1

The Phenomenology of Mathematical Beauty

Gian-Carlo Rota

Whereas painters and musicians are likely to be embarrassed by references to the beauty of their work, mathematicians enjoy discussions of the beauty of mathematics. Professional artists stress the technical rather than the aesthetic aspects of their work. Mathematicians, instead, are fond of passing judgment on the beauty of their favored pieces of mathematics. A cursory observation shows that the characteristics of mathematical beauty are at variance with those of artistic beauty. Courses in "art appreciation" are fairly common; it is unthinkable to find courses in "mathematical beauty appreciation." We will try to uncover the sense of the term "beauty" as it is used by mathematicians.

What Kind of Mathematics Can Be Beautiful?

Theorems, proofs, entire mathematical theories, a short step in the proof of some theorem, and definitions are at various times thought to be beautiful or ugly by mathematicians. Most frequently, the word "beautiful" is applied to theorems. In the second place we find proofs; a proof that is deemed beautiful tends to be short. Beautiful theories are also thought of as short, self-contained chapters fitting within broader theories. There are complex theories that every mathematician agrees to be beautiful, but these examples are not the ones that come to mind in making a list of beautiful pieces of mathematics. Theories that mathematicians consider beautiful seldom agree with the mathematics thought to be beautiful by the educated public. For example, classic Euclidean geometry is often proposed by non-mathematicians as a paradigm of a beautiful mathematical theory, but I have not heard it classified as such by professional mathematicians.

It is not uncommon for a definition to seem beautiful, especially when it is new. However, mathematicians are reluctant to admit the beauty of a definition. It would be interesting to investigate the reasons for this reluctance. Even when not explicitly acknowledged as such, beautiful definitions give themselves away by the success they meet. A peculiarity of twentieth-century mathematics is the appearance of theories where the definitions far exceed the theorems in beauty.

The most common instance of beauty in mathematics is a brilliant step in an otherwise undistinguished proof. Every budding mathematician quickly becomes familiar with this kind of mathematical beauty.

These instances of mathematical beauty are often independent of one another. A beautiful theorem may not be blessed with an equally beautiful proof; beautiful theorems with ugly proofs frequently occur. When a beautiful theorem is missing a beautiful proof, attempts are made by mathematicians to provide new proofs that will match the beauty of the theorem, with varying success. It is, however, impossible to find beautiful proofs of theorems that are not beautiful.

Examples

The theorem stating that in three dimensions there are only five regular solids (the Platonic solids) is generally considered to be beautiful. None of the proofs of this theorem, however—at least none of those known to me—can be said to be beautiful. Similarly, the prime number theorem is a beautiful result regarding the distribution of primes, but none of its proofs can be said to be particularly beautiful.

Hardy's opinion that much of the beauty of a mathematical statement or of a mathematical proof depends on the element of surprise is, in my opinion, mistaken.[1] True, the beauty of a piece of mathematics is often perceived with a feeling of pleasant surprise; nonetheless, one can find instances of surprising results that no one has ever thought of classifying as beautiful. Morley's theorem, stating that the adjacent trisectors of an arbitrary triangle meet in an equilateral triangle, is unquestionably surprising, but neither the statement nor any of the proofs are beautiful, despite repeated attempts to provide streamlined proofs. A great many theorems of mathematics, when first published, appear to be surprising; some twenty years ago the proof of the existence of nonequivalent differentiable structures on spheres of high dimension was thought to be surprising, but it did not occur to anyone to call such a fact beautiful, then or now.

Instances of theorems that are both beautiful and surprising abound. Often such surprise results from a proof that borrows ideas from another branch of mathematics. An example is the proof of the Weierstrass approximation theorem that uses the law of large numbers of probability.

An example of mathematical beauty upon which all mathematicians agree is Picard's theorem, asserting that an entire function of a complex variable takes all values with at most one exception. The limpid statement of this theorem is matched by the beauty of the five-line proof provided by Picard.

Axiom systems can be beautiful. Church's axiomatization of the propositional calculus, which is a simplified version of the one previously given by Russell and Whitehead in *Principia Mathematica*, is quite beautiful. Certain re-elaborations of the axioms of Euclidean geometry that issue from Hilbert's *Foundations of Geometry* are beautiful (for example, Coxeter's).[2] Hilbert's original axioms were clumsy and heavy-handed and required streamlining; this was done by several mathematicians of the last hundred years.

The axiomatization of the notion of category, discovered by Eilenberg and Mac Lane in the forties, is an example of beauty in a definition, though a controversial one. It has given rise to a new field, category theory, which is rich in beautiful and insightful definitions and poor in elegant proofs. The basic notions of this field, such as adjoint and representable functor, derived category, and topos, have carried the day with their beauty, and their beauty has been influential in steering the course of mathematics in the latter part of the twentieth century; however, the same cannot be said of the theorems, which remain clumsy.

An example of a beautiful theory on which most mathematicians are likely to agree is the theory of finite fields, initiated by E. H. Moore. Another is the Galois theory of equations, which invokes the once improbable notion of a group of permutations in proving the unsolvability by radicals of equations of degree greater than four. The beauty of this theory has inspired a great many expositions. It is my opinion that so far they have failed to convey the full beauty of the theory, even the renowned treatise written by Emil Artin in the forties.[3]

This example shows that the beauty of a mathematical theory is independent of the aesthetic qualities, or the lack of them, of the theory's rigorous expositions. Some beautiful theories may never be given a presentation that matches their beauty. Another such instance of a beautiful theory that has never been matched in beauty of presentation is Gentzen's natural deduction.

Instances of profound mathematical theories in which mathematical beauty plays a minor role abound. The theory of differential equations, both ordinary and partial, is fraught with ugly theorems and awkward arguments. Nonetheless, the theory has exerted a strong fascination on many mathematicians, aside from its applications.

Instances can also be found of mediocre theories of questionable beauty which are given brilliant, exciting presentations. The mixed blessing of an elegant presentation will endow the theory with an ephemeral beauty that seldom lasts beyond the span of a generation or a school of mathematics. The theory of the Lebesgue integral, viewed from the vantage point of one hundred years of functional analysis, has received more elegant presentation than it deserves. Synthetic projective geometry in the plane held great sway between 1850 and 1940. It is an instance of a theory whose beauty was largely in the eyes of its beholders. Numerous expositions were written of this theory by English and Italian mathematicians (the definitive one being the one given by the Americans Veblen and Young). These expositions vied with one another in elegance of presentation and in cleverness of proof; the subject became required by universities in several countries. In retrospect, one wonders what all the fuss was about. Nowadays, synthetic geometry is largely cultivated by historians, and an average mathematician ignores the main results of this once flourishing branch of mathematics. The claim raised by defenders of synthetic geometry, that synthetic proofs are more beautiful than analytic proofs, is demonstrably false. Even in the nineteenth century, invariant-theoretic techniques were available that could have provided elegant, coordinate-free analytic proofs of geometric facts without resorting to the gymnastics of synthetic reasoning and without having to stoop to using coordinates.

Beautiful presentations of entire mathematical theories are rare. When they occur, they have a profound influence. Hilbert's *Zahlbericht*,[4] Weber's

Algebra,[5] Feller's treatise on probability, and certain volumes of Bourbaki have influenced the mathematics of our day; one rereads these books with pleasure, even when familiar with their content. Such high-caliber expository work is more exploited than rewarded by the mathematical community.

Finally, it is easy to produce examples of a particular step in a theorem that is generally thought to be beautiful. In the theory of noncommutative rings, the use of Schur's lemma has often been thought of as a beautiful step. The application of the calculus of residues in the spectral theory of linear operators in Hilbert space is another such instance. In universal algebra, the two-sided characterization of a free algebra in the proof of Birkhoff's theorem on varieties is yet another such instance.

The Objectivity of Mathematical Beauty

The rise and fall of synthetic geometry shows that the beauty of a piece of mathematics is dependent upon schools and periods. A theorem that is in one context thought to be beautiful may in a different context appear trivial. Desargues's theorem is beautiful when viewed as a statement of synthetic projective geometry but loses all interest when stated in terms of coordinates.

Many occurrences of mathematical beauty fade or fall into triviality as mathematics progresses. However, given the historical period and the context, one finds substantial agreement among mathematicians as to which mathematics is to be regarded as beautiful. This agreement is not merely the perception of an aesthetic quality superimposed on the content of a piece of mathematics. A piece of mathematics that is agreed to be beautiful is more likely to be included in school curricula; the discoverer of a beautiful theorem is rewarded by promotions and awards; a beautiful argument will be imitated. In other words, the beauty of a piece of mathematics does not consist merely of the subjective feelings experienced by an observer. The beauty of a theorem is an objective property on a par with its truth. The truth of a theorem does not differ from its beauty by a greater degree of objectivity.

Mathematical truth is endowed with an absoluteness that few other phenomena can hope to match. On closer inspection, one realizes that this definitiveness needs to be tempered. The dependence of mathematical truth upon proof is its Achilles' heel. A proof that passes today's standard of rigor may no longer be considered rigorous by future generations. The entire theory upon which some theorem depends may at some later date be shown to be incomplete. Standards of rigor and relevance are context-dependent, and any change in these standards leads to a concomitant change in the standing of a seemingly timeless mathematical assertion.

Similar considerations apply to mathematical beauty. Mathematical beauty and mathematical truth share the fundamental property of objectivity, that of being inescapably context-dependent. Mathematical beauty and mathematical truth, like any other objective characteristics of mathematics, are subject to the laws of the real world, on a par with the laws of physics. Context-dependence is the first and basic such law.

A Digression into Bounty Words

A psychologist of my acquaintance received a grant to study how mathematics works. She decided that creativity plays a crucial role in mathematics. She noticed that an estimate of a mathematician's creativity is made at crucial times in his or her career. By observation of mathematicians at work, she was led to formulate a theory of mathematical creativity, and she devised ways of measuring it. She described how creativity fades in certain individuals at certain times. She outlined ways of enhancing creativity. In her final report, she made the recommendation to her sponsors that mathematics students should, at some time in their careers, be required to register for a course in creativity. Some college presidents took her suggestion seriously and proceeded to hire suitable faculty.

Our friend was seriously in error. It is impossible to deal with mathematical creativity in the way that she suggested. It is impossible to measure or teach creativity for the simple reason that creativity is a word devoid of identifiable content. One can characterize a mathematical paper as "creative" only after the paper has been understood. It is, however, impossible to produce on commission a "creatively" written mathematical paper. Creativity is what we propose to call a "bounty word," a word that promises some benefit that cannot be controlled or measured and that can be attained as the unpredictable by-product of some identifiable concrete activity.

Other bounty words are "happiness," "saintlihood," and "mathematical beauty." Like creativity and happiness, mathematical beauty cannot be taught or sought after; nevertheless, any mathematician may come up with some beautiful statement or some beautiful proof at unpredictable times. The error my friend made might be called "the bounty error." It consists of endowing a bounty word with measurable content.

It is unlikely that a mathematician will commit the bounty error in regards to mathematical beauty. Passing judgment on a piece of mathematics on the basis of its beauty is a risky business. In the first place, theorems or proofs that are agreed to be beautiful are rare. In the second place, mathematical research does not strive for beauty. Every mathematician knows that beauty cannot be sought *directly*. Mathematicians work to solve problems and to invent theories that will shed new light, not to produce beautiful theorems or pretty proofs.

Even in the teaching of mathematics, beauty plays a minor role. One may lead a class to a point where the students appreciate a beautiful result. However, attempts to arouse interest in the classroom on the basis of the beauty of the material are likely to backfire. Students may be favorably impressed by the elegance of a teacher's presentation, but they can seldom be made aware of beauty. Appreciation of mathematical beauty requires familiarity with a mathematical theory, which is arrived at at the cost of time, effort, exercise, and *Sitzfleisch* rather than by training in beauty appreciation.

There is a difference between mathematical beauty and mathematical elegance. Although one cannot strive for mathematical beauty, one can achieve elegance in the presentation of mathematics. In preparing to deliver a mathematics lecture, mathematicians often choose to stress elegance and succeed in recasting the material in a fashion that everyone will agree is elegant. Mathematical elegance has to do with the presentation of mathematics, and only tangentially does it relate to its content. A beautiful proof—for example, Hermann Weyl's proof of the equidistribution theorem—can be presented elegantly and inelegantly. Certain elegant mathematicians have never produced a beautiful theorem.

Mathematical Ugliness

It may help our understanding of mathematical beauty to consider its opposite. Lack of beauty in a piece of mathematics is a frequent occurrence, and it is a motivation for further research. Lack of beauty is related to lack of definitiveness. A beautiful proof is more often than not the definitive proof (though a definitive proof need not be beautiful); a beautiful theorem is not likely to be improved upon, though often it is a motive for the development of definitive theories in which it may be ensconced.

Beauty is seldom associated with pioneering work. The first proof of a difficult theorem is seldom beautiful. Strangely, mathematicians do not like to admit that much mathematical research consists precisely of polishing and refining statements and proofs of known results. However, a cursory look at any mathematics research journal will confirm this state of affairs.

Mathematicians seldom use the word "ugly." In its place are such disparaging terms as "clumsy," "awkward," "obscure," "redundant," and, in the case of proofs, "technical," "auxiliary," and "pointless." But the most frequent expression of condemnation is the rhetorical question, "What is this good for?"

Observe the weirdness of such a question. Most results in pure mathematics, even the deepest ones, are not "good" for anything. In light of such lack of applications, the disparaging question, "What is this good for?" is baffling. No mathematician who poses this rhetorical question about some mathematical theorem really means to ask for a list of applications. What, then, is the sense of this question? By analyzing the hidden motivation of the question, we come closer to the hidden sense of mathematical beauty.

The Light Bulb Mistake

The beauty of a piece of mathematics is frequently associated with shortness of statement or of proof. How we wish that all beautiful pieces of mathematics shared the snappy immediacy of Picard's theorem. This wish is rarely fulfilled. A great many beautiful arguments are long-winded and require extensive buildup. Familiarity with a huge amount of background material is the condition for understanding mathematics. A proof is viewed as beautiful only after one is made aware of previous, clumsier proofs.

Despite the fact that most proofs are long, and despite our need for extensive background, we think back to instances of appreciating mathematical beauty as if they had been perceived in a moment of bliss, in a sudden flash like a lightbulb suddenly being lit. The effort put into understanding the proof, the background material, the difficulties encountered in unraveling an intricate sequence of inferences fade and magically disappear the moment we become aware of the beauty of a theorem. The painful process of learning fades from memory, and only the flash of insight remains.

We would *like* mathematical beauty to consist of this flash; mathematical beauty *should* be appreciated with the instantaneousness of a light bulb being lit. However, it would be an error to pretend that the appreciation of mathematical beauty is what we vaingloriously feel it should be, namely, an instantaneous flash. Yet this very denial of the truth occurs much too frequently.

The lightbulb mistake is often taken as a paradigm in teaching mathematics. Forgetful of our learning pains, we demand that our students display a flash of understanding with every argument we present. Worse yet, we mislead our students by trying to convince them that such flashes of understanding are the core of mathematical appreciation.

Attempts have been made to string together beautiful mathematical results and to present them in books bearing such attractive titles as *The One Hundred Most Beautiful Theorems of Mathematics*. Such anthologies are seldom found on a mathematician's bookshelf.

The beauty of a theorem is best observed when the theorem is presented as the crown jewel within the context of a theory. But when mathematical theorems from disparate areas are strung together and presented as "pearls," they are likely to be appreciated only by those who are already familiar with them.

The Concept of Mathematical Beauty

The lightbulb mistake is our clue to understanding the hidden sense of mathematical beauty. The stark contrast between the effort required for the appreciation of mathematical beauty and the imaginary view mathematicians cherish of a flashlike perception of beauty is the *Leitfaden* that leads us to discover what mathematical beauty is.

Mathematicians are concerned with the truth. In mathematics, however, there is an ambiguity in the use of the word "truth." This ambiguity can be observed whenever mathematicians claim that beauty is the raison d'être of mathematics, or that mathematical beauty is what gives mathematics a unique standing among the sciences. These claims are as old as mathematics and lead us to suspect that mathematical truth and mathematical beauty may be related.

Mathematical beauty and mathematical truth share one important property. Neither of them admits degrees. Mathematicians are annoyed by the graded truth they observe in other sciences.

Mathematicians ask "What is this good for?" when they are puzzled by some mathematical assertion, not because they are unable to follow the proof or the applications. Quite the contrary. Mathematicians have been able to verify its truth in the logical sense of the term, but something is still missing. The mathematician who is baffled and asks "What is this good for?" is missing the *sense* of the statement that has been verified to be true. Verification alone does not give us a clue as to the role of a statement within the theory; it does not explain the *relevance* of the statement. In short, the logical truth of a statement does not enlighten us as to the *sense* of the statement. *Enlightenment*, not truth, is what the mathematician seeks when asking, "What is this good for?" Enlightenment is a feature of mathematics about which very little has been written.

The property of being enlightening is objectively attributed to certain mathematical statements and denied to others. Whether a mathematical statement is enlightening or not may be the subject of discussion among mathematicians. Every teacher of mathematics knows that students will not learn by merely grasping the formal truth of a statement. Students must be given some enlightenment as to the *sense* of the statement or they will quit. Enlightenment is a quality of mathematical statements that one sometimes gets and sometimes misses, like truth. A mathematical theorem may be enlightening or not, just as it may be true or false.

If the statements of mathematics were formally true but in no way enlightening, mathematics would be a curious game played by weird people. Enlightenment is what keeps the mathematical enterprise alive and what gives mathematics a high standing among scientific disciplines.

Mathematics seldom explicitly acknowledges the phenomenon of enlightenment for at least two reasons. First, unlike truth, enlightenment is not easily formalized. Second, enlightenment admits degrees: some statements are more enlightening than others. Mathematicians dislike concepts admitting degrees and will go to any length to deny the logical role of any such concept. Mathematical beauty is the expression mathematicians have invented in order to admit obliquely the phenomenon of enlightenment while avoiding acknowledgment of the fuzziness of this phenomenon. They say that a theorem is beautiful when they mean to say that the theorem is enlightening. We acknowledge a theorem's beauty when we see how the theorem "fits" in is place, how is sheds light around itself, like *Lichtung*—a clearing in the woods. We say that a proof is beautiful when it gives away the secret of the theorem, when it leads us to perceive the inevitability of the statement being proved. The term "mathematical beauty," together with the lightbulb mistake, is a trick mathematicians have devised to avoid facing up to the messy phenomenon of enlightenment. The comfortable one-shot idea of mathematical beauty saves us from having to deal with a concept that comes in degrees. Talk of mathematical beauty is a cop-out to avoid confronting enlightenment, a cop-out intended to keep our description of mathematics as close as possible to the description of a mechanism. This cop-out is one step in a cherished activity of mathematicians, that of building a perfect world immune to the messiness of the ordinary world, a world where what we think *should* be true turns out to *be* true, a world that is free from the disappointments, ambiguities, and failures of that other world in which we live.

Notes

1. G. H. Hardy, *A Mathematician's Apology* (Cambridge: Cambridge University Press, 1967).

2. D. Hilbert, *Die Grundlagen der Geometrie*, 7th ed. (Leipzig: B. G. Teubner, 1930).

3. E. Artin, *Galois Theory*, Notre Dame Mathematical Expositions (Notre Dame, Ind.: University of Notre Dame, 1941).

4. D. Hilbert, "Die Theorie der algebraischen Zahlkörper," *Jahresbericht der Deutschen Mathematikvereinigung* 4 (1897): 175–546.

5. H. Weber, *Lehrbuch der Algebra*, 3 vols. (Braunschweig: Vieweg, 1895–1896).

Editor's note: Gian-Carlo Rota was born on 27 April 1932, to a prominent family in Vigevano, Italy. He died at the age of 66 on 19 April 1999. Dr. Rota was the only MIT faculty member ever to hold the title of professor of applied mathematics and philosophy. His uncle by marriage, Ennio Flaiano, wrote scripts for Federico Fellini's films, including *La Dolce Vita*. The wife of Flaiano, Rosetta, was a mathematician at the University of Rome. Flaiano was my godfather, and he wrote several scripts for the films of my father, Luciano Emmer.

The paper "The Phenomenology of Mathematical Beauty" has been published as chapter 10 in Rota's book *Indiscrete Thoughts*, edited by Fabrizio Palombi (Basel: Birkhäuser Verlag AG, 1997), 121–133. The editor wants to thank Birkhäuser Verlag for permission to republish it.