Journal of The Electrochemical Society, **149** (5) S35-S58 (2002)
0013-4651/2002/149(5)/S35/24/$7.00 © The Electrochemical Society, Inc.

An ECS Centennial Series Article

An Electronics Division Retrospective (1952-2002) and Future Opportunities in the Twenty-First Century

Howard R. Huff*

International SEMATECH, Austin, Texas 78741, USA

The emergence of crystalline silicon and silicon-based materials such as silicon-germanium as the premier materials and the personnel driving the integrated circuit (IC) microelectronics revolution will be reviewed. The major threshold events from the 1940s through the mid-1960s, presaging the onset of the large scale integration microelectronics era, will be highlighted. The major silicon material challenges such as dislocation-free single-crystal growth, plastic deformation, the point-defect dilemma, gettering, oxygen in silicon, carrier lifetime, and controlled point-defects in the silicon crystal during the evolution of silicon microelectronics from large scale integration through the very large scale integration era in the 1970s and the 1980s into the ultralarge scale integration era of the 1990s will then be reviewed. Opportunities in epitaxy, wafer cleaning, silicon-on-insulator, silicon-germanium, IC scaling and potential changes in device configuration and IC architecture in the evolution towards the 64 Gbit DRAM and 9 G transistor high-performance MPU logic era in 2016 (per the 2001 edition of the International Technology Roadmap for Semiconductors) will be discussed in the context of silicon-based microelectronics. The complementary role of compound semiconductors, nanoelectronics and the continuing initiative to obtain an optoelectronic system compatible with silicon will also be discussed. Finally, nonsilicon materials and device configurations will briefly be noted.
© 2002 The Electrochemical Society. [DOI: 10.1149/1.1471893] All rights reserved.

Available electronically April 11, 2002.

The computer and communications age has catapulted electronics to its current status as the dominant global industry. Indeed, we are in the midst of a revolution brought about by the availability of inexpensive information acquisition, manipulation, and distribution systems. Exploitation of electron and hole conduction in silicon has resulted in electronic memory and logic circuits such as the dynamic random access memory (DRAM) and microprocessor. Control of the interaction of photons with compound semiconductors has resulted in optical devices such as the laser and optical fiber networks. This revolution has been and will continue to be dependent on our ability to control electronic and photonic material processing techniques for the manufacture of useful devices, circuits and systems. The preparation and detailed processing sequence of a material from crystal growth through device and circuit fabrication determines the microstructure and, therefore, the electronic properties of the material and resulting device and circuit performance, yield, and reliability. Electronic materials include semiconductors, dielectrics, magnetics, piezoelectrics, optoelectronic materials, and optical fibers which may be utilized in crystalline, polycrystalline, or amorphous form. Materials are indeed the *sine qua non* of electronic and optical devices and circuits.

An extensive list of relevant citations through 1997 is noted in Ref. 1, a small number of which will be referred to in the course of this review. A review from 2001 contains an additional set of updated references.[2] Finally, a complementary set of relevant references is also cited.[3,4] Of course, only a small subset of the Electronics Division (ED) thrust could possibly be included in this brief review. Needless to say, retrospective assessments are very subjective and projections are the result of less than prescient skill, so this article should, perhaps, be pursued cautiously. ED initiatives very relevant to integrated circuit (IC) fabrication, yet not significantly covered, moreover, include chemical vapor deposition, the Si/SiO_2 interface, ion implantation, low-k and high-k dielectric materials, metallization, patterning science and technology, plasma etching, analytical and diagnostic techniques, process modeling and control, and packaging, as well as diamond and diamond-like films, high-temperature superconductors, sensors, and quantum confinement and nanostructures.

* Electrochemical Society Fellow.

Status-1952

Ralph Hunter, in his 1952 Electrochemical Society (ECS) Presidential Address,[5] noted that ECS Past President R. M. Burns, upon delivering the Perkin Medal address, predicted "a revolution in the electronics industry as a result of the developments of germanium", and "the tremendous reduction in size and power requirements for the transistor when it replaces the vacuum tube." [6] Interestingly, 46 years later, Ian Ross recalled related discussions during the similar time frame at Bell Telephone Laboratories (BTL) as to the dichotimy of combining the smaller size (and increased operating frequency of the transistor) while seeking higher power at the higher frequency.[7] Ross quoted Bob Wallace of BTL as:

"Gentlemen, you've got it all wrong! The advantage of the transistor is that it is inherently a small size and low power device. This means that you can pack a large number of them in a small space without excessive heat generation and achieve low propagation delays. And that's what we need for logic applications. *The significance of the transistor is not that it can replace the tube but that it can do things that the vacuum tube could never do*!" [emphasis added].

Ross continued "and this was a revelation to us all. We realized that in chasing the vacuum tube, we had the wrong emphasis." [7] Rather, the opportunity was created by focusing on the transistor *in its own right*. The application of an invention is a powerful stimulus for innovation and development should not be restricted to the originally intended application as it may not be the most important. In a similar vein, Robert Lucky has recently noted "moreover there is no *a priori* way to determine what will tip a market. It's a fundamental instance of chaos in group dynamics. And that makes it fundamentally difficult to predict future societal behaviors in the adoption of technologies." [8]

Indeed, neither John Bardeen or Walter Brattain (Fig. 1), who co-discovered transistor action on December 16, 1947 in polcrystalline germanium[4,9,10] nor William Shockley (Fig. 1), inventor of the minority-carrier injection concept on January 23, 1948 and the bipolar junction transistor,[4,11-14] anticipated the ensuing impact of the transistor. Nor did the co-inventor's of the integrated circuit (IC) in 1958, Jack Kilby (Fig. 2)[15,16] and Bob Noyce,[17] initially anticipate the ensuing microelectronics revolution. An excellent description of the differences between Kilby's and Noyce's approaches to the IC has been described.[18] On the other hand, Patrick Haggerty of Texas Instruments clearly foretold the impact of the learning curve in de-

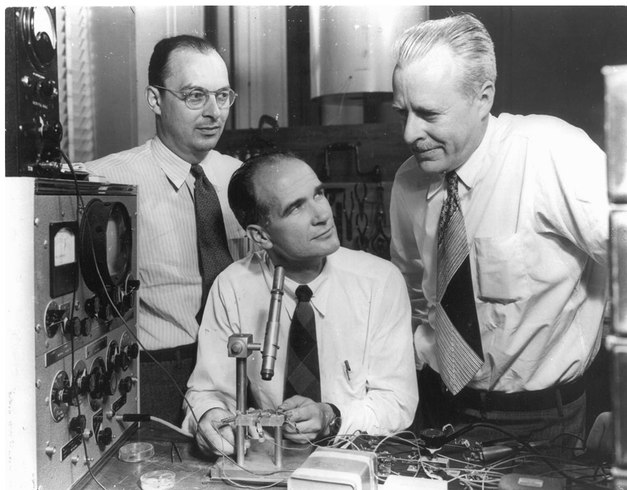

Figure 1. (*From left to right*) John Bardeen, William Shockley, and Walter Brattain. (*Property of AT&T Archives. Reprinted with permission of AT&T.*)

creasing the cost of manufacturing ICs and its increased pervasiveness in the generation of new market opportunities and the growth of the IC industry.[15,19] Gordon Moore of Intel (Fig. 2), initially projected the growth of the number of components per chip quadrupling every two years in 1965 while at Fairchild Semiconductor Corp.,[20] subsequently modifying his prognosis in 1995 to about every three-to-four years.[21] This observation, enshrined as Moore's law, became the productivity curve criterion by which the IC industry grew at approximately 25-30% compound annual growth rate (CAGR) (*i.e.*, increased number of transistors, decreased cost per transistor.) Although Moore's law will eventually saturate, it has been the cornerstone by which the IC industry gauges its growth, as, for example, in the International Technology Roadmap for Semiconductors (ITRS).[22-24]

As a matter of historical interest, Walter Schottky comprehended the existence of a depletion region near the surface of a semiconductor in 1938, contemporaneously with Neville Mott, Boris Davidov, and shortly, thereafter, Hans Bethe.[4] These personnel, however, focused on explaining rectification in terms of the contemporaneous space-charge theory of rectification, also referred to as the "one current" theory of rectifiers, making further scientific discussion of minority carriers at the time moot (although Davidov apparently had developed the rectification equation in 1938 ten years before Shockley[4,25]).

J. W. Marden, chairman of the ECS Electronics Division in 1952, noted in his Editorial of The Electronics Division of the ECS that "germanium has, in like manner, found its way into the medical field and industrial applications in the production of transistors." [26]

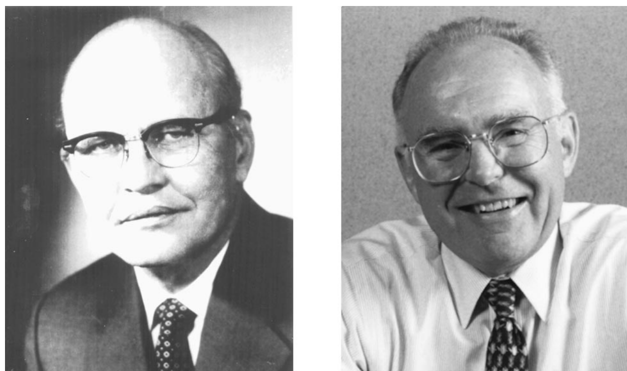

Figure 2. (*left*) Jack Kilby and (*right*) Gordon Moore.

Figure 3. (*left*) Robert Hall and (*right*) Nick Holonyak.

Additionally, Marden noted in his History of The Electronics Division that "today, the field of semiconductors has been studied very little. It is possible that by alternating currents, or otherwise, electrons may be moved from one energy level to another within atoms, in such manner that during their return to a stable condition light is emitted. Who can say that a far more efficient light source will not ultimately be produced. The function of the Electronics Division of The Electrochemical Society has been and will be to foster such new developments." [27]

Rudolf Nagy, in his assessment of the Growth of the Luminescence Section of the Electronics Division of the ECS, noted that "there is a great deal to be learned regarding the mechanisms of luminescence . . . . It is predicted that many of these new discoveries will be reported before The Electronics Division and thus The Electrochemical Society will play a major role in this field." [28] What is truly significant was that Bardeen and Brattain's observation of transistor action on December 16, 1947, the first minority-carrier device to increase the minority-carrier concentration in a semiconductor by a current, whereas previous methods involved optical or thermal processes,[4] facilitated both the subsequent transistor revolution and light emission from a current-driven semiconductor (although the importance of the material's band-structure was yet to be comprehended.) The development of the maser, gaseous laser, and, especially, the solid-state laser, revolutionized the communications industry, The solid-state laser was concurrently reported by four groups in 1962; Robert Hall (Fig. 3) (GaAs), Marshall Nathan (GaAs), Nick Holonyak (Fig. 3) (GaAsP), and Robert Rediker (GaAs).[29] In particular, the recent generation of III-V nitride-based Light Emitting Diodes (LEDs) and lasers, pioneered by Shuji Nakamura,[30] has facilitated the generation of white lighting as well as tunable wavelength lighting.

Scientific concepts developed during the quantum mechanical revolution in the early decades of the twentieth century (and during the first 50 years of the Society) subsequentially led to successful commercial applications (and an extensive number of scientific publications in the *Journal of The Electrochemical Society*) during the second 50 years of the ECS. These early achievements included Albert Einstein's mathematical derivation in 1915 of his A and B coefficients describing spontaneous emission (relevant to laser operation in 1916-1917), Clinton Davisson and Lester Germer's description of matter waves via the diffraction pattern observed by electron scattering from nickel in 1927 (relevant to electron and hole transport in semiconductors as well as electron transport in metals), and the prediction of holes in semiconductors in 1931 by Allen H. Wilson, although it appears that Rudolf Peierls had anticipated their existence in 1928.

In that regard, Lawrence Addicks in his 1952 address on The Founders noted that "now, while The Electrochemical Society by definition merely stands between the chemists and the electrical engineers, in practice we seem to have become the homeland for the refugee scientists who find themselves in the misty no-man's land

between the older disciplines." [31] Indeed, Marden noted "Thus the Society opened its doors for new subjects and made its scientific publications available to those working in the pioneer fields of the electronic arts. This forward looking policy has and can be expected to result in the further growth of The Society, not only in membership, but also in scientific leadership." [27] Finally, it was quite prescient of Thomas Callinan in his Editorial on The Electronics Insulation Division to note that "since the relative conductivity of each of these items [materials considered relevant to the Electric Insulation Division] is a function of the purity and processing involved in their manufacture, preparative methods are considered parts of Materials Dielectrics." [32] Even 50 years ago, the close symbiosis between the Electronics Division and the Dielectric Division (currently the Dielectric Science & Technology Division) was recognized.

The key unifying concept that evolved from Addick's and Callinan's observations has been especially epitomized by The Electronics Division during the last 50 years. This principle may be described in terms of the process-structure-property methodology as the unifying principle in describing solid-state (condensed) material characteristics. The process-structure-property approach, wherein the fabrication process determines the structure which, therefore, uniquely determines the material properties (electrical, mechanical, optical, chemical, thermal, etc.) is critical for the effective fabrication of ICs, optical components (lasers) as well as sub-three dimensional material configurations (*i.e.*, two-dimensional planar interfaces, one-dimensional linear arrays and point structures, either singularly or incorporated within higher-dimensional host matrices) to ensure the requisite performance, yield, and reliability.

Threshold Events

Background.—Transistor action was experimentally observed by Bardeen and Brattain in n-type polycrystalline germanium on December 16, 1947 (and subsequently polycrystalline silicon) as a result of the judicious placement of gold-plated probe tips in nearby single crystal grains of the polycrystalline material, *i.e.*, the point-contact semiconductor amplifier, often referred to as the point-contact transistor.[4,9,10] The device configuration exploited the inversion layer (in their opinion) as the channel through which the majority of the emitted (minority) carriers were transported from the emitter to the collector. Indeed, it was Bardeen's surface state theory[33] which clarified the origin of the surface space charge region (SSCR) (*i.e.*, Schottky's depletion region) and its associated inversion layer. The grown junction single crystal germanium (bipolar) transistor followed in 1951 under Gordon Teal's leadership,[3,34] based on Shockley's minority-carrier injection concept.[4,11-14] Indeed, Shockley's injection patent was filed later than both transistor patents (*i.e.*, the point-contact semiconductor transistor patent and the bipolar junction transistor patent) and yet was awarded on April 4, 1950, earlier than the transistor patents, attesting to the importance (and recognition) of the injection concept.[4,35]

The point-contact transistor was manufactured for 10 years starting in 1951 by the Western Electric Division of AT&T.[7] The *a priori* tuning of the point-contact transistor parameters, however, was not simple inasmuch as the device was dependent on the detailed surface structure and, therefore, very sensitive to humidity and temperature as well as exhibiting high noise levels. Accordingly, the devices differed significantly in their characteristics and electrical instabilities leading to "burnout" were not uncommon.[36] With the implementation of crystalline semiconductor materials in the early 1950s,[1,3,4] however, p-n junction transistors began replacing the point-contact transistor, silicon began replacing germanium[36] and the transfer of transistor technology from the lab to the fab accelerated.

Single-crystal growth.—Polycrystalline germanium and silicon were the basic materials used at BTL and elsewhere for transistor research and development inasmuch as the utilization of single crystals of germanium and silicon for the transistor was a very controversial matter at that time. Gordon Teal (Fig. 4), however, was a proponent of the criticality of single-crystal materials for the electronics era, recognizing that Shockley's bipolar junction transistor characteristics in single-crystal germanium[37,38] and silicon[39] would be substantially better and more reproducible than those of polycrystalline material.[1,3,4,40] Teal believed the fundamental property of a crystalline semiconductor which would result in its technological importance was the easily controllable and spatially variable concentration, type and mobility of free carriers, which was indeed found to be the case.[1,3,4,40]

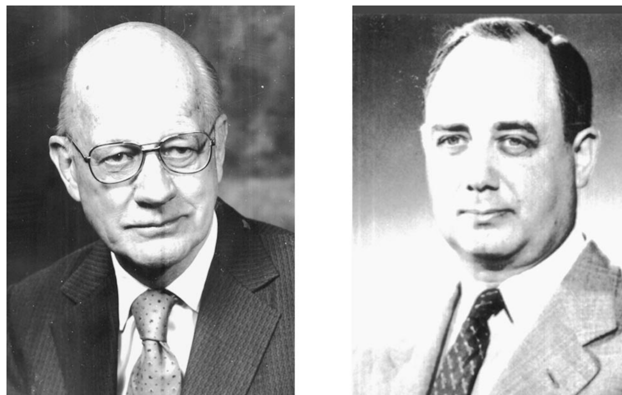

Figure 4. (*left*) Frederick Seitz and (*right*) Gordon Teal.

The successful initial results obtained in the joint research program between Teal and Little, begun in September, 1948, resulted in several germanium rods with some large single crystals[37] as well as grown germanium p-n junctions[38] by pulling from a melt, often referred to, somewhat inappropriately, as the Czochralski (CZ) technique.[1] The use of a pulling method for germanium, a single-crystal seed and the employment of new process techniques, such as precise temperature control of the melt-solid interface, were vital factors in the attainment of single-crystals in which the essential semiconducting properties became highly controlled. The preparation and characterization of single crystal material, furthermore, facilitated experimental verification of a number of quantum theoretical concepts developed for electrons and holes in crystalline semiconductors such as effective mass, drift and conductivity mobility, carrier lifetime and tunneling[3] and clarification of number of phenomena in p-n junctions[41] and, indeed, exhibited significantly improved characteristics compared to polycrystalline samples. The conversion of germanium and silicon ores to metallurgical grade material and their subsequent purification during the 1940s has been reviewed by Frederick Seitz (Fig. 4) and colleagues.[42,43] By the early 1950s, all investigators of the semiconducting properties of germanium and silicon preferred to use pulled single crystals.[1,3]

The description of dopant distribution during single-crystal growth by normal freezing was described by William Pfann (Fig. 13), ECS Solid-State Science and Technology Awardee in 1973, via the related zone-refining techniques and classic textbook, *Zone Melting*.[44-46] The relationship between the effective distribution coefficient, k_e, and the equilibrium, k_o, distribution coefficient during CZ growth was identified[47] and subsequently related to the microscopic crystal growth rate as described by Harry Gatos, ECS President in 1967-1968 and Solid-State Science and Technology Medalist in 1975, and colleagues.[48,49] An extensive summary of the equilibrium distribution coefficients and solubilities for a variety of elements in germanium and silicon were summarized by Forrest Trumbore (Fig. 5), the first (1970) Electronics Division Awardee.[50] The path by which the role of group III and V impurities were deduced as p- and n-type dopants, respectively, in silicon was reviewed by Jack Scaff.[1,51] The utilization of growth rate variations in germanium by Robert Hall, Solid-State Science and Technology Medalist in 1977,[52,53] was essential to the fabrication of p-n junctions and transistors. Since it was easier at the time to make good

Figure 5. (*left*) Willis Adcock and (*right*) Forrest Trumbore.

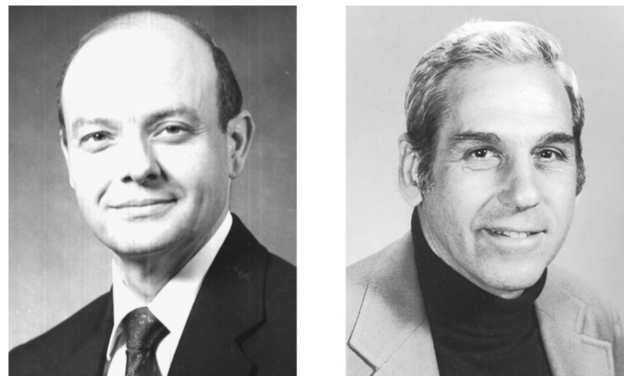

Figure 7. (*left*) Dennis Hess and (*right*) Ed Nicollian.

contact to a p-type base in an n-p-n transistor rather than to an n-type base in a p-n-p transistor, the former became commercially available, subsequently followed by the p-n-p transistor using a more complicated process.[54]

Transistor fabrication.—Silicon and germanium transistors were subsequently fabricated by solid-state diffusion processes in a mesa structure.[55,56] The grown-junction silicon transistor, introduced by Teal in 1954,[57] was subsequently described by Willis Adcock (Fig. 5) in conjunction with Mort Jones and colleagues.[58] Rate-growth single crystals of silicon, containing up to five n-p-n regions (with gallium and antimony dopants), was described for silicon transistors in 1955.[59] Silicon rapidly replaced germanium for transistor fabrication[36] as a result of silicon's larger energy gap which facilitated higher-temperature device operation and lower-reverse current. Additionally, silicon's oxide, SiO_2, was insoluble in water, whereas germanium's oxide was water soluble.[60] This attribute of silicon led to the fabrication of the planar silicon transistor by Jean Hoerni,[61,62] and facilitated its utilization in the planar process as a diffusion mask for p-n junction fabrication as developed by Carl Frosch and Link Derrick,[63] passivation of the silicon surface and p-n junctions intersecting the surface by Mohammed (John) Atalla and collegues,[64] and a dielectric layer for supporting metallic conductor overlayers.[60] All the elements were now available (oxidation, diffusion, photolithography, and thermocompression bonding[7,65] for the fabrication of junction transistors and the silicon controlled rectifier (SCR) (also referred to as the four-layer switch or thrysistor), in John Moll's laboratory,[66] in conjunction with Nick Holonyak, Jr., 1983 Solid State Science and Technology Awardee.[67] The SCR, developed by John Moll (Fig. 6) in conjunction with Holonyak and colleagues has a rich history.[68,69] Jim Early concurrently improved our understanding of the static characteristics of conventional bipolar transistors by utilizing a heavily doped, thin base such that the space-charge widening in the collector enhanced the bipolar transistor's transist time.[70,71] The mesa and planar processes paved the way for the fabrication of the IC by Jack Kilby[15,16,18] and Robert Noyce,[17,18] respectively, in 1958. Henry Theurerer and colleagues[72] expanded the applicability of epitaxial structures[1] by implementing Bernard Murphy's localized, high-concentration sub-collector diffusion in the silicon substrate,[73,74] before epitaxial deposition, which enhanced bipolar IC performance.

Precursors to the IC era.—The concurrent announcement of a practical silicon MOS (unipolar) transistor by Dawon (David) Kahng and Attala[75,76] was based on Shockley's and earlier patents,[1,4] quickly followed by Stefan Hofstein and Frederick Heiman's MOS IC consisting of 16 silicon n-channel MOS transistors.[77] The description of the oxidation process and methodologies for controlling the silicon/silicon dioxide interface were essential for the subsequent successful commercialization of the MOSFET and implementation of the DRAM memory era in the early 1970s. Bruce Deal (Fig. 6), ECS President in 1988-1989, ECS Fellow, recipient of the Solid-State Science and Technology Award in 1993 and the 1974 Electronics Division Award as well as the 1982 DS&T Callinan Award and Andrew Grove described the oxidation kinetics of silicon.[78] Dennis Hess (Fig. 7), ECS President in 1996-1997, ECS Fellow, and 1993 DS&T Callinan Awardee,[79] Eugene Irene, 1988 DS&T Callinan Awardee,[80] Hisham Massoud,[81] co-editor of the *The Physics & Chemistry of SiO_2 and The Si-SiO_2 Interface*, and Stanley Raider, 1992 DS&T Callinan Awardee[82] and their colleagues extended this research. Concurrently, Vik Kapoor, ECS Fellow and 1991 DS&T Callinan Awardee and William Brown, 1996 DS&T Callinan Awardee, were co-editing the ECS *Silicon Nitride and Silicon Dioxide Thin Insulating Films* proceedings, with their colleagues. Bruce Deal described the Si-SiO_2 electrical interface stability in terms of 17 types of charge mechanisms[1] and intro-

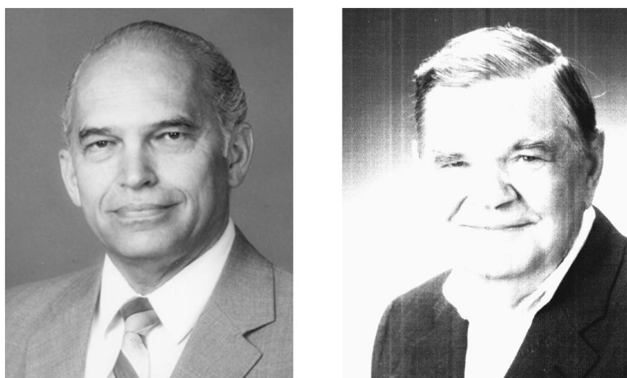

Figure 6. (*left*) Bruce Deal and (*right*) John Moll.

Figure 8. (*left*) Ken Bean and (*right*) Else Kooi.

Figure 9. (*left*) Bruce Hannay and (*right*) Walter Runyan.

duced the standard description for the charge notation associated with thermally oxidized silicon (SiO_2) and the Si-SiO_2 interface.[83-85] Pieter Balk described the significance of a post SiO_2 anneal in a hydrogen bearing ambient and a nitrogen anneal in the case of the Al-SiO_2-Si system[86] to stabilize the Si-SiO_2 interface and reduce the fixed charge, Q_f. Edward Nicollian (Fig. 7), ECS Fellow and 1986 Electronics Division Awardee, and co-author of the classic textbook *MOS (Metal Oxide Semiconductor) Physics and Technology* with John Brews, and Adolf Goetzberger described the Si-SiO_2 interface electrical properties[87–88] even as others described the p-n junction under nonequilibrium conditions.[1] More recently, Yves Chabal has edited a monograph on the *Fundamental Aspects of Silicon Oxidation*, emphasizing the chemical, morphological, and electronic configuration of the Si-SiO_2 interface accompanying the silicon oxidation process, complementing the Nicollian and Brews' compendium. The LOCOS process, developed by Else Kooi (Fig. 8) and colleagues in the late 1960s,[89–91] was instrumental in the fabrication and achievement of superior MOS IC characteristics. Kooi summarized this enhanced IC process methodology, a mainstay for CMOS IC fabrication for more than 30 years, in the first volume of the *IEEE Case Histories of Achievement in Science and Technology*.[90]

It was during this time that the Society's International Symposia on Silicon Materials Science and Technology series, the foremost Symposium describing the status, challenges, and opportunities in the advancement of silicon materials consonant with the advances in IC performance, yield and reliability, was initiated (1969), with the Ninth Symposium scheduled during the centennial anniversary of the ECS in May, 2002. Upon the occasion of the ECS's 75th anniversary, Deal and Early summarized the various device/IC configurations appropriate to that juncture.[92] Bean (Fig. 8), Electronics Division 1988 Awardee[93] described his epoch research on dielectric isolation for bipolar ICs with Paul Gleim and Walter Runyan (Fig. 9).[94,95] It was also during this time that Runyan and Grove published their classic textbooks, *Silicon Semiconductor Technology*[96] and *Physics and Technology of Semiconductor Devices*, respectively.[97] The exploding research on the physical chemistry and resulting electronic properties of single-crystal germanium and silicon (*i.e.*, crystal growth, control of composition, defects, structure, diffusion, electrical and thermal transport properties as well as surface and electrode properties) were summarized by Bruce Hannay (Fig. 9), ECS President in 1973-1974 in the classic monograph *Semiconductors*.[98] Several of the more relevant material properties of silicon and germanium were also summarized[99-103] as well as a detailed review of the physics of silicon.[104]

Finally, it should be noted that a rather innocuous observation that silicon crystals grown by the CZ method contain approximately 1.5×10^{18} cm^{-3} of oxygen, corresponding to 30 parts per million atomic (ppma) of oxygen[105-107] (due to the etching of the quartz crucible by the molten silicon), where the oxygen concentration quoted, based on the old ASTM calibration factor, F 121-79, Annual Book of ASTM Standards, Part 43, Electronics (1979), was to have repercussions to the present day as regards wafer warpage, plastic deformation and gettering (see Oxygen in Silicon in both the Plastic Deformation and Gettering sections below).

Plastic Deformation in Silicon

Electrical effects.—Although the Teal-Little or CZ technique resulted in the growth of single crystals, there were often residual dislocations and, as it turned out, uncontrolled point defect and point-defect complexes. The detrimental influence of dislocations on planar diffused diode and transistor performance was identified by Gunther Schwuttke (Fig. 10) and colleagues to be due to excessive leakage currents (*i.e.*, soft I-V characteristics) at sub-avalanche voltages related to the precipitation of metallics onto dislocations residing in the device space charge region (SCR).[108-110] Schwuttke also invented the scanning X-ray oscillator technique,[111] an indispensable diagnostic procedure for studying defects in semiconductors. In a related manner, the emitter-collector leakage in narrow-base devices was shown by Fred Barson and colleagues to result from the enhanced emitter dopant diffusion along dislocations threading the emitter and collector of a transistor (pipes), which resulted in electrical shorts.[112-114]

The degradation of gain, increased noise of PNP transistors, and increased transistor second breakdown as well as related effects in the degradation of p-n junction characteristics due to hydrostatic pressure, uniaxial stress, anisotropic stress, mechanical damage and other structural defects have also been summarized.[1] The recovery of p-n junction characteristics due to thermal processing and the beneficial effects of gettering procedures by Goetzberger and Shockey on device[115] and Michael Poponiak and colleagues IC characteristics[116] were also observed.

It was eventually recognized that the device degradation associated with dislocations was dependent on the detailed nature of the method and temperature of dislocation introduction during crystal growth and/or subsequent device/IC fabrication. In the latter case, especially important factors included the temperature and temperature gradient during IC fabrication processes including the wafer cumulative history, the degree of dislocation dissociation, the structure of the dislocation core, the type and degree of impurity capture,[117] stress effects due to the difference in size of the silicon and dopant atoms, residual peripheral latent sources of cracked dislocations, etc. John Lawrence showed that dislocations dynamically introduced during device or IC fabrication, due to the stress induced by the size difference between the silicon and dopant atom's size, were more detrimental in modifying the device microstructure and electrical characteristics than grown-in dislocations.[118-120] Freshly formed dislocations, for example, were found to enhance the emitter dopant diffusion along the dislocation, forming dislocation pipes and, thereby, resulting in emitter-collector shorts.[112-114,120] The large number of variables noted above also contributed to the significant variability of carrier lifetime with the "same" dislocation density

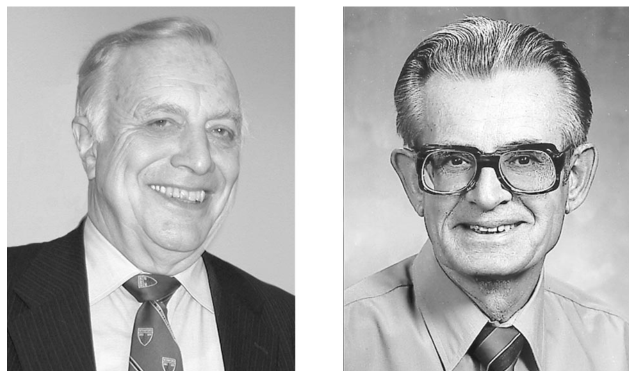

Figure 10. (*left*) Hans Queisser and (*right*) Gunther Schwuttke.

Figure 11. (*left*) Bob Brown and (*right*) Wilfred von Ammon.

and clarified why the direct effect of dislocations in silicon were often not observed.[1,121,122] The role of dislocations in silicon was eventually recognized to be an extremely complex topic and interrelated with point defects and impurities.[117,121]

The sum total of a large number of such observations, from the mid-1950s to the mid-1960s, was interpreted as requiring the use of dislocation-free silicon. The availability of dislocation-free material, however, did not necessarily improve the device/IC performance or yield,[123] which led to the "point-defect dilemma" enunciated by Erhard Sirtl, Electronics Division 1979 Awardee.[124] Indeed, there were indications that point-defect clusters were dominant and detrimental features of low-dislocation density wafers, consistent with earlier research that the structural perfection of dislocation-free silicon may be inferior for device characteristics compared to silicon with a moderate dislocation density, useful as gettering sinks for point defects as described by Hans Queisser (Fig. 10).[125-128]

Dislocation-Free Crystal Growth

Dash methodology.—The commercial introduction of dislocation-free silicon crystals in the mid-to-late 1960s was pioneered by William Dash[129-131] and immediately recognized as a major technological achievement. The design of heat-shielding components in the hot-zone of the crystal puller was subsequently shown to be of immense importance in controlling the temperature distribution at the melt-crystal interface (and point-defect incorporation in the growing crystal), as studied by Wilfred von Ammon (Fig. 11) and Eric Dornberger[132] and Robert Brown (Fig. 11)[133] and their colleagues. Although macroscopic dislocation-free growth is maintained because the Peierls stress[103] to initiate dislocation movement is sufficiently high ($\sim 10^{-2}$ of the bulk modulus), temperature fluctuations can induce both microscopic point-defect fluctuations in silicon, variously described as growth spirals, growth cores, swirl

Figure 12. (*left*) Ken Benson and (*right*) K. V. Ravi.

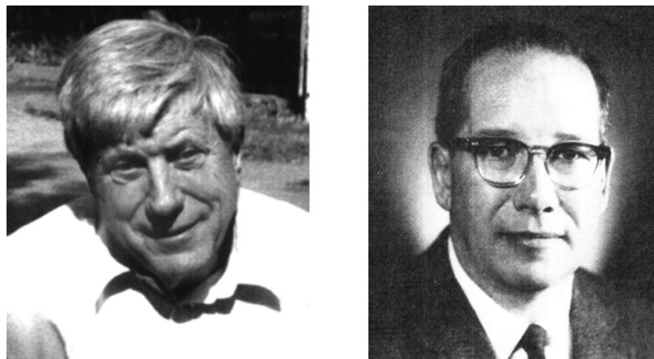

Figure 13. (*left*) Ron Newman and (*right*) William Pfann.

and striations[1,121] as well as fluctuations in diameter. The diameter fluctuations could be a considerable source of raw silicon yield loss, because a centerless grinding process was required to subsequently shape the grown crystal to a constant diameter. Sensing mechanisms and feedback procedures, therefore, to minimize temperature fluctuations at the melt-crystal interface received considerable attention.

Swirl is a generic term referring to a class of defect structures in silicon developed as a result of the condensation of point-defect complexes formed via solid-state reactions in the absence of readily available line and/or surface sinks and garnered extensive attention.[134-139] The defects distribute in a swirl pattern, reflecting the combined effects of melt convection, rotation of the crystal and temperature fluctuations at the crystal-melt interface during crystal growth. The observed defects reflect conditions subsequent to the complete crystal growth process, resulting from the complicated set of microscopic defect reactions in silicon. These complexities in CZ crystal growth and the dynamical properties of the silicon melt were extensively studied by Wen Lin and Kenneth Benson (Fig. 12) and colleagues,[140,141] Sam Rea,[142] and a plethora of other personnel.[1] A thermal oxidation process followed by an oxide strip and a preferential etchant is generally required to observe these defects, although electrical analysis such as carrier-lifetime mapping has also proved useful. The microdefects in silicon have distinct electrical effects as shown, for example, by K. V. Ravi (Fig. 12) and Charles Varker's studies[143,144] and, for example, could significantly impact CCD imagers[145,146] as discussed by Lubek Jastrzebski and colleagues.

The utilization of a magnetic field of about 0.3 T ($1\ \mathrm{T} = 10^4$ gauss) perpendicular to the crystal growth axis was shown to stabilize the temperature at the melt-crystal interface. Cusp magnetic field configurations have also been utilized. The magnetic field also reduces the erosion of the quartz crucible as a source of oxygen as well as reduce the thermal convection currents in the metallic-like silicon melt transporting oxygen to the solidifying crystal, thereby reducing both the magnitude of the oxygen content and increasing its uniformity.[147]

Oxygen in silicon.—Dislocation-free samples plastically deformed similarly as samples with 10^3 dislocations/cm^2 as shown by Witold Sylwestrowicz.[148] The nucleation of new dislocations during plastic deformation was deduced to overwhelm the grown-in dislocations.[148] Macroscopic dislocation-free silicon crystals, however, exhibited a larger spread in the ratio of the lower-to-upper yield points compared to crystals with grown-in dislocations.[148] The observation that dislocation-free silicon deforms plastically, furthermore, indicated that the Frank-Read mechanism was not the only possible source of dislocations. Rather, the oxygen content, its state of aggregation and the detailed thermal process conditions, was deduced to play a dominant role in plastic deformation of CZ crystals.[149] Jim Patel (Fig. 13)[150-152] and Ronald Newman (Fig.

13)[153-156] and their colleagues were just two of an extensive array of personnel studying oxygen in CZ silicon, to be further discussed below.

The disappearance of the upper yield point between approximately 1100-1300°C was interpreted as a result of the formation of SiO_x precipitates due to the supersaturated oxygen content.[148] Above 1300°C, however, the SiO_x structures dissociate as the degree of super-saturation decreases and the oxygen returns to its atomic form, with the concurrent recovery of both the upper and lower yield points. Detailed studies of the correlation of the precipitated oxygen on both the upper and lower yield point in CZ silicon were pursued by Koji Sumino[121,157-160] and Yojiro Kondo[161] and their colleagues. The functional dependence of the yield point on the dislocation speed was found to be related to the charge state of the dislocation kink and, therefore, dependent on the Fermi energy (*i.e.*, the doping level).[162] The comparative values of silicon's critical shear stress at different temperatures and varying oxygen precipitate concentrations have been summarized.[1,121,163-165] The critical stress increases with both reduced wafer processing temperature and reduced oxygen precipitation. Alternatively, oxygen precipitation can reduce the critical stress for subsequent plastic deformation by the generation of internal stresses around the precipitates. The oxygen microstructure and extent of precipitated oxygen was shown to be dependent on both material and IC thermal processes as well as equipment issues, such as the insert rate of the boat into the furnace (set a given dwell temperature), wafer spacing, thermal design of the furnace, etc.[1,121,165,166] The dislocation structure was noted to preferentially form on the concave side of a bowed wafer by Bernard Leroy, Lawrence Dyes and Howard Huff (Fig. 14) and their colleagues.[164,167] Excessive oxygen precipitation can also preclude effective lithographic printing due to uncontrolled wafer warpage as shown by Hirofumi Shimizu and colleagues.[168,169] The increased ultilization of rapid thermal annealing (RTA) requires attentiveness to minimize global alignment errors as well as to monitor warpage so as to minimize global alignment errors in lithography. Site flatness degradation may not be as severe, and may be corrected for by site-to-site realignment. Although a number of these observations were for significantly different wafer diameters and oxygen content, the comments were generically applicable to 200 mm and may be useful guidelines for the approaching 300 mm diam era.

In view of the importance of the oxygen content for internal gettering (IG) (see Gettering section below) and considering the influence of uncontrolled oxygen content in plastically deforming and warping the wafer, the comprehension of oxygen in silicon became very critical. Detailed investigations of the mechanisms and concentration of oxygen incorporated in the grown CZ crystal were implemented[1,121,149,165,170-174] The oxygen microstructure and extent of precipitated oxygen was pursued by numerous investigators, covering the oxygen range as high as 2.0×10^{18} cm^{-3}, corresponding to 40 ppma (though typically 1.5×10^{18} cm^{-3}, corresponding to 30 ppma) to as low as $2\text{-}5 \times 10^{17}$ cm^{-3}, corresponding to 4-10 ppma.[1,121,149,165,170-174] The importance of concurrently reducing the carbon concentration to less than 1×10^{16} cm^{-3}, 0.2 ppma to minimize its catalytic effects on oxygen precipitation was also noted.[175,176]

Precipitated oxygen enhanced plastic deformation in CZ wafers, compared to float-zone (FZ) silicon (utilized for power devices), where the oxygen content is typically less than 10^{16} cm^{-3}.[172] Small clusters of nitrogen in FZ crystals, however, were shown to be even more effective as regards work-hardening compared to oxygen in CZ crystals.[172,177] Nevertheless, FZ material has not replaced CZ material as the dominant material for IC fabrication, probably as a result of the extensive experience already in place with the utilization of CZ material, the wide installed base of IC processes consonant with CZ material and the larger wafer diameter available with CZ, compared to FZ, material.

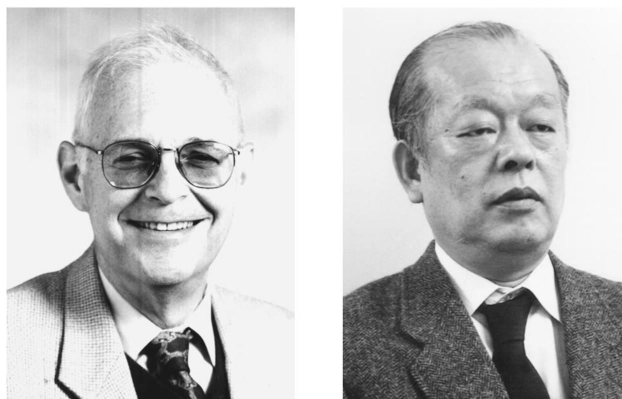

Figure 14. (*left*) Howard Huff and (*right*) Shin-ichiro Takasu.

Wafer Mechanics

The commercial introduction of dislocation-free crystals in the mid-1960s clarified the differences among dislocations; grown-in during crystal growth, introduced into the crystal during the mechanical conversion of the crystal into polished wafers, and generated in the wafer during epitaxial deposition and IC fabrication. A silicon wafer prepared for microelectronics applications requires a polished, flat upper surface free from residual edge mechanical damage and edge chips, etc. (due to the shaping processes in the conversion from the crystal to a polished wafer) and minimal residual polish damage, chemical contamination, and particulates, as well as exhibit controlled surface microroughness.[1,121,178,179] It was experimentally found that extreme care is required to maintain the optimally prepared silicon wafer during device and IC fabrication. It was noted as early as 1966 by Schwuttke that "one should handle wafers as little as possible and use extreme care when it is necessary."[108] Today's 300 mm diam silicon wafers have both a polished wafer back-surface and edge. The formation of dislocations was shown to be an extremely process- and structure-sensitive parameter. Designing a wafer from a process-structure-property perspective, in order to avoid plastic deformation during IC fabrication, was usefully employed to assess, monitor and control the above variables in the multidimensional wafer and IC process space.[1,121,165]

Radial thermal gradients developed within silicon wafers as a result of nonuniform heating during the approach to (or cooling during the descent from) the epitaxial or furnace process temperature, the insert and dwell temperatures (and strain rate due to the boat insertion/withdrawal speed) and the furnace ramp rate to the processing temperature during the furnace process. The location of wafer support points and wafer spacing in conjunction with the wafer diameter, type of boat material, and related equipment and process parameters were also critical. The furnace boat shape and the material contacting the wafers also significantly influenced the thermal gradient within the wafers. A completely enclosed boat considerably reduced the thermal gradient across the wafer by reducing the rate of heat transfer to or from the wafers.[180] The portion of the wafer near the contact point with the rails of the boat can exhibit a significant temperature gradient and, therefore, activate dislocation sources. This gradient is the result of both local heat flow nonuniformities due to differing thermal conductivities between silicon and the quartz glass boat[165] as well as differences in their coefficients of thermal expansion. The latter source acts when the silicon wafer sticks to the quartz glass boat. Plastic deformation has been noted to be more difficult to initiate at these points of contact by the utilization of silicon carbide or silicon, rather than quartz glass, boats due to the more favorable matching of the coefficients of thermal expansion.[165]

The mathematical assessment of plastic deformation during IC fab oxidation processes was initiated by Shih-Ming (Jimmy) Hu