the Spirit of Play

a memoir for Stan Ulam

by David Hawkins

Sometime in early 1944 I passed the open door of a small office near my own: S. ULAM. He had arrived at Los Alamos only a few days before and seemed unoccupied. We introduced ourselves—he a young mathematician, I an even younger philosopher, one with mathematical leanings. My field of work was the philosophy of mathematics and science. I had listened in on the shop talk of the theoretical physicists at Berkeley and knew their style. They thought of me for managerial chores in the newly created Los Alamos laboratory. So I came, as an administrative assistant to Robert Oppenheimer. Only later was I given the job of writing a wartime history. I was in fact the sole representative of my trade at Los Alamos, and the label "philosopher" usually caught curious attention. But Stan ignored it. He had come as a new member of the Theoretical Division, although no one (he slyly suggested) knew quite why. I later guessed that he had indeed been invited for no particular reason other than the urging of John von Neumann. Stan's version was characteristic: "Physicists don't know what to do with mathematicians."

It was the beginning of a long personal and family friendship. But here I shall restrict my recollections to associations of the thinner, more mathematical kind. We soon discovered one strong common interest, in the foundations and uses of probability theory. Some of Stan's work (Lomnicki and Ulam 1934) had preceded that of Kolmogorov on the measure-theoretic formulation of probability. Mine had been on the conceptual foundations, battled over since the time of Bernoulli and Leibnitz and closer to the philosophy of physics.
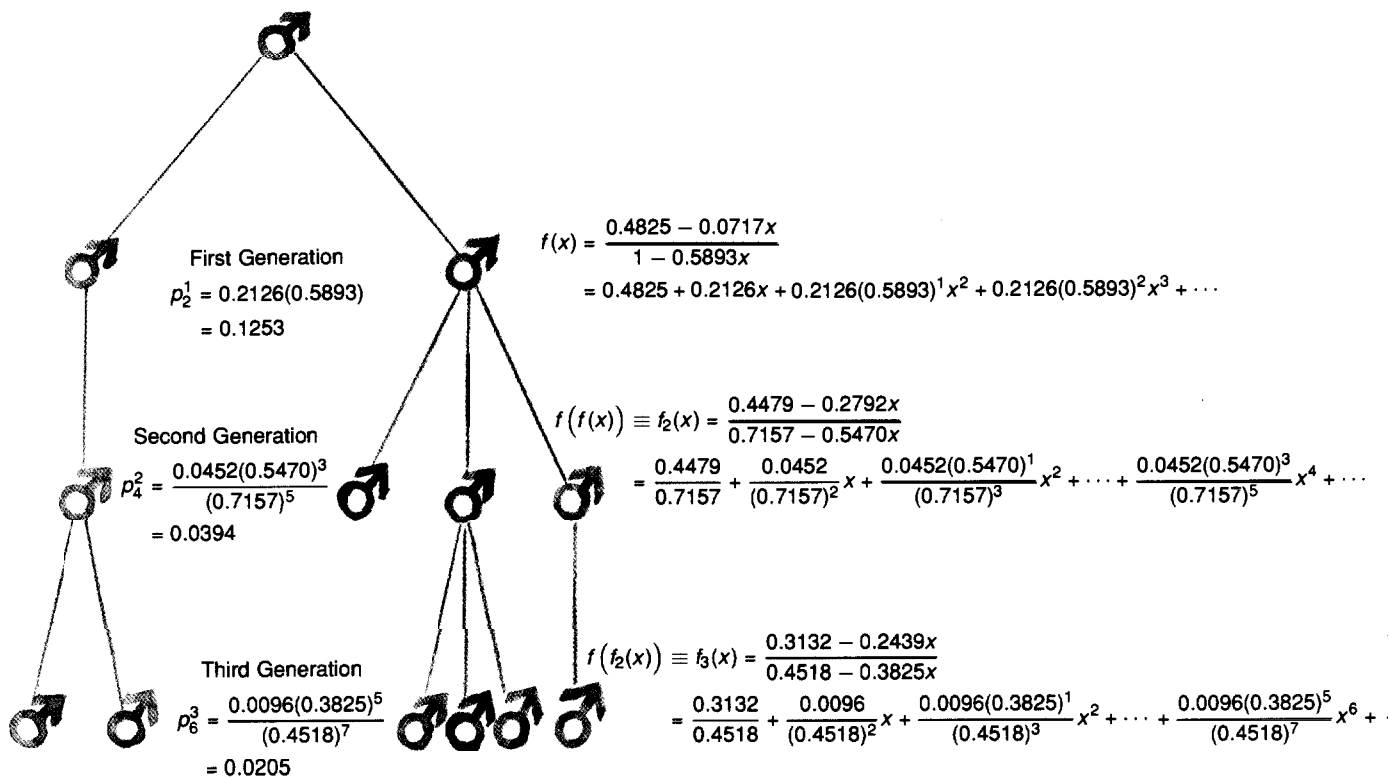
One day Stan threw a problem at me, as if to bring our academic discussions back to the concerns of a wartime laboratory. In the chain reaction that was to power the atomic bomb, some fraction of the neutrons liberated by a fission induce other fissions, which in turn liberate more neutrons that induce more fissions, and so on. Suppose the number of induced fissions per fission is a random variable that can take on the values i = O, 1, 2, . . . with probability P_i. (That is, p_0 is the probability that the neutrons from a single fission induce no further fissions, p_1 is the probability that they induce one further fission, and so on.) What then is the probability distribution of the number of fissions occurring in the nth "generation" of such a process started by a single fission? Although we didn't know it at the time, the same problem—stated differently—had been solved long before. One earlier version had been posed in terms of the proliferation of a family name through male descendants. Assume that each male Jones produces i male offspring with probability p_i, $i = O, 1, 2, . .$ (and that this probability does not vary from one generation to another). What then is the probability that a given Jones has k males in the nth generation of his descendants?

I spent several evenings on the problem. By persistence rather than insight I found the very simple solution (Hawkins and Ulam 1944). A lot of algebraic solvent evaporated and left behind an unexpected little crystal of a formula, the sort of outcome that makes you ask why it hadn't been obvious all along.

Let $f(x)$ be the Laplace generating function of the sequence of probabilities $\{p_0, p_1\ m.. \}$ (That is, let $f(x)$ be the function to which the infinite series $p_0 + p_1 x + p_2 x^2 + converges.$) Then the probability that Jones has k grandsons (or k second-generation male descendants) is the coefficient of x^k when $f_2(x) \equiv f(f(x))$ is expanded in powers of x. And in general the probability that Jones has k rzth-generation male descendants is the coefficient of x^k when $f_n(x) \equiv f(f_{n-1}(x))$ is similarly expanded. Thus, to the biological process, that of reproduction, there corresponds an algebraic one, that of iteration, in which the argument of a function is replaced by the function itself. 1'11 mention other related results and further applications later, but this was the essence of our first venture into what was to develop into the theory of branching (we said "multiplicative") processes. *

Stan was delighted with my solution, and I, the rank amateur, was flattered. He already knew quite a lot about the deceptively simple operation of iteratively substituting a function for its own argument, and I got a lesson or two. In the course of these discussions, we got on to such topics as space-filling curves, turbulence, and what have recently come to be called catastrophes, in which deterministic laws lead rigorously to results we can only describe as chaotic. A good many years later when we were reminiscing about all of this, I complained that we had almost been pioneers in such matters. Why hadn't we pursued them? Stan's reply: "It's because there are so many of them guys and so few of us!"

*I should also mention a prior Los Alamos paper by S. Frankel, which may lie buried in the 1943 series of Los Alamos reports. Frankel had thought in terms of a continuous time parameter instead of discrete generations. That approach leads to a one- parameter family of generating functions embedding our $f_n(x)$. The problem actually has an even earlier origin. It was discussed by Darwin's cousin Francis Galton in 1889 and then by A. Lotka in 1939. Later, in 1945, Erwin Schrödinger addressed the problem, and I recall seeing the title of a relevant Russian paper (obviously declassified!) of about the same date. A section of Feller's classic text on probability theory (Feller 1968) is devoted to branching processes; a full development is that of T. E. Harris (Harris 1963).

First Generation

$$p_2^1 = 0.2126(0.5893)$$
$$= 0.1253$$

Second Generation

$$p_4^2 = \frac{0.0452(0.5470)^3}{(0.7157)^5}$$
$$= 0.0394$$

Third Generation

$$p_6^3 = \frac{0.0096(0.3825)^5}{(0.4518)^7}$$
$$= 0.0205$$

$$f(x) = \frac{0.4825 - 0.0717x}{1 - 0.5893x}$$
$$= 0.4825 + 0.2126x + 0.2126(0.5893)^1 x^2 + 0.2126(0.5893)^2 x^3 + \cdots$$

$$f\big(f(x)\big) \equiv f_2(x) = \frac{0.4479 - 0.2792x}{0.7157 - 0.5470x}$$
$$= \frac{0.4479}{0.7157} + \frac{0.0452}{(0.7157)^2}x + \frac{0.0452(0.5470)^1}{(0.7157)^3}x^2 + \cdots + \frac{0.0452(0.5470)^3}{(0.7157)^5}x^4 + \cdots$$

$$f\big(f_2(x)\big) \equiv f_3(x) = \frac{0.3132 - 0.2439x}{0.4518 - 0.3825x}$$
$$= \frac{0.3132}{0.4518} + \frac{0.0096}{(0.4518)^2}x + \frac{0.0096(0.3825)^1}{(0.4518)^3}x^2 + \cdots + \frac{0.0096(0.3825)^5}{(0.4518)^7}x^6 + \cdots$$

A Biological Chain Reaction

A study in the thirties by A. Lotka showed that the probability p_i^1 of an American male having i male children was described by the sequence of probabilities

$$\{p_0^1, p_1^1, p_2^1, \ldots\} = \{0.4825, 0.2126, 0.2126(0.5893), 0.2126(0.5893)^2, \ldots\}$$

Assuming that this sequence is also applicable to American males of later generations, we can treat the production of male descendants (and the proliferation of a surname, assuming further that male children bear the surname of their father) as a simple branching process—a chain reaction. Let $f(x)$ be the generating function for Lotka's sequence of probabilities; that is,

$$f(x) \equiv 0.4825 + 0.2126x + 0.2126(0.5893) + 0.2126(0.5893)^2 + \cdots$$
$$= \frac{0.4825 - 0.0717x}{1 - 0.5893x}.$$

Then p_k^n, the probability that the nth-generation descendants of an American male includes k males, is the coefficient of x^k in the expansion of $f_n(x) \equiv f\big(f_{n-1}(x)\big)$. Illustrated here are the first three generations of male descendants of some ancestral male; the first generation includes two males, the second four, and the third six. Listed for each generation are $f_n(x)$ and its expansion in powers of x. The highlighted probabilities p_k^n depend only on n and k and not on the particular family tree. ∎

I know very little in detail of the wide range of Stan's work and his repertoire. In this memoir I shall confine myself to matters we corresponded about or worked on jointly. I do this partly because some of these may not be otherwise known and partly because they affected my own mathematical avocation in a way that throws some light on the character of Stan Ulam, teacher. I never sat in on any of his courses, to be sure, though I sometimes heard him lecture. The teaching I shall speak of is that I occasionally received, over many years, one-to-one. In talking about all this I shall refer to some work of mine that shows the nature of the Ulam influence; it is minor work but still a mirror of our associations. And I enjoy bringing these pieces together for the first time.

Stan was indeed a superb teacher, of a kind not very common. One part of his secret was a quite extraordinary talent for turning forbidding topics into attractive problems, attractive because they seemed promising, seemed to open up some larger area. Another part is a quality I am tempted to describe as meritorious laziness. Though Stan could, on occasion, himself engage in intense and concentrated work, as a teacher he would give you the challenge and then—let you do the work. I remember feeling a bit resentful. I did all the work on that first little paper, and he could have added more! But what he really added was to my confidence. For Stan no ego was invested.

Later, when I was at the University of Colorado, Stan and I both did some further work on branching processes. He, with C. J. Everett, had generalized the whole scheme by including "particles" of different types (Everett and Ulam 1948). This generalization, in its physical applications, allowed offspring and progenitors to differ from each other, for example, in their spatial or dynamical, and hence also in their reproductive, characteristics.

My own related work was inspired partly by a conversation we had about one of the great and vital mysteries of mathematics. The Greeks got on to it, long before Euclid, in the discovery that geometrical facts could be represented arithmetically, while those of arithmetic could be seen in the mirror of geometry. In our own day the pendulum has swung far toward the arithmetical, whether analytic or digital, side. Rather typically Stan took the "wrong" side, that of geometry. "Draw a curve," he said, "of a nice simple function. Now draw another curve parallel to it. The relation is very simple to see and understand, but algebraically it can be quite messy." How is it possible that relationships that are so complicated in one domain can be mapped into another where they appear so simple, or vice versa?

The generating-function transformation I had used in that first problem of ours is an elegant elementary example; it belongs to a wider family with many applications in applied mathematics, including probability theory. We had extended its use a bit, and it was Stan's challenge to extend it further, as he did in the work with Everett. To me the challenge was to explore the relevance of this transformation to other operations of a stochastic nature. Long known of course is the fact that addition of independent random variables corresponds to multiplication of their generating functions. What could one say about other arithmetical operations-division, say, or the logarithm—when random variables take the place of simple numbers?

Consider the following example. Physics students learn that the number of alpha particles emitted per unit time by a bit of uranium is a random variable (call it F) described by the Poisson distribution, whose Laplace generating function $f(x)$ is

$e^{-\lambda(1-x)}$, where λ is a decay constant characteristic of uranium. The time between emission of successive alpha particles is also a random variable (call it G); it is described by the exponential distribution, whose generating function $g(x)$ is $\lambda \int_0^\infty e^{-\lambda u} x^u du = \lambda/(\lambda - \ln x)$. Now F and G are reciprocals of each other. What then is the relation between f and g, the generating functions of their probability distributions? The answer is that they satisfy the remarkably simple relation

$$f^{-1}(x)g(x^{-1}) = f(x^{-1})g^{-1}(x) = 1,$$

where f^{-1} and g^{-1} denote the inverses of f and g: $f^{-1}(x) = (\ln x + \lambda)/\lambda$ and $g^{-1}(x) = e^{\lambda(1-1/x)}$.

One can show that this functional relation holds quite generally. Whatever the probability distribution may be for A per unit B, its generating function and that of the distribution for the reciprocal measure B per unit A will satisfy the above identity. (A mathematical nicety is that the inverses of such functions always exist.) Thus one can easily calculate means, variances, and higher moments of one distribution from those of the other.

A related topic is the "random logarithm": Find the probability distribution, for example, of the time required for a chain reaction to produce a given population size. I was in fact looking for some old notes on these matters, which Stan had asked about, when I learned of his death.

After the war I was absented from weaponeering—first from choice and then by the F.B.I. I became politically opposed to the arms race that supported it, but not to wartime friends. Over the following years Stan and I corresponded or talked about a good many different topics, and again it was he who got me thinking about some of these. As I write now, I realize they all concerned iterative processes, deterministic or stochastic or mixed, that seemed to lie beyond the range of "standard methods." So although it might seem a bit of a jump to go from chain reactions to prime numbers, both fitted that general category. There is an iterative definition of the prime numbers, the sieve of Eratosthenes. The process is completely deterministic, but the way the primes are scattered among the other numbers has a very chancy look that has stimulated generations—centuries—of investigation.

First there came from Stan some rolls of print-out: very long lists of primes, of twin primes, of successive differences between them, and so forth. All these of course were computer-generated. Others may have computed even longer lists; Stan was one of the first to do so. But soon the pattern changed. The theory of primes is a high-order specialty for number theorists, and a happy hunting ground for amateurs like me. Stan was neither, or both. I think he may have been the first to think of the sieve of Eratosthenes as merely one among many sorts of iterative processes whose products lie beyond the range of standard methods. He thought of a good many other ways of generating number sequences that were more or less prime-like in their frequency and distribution. It was a flanking maneuver: If you can't solve the original problem, think of others that resemble it, and may be easier. Some of Stan's schemes seemed to me far-fetched, and I said so. His reply: "Yes, but I am the village idiot!"

Indeed, I think that Stan often did not care whether he got to the essence of a

SIEVE OF ERATOSTHENES
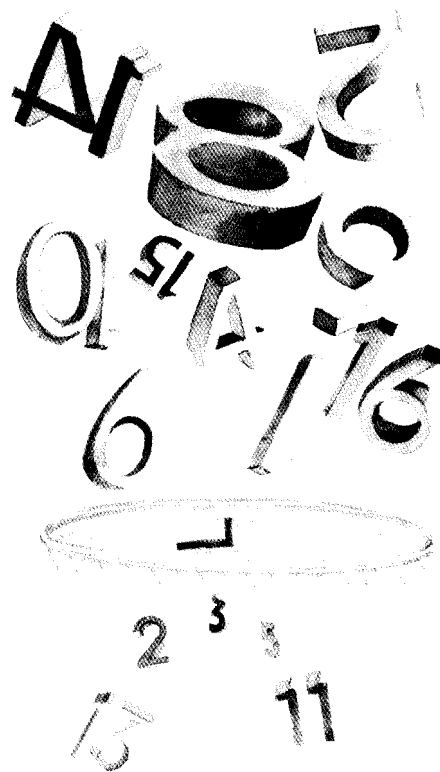
SOME RANDOM NUMBERS

The random numbers below are a small fraction of the 10,000 generated in the early 1900s by L. H. C. Tippett, then a member of the Biometric Laboratory at University College, London. Tippett generated the numbers, which were used in statistical sampling procedures, by selecting 40,000 single digits from census reports and combining them by fours. The collection of numbers was originality handwritten; the excerpt here is reprinted, with permission, from a version published in 1959 by Cambridge University Press (Random Sampling Numbers, Tracts for Computers, edited by E. S. Pearson, Number 15),

```
3684  8486  2617  5488
4798  1311  8701  2210
4776  4512  1746  2593
 780  0753  1232  7181
 473  3099  0729  4950

 745  9646  8105  3133
 834  8710  8829  8637
 123  2232  5039  5274
3403  3971  0899  0032
9255  5780  8728

9878  4918  0987  9118
0112  6523  8667  4707
6145  5895  0726  3883
9902  3764  7388  2729
7945  4847  4317  9636

4132  6668  0799  6137
1294  8988  1527  1467
7640  7128  1023  2418
5666  7671  1184  2328
0506  6473  5356  8940

0852  8066  5706  1940
3512  8981  7468  3836
3946  6777  1700  0323
8645  3290  7191  6152
2940  0237  7670  2852

7699  4151  8152  7271
6002  0857  5784  4480
6932  0871  1988  9881
0230  1337  7413  8840
1037  2978  6075  0971

0632  0693  2263  5290
1409  7624  4903  4692
2099  8290  3611  6501
4160  0429  3488  3741
7874  7264  4587  9591
```

particular problem. There are many things you can do with problems besides solving them. First you must define them, pose them. But then of course you can also *refine* them, *depose* them, or *repose* them, even *dissolve* them! A given problem may send you looking for analogies, and some of these may lead you astray, suggesting new and different problems, related or not to the original. Ends and means can get reversed. You had a goal, but the means you found didn't lead to it, so you found a new goal they did lead to. It's called play. Cyril Smith has argued persuasively, and with good historical evidence, that play, not utility, has long been the mother of invention (Smith 1981). Utility has been only the nursemaid. Creative mathematicians play a lot; around any problem really interesting they develop a whole cluster of analogies, of playthings.

One of Stan's playthings, his "lucky numbers," got considerable attention (Gardiner, Lazarus, Metropolis, and Ulam 1956). These numbers, which are generated by a sieve quite like that for the primes, have no particular arithmetical properties; they are just lucky to survive the sieving. A number of us got involved in studying their long-run distribution, which turned out to very close to that of the primes (Hawkins and Briggs 1957).

It is here I should mention an important one of Stan's contributions in the general grouping of nonstandard iterative processes. There was nothing lazy about his pursuit of a really good idea. "Pursuit" is probably the wrong word; it implies he already had the idea on the run. In the beginning it maybe was more like a roundup, a nudging together of possible example after possible example. I recall his ruminations about the Monte Carlo method in 1944, when already he was talking about it. More than fifty years ago there began to appear small compilations of random numbers. I remember the incredulity of a good physicist friend when I showed him such a listing. He knew Jahnke-Emde, of course, an old book of tables of almost every then-standard function. But *random* numbers? Statisticians were the wave of the future in those days. They alone used random numbers—for honest sampling procedures—even though nobody quite knew what "random" meant. But Stan foresaw that when high-speed computers should come along, they might well be used to imitate various deterministic or stochastic processes. Stan's phrase was "playing the game." The question was how to provide a computer with random numbers. I favored the built-in alpha counter, but that violated the sensibilities of mathematicians. There is indeed a grey area between chance and determinism, occupied by pseudo-random sequences. Even that early Stan was exploring it.

I think the most original part of Stan's early thinking about such matters was the idea that you could transform the equations describing a completely deterministic process into a mathematical form that also describes a stochastic one, and then you could get approximate solutions by playing the game repeatedly on a high-speed computer. But Monte Carlo is not my topic, except again as it affects my picture of Stan Ulam, teacher.

One of the topics we got on to later was a Monte Carlo approach to the theory of prime numbers. Pick a number at random from the neighborhood of some number N. The probability of picking a prime there is about $\frac{1}{\ln N}$, their approximate frequency. Why not reverse the process and produce a *random* sequence of numbers that mimic the primes? Try out each number $N = 2, 3, 4, 5, \ldots$ against a game of chance for which the probability of "winning" is $\frac{1}{\ln N}$, and select it for the sequence only if it wins. The

properties of such a sequence, or rather of many such sequences examined together, might throw light on some aspects of prime-number theory. The statistician Harald Cramer, it turned out, had already written about that. My own next bright idea was to turn the sieve of Eratosthenes itself into a Monte Carlo device. Drop out half the numbers beyond 2, namely those that "lose" a game of chance for which the probability of losing is $\frac{1}{2}$. Let N be the first survivor. Now drop out $\frac{1}{N}$ of the numbers beyond N, those that lose another game of chance for which the probability of losing is $\frac{1}{N}$. Keep repeating the process indefinitely, each time basing the sieving on the first survivor. One can play this game on a computer, which I did. But the theory of these "random primes" turned out to be not too difficult (Hawkins 1974), and it showed that the prime-number sequence could be regarded as one of an infinite family of sequences very much like it in their average properties. It supports some familiar conjectures about the primes and suggest others. The best result, I think, is that the famous and unproved Riemann hypothesis turns out to be true of "almost all" sequences generated by the random sieve. This hypothesis is a more recondite example of the kind of transformation I have talked about. It concerns the zeros of a certain function in the complex domain and, if true, implies a whole batch of propositions in number theory. Many of these can be proved independently, and none have been disproved. But some seem to be beyond the range of simple methods. That the Riemann hypothesis can be shown to be true of the random primes, and thus of almost all prime-like sequences, surely makes even more unlikely the possibility that the primes themselves should prove an exception.

Here, finally, I should mention another component of Stan's work, one that I can also trace back to early Los Alamos days. It grew later to very substantial proportions. One beginning I recall was to discuss a stochastic branching process that requires the "mating" of two "particles" from one generation in order to produce "offspring" for the next: sexual reproduction. Here the branching goes in both time directions, backward genealogically and forward by descent. The theory of this branching is essentially nonlinear. "Sex," Stan said, "is quadratic!" I had indeed examined one kind of nonlinear stochastic process, a chain reaction in which depletion of fuel, or of nutrient in the case of bacterial reproduction, is a factor. This led to a stochastic version of the well-known logistic curve of growth, which at first rises exponentially and then tapers off to a zero or negative slope. My work had a certain mathematical interest because it showed that the statistical fluctuations in such a nonlinear process can also change its average character; they don't "average out."

Such work as this might have stayed in abeyance except for Stan's development of other and much broader interests, namely in mathematical models of growth and reproduction. I remember approaching him with my own new-found interest in Claude Shannon's work on information theory and in the discovery of Watson and Crick. I wanted to define a measure of biological complexity, or organization, in information-theory terms, and we were immediately at loggerheads, He wanted to insist that very simple instructions could produce very complex patterns and I that such simplicity would nevertheless limit the *variety* of such patterns. Each of us was defending a different meaning of "complex." I already knew of his work (or play) with computer-generated growth patterns (Ulam 1962) but hadn't realized fully the range of ideas he was bent on exploring. Once more it was that flanking move. The genetic instruction

"SEX," Stan said, "IS QUADRATIC"

In this quote Stan was expressing a broad mathematical view of sex as a branching process in which some interaction, or "mating," between "male" and "femaie" members of a species is required for reproduction. An exampie is the deadly mating of male and female biack widow spiders.

Latrodectus mactans
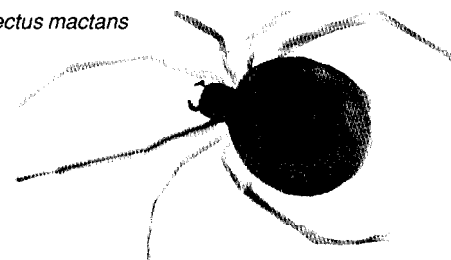
Oddiy enough, the animai kingdom inciudes some species, nameiy, a few of the tapeworms, that reproduce without any mating.

Tapeworm

The demography of a sexually reproducing species depends on (among other factors) a product of the maie and femaie populations-hence the adjective' 'quadratic." For mathematical simplicity mating is often assumed to be random, as it is for the ornamental ginkgo, or maidenhair, tree.

Ginkgo biloba
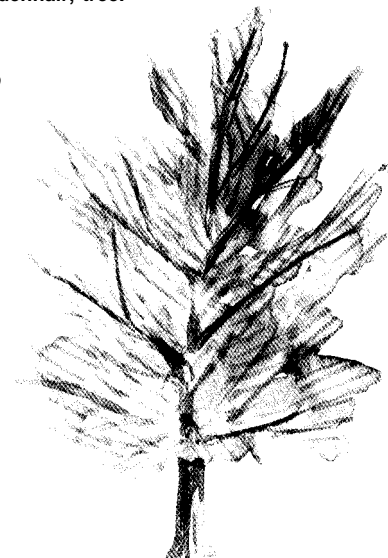
of biological growth and reproduction is a vast and still mostly uncharted domain for investigation. But once more the "village idiot" could invent all kinds of very simple processes bordering that domain. The idea of "growing" elaborate dendritic patterns, "organisms," by the endless repetition of a few simple "genetic" instructions, applied in each cycle to the results of previous cycles, was another in the category of iterative processes that lay beyond the range of standard methods. It later became the basis for the famous "game of life' '"—was Stan its first inventor? I don't know. I connect this work also with Stan's important work on the nature of and approach to equilibrium in even slightly nonlinear iterative processes. In the years following he became quite deeply involved in more realistic problems of genetics, but I mostly lost touch.

One of these problems, now well known and used in molecular genetics, came from Stan's deep familiarity with measure theory. Suppose a deck of cards can be shuffled only by several allowable operations. Knowing these and the end result of a shuffling, find the smallest number of allowable operations that accomplishes the given result, and call it the "distance" between the two orderings. Two decks of cards, or two nucleic acid strands, might appear very different in an item-by-item comparison yet be by shuffling history very close. Stan was a visiting professor at the University of Colorado's medical school when he worked on this, and I have a nice story from Theodore Puck. Stan got so interested in the mathematics (now *not* an iterative process) that he seemed to be ignoring the relevant biology. Reproached, he mended his ways. But he began his final talk on the subject with an imperative: "Ask not what mathematics can do for biology; ask rather what biology can do for mathematics!"

In the sixties and seventies I became more and more concerned with practical and theoretical work relating to elementary-school education in mathematics and science, to "school-doctoring." Toward this new career of mine Stan was-tolerant. We enjoyed good conversations but little time for shared work. It was only last year that I was suddenly recalled to our earliest association, catching up on some work he had done in population genetics and related matters. With characteristic initial disregard for humdrum scholarship, he had reinvented *and* extended some of the existing theory, developed first by R. A. Fisher and Sewall Wright.

I had known generally about this work but had missed one small paper, one in which he and Jan Mycielski formulated the basic theory of stochastic pairing, the branching process involved in sexual reproduction (Mycielski and Ulam 1969). Its main focus was not, however, on the fluctuational aspect of the process but on the average distribution and evolution of mutations within a species. The paper set forth three measures of the "distance" between two individuals. I shall mention only one of these, proposed by Mycielski. It is simply the sum, over the present generation and all past generations, of symmetric differences in genealogy; that is, the number of entries present in one family tree and absent from the other, plus the number present in the other and absent from the one. Since sexual reproduction is already a stochastic process, this measure is genetically crude (for example, it ignores sibling diversity). But it is surely a plausible first (or if you wish, zeroth) approximation—a measure of purely genealogical, not yet of genetic, distance.

Stan had done (as he often had for other problems) some Monte Carlo simulations assuming a constant population size of *2N,* random pairing between the N males and

females in each generation, and two offspring, one male and female, per pairing. The simulations had told him that the genealogical distance (as defined above) between two randomly selected individuals from the same generation was, on the average, about $4N$. It was shown subsequently, by Kahane and Marr, that the average distance is precisely $4(N - 1)$. Intrigued by all this and some explanation by Jan Mycielski, I found myself recovering some of our ancient lore, or reinventing it, and realized that we had been within an inch of this more recent thought, and then lost it, forty years ago!

Stan was an accomplished Latinist, but he deferred to my own (rather slight) knowledge of Greek. "Tell me," he once asked, "why is it that people are always saying 'Eureka!' and never—what is the Greek?—'I *lost* it!' After all, it's much more likely." I got the Greek for him, something like "$O\lambda\omega\lambda\alpha$!." It is by chance pronounced very like the modern French "Oh là là!," which can have a related meaning, "I am undone!" At any rate it did come back, and we had indeed lost it.

Not long after Stan retired from Los Alamos and came to Boulder, we touched on the subject again. If we trace the branching of ancestors back far enough, we of course share almost all of them. All men overwhelmingly are indeed almost-brothers; they are N-fold nth cousins. We jokingly speculated about the distance back to some common ancestor—Stan a Polish Jew and I an Anglo-American. Very likely less than twenty generations. All this was play, rediscovering the obvious along a pathway paved with numbers. I also remember that Stan the set-theorist poked his head out here with a "little remark." In a population of fixed finite size and infinite duration, everyone could be assigned an integer as a proper name, but the set of all sets of ancestors, of the genealogical "names," would be uncountable. We also noted more mundanely that the backward count of ancestors would fit some logistic curve, going up exponentially at first and then flattening out.

When I read the above paper, I was reminded of all this and noticed that the random pairing process was after all quite tractable. ("Noticed" is one of Stan's words; for me it represented hours of fussing about.) If, among the $2N$-member population that existed n generations ago, your mother had r female ancestors and your father had s, then the probability q_k that they share k female ancestors in that generation is given by

$$q_k = \frac{\binom{r}{k}\binom{N-r}{s-k}}{\binom{N}{s}}. \tag{1}$$

The sequence of probabilities defined by Eq. 1 is the hypergeometric distribution, a standard textbook entry. The applicability of this distribution to the problem at hand can perhaps be more easily seen by recasting it in textbook terms. Suppose we have an N-element set (the ancestral female population), r members of which are white (your mother's female ancestors) and the remainder are non-white (say black). Suppose we choose, randomly from among that set, an s-element set (your father's female ancestors, an identification allowed by the assumption of random pairing). The problem then is to find the probability q_k that exactly k elements of the chosen set are white (shared ancestors). Now the k white elements can be selected for the s-element set in $\binom{r}{k}$ ways and the $(s-k)$ black elements in $\binom{N-r}{s-k}$ ways. Since any choice of k white elements can be combined with any choice of black elements, the k white elements can be selected

GROWTH CURVE FOR EXPECTED NUMBER OF FEMALE ANCESTORS *n* GENERATIONS AGO

As *n* increases, the expected number of female ancestors departs more and more from 2^n because of the increase in the number of shared female ancestors.

$$A_n = N\left[1 - \left(1 - \frac{1}{N}\right)^{2^{n-1}}\right]$$

$N = 100$
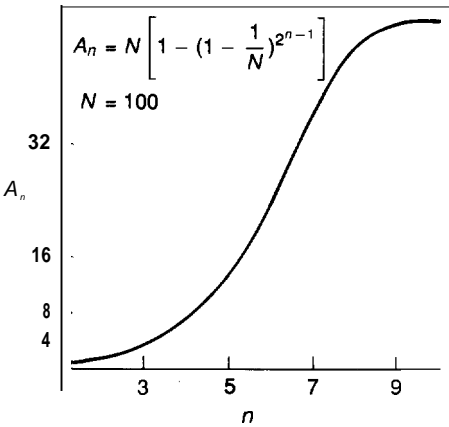
in a total of $\binom{r}{k}\binom{N-r}{s-k}$ ways. To obtain q_k, this number is then divided by $\binom{N}{s}$, the total number of ways of choosing the s-element set. (I should add that if the words "mother" and "father" are interchanged, along with r and s, the answer is the same; though the resulting formula for q_k will look different, it is not different in value.)

Using Eq. 1, we can now deduce the probability $p_{n+1,t}$ that, $(n + 1)$ generations back, you yourself have $t = r+s-k$ female ancestors. Let $p_{n,r}$ and $p_{n,s}$ be the respective probabilities that, among that generation, your mother has r female ancestors and that your father has s. Since various values for r, s, and $k = r+s - t$ can yield a particular t value, $p_{n+1,t}$ is a sum over those variables:

$$p_{n+1,t} = \sum_{s=2}^{t}\sum_{r=2}^{t} p_{n,r}p_{n,s} \frac{\left(\begin{array}{c} r \\ r+s-t \end{array}\right)\left(\begin{array}{c} N-r \\ t-r \end{array}\right)}{\left(\begin{array}{c} N \\ s \end{array}\right)}.$$

The Laplace generating function for this sequence of probabilities, call it $f_{n+1}(x)$, is therefore given by

$$f_{n+1}(x) = \sum_{t}\sum_{s}\sum_{r} p_{n,r}p_{n,s} \frac{\left(\begin{array}{c} r \\ r+s-t \end{array}\right)\left(\begin{array}{c} N-r \\ t-r \end{array}\right)}{\left(\begin{array}{c} N \\ s \end{array}\right)} x^t. \tag{2}$$

Equation 2 does not lend itself to derivation of an elegant recurrence relation between $f_{n+1}(x)$ and $f_n(x)$, but it does provide such a relation between A_{n+1} and A_n, where A_n is the expected, or average, number of female ancestors n generations back. This relation is

$$A_{n+1} = 2A_n - A_n^2/N, \tag{3}$$

in which
$$A_n = N\left[1 - \left(1 - \frac{1}{N}\right)^{2^{n-1}}\right] \tag{4}$$

(Interestingly enough, the right side of Eq. 4 is also the answer to a much simpler problem: If 2^n objects are distributed randomly among N boxes, what is the expected number of non-empty boxes?) If we identify the term A_n^2/N as the average number of shared female ancestors n generations back, then Eq. 4 defines just the logistic curve Stan and I had seen to describe the expected loss of ancestry; the difference between 2^n and A_n (the average number of your female ancestors n generations back) is just A_{n-1}^2/N (the average number of female ancestors among that generation shared by your parents).

Now Mycielski's definition of the expected genealogical distance between two randomly chosen individuals of the same generation can be written $\sum_0^\infty 2(A_n - A_n^2/N)$. We can evaluate this distance by using Eqs. 3 and 4:

$$\sum_1^\infty 2(A_n - A_n^2/N) = 2\sum_1^\infty (A_{n+1} - A_n) = 2(N - 1),$$

which, when doubled to include male ancestors, is just the result obtained by Kahane and Marr.

After I had "noticed" these simple relations (with Jan Mycielski's forbearance), I went to some of the literature of mathematical demography and population genetics and learned, of course, that it dealt with much more recondite problems, which I was loth to become involved in. Not equipped to make judgements, I nevertheless wondered why it seemed to skip over these simple zeroth approximations. And then I realized why I wished to talk about all this in a personal memoir about Stan Ulam. He, the "village idiot," the one who had the necessary "don't-know-how," did *not* skip over them. It was his style to value the art of successive approximation, of evading the big complexities until he was ready for them, the art sometimes called common sense. Many of his computer simulations were rough sketches of this kind yet could lead into deep water, such as his work on iterated nonlinear transformations. [See "Iteration of Maps, Strange Attractors, and Number Theory—An Ulamian Potpourri."]

With a few further modifications this mathematical genealogy begins to resemble a real biological story, possibly our own, and with all kinds of further questions in tow. I bring a closure to this writing by mentioning two such modifications, neither of which is so complex as to obscure the essential simplicity. The first recapitulates our early work on branching processes (Hawkins and Ulam 1944). Such processes appear *within* the scheme of sexual reproduction as soon as we shift from pure genealogy to genetics and to an interest in evolution. I shall describe briefly the simplest example. The second modification is necessary to give context for the first. It generalizes the original scheme, moving it away from the unrealistic assumption of random pairing toward a pattern of "assortative" mating. This move is curiously parallel to the later work of Everett and Ulam on branching processes in several dimensions (Everett and Ulam 1948).

Genetically considered, sexual reproduction is not only quadratic but also bi-quadratic: Each partner contributes to an offspring half of a diploid genome. But once inherited, the genetic makeup of the offspring remains constant, apart from mutations. Consider then the fate, within our model, of any individual genetic token, taken to be the only one of its kind. It will or will not be transmitted to an offspring with probability $\frac{1}{2}$. So the probability of its transmission to 0, 1, or 2 offspring is the coefficient of the corresponding power of x in the generating function $g(x) = (\frac{1}{2} + \frac{1}{2}x)^2$. Its appearance in subsequent generations is described by a simple chain reaction with $g_n(x) = g(g_{n-1}(x))$, one just at the level of transition from a subcritical to a super-critical condition. In any later generation the expected number of descendants with the token is a constant, namely 1. The probability that the token eventually disappears is 1, but its expected lifetime is infinite. The model itself forbids any evolutionary consequences. All of the model's essential properties are preserved, however, by allowing a *variation* of family size, insisting only on a mean value of 2. (Indeed even a slow exponential rate of population growth leaves essentials unchanged.) Then inheritance of any given "bad" gene will be decreased, and that token will have a finite expected lifetime. For a "good" gene the chain goes supercritical; with probability greater than $\frac{1}{2}$ the number of descendants with the "good" gene will grow exponentially with time and eventually dominate the population.

In such a way we can mimic stochastic adaptation. That is a necessary condition for evolution, but not sufficient. *Diveregent* adaptation is also necessary. If different environmental conditions face two subpopulations, "good" genetic changes in one might

be "bad" in the other. If the two are long separated, genealogical distances become very great, and the original gene pool may finally fission into those of separate species.

For such reasons we may consider a pattern of assortative mating that involves random pairing within subpopulations and rates of migration between them that decrease with some measure of distance. Successive generations in one subpopulation will gradually acquire more ancestors in the others. In the long run complete mixing will occur, but genealogical distances can now spread over a wide range. If the rate of mutation is assumed to be low but constant, genetic distances will increase with genealogical.

All this seemed at first quite difficult to mathematize, but surprisingly it is not. Shared ancestries and genealogical distances can be expressed in closed algebraic forms that depend only on the rates of diffusion between the subpopulations and their sizes. I leave the subject at this point. Stan's work in biomathematics went further in other areas, but this extension of early work I think would have pleased him.

I mentioned above that Stan was a bit standoffish about my involvement in work relating to the education of children. I was playing with them instead of him, my mathematical mentor! But I heartily forgive him. Some of what I had learned from him, that very spirit of play, I could take to the struggles for better science and mathematics teaching in the schools. Children don't have to be taught how to engage in serious play, usually, but teachers and other "educators" frequently do. They too often have *lost* the art, overwhelmed by mistaken notions of some puritan or utilitarian origin. Stan never lost it. ■
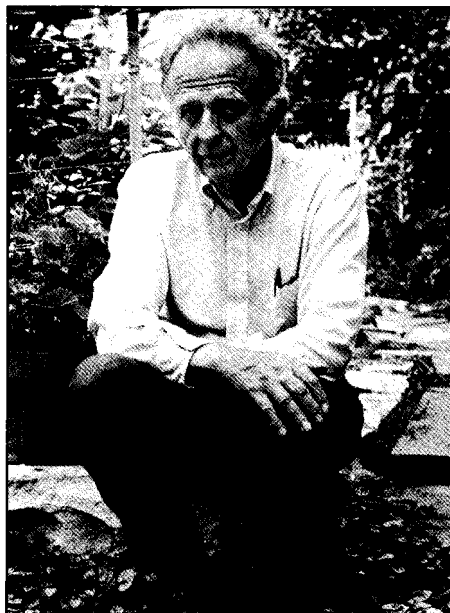
David Hawkins earned his academic degrees in philosophy: an A.B. and M.A. from Stanford University and a Ph.D. from the University of California, Berkeley. (The title of his doctoral dissertation, "A Causal Interpretation of Probability," reflects a combined interest in the humanities and science that continues to this day,) In 1943, after short teaching stints at Stanford and Berkeley, he joined the newly created Los Alamos laboratory, serving first as administrative aide to J, Robert Oppenheimer and later as historian, A year at George Washington University was followed in 1947 by a move, which proved permanent, to the University of Colorado, Boulder. He is now a Distinguished Professor Emeritus at that institution. Hawkins has devoted much of his professional life to projects concerning the teaching of mathematics and science. In 1970 he helped create the University of Colorado's Mountain View Center for Environmental Education, an advisory center for preschool and elementary teachers, and is still a participant in its activities. He has enjoyed leaves of absence at several colleges and universities in the United States and abroad and has been honored with a fellowship at the Institute for Advanced Study, a MacArthur Fellowship, membership in the Council of the Smithsonian Institution, and chairmanship of the Colorado Humanities Program. In addition to numerous journal articles, he has written four books: *Science and the Creative Spirit: Essays on Humanistic Aspects of Science* (Harcourt Brown, editor; 1958), *The Language of Nature: An Essay in the Philosoophy of Science (1964), The Informed Vision: Essays on Learning and Human Nature (1974),* and *The Science and Ethics of Equality (1977).*

Further Reading

David Hawkins. 1947. Manhattan District History: Project Y, The Los Alamos Project. Los Alamos Scientific Laboratory report LAMS–2532, vol. 1. Also in *Project Y: The Los Alamos Story.* The History of Modern Physics, 1800–1950, vol. 2. Los Angeles: Tomash Publishers, 1983.

Z. Lomnicki and S. Ulam. 1934. Sur la theorie de la mesure clans les espaces combinatoires et son application au calcul des probabilités. I. Variables indépendantes. *Fundamenta Mathematicae* 23: 237–278. Also in *Stanislaw Ulam: Sets, Numbers, and Universes,* edited by W. A. Beyer, J. Mycielski, and G.-C. Rota. Cambridge, Massachusetts: The MIT Press, 1974.

D. Hawkins and S, Ulam, 1944. Theory of multiplicative processes: Part 1. Los Alamos Scientific Laboratory report LA–171.

William Feller. 1968. *An Introduction to Probability Theory and Its Applications.* Third edition. New York: John Wiley & Sons, Inc.

Theodore E. Harris, 1963. *The Theory of Branching Processes.* Die Grundlehren der Matbematischen Wissenschaften in Einzeldarstellungen, vol. 119. Berlin: Springer-Verlag.

C. J. Everett and S. Ulam. 1948, Multiplicative systems, I. *Proceedings of the National Academy of Sciences 34: 403405.* Also in *Stanislaw Ulam: Sets, Numbers, and Universes,* edited by W, A. Beyer, J. Mycielski, and G.-C. Rota. Cambridge, Massachusetts: The MIT Press, 1974,

Cyril S. Smith, 1981. *A Search for Structure: Selected Essays on Science, Art, and History.* Cambridge, Massachusetts: The MIT Press.

Verna Gardiner, R, Lazarus, N. Metropolis, and S, Ulam. 1956. On certain sequences of integers defined by sieves. *Mathematics Magazine 29:* 117–122. Also in *Stanislaw Ulam.' Sets, Numbers, and Universes,* edited by W. A. Beyer, J. Mycielski, and G.-C. Rota, Cambridge, Massachusetts: The MIT Press, 1974.

D. Hawkins and W. E. Briggs. 1957, The lucky number theorem. ,*Mathematics Magazine 31: 277–280.*

D. Hawkins, 1974, Random sieves: Part 11, *Journal of Number Theory 6(3): 192–200.*

S. Ulam. 1962. On some mathematical problems connected with patterns of growth of figures. In *Proceedings of the Symposium on Mathematical Problems in the Biological Sciences,* pp. 215–224. American Mathematical Society Symposia in Applied Mathematics, vol. 14. Providence, Rhode Island: American Mathematical Society.

Jan Mycielski and S. M. Ulam. 1969. On the pairing process and the notion of genealogical distance, *Journal of Combinatorial Theory 6; 227–234.* Also in *Stanislaw, Ulam: Sets, Numbers, and Universes,* edited by W. A. Beyer, J, Mycielski, and G.-C. Rota. Cambridge, Massachusetts: The MIT Press, 1974.

Stanislaw M. Ulam. 1986. *Science, Computers, and People: From the Tree of Mathermatics.* Edited by Mark C. Reynolds and Gian-Carlo Rota, Boston: Birkhauser. This posthumous work of Stan's, which I received after writing this memoir, contains references to some of the matters I have discussed, notably in the essays on biomathematical topics.