Structure in Large Sets: Two Proofs Where There Were None

At times mathematicians are taunted by their failure to solve a problem. Their failure confronts them with their lack of understanding of a particular field of research. This has happened, for example, in the field of combinatorics, which is a branch of mathematics that deals with configurations of elements of finite sets. For more than 40 years a major problem resisted all attempts to solve it. The Hungarian mathematician Paul Erdös, who, along with the late Paul Turán, first posed this difficult and important problem, offered $1000 to anyone who could solve it. Now, surprisingly, there are two solutions to the problem. Even more surprisingly, one of the solutions came out of ergodic theory—a field of mathematics whose origins, style, and even language are completely distinct from those of combinatorics.

The problem of Erdös and Turán had its origins in a theorem that was proved about 50 years ago by the Dutch mathematician B. L. van der Waerden. This theorem influenced the development of a branch of combinatorics known as Ramsey theory. It states that the set of all positive integers has an intrinsic structure that cannot be destroyed when the set is divided into a finite number of subsets. The particular structure referred to in van der Waerden's theorem is the existence of arithmetic progressions of all lengths. An arithmetic progression is a sequence of numbers with the property that the difference between any member of the sequence and the member following it is a fixed integer. The length of an arithmetic progression is the number of members of the sequence. Thus 5, 8, 11, 14, 17 is an arithmetic progression of length 5.

According to van der Waerden's theorem, any division of the positive integers into a finite number of subsets will lead to at least one subset that contains arithmetic progressions of any arbitrarily chosen length. This result was seminal to the development of Ramsey theory, the basic premise of which, according to Ronald Graham of Bell Laboratories in Murray Hill, New Jersey, is that "no matter how you break up a large structure into relatively few pieces, you will still leave a lot of smaller structures intact." An alternate explanation of Ramsey theory, Graham says, is that "complete disorder is impossible."

Ramsey theorists do not always work with infinite sets—they often work with very large finite sets instead. One problem that concerns them is to find the minimum size of a set that will ensure that a particular structure remain intact. A primary goal of Ramsey theorists is to discover why the structures in large sets are so resistant to dissolution. In the case of van der Waerden's theorem, the goal was to find the properties of these sets of integers that cause them to contain arithmetic progressions of arbitrary length. Erdös and Turán proposed an answer to this question. They conjectured that sets of integers that have a property called positive upper density—which basically means that the numbers in the set cannot be too sparsely distributed—will contain arithmetic progressions of arbitrary length.

The First Breakthrough

The conjecture seemed reasonable, but the search for its proof became a quest. The conjecture assumed a central position in Ramsey theory and its proof was attempted by some of the best mathematicians. The first breakthrough occurred in 1953, when Klaus Roth of the University of London, who had previously been awarded the Fields Medal (often called "the Nobel Prize of mathematics") showed that if a set of integers has positive asymptotic density, it must contain arithmetic progressions of length 3. No one could generalize Roth's proof to apply to larger progressions until, in 1968, E. Szemerédi of the Mathematical Institute of the Hungarian Institute of Sciences in Budapest found a completely different way to prove this result for arithmetic progressions of length 4. This proof, too, could not be generalized and, for a time, Szemerédi suggested that the conjecture might be false for progressions of length 5.

Finally, a few years ago, Szemerédi developed a long and exceedingly difficult proof of the conjecture of Erdös and Turán. Szemerédi himself never wrote out his proof but instead dictated it to a colleague. This version of the proof was nearly incomprehensible. Graham then took it upon himself to rewrite Szemerédi's proof so that it could be understood by others. Graham worked for several months on the proof; he lived with it, he says, until he finally felt that he understood it. It was not just the length of the proof—the final version was 100 typed pages—but the gestalt that was so hard to grasp. Erdös, who awarded the $1000 to Szemerédi, says, "The proof is one of the most difficult I have ever seen." Even though the proof was finally published last year, very few people actually examined it in its entirety. They relied instead on the word of such people as Graham, Erdös, and Roth that the proof was correct.

Now, Harry Furstenberg of the Hebrew University of Jerusalem has managed to prove the conjecture of Erdös and Turán with ergodic theory. This branch of mathematics arose from problems in a branch of physics, statistical mechanics, that deals with such things as the motions of idealized gas molecules. It is concerned with the average behavior of large collections of molecules that move randomly for indefinite periods of time. Although ergodic theory seems far removed from Ramsey theory, Furstenberg noticed an analogy between the concept of positive upper density and that of "measure" or size in a probability space. Ergodic theorists commonly deal with measure in probability spaces and have developed powerful theorems involving ramification of these ideas. Furstenberg translated Erdös' conjecture into the language of ergodic theory and used some of its powerful theorems in his proof.

Furstenberg's proof is still long—82 typed pages—and difficult to follow. Whether it is as difficult as Szemerédi's proof remains uncertain because no mathematician has yet managed to read both proofs. Furstenberg recently traveled about the United States and lectured on his proof. Because of this and because he is recognized as a credible mathematician, the validity of his result is accepted by ergodic theorists. His proof has not yet been published in a journal, however.

According to Donald Ornstein, an ergodic theorist from Stanford University, Furstenberg's proof represents the first time that an important result in combinatorics has been proved by means of ergodic theory. It means that ergodic theory and Ramsey theory may be more closely related than was previously suspected. Graham speculates that it may lead to new insights into both fields of mathematics, but it is still too soon to know what the long-term effects will be. It remains possible that these two fields

will touch and then drift apart again. But if there is something deeper in the connection between the two fields, sparks may fly.

Now that two proofs have been reported, combinatorial mathematicians feel that they are gaining insight into the theoretical underpinnings of their field. Erdös, however, has made a still stronger conjecture. He proposes that if the sums of the reciprocals of the integers in a subset diverge (that is, if the reciprocals are added up and, as more and more terms are added, the sum grows larger and larger without bounds) the subset will contain arithmetic progressions of arbitrary length.

If true, Erdös' latest conjecture would imply his original conjecture and would also solve a long-standing problem about the distribution of prime numbers. It would indicate that arbitrarily long arithmetic progressions of primes exist. So far, because the primes are so sparsely distributed among the rest of the integers, only short progressions of less than about length 16 have been found. Neither Erdös nor anyone else has any idea of how to prove or disprove this most recent conjecture. For the solution to this problem, Erdös is offering a $3000 reward. If the size of Erdös' reward indicates the relative difficulty of proving his new conjecture, nothing short of Revelation may allow the reward to be claimed.—GINA BARI KOLATA

Multiple Sclerosis: Two or More Viruses May Be Involved

Until recently, the strongest evidence that multiple sclerosis (MS) might be caused by a virus was epidemiological. The pattern of occurrence of the disease suggests that MS is very likely caused by a viral infection early in life. But there has been only indirect immunological evidence to support such a possibility. No one has ever isolated a virus that, when injected into animals, produces MS. Until 5 years ago, no one had even isolated from MS patients a virus that might be associated with the disease.

In the past 5 years, though, much more direct evidence has been obtained. Many investigators have found traces of the measles virus at different sites in the bodies of MS patients, and one group recently identified a persistent measles infection in MS patients. Another group has isolated from MS patients a virus that is serologically related to measles virus. And two other groups have identified a second, unrelated virus, that they think is associated with MS. Some of the results conflict with each other, and some are controversial. Nonetheless, these findings promise that a firm identification of the causative agent of MS may be made in the foreseeable future and that steps can be taken to prevent future infections.

Measles (rubeola) virus has been a prime suspect in MS at least since 1962, when John M. Adams and David T. Imagawa of the University of California School of Medicine at Los Angeles identified antibody to the measles virus in the blood and cerebrospinal fluid of MS patients. Other investigators, including Martin Panelius and Aimo A. Salmi of the University of Turku in Finland, Kenneth B. Fraser and his associates at the Queen's University of Belfast, and Jacob A. Brody and John L. Sever of the National Institute of Neurological Diseases and Stroke (NINDS), have subsequently confirmed that higher than normal concentrations of antibodies to measles virus occur in the blood of most MS patients and that the antibodies are present in cerebrospinal fluid from many MS patients. Panelius and Salmi and D. Carleton Gajdusek of NINDS have shown that antibodies to measles virus can also be isolated from the brains of some patients who died from MS.

These results are considered significant because viral antibodies are not generally found in the brain or cerebrospinal fluid. They thus suggest not only that a large number of MS patients had a measles infection, but also that the virus proliferates in the bodies of MS patients at one or more sites at which it does not proliferate in healthy individuals.

Antibodies to Many Viruses

The significance of these findings is clouded somewhat by the discovery of other antiviral antibodies in cerebrospinal fluids from MS patients. C. Henry Kempe and his associates at the University of Colorado School of Medicine, for example, found high concentrations of antibodies to vaccinia virus (which is used to vaccinate against smallpox) in cerebrospinal fluid from about half the patients they studied. And Natalie E. Cremer of the California State Department of Health has observed antibodies against many viruses, including measles, rubella, vaccinia, and herpes simplex. It has thus been reasoned that MS patients have an inborn defect in their immune systems that allows viruses to proliferate in the central nervous system.

Even though MS patients produce larger than normal quantities of antibody against measles and other viruses, they may have a deficient cellular immunity—that is, the virus-sensitized leukocytes that would normally abolish a viral infection may be produced in insufficient quantities or may not respond to an infection properly. John B. Zabriskie of Rockefeller University and Virginia Utermohlen of Cornell University have found, for instance, that leukocytes from MS patients do not respond to measles virus as effectively as leukocytes from healthy individuals. This impairment was not observed for any other virus tested. A similar defect in cellular immunity has been observed by Caspar Jersild, Torben Fog, and their associates at the Copenhagen University Hospital. Jersild and Fog, however, also observed a deficient response to other paramyxoviruses, members of the same family of viruses as measles.

Zabriskie and Utermohlen's results have been confirmed by William Sheremata of the Montreal Neurological Institute, but other investigators have not been successful in reproducing the results. Why different investigators have obtained different results is still unknown. Zabriskie's results suggest that the cellular immune system of MS patients has a specific defect which prevents it from responding adequately to an infection by measles virus. The results of other investigators suggest that there is a more general defect in the cellular immune system, and that the reduced response to measles virus is a particularly sensitive indicator of that defect.

That the latter possibility might be correct is indicated by experiments with transfer factor, a substance that is thought to provide a way to transfer cellular immunity. (Transfer factor is, itself, a controversial subject, and many scientists believe that there is no firm evidence to support its existence.) Transfer factor is thought to be a small nucleo-