continued from page 15

ently did not often display.

I got into the transistor business a decade after the events described by Holonyak and therefore do not have the benefits of some of his personal insights. However, in reading the literature without reference to any personal knowledge of the authors, it seemed to me that the record was made crystal clear by two publications on the transistor in 1949: the first by Bardeen and Brattain in April,[1] the second by Shockley in July.[2] They are as different as night and day.

In the first paper, all is still murky. It states that "a complete quantitative theory is still not available." In fact, the theory given does not even qualitatively describe the essential transistor effects. For example, Bardeen and Brattain assign a significant role to electric field rather than diffusion in transport across the base, attribute current flow in the collector to electrons, and fail to dismiss surface conduction as a diversion and not an essential part of the real thing. As Holonyak emphasizes, they did recognize the importance of hole injection from the emitter, but they evidently did not understand, and certainly did not explain, its real role in transistor action.

In Shockley's paper, daylight streams in. He sees clearly the essential features of the original discovery of the transistor effect by Bardeen and Brattain, and he introduces the necessary simplifying assumptions for a satisfactory theory that leads to notions of how a bipolar transistor should really be implemented. These ideas have stood the test of time, and present bipolar-transistor theory and structure stem from that paper rather than the first. As if that were not enough, Shockley went on to provide us with a basic unipolar-transistor theory,[3] again as sole author. It is not overstating things to say that most if not all of modern transistor theory has roots in those two seminal papers.

Why Shockley's ideas were not shared with or by his colleagues, and they instead put into the record a publication that showed that they just didn't have the insights he brought to bear on the problem, is perhaps best understood in the context of Conyers Herring's remarks in another article in your special issue (page 26). These, together with accounts of Shockley's dealings with his colleagues in his transistor company, and his grossly "politically incorrect" social theories, suggest that Shockley was not exactly interested in winning popularity contests. He may or may not have been lovable to those around him, but it is clear that he showed even an outstanding physicist like John Bardeen a clean pair of intellectual heels on the transistor project, and that we owe him a great debt for the brilliance of the insights he gave us.

References

1. J. Bardeen, W. H. Brattain, Phys. Rev. **75**, 1208 (1949).
2. W. Shockley, Bell Syst. Tech. J. **28**, 435 (1949).
3. W. Shockley, Proc. IRE **40**, 1365 (1952).

Bernard T. Murphy
4/92 *Mertztown, Pennsylvania*

Holonyak replies: My article on John Bardeen is not concerned with William Shockley or whether junction transistors are better understood than point-contact transistors. The article, I think, makes clear the real issue: How and when did the idea of carrier injection with a current, and as a consequence the transistor, occur? In my 40-year association with Bardeen, I had ample opportunity to learn the answer. Nevertheless, I interrupted Bardeen in his June 1990 interview with NHK, the Japanese television company, when he stated that it was known in 1947 that increasing the temperature of or shining light on a semiconductor increased the electron–hole population. I asked specifically if anyone knew at the time that this could be done with a current. He shook his head and told us that the Bardeen and Brattain experiments revealed this on 16 December 1947, and a demonstration of the effect to the Bell Labs "brass" occurred on 23 December 1947, which is taken as the official date for the "birth" of the transistor. The point is, a new principle had been established for an amplifying device. The rest follows: the semiconductor electronics around us.

Is it necessary to say more? Is it necessary to mention that the point-contact device, because of how it functioned, gave rise to the name "transistor"? Is it necessary to repeat that Bardeen and Brattain recognized that a forward-biased "emitter" (John's terminology), a hole injection current into n-type germanium, gave rise to a current change in a reverse-biased "collector"? Who would have believed that the Ge band structure, then unknown, and carrier lifetime would have permitted carrier injection and collection, even if the idea existed? Accordingly, is it not correct that the point-contact device, whether complicated or not, or even still mysterious in its detailed behavior, is a bipolar transistor—in fact, the original bipolar transistor?

Perhaps I should quote directly from Bardeen and Brattain's US patent 2 524 035, filed on 17 June 1948 and granted on 3 October 1950: "When operated as an amplifier, the emitter is normally biased in the direction of easy current flow with respect to the body of the semiconductor block. The nature of the emitter electrode and of that portion of the semiconductor which is in the immediate neighborhood of the electrode contact is such that a substantial fraction of the current from this electrode is carried by charges whose signs are opposite to the signs of the mobile charges normally in excess in the body of the semiconductor. The collector is biased in the reverse, or high resistance direction relative to the body of the semiconductor."

No matter what we think of the junction transistor, there is no way we can pretend it had an origin apart from or independent of the point-contact transistor, which came first, revealed the phenomenon of minority-carrier injection (which Bardeen identified) and, as I already said, was the original bipolar transistor. Whether we like it or not, the point-contact transistor was the prototype for all succeeding bipolar (injection) devices.

There was no predecessor to Bardeen on the transistor, which is not true of others. All that Bernard T. Murphy's letter does is to make it apparent why I wrote about John Bardeen in the context of the point-contact transistor.

Nick Holonyak Jr
University of Illinois
6/92 *at Urbana–Champaign*

What Accelerates Auroral Particles?

In their appropriately enthusiastic account of several very welcome new initiatives in NASA's Small Explorer Program (December 1991, page 44), Daniel N. Baker, Gordon Chin and Robert F. Pfaff Jr include as their figure 5 a drawing (adapted from reference 1) depicting in one plane a set of electrostatic equipotentials apparently hovering above the Earth's auroral zone. The lines, representing several different levels of potential, seem, inconsistently, to be joined together at one end and enigmatically left open at the other. While it is true that there is still a school of thought that holds that the electrons accelerated toward the Earth's auroral regions gain their energy from a potential difference, as the figure is clearly intended to imply, readers acquainted with the fact that equipotentials are nonconnecting closed surfaces will realize at once that the joined, open-ended sketched lines do not represent anything physically meaningful and in particular do not define a particle accelerator.

Even if the supposed equipotentials are meant to be seen as approaching infinitely closely without actually touching, there is still a fundamental problem, quite apart from the impracticalities of infinitely strong electric fields and maintenance of the necessary space-charge configuration. It has been clear since being pointed out by T. G. Cowling[2] fully 50 years ago that traversal of static space-charge fields cannot account for auroral particle acceleration, on the irrefutable grounds that such fields are conservative. Poisson's equation ensures that in all such cases the line integral of the electric field along a trajectory passing through these fields completely vanishes. The potential-difference theory of auroral particle acceleration seems therefore to be fundamentally untenable. The essentially dynamic nature of wave, or equivalent, motion is required, of course, before energy can be transferred, as in a linear accelerator. Alternatively, a nonconservative element needs to be introduced, as in a battery- or dynamo-driven circuit or via the mechanically driven charging belt of a Van de Graaff generator. A full discussion of these surprisingly long-standing issues has recently been published.[3]

While the Fast Auroral Snapshot satellite can confidently be expected to produce many high-resolution measurements crucial to the understanding of auroral plasma physics, it also can safely be predicted that FAST will find neither interconnecting equipotentials nor energization by static conservative fields.

References

1. M. A. Temerin, C. W. Carlson, C. A. Catell, R. E. Ergun, J. P. McFadden, F. S. Mozer, D. M. Klumpar, W. K. Peterson, E. G. Shelley, R. C. Elphic, in *Physics of Space Plasmas*, SPI Conf. Proc. and Reprint Ser. 9, T. Chang, G. B. Crew, J. R. Jasperse, eds., Scientific Publishers, Cambridge, Mass. (1989), p. 343.
2. T. G. Cowling, Terr. Magn. Atmos. Electr. **47**, 209 (1942).
3. D. A. Bryant, R. Bingham, U. deAngelis, Phys. Rev. Lett. **68**, 37 (1992).

Duncan A. Bryant
Rutherford Appleton Laboratory
Chilton, Oxfordshire, England

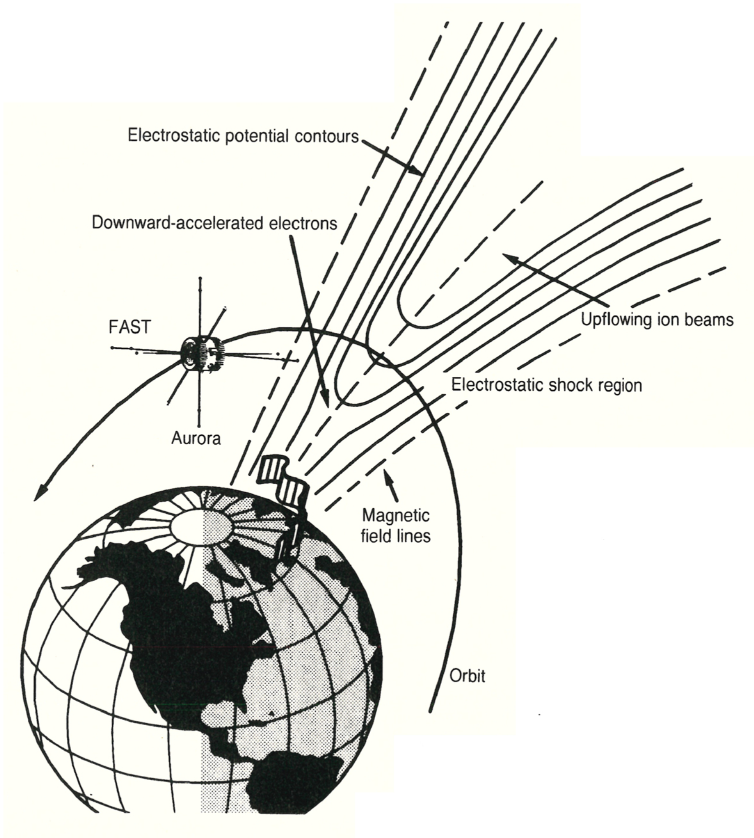
1/92

Baker, Chin and Pfaff reply: Duncan Bryant is correct to point out the deficiencies of the illustrative sketch of electrostatic potentials and the Earth's high-latitude (auroral) magnetic field lines in our article. It was not our intent to suggest that the potential contours actually "touch" each other. However, it was a conscious choice to depict the high-altitude portion of the electrostatic potential lines as "open," given the purposes and limitations of the sketch. Of course such potential contours *must* close eventually, but depending on the physical mechanism, one might imagine closure on a scale much larger than that of our diagram. (Similarly, magnetic field lines in the Earth's magnetotail, stretched by the solar wind, are often drawn as if they were open on diagrams of the magnetosphere that focus on regions near the Earth, even though such field lines also must eventually close.) An improved sketch is shown below.

The much more fundamental question that Bryant raises concerns how charged particles may be "energized" by their traversal of static space-charge fields. By this energization we refer to a mechanism by which the *kinetic* energy of an electron or ion is increased as it encounters a localized gradient in the ambient space-charge potential. We agree that such static fields are conservative, but since we are not arguing that the *total* energy of the particle is increased in this manner, we fail to see any fundamental problem here.

In our view, the situation is analogous to a ball rolling slowly along a level field that encounters a downward slope. The ball will be accelerated by the Earth's gravitational field and its kinetic energy will rapidly increase. (That is, we say it is "energized" by the potential well, but clearly we mean that some of its potential energy is converted to kinetic energy.) If the ball then encounters another level field, it will continue to roll at a higher velocity than it had on the upper level field, until it is slowed by friction. Although an instrument designed to measure the speed of balls on the lower field will record the fast speed of this particular ball, without knowledge of the existence of the steep slope or this ball's history, knowledge of how it was "accelerated" may only be hypothesized.

Similarly, at high latitudes in the Earth's auroral zone, scientists have observed distinctly accelerated charged particles amid a sea of slow, or "thermal," ones in the low-altitude ionosphere, and have wondered how such fast particles were energized. Although several possible mechanisms might produce this effect, we



see nothing wrong with the acceleration of charged particles by a localized space-charge potential gradient. A more detailed discussion refuting the ideas suggested in Bryant's letter has recently been published.[1]

Although probes on satellites have observed such localized electric potentials,[2] several key questions remain concerning their origins, their physical properties and their interaction with plasma waves and particles. Collecting the data necessary to address these questions is one of the chief goals of the FAST Small Explorer satellite.

References

1. J. E. Borovsky, Phys. Rev. Lett. **69**, 1054 (1992).
2. See, for example, F. S. Mozer *et al.*, Phys. Rev. Lett. **38**, 292 (1977); L. P. Block, C. G. Falthammar, J. Geophys. Res. **95**, 5877 (1990); and refs. in ref. 1.

Daniel N. Baker
Gordon Chin
Robert F. Pfaff Jr
NASA Goddard Space Flight Center
Greenbelt, Maryland

9/92

Supersymmetry History Addendum

Pierre Ramond's letter on the development of supersymmetry (May 1992, page 13) is a very appropriate complement to the article by Savas Dimopoulos, Stuart A. Raby and Frank Wilczek, but it does not give a complete history as yet.

André Neveu, John H. Schwarz and Ramond did indeed make the key step of discovering closed algebras of graded commutators, but the field theoretic aspect was completely absent from their discussions. We introduced this aspect in an article entitled "Field Theory Interpretation of Supergauges in Dual Models."[1] There we gave the first example of a supersymmetric Lagrangian. It involved two-dimensional spin-0 and spin-$\frac{1}{2}$ fields, and we formulated the corresponding superconformal tranformations. This supersymmetry was called supergauge because its parameters depend upon the (two-dimensional) space–time coordinates. Moreover, this work led us to consider, for the first time, transformations with anticommuting parameters, another key step toward supersymmetry as we now know it. We believe that our work played a significant role in the subsequent developments by Julius Wess and Bruno Zumino.

Reference

1. J.-L. Gervais, B. Sakita, Nucl. Phys. B