contributions to ionizing dose and displacement damage for spacecraft in interplanetary space. Also, the relative abundance of protons to heavy ions is large enough to compensate for the small cross section for producing single event effects by indirect ionization, so solar event protons can be an important contribution to single event effects in proton-susceptible devices.

Heavy-ion measurements were (and still are) sparse, so it is necessary to use credible assumptions regarding heavy ions. Early models assumed that heavy ions have the same

spectral shape as the protons, and that the relative abundance of heavy ions to protons is comparable to that for GCR. However, predicted SEU rates did not agree with flight observations during recent solar events, so it became necessary to abandon the early models. There is not yet a consensus on heavy ion models, but the Jet Propulsion Laboratory has used a model that is documented in internal memos and frequently called the "design case flare" (DCF). The relative abundance of heavy ions to protons is on the order of 1/100 of the relative abundance for GCR, and the heavy ions have different (softer) spectral shapes than the protons. The softer spectral shapes imply that spacecraft shielding is much more effective at stopping heavy ions than protons. The result is that the relative abundance of heavy ions to protons (already small even without shielding) becomes progressively smaller with increasing spacecraft shielding. The exact ratio of DCF heavy-ion induced SEU rates to solar minimum GCR heavy-ion induced SEU rates depends, as usual, on the device, but typical numbers can be given. At 1 AU and behind a 100 mil aluminum shield, the number of heavy-ion induced SEUs accumulated over the duration of one DCF is equal to the number accumulated from a few (between one and ten) years of solar minimum GCR. The smaller numbers of equivalent GCR years apply to the harder devices.

Mass shielding is very effective against the heavy-ion component of solar events as described by the DCF model. As is usually the case, the exact comparison between SEU rates depends on the device, but typical ratios can be given for comparing thicknesses of centered aluminum spherical shields. The rate behind 60 mils is roughly one order of magnitude greater than the rate behind 100 mils, which in turn is roughly one order of magnitude greater than the rate behind 250 mils. Shielding is less effective against the more penetrating protons, but shield thicknesses practical for spacecraft use can still provide some relief, particularly if spot shielding is used, from events similar to the August 1972 event. However, some other events, such as the February 1956 event, are much harder and shielding is much less effective [2.5]. As with the trapped protons, shielding provides diminishing returns even when it is effective.

References

[2.1] J.L. Barth, "Modeling Space Radiation Environments," *1997 IEEE Nuclear and Space Radiation Effects Conference Short Course.*

[2.2] E. Stassinopoulos, "Radiation Environment of Space," *1990 IEEE Nuclear and Space Radiation Effects Conference Short Course.*

[2.3] A.J. Tylka, J.H. Adams, Jr., P.R. Boberg, W.F. Dietrich, E.O. Flueckiger, E.L. Petersen, M.A. Shea, D.F. Smart, and E.C. Smith, "CREME96: A Revision of the Cosmic Ray Effects on Micro-Electronics Code," presented at the IEEE/NSREC Conference in July 1997.

[2.4] J. Feynman and S.B. Gabriel, "High-Energy Charged Particles in Space at One Astronomical Unit," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 2, pp. 344-352, April 1996.

[2.5] J.H. Adams, Jr., and A. Gelman, "The Effects of Solar Flares on Single Event Upset," *IEEE Trans. Nucl. Sci.*, vol. 31, no. 6, pp. 1212-1216, Dec.1984.

3. BASIC RADIATION EFFECTS AND ENVIRONMENTAL MEASURES

The environment at a given time and location can be described without any reference to radiation effects if the environment is described with the greatest possible detail; flux versus energy (and direction if relevant) for each and every particle species in the environment. However, this level of detail is not always the most convenient way to describe or measure the environment. Other measures (e.g., dose, which will be defined later) are much more convenient when relevant, but understanding the meanings of such measures requires an introductory-level understanding of radiation effects. Also, some understanding of radiation effects is needed in order to understand which environmental components (e.g., protons in Earth's radiation belts versus GCR heavy ions) are the greatest concern for a given reliability issue. Radiation effects are discussed in this section in only enough detail to explain why various environmental measures are useful. Radiation effects will be discussed in more detail in the more specialized topics in Chapter 5.

The radiation effects discussed in this course can be placed in one of three broad categories; cumulative ionization, single event effects, and displacement damage. There is a convenient environmental measure for each of these categories, and each measure is discussed separately below, after some general discussion in Section 3.1. It should be noted that the measures discussed below are not always useful substitutes for flux versus energy spectra. For example, none of these measures are useful for describing dielectric charging produced by secondary electron emission. However, these measures are useful for the radiation effects discussed in this course.

It should be noted that a given particle sometimes creates several effects in a device. A notable example is protons making significant contributions to both cumulative ionization and displacement damage. This has very important consequences in some practical situations. However, the different effects can be discussed separately because they are distinguished in terms of the physical mechanisms involved.

3.1 Interaction of Ionizing Radiation with Matter

Photons, electrons, protons, and heavy ions are all able to produce ionization when they travel through matter. Photons are discussed separately at the end of this section because they have some unique properties. Until photons are specifically mentioned, "ionizing particle" will refer to electrons, protons, or heavy ions.

The energy loss of an ionizing particle as it travels through matter can be divided into two components: electronic energy loss and nuclear energy loss. The electronic energy loss is caused by interactions with the electrons in the atoms of the medium, and the nuclear energy loss is caused by interactions with the nuclei in the atoms of the medium. The latter interactions are usually through Coulomb forces. An incident proton

occasionally interacts with a nucleus through nuclear forces (hard collisions) producing nuclear fragments, but these events are relatively infrequent and therefore considered separately from the other interactions that produce a more systematic slowing down of the incident particle. Another reason for placing the hard collisions in a separate category is that they can also be produced by neutrons, which are not classified as ionizing radiation.

Except for ionizing particles near the end of their range, nearly all of the energy loss is electronic. The deposited energy accumulated over the particle trajectory is nearly all electronic if the incident energy is sufficiently large, which is almost always the case in practical applications. Therefore the total energy loss provides a good approximation for the electronic energy loss. The total energy loss per unit distance of travel is called the linear energy transfer, or LET. The LET is usually normalized by dividing by the density of the medium, and the most popular units are MeV-cm^2/mg. The reason for this normalization is that it makes LET for a given particle and energy more nearly the same in different materials (still not exactly the same, but more nearly the same). LET depends also on the incident particle species and energy. Plotting LET against energy shows a peak, so LET increases with energy when the energy is sufficiently small, and decreases with energy at higher energies.

Some of the electronic energy loss in a semiconductor or insulator results in ionization, i.e., the creation of electron-hole (e-h) pairs. What happens to these e-h pairs after they are created is discussed in the sections that follow. The present discussion is only concerned with the initial creation. Most of these e-h pairs are created in the second step of a two-step process. The first step is the transfer of energy from the incident particle into high-energy primary knock-on electrons called delta rays. This first step controls the energy loss, or LET, of the incident particle, but the second step strongly influences the final ionization. In the second step, the delta rays, which are themselves ionizing particles, deposit their energy in a variety of ways, including phonons (heat), excitation, and the creation of many e-h pairs by each delta ray. Although the details are quite complex, there is a simple end result in that the number of e-h pairs created is proportional to the electronic energy loss by the incident particle. For silicon, the proportionality constant is 1/3.6eV, i.e., the number of liberated e-h pairs is the deposited electronic energy divided by 3.6 eV. This is the same conversion that would be used for the hypothetical case in which all deposited energy is converted to e-h pairs, but the energy required to create an e-h pair is 3.6 eV. Similar considerations apply to SiO_2, except that we use 18 eV instead of 3.6 eV [3.1].

A convenient memory aid for silicon, using the 3.6 eV conversion, is that an LET of 97 (approximately 100) MeV-cm^2/mg liberates 1 pC/μm when each e-h pair is counted as one elementary charge.

A few points regarding the influence of spacecraft shielding on particle spectra are worth noting. Shielding tends to modify a particle spectrum in such a way that the shielded spectrum displays a "replacement effect". If the shielded spectrum is transported through a thin (compared to the original shield) layer of additional matter, this additional

matter will stop some low energy particles but also slow down some higher energy particles to replace those that were stopped. The result is that the spectrum behind the additional matter is nearly the same as the spectrum in front of it. This becomes intuitively obvious if we note that the effect of the additional matter is equivalent to making a minor change in the original shield thickness, which produces only a minor change in the spectrum. Another point is that the modification produced by shielding on a particle spectrum also tends to remove the very low energy particles. The reason is that the initial particle energy would have to be very precise in order for the particle to go through a thick shield and emerge with almost no energy left. Very few particles have initial energies in such a narrow range. The scarcity of very low energy particles and the replacement effect each tend to produce a spatially uniform spectrum in a small object, such as a microelectronic device, protected by spacecraft shielding. The important conclusion from this paragraph is that particle spectra tend to be spatially uniform within a microelectronic device protected by spacecraft shielding.

We now consider photons, which are distinguished from the particles previously discussed because deposited energy is due to photons being absorbed rather than slowed down. Therefore deposited energy per unit distance is not calculated from LET. Instead, it is calculated, for a photon beam, as being proportional to the beam intensity at the depth in question multiplied by a quantity called an absorption coefficient, which is a measure of the probability of photon absorption. The absorption coefficient is a function of the medium and photon wavelength. Although deposited energy is calculated differently for photons, the number of liberated e-h pairs is calculated from this deposited energy in the same way as for the particles previously discussed; subject to the requirement that the photons are high-energy (e.g., gamma rays). This is because high-energy photons create high-energy delta rays, which then behave like the delta rays created by any other incident particle. Therefore the conversion factors (3.6 eV for Si or 18 eV for SiO_2) also apply to high-energy photons. However, a different simplification applies to the opposite extreme of low energy primary knock-on electrons, such as created by photons with energies not much greater than the material bandgap energy. This is relevant to illuminated solar cells and to laboratory experiments using a laser for an ionization source. In this case, the primary knock-on electrons do not create additional e-h pairs, so the total number of e-h pairs is the number of primary knock-ons. This is the number of absorbed photons, which is the deposited energy divided by the photon energy. Therefore, for this opposite extreme, the conversion factor is the photon energy instead of 3.6 eV (for Si) or 18 eV (for SiO_2). An equivalent conversion, which refers to numbers of photons instead of deposited energy, is one e-h pair for each absorbed photon.

An interaction not previously discussed is the radiative scattering of a high-energy incident electron in a material, producing a high-energy "bremsstrahlung" photon. Such events are sufficiently infrequent so that they can be ignored when the objective is to estimate electronic energy deposition at a given location in a material when the environment at that location is given. However, these events cannot always be ignored when the objective is to estimate the environment at a given location. An example is provided by electrons incident on a thick shield. Bremsstrahlung photons are very penetrating and are therefore additive in the sense that photons created anywhere within

the shield contribute to the environment emerging from the shield. If the shield is thick enough to stop most of the incident electrons, the bremsstrahlung photons can be the dominant environment emerging from the shield. Note that the bremsstrahlung photons emerging from the shield are not significant (in terms of radiation effects) compared to the incident electron spectrum, but can still be an important environment behind a shield if the electron flux in front of the shield is large enough.

3.2 Cumulative Ionization

When a photon or charged particle travels through a material, it interacts with electrons in the material and causes some of the atoms to become ionized, creating e-h pairs. This occurs in both semiconductors and insulators, but one distinction between the two types of materials is in terms of whether this effect accumulates or quickly dissipates. The present discussion will focus on insulators (e.g., a gate oxide in a CMOS device), which exhibit a cumulative effect. The radiation response of oxides in semiconductor devices is a very complex subject that is still an area of active research, but the present objective is merely to understand why a particular type of environmental description is useful. This objective can be met by considering a few basic concepts applicable to charge trapping in the oxide interior.

The electrons liberated by an ionizing particle in an oxide are sufficiently mobile so that they move in response to any electric field present in the oxide. Biasing voltages can induce such an electric field, but even without biasing voltages there are built-in fields that can move electrons out of the oxide. In contrast, the holes are very much less mobile, and most that survive recombination with the electrons become trapped and are left behind after the electrons are driven out. The result is a net positive trapped charge in the oxide interior. This trapped charge can affect device performance, e.g., create a threshold voltage shift in MOSFET devices. The trapped holes can eventually become neutralized (the oxide anneals), but the anneal time can be very long; possibly years at room temperature, depending on the electric field during annealing and how complete the annealing is. For time scales shorter than the anneal time, trapped charges accumulate while the irradiation continues. Therefore, one useful environmental measure is one that will describe the ionization accumulated within the oxide over a given period of time.

Temporarily ignoring some complications (discussed shortly), the accumulated trapped charge is measured by the accumulated ionization, which in turn is measured by the sum (over particles) of the energy lost by the particles to the material via interactions with the electrons. Therefore a useful measure is the total energy, per unit mass of material, transferred to the material via ionization from all ionizing particles, which is called the total ionizing dose (TID). This dose is typically measured in rads, with one rad defined as 100 ergs of deposited energy per gram of material.

One complication that was temporarily ignored is recombination between holes and electrons immediately after they were liberated. This recombination, which occurs simultaneously with electrons being driven out of the insulator, profoundly influences the

amount of trapped charge. In some cases, nearly all of the charge recombines or is collected at contacts in such a way that there is little effect of the ionizing radiation of the device. This is usually the case with III-V material-based technologies such as GaAs laser diodes, LEDS, and MMIC devices. For other cases, such as a CMOS gate oxide, the trapped charge is more important, but as previously stated, the amount is profoundly influenced by recombination. This is measured by a "yield function", defined to be the fraction of e-h pairs that survive the initial recombination. One factor influencing recombination (hence the yield function) is the biasing voltage, because the electric field strength influences how quickly the electrons are driven out, and this determines how much time the electrons and holes are near each other. Consequently, the yield function is a function of the electric field in the oxide, which is why it is called a function. However, this is a device characteristic, not an environmental characteristic, and can be described by saying that device susceptibility to TID is influenced by the biasing voltage. What *is* an environmental characteristic is that recombination is also influenced by the type of radiation, because the spatial distribution of e-h pairs can be different for different types of radiation (e.g., electrons versus heavy ions).

For a sufficiently small e-h pair density, the electron and hole in a given e-h pair will see only each other instead of seeing other e-h pairs. Therefore the fractional recombination loss is insensitive to the e-h pair density. Recombination under these low-density conditions is called geminate. The recombination is called columnar when the density is high enough for e-h pairs to see other e-h pairs. The higher density not only leads to greater recombination loss, but also makes the situation more complicated, because the recombination loss becomes strongly dependent on the density. The same deposited energy can produce different device effects depending on whether the radiation source induces geminate or columnar recombination. Consequently, the yield function can be a different function for different types of irradiation. In particular, the yield is very different (much smaller) for heavy ions (columnar recombination) than for electrons or high-energy protons (geminate recombination). Fortunately, heavy ions are not an important dose contributor in space environments (excluding heavy ions trapped in Jupiter's radiation belts, and also excluding microdose discussed later). The most important dose contributors are typically electrons, high-energy protons, and gamma rays (bremsstrahlung or from radioactive generators if present) and are sufficiently similar with regards to recombination (because it is geminate) so that "a rad is a rad". In other words, x number of rads from high-energy protons is (at least roughly) equivalent (in terms of TID effects) to x number of rads from electrons, which is also equivalent to a total of x number of rads from a mixture of the particle types. A set of spectral curves (fluence versus energy for each particle species) can therefore be reduced to a single number; the number of rads.

Readers that intend to apply concepts discussed in this course to other areas of research should be aware that the above assertion, that "a rad is a rad", has limited applications, even when comparing electrons to high-energy protons. Previous discussions ignored the fact that electrons and protons have different masses. Coulomb collision theory shows that this difference implies that the directional distribution of knock-on electrons is different for incident protons than for incident electrons (in terms of dose effects, incident

electrons have more in common with gamma rays than with protons). This difference does not appear to have important consequences when the subject of interest is microelectronic devices, and the above assertion appears to be an adequate approximation for this application when comparing electrons to high-energy protons. However, the above difference does have important consequences in some other areas of research, such as dose effects in some biological samples [3.2]. Readers that may enter such areas of research should become aware of the limitations of the above assertion.

Another complication that was temporarily ignored is the time dependence of device responses. Even when the time required for a nearly complete anneal is very long, the time required for a partial annealing is still finite, so the degradation of a device at a given time after accumulating a given dose can be different for different time profiles of dose accumulation. Furthermore, device characteristics can continue to change in a complex way (not explainable from trapped-hole neutralization alone) long after the irradiation has stopped. Also, a phenomenon called "enhanced low dose rate effects" (discussed in Chapters 4 and 5) can result in different dose rates producing profoundly different device responses. It may therefore be relevant whether a given dose was produced by one or a few short bursts (e.g., from solar events), or from a roughly uniform distribution of many short bursts (e.g., from periodic passes through the SAA), or from a nearly constant dose rate (e.g., from man-made sources or from trapped particles in some Earth orbits). Ideally, a dose estimate would be presented as a function of time. However, one objective of test methods for measuring device susceptibility to TID (a sophisticated, but still evolving subject) is to predict device response for either the time profile of the expected environment, or for a worst-case time profile, so that dose accumulated over the mission duration becomes an adequate environmental description for TID.

An issue not yet considered is that dose deposition may appear homogeneous or not, depending on the level of magnification used to view the deposition. A high level of magnification may see statistical variations that are averaged out in lower levels of magnification. In particular, if gates or other susceptible structures in a device are sufficiently small, the dose averaged over the structure can differ from one structure to another. This is less of a concern when a structure requires many particle hits for significant degradation, because statistical fluctuations are smaller (when measured as a fraction or percent) for larger numbers. For this reason, such fluctuations are not normally considered to be important for electrons or protons in space, or for gamma rays used in the laboratory. The exceptional case is heavy ions. Some devices contain numerous structures that are individually so small that one or a few heavy ion hits to the same structure can cause significant degradation in that structure. TID from heavy ions that is very small when averaged over the entire device can mean that most structures were never hit, but the small fraction that were hit saw the equivalent of a severe TID environment. This effect has been called microdose. While heavy ions are not normally important to TID in most space environments, they are important to devices sufficiently miniaturized to be susceptible to microdose effects [3.3, 3.4].

The discussion of dose calculations is deferred to the next section, which explains LET spectra.

3.3 Single Event Effects

A single event effect (SEE) occurs when a single energetic particle is capable of creating an observable effect in a device. Some laboratory experiments use a short (in time duration) and spatially narrow laser pulse to mimic an incident particle. If we regard the entire pulse (as opposed to an individual photon contained in it) as a single entity, or "particle", then this is also under the category of SEE. Excluding the microdose effect discussed in the previous section (a hybrid between TID and SEE), all other presently known types of SEE differ from TID not only in that a single particle is important, but also in that disturbances created by different particles are not added together to create an event. For TID, even the most lightly ionizing particles are important if they are sufficiently abundant to collectively produce an effect, and contributions are summed over all ionizing particles. In contrast, for SEE (excluding microdose), only the particles individually able to produce an event are relevant, and contributions from different particles are not summed to determine whether the conditions for an event are satisfied. A useful environmental measure for SEE is one that quantifies the number of particles individually able to produce an event. Such a measure requires that the relevant particle characteristics be identified, which in turn requires an understanding of some basic concepts pertinent to SEE.

One category of SEE involves the destruction or degradation of a gate oxide or other dielectric. The relevant particle characteristics for this category is still an area of some controversy, so we will focus on another category, which involves charge collection following the liberation of mobile e-h pairs by an energetic particle in a semiconductor. This category can be further divided into direct ionization, indirect ionization, or a combination of both. Direct ionization (always the dominant process for heavy ions) occurs when the incident particle creates e-h pairs. Indirect ionization occurs when the incident particle (usually a proton or neutron) produces an energetic recoil particle, a fragment of (or the entire) nucleus from a target atom, and the latter particle creates e-h pairs. This differs from displacement damage discussed in the next section in terms of the energies involved. The process discussed here involves hard collisions (nuclear interactions), produces high-energy (and highly ionizing) recoils, but also requires that incident proton energies be large. Required proton energies start at about 7 MeV, while 100 MeV is much more effective. The most useful environmental description for the indirect process is flux or fluence versus energy for each particle species capable of producing this effect. From this point on, all discussions will refer to charge collection induced by heavy ions (i.e., direct ionization) in a semiconductor.

The most familiar example is single event upset (SEU) in digital devices (e.g., SRAMs) from heavy ions. However, even SRAMs (discussed later in Section 5.6) are unnecessarily complicated for the present objective, which is merely to understand why a particular environmental measure is useful, so we consider the simplest arrangement that illustrates some basic concepts that are also relevant to practical devices. The simplest arrangement consists of a reverse-biased p-n junction silicon diode connected to a monitoring circuit, with this circuit designed to produce a count (record an upset) when charge collected by the diode (time integral of the current through the diode) following an

ion hit to the diode exceeds some critical value Q_c. The critical charge Q_c for a practical device is more precisely defined for some cases than others, depending on how various time constants compare. When it is defined, it is a property of the device and biasing conditions.

The first step leading to SEU is the liberation of e-h pairs, which is controlled by ion LET (as discussed in Section 3.1). The column of e-h pairs, which are mobile carriers, is called the ion track. After the initial track formation, the next step is charge collection from these mobile carriers. The process is qualitatively similar to the photocurrents produced in an illuminated solar cell under either short-circuit or reverse-bias conditions. There are quantitative differences because charge collection from an ion track is under high-density conditions (the carrier density exceeds the doping density), but, like the solar cell, collected charge includes charge liberated both within and below the depletion region. Early investigations attributed charge collection from below the depletion region to a poorly defined object called a "funnel", which was often described as a funnel-shaped extension of the depletion region; a strong-field drift region that promptly collects all charge contained within by drift. The phrase "funnel length" (the length of a funnel) is still prevalent in the literature, even though it is now known that regions identified as funnels in the literature are actually regions where the electric field is weakest. Instead of a strong-field drift region, a better description is a weak-field ambipolar diffusion region. This led to a different description of the charge-collection process [3.5-3.7], and a funnel length is now regarded as a "quasi-physical 'fudge factor' in error rate calculations" [3.8].

A correct understanding of charge collection physics may, someday, reduce the error bars associated with angular effects discussed later. But until that day comes, it is enough to know that device susceptibility to heavy-ion induced SEU, measured as a cross section, is determined by the ion LET, subject to the important qualification that the ions have adequate range. Charge collection has not yet been found to be sensitive to the initial radial profile of the ion track. This indicates that most e-h pair recombination occurs at times late enough so that radial diffusion has erased most memory of the initial profile (a characteristic of radial diffusion is that the final radial profile is insensitive to the initial profile after a sufficiently long time). Note that this is very different than recombination in an oxide, which is important at the earliest times. Therefore it is not necessary to introduce a parameter (except possibly to describe second-order effects) analogous to the yield function used for oxides. However, track length is important. Ions having the same average LET within some device region can produce different device responses unless they all have a sufficiently long range. The vast majority of heavy ions found in space do have a sufficiently long range, but this is not always true for ions used in laboratory experiments. Ion range is an important issue in laboratory experiments. Desired ranges are at least 40 μm, and longer is better.

The importance of ion range was illustrated by an experiment using fission fragments produced by Cf^{252} to induce SEUs in a CMOS SRAM containing an epitaxial (epi) layer. Time of flight measurements identified the fission fragments and eliminated the problem associated with the fact that there is a mixture of fission fragment species and energies. Also, the energy deposited by a given particle in the epi layer was carefully calculated.

This calculation was intended to eliminate the problem associated with the fact that the particles have short ranges (about 18 μm for one major group of particles), and therefore a non-uniform LET. In spite of the care that went into this experiment, test results still did not agree with results obtained from longer-range ions from an accelerator [3.9]. The device was much less susceptible to the fission fragments than to accelerator ions that deposit the same energy in the epi layer. It was later discovered [3.10] that even devices containing an epi layer can be influenced by disturbances below the epi, so ion range is relevant even for these devices.

However, given that all ions considered do have adequate range, it is not necessary to distinguish between different ions having the same LET. All such ions can be added together in a single integral LET flux, which is the flux consisting of all ions having an LET that exceeds a specified value. A large set of spectral curves (flux versus energy for each heavy-ion species) can therefore be reduced to a single curve; flux versus LET. This curve is one environmental measure useful for heavy-ion induced SEU. Another useful measure is effective flux, which is discussed below after some discussions of cross section.

Device SEE susceptibility from a specified type of particle is usually described in terms of a cross section. The device cross section for the specified particle type is defined by the property that this cross section, multiplied by the flux of such particles, produces the statistical average SEE rate produced by that flux. For heavy-ion-induced SEU, the particle type is specified by LET, and the cross section is a function of LET. The cross section at a given LET is the area of the portion of the device that ions having that LET must hit in order to produce an SEU. The set of hit locations leading to SEU is larger for the higher-LET ions, so the cross section increases with increasing LET. However, the cross section is also a function of the direction of ion trajectories relative to the device normal.

An omnidirectional flux versus LET spectrum would be the most convenient environmental description regarding SEU by direct ionization if the heavy-ion environment is isotropic (which is usually an adequate approximation) and if device susceptibility data were presented as a directional average of the directional cross section versus ion LET. However, device susceptibility data are rarely presented this way. Such data are usually presented as a cross section versus LET curve without any reference to the direction of ion trajectory relative to the device, but it is not a directional average cross section. It is therefore not a surprise that nearly everyone makes the same mistake when looking at a cross section curve for the first time. The mistake is to assume that either there is no directional dependence, or there is but the cross section is a directional average, so SEU rates are calculated by integrating the cross section curve with an omnidirectional LET flux curve. In reality, SEU cross section curves traditionally represent device susceptibility at normal incidence. Even when tests are performed at angles, the data are processed in such a way as to describe device susceptibility at normal incidence (at least this is the intention, but there is sometimes some error in this). The unfortunate reality is that a cross section curve, as traditionally presented, does not

completely characterize a device. The directional dependence of device susceptibility is a separate and independent characterization that must supplement the cross section curve.

This characterization often contains a great deal of uncertainty, but a simple relationship, called the "cosine law" is often observed if the angle between ion trajectory and device normal is not too large. A mathematical statement of the cosine law is given in Section 7.2, but a simple illustration is adequate for the present discussion. For this illustration we consider a hypothetical device having a normal-incident cross section curve described by just two parameters; a threshold LET L_{th} and a saturation cross section σ_0. Normal-incident ions having LET less than L_{th} do not cause upsets, while all normal-incident ions having LET greater than L_{th} see the same cross section σ_0 for causing upsets. The cosine law states that for ions that hit the device at an angle θ measured from the device normal, the threshold value of the ion LET is $L_{th}\cos\theta$, and the saturation cross section is $\sigma_0\cos\theta$. In other words, ions having a smaller LET are capable of causing upsets when the angle is larger, but the cross section is smaller.

The importance of angular effects can be seen from a numerical example, which compares two hypothetical devices having the same characteristics at normal incidence. One device is isotropic (the threshold LET and saturation cross section are the same for all angles) while the other satisfies the cosine law. The LET spectrum selected for this example is shown in Fig. 3.1, which refers to solar maximum GCR in interplanetary space. Iron has a special significance because it is the most abundant of the very heavy ions, and the first rapid dip in the flux curve (at LET≈25 MeV-cm^2/mg) is due to the inclusion of iron at lower LETs and exclusion of iron at higher LETs. Suppose both devices have a normal-incident threshold LET equal to 40 MeV-cm^2/mg and a normal-incident saturation cross section equal to some arbitrary σ_0. Now compare the two device responses to particles with trajectories within a solid angle increment centered about a 60° angle from the device normal. For the isotropic device, the cross section seen by these particles is σ_0, and the relevant flux is the flux evaluated at LET=40 (scaled by the size of the solid angle increment), which excludes iron. For the cosine-law device, the cross section seen by these particles is only half of σ_0, but the relevant flux is the flux evaluated at LET=20 (scaled by the size of the solid angle increment), which includes iron. This flux is about two thousand times the flux applicable to the first device, which more than compensates for the factor of two reduction in cross section. When all angles are considered, the SEU rates for these two devices differ by several orders of magnitude. It is therefore essential that angular effects be included.

One way to include the directional dependence of device susceptibility (when known or assumed) is to include this in the cross section to obtain a directional-average cross section, which is used with the ordinary LET flux for SEU rate predictions. Although a reasonable approach, device data are not traditionally presented this way. Another approach is to include device directionality (when known or assumed) in the flux to obtain an effective flux, which is used with normal-incident device cross sections (the traditional cross section curves) for rate predictions. Two examples of effective flux are shown in Fig. 3.2 (calculated from methods discussed in Section 8.2). One curve is the effective flux appropriate for isotropic devices, which is the same as the ordinary LET

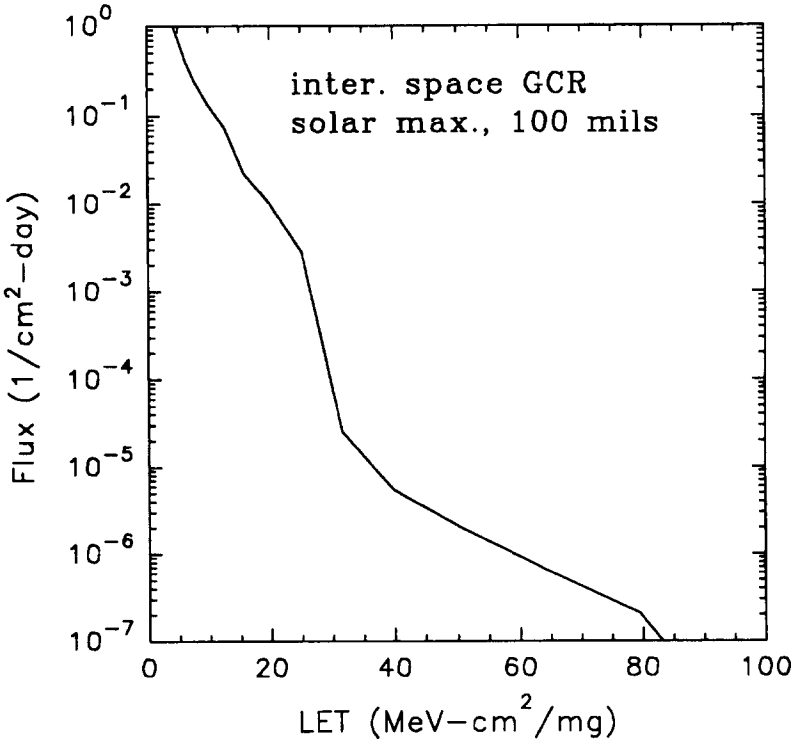

Fig. 3.1: The LET spectrum for solar maximum GCR in interplanetary space.

flux. The other curve is the effective flux appropriate for cosine-law devices. SEU rates for the hypothetical devices previously considered are determined by evaluating the appropriate effective flux (isotropic for the isotropic device and cosine law for the cosine-law device) at the normal-incident threshold LET, which is 40. This is off-scale in Fig. 3.2 for the isotropic device, but the same curve is in Fig. 3.1 and the flux is about 5×10^{-6} /cm^2-day. For the cosine-law device, the effective flux is about 7×10^{-2}/cm^2-day. The SEU rate for each device is determined by multiplying the appropriate effective flux by the same normal-incident cross section σ_0. Note that the two device rates differ by four orders of magnitude.

The cosine law is an adequate approximation for calculating SEU rates for some real devices, while some other real devices are very nearly isotropic, but most devices are somewhere between these cases. Additional effective flux options applicable to other cases are discussed in Section 8.2. However, the two cases shown in Fig. 3.2 give an idea of the error bars (measured by the spread between the curves) encountered when the directional dependence of device susceptibility is completely unknown. SEU rate estimates for soft devices (small threshold LET) are relatively insensitive to the assumed directional dependence of device susceptibility. For the harder devices, uncertainties in the directional dependence of device susceptibility become large uncertainties in rate predictions. Incidentally, the relative insensitivity for the softer devices helps to explain why SEU rate estimates reported in the published literature almost always agree with reported flight observations (although the general tendency for published work to report agreement between predictions and observations may also have other explanations). Meaningful comparisons can only be made if the number of detected SEUs in flight is

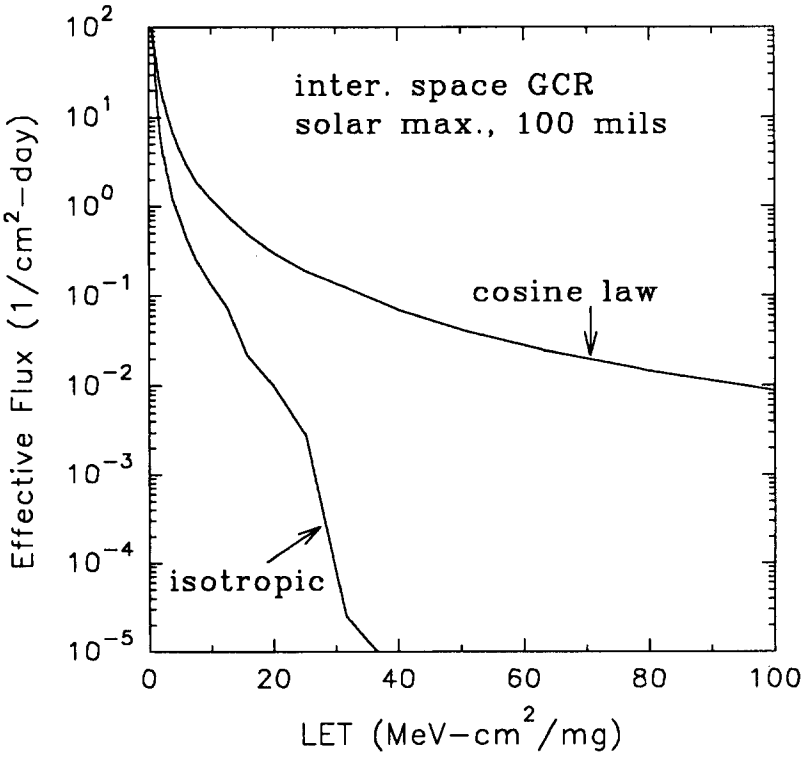

Fig. 3.2: Two examples of effective flux curves corresponding to the LET spectrum in Fig. 3.1.

large enough to be statistically significant. This generally means that the device is soft, which in turn means that predicted rates are insensitive to assumptions regarding the directional dependence of device susceptibility.

Depending on the application, flux or fluence (or both) could be relevant. For example, if a particular type of SEE is destructive, and if the objective is to estimate the probability that the event will occur sometime during a mission, the relevant environmental description is mission integrated fluence (scaled by a duty factor if the device is not susceptible all of the time). In contrast, if a device is protected by an error detection and correction circuit, so that an observable SEU requires two upsets within a common word and a common time window, peak flux is important for determining the event probability. Sometimes a device is part of a science instrument that is not required to operate during a solar event, and the only concern is how much noise (SEUs) will be produced by the GCR environment, or by trapped protons via indirect ionization. Therefore, each environmental component having a distinctly different time profile (e.g., solar events versus GCR) should be individually characterized in an environmental description.

As stated earlier, a primary difference between TID and SEE is that one is cumulative (sums over particles) while the other is not, but one thing they have in common is that they both result from energy transferred from an incident particle into ionization. Because dose and LET both refer to ionizing energy loss, we can calculate proton or heavy ion

dose from omnidirectional LET spectra. Dose from protons should be listed separately from dose from heavy ions because the yield functions are different (i.e., the same dose from the different particles can produce different device responses), but dose is calculated the same way for all. For each LET fluence spectrum, with LET referring to the material in which dose is to be calculated, an integration that sums deposited energy from all particles in the spectrum produces dose. However, we cannot go the other way because an LET spectrum (a curve) contains more information than dose (a number). Dose does not distinguish a large number of low LET particles from a smaller number of higher LET particles, so we cannot calculate an LET spectrum from it.

3.4 Displacement Damage

Displacement damage effects occur in semiconductor materials due to scattering interactions of incident particles, such as protons, neutrons, and even high-energy electrons, with the atoms of the semiconductor lattice. In the initial scattering event, the bombarding particle displaces an atom from its lattice site, and this primary knock-on can produce an additional cascade of displacements, the magnitude of which depends on the amount of kinetic energy transferred to the primary knock-on. For protons and heavy ions (including knock-ons), the energy transfer is usually by Coulomb interactions with the target nucleus. This is associated with the nuclear energy loss discussed in Section 3.1, and is distinguished from the hard collisions (producing nuclear fragments) in terms of frequency of occurrence and the energies involved. Compared to the hard collisions, these softer interactions are much more frequent, the secondary particles have much less energy (not enough to produce an ionized track leading to SEU), and the required energy of the incident particle is also much less. The energy transfer from a proton or heavy ion to a knock-on via this type of interaction is greatest when the incident particle energy is low in the sense that the particle is near the end of its range. Therefore the production of displacements is greatest near the end of range for such particles.

For a low-energy monoenergetic proton beam (such as might be used in a laboratory experiment), the displacement damage is much greater at some depths within the device than at other depths. For electrons, the energy transfer is still by Coulomb interactions, but the transfer is greater at higher electron energies. These are the energies at which electrons are penetrating enough to produce a nearly uniform spectrum throughout a microelectronic device, so the displacement damage is fairly uniform throughout a device region consisting of a homogeneous material. Neutrons are also very penetrating, so displacement damage produced by these particles also tends to be insensitive to depth. Note that neutrons interact via nuclear forces, and earlier discussions of protons interacting via nuclear forces (hard collisions) might give the impression that these are high-energy interactions leading to highly ionizing reaction products. A distinction is that protons have to overcome a Coulomb barrier before interacting via nuclear forces, requiring a high-energy incident particle and leading to high-energy interactions. While high-energy interactions occasionally occur from neutrons (and can cause a SEE), most interactions involve low energy transfers, producing displacement damage without significant ionization. This is useful for laboratory work because it is possible to use

neutrons to induce displacement damage in semiconductors with only minimal TID induced in nearby insulators. In contrast, ionization by protons (even without hard collisions) can significantly contribute to TID in nearby insulators.

The first result from displacement damage is that interstitials (displaced silicon atoms) and vacancies are created in the semiconductor. The details that follow are complicated because the interstitials and vacancies are initially dynamic (they can move, recombine, or produce stable defects), but for this overview it is enough to know that the end result is the creation of carrier traps and recombination centers (discussed in more detail in Chapter 4). This reduces minority carrier lifetime in a semiconductor. Displacement damage in sufficient concentration can also produce other effects, such as a reduced carrier mobility or compensation of donors or acceptors.

A reduced lifetime removes some carriers, via recombination, that would otherwise contribute to a desired current in a device. This reduces, for example, the short circuit current in an illuminated solar cell. Also, non-radiative recombination removes some carriers that would otherwise participate in a desired (for light emitting diodes) radiative recombination. Note that carrier recombination acts in reverse, i.e., becomes carrier generation, when the carrier density is less than the equilibrium density. This occurs, for example, in a reverse-biased depletion region, leading to an increased leakage current in a reverse-biased diode. Displacement damage can alter the electrical or optical properties of a variety of semiconductor devices, such as charge-coupled devices, photodetectors, light-emitting diodes, optocouplers, solar cells, and high-precision linear devices.

One environmental measure related to displacement damage is an equivalent fluence. The concept is easily understood by imagining a laboratory experiment. In this imagined experiment we irradiate a particular device with 500 KeV protons and an identical device with 10 MeV protons and record some device property (e.g., the gain of a bipolar transistor) as a function of fluence for both cases. Then we look through the data and observe that some fluence, call it x, of 500 KeV protons produced the same degradation, measured by a change in the device property being observed, as some other fluence, call it y, of 10 MeV protons. We have an equivalent fluence; x fluence of 500 KeV protons is equivalent to y fluence of 10 MeV protons. The ratio y/x is the conversion factor that converts x into y. By repeating this experiment for different particle energies and species, we obtain conversion factors for converting fluences at each energy and species into an equivalent 10 MeV proton fluence. This allows the fluence from a mixture of particles, e.g., found in the space environment, to be expressed as an equivalent 10 MeV proton fluence. A set of spectral curves (fluence versus energy for each particle species) can therefore be reduced to a single number; the equivalent 10 MeV proton fluence.

However, there are complications. The equivalent fluence conversion factors applicable to protons having low energies (the most damaging energies for protons) may be unique to the device. The reason is that low-energy protons slow down as they travel through the device. It is experimentally convenient to identify a particle by its energy outside the device, but this may not be the energy it has when it reaches the device location at which the displacements are most important. This is particularly true if the

incident particles hit a device at an angle so that they cannot penetrate to the same depth as the same particle at normal incidence. We might find different equivalent fluence conversion factors for devices that provide different amounts of self-shielding around the active regions. In fact, we can find different conversion factors for different types of device properties if these properties are influenced by displacement damage in different device regions. For example, the short-circuit current from an illuminated solar cell is primarily affected by displacement damage in the substrate, while the open circuit voltage and maximum power are both affected by displacement damage in the p-n junction depletion region. When the spectrum (produced in a laboratory experiment) in one region differs from that in the other, the equivalent fluence conversion factors are different depending on whether we are monitoring short-circuit current or open-circuit voltage.

When using the above experimental definition, equivalent fluence can depend on incident particle direction relative to the device, device construction, and on the type of device property (e.g., short-circuit current or open-circuit voltage for a solar cell) being investigated. One approach is to accept this state of affairs and define different kinds of equivalent fluences for different devices or device properties. This is the approach used by Tada *et al.* in the "Solar Cell Radiation Handbook" [3.11]. Actual fluences are first transported through the cover glass before being converted into equivalent fluence, so the cover glass thickness does not affect the conversion, but it is still necessary to distinguish short-circuit current (affected only by displacement damage in the substrate) from open-circuit voltage or maximum power (affected by displacement damage in the depletion region). This distinction is enough as long as solar cells are sufficiently similar with regards to the depths of the depletion regions, so only two kinds of equivalent fluences are needed; one for short-circuit current and another for maximum power. This is a convenient approach for solar cells, because there are only two kinds of conversions, and also because there are computer codes which include the effects of incident angle to calculate equivalent fluence for an isotropic environment [3.11]. Furthermore, there are many cases in which the user does not have to measure anything, because an extensive set of measurements representing many solar cell constructions were already performed [3.11]. However, neither conversion can be expected to apply universally to all semiconductor devices susceptible to displacement damage.

An alternative to accepting the above state of affairs is to replace the experimentally convenient but logically awkward definition of equivalent fluence with a more logically appealing but experimentally awkward definition. A more universal equivalent fluence conversion might be expected if particle energy refers to the energy at the location where the damage is being assessed, instead of energy outside the device. Damage now refers to local damage, which might be defined, for example, in terms of a local carrier lifetime or diffusion length. Using this definition, equivalent fluence conversions can be expected to be more of a material property and less of a device property. We temporarily set aside the question of how equivalent fluence conversions might be determined and assume that the conversions are (somehow) already known, and consider the question of how this information can be used to predict device response. As pointed out in Section 3.1, the spectra are reasonably uniform (even including low-energy particles, because of a

replacement effect) within a microelectronic device surrounded by spacecraft shielding and exposed to the natural space environment, so an equivalent 10 MeV proton (for example) fluence will be reasonably uniform throughout the device. Let y be the equivalent 10 MeV proton fluence, so the local damage throughout the device is the same as would be produced by a fluence y of 10 MeV protons at the same location. But 10 MeV protons are penetrating enough, at normal incidence, to produce a uniform spectrum throughout a microelectronic device (even though there is no replacement effect), so the actual damage throughout the device is the same as produced by a fluence y of normal incident protons having 10 MeV of energy outside the device. The device response to the latter damage can be measured, so the device response to the actual environment can be predicted. Therefore, if equivalent fluence is defined locally so that it becomes less of a device property than when defined by the previous experimental definition, it can still be combined with measured device data to predict device responses to the actual environment. These are desirable characteristics of an environmental measure; the measure is not a device property but can still be combined with measured device data to predict device behavior.

There still remains the question of how to determine equivalent fluence conversions when particle energy refers to energy at the site. This is experimentally difficult for low-energy protons, so a theoretical treatment has been becoming increasingly popular over the last ten years. This is based on the non-ionizing energy loss (NIEL) rate, which is just like LET for total (and almost all electronic) energy loss, except that it refers to the nuclear energy loss. Note that not all of this energy loss is converted into stable defects, partly because not all of this energy is converted into vacancy-interstitial pairs, but primarily because most of the initial vacancy-interstitial pairs recombine instead of becoming stable defects. However, the only requirement for this treatment to be useful is that damage (measured in terms of a quantity considered relevant, such as change in the reciprocal carrier lifetime) be proportional to the NIEL. A similar consideration was already discussed for LET. Not all energy transfer via interactions with electrons is converted to electron-hole pair creation, but LET is still a useful measure of ionization because electron-hole pair creation is proportional to this energy (recall that the proportionality constant is 1/3.6eV for silicon).

However, it is not clear from first principles that damage is proportional to NIEL, so this question must be answered experimentally. The experimental answer is that there are limitations [3.12]. In particular, lower energy protons appear to be more effective at producing displacement damage in GaAs as compared to higher energy protons than the NIEL correlation would indicate [3.12]. Also, the use of NIEL correlation to predict proton effects from neutron data has only a marginal performance [3.12]. A possible explanation is that different incident particles produce a different energy distribution of primary knock-on atoms, which affects the density of vacancy-interstitial pairs, which in turn influences the fraction of these pairs that recombine. This is similar to the problem discussed under TID, where significant differences in e-h pair recombination causes heavy ions to produce a different device response than the same dose of electrons. A yield function was needed to improve the TID correlation between different particles. Improving the displacement damage correlation between different particles is still an area

of active research. A practical approach in the meantime is to test device degradation under conditions not too dissimilar to the space environment. For example, if nearly all displacement damage from a given space environment is from protons in a given energy range, then device degradation should be measured using protons having energies that are representative, but also high enough to avoid a non-uniform spectrum within the device. NIEL curves would then be needed only to correlate particles that are somewhat similar to each other.

The NIEL concept provides some options that are based on the same physical assumptions, so the choice is a matter of preference. NIEL curves can be used to find fluences equivalent to an actual fluence, so an environment can be measured by an equivalent fluence, or we can use NIEL curves to convert all fluences into a "displacement dose". This is calculated for an environment in the same way as ionizing dose, except that NIEL curves are used instead of LET curves.

In summary, there are several measures used to describe an environment with regards to displacement damage. One is in terms of equivalent fluences applicable to solar cells. There are two kinds; one for short-circuit current and another for maximum power. Another measure is an equivalent fluence derived from NIEL curves, and a third is displacement dose, also derived from NIEL curves.

3.5 Pertinent Environments

The various environments that significantly contribute to a given radiation effect, or, conversely, the various radiation effects associated with a given environment are summarized in Table 3.1. A "yes" in the table means that the indicated environment is important to the indicated effect. A "no" means that, while the indicated environment may contribute to the indicated effect, this contribution is dominated by the contributions from other environments.

Table 3.1. Radiation effects associated with various environments.

A "yes" indicates that an environment is an important contribution to an effect, while a "no" indicates that the contribution is not important.

	Trapped Particles around Earth	**GCR**	**Solar Events**
TID	Yes	No**	Yes
Single Event Effects (direct ionization)	No*	Yes	Yes
Single Event Effects (indirect ionization)	Yes	No	Yes
Displacement Damage	Yes	No	Yes

* There are a few exceptions. Some very susceptible devices respond to protons by direct ionization.
** Excluding microdose effects.

References

[3.1] T.P. Ma and P.V. Dressendorfer (editors), *Ionizing Radiation Effects in MOS Devices and Circuits*, John Wiley and Sons, p. 91, 1989.

[3.2] H. Johns and J. Cunningham, *The Physics of Radiology*, Charles Thomas, Springfield, pp. 363-365, 1974.

[3.3] C. Dufour, P. Garnier, T. Carriere, J. Beaucour, R. Ecoffet, and M. Labrunee, "Heavy Ion Induced Single Hard Errors on Submicronic Memories," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 6, pp. 1693-1697, Dec.1992.

[3.4] T. Oldham, K. Bennett, J. Beaucour, T. Carriere, C. Polvey, and P. Garnier, "Total Dose Failures in Advanced Electronics from Single Ions," *IEEE Trans. Nucl. Sci.*, vol. 40, no. 6, p. 1820, Dec.1993.

[3.5] L.D. Edmonds, *A Theoretical Analysis of Steady-State Photocurrents in Simple Silicon Diodes*, JPL Publication 95-10, March 1995.

[3.6] L.D. Edmonds, "Charge Collection from Ion Tracks in Simple EPI Diodes," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pp. 1448-1463, June 1997.

[3.7] L.D. Edmonds, "Electric Currents Through Ion Tracks in Silicon Devices," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pp. 3153-3164, Dec. 1998.

[3.8] P.E. Dodd, "Basic Mechanisms for Single Event Effects," *1999 IEEE Nuclear and Space Radiation Effects Conference Short Course.*

[3.9] M. Reier and G. Swift, "A 252Cf Time-of-Flight System for SEU Testing," *RADECS 93*, pp. 93-96, 1993.

[3.10] H. Dussault, J.W. Howard, Jr., R.C. Block, M.R. Pinto, W.J. Stapor, and A.R. Knudson, "High Energy Heavy-Ion-Induced Single Event Transients in Epitaxial Structures," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 2018-2025, Dec.1994.

[3.11] H.Y. Tada, J.R. Carter, Jr., B.E. Anspaugh, and R.G. Downing, *Solar Cell Radiation Handbook*, third edition, JPL Publication 82-69, Nov.1, 1982.

[3.12] P.W. Marshall and C.J. Marshall, "Proton Effects and Test Issues for Satellite Designers," *1999 IEEE Nuclear and Space Radiation Effects Conference Short Course.*

4. BULK TRAPS AND INTERFACE STATES

Chapter 3 presented some of the basic concepts that are prerequisites for understanding the more specialized subjects in Chapter 5, but one important omission is the subject of charge trapping. It was noted in Section 3.2 that holes liberated by ionizing radiation can become trapped in the interior of an insulator, but there was no explanation of how, and no mention of the fact that there is also a charge (which can have either polarity) associated with interface states. The present chapter begins with a preliminary discussion of states. Specific illustrations refer to semiconductors, because these illustrations will be more familiar to readers that have not already learned about radiation effects, and also because conceptually simple comparisons (both similarities and differences) can be made with other types of states. The next two sections discuss bulk trapping and interface states. To be definite, the insulator selected for these discussions is SiO_2. The last two sections discuss time-dependent effects on the level of fundamental principles, with the specifics deferred to Chapters 5 and 6.

4.1 A Preliminary Discussion of States

A state of an electron in a material has an energy level specified by a set of internal quantum numbers (e.g., indicating that the electron is on the valence band edge or conduction band edge), and a location-dependent electrostatic potential. It is important to make a distinction between an electron state and the occupancy of the state. This is analogous to distinguishing between a bear trap and a trapped bear. A trapped bear is an occupied bear trap. A state can be visualized as a container that is either occupied or unoccupied. When in thermodynamic equilibrium, an electron state is occupied nearly all the time (or, equivalently, nearly all identical states in a volume element are occupied at any given time) if the energy level of the state is far below the Fermi level. The state is unoccupied nearly all the time if the energy level is far above the Fermi level. A state is called "donor-like" if it is neutral when occupied by an electron and positively charged when unoccupied. It is "acceptor-like" if negatively charged when occupied with an electron and neutral when unoccupied [4.1].

The occupancy of a state can temporarily and randomly (by thermal interactions) change via a two-step recombination process. For example, consider a donor-like state that is initially unoccupied (positively charged). The first step can be the capture of a conduction electron, which occupies the previously unoccupied state, and the second step gives up this electron to the valence band by filling a hole. The final result is that the state is back to being unoccupied (positively charged), but an e-h pair has been annihilated. For another example, consider a donor-like state that is initially occupied (uncharged). The first step can be the capture of a hole, which is another way of saying that the electron initially occupying the state was given up to the valence band by filling a hole. The second step can be the capture of a conduction electron, so the state is back to being occupied (uncharged), but an e-h pair has been annihilated. Analogous considerations apply to acceptor-like states. Note that during the time between the two steps of the

recombination process, the occupancy (and charge) is different than before or after the process, so a state is occupied some fraction of the time and unoccupied the rest of the time. Mechanisms other than the recombination steps discussed here can also change a state's occupancy. In any case, when in thermodynamic equilibrium, the fraction of time that a state is occupied (or the fraction of occupied states in a volume element at a given time) is determined by the position of the energy level relative to the Fermi level.

Some states are charged almost all of the time, some other states are uncharged almost all of the time, while still other states are neither. An example of states that are charged almost all the time are those produced by doping a semiconductor. For the n-type case, the impurity is selected so that the states it introduces are donor-like and unoccupied almost all the time. The occupancy requirement is satisfied if the state energy level is close to the conduction band edge. With this condition satisfied, it is difficult to change the occupancy of the states, even when ionizing radiation creates e-h pairs, except for a small fraction of the time during recombination. The state is uncharged between the two recombination steps, and e-h pairs produced by ionizing radiation increases the frequency of occurrence of this two-step process. However, even when exposed to ionizing radiation, the fraction of time spent in this uncharged condition is too small to significantly affect the spatial density of immobile ions. In spite of this recombination, the density of immobile ions is simply the doping density. However, the recombination can still have an important effect on carrier lifetime. Analogous considerations apply to acceptor-like states that are almost always occupied in a p-type semiconductor.

An example of a state that is uncharged almost all the time is a donor-like state in a semiconductor with energy level near the valence band so that it is occupied almost all the time. There is some fraction of time in which the state is unoccupied (charged) between the two recombination steps, and e-h pairs produced by ionizing radiation increases the frequency of occurrence of this two-step process. However, even when exposed to ionizing radiation, the fraction of time spent in this charged condition is too small to significantly contribute to the spatial density of immobile ions. In spite of this recombination, the density of immobile ions produced by these states is negligible. However, the recombination can still have an important effect on carrier lifetime. Analogous considerations apply to acceptor-like states that are almost always unoccupied.

Examples of states that are in neither of the above categories can be acceptor-like or donor-like. Analogous discussions apply to the two cases. To be definite, consider a normally-occupied (uncharged), donor-like state in a semiconductor going through the same two-step recombination process described in the previous paragraph, except that the time between the two steps (the time the state is charged) is a significant fraction of the total time. After the initial hole capture, the state remains charged for an extended time, and the hole is said to be "trapped", until the second (neutralization) step. Because this condition persists for an extended time (or, equivalently, because many identical states in a volume element are charged at a given point in time), this type of state *does* affect the spatial density of immobile ions. The immobile ions now include trapped holes. As stated earlier, when in thermodynamic equilibrium, the average fraction of states in a volume

element that are unoccupied (charged for this case) is determined by the state energy level relative to the Fermi level, so the trapped hole density can be calculated under equilibrium conditions. A condition of zero current but not thermodynamic equilibrium (e.g., a MOS capacitor with an applied bias voltage) mimics thermodynamic equilibrium if the state energy level at a given location includes the potential at that location produced by biasing voltages. The spatial density of the trapped holes depends on location within the device and on biasing voltage. The occupancy can change under time-varying conditions (e.g., by changing biasing voltage or by disturbing equilibrium by using ionizing radiation to create e-h pairs), but even when the average occupancy does not change, the recombination process reduces carrier lifetime.

4.2 Bulk Traps

We now consider charge trapping in the interior of an oxide. Volume defects (e.g., oxygen vacancies) in the oxide produce states in the forbidden band that are able to trap holes. Electrons can also be trapped in an oxide, but they are usually outnumbered by the trapped holes, so the focus of this discussion will be on trapped holes. The volume defects can produce donor-like states that are normally occupied by electrons (uncharged) when in equilibrium. Ionizing radiation disturbs equilibrium conditions by creating holes near these defects, so that there are many more opportunities for changes in state occupancy than would be the case under equilibrium conditions. Many previously and normally occupied donor-like electron states, called trapping sites or traps, give up an electron to the valence band by filling a hole, so the hole is removed and the state is positively charged. This unoccupied electron state, or charged trapping site, is an immobile ion called a trapped hole. Note that the ionizing radiation also creates free electrons. Not all survive long enough to be driven out of the oxide by the electric field. Some recombine directly with holes (the probability of this determines the yield function) before the hole is trapped. Some others are captured by an unoccupied electron state, annihilating the trapped hole, i.e., the free electron recombines with a trapped hole. An increasing trapped hole density, produced by an accumulating dose, increases the probability that a free electron will recombine with a trapped hole, so the trapped hole density tends to saturate with increasing dose after a sufficiently large dose level is reached.

After the normally uncharged trapping sites become trapped holes, many remain in this charged condition for an extended time; enough time to produce a significant spatial density of trapped holes, so the oxide contains a space charge. This condition persists until thermodynamic equilibrium is re-established, which occurs through annealing. Annealing is a very slow process at room temperature (hours to years, depending on the electric field and depending on how complete the annealing is) because charge exchange between the oxide and semiconductor is very slow after the initially liberated mobile e-h pairs are gone. Annealing is faster at elevated temperatures.

If the oxide is in contact with silicon, one mechanism for annealing is the tunneling of electrons through the oxide-silicon barrier, from the silicon to the oxide, and these

electrons fill the charged (unoccupied) donor-like states in the oxide [4.2]. This is similar to the recombination process, discussed at the end of Section 4.1, but there is a very long delay between the two steps. After the first step (hole capture, or trapping), the trapping site keeps the hole for a very long time before the second step of the process (electron capture) is complete. Another annealing mechanism is the thermal emission of electrons from the valence band to the traps [4.2]. This is similar to recombination except that the captured electron is from the valence band, leaving a hole. After the initial hole trapping, another hole appears later. This is hole de-trapping.

For oxides in contact with silicon, the spatial distribution of trapping sites is a superposition of two components. One component is a comparatively uniform distribution associated with volume defects such as oxygen vacancies. The other component is concentrated near the oxide/silicon boundary; in a disturbed region (characterized by strained bonds) that is the transition between the crystalline silicon and the amorphous silicon dioxide [4.2]. Trapped charge at or near the oxide/silicon boundary has a stronger influence on the operation of a device than the same amount of charge trapped at a different location. For the purpose of calculating the electric field distribution in the oxide, this latter component can be approximated as being at the boundary. However, it is still necessary to recognize that these trapping sites are in the bulk in the sense that they do not freely exchange charge with the silicon, so they are distinguished from the interface states discussed in the next section.

Damage from ionizing radiation can contribute to the spatial density of trapping sites, but the most immediate effect of ionizing radiation, regarding the oxide bulk, is in terms of the occupancy of states rather than the number of states. The opposite applies to interface states discussed in the next section. The primary effect of ionizing radiation on these states is in terms of the number of states rather than the occupancy. The significance of this is discussed in the next section.

4.3 Interface States

Consider an oxide in contact with silicon. A MOS capacitor is a simple prototype arrangement. Interface states are produced by missing (dangling) bonds at the oxide-silicon interface [4.2]. These states are already present prior to radiation damage, and belong as much to the silicon as to the oxide. Therefore, these states readily exchange charge with the silicon. The significance of this is that when conditions (e.g., biasing voltages applied to a MOS capacitor) do not change too fast (with time scales on the order of a microsecond for the "rapid states", which are usually the most common [4.2]), the charge exchange can keep up with changing conditions in the sense that quasi-equilibrium is maintained. In other words, the occupancy of these states is in thermodynamic equilibrium with the silicon. In this respect, these states are similar to states in the silicon discussed in Section 4.1.

Experiments have indicated that interface states below the middle of the silicon band gap are donor-like, while those above are acceptor-like [4.3]. Therefore, if biasing

conditions are chosen so that the midgap potential at the interface (which changes with biasing voltage) is at the Fermi level, the donor-like states are occupied and the acceptor-like states are unoccupied, so the interface states are uncharged. This is the neutral point for the interface states. If bias voltage is changed from this neutral point, some interface state occupancies change and contribute a net surface charge. Recall that biasing voltages can cause a MOS capacitor to be in either of several modes. The accumulation mode is characterized by majority carriers in the silicon accumulating in the silicon interior near the oxide-silicon interface. The depletion mode is characterized by majority carriers being driven away from the interface, leaving uncompensated immobile ions to create a space charge in the silicon interior near the interface. The inversion mode is characterized by minority carriers accumulating near the interface. The neutral point for interface states occurs when the midgap potential is at the Fermi level, and this is a point in the depletion mode. If the bias voltage is changed from this point in the direction of inversion, there is a net interface state charge, having the same polarity as the charge in the silicon bulk during inversion. If the bias voltage is changed from the neutral point in the direction of accumulation, there is a net interface state charge having the same polarity as the charge in the silicon bulk during accumulation.

The above paragraph leads to the following conclusions:

A) If the silicon under the oxide is n-type, interface states contribute a positive surface charge if biasing voltages produce the inversion mode (an accumulation of holes in the bulk silicon near the interface). The contribution is negative if biasing voltages produce the accumulation mode (an accumulation of electrons in the bulk silicon). The neutral point for the interface state charge is a point in the depletion mode (a depletion of electrons leaving uncompensated positive impurity ions in the bulk silicon).

B) If the silicon under the oxide is p-type, interface states contribute a negative surface charge if biasing voltages produce the inversion mode (an accumulation of electrons in the bulk silicon near the interface). The contribution is positive if biasing voltages produce the accumulation mode (an accumulation of holes in the bulk silicon). The neutral point for the interface state charge is a point in the depletion mode (a depletion of holes leaving uncompensated negative impurity ions in the bulk silicon).

In addition to contributing to charge, interface states (like other states discussed in Section 4.1) can also contribute to recombination, reducing the lifetime of carriers in the bulk silicon near the interface. The charged interface states can also reduce the mobility of carriers in the bulk silicon near the interface, due to scattering of the carriers by the surface charges.

There has not yet been any mention of radiation effects, because such effects do not alter the above discussions. In particular, the occupancy of interface states is still in thermodynamic equilibrium with the silicon, which allows the charge to have either polarity (depending on biasing conditions) and is therefore very different than the bulk traps discussed in the previous section. What ionizing radiation *will* do is increase the

number of interface states. The mechanisms are still areas of active research. Some models refer to hydrogen, while some other models refer to the detrapping of holes from the oxide (a delayed reaction that can occur long after the irradiation exposure), which release energy when they reach the interface. Whatever the mechanisms are, an experimental observation is that the number of interface states can continue to increase long after the irradiation has ceased. Once created, an "annealing" of interface states would be the annihilation of the states, as opposed to annealing of bulk trapped charge, which is a change in state occupancy. Interface states do not anneal as readily as bulk trapped charge. In fact, the number of interface states can be important and still increasing even after the bulk trapped charge has significantly annealed. This leads to some interesting effects discussed in Section 4.5.

4.4 Time Regimes

Each time an ionizing particle liberates e-h pairs in an insulator, the carrier creation initiates a set of processes that can be partitioned into three distinct and consecutive time regimes. Two processes occur simultaneously during the first time regime. One is the recombination of electrons with mobile or trapped holes, and the other is the removal of the surviving electrons via any electric field that is present. This first time regime is short enough (on the order of a picosecond) to be effectively instantaneous from the point of view of laboratory measurements. The processes during this time regime determine the yield discussed in Section 3.2. Recall that the yield depends on the electric field, so it depends on the biasing voltage that was present during the first time regime.

The second time regime refers to the transport and trapping of holes. The total number of holes (mobile plus trapped) is nearly constant during this time regime, so the net charge in the insulator interior is constant and can be calculated from the yield. However, the spatial distribution of the charge density is changing as the holes move, until they become trapped and the spatial distribution becomes semi-permanent. In this context, "semi-permanent" means that subsequent variations are over times much longer than the duration of the second time regime. The duration of this time regime varies from case to case, but is on the order of a second for oxides at room temperature and having typical thicknesses and electric fields. The semi-permanent spatial distribution of charge having the greatest impact on device (e.g., a MOS FET) functionality is that in which the charge is concentrated near the insulator-semiconductor interface. The spatial distribution having the least impact is that in which the charge is concentrated near the insulator-electrode interface, where it is shielded by image charges in the electrode. The spatial distribution depends on the applied fields during the second time regime, so the impact on device functionality depends on the polarity of the biasing voltages that were present during this time regime. The most relevant quantitative measure of the spatial distribution is the first moment, which is the volume integral of the product of the charge density multiplied by the distance from the electrode. An equivalent measure (equivalent in the sense that one measure can be calculated from the other when the insulator thickness is known) is an effective charge. This effective charge is the amount of charge which, when concentrated near the insulator-semiconductor interface, produces the same first moment as the actual

charge distribution. A given amount of actual charge near the insulator-semiconductor interface produces a larger effective charge than the same amount of actual charge near the insulator-electrode interface. A larger effective charge has a greater impact on device functionality. The first and second time regimes together determine the semi-permanent effective charge.

The third time regime refers to two processes that occur simultaneously. One is the annealing of the trapped charge in the insulator interior, and the other is the creation of interface states. As stated in Section 4.2, the duration of this time regime can be hours or even years. When device functionality is measured and recorded on an hourly basis, we are comparing various time points in the third time regime. Some time-dependent effects can be explained in terms of the two above processes that occur during the third time regime, as discussed in the next section.

4.5 Conventional Time-Dependent Effects

A simple prototype device used for illustration in this discussion is a MOS FET. A p-channel device has a p-type source and drain, with the region between them being n-type silicon under an oxide. Similarly for an n-channel device, but with p-types and n-types interchanged. The threshold voltage is the voltage at the gate above the oxide at which the device switches on, and this is the gate voltage at which an inversion layer forms under the oxide between the source and drain. One measure of the radiation response of such a device is the change in threshold voltage produced by irradiation. For this measure, carrier lifetime and mobility in the silicon are not relevant, but oxide charge (both bulk and at the interface) are relevant.

Recall from Section 4.3 that an inversion layer in n-type silicon under an oxide (the p-channel device) is accompanied by a positive surface charge at the interface. Therefore, when at the threshold voltage, the interface states produce a net positive charge for the p-channel device. This adds to the positive charge in the oxide bulk, i.e., bulk charge and interface charge have additive effects on the threshold voltage for the p-channel device.

The analogous statement for the n-channel device is that bulk charge and interface charge have competing effects on the threshold voltage. Therefore, a perturbation, or offset, in the threshold voltage produced by ionizing radiation can be in either direction depending on whether the bulk charge or interface charge has the stronger influence. Suppose an n-channel device is irradiated at a sufficiently high dose rate so that a prescribed amount of dose is accumulated during a time that is short compared to the bulk oxide anneal time. The threshold voltage is then measured at various times after the irradiation. Early after irradiation, the bulk oxide charge has the stronger influence. However, this influence weakens as the bulk oxide charge anneals, while the influence from the interface charge still remains (and possibly even increases due to delayed creation of interface states). Instead of a recovery (no offset) at late times, there remains an offset from the interface charge, and in the opposite direction of the initial offset. This is the rebound effect. The perception from device behavior is that the threshold voltage

recovers too much from annealing, i.e., it overshoots. This applies only to the n-channel device. For the p-channel device there is no competition and no rebound.

Another interesting effect is seen when comparing different dose rates for the n-channel device. The relevant concepts are the same as in the previous discussion, except for recognizing that irradiation and bulk oxide charge annealing can occur together. First consider the case in which the n-channel device is irradiated at a sufficiently high dose rate so that a prescribed amount of dose is accumulated during a time that is short compared to the bulk oxide anneal time. The threshold voltage is measured, not a long time after the irradiation has terminated, but immediately after the prescribed dose was accumulated. At this time, the bulk charge is still strong and competes with the interface charge. This competition gives the illusion that there is only a mild radiation effect, when in reality there can be strong separate effects that are not individually observable from this single measurement. Now consider another case in which an identical device is exposed to the same dose but at a much lower rate. For this second case we measure the threshold voltage immediately after the prescribed dose was accumulated, so the exposure time prior to measurement is much longer for the second device. For this device, significant bulk oxide annealing was occurring even during the irradiation, while the number of interface traps was increasing up to the time of measurement. The result is the same as the rebound effect, except that it was observed as soon as the prescribed dose had accumulated, instead of having to wait for a post-irradiation annealing. The end result is that the two dose rates produce different device responses measured at the time the prescribed dose level was reached.

The previous two paragraphs compared device properties at different times and with different combinations of irradiation with annealing. An annealing followed irradiation (with negligible annealing during the irradiation) in one case, while annealing and irradiation occurred simultaneously in another case. In the examples considered, bulk oxide charge was in competition with interface charge. These are specific examples of the more general subject of time-dependent effects. Other examples include other combinations of irradiation and annealing, or those cases (doping types or biasing voltages at which device properties are measured) in which bulk charge reinforces the interface charge. The subject of time-dependent effects is sufficiently well understood so that, for many devices, it is possible to predict device response to arbitrary dose rates (excluding extreme cases, such as very-high dose rates) from laboratory measurements utilizing high dose rates (convenient for laboratory use) and anneals [4.4, 4.5]. Details are given in Section 6.1. However, there is an exceptional case (discussed in the next section) that is not well understood, but can be very important to space missions.

4.6 Enhanced Low Dose Rate Effects

If the electric field in an oxide is sufficiently weak and if the oxide is soft (contains a large density of defects such as oxygen vacancies), there can be another type of time-dependent effect, called the enhanced low dose rate (ELDR) effect. The first word "enhanced" distinguishes this from the more conventional time-dependent effects

discussed in the previous section. The cause of this effect is still an area of active research, but a result of this effect is illustrated by the following example. Two identical devices are irradiated at different dose rates, and bulk oxide charge measurements are performed after a prescribed dose level was reached. At the time of measurement, the accumulated dose is the same for both devices, but the exposure time (and the annealing during irradiation) is larger for the low-dose-rate case. The expected result is that the bulk oxide charge should be less for the low-dose-rate case. The surprising experimental observation is that this charge can actually be greater for the low-dose-rate case [4.5]. A particular experiment using zero biasing conditions also found that high-rate irradiation followed by room temperature anneals could not simulate low-rate responses for a soft oxide, even though it can for harder oxides [4.5]. The ELDR effect was not observed, for either soft or hard oxides, for other biasing voltages used in these experiments [4.5].

The physical explanation for the ELDR effect has not yet been conclusively established, and the subject is still an area of controversy, but it appears to be connected with the dynamics of charge motion in thick oxides. However, this effect is expected to be relevant to the dose rates encountered by many space missions. A particular concern is the experimental characterization of a device intended for spacecraft use. Duplicating the dose rate seen by a spacecraft over a period of several years would require an experiment of equal duration, which is not practical. An important task that remains to be done is to understand ELDR effects well enough so that a device response to a known space environment can be predicted from laboratory measurements utilizing practical dose rates.

References

[4.1] T.P. Ma and P.V. Dressendorfer (editors), *Ionizing Radiation Effects in MOS Devices and Circuits*, John Wiley and Sons, p. 17, 1989.

[4.2] Jean-Luc Leray, "Total Dose Effects: Modeling for Present and Future," *1999 IEEE Nuclear and Space Radiation Effects Conference Short Course.*

[4.3] T.P. Ma and P.V. Dressendorfer, *op. cit.*, p. 196.

[4.4] T.P. Ma and P.V. Dressendorfer, *op. cit.*, p. 287.

[4.5] D.M. Fleetwood, S.L. Kosier, R.N. Nowlin, R.D. Schrimpf, R.A. Reber, Jr., M. DeLaus, P.S. Winokur, A. Wei, W.E. Combs, and R.L. Pease, "Physical Mechanisms Contributing to Enhanced Bipolar Gain Degradation at Low Dose Rates," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 1871-1883, Dec.1994.

5. A MORE DETAILED LOOK AT DEVICE RESPONSES

This chapter looks at real devices, which exhibit a greater variety of radiation effects than discussed in earlier chapters. While earlier chapters focused on fundamental concepts, the present chapter discusses specific device technologies.

5.1 TID Effects in MOSFETs, Bipolar Circuits, and Miscellaneous Devices

The most widely known and widely studied structures that exhibit sensitivity to TID is SiO_2 and the oxide-silicon interface. These are found in all silicon CMOS devices, and also in silicon bipolar circuits with oxide isolation. The susceptibility of CMOS circuits to TID can be attributed almost entirely to these structures, via trapped charge and interface states discussed in Chapter 4. In particular, MOSFET gate oxides, and field oxides used to isolate portions of the circuit from each other, can cause large changes in device properties (MOSFET threshold voltage shift, increased leakage current, and degraded timing parameters). As a rough rule for MOSFET-based circuits, unhardened commercial technologies exhibit failure doses from a few krads(Si) to a few tens of krads(Si), "radiation tolerant" technologies are in the range of 20 to 100 krads(Si), and intentionally hardened technologies can be in excess of 1 Megarad(Si). In the future, scaling effects (i.e., reduced feature sizes) may lead to greater gate oxide hardness because these oxides will be thinner, but there will be continuing problems with field oxides resulting in current leakage within the circuit.

Digital bipolar circuits tend to be less susceptible to TID effects than equivalent MOSFET circuits, because the bipolar circuits do not contain sensitive gate oxides directly over carrier channels. However, digital bipolar circuits are not completely immune to TID, and some are more susceptible than others, because they do employ oxide isolation techniques. In contrast, silicon bipolar linear circuits can be very susceptible to TID, in the range of a few tens of krads(Si). Devices such as operational amplifiers, comparators, and voltage references can exhibit degradation in the form of changes in offset voltage, offset current, bias current, and open loop voltage gain.

The relative hardnesses of various MOSFET and bipolar technologies to TID are indicated in Figs. 5.1 and 5.2, which were taken from Ref.[5.1]. Fig. 5.2 also includes GaAs digital circuits, indicating that they are very hard. The reason is that these circuits do not contain SiO_2 oxides.

For a broader comparison of TID hardness levels, Fig. 5.3 (taken from Ref.[5.1]) includes a variety of different technologies including those discussed briefly above, and also silicon discrete devices, GaAs devices, charge coupled devices (CCDs), crystal resonators, and optical fibers. It is interesting to note that in this bar chart, created in 1990, silicon bipolar transistors do not show failure until a dose level of about 1 Megarad(Si) is reached. However, tests performed by the Jet Propulsion Laboratory on transistors for the Cassini mission found significant degradation at tens of krads(Si).

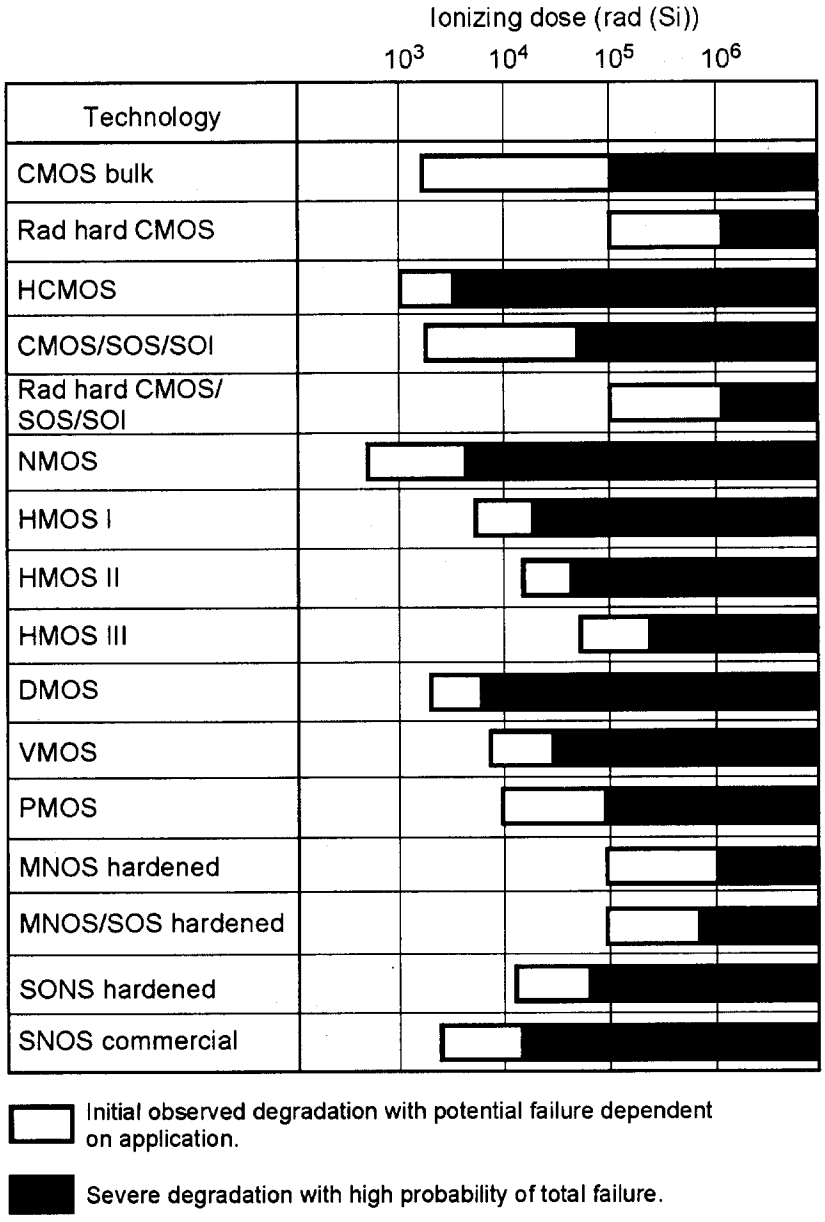

Fig. 5.1: Ionizing dose failure levels for MOSFET integrated circuits from Ref.[5.1].

The reason for this is presumably that performance "improvements" in the transistors, an ability to operate over a wider collector current range (for example), have increased the TID susceptibility. In fact, performance enhancements in silicon technologies have often led to increased radiation sensitivity. Finally, we note that the GaAs circuits in Fig. 5.3 are among the hardest shown in this bar chart.

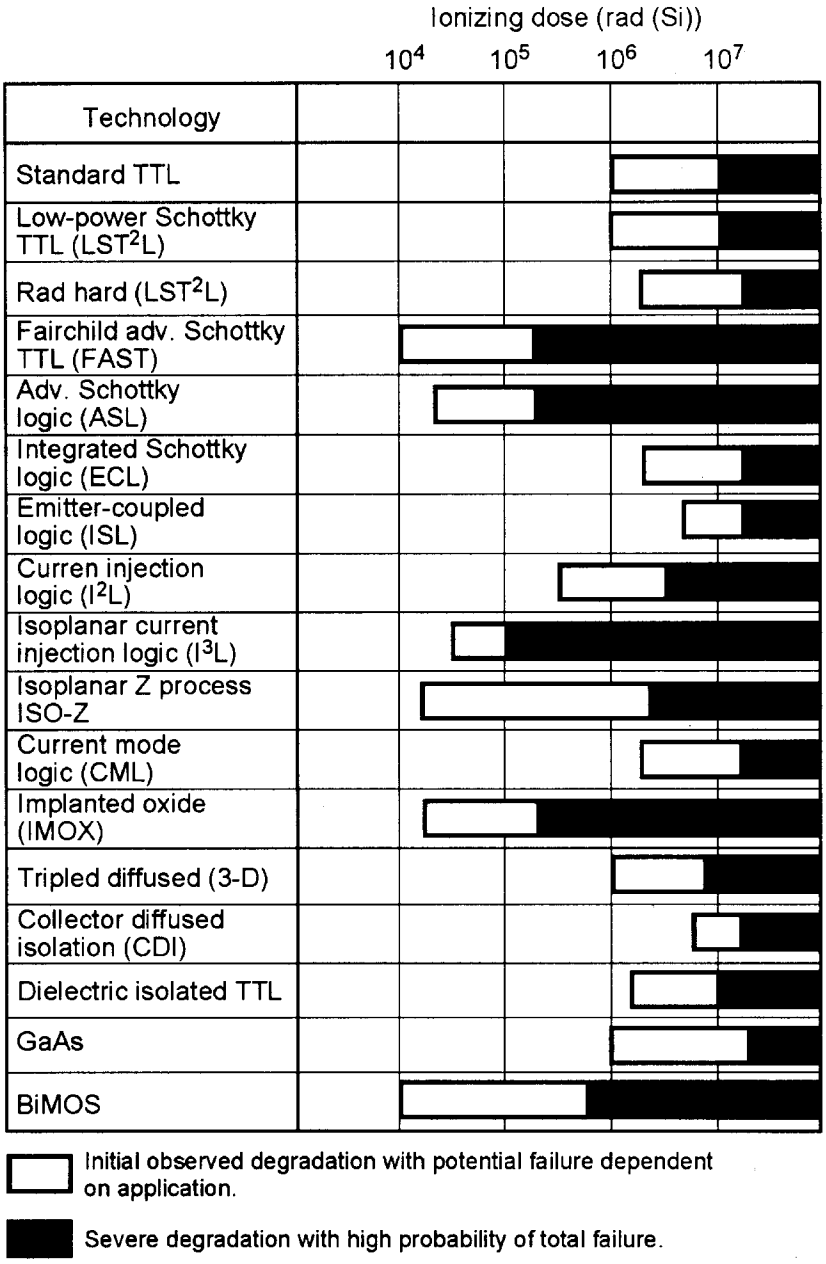

Fig. 5.2: Ionizing dose failure levels for bipolar integrated circuits from Ref.[5.1].

5.2 ELDR Effects in Linear Bipolar Circuits

Because silicon linear bipolar circuits are very susceptible to TID, the ELDR effects discussed in Section 4.6 will also be important, if they occur. The processing steps used to produce these devices tend to result in soft oxides. Also, the electric fields in the oxides are small because there are no gates that directly bias the oxides. Therefore, the ELDR effects do occur. The required dose rate is generally less than about 1 rad(Si)/sec. Sometimes the required dose rate is several orders of magnitude less than this, in which

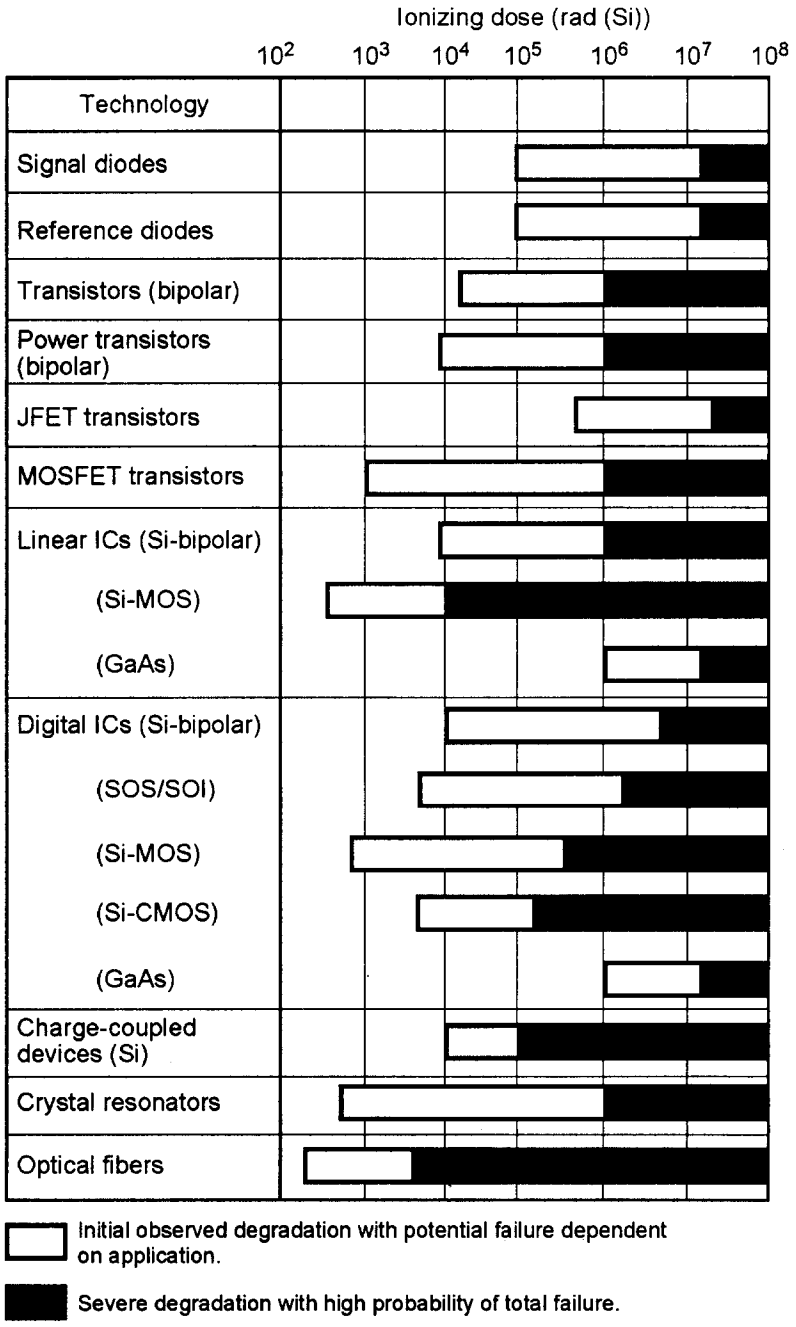

Fig. 5.3: Ionizing dose failure levels for discrete, linear, and digital device families from Ref. [5.1].

case the effect is difficult to observe in the laboratory, but can still occur in some space environments. This is an important concern for space missions.

Device susceptibility to ELDR tends to increase with increasing oxide thickness. Experimental measurements have found that different circuit types produced by the same manufacturer, or the same circuit type produced by different manufacturer, can have different ELDR susceptibilities because of variations in oxide thickness, but show a consistent correlation between device susceptibility and oxide thickness [5.2].

5.3 Displacement Damage Effects

While TID primarily affects oxides and interfaces within a circuit, displacement damage affects the bulk properties of a semiconductor as discussed in Section 3.4. The most immediate effect (i.e., seen at low damage levels) is to reduce minority carrier lifetime in a semiconductor, but displacement damage in sufficient concentration can also produce other effects, such as a reduced carrier mobility or compensation of donors or acceptors. Thus, the types of devices and circuits that are most sensitive to displacement damage are those that depend on large values of minority carrier lifetime for satisfactory performance: bipolar transistors (excluding very high speed switching transistors), linear bipolar circuits, BiCMOS circuits, photodiodes, phototransistors, LEDs, laser diodes, and solar cells. Note that MOSFETs are majority carrier devices that are usually relatively heavily doped, so they are not very sensitive to displacement damage because large levels of damage are needed to cause significant majority carrier removal via compensation of donors or acceptors. Thus, silicon CMOS circuits are not particularly susceptible to displacement damage.

A major contribution to displacement damage relevant to space missions is protons, which also contribute to TID, producing a mixture of effects (the next section discusses an example). For those cases in which displacement damage can be observed separately from other effects, and assuming equivalent fluence conversions (see Section 3.4) are applicable, device susceptibility to displacement damage in the space environment can be measured in terms of susceptibility to neutrons. The susceptibility to neutrons of various types of discrete devices and integrated circuits is shown in Figs. 5.4 and 5.5, taken from Ref.[5.1]. Note that GaAs circuits are relatively hard. This is because of their relatively short minority carrier lifetimes and high majority carrier concentrations.

Until recently, displacement damage in microelectronic devices in the space environment was sometimes given inadequate attention. An example is the 4N49 optocoupler used on the TOPEX/Poseidon satellite. These devices were previously tested for TID using gamma rays, which produce negligible displacement damage. Data indicated that these devices do not degrade significantly at levels like 100 krads(Si), so the expectation was that a proton environment producing this dose or less should not be a problem. However, there were observed failures on the satellite at doses less than 30 krads(Si). It was later determined that the failures were caused by proton-induced displacement damage in the LED and phototransistor making up the 4N49, which produces a large degradation in the current transfer ratio [5.3]. Although displacement damage effects in solar cells are normally anticipated, this example helped to increase community awareness of the importance of displacement damage in other devices, even behind considerable spacecraft shielding.

Displacement damage is also important to precision reference devices. Because of very demanding requirements, such devices can be affected by changes in internal parameters that are considered to be second-order effects for more conventional devices. Circuit simulations (computer calculations) and measurements performed on several bipolar precision reference devices demonstrate that high-precision circuits can be

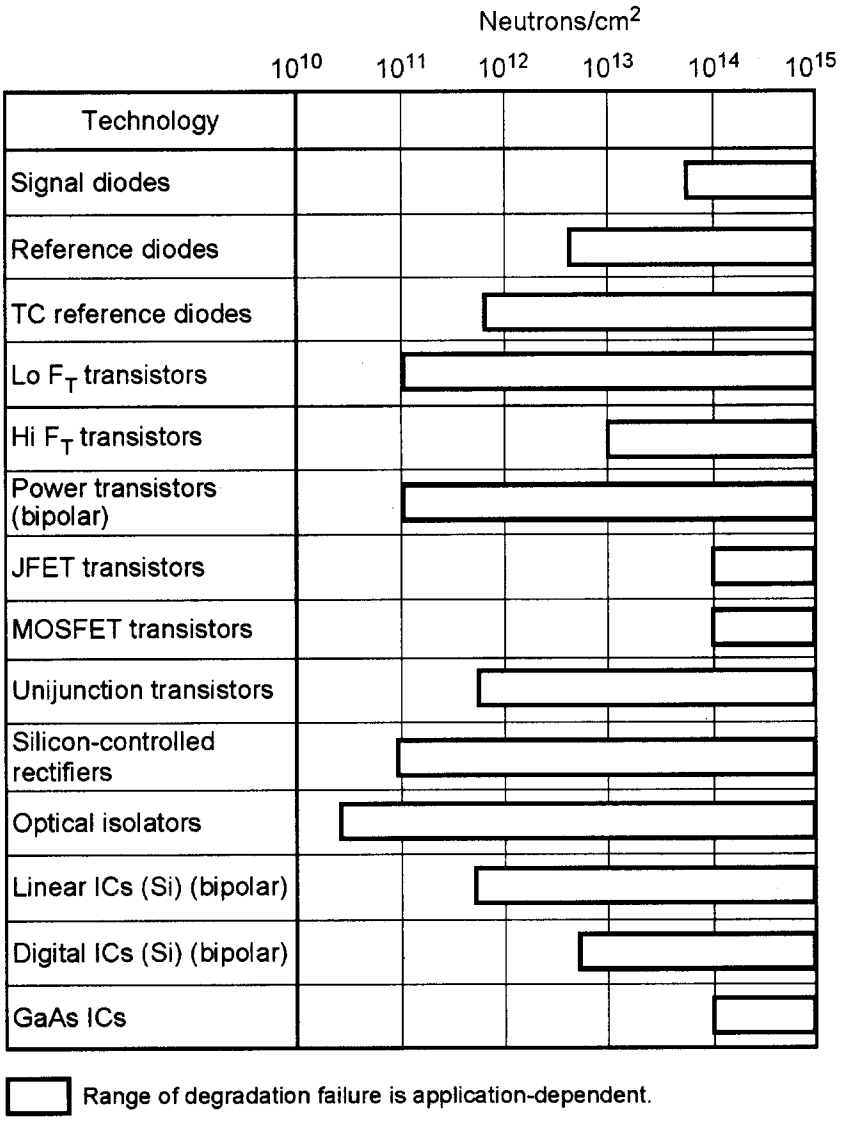

Fig. 5.4: Neutron hardness levels for discrete devices from Ref. [5.1].

sensitive to low radiation levels, even if they are fabricated with components that are relatively unaffected by radiation [5.4]. Furthermore, simulations and measurements both show that protons may produce considerably more degradation than equivalent (in terms of TID) gamma irradiation because of displacement damage [5.4].

5.4 Multiple Failure Modes in Mixed Signal Devices

Mixed signal devices, such as analog to digital converters (ADCs), can exhibit multiple failure modes, particularly when exposed to protons which contribute to both TID and displacement damage. For devices containing both CMOS and bipolar elements (BiCMOS), there can be at least four radiation-induced failure modes:

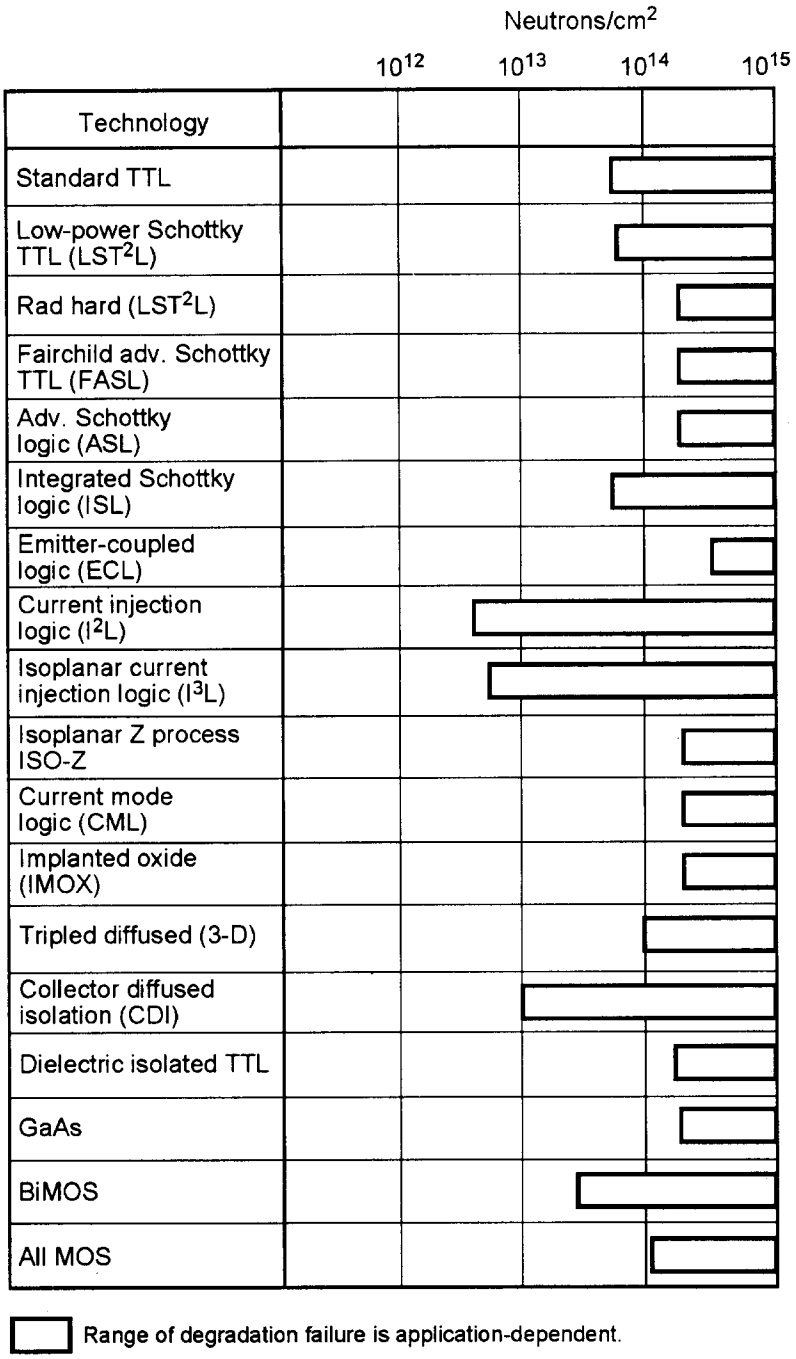

Fig. 5.5: Neutron hardness levels for integrated circuit families from Ref.[5.1].

1. Failure due to trapped holes in the gate and field oxides of the MOSFET elements.
2. Failure due to interface state effects at the SiO2-Si interface of MOSFETs.
3. Failure due to ELDR effects in bipolar sections of the device.
4. Failure due to displacement damage effects in bipolar transistors.

In addition, the failure mode can depend on the environment and how the device is operated while it is undergoing exposure. Thus, it is difficult to draw general conclusions about the hardness levels of such devices. Also, test data can show inconsistent trends

regarding dose rate effects. For dose rates that are not too low, device degradation at the time a prescribed dose has accumulated can be worse at higher dose rates, as expected for TID in CMOS devices not exhibiting rebound. However, sufficiently low dose rates can show the opposite trend, consistent with ELDR effects in the bipolar sections. Therefore, if limited test data show only one trend, it is risky to conclude that this trend applies to all conditions. Furthermore, there can be conditions such that device degradation is dominated by displacement damage effects in the bipolar sections. It is very difficult to characterize such devices sufficiently well so that their response to a space environment can be predicted.

5.5 A Preliminary Discussion: The Meaning of a Collapsed DR

The remaining sections discuss various types of single event effects (SEE). Several sections refer to a collapsed reverse biased depletion region (DR), so an explanation is given here of what that means. A reversed biased DR in almost any type of device can collapse, and the source of charge carriers leading to this is irrelevant, but for illustration we start with a simple reverse biased p-n junction diode. The source of carriers for this illustration is an ion track that liberates carriers directly within the DR and/or outside the DR but with the carriers moving into the DR via drift/diffusion (drift becomes relevant after the collapse has already been initiated).

Prior to the ion hit, all of the applied diode voltage is across the DR. After the ion hit, the DR can become flooded with carriers so that it partially collapses. This means that carriers occupying the volume defined by the pre-ion-hit DR are sufficiently numerous so that, following a spatial rearrangement of carriers produced by the electric field within the DR, a significant number of uncompensated ionized dopants become compensated or screened by majority carriers on either side of the metallurgical junction. The region containing uncompensated ions becomes narrower, i.e., the post-ion-hit DR is narrower than the pre-ion-hit DR. The reduced DR width is often (there are exceptions) accompanied by a reduced voltage across the DR, i.e., the collapsed DR does not support the full device voltage. If the device biasing voltage is fixed, the voltage lost by the DR must appear somewhere else in the device. For the simple diode just considered, this "somewhere else" is the device substrate, producing an effect that has been called funneling [5.5]. For some devices containing other p-n junctions, the voltage lost by the collapsed DR can appear as a change in biasing conditions for another DR. This will be seen to be relevant to the bipolar transistor effect (Section 5.6) and to latchup (Section 5.8). In some structures, the voltage lost by a DR can appear across an oxide, which is relevant to SEGR (Section 5.9).

As previously acknowledged, a DR collapse in the sense of a reduced width is not always accompanied by a reduced voltage. This is because the collapse process is not only a compensation of ionized dopants by majority carriers, reducing the width of the region (the post-ion-hit DR) still occupied by uncompensated ions. The process can also populate the post-ion-hit DR with an excess of minority carriers on either side of the metallurgical junction. The charge of these minority carriers adds to the charge of the

uncompensated ions, increasing the net charge density in the DR. Under certain conditions, the increased charge density in the DR combined with the reduced DR width can result in the voltage across the DR being very nearly the same as it was before the ion hit.

Under steady-state conditions, it is possible to analytically solve the nonlinear drift/diffusion equations (even for an arbitrary three-dimensional geometry if solutions are expressed as nonlinear functions of solutions to linear boundary value problems), so that the voltage lost by the DR can be calculated [5.6]. For such steady-state conditions, there is no impulsive ion track, but there can be an arbitrary steady-state carrier injection and/or photogeneration. An interesting conclusion is that there is a profound difference between the p^+-n junction DR and the n^+-p junction DR in terms of the lost voltage produced by carrier flooding. For a p^+-n diode, the voltage lost by the DR from photogeneration is negligible unless the photogeneration is spatially concentrated near the DR. In particular, the DR will support virtually all of the device voltage if the photogeneration is spatially uniform throughout the diode, even if the excess carrier density greatly exceeds the doping density. In contrast, the same conditions will cause the voltage across the DR in a n^+-p diode to be significantly less than the device voltage. Most of the voltage lost by the DR appears across a narrow region near the p contact, producing an intense electric field there.

Analytical solutions have not yet been found for the transient problem of an impulsive ion track, but computer simulation results provide some insight regarding the voltage lost by a collapsed DR [5.7, 5.8]. The results are qualitatively consistent with the steady-state predictions. For a p^+-n bulk diode hit by an ion track, the DR voltage is greatly reduced (compared to the pre-ion-hit voltage) immediately after the collapse, but there is also a fast recovery. The DR regains most of its prior voltage before very much charge collection occurs. For a p^+-n-n^+ epi diode, the recovery (of voltage but not DR width) is virtually 100% and virtually instantaneous from the point of view of charge collection analysis. In contrast, the DR voltage remains significantly reduced during most of the charge collection process for the opposite doping type. For the n^+-p bulk diode containing an ion track not long enough to go clear through, most of the voltage lost by the DR appears across the substrate section that is below the lower track end, where the carrier-modulated conductivity is least. For the n^+-p-p^+ epi diode, most of the voltage lost by the DR appears across the high-low junction, producing an intense electric field there.

5.6 SEU in Digital Circuits

The earliest and best-known example of SEE is single event upset (SEU) in digital circuits, in which a heavy ion (or proton- or neutron-induced nuclear reaction product) deposits enough energy in a bistable element to switch the logic state of that element. A common example is an SRAM. The process for upset is shown in Fig. 5.6 for a 6-transistor memory cell as might be found in a CMOS SRAM. At the top of the figure, a cross-sectional view of one inverter of the memory cell is shown along with an ion strike by a GCR particle on a sensitive node (the junction of the OFF p-channel transistor) of

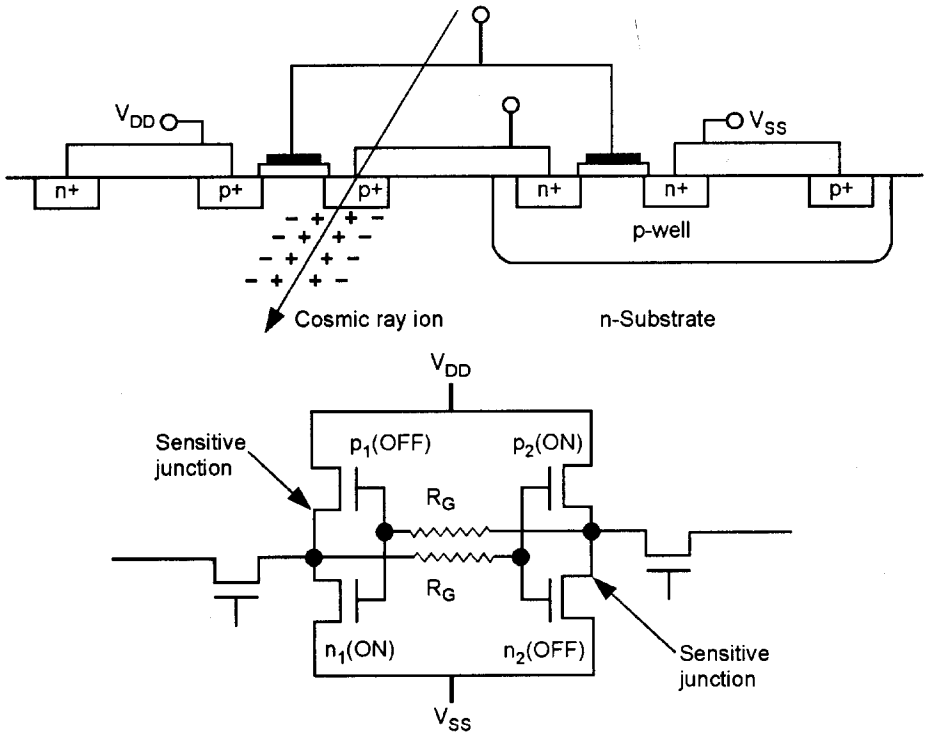

Fig. 5.6: Single event upset in a typical SRAM memory cell.

the memory cell. Mobile e-h pairs liberated by the ion are transported through a combination of drift and diffusion, so that some charge is collected by the junction, i.e., a current flows across the junction. In this illustration, the hit transistor is the p_1(OFF) transistor in Fig. 5.6, so there is a transient current pulse through this transistor. The direction of the current is the same as the current produced without an ion hit but with the p_1 transistor temporarily shorted (turned on). In this respect, the ion hit mimics a temporary turn-on of the p_1 transistor. The transient current pulse through the p_1 transistor also flows through the n_1(ON) transistor, but the latter transistor has a finite channel conductance, so the current pulse through it produces a transient voltage at the n_1 transistor drain, which is conducted to the gates of the p_2 and n_2 transistors in the figure. This voltage transient at the gate tends to reverse the states of the p_2 and n_2 transistors. If the transient is strong enough to reverse these states for a sufficiently long time, feedback to the p_1 and n_1 transistors will reverse the states of these transistors, and the entire memory cell stabilizes in the new state.

The process described above is a contest between a voltage transient that tends to change the device state, and circuit interactions that tend to hold the device in its present state. Unlike a typical unhardened memory cell, the coupling lines between the two inverters in Fig. 5.6 contain resistors R_G. These resistors slow the feedback between inverters via an RC time constant. A longer time constant (i.e., a smaller current through a coupling line) requires a longer duration voltage transient at one inverter in order to produce a specified voltage transient at the gate of the opposite inverter. For a fixed transient duration at the struck inverter (the time required for the ion track to dissipate), a larger R_G results in a smaller voltage transient at the gate of the opposite inverter. This

hardens the cell against SEU, but at the expense of reducing the speed of the memory. Also, the polysilicon resistors are quite temperature sensitive.

When analyzed in complete detail, SEU in an SRAM is extremely complex, requiring computer solutions for both circuit interactions and the charge collection process. However, for conceptual purposes (and also for quantitative modeling when high accuracy is not required), the end result of the circuit interactions can be approximated by assuming the existence of a critical charge and a device time constant, having the property that an SEU occurs if the charge collected from the ion track, over a time equal to the device time constant, exceeds the critical charge. The critical charge and device time constant are both functions of device construction and operating conditions (and can also be different for the p- and n-transistors in Fig. 5.6), but are modeled as being independent of the ion track.

Even when complex circuit interactions can be replaced by the above simple assumption, there still remains the issue of charge collection from the ion track, which is also complex. The lowest level of complexity is often applicable. This involves only the susceptible node, e.g., the p-drain in the upper part of Fig. 5.6. In this case, the charge collection process resembles that in a simple reverse biased diode (as discussed in Section 3.3), with no other structures participating.

The next level of complexity includes interactions with other junctions, e.g., the p-source in Fig. 5.6. Charge collected from the ion track by the drain can cause the drain junction DR to become flooded with carriers, so that it partially collapses. The DR supports less device voltage than it did prior to the ion hit as discussed in Section 5.5. If the device terminal voltages are sufficiently clamped, some of this voltage lost by the drain DR must appear somewhere else in the device. Under certain conditions, some of this voltage can appear as a forward biasing of the source DR, producing emitter action. There is now another supply of carriers in addition to those supplied by the track, giving the appearance of an "amplified track". This phenomenon has been given several names, including barrier lowering [5.9] and the bipolar transistor effect [5.10]. The strength of this effect increases with decreasing distance between source and drain. For CMOS devices, the effect appears to be more significant for nodes in a well (e.g., the n-drain in the p-well in Fig. 5.6) than for nodes outside (e.g., the p-drain in Fig. 5.6) [5.10]. This effect is very important in SOI devices.

Another type of device extensively studied with regards to SEU is the DRAM. There are a variety of potential SEU mechanisms in this device, which are too complex to discuss here. The interested reader is referred to Ref. [5.11]. We discuss here only the simplest type of SEU, but this is also the most common type. This type is a cell error created by the discharge, via an ion track, of a charged capacitor. A DRAM cell uses the storage of charge in a storage capacitor to represent information, e.g., a charged capacitor might be interpreted as a 1, and an uncharged capacitor might be interpreted as a 0. The method by which the capacitor becomes charged or uncharged during normal operation involves device structures that are relevant to some types of SEU, but these structures are not relevant to the simple process discussed here. We will assume that a capacitor is

initially charged, but ignore the presence of the structures that produced this initial condition. The capacitor is floating in the sense that there is no biasing voltage that actively maintains this charge. The charge is held in place by being in the form of majority carriers above a reverse biased p-n junction. The lower side of this junction is the device substrate. A circuit equivalent is a charged capacitor in series with a diode, with the diode oriented to block the current when in the direction that would discharge the capacitor. However, the rectifying property of a p-n junction is lost when minority carriers are present. When reversed biased, the junction will block the majority carrier current, so there is normally (without an ion track) only a small leakage current associated with minority carriers. This leakage current tends to discharge the capacitor, so the capacitor charge has to be refreshed periodically. If an ion track produces e-h pairs, the holes on the n-side and/or electrons on the p-side can cross the junction, producing an enhanced leakage current that quickly discharges the capacitor.

The concept of a critical charge is quite literal for this simplest type of SEU in DRAMs. It is the amount of charge that has to be lost by the capacitor in order for its state to be interpreted (by other circuitry) as being uncharged. An SEU occurs if the charge collected at the junction from an ion track exceeds this critical value. The critical charge is comparatively small for DRAMs, and the device time constant is effectively infinite (because the refresh time is much longer than the life of the track), so DRAMs are particularly soft with regards to SEU. Also, the amount of charge that can be collected at a given node is limited, leaving more charge available to be shared by other nodes. The result of this, combined with the critical charge being small and the device time constant being effectively infinite, is that DRAMs are notorious for multiple-bit upsets; one ion track can upset many bits. However, because DRAMs have very large memory capabilities, their use in space systems is becoming more common in spite of their susceptibility to SEU. A common practice is to use an error detection and correction circuit (EDAC) that corrects single-bit upsets. Although multiple-bit upsets within a single word are not corrected, EDAC is still effective if the device architecture is arranged so that multiple-bit upsets produced by one ion hit (which are grouped in terms of physical location, unless the ion trajectory is at a grazing angle) are distributed among different words, i.e., there are no multiple-bit upsets within a common word from one hit. The device is still susceptible to uncorrectable SEUs from multiple ion hits.

SEU can occur in almost any integrated circuit containing bistable (digital) elements. For devices containing both digital and analog components, such as an ADC, an SEU can occur from either an ion hit to a digital component, or an ion hit to an analog component. In the latter case, the ion hit creates a voltage transient in an analog component, which in turn changes the state of a digital component.

5.7 SET in Analog Circuits

Recall from the SRAM discussion in the previous section that SEU starts with an ion-induced transient, which then changes the state of a digital element. For an analog device containing no digital elements (e.g., a photodiode), there can still be an ion induced

transient, called a single event transient (SET). If this device is electrically connected to a digital device, the SET produced in the analog device can produce an upset (state change) in the digital device. This is similar to SEU in an ADC previously discussed, except that analog and digital sections are now different devices instead of different parts of the same device. This distinction makes analog devices difficult to characterize in terms of SET, because it may not be known in advance what the companion digital circuit will be. The transient amplitude and pulse width required to upset a companion digital circuit depends on the companion circuit. Furthermore, the transient amplitude and pulse width produced by a given ion hit to the analog circuit may depend on loading conditions seen by this circuit, which also depends on the companion circuit.

For a given input voltage, output loading, and incident particle type (species and energy), a distribution of ion hit locations within the device sometimes produces a distribution of output pulses of varying amplitude and width. Therefore, the occurrence of SET does not become a simple "yes or no" until after some criteria is selected to define an event. The criteria will be based on pulse amplitude, pulse width, or both (e.g., time integral of the pulse). If the companion circuit is not known in advance, then each of several criteria and loading conditions might be considered, in which case device susceptibility to SET becomes a function of the selected criteria and loading conditions.

5.8 SEL

Latchup can occur in a device containing a four layer n-p-n-p structure which creates a bistable state. One state is a low-current state, and the other (the latched state) is a high-current state. A transition from one state to the other can be triggered by a transient in biasing conditions or by other kinds of disturbances in the device. The event in which a heavy ion (or proton- or neutron-induced nuclear reaction product) causes a transition to the high-current state is called a single event latchup (SEL). After the latched state has been initiated, it would be stable, except that large currents can heat the device to destruction. The current during a latched state is not always large enough to destroy the device, in which case the device remains in this state until the power is turned off. Whether or not a self-sustaining latched state exists can depend on the external circuit. If either the maximum current or maximum voltage that can be supplied by the external circuit are sufficiently limited, the device will be immune to latchup. It is therefore possible for a device to be susceptible to latchup at one operating voltage, but immune at some other, lower, operating voltage. Immunity in this respect means that no triggering mechanism can induced a latchup (although transients can be induced) because there is no self-sustaining latched state. This will be called inherent immunity, which is distinguished from immunity to a particular triggering mechanism.

The most familiar example of latchup is in a CMOS inverter used in an SRAM. An example is shown in Fig. 5.7. Superimposed in this figure is a lumped element circuit frequently used to describe the device. The use of a two-transistor equivalent circuit originated in the early days of semiconductor physics with the thyristor (limitations of the equivalent circuit were also known a long time ago [5.12]). Recognizing that the CMOS

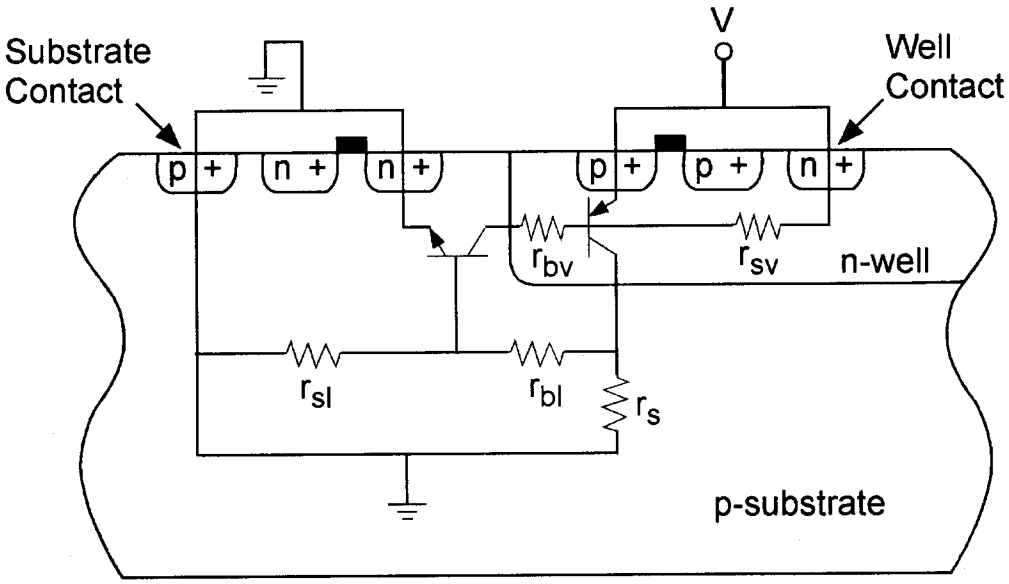

Fig. 5.7: Parasitic latchup structure inherent to CMOS devices.

device also contains a four-layer structure inspired the use of a similar representative circuit for this device [5.13], although various resistors are included to represent various current paths. The parasitic transistor in the substrate is called the lateral transistor, and the one in the well is called the vertical transistor.

The static representative circuit in Fig. 5.7 contains no elements describing triggering mechanisms and is only intended to describe the two (if there are two) steady states that the device can be in. To the extent that the circuit applies, it would predict the device holding voltage, which is the voltage that an external circuit must be able to provide in order for the device to have a self-sustaining latched state. However, with the present availability of modern computer codes, it is no longer necessary to rely on lumped circuit models. This is fortunate because such models do not represent the distributed effects that occur in real devices. An interesting observation from computer simulation results is that a latchup produces a total collapse of the well-substrate junction DR [5.14-5.18]. The DR is completely obliterated and supports virtually zero voltage. It was shown that an implication of this for epi devices is that, when in the latched state, the device behaves like a diode with a modulated conductivity. The conductivity is modulated by excess carrier density. This is very important (though not included in lumped circuit representations) because the carrier density during a latchup (or during transient conditions leading to a latchup) can be much greater than the doping density, profoundly influencing the conductivity throughout the device. An analytical model reflecting this observation was developed [5.14], and was found [5.14-5.16] to be very successful (predictions agree with both computer simulations and measurements) at predicting holding voltages for a wide variety of device constructions.

A knowledge of the holding voltage can determine whether a device is inherently immune to latchup in a specified application, but does not answer the question of what kinds of triggering mechanisms can produce a latchup in a given device that is not

inherently immune. In particular, there is still the question of whether a particular ion hit can produce SEL in a device that is not inherently immune. This requires a time-dependent analysis. A quantitative transient analysis of SEL is complex and presently can only be performed by a computer simulation. However, computer simulations of the transient problem [5.17, 5.18] reveal some basic processes that make a *qualitative* explanation of SEL fairly simple, as explained below.

During normal operation, i.e., before the ion hit, all of the device voltage (V in Fig. 5.7) is across the reverse biased well-substrate junction. An ion hit near this junction causes the DR to become flooded with carriers, so that it partially collapses and supports less voltage than it did prior to the ion hit. This voltage lost by the DR must appear somewhere else in the device. Part of this lost voltage appears as a forward biasing of the p-source DR in the n-well, and another part appears as a forward biasing of the n-source DR in the substrate. The remainder of this voltage is across quasi-neutral regions between structures. The way the voltage is divided between DRs and quasi-neutral regions varies with time, because the conductivity (which is modulated by the carrier density) varies with time as carriers move. The result is a delay before the n-source DR becomes forward biased. However, as soon as either source DR becomes forward biased, emitter action occurs. The emitted holes by the p-source (minority carriers in the well), and the emitted electrons by the n-source (minority carriers in the substrate) are able to cross the well-substrate junction. These carriers contribute to the flooding of the well-substrate junction DR, maintaining (and even increasing) the partial collapse initiated by the ion track. The state becomes self-sustaining; the voltage lost by the collapsed well-substrate junction DR contributes to forward biasing of other junctions, leading to emitter action, which maintains the DR collapse. However, this assertion that the state is self-sustaining merely noted a qualitative consistency. Quantitative consistency applies only to devices having latched states, and even then involves more details (such as a general requirement for stability and the ability of conductivity modulation to satisfy this requirement [5.14]) than described here.

A device can be made to be less susceptible to latchup, even if it is not made to be inherently immune, when a design change requires a stronger triggering mechanism to cause a latchup. Susceptibility can be reduced by reducing the tendency for various junctions to become forward biased. One way to do this is to weaken the mechanism contributing to the forward biasing, by increasing the distance between the well-substrate junction and either (or both) FETs. Another way is to more strongly clamp the junction DR biasing voltages so that the biasing is more difficult to change. One way to do this is to decrease the distance between the substrate contact and n-source, and/or the distance between the well contact and p-source in Fig. 5.7. Another way to more strongly clamp the n-source junction biasing is by building the device in a thin epi layer so that the region outside the junction is close to the heavily-doped region where the potential is effectively clamped. The use of an epi layer is a particularly effective hardening approach for SEL, not only because the junction biasing voltage is more strongly clamped, but also because it reduces charge collection from an ion track, weakening the triggering mechanism. However, this does not guarantee immunity in all devices.

5.9 SEGR

The best-known example of single event gate rupture (SEGR) is in a power MOSFET such as shown in Fig. 5.8. For the case illustrated, the n^+ diffusion is the source, the p-region is the body, and the n-epi is the drain. SEGR is an ion-hit-induced rupture of the gate oxide. This is a destructive event because the dielectric and gate electrode material melt and mix, producing either an ohmic short or a rectifying contact through the dielectric [5.19].

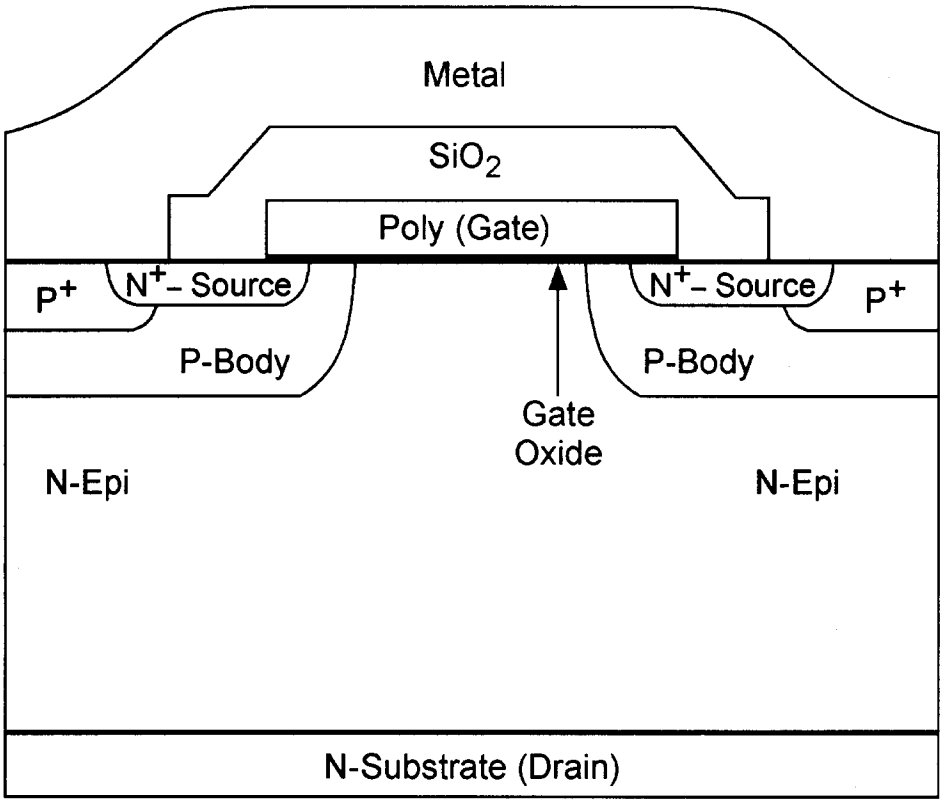

Fig. 5.8: Illustration of a power MOSFET.

The device is susceptible when in the off state, which is the state in which a high voltage from a power supply is across the device. In this state, there may be a moderate potential difference between the gate and source, with polarity consistent with the device being off, and a large voltage between the source and drain. For the doping type illustrated in Fig. 5.8, the drain is positive relative to the other contacts. The high voltage is across a DR in the epi layer. This DR can be described in several ways. If we look at the n-epi under the oxide and use MOS capacitor terminology, this region is the DR produced by a MOS capacitor in the depletion mode. If we look in the n-epi near the p-region and use p-n junction terminology, this region is the DR of a reverse biased p-n junction. By using a thick and lightly doped epi so that the DR is thick, it can support a large voltage without a breakdown. There is never a large enough voltage across the oxide for a breakdown to occur there during normal operation, but an ion hit can change this situation.

Whether or not an SEGR occurs depends on which of two quantities is larger. One quantity is the electric field required to break down the oxide (called the critical electric field), and the other quantity is the electric field that is actually present in the oxide. An ion hit to the oxide has detrimental effects on both quantities. Although the details are not yet well understood, an ion hit weakens the oxide in the sense that the critical electric field is temporarily reduced by the hit. The ion hit also increases the electric field actually present in the oxide, by collapsing the DR so that it gives up some of its voltage to the oxide as explained later.

There is still some controversy regarding the oxide weakening effect. One of the earliest investigations of this subject correlated the weakened critical electric field with ion LET and incident angle, and noted that the correlation is independent of past irradiation history of the oxide [5.19]. The controversy that came later is summarized in the most recent papers on the subject [5.20-5.23] and references therein. There appears to be a consensus, at this time, that breakdown is the result of a single ion interaction, with no dependence on residual damage from previous ion strikes. There are still disputes regarding the mathematical equation expressing the critical electric field in terms of the pertinent ion parameter, and there are even disputes regarding what the pertinent ion parameter is (e.g., LET versus atomic number).

The physical origin of the enhancement of the electric field present in the oxide is less controversial (although a simplified analytical model representing this effect is still an area of active research) because computer simulations can treat this subject. For illustration, consider a normal-incident ion track along the centerline in Fig. 5.8. The ion hit causes the DR in the epi to collapse in the sense of a reduced width. A reduced width does not, by itself, imply that the DR gives up any of its voltage to another device structure. It does so in this case, but additional explanation is needed. Simulations show that the collapse causes the width to be reduced more than the DR voltage, in the sense that the spatial average electric field in the DR is larger after the collapse. Also, the electric field in the DR drives minority carriers up to the oxide, producing a layer of minority carriers (holes in Fig. 5.8) in the silicon just below the oxide. A charge layer produces a spatial change in an electric field, i.e., the field above the layer differs from the field below by an amount proportional to the amount of charge in the charge layer. Therefore the electric field in the silicon at the upper boundary of the charge layer equals the electric field below the charge layer plus a contribution from the charge layer. The polarities are such that the charge layer strengthens the field above it. Therefore, the electric field at the oxide boundary in the silicon above the charge layer has one enhancement from a stronger field in the DR below the charge layer, and a second enhancement from the charge layer. Taken together, the enhancement is very strong. The electric field in the oxide is proportional (the proportionality constant is a ratio of dielectric constants) to the electric field in the silicon at the oxide boundary above the charge layer, so the electric field in the oxide is greatly enhanced. This increases the voltage in the oxide, so the conclusion is that the collapsed DR does give up some of its voltage to the oxide.

If the p-region in Fig.5.8 were far removed so that the middle of the oxide could be viewed as an isolated MOS capacitor, the only relevant device terminal voltage would be the gate-to-drain voltage V_{GD}. However, the close proximity of the p-region will produce emitter action if the potential in the silicon just below the oxide is less than the source potential, by forward biasing the junction between the n-epi and p-region. The result of this interaction is that the nominal (pre-ion-hit) voltage across the oxide is approximately the gate-to-source voltage V_{GS}, leaving the drain-to-source voltage V_{DS} as the remaining voltage to appear across the DR in the epi. The post-ion-hit electric field in the oxide is the nominal field, which is controlled by V_{GS}, plus the enhancement which depends on the voltage V_{DS} formally across the DR. Therefore, the post-ion-hit electric field in the oxide depends on both V_{GS} and on V_{DS}.

As explained above, the electric field in the oxide shortly after an ion hit depends on both V_{GS} and V_{DS}. Therefore, device susceptibility to SEGR depends on the ion and on both voltages. For a given ion type (species and energy), threshold conditions for SEGR can be represented as a curve in the V_{GS}, V_{DS} plane. A point on one side of the curve indicates a value of V_{GS} and V_{DS} such that the given ion can create a SEGR, and a point on the other side of the curve indicates a value of V_{GS} and V_{DS} such that the given ion cannot create a SEGR.

Device susceptibility to SEGR can be reduced by increasing the oxide thickness, because this decreases the electric field in the oxide corresponding to a given voltage across the oxide. However, increasing the oxide thickness increases device susceptibility to TID, so the practicality of this hardening approach depends on whether SEGR or TID is the greatest concern for a given application. Another way to reduce the risk of SEGR is to reduce the voltages applied to the device. However, an approach that is *not* effective is to limit the current through the device. Various capacitances in the device can store enough electrical energy to produce a SEGR, once initiated by an ion hit. For this reason, SEGR is destructive even under carefully controlled laboratory conditions.

It should be noted that SEGR is not limited to power MOSFETs. A simple MOS capacitor can be susceptible, although the required voltages are typically on the order of those applied to some of the higher-voltage devices, such as power MOSFETs. However, if progressing miniaturization results in decreasing gate oxide thicknesses, SEGR could become a problem in other devices. In fact, SEGR might have already been observed in a DRAM. The failure mechanism was not conclusively identified, but appears to be more consistent with SEGR than with any other known failure mechanism [5.24].

A fortunate property of SEGR is the dependence of device susceptibility on the direction of the ion trajectory. The directional dependence of the critical electric field, which reflects oxide weakening, appears to depend on oxide thickness [5.22]. The critical electric field was found to be nearly isotropic for the thinner oxides, but a stronger function of direction for thicker oxides [5.19, 5.22]. However, for those cases in which the directional dependence is significant, the worst-case direction is at or close to normal incidence. This is in contrast to SEU, SEL, and SET, in which a given ion is often more capable of producing an event at the larger angles. For SEGR, the susceptibility measured

in laboratory tests utilizing small angles is expected to be either the worst-case susceptibility or close to it. Therefore, SEGR rates or probabilities in known heavy ion environments are sometimes estimated by assuming an isotropic cross section equal to the cross section measured at normal incidence. Such estimates are expected to be upper bounds for the actual rates or probabilities. Better estimates would utilize the measured directional dependence of device susceptibility, but it is difficult to obtain such an extensive data set with adequate counting statistics because of the destructive nature of SEGR.

5.10 SEB

Single event burnout (SEB) has been observed in bipolar transistors and in n-channel power MOSFETs (e.g., Fig. 5.8) in the off state, supporting a large drain to source voltage. SEB has not yet been reported in p-channel power MOSFETs, but it is possible that some such devices are susceptible.

All types of SEE discussed in the earlier sections are quantitatively difficult, but qualitative explanations could be made simple enough to be appropriate for this introductory course. For SEB, unfortunately, even a qualitative explanation is difficult if thorough enough to explain all pertinent aspects (e.g., why an intense electric field is located where it is). Interested readers can find an explanation in Ref.[5.25]. Instead of attempting to give a complete explanation of SEB, the discussion below merely states some characteristics of this phenomenon.

A characteristic that SEB has in common with SEL is that an ion hit to a susceptible device can trigger a transition from the normal operating state to a high-current state, which leads to junction heating and eventual burnout of the device. Device susceptibility depends on the ion and on biasing voltages, with greater susceptibility produced by larger voltages.

A distinguishing characteristic of SEB is that it involves a regenerative feedback between a bipolar transistor action and avalanche multiplication. This is one important difference between SEB and SEGR. For SEGR, the presence of the p-body in Fig. 5.8 has quantitative effects on device susceptibility (by influencing the nominal voltage across the oxide), but is not qualitatively essential. In fact, SEGR can occur in a simple MOS capacitor. In contrast, the p-body and n^+-source in the figure are both essential for SEB. The n^+-source, p-body, and n-epi form the emitter, base, and collector (respectively) of a parasitic bipolar transistor. When the device is in the high-current state produced by SEB, the emitter-base junction is forward biased, producing emitter action. Electrons from the emitter cross the base-collector junction, producing an electron current in the n-epi. In addition, a region in the n-epi contains an electric field strong enough to produce avalanche multiplication, which creates holes that are able to move up and cross the base-collector junction. This maintains the forward biasing of the emitter-base junction.

5.11 Other Types of SEE

State-of-the-art DRAMs are complex devices that often include basic DRAM cells and the necessary peripheral circuitry for accessing and writing/reading these cells. In particular, several operating modes may be built into the device in order to facilitate various test modes of operation. Normally, a command sequence is used to place the DRAM in a test mode, but an ion hit to a circuit component that controls the operating mode can throw the DRAM out of the normal operating mode and into one of its test modes [5.26]. These single event functionality interrupt (SEFI) events render the device inoperable until it is placed back in normal mode by specific commands. Unlike SEU, which produces an error in one logic element, SEFIs manifest themselves as bursts of large numbers of errors from one ion hit. Although the cross section for SEFI is smaller than that for SEU, the effect has much greater impact on DRAM functionality. The DRAM is probably the first device to be extensively studied with regards to SEFI, but SEFI can also occur in other devices, such as microprocessors.

Single particle induced errors have been observed in DRAMs and in some SRAMs which are permanent or semi-permanent in nature, in contrast with the more conventional SEU. These "stuck bits" or "hard errors" remain in the device despite power cycling or reloading of a memory. Analysis of data obtained from a DRAM revealed that these hard errors fell into two clearly distinguished classes [5.24]. One class is believed to be caused by SEGR as discussed in the Section 5.9. The other class is believed to be caused by the microdose effect discussed in Section 3.2. Transistors in advanced memories are now small enough so that a single ion can deposit enough ionizing energy in an oxide to create the equivalent of a TID effect within a single transistor. Some microdose-induced hard errors anneal, as might be expected from TID effects, but annealing was not observed in the SEGR-induced hard errors.

SEE testing and research [5.27, 5.28] on field programmable gate arrays (FPGAs) found a destructive failure mechanism called single event dielectric rupture (SEDR). The programming of some FGPAs utilizes thin dielectrics called antifuses. Such a device contains a grid of conducting lines that are initially electrically isolated from each other. Where two perpendicular lines cross, they are separated from each other by an antifuse. By applying an electrically induced dielectric breakdown in selected antifuses, which establishes a low resistance connection between the conductors, connections are made between selected conductors and the device becomes programmed. A typical programmed FPGA will have a few percent of the antifuses connected to establish the intended functionality of the device. The concern is with the remaining unconnected antifuses, because an unintentional connection of one of these will alter the programming. A given unconnected antifuse will have a voltage across it when the logic levels on the two crossing conductors are different, the occurrence of which depends on the duty cycle and phase of the two signals on these lines. When a voltage is present, a thin antifuse contains a strong electric field, making it susceptible to SEDR, which is an ion-hit-induced rupture similar to SEGR. This rupture establishes an undesired connection in the same way as the intentional over-voltage used for programming.

References

[5.1] M. Rose, "Updated Bar Charts of Device Radiation Thresholds," Physitron Corp. San Diego, CA, 1990.

[5.2] A.H. Johnston, C.I. Lee, B.G. Rax, and D.C. Shaw, "Using Commercial Semiconductor Technologies in Space," *RADECS 1995 Proceedings*, p. 175.

[5.3] B. Rax, C. Lee, A. Johnston, and C. Barnes, "Total Dose and Proton Damage in Optocouplers," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, p. 3167, Dec.1996.

[5.4] B.G. Rax, C.I. Lee, and A.H. Johnston, "Degradation of Precision Reference Devices in Space Environments," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 6, pp. 1939-1944, Dec. 1997.

[5.5] C.M. Hsieh, P.C. Murley, and R.R. O'Brien, "A Field-Funneling Effect on the Collection of Alpha-Particle-Generated Carriers in Silicon Devices," *IEEE Electron Device Lett.*, vol. 2, no. 4, pp. 103-105, 1981.

[5.6] L.D. Edmonds, *A Theoretical Analysis of Steady-State Photocurrents in Simple Silicon Diodes*, JPL Publication 95-10, March 1995.

[5.7] L.D. Edmonds, "Electric Currents Through Ion Tracks in Silicon Devices," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pp. 3153-3164, Dec. 1998.

[5.8] L.D. Edmonds, "Charge Collection from Ion Tracks in Simple EPI Diodes," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pp. 1448-1463, June 1997.

[5.9] R.L. Woodruff and P.J. Rudeck, "Three-Dimensional Numerical Simulation of Single Event Upset of an SRAM Cell," *IEEE Trans. Nucl. Sci.*, vol. 40, no. 6, pp. 1795-1803, Dec.1993.

[5.10] P.E. Dodd, "Basic Mechanisms for Single Event Effects," *1999 IEEE Nuclear and Space Radiation Effects Conference Short Course*.

[5.11] L.W. Massengill, "Cosmic and Terrestrial Single-Event Radiation Effects in Dynamic Random Access Memories," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 2, pp. 576-593, April 1996.

[5.12] A. Herlet and K. Raithel, "Forward Characteristics of Thyristors in the Fired State," *Solid-State Electron.*, vol. 9, p. 1089, 1966.

[5.13] D.B. Estreich, "The Physics and Modeling of Latch-Up and CMOS Integrated Circuits," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1980.

[5.14] J.A. Seitchik, A. Chatterjee, and P. Yang, "An Analytic Model of Holding Voltage for Latch-Up in Epitaxial CMOS," *IEEE Electron Device Lett.*, vol. 8, no. 4, pp. 157-159, April 1987.

[5.15] A. Chatterjee, J.A. Seitchik, J.H. Chern, P. Yang, and C.C. Wei, "Direct Evidence Supporting the Premises of a Two-Dimensional Diode Model for the Parasitic Thyristor in CMOS Circuits Built on Thin Epi," *IEEE Electron Device Lett.*, vol. 9, no. 10, pp. 509-511, Oct. 1988.

[5.16] D.J. Sleeter and E.W. Enlow, "The Relationship of Holding Points and a General Solution for CMOS Latchup," *IEEE Trans. Elect. Dev.*, vol. 39, pp. 2592-2599, Nov. 1992.

[5.17] T. Aoki, R. Kasai, and M. Tomizawa, "Numerical Analysis of Heavy Ion Particle-Induced CMOS Latch-Up," *IEEE Electron Device Lett.*, vol. 7, no. 5, pp. 273-275, May 1986.

[5.18] T. Aoki, "Dynamics of Heavy-Ion-Induced Latchup in CMOS Structures," *IEEE Trans. Elect. Dev.*, vol. 35, pp. 1885-1891, Nov. 1988.

[5.19] T.F. Wrobel, "On Heavy Ion Induced Hard-Errors in Dielectric Structures," *IEEE Trans. Nucl. Sci.*, vol. 34, no. 6, pp. 1262-1268, Dec.1987.

[5.20] J.L. Titus, C.F. Wheatley, K.M. Van Tyne, J.F. Krieg, D.I. Burton, and A.B. Campbell, "Effect of Ion Energy Upon Dielectric Breakdown of the Capacitor Response in Vertical Power MOSFETs," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pp. 2492-2499, Dec.1998.

[5.21] A.J. Johnston, G.M. Swift, T. Miyahira, and L.D. Edmonds, "Breakdown of Gate Oxides During Irradiation with Heavy Ions," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pp. 2500-2508, Dec.1998.

[5.22] F.W. Sexton, D.M. Fleetwood, M.R. Shaneyfelt, P.E. Dodd, G.L. Hash, L.P. Schanwald, R.A. Loemker, K.S. Krisch, M.L. Green, B.E. Weir, and P.J. Silverman, "Precursor Ion Damage and Angular Dependence of Single Event Gate Rupture in Thin Oxides," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pp. 2509-2518, Dec.1998.

[5.23] L.E. Selva, G.M. Swift, W.A. Taylor, and L.D. Edmonds, "On the Role of Energy Deposition in Triggering SEGR in Power MOSFETs," presented at the IEEE/NSREC Conference in July 1999.

[5.24] G.M. Swift, D.J. Padgett, and A.H. Johnston, "A New Class of Single Event Hard Errors," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 2043-2048, Dec.1994.

[5.25] G.H. Johnson, J.M. Palau, C. Dachs, K.F. Galloway, and R.D. Schrimpf, "A Review of the Techniques Used for Modeling Single-Event Effects in Power MOSFET's," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 2, pp. 546-560, April 1996.

[5.26] G. Swift, L. Smith, and H. Schafzahl, "Single Event Functionality Interrupt: A New Failure Mode for Semiconductor Memories," presented at the *9th Single Event Effects Symposium*, Manhattan Beach, CA, April 1994.

[5.27] R. Katz, R. Barto, P. McKerracher, B. Carkhuff, and R. Koga, "SEU Hardening of Field Programmable Gate Arrays (FPGAs) for Space Applications and Device Characterization," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, p. 2179, Dec. 1994.

[5.28] G. Swift and R. Katz, "An Experimental Survey of Heavy Ion Induced Dielectric Rupture," p. 425 in *Proc. of Third European Conf. On Radiation and its Effects on Components and Systems (RADECS 95)* (1995).

6. TID AND DISPLACEMENT DAMAGE: TEST METHODS, INTERPRETATION AND LIMITATIONS OF DATA

The objective of radiation testing at ground-based facilities is to provide test results that can be used to predict a device response in the actual space environment. A difficulty is that ground-based facilities do not duplicate the environments found in space. The best situation is that in which a radiation effect is understood well enough so that a measured device response in an easily arranged laboratory environment can be used to predict the response in other environments (e.g., the space environment). This is the case for some of the more classical TID effects. If a radiation effect is not so well understood, the laboratory environment is arranged to resemble the space environment as closely as practical considerations allow. However, practical considerations (equipment limitations, test time and/or cost) generally do not permit an accurate reproduction of the space environment, and the interpretation of test data sometimes contains uncertainties. An improved understanding of radiation effects is always a goal. Test methods evolve as more is learned about radiation effects. The sections below outline some of the methods commonly used at this time.

6.1 TID

As mentioned in Section 3.2, the most important TID contributors in most space environments are sufficiently similar to gamma rays so that "a rad is a rad". Device degradation in space from TID alone (i.e., excluding displacement damage discussed in the next section) can be determined by exposing the device to gamma rays produced by a Co^{60} source; subject to the requirement that either the laboratory dose rate versus time duplicates that in the space environment (rarely a practical approach), or that time dependent effects be accounted for when interpreting the data. As mentioned in Section 4.5, the more classical TID effects are sufficiently well understood so that data utilizing moderately high dose rates followed by anneals can be used to calculate device degradation from other dose versus time profiles.

The theory for predicting TID responses in arbitrary environments from measurements in a laboratory environment is fairly simple for those devices (hypothetical or real) in which a particular assumption applies. To state this assumption, let the device be in an environment described by a dose rate $\gamma'(t)$ which is an arbitrary function of time t. The prime denotes differentiation, so $\gamma(t)$ is the dose accumulated up to the time t. The device can be arbitrarily complex, containing many components and having many different measures of functionality (e.g., an offset voltage in one component and a transistor gain in another component). Each measure of functionality is considered separately, so we imagine that a particular measure has been selected. The assumption is that at any time t there exists a parameter S(t) corresponding to the selected measure of functionality and having the following two properties:

Property 1: The selected measure of functionality is some function of S, but not necessarily a linear function. Furthermore, larger values of S indicate greater degradation of the functionality measure.

Property 2: S is a linear functional of the dose rate of the form

$$S(t) = \int_0^t \gamma'(\tau) H(\tau,t)\, d\tau \tag{1}$$

for some function H.

Whether or not the above properties are satisfied depends not only on the device but can also depend on which measure of functionality is being investigated. For grammatical brevity, we will assume that only one measure of functionality is considered relevant, so the above properties are associated with a device. Devices having the above properties will be called ideal. Functionality degradation of ideal devices in an arbitrary environment can be calculated once the function H is known. Experimental determination of H can be done the following way. Select some dose level γ_0 and expose the device to this dose level using a dose rate large enough so that the exposure time is short enough for negligible annealing during this time. After the exposure, the device is allowed to anneal. Let t=0 be the start of the exposure and let $S_0(t)$ denote the function S(t) produced by this particular choice of dose rate function. If the experiment is then modified by starting the exposure at some time t_0 instead of at t=0, the corresponding S function would obviously be $S_0(t-t_0)$. Because the exposure time is short compared to the anneal time, the dose rate can be regarded as being proportional to a Dirac delta function for the purpose of evaluating the integral in (1). Therefore $S(t)=S_0(t-t_0)$ when γ' is γ_0 multiplied by the Dirac delta function with singularity at t_0. Substituting into (1) gives $S_0(t-t_0)=\gamma_0 H(t_0,t)$ so (1) can be written as

$$S(t) = \int_0^t \gamma'(\tau) \frac{S_0(t-\tau)}{\gamma_0} d\tau \tag{2}$$

where the function S_0 can be measured.

In practice, calculations using (2) require knowledge of the dose rate as a function of time in the space environment, and a knowledge of the appropriate function relating the functionality measure (e.g., offset voltage) to S. However, the theory also predicts a simple method for worst-case estimates of device degradation, requiring that the dose accumulated over the mission duration be known, but not requiring any knowledge of how the dose rate varies with time. This application does not even require that we know what the parameter S is, as long as there exists some parameter S (e.g., a measure of the combined charge from bulk trapping and from interface states for a particular set of biasing conditions) having the two properties stated above. To see how this is done, let $S_{0,max}$ be the maximum in t of $S_0(t)$. Let the dose rate γ' be any non-negative function of time producing an accumulated dose over a given mission duration equal to γ_0. For any time t up to the mission duration, an upper bound for S(t) is given by

$$S(t) = \int_0^t \gamma'(\tau)\frac{S_0(t-\tau)}{\gamma_0}d\tau \le \int_0^t \gamma'(\tau)\frac{S_{0,\max}}{\gamma_0}d\tau = \frac{S_{0,\max}}{\gamma_0}[\gamma(t)-\gamma(0)] \le S_{0,\max}\,.$$

The equality applies if all the dose was accumulated over a time interval short compared to the anneal time, and if the observation time t is the worst possible time. This is the worst functionality degradation that can occur in an ideal device exposed to an accumulated dose γ_0.

The above conclusion leads to a simple experimental method, valid for ideal devices, for worst-case estimates of device functionality degradation when accumulated dose is known but the time distribution is unspecified. The method is to apply worst-case biasing conditions while exposing the device to the specified dose, at a dose rate large enough so that insignificant annealing occurs during the exposure. Then let the device anneal while periodically measuring its properties. Assume for illustration that the property of interest is offset voltage. During the annealing, the offset voltage will vary with time. The largest offset voltage that is ever observed is a worst-case estimate of the largest offset voltage that will occur in the space environment. In practice, elevated temperatures may be used to accelerate the annealing so that the test is less time consuming. Also, in practice, a dose level γ_0 might not be specified in advance, but multiple tests can be performed using various dose levels, producing a plot of worst-case functionality degradation as a function of dose level. Some details of the test method are discussed in the military standard MIL-STD-883, Method 1019.5, but other details are deferred to test plans which may vary from case to case.

All of the above discussion refers to ideal devices. Unfortunately, not all real devices can be approximated as ideal. A notable example is a device exhibiting the ELDR effect discussed in Section 4.6. A method to deal with this has been suggested [6.1]. The method utilizes a dose rate that is convenient for laboratory work combined with an elevated temperature as an attempt to duplicate the effects of a very low dose rate at the actual temperature. This method may become useful, but there are presently some problems that need to be worked out [6.2]. However, any statements made here on the subject may soon become obsolete as research continues, so the subject will not be discussed further here, except to point out that the ELDR effect is presently a source of uncertainty for some devices.

6.2 Displacement Damage

Nearly pure displacement damage effects (i.e., TID is minimal) can be produced by irradiating a device with neutrons. In one respect, such testing is simpler than TID testing because we do not expect complex time dependent effects from pure displacement damage. While some annealing can occur, this is expected to be minor in most cases, except at elevated temperatures. Therefore it is only necessary to pay attention to fluence, without concern for flux as a function of time. However, there are difficulties regarding the interpretation of data obtained from neutrons. Recall from Section 3.4 that there are some limitations regarding the use of NIEL curves for calculating equivalent fluences.

For this reason, it is best to test with particles representative of those producing the greatest displacement damage in space. For parts protected by typical spacecraft shielding, most of the displacement damage is usually from protons. This is one motivation for using protons for displacement damage tests. Note that protons also produce TID effects in devices susceptible to TID. From the point of view of scientific studies, it may not be desirable to mix the two effects. From the point of view of engineering applications, it is not always clear how to combine data representing separate effects, so if a mixture of effects occurs in the space environment then we should attempt to duplicate this in the laboratory. This is another motivation for using protons. Unfortunately, it is not valid to use NIEL curves to correlate the laboratory environment with the space environment when TID effects are significant. However, if the protons used in the laboratory are sufficiently representative of the proton spectrum in the space environment, we can obtain at least a rough estimate of displacement damage effects in the space environment without the use of NIEL curves. For those devices in which TID effects are either insignificant or can be distinguished from displacement damage effects (e.g., if a device has several components with each susceptible to only one effect), NIEL curves can be used for a more accurate estimate.

Proton-induced displacement damage effects tests are performed at accelerators that provide proton beams of various energies and fluxes. As discussed in Section 3.4, the proton energies used should be large enough so that the energy is not significantly degraded before reaching the device active region. Fortunately, this is consistent with selecting an energy that is most representative of the distributed energy spectrum in the space environment.

Neutron tests can also provide useful information, and this is the obvious choice when the objective is to produce displacement damage with minimal TID in devices susceptible to TID. Such tests often use a nuclear reactor as a neutron source. Reactors with undegraded fission neutron energy spectra are better than reactors with some form of neutron moderation, such as water-moderated reactors, for several reasons: 1) the neutron energies are higher, producing more efficient displacement, 2) there is much less TID from gamma ray and electron contamination, 3) there is less radioactivity induced in the device from thermal neutrons, and 4) bare, fast burst reactors allow easier experimental setup with large open spaces around the reactor core. Because room temperature annealing is less effective for displacement damage than for TID, post-irradiation device characterization measurements can wait until the radioactivity in the device dies out.

References

[6.1] D.M. Fleetwood, S.L. Kosier, R.N. Nowlin, R.D. Schrimpf, R.A. Reber, Jr., M. DeLaus, P.S. Winokur, A. Wei, W.E. Combs, and R.L. Pease, "Physical Mechanisms Contributing to Enhanced Bipolar Gain Degradation at Low Dose Rates," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 1871-1883, Dec.1994.

[6.2] A.H. Johnston, C.I. Lee, and B.G. Rax, "Enhanced Damage in Bipolar Devices at Low Dose Rates: Effects at Very Low Dose Rates," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, pp. 3049-3059, Dec.1996.

7. SEE: TEST METHODS, INTERPRETATION AND LIMITATIONS OF DATA

7.1 Basic Concepts

The design and operation of support equipment that will exercise a device, detect and record SEE when they occur, and not interfere with device susceptibility to SEE, is a sufficiently complex and specialized subject to be beyond the scope of this course. Excluding this topic, the basic idea behind SEE testing is simple. The device under test is placed in front of a particle beam, usually produced by an accelerator, and device response is recorded as a function of beam characteristics (LET for a heavy-ion beam, or energy for a proton beam). The more affordable heavy-ion tests, performed at moderate-energy accelerators, require the device to be in a vacuum and with the upper packaging material removed so that the particles can reach the device active regions. This is not required for the more penetrating particles, such as high-energy protons or for the (rarely used because of cost) heavy-ions produced at the very-high-energy accelerators.

The objective is to obtain a cross section measurement for each type of particle and for each type of SEE (e.g., SEU or SEL) being investigated. A cross section is determined by counting the number of events accumulated from a given beam fluence and dividing by this fluence. In the case of SET, we select a loading condition and measure a histogram of events. The number of events satisfying any given criteria (e.g., that the pulse width exceed some value) is determined from the histogram and used to calculate the cross section for such events. Some types of SEE (e.g., SEGR) are so destructive that the first event can destroy a device even under carefully controlled laboratory conditions. Many devices must be tested to obtain results that are statistically significant. For other types of SEE, it is possible to accumulate many events (e.g., many bit flips when testing a large-memory device for SEU), or reset the device after counting an event (e.g., for SEL tests) so that a statistically significant number of events can be recorded from one device during one exposure to a particle beam.

Device susceptibility is often a function of the direction of the particle trajectory relative to the device normal (sometimes also a function of azimuth angle, but this is usually ignored). The directional cross section is the average number of events per unit fluence with fluence measured perpendicular to the particle beam. A device is sufficiently characterized, for the purpose of predicting SEE rates in a known environment, when the directional cross section is a known function of angle and incident-particle characteristics.

High-energy protons can indirectly create many of the same types of SEE as heavy ions by producing high-energy recoils via inelastic nuclear reactions, which in turn create ionization. The spectra of the heavier particles created by a proton beam is not isotropic [7.1], but the ranges of most of these particles are short enough (a few microns) so that the direction of their travel is less important than the location at which they were created.

The result is that proton cross sections applicable to the indirect process have only a weak dependence on proton direction. It is usually adequate to measure these cross sections as a function of proton energy at normal incidence and assume an isotropic cross section when calculating rates in space. Protons can also produce some types of SEE in exceptionally susceptible devices by direct ionization. Such events often require nearly as much ionization as a proton is capable of creating, which requires sufficiently low energies so that the LET is as large as possible, but this also causes the LET to vary considerably along the trajectory in the device. This is unfortunate because it makes algorithms developed for predicting rates from very penetrating heavy ions (uniform LET) found in space inappropriate. The mixed case in which direct ionization is added to indirect ionization is even more complicated. The direct ionization and mixed cases are relatively new subjects and there is not yet a consensus on how to use laboratory data to calculate rates in space. In contrast, while the indirect process is complicated on a physics level, it is simple to use laboratory data to predict rates in space.

For some types of SEE created by heavy ions, e.g., SEGR, the most important (most damaging) trajectories are close enough to normal incidence so that the directional cross section can be directly measured. For heavy-ion-induced SEU, SEL, and SET, special experimental methods take advantage of the assumed angular dependence of the directional cross section. A separate section (below) is dedicated to this subject.

7.2 SEU, SEL, and SET from Heavy Ions

For heavy-ion-induced SEU, SEL, and SET, cross section data are usually measured and reported in a manner that conforms to the cosine law. The statement of the cosine law is that the directional cross section $\sigma(\theta,L)$ is related to the normal incident cross section σ_N according to

$$\sigma(\theta, L) = |\cos\theta|\,\sigma_N(L/|\cos\theta|). \quad (1)$$

The same equation can also be written as

$$\sigma_N(L) = |\sec\theta|\,\sigma(\theta, L|\cos\theta|). \quad (2)$$

When (2) is valid, the normal incident cross section can be plotted as a function of LET L by plotting the right side of (2) as a function of L. For values of L such that no laboratory ion has an LET equal to that value, θ can be selected so that the argument $L|\cos\theta|$ of σ on the right side of (2) is an available ion LET. This allows arbitrary values to be assigned to L. In particular, it is possible to select values of L that are larger than any available ion LET.

A sufficient condition for the cosine law to apply consists of the simultaneous conditions: (a) the target (sensitive region) is flat, so the projected area seen by an ion beam scales with the cosine of the angle between the beam and device normal, and (b) an

ion hit with LET L and angle θ produces the same device response as a normal-incident hit at the same location but with LET L|secθ|. Condition (b) is often satisfied (approximately) because the liberated charge per unit depth (measured perpendicular to the device) is proportional to |secθ| when the ion range is effectively infinite. The conditions (a) and (b) taken together produce a sufficient condition for the cosine law to apply, but not a necessary condition. For a hypothetical example, consider a device that is both isotropic, i.e., $\sigma(\theta,L)=\sigma_N(L)$, and also satisfies the condition that the normal-incident cross section is proportional to L. These assumptions can be used to show that (1) is satisfied. When the normal-incident cross section is proportional to LET, an isotropic cross section is also a cosine-law cross section. If the normal-incident cross section is approximately proportional to LET over a restricted LET range, then an isotropic cross section can simultaneously satisfy the cosine law for restricted values of LET.

For SEU, SEL and SET in real devices, the cosine law is often an adequate approximation for angles that are not too large, so it is customary to use angles to mimic a different ion LET. The traditional practice is to assume that the cosine law applies up to 60°, although it frequently does not and the resulting data sometimes contain "cosine-law errors". The cosine law is particularly inappropriate for a few cases. For example, the cross section for SEU in some DRAMs is very nearly isotropic, although it can sometimes simultaneously satisfy the cosine law for restricted values of LET as discussed in the previous paragraph. Outside this limited LET range, the cosine law can produce wildly scattered data points for such devices, while the same data plotted in the format of directional cross section versus ion LET clearly shows a smooth and isotropic curve. Another source of cosine-law error is from ion range limitations. The cosine law might apply to a particular device when all ions have adequate range, but may still not apply to ions used in a laboratory test and having inadequate range. In general, the cosine law can fail either because of ion range limitations, or because it does not apply even to long-range ions, or a combination of both. While it is common practice to assume (not always correctly) that the cosine law is valid for angles up to about 60° when reporting test data, it is also good practice to keep records of which data points were measured at which angles.

One limitation of test data is due to the fact that ions available at affordable test facilities have limited ranges. The cosine law is not expected to apply to the largest angles (approaching 90°), but the real problem is that (excluding those cases in which the cosine law fails at small angles, e.g., an isotropic cross section) we cannot determine what law should be used *instead* of the cosine law at the largest angles. The reason is that observed trends in measured data with increasing angle can be an artifact of ion range limitations. The same trends might not be produced by the very penetrating particles found in space. Not only can the cosine law fail at large angles, the ability of measured data to represent cross sections seen by very penetrating particles can also fail at large angles. Excluding some exceptional cases, the directional dependence of device susceptibility, seen by very penetrating particles found in space, is unknown at large angles. This sometimes produces the greatest uncertainties in SEU rate calculations.

The calculation of heavy-ion induced SEU, SET, and SEL rates in space might be included in this section discussing data interpretation, but a separate chapter (the next chapter) is dedicated to this because the subject is fairly lengthy.

References

[7.1] P.M. O'Neill, G.D. Badhwar, and W.X. Culpepper, "Risk Assessment for Heavy Ions of Parts Tested with Protons," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 6, pp. 2311-2314, Dec.1997.

8. CALCULATION OF HEAVY ION INDUCED SEU, SEL, AND SET RATES IN SPACE

8.1 The Model

SEU and SET are very similar with regard to the underlying physics. Some criteria (e.g., pulse width) must be selected to define a SET, but having done that, SEU and SET rates are calculated the same way. Although there are some differences between the physics producing SEL and that producing SEU, the same method is also used to calculate SEL rates. Therefore, SEU serves as the prototype for heavy-ion-induced rate calculations for these types of SEE.

The standard method of calculation starts with the "sensitive volume model". This model assumes that charge collected from an ion track at a device node equals the charge liberated by the track in some well-defined sensitive volume associated with that node. The collected charge is then proportional to ion LET multiplied by the length of track section contained within the volume. In this model, transport equations are replaced by a geometry problem; the problem of calculating the lengths of track sections. The model ignores the fact that an imagined volume initially devoid of carriers (the track misses the volume) can become populated with carriers a short time later due to diffusion. The complete calculation combines the sensitive volume model with the following two additional assumptions [8.1]: *Assumption 1* - A device is assumed to contain a collection of geometrically identical rectangular parallelepiped (RPP) shaped sensitive volumes, and a critical charge is associated with each one. An SEU occurs if the charge liberated by an ion track in any one of these RPPs exceeds the critical charge associated with that RPP. *Assumption 2* - The gradual rise in a device cross section curve with increasing LET is attributed to a statistical distribution of critical charges. In other words, different RPPs can have different sensitivities to SEU by having different critical charges, and high-LET ions are capable of upsetting a larger fraction of the RPPs than low-LET ions.

The above assumptions are used, not because of physical validity, but because there is not yet a community-accepted alternative method for rate calculations. Many (not all) investigators previously thought that the assumptions were physically realistic, and computer algorithms were developed to perform the calculations. Because the computer codes already exist, the same calculations are still used today. Evidence of errors in these assumptions can be found in the following observations:

1) Laser tests show that upsets can be induced by hits at remote locations, indicating that charge collection is not limited to a local sensitive volume.

2) Multiple-bit upsets are sometimes observed from heavy-ions at normal incidence, again indicating that charge collection is not limited to a local sensitive volume.

3) Saturation cross sections can be as large as the entire device area (or larger depending on how multiple-bit upsets are counted), again indicating that charge collection is not limited to a local sensitive volume.

4) Computer simulations show that, excluding devices utilizing physical boundaries for isolation, charge collected at a circuit node from a given ion gradually diminishes as the ion hit location is moved further from the node. This contradicts the model, which predicts that collected charge abruptly becomes zero as the hit location is moved outside the sensitive volume.

5) Experimental charge-collection measurements find distributions of collected charges from the same ion species and energy. This contradicts the model, which assumes that only one non-zero value of collected charge is possible from a given LET at normal incidence. The interpretation of the experimental results is that a distribution of hit locations produces a distribution of collected charges, consistent with simulation predictions in (4) above.

6) Computer simulations show that long tracks can produce a different collected charge than short tracks, even when both have the same average LET in the upper device region where a sensitive volume would be imagined. This contradicts the model, which assumes that the section of ion track contained within the sensitive volume is the only relevant section.

7) It is an experimental observation that long-range ions can produce different cross section measurements than short-range ions, even when both have the same average LET in the upper device region where a sensitive volume would be imagined. This is why long-range ions are preferred for SEU testing. This is consistent with simulation predictions in (6) above. Again, this contradicts the model, which assumes that the section of ion track contained within the sensitive volume is the only relevant section.

8) Computer simulations, laser tests, and microbeam experiments show that the gradual increase in device cross section with increasing LET is largely, if not entirely, due to a gradually increasing cross section for each bit. This is consistent with (4) and (5) above, but contradicts the model, which interprets the gradual increase in the device curve as an increasing number of contributing RPPs.

Although the model is not physically realistic, it can still be used for rate calculations as long as it satisfies the bottom-line requirement: that the predicted directional dependence of device susceptibility, when exposed to the long-range ions found in space, agrees with the actual directional dependence. If the actual directional dependence is known, model parameters (RPP dimensions and critical charges) can be selected accordingly. However, because the model is unrealistic, parameters satisfying this bottom-line requirement need not correspond to real physical quantities. In particular, charge collection depths defined in terms of collected charge from a given ion and determined from either charge collection measurements or computer simulations are often much larger than the RPP thicknesses typically assumed for rate calculations. Also,

the critical charge that should be assigned to an RPP associated with a given node is not the actual critical charge as predicted by a circuit analysis, but is instead chosen to be compatible with the threshold LET and RPP thickness. It is clear that neither the model nor the optimum model parameters are realistic, but taken together they can satisfy the bottom-line requirement.

Because the model is analogous to a curve fit, containing no analysis of charge transport equations, it is not clear how to identify those model parameters that will satisfy the bottom-line requirement; other than by fitting data. However, a problem with fitting data is that observed trends in the directional dependence of device susceptibility (even for test angles up to only 60°) can be an artifact of ion range limitations, which would not be observed in the high-energy space environment. Therefore, selecting an RPP thickness that satisfies the bottom-line requirement is often the greatest uncertainty in rate calculations. Several rate estimates should be calculated, using different assumed thicknesses, to provide a sensitivity test. This will be illustrated in Section 8.4. Depending on the individual case, calculated rates can be insensitive to the assumed thickness (typical of parts with a low threshold LET) or very sensitive (typical of parts with a sufficiently high threshold LET so that most upsets are from ions hitting the part at large angles).

8.2 Effective Flux

The actual number crunching can be simplified by using the concept of an effective flux. This was first introduced for cosine-law devices [8.2], but can be generalized to include a broad category of models, including the RPP model. Note that there was previously some controversy regarding effective flux [8.3], but the objection was directed towards the physical assumptions used to derive an effective flux. In the context used here, effective flux is just a mathematical trick. It is still up to the user to decide what the physical assumptions are going to be (e.g., the cosine law, the RPP model, or something else). The only requirement is that we find a function $K(L',L,\theta,\phi)$, which is consistent with the physical assumptions (whatever they are) and expresses the directional cross section σ in terms of the normal incident cross section σ_N according to

$$\sigma(L,\theta,\phi) = \int_0^\infty K(L',L,\theta,\phi)\frac{\partial\sigma_N(L')}{\partial L'}dL'. \qquad (1)$$

The method is convenient when calculating rates for several devices in the same environment and having the property that the same function K applies to each device. If the cosine law is the physical assumption, then K satisfying (1) is given by

$$K(L',L,\theta,\phi) = |\cos\theta|\,U(L|\sec\theta| - L')$$

where U is the unit step function, i.e., $U(x)=0$ if $x\le 0$ and $U(x)=1$ if $x>0$. If the physical assumption is the RPP model with specified ratios of dimensions, K is the directional cross section for an RPP having these dimension ratios and having threshold LET L',

divided by the area of the upper face (the face seen at normal incidence). This can be seen by considering a single RPP having upper-face area A, normal-incident threshold LET L_{th}, and dimension ratios that produce the angular dependence of directional cross section described by K (note that when threshold LET, instead of critical charge, is specified, the function K describing the angular dependence of the directional cross section is determined by RPP dimension ratios, independent of the size of the RPP). The normal-incident cross section for this RPP is a step function with height A and step location at L_{th}. The directional cross section is given by (1) with the derivative of the normal-incident cross section on the right equal to A times the Dirac delta function with singular point at $L'=L_{th}$. Therefore the directional cross section for this RPP is A times $K(L_{th},L,\theta,\phi)$. Conversely, $K(L_{th},L,\theta,\phi)$ is the directional cross section for this RPP divided by A.

To obtain an effective flux, we start with

$$rate = \int_0^\infty \int_{-1}^{1} \int_0^{2\pi} f(L,\theta,\phi)\, \sigma(L,\theta,\phi)\, d\phi\, d(\cos\theta)\, dL$$

where f is the differential (in LET) directional flux. Note that the flux is zero for sufficiently large L and the cross section is zero for sufficiently small L, so it is okay for the L integration limits to extend from zero to infinity. Substituting (1) into the above equation gives

$$rate = \int_0^\infty F_{eff}(L') \frac{\partial}{\partial L'} \sigma_N(L')\, dL' \qquad (2)$$

where F_{eff} is the integral effective flux defined by

$$F_{eff}(L') \equiv \int_0^\infty \int_{-1}^{1} \int_0^{2\pi} f(L,\theta,\phi)\, K(L',L,\theta,\phi)\, d\phi\, d(\cos\theta)\, dL. \qquad (3)$$

The final equations are (3), which calculates effective flux, and (2), which calculates the rate from the effective flux. For the RPP model with specified dimension ratios, the effective flux $F_{eff}(L')$ is the SEU rate for a single RPP having these dimension ratios and having a normal-incident threshold LET L', divided by the area of the upper face. One way to reach this conclusion is to recall that K is the normalized directional cross section for this RPP, so the right side of (3) is the normalized rate for this RPP. Another way to reach the same conclusion is to consider a single RPP having upper-face area A, normal-incident threshold LET L_{th}, and dimension ratios that produce the angular dependence of directional cross section described by K. The normal-incident cross section for this RPP is a step function with height A and step location at L_{th}. The rate for this RPP is given by (2) with the derivative of the normal-incident cross section on the right equal to A times the Dirac delta function with singular point at $L'=L_{th}$. Therefore the rate for this RPP is A times $F_{eff}(L_{th})$. Conversely, $F_{eff}(L_{th})$ is the rate for this RPP divided by A.

This rate can be calculated using a code such as CREME, so the effective flux can be tabulated against L_{th} (or L if we change symbols). An example of such a tabulation is in Table 8.1, which lists several effective fluxes for several RPP dimension ratios, but all referring to galactic cosmic ray heavy ions in interplanetary space during the solar maximum time period and behind a 100 mil aluminum shield. In all cases, the RPP upper faces are square, but the different effective fluxes refer to different ratios of thickness to lateral dimensions. This ratio is denoted by r in the table. For devices having a nearly isotropic cross section, we use the r=1 effective flux. This applies to RPPs that are cubes. A cube does not have a perfectly isotropic cross section, but this is as isotropic as the RPP model can be. To mimic the cosine law, we use the smallest value of r listed, which is 0.01. The RPPs are now thin slabs, which conform well to the cosine law. For r=0.2 (not listed in the table, but rates can be calculated for neighboring values of r and then interpolated), the directional cross section roughly conforms to the cosine law for angles up to about 60°, but diverges from the cosine law at larger angles. Many real devices appear to behave this way, and r=0.2 often produces agreement between calculated rates and flight observations.

To calculate rates from the effective flux table, we numerically evaluate the integral in (2) by selecting a set of LET values $L_1<L_2<\ldots<L_M$ and use

$$rate = \sum_{i=1}^{M} F_i \left[\sigma_{i+1} - \sigma_i\right] \tag{4}$$

where F_i is the effective flux evaluated at LET L_i, and σ_i is the normal incident cross section evaluated at LET L_i. L_1 must be at or below the device threshold LET (below is okay if the cross section is set equal to zero for any LET less than the threshold), and L_M must be large enough so that upsets from particles having a larger LET can be ignored. Otherwise, the choice of the Ls is fairly arbitrary. Selecting them to be a subset of the LETs listed in the effective flux table eliminates the need for interpolating the flux table. The right side of (4) contains F_i, instead of a numerical average $(F_i+F_{i+1})/2$, so that the user has the option of being conservative and lazy. Subject to the above requirements regarding L_1 and L_M, a coarse numerical integration (small M) will produce a conservative estimate, but the work is easy because a few cross section values can be read directly from the cross section curve. A fine numerical integration, using many closely spaced LET values, is more accurate, but also a lot of work unless the cross section curve is fit with an analytic expression so that we don't have to read a lot of points from the curve. This brings us to the next section.

8.3 Curve Fitting

A common practice is to fit cross section data with a Weibull function. This practice originated from an old belief by many (not all) investigators that the gradual rise in a device cross section curve with increasing LET is due to a statistical distribution of critical charges [8.1]. The Weibull function, which is derived from statistics, was thought to have a physical basis. It was later realized that the gradual rise is largely, if not

Table 8.1. Effective flux for solar maximum GCR in interplanetary space for several ratios (denoted by r) of RPP thickness to lateral dimensions.

Effective flux in units of 1/cm^2-day
LET in units of MeV-cm^2/mg

LET	r = 1.00	0.40	0.25	0.16	0.10	0.01
1.0E-01	6.8E+02	5.1E+02	4.6E+02	4.3E+02	4.1E+02	3.7E+02
1.3E-01	5.7E+02	4.7E+02	4.3E+02	4.0E+02	3.8E+02	3.5E+02
1.6E-01	4.5E+02	4.1E+02	3.8E+02	3.6E+02	3.4E+02	3.2E+02
2.0E-01	3.5E+02	3.4E+02	3.2E+02	3.1E+02	3.0E+02	2.8E+02
2.5E-01	2.7E+02	2.8E+02	2.7E+02	2.6E+02	2.5E+02	2.4E+02
3.2E-01	2.0E+02	2.2E+02	2.2E+02	2.1E+02	2.0E+02	1.9E+02
4.0E-01	1.5E+02	1.7E+02	1.7E+02	1.7E+02	1.6E+02	1.6E+02
5.0E-01	1.1E+02	1.3E+02	1.3E+02	1.3E+02	1.3E+02	1.2E+02
6.3E-01	8.4E+01	9.7E+01	1.0E+02	1.0E+02	1.0E+02	9.7E+01
7.9E-01	6.2E+01	7.3E+01	7.7E+01	7.9E+01	7.9E+01	7.6E+01
1.0E+00	4.4E+01	5.5E+01	5.9E+01	6.1E+01	6.2E+01	6.1E+01
1.3E+00	2.6E+01	3.9E+01	4.2E+01	4.5E+01	4.6E+01	4.6E+01
1.6E+00	1.2E+01	2.6E+01	2.8E+01	3.0E+01	3.1E+01	3.2E+01
2.0E+00	6.2E+00	1.6E+01	1.8E+01	2.0E+01	2.1E+01	2.2E+01
2.5E+00	3.4E+00	9.7E+00	1.2E+01	1.3E+01	1.4E+01	1.5E+01
3.2E+00	2.0E+00	5.3E+00	7.2E+00	8.1E+00	8.8E+00	9.8E+00
4.0E+00	1.2E+00	2.7E+00	4.3E+00	5.1E+00	5.6E+00	6.5E+00
5.0E+00	6.6E-01	1.5E+00	2.4E+00	3.2E+00	3.5E+00	4.3E+00
6.3E+00	3.8E-01	8.3E-01	1.2E+00	1.9E+00	2.2E+00	2.8E+00
7.9E+00	2.2E-01	4.8E-01	6.7E-01	1.0E+00	1.4E+00	1.8E+00
1.0E+01	1.2E-01	2.8E-01	3.8E-01	5.4E-01	8.2E-01	1.2E+00
1.3E+01	6.4E-02	1.6E-01	2.2E-01	3.0E-01	4.5E-01	7.5E-01
1.6E+01	3.0E-02	8.7E-02	1.3E-01	1.7E-01	2.3E-01	4.8E-01
2.0E+01	1.1E-02	4.5E-02	6.7E-02	9.4E-02	1.3E-01	3.0E-01
2.5E+01	3.8E-03	2.3E-02	3.6E-02	5.1E-02	6.9E-02	1.9E-01
3.2E+01	9.5E-04	1.1E-02	1.8E-02	2.7E-02	3.8E-02	1.2E-01
4.0E+01	8.7E-05	4.3E-03	8.9E-03	1.4E-02	2.0E-02	7.0E-02
5.0E+01	2.9E-06	1.4E-03	4.1E-03	7.1E-03	1.1E-02	4.2E-02
6.3E+01	8.7E-07	4.6E-04	1.7E-03	3.5E-03	5.6E-03	2.5E-02
7.9E+01	2.8E-07	9.3E-05	5.4E-04	1.6E-03	2.8E-03	1.5E-02
1.0E+02	6.8E-08	4.3E-06	1.8E-04	6.5E-04	1.4E-03	8.8E-03
1.3E+02	6.5E-09	3.5E-07	3.6E-05	2.1E-04	6.4E-04	4.8E-03
1.6E+02	1.1E-10	1.1E-07	1.6E-06	7.0E-05	2.6E-04	2.4E-03
2.0E+02	2.8E-13	3.3E-08	1.4E-07	1.4E-05	8.5E-05	1.3E-03
2.5E+02	0.0E+00	6.7E-09	4.4E-08	6.1E-07	2.8E-05	6.9E-04
3.2E+02	0.0E+00	3.2E-10	1.3E-08	5.3E-08	5.6E-06	3.8E-04
4.0E+02	0.0E+00	3.8E-12	2.6E-09	1.7E-08	2.4E-07	2.0E-04
5.0E+02	0.0E+00	0.0E+00	1.2E-10	5.0E-09	2.1E-08	1.1E-04
6.3E+02	0.0E+00	0.0E+00	1.4E-12	1.0E-09	6.9E-09	5.6E-05
7.9E+02	0.0E+00	0.0E+00	0.0E+00	4.6E-11	2.0E-09	2.8E-05
1.0E+03	0.0E+00	0.0E+00	0.0E+00	5.2E-13	4.0E-10	1.4E-05

entirely, due to charge-collection physics having nothing to do with statistics, but the Weibull function continues to be popular because it can fit many data sets by virtue of its many adjustable parameters. It should be noted that the Weibull function is just a fit, having no demonstrated physical basis relevant to charge collection, so use of this function is optional. Other fitting functions can also be used. One example given by

$$\sigma_N(L) = A\, e^{-\frac{B}{L}} \tag{5}$$

has only two adjustable parameters A and B, and does have a physical basis [8.4]. However, a physical basis in lieu of a multitude of adjustable parameters can actually be a disadvantage, because this implies that there are also limitations. One limitation is that a sum of such functions having different values of B is not another function of the same type. Therefore, the fit is limited to those cases in which the device cross section is dominated by contributions from similar components (e.g., an SRAM in an all 1's pattern or all 0's pattern, but not a mixture of 1's and 0's unless the two contributions to the cross section refer to the same B, or one contribution is negligible compared to the other). Another characteristic of (5), which is more of a virtue than a limitation, is that it will not fit cosine-law errors. Cosine-law errors are often visible as scatter in the data when different ions produce overlapping data sets. When there are no overlapping data, cosine-law errors can be systematic and difficult to recognize. The Weibull function will fit systematic errors (not a virtue), but (5) will not, so it is often necessary to distinguish data taken at normal incidence from data taken at angles and give preference to the normal-incident data when selecting A and B. It is good practice to distinguish data taken at normal incidence from the other points anyway, because this frequently identifies the source of scatter. Much of the scatter in cross section data not due to poor counting statistics is often due to cosine-law errors. Subject to the above qualifications, the fit given by (5) is often quite good and is very convenient. Cross section is plotted against reciprocal LET on semi-logarithmic paper (cross section on the log scale and reciprocal LET on the linear scale) and the points are fit with a straight line. There is no well-defined threshold LET using this method, but it can be selected to be any LET at which the cross section is negligibly small.

As long as data points are sufficiently abundant and the only purpose of a fitting curve is to interpolate between points, the choice of a fitting function is a matter of preference. A fitting function having a physical basis is only needed for those cases in which data points are scarce enough so that interpolation/extrapolation is uncertain. One convenient approach is to try the simplest fit first. Plot the cross section against reciprocal LET on semi-logarithmic paper and look for a straight line while giving preference to normal incident data having adequate counting statistics. If this plot does not produce a straight line, it is time to look for other fits, such as the Weibull function. Sometimes a straight line is evident in all points except the lowest LET (largest reciprocal LET) points. If the cross section corresponding to these points is several orders of magnitude smaller than the largest measured cross section, these points will not strongly influence the calculated SEU rate, so the straight line conforming to the other points can be used. Two points lying on the straight line can be used to determine the parameters A and B in (5), which is used for LET greater than the measured threshold LET. Below this LET, we set σ_N equal to zero.

8.4 An Example Calculation

Latchup data measured by the Jet Propulsion Laboratory for an AD12062 analog to digital converter are used to illustrate a heavy-ion rate calculation. Plotting cross section against reciprocal LET on semi-logarithmic paper produces a nearly straight line, so the simple fit given by (5) applies (if it did not, we would have to look for a different fit). The equation is then used to plot cross section versus LET. Data (points) and the fit (curve) are shown in Fig. 8.1. A perfect fit through the two points measured at normal incidence also produces (in this case) a good fit to all points, except one. The one "misfit" point at LET=35 was measured at a 45° angle (the largest angle used for this test), so this point may be influenced by cosine-law error. All other points, including those measured at angles, conform to the curve given by (5) with $A=1.19\times10^{-3}cm^2$, B=65.3902MeV-cm^2/mg. The fit (5) has no threshold LET, but the cross section is small enough at LET=12 so that this can be taken to be the threshold. We therefore use $\sigma_N=0$ for LET less than 12, and use (5) with the stated values for A and B for larger LET.

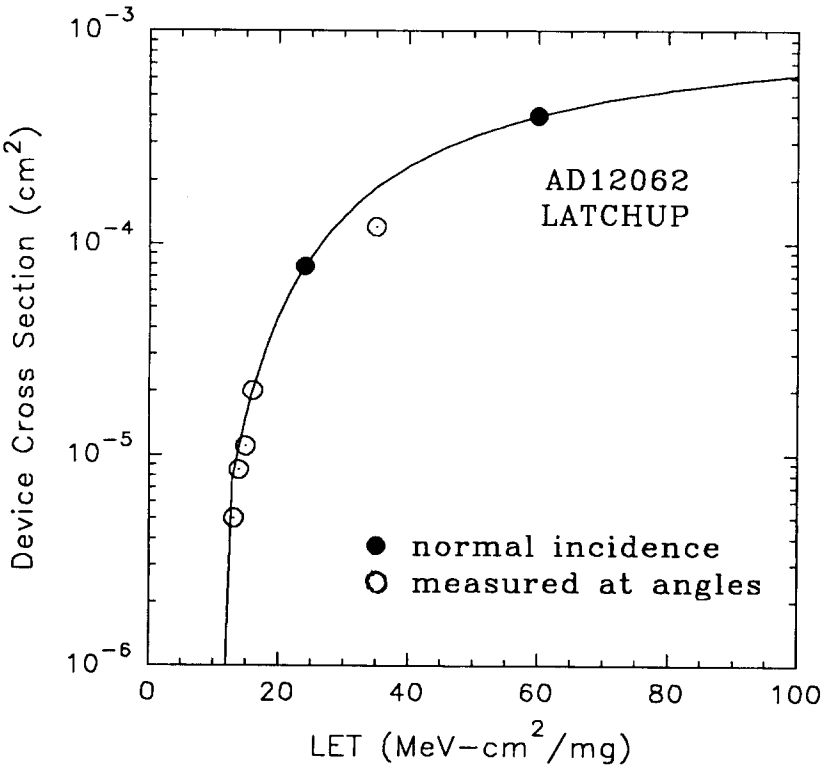

Fig. 8.1: Cross section data and a fitting curve for illustration of the rate calculation method.

With an analytic fit available for the cross section, it is easy to perform a moderately fine numerical integration by using the same LET values in (4) as those used in the effective flux tabulation. A finer integration requires interpolation of the flux table, but a dedicated computer code, which accepts effective flux tables as input, is trivial to write

and makes these calculations easy to do. A tabulation of cross section using the stated LET values is shown in Table 8.2. As is usually the case, it is not clear which value of r (i.e., which effective flux in Table 8.1) best describes device susceptibility, seen by the very penetrating ions found in space, at large angles. We therefore consider credible values (r=0.16 and r=0.25 are credible, based on experience), but also include an extreme value (r=0.01) to furnish a sensitivity test. For each selected value of r, the column in Table 8.1 corresponding to that r is combined with Table 8.2 via (4). To be more explicit, consider the r=0.16 column. The numbers are:

$$\text{rate} = 0 + \ldots + 0 + 5.4\times10^{-1} \times [7.8\times10^{-6} - 0] + 3.0\times10^{-1} \times [2.0\times10^{-5} - 7.8\times10^{-6}] + \ldots + 4.6\times10^{-11} \times [1.1\times10^{-3} - 1.1\times10^{-3}].$$

The results are:

$$\text{rate} = 1.7\times10^{-5}/\text{day} \ @ \ r = 0.25$$
$$\text{rate} = 2.4\times10^{-5}/\text{day} \ @ \ r = 0.16$$
$$\text{rate} = 2.0\times10^{-5}/\text{day} \ @ \ r = 0.20$$
$$\text{rate} = 7.9\times10^{-5}/\text{day} \ @ \ r = 0.01.$$

The r=0.2 rate was estimated by interpolating between the rates for the neighboring values of r. The r=0.01 rate is almost certainly too large, so this furnishes an upper bound rate estimate. The other, more credible, estimates are within a factor of roughly 4 of the upper bound, so the sensitivity is not bad in this case. Rate estimates that are within a factor of several of each other are usually considered to be in good agreement. The r=0.2 rate is a reasonable estimate. For other examples in which the sensitivity is much greater, it is necessary to acknowledge that the rate estimate contains a great deal of uncertainty.

Table 8.2. A tabulation of the cross section data in Fig. 8.1. The cross section is taken to be zero for LET less than 12.

Cross section in units of cm^2
LET in units of MeV-cm^2/mg

LET	Cross Section	LET	Cross Section	LET	Cross Section
1.0E-1	0	2.5E+0	0	6.3E+1	4.2E-4
1.3E-1	0	3.2E+0	0	7.9E+1	5.2E-4
1.6E-1	0	4.0E+0	0	1.0E+2	6.2E-4
2.0E-1	0	5.0E+0	0	1.3E+2	7.2E-4
2.5E-1	0	6.3E+0	0	1.6E+2	7.9E-4
3.2E-1	0	7.9E+0	0	2.0E+2	8.6E-4
4.0E-1	0	1.0E+1	0	2.5E+2	9.2E-4
5.0E-1	0	1.3E+1	7.8E-6	3.2E+2	9.7E-4
6.3E-1	0	1.6E+1	2.0E-5	4.0E+2	1.0E-3
7.9E-1	0	2.0E+1	4.5E-5	5.0E+2	1.0E-3
1.0E+0	0	2.5E+1	8.7E-5	6.3E+2	1.1E-3
1.3E+0	0	3.2E+1	1.5E-4	7.9E+2	1.1E-3
1.6E+0	0	4.0E+1	2.3E-4	1.0E+3	1.1E-3
2.0E+0	0	5.0E+1	3.2E-4		

8.5 Miscellaneous Rate Estimates for Hypothetical Devices

One way to compare different heavy-ion environments is by comparing effective fluxes, or, equivalently, by comparing calculated SEU rates for hypothetical devices having a step-function for a normal-incident cross section curve (in RPP terminology, the device is described by a single RPP). The normal-incident cross section for such a device is zero for ion LET less than some device threshold LET L_{th}, and the cross section is a nonzero constant (which will be called the saturation cross section) for larger ion LET. An exact comparison of rates in different environments depends on which effective flux (i.e., which r value) is used. The r=0.2 rates are believed to adequately approximate a fairly large number of parts (although a step-function cross section curve does not, and leads to very conservative rate estimates for devices having cross sections that gradually rise from the assumed threshold to the assumed saturation value), so this value of r may be a good choice for comparing environments. An advantage of selecting a definite r value is that rates for RPPs having different sizes but the same r value and the same threshold LET (not the same critical charge) scale with RPP area. Therefore we can arbitrarily select the saturation cross section to be 1 cm^2. The rate for a 10^{-3} cm^2 (for example) cross section is calculated by multiplying the rate for the 1 cm^2 cross section by 10^{-3}.

Rates for some hypothetical devices in several environments are listed in Table 8.3. The following notes apply:

1. The hypothetical devices are characterized by r=0.2 and a step function for the cross section curve, with a saturation cross section of 1 cm^2. Rates for such hypothetical devices having different saturation cross sections can be scaled from these rates.

2. DCF rates are peak rates in interplanetary space at 1 AU. The accumulated number of upsets over the duration of a DCF is the peak rate multiplied by 0.83 days.

3. GCR rates are at 1 AU and behind a 100 mil aluminum shield, but are not sensitive to either mass shielding or AU distance.

4. DCF rates are sensitive to mass shielding, so three aluminum sphere thicknesses are shown.

5. A worst-case DCF rate at locations other than at 1 AU is determined by taking the radial dependence of the rate to be $1/R^3$ inside of 1 AU, and $1/R^2$ outside of 1 AU. For example, the rate at 0.8 AU from the sun is obtained by multiplying the 1 AU rate by $(1/0.8)^3=1.95$.

6. Rates include only interplanetary space heavy ions. GCR heavy ion rates should dominate GCR proton rates, but solar event proton rates (not included here) could dominate the DCF heavy ion rates for proton susceptible devices, particularly behind the thicker shields.

Table 8.3. SEU Rates in Interplanetary Space from Heavy Ions (only) for Hypothetical Devices

The hypothetical devices have a saturation cross section of 1 cm^2.
Rates can be scaled for other saturation cross sections.

Device Threshold LET (MeV-cm^2/mg)	GCR Solar Max. Rate (1/day)	GCR Solar Min. Rate (1/day)	DCF @ 250 mils Rate (1/day)	DCF @ 100 mils Rate (1/day)	DCF @ 60 mils Rate (1/day)
1	6.0E+1	1.8E+2	1.6E+5	6.2E+5	4.3E+6
3	8.4E+0	3.2E+1	8.6E+3	4.0E+4	3.3E+5
6	1.7E+0	7.6E+0	8.4E+2	5.4E+3	6.6E+4
10	4.5E-1	2.3E+0	1.1E+2	1.2E+3	1.9E+4
15	1.7E-1	8.5E-1	2.6E+1	3.8E+2	7.2E+3
20	7.9E-2	4.1E-1	1.1E+1	1.8E+2	3.5E+3
30	2.6E-2	1.4E-1	2.5E+0	4.8E+1	1.0E+3
50	5.4E-3	3.1E-2	2.6E-1	7.3E+0	1.7E+2
70	1.6E-3	1.0E-2	4.7E-2	1.7E+0	4.5E+1

References

[8.1] E.L Petersen, J.C. Pickel, J.H. Adams, Jr., and E.C. Smith, "Rate Prediction for Single Event Effects-- a Critique," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 6, pp. 1577-1599, Dec.1992.

[8.2] D. Binder, "Analytic SEU Rate Calculation Compared to Space Data," *IEEE Trans. Nucl. Sci.*, vol. 35, no. 6, pp. 1570-1572, Dec.1988.

[8.3] E.L. Petersen, "SEE Rate Calculations Using the Effective Flux Approach and a Generalized Figure of Merit Approximation," *IEEE Trans. Nucl. Sci.*, vol. 42, no. 6, pp. 1995-2003, Dec.1995.

[8.4] L.D. Edmonds, "SEU Cross Sections Derived from a Diffusion Analysis," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, pp. 3207-3217, Dec.1996.