The Language of Creation

Why Physics Generates Mathematics

William de France

Clementinium Editions

Clem

The Language of Creation

Why Physics Generates Mathematics

William de France

Clementinium Editions

ESSAI Collection – REV260520 – 10pt-print

Clementinium Editions
`https://clementinium.com`

Typeset with LaTeX, in EB Garamond.
Format: 105 mm × 170 mm (ESSAI Collection, 10 pt print).

"Then what I cannot forget, cannot communicate happened. It was the union with the divinity, with the universe (I do not know if these words differ)."

Jorge Luis Borges,
The Writing of the God.

Contents

Foreword

This book has a literary ancestor. He is neither a physicist nor a mathematician but a writer who lived from 1899 to 1986 and whose name appears on the epigraph that opens the volume.

Jorge Luis Borges returns, across his sixty years of work, again and again to a single idea. The idea is that the structure of the world might be inscribed, in its entirety, in a single object that one could read, if only one had learned how. In The Writing of the God, *the Mayan priest Tzinacán spends his years in prison searching for a sentence his god encoded in the world at the dawn of creation; he finds it at last in the pattern of spots on the skin of a jaguar that has been pacing in his cell from the beginning. In* The Aleph, *a man descends into the basement of an ordinary house in Buenos Aires and contemplates a small sphere in which the entire universe, all of it at once, is present, visible, and given. In* The Library of Babel, *every possible book exists, distributed across the infinite shelves of a building whose arrangement no librarian has charted but whose contents include, by construction, the answer to every question that has been asked or that will ever be asked. In* The Immortal, *a city composes itself, stone by stone, in the mind of the traveller who reads it. The same intuition runs through all four. There is a structure. It is everywhere. The work of being conscious is the work of learning to see it.*

Borges treats this intuition as literature. He does not propose it as a thesis about how the world actually is; he presents it as the kind of vision that a mystic, a librarian, a priest, or a translator might be granted, and that no scientific discipline could either confirm or refute. The vision is, in his pages, what literature is for: it shows the reader something that argument cannot.

The central idea of this book arose in one of those late readings of

Borges in which the borders between disciplines look, for a moment, artificial. The writer had been turning over a fact that had been present for some time: the mathematics generated by recent physics has converged, in unexpected places, with mathematics that mathematicians had developed for entirely internal reasons. The fact was waiting, like the spots on the jaguar's skin, to be read in a particular way. The way in which it asked to be read was Borges's way. What if the convergence of the two disciplines were not a coincidence to be explained away, but a sign of something neither discipline was equipped, on its own, to acknowledge? What if mathematics and physics were two methods of doing what Tzinacán was doing in his stone cell?

The book that follows is the long argument that this question deserves. It treats Borges's recurring vision not as a literary device but as a candidate description of the actual relation between mathematical and physical knowledge. It does not pretend that this is what Borges himself would have said. He was a writer, not a philosopher of science; the move from his stories to a scientific thesis is mine, not his. But the move is not arbitrary. It rests on a body of work in mathematics and physics, accumulated over the last sixty years, whose pattern Borges's stories anticipate with disquieting accuracy.

A word about what this book is not. It is not a popularisation of string theory; it does not aim to teach the reader how to do physics or mathematics. It is not a refutation of the various positions philosophers of science take on the relation of mathematics to nature; it engages those positions but does not pretend to settle them. It is not, finally, a defence of Borges or a meditation on his fiction; the writer of these pages has read Borges with great pleasure and great profit, but the book belongs to the territory in which mathematicians and physicists work, and the questions it asks belong to the philosophy of science. Borges is the figure at the door who lets the reader in. Once one is inside, the work is one's own.

A word, finally, for the reader who may have encountered the present author before. The earlier book Langlands: The Secret Unity of Mathematics *(2024) treats, at much greater length than the tenth chapter of the present volume can afford, the mathematical unification programme that is one of the three stories of Part III. The tenth chapter of the present book is self-contained; it presupposes nothing of the earlier volume, and a reader new to Langlands will find in it everything they need. A reader who knows the earlier book will recognise the condensation of that material and may wish to revisit it for the fuller treatment. The two books form a diptych, written in different keys. The earlier one follows a single programme through six centuries of mathematics; the present one uses that programme, and two others alongside it, to argue a thesis about the relation between mathematics and physics. Neither presupposes the other. The reader may take them in either order, or alone.*

The reader who turns the page now will find a prologue set in Kyoto in August 1990, on a morning when a physicist received the highest honour of the mathematicians. The prologue does not mention Borges. It does not need to. The whole book that follows is one long way of saying what Borges said, in more pages, and without the alibi of fiction.

Prologue

Kyoto, 21 August 1990

Kyoto International Conference Hall, the opening morning of the International Congress of Mathematicians. A few thousand mathematicians have come, this year, from nearly every country. The Congress is held once every four years, and its first formal act is the announcement of the Fields Medal, the most prestigious distinction in the discipline. The medal is given to up to four mathematicians under the age of forty. This morning there will be four.

Three of the names have circulated in advance. Vladimir Drinfeld of Kharkov, for his work on quantum groups. Vaughan Jones of Berkeley, for his polynomial invariant of knots. Shigefumi Mori of Nagoya, for his classification of algebraic threefolds. Mathematicians in the usual sense: they publish in mathematical journals, they cite their predecessors in mathematical articles, they were trained in mathematics departments. The fourth name is Edward Witten. He is thirty-eight. He holds a professorship at the Institute for Advanced Study in Princeton. He writes for Nuclear Physics B *and* Communications in Mathematical Physics. *He was trained under a theoretical physicist. He is, at this moment, the only theoretical physicist ever to have been awarded the Fields Medal. He has remained the only one since.*

The official report on Witten's work is the work of Michael Atiyah. In the decade before 1990, Atiyah has done more than any other mathematician to open the boundary between mathematics and theoretical physics, and the task of reporting, to his profession and on its behalf, on the reasons for a departure from its usage falls to him. In his absence, the address to the Congress is delivered by Ludwig Faddeev, who takes up its argument. Atiyah's written report, *circulated to the Congress members, describes Witten's command of modern mathematics as rivalled by few mathematicians, and notes that he has, over the preceding years, time and again surprised the mathematical community by the application of physical insight to problems the mathematicians had been unable to solve on their own.*

The report takes the reader through the work. A supersymmetric proof of the index theorem, discovered by treating the quantum mechanics of a particle on a Riemannian manifold. A derivation of the Jones polynomial from a three-dimensional Chern-Simons gauge theory, yielding along the way a family of invariants of three-manifolds that mathematicians had not known to define. A topological quantum field theory that recovers, and in some cases goes beyond, the Donaldson invariants of four-dimensional manifolds. In each case the logical path is the same. An argument borrowed from theoretical physics produces a mathematical result that the mathematicians did not have, and that, in some cases, their own methods could not have reached.

In his report, Atiyah chooses his words carefully. He does not call Witten's arguments proofs in the usual sense. He notes that they depend on notions, the functional integral above all, whose rigorous mathematical foundation has not been established. But he is clear on what they have accomplished. The theorems are true. Mathematicians have verified them by their own methods, one after another, in the years since they were first announced.

What Atiyah is describing is not applied mathematics, where established mathematical tools are brought to bear on problems posed by physics. Nor is it mathematical physics in the older sense, where one gives rigorous foundations to theories that physicists use informally. It is something else. It is the arrival of new mathematical theorems, never proved before, obtained through a physical reasoning whose empirical content is unknown or untestable. Quantum field theory, the framework Witten works in, has not been given a full mathematical foundation. String theory, from which several of his results come, has no experimental confirmation at the energies where *its predictions matter. And yet, starting from these incompletely founded and experimentally unverified frameworks, Witten and others have in the past decade produced theorems that the mathematicians must now prove on their own terms, because there is no room left for doubt that the theorems are true.*

When Witten accepts the medal he does not make the case for his own importance. He thanks his colleagues, his students, his institution. When reporters ask him afterwards whether he feels, today, like a mathematician, he answers that he feels as he has always felt: like a physicist. He works on physics. The mathematics that comes out of his work is, to him, a consequence, not an aim.

The ceremony ends. Over the week of the Congress that follows, in the lecture rooms and the cafés of Kyoto, the mathematicians discuss what has happened. Some are delighted. Some are troubled. Many are simply curious. What is the relation, they ask one another, between a theory of nature that cannot yet be tested and a body of mathematical results that can be verified independently of nature? Is what happened this morning a happy accident, a one-off that will not repeat? Or is there something to be learned from it about what mathematics and physics actually are?

This book is written out of an answer to the second question. The award of August 1990 was not an isolated event. It was one visible moment in a longer process that had been gathering for three decades, and that has continued, accelerating, in the decades since. Over most of the second half of the twentieth century, and continuing today, theoretical physics has been generating new mathematics, faster than the mathematicians alone could have produced it. The direction of influence, for the first time since Newton and Euler, is running the other way. And this inversion, when one considers it long enough, begins to look less like a historical anomaly than like an indication, discreet but insistent, about what physics and mathematics are, and why they have the relation they have.

The chapters that follow tell what had to happen for that morning in Kyoto to become possible, and what has come of it since. The last *of them attempts an answer to the question the audience asked one another in the cafés that week.*

Part I

The Question

1

The Miracle of Wigner

Eugene Wigner opens the most famous philosophical essay of twentieth-century physics with a small story. Two former classmates, high-school friends who have lost touch, meet again years later. One has become a statistician and is working on population trends. He takes out, to show his friend, a reprint of his latest paper. The reprint begins, as such papers do, with the Gaussian distribution. The statistician patiently explains what each symbol means: the actual population, the mean, the spread. The friend, who has no mathematical training, listens politely at first and then, increasingly incredulous, interrupts. "And that one?" "Oh," says the statistician, "that is π." "What is that?" "The ratio of the circumference of a circle to its diameter." The friend is silent for a moment, then speaks. "Surely you are pulling my leg. The population has nothing to do with the circumference of a circle."

The friend is right. The population has, in fact, nothing whatever to do with the circumference of a circle. And yet π appears in the formula, as it must, because the distribution of values in a large collection of random samples is governed by the curve of a bell, and that curve, under analysis, cannot be written without referring to the number that was first introduced as the ratio of two geometric quantities of an entirely different object. The friend's puzzlement is the correct puzzlement. The statistician, who has learned not to puzzle, has lost sight of something.

This story opens the 1960 essay *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*. Its author was not a philosopher. Eugene Paul Wigner was a theoretical physicist of the first rank, born in Budapest in 1902, one of the cohort of Hungarian émigrés (Leo Szilárd, Edward Teller, John von Neumann, and Wigner himself) who left Europe in the 1930s and remade American physics over the decades that followed. Wigner had done his most important work in the 1920s and 1930s, showing how the abstract theory of group representations, developed by Frobenius and Schur for purely algebraic reasons, governed the behaviour of atoms and molecules under rotation and reflection. For that work he would receive the Nobel Prize in Physics in 1963, three years after the essay was published. He was, when he gave the lecture, fifty-six years old, teaching at Princeton, and revered as one of the founders of modern theoretical physics.

The essay began as a lecture. On May 11, 1959, Wigner was invited to give the Richard Courant Lecture in Mathematical Sciences at New York University. The lecture was, and still is, a distinguished annual event, in which a physicist or mathematician of the first rank is asked to reflect, before an audience that includes both disciplines, on some matter of common interest. Wigner chose, rather than present a new technical result, to raise a philosophical question. The choice was not universally well received at the time; some of his colleagues in the audience thought the problem too vague to discuss seriously. In the decades that followed, however, his talk became the single most cited essay in the philosophy of physics, and its title entered common usage. The word "unreasonable" attached itself to the question, and has not left it.

Wigner's question has two parts. It is first the question of why mathematics applies to nature at all. It is second, and more sharply, the question of why it applies with such precision. The second question is harder. That nature should admit some kind of description, in some formal language, is not in itself surprising; any regularity whatever would count as such. What is strange is that the formal language which describes nature should be the same formal language that mathematicians, working in their studies and with no interest in nature, have developed for reasons internal to their own discipline. The mathematician pursues a structure for its own sake, because it seems fertile or beautiful or inevitable, and discovers, a generation later or a century later, that it is precisely the structure the physicist was about to need.

The cases Wigner cites have become, in the years since his essay, the classic examples of the phenomenon. Isaac Newton's law of universal gravitation, written down in the 1680s on the basis of terrestrial mechanics and a few astronomical observations, predicts the motions of the planets to a precision its author could scarcely have imagined. Within a century, the planet Neptune is found by calculation before it is observed. Le Verrier, in 1846, computes the position of a perturbation in the orbit of Uranus and tells the astronomer Johann Galle, in Berlin, where to point his telescope; and there it is. Within three centuries, the Global Positioning System applies the same law, with small corrections from general relativity, to locate a vehicle to within a few metres on the surface of the Earth. The mathematical structure has not changed. The formula is the one Newton wrote.

James Clerk Maxwell's equations of electromagnetism, formulated in the 1860s to unify what had been known separately about electric and magnetic phenomena, turn out to imply the existence of a new kind of wave, propagating through the vacuum at a speed that can be calculated from two parameters measured in laboratory experiments on charges and currents. The speed, to the precision of the measurements then available, comes out equal to the speed of light. The implication is staggering, and Maxwell draws it: light is an electromagnetic wave. Heinrich Hertz, in 1887, confirms this by producing in his laboratory waves at radio frequencies (invisible, inaudible, entirely new) and detecting them across the room. Within a decade Marconi is broadcasting across the English Channel. Within a century every one of the communication systems on which modern life depends, from radio and television to radar and cellular telephony and optical fibre, rests on Maxwell's equations. Maxwell had not been looking for radio waves. He was organising equations.

Paul Dirac's equation, written down in 1928 as an attempt to reconcile the Schrödinger equation with the principles of special relativity, has a mathematical property that, at first glance, looks like a disaster. Its solutions include states with negative energy. Negative energy corresponds to no observed particle, and a physical theory that predicts unobservable things is usually wrong. Dirac, twenty-five years old and working at St John's College, Cambridge, declines to throw the solutions away. He proposes, after some reflection, that they describe a new kind of particle: the antiparticle of the electron, identical in mass but opposite in charge. In 1932, Carl Anderson, examining tracks of cosmic rays in a cloud chamber at Caltech, finds a particle that curves in a magnetic field as an electron would but in the wrong direction. He calls it the positron. The antimatter that Dirac's mathematics had implied, and that no experiment before 1928 had so much as hinted at, is real. Dirac receives the Nobel Prize the following year, at the age of thirty-one.

More recently, in 1964, three theorists working independently (François Englert and Robert Brout in Brussels, and Peter Higgs in Edinburgh) noted that the mathematical structure of what would become the Standard Model of particle physics required a scalar field with certain properties, the quanta of which would appear as a new particle of definite character. The Standard Model, without this field, is mathematically inconsistent; with it, it is consistent and predictive. For forty-eight years the particle was searched for in increasingly powerful accelerators. In July 2012, at CERN, two experiments independently announced its detection, with a mass of 125 gigaelectronvolts and properties matching the theoretical prediction to the precision of the measurements. A consistency argument made on paper had forced the existence, in the real world, of a particle that the world then produced.

These are the cases Wigner's essay assembles, to which the Higgs discovery added, half a century later, one more instance. They share a pattern. In each case, a mathematical structure, developed or formulated for reasons internal to either mathematics or theoretical physics, turns out to describe nature not approximately but exactly, not qualitatively but quantitatively, and not as a convenient fiction but as a predictive instrument that permits the discovery of entities the theorist could not otherwise have found. The pattern admits no trivial explanation. That nature should be describable is one thing; that it should be describable by structures developed for their own sake, and describable to eight decimal places, is another.

The paradigm case, the case in which the mystery is sharpest, is the case of complex numbers. The symbol $\sqrt{-1}$ was introduced in sixteenth-century Italy to resolve an awkwardness in the formula for the roots of a cubic equation. Girolamo Cardano, physician, mathematician, astrologer, gambler, published in 1545 the solution of the general cubic in his *Ars Magna*. His formula gave, for the equation $x^3 = 15x + 4$, a root that required the square root of a negative number: the square root of minus 121. Any student of arithmetic knew that the square of a real number was always positive, and therefore that the square root of a negative number was impossible. And yet the formula was correct: if one handled the impossible quantities according to the ordinary rules of algebra, and carried them through the calculation, one obtained at the end the perfectly real root $x = 4$. Cardano noted the oddity, described the rules for working with such fictitious quantities, and moved on.

Rafael Bombelli, in his *L'Algebra* of 1572, gave a more careful treatment. He wrote out the rules of arithmetic for what he called *piu di meno* ("more of minus") quantities. He showed that one could add, subtract, multiply, and divide them coherently, and that calculations involving them gave consistent results. The square roots of negatives became, for Bombelli, a family of formal objects with their own arithmetic. They were called imaginary numbers. They remained, for two more centuries, a calculational device of uncertain status. Mathematicians used them when they had to, apologised for them afterwards, and avoided them when they could.

In the early nineteenth century Carl Friedrich Gauss rehabilitated them. He gave them a geometric interpretation: a complex number $a + bi$ corresponds to a point in a plane, with a the horizontal coordinate and b the vertical. Addition corresponds to translation. Multiplication by i corresponds to rotation through a right angle. Multiplication by $a + bi$ corresponds, in general, to a rotation combined with a scaling. Under this reading, the impossibility of imaginary numbers dissolves. They are no more imaginary than arrows drawn on a sheet of paper. Gauss used complex numbers in his proof of the fundamental theorem of algebra, in his work on quadratic forms, in his theory of modular forms. By the end of the nineteenth century, complex numbers were indispensable to the theory of analytic functions, the theory of differential equations, the theory of algebraic curves. They had become a mathematical object of the first importance. Nothing in physics, as it was taught in 1900, required one to use them.

This changed in 1926, when Erwin Schrödinger wrote down the equation that bears his name. The equation describes the time evolution of the wave function of a quantum system. It involves, on the left-hand side, the time derivative of the wave function, multiplied by a factor of $i\hbar$: Planck's reduced constant, multiplied by the imaginary unit. The presence of i is not cosmetic. The Schrödinger equation, with a real coefficient in place of i, would not describe quantum mechanics; it would describe a diffusion process, which is something entirely different. It would describe, in rough terms, the spreading of a drop of ink in water. The whole phenomenon of quantum interference, with its fringes and its correlations and its probabilistic structure, depends on the fact that the wave function takes complex values. The imaginary numbers Cardano tolerated as a computational trick, and Bombelli codified as a formal arithmetic, had become, by 1926, the indispensable language of the matter from which we and everything else are made.

There is no way out of this dependence. It is not the case that quantum mechanics can be formulated in several mathematically equivalent ways, of which the complex-number formulation happens to be convenient. It is the case that quantum mechanics requires complex numbers, and that physical attempts to build an equivalent theory without them run into inconsistencies. A formal device noted by Italian algebraists in 1545, because the formula for the cubic demanded it, turned out, nearly four centuries later, to be the only language in which the fundamental theory of matter could be written down. The interval between Cardano's *Ars Magna* and Schrödinger's equation is about the interval between the first flintlock musket and the invention of the aeroplane; it is long. During that interval, nothing in the physical world hinted that complex numbers would play a fundamental role. Mathematicians developed them for their own reasons.

Wigner, discussing this case, writes that it is one of the most unsettling features of modern physics. He does not offer an explanation. He writes, in a sentence that has been quoted so often since that it has become a motto of the philosophy of physics, that "the miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve. We should be grateful for it, and hope that it will remain valid in future research." The tone is characteristic of Wigner: formal, slightly archaic, refusing both the optimism of full rationalist explanation and the pessimism of mystical acceptance. The miracle is noted. It is not explained.

The terms in which Wigner poses the question are worth attending to. Mathematics, in his framing, is the active term. It is developed, by human beings, for reasons internal to the human mind and to the sense of formal beauty that mathematicians cultivate. Its connection to nature is, from this vantage, a happy accident that ought, in principle, to fail; that it keeps not failing is the miracle. The direction of the arrow is mathematics toward nature. The world is something onto which mathematics is cast, with a fit that is surprising given that nothing connected the two domains in advance. This framing has been so influential that it has become difficult to see it as a framing. But it is one.

An alternative deserves consideration from the outset: one in which the arrow runs the other way. In that alternative, it is not mathematics that, having been developed, turns out to describe the world. It is the world, or more precisely the structure the world has and that the physicist probes, that generates the mathematics we use to describe it. The mathematicians' interior pursuits are not arbitrary. They are responding, without knowing it, to a pressure that also generates the physics. The mathematics and the physics are not two independent products that happen to match. They are two distillations of the same substance. From this vantage, Wigner's miracle becomes something different. It becomes less of a miracle and more of an indication of what mathematics, and physics, are for.

The case for this inversion is assembled over the rest of this book. The chapter that follows describes the reversal in its most visible modern form. Since the 1970s, and with increasing force since the 1990s, theoretical physics has been producing mathematical results of the first importance: results that mathematicians now accept as theorems, but that no mathematician, working by the tools of mathematics alone, could or did prove. The 1990 Fields Medal ceremony in Kyoto, with which this book opened, was one public acknowledgment of this process. The rest of the book will examine what it means.

For the moment, one observation suffices. Wigner asked the right question. He asked it in the wrong direction.

2

The Reversal

Eugene Wigner's essay, when it appeared in 1960, described a one-way movement. Mathematicians developed structures for their own reasons; physicists, later, found themselves obliged to use those structures to describe nature. The movement went from mathematics to physics. The essay treated that direction as the whole of the phenomenon to be explained. For most of the period between 1600 and 1960, the essay was right.

That period ended around the time Wigner's essay was published. In the decades that followed, a second movement became visible, running in the opposite direction. Physicists, working on theories whose experimental verification remained out of reach, produced mathematical results of the first importance. Not applications of mathematics they already possessed, but new mathematics: theorems that had not been stated, let alone proved, by the mathematicians who would turn out to specialise in the areas concerned. Those theorems, once stated by the physicists, had to be verified by the mathematicians according to their own standards of proof. In every case in which the verification was carried out, it confirmed the physicists' result.

The phenomenon was eventually acknowledged by the mathematical community itself, though not without delay. In his 2002 retrospective on twentieth-century mathematics, published in the *Bulletin of the London Mathematical Society*, Michael Atiyah (the same Atiyah who had given the laudatio for Witten at Kyoto twelve years earlier) wrote that the most striking development of the last quarter of the century had been the influx of new mathematics from theoretical physics. He devoted a full section of the essay to it, under the heading "Impact of Physics." He noted, in the careful language appropriate to an author who knew he would be read by mathematicians of several generations, that this influence was without precedent since the eighteenth century, and that it had arrived from theories whose empirical status was either unknown or disputed. The conclusion he drew was measured. It was also unambiguous. The boundary between mathematics and theoretical physics, he wrote, had become porous in a way that it had not been within living memory, and the most interesting mathematics of the generation just then retiring had flowed across that boundary in a direction no one had anticipated.

The label "Physical Mathematics," which has since attached itself to the field Atiyah described, was proposed in 2014 by Gregory Moore, a theoretical physicist at Rutgers, in a vision paper written on the occasion of his election to the American Academy of Arts and Sciences. Moore's point was that neither of the existing labels fit what was happening. "Applied mathematics" named what happens when already-developed mathematical tools are brought to bear on physical problems. "Mathematical physics" named what happens when mathematicians give rigorous foundations to theories that physicists use informally. Neither label described what Witten, Vafa, Segal, Strominger, Seiberg, and many others were doing. What they were doing was this: using physical reasoning, on theories not yet experimentally confirmed, to prove or conjecture mathematical statements that mathematicians then recognised as theorems of their own discipline. Moore called the activity Physical Mathematics. The term has stuck.

The clearest single illustration of the phenomenon, and the one most often cited to convey its strangeness to outsiders, is the 1991 prediction of Philip Candelas and his collaborators Xenia de la Ossa, Paul Green, and Linda Parkes. Candelas, a physicist then at the University of Texas at Austin, was working in string theory on a problem in the geometry of what are called Calabi-Yau manifolds: six-dimensional shapes used in the theory to accommodate the extra spatial dimensions beyond the three we observe. The simplest Calabi-Yau manifold of a certain standard kind is the quintic threefold, a three-complex-dimensional variety defined by a single polynomial equation of degree five in four-complex-dimensional space. On this variety, an old question can be posed. How many rational curves of a given degree lie on it? Mathematicians call this a question of enumerative geometry. It had been a prized problem in algebraic geometry since the late nineteenth century.

The answer for degree one, known since Schubert in the 1870s, is 2875. The answer for degree two, obtained by Sheldon Katz in 1986, is 609250. The answer for degree three was a major open problem at the time of Candelas's work. It was expected, on the basis of general structural considerations, to have a large integer as its answer, but no one knew what integer. An independent calculation that had been under way at Oslo, by Geir Ellingsrud and Stein Arild Strømme, using the most powerful modern techniques of algebraic geometry, had produced in 1991 a number running to nine digits, which after verification was publicly cited as the presumed answer. For degree four and above, the problem was believed to be beyond reach.

Candelas and his coauthors, working from string theory, took a different route. They used a hypothesis that had emerged in the late 1980s from the physics of superstring compactifications, and that physicists called *mirror symmetry*: the hypothesis that every Calabi-Yau manifold has a "mirror" partner, another six-dimensional manifold, on which the string theory is equivalent to the one on the first but with certain kinds of quantities exchanged. The exchange has the property that questions hard to solve on one side become easy on the other. Counting rational curves on a quintic is very hard. Computing a certain integral on the mirror of the quintic, as it turns out, is not. Candelas and his coauthors computed the integral on the mirror. They then applied the dictionary of mirror symmetry to translate the result back.

And they obtained, for every positive integer degree, a prediction for the number of rational curves of that degree on the original quintic.

Their number for degree one was 2875. Agreement. Their number for degree two was 609250. Agreement. Their number for degree three was 317206375. At the time, the mathematicians at Oslo had an independent calculation in progress for degree three, and they had arrived at a different, nine-digit answer. It turned out, after the Candelas prediction was circulated, that the mathematicians' calculation contained a programming error. The corrected mathematical calculation reproduced 317206375. Within five years, Alexander Givental, working at Berkeley, proved rigorously that the physicists' formula gave the correct answer for all degrees. Bong Lian, Kefeng Liu, and Shing-Tung Yau reached the same conclusion shortly after by a different route. The mathematical community had, by then, accepted that a physics argument about six-dimensional spaces, arising from a theory that had not been and still has not been tested at the energies where its predictions matter, had answered a century-old question in algebraic geometry that mathematicians had been unable to answer on their own.

The Candelas prediction is not isolated. A comparable story can be told about Simon Donaldson's invariants of four-dimensional manifolds, introduced in 1982 and then illuminated by Witten's reformulation as a topological quantum field theory, and transformed again by the Seiberg-Witten equations of 1994, which simplified the computation of the invariants to a degree no one had imagined possible; the Jones polynomial of knots, which the mathematician Vaughan Jones had discovered in 1984 by ingenious but ad hoc algebraic means, and which Witten in 1988 showed to be the natural output of a three-dimensional quantum field theory of a particular simple type, belongs in the same chapter, since both reformulations were products of the same year. A further story can be told about the geometric Langlands programme, in which the deepest conjectures of a field of pure mathematics turned out, in work by Anton Kapustin and Edward Witten published in 2006, to correspond exactly to a symmetry of a four-dimensional gauge theory. These three stories (mirror symmetry, the topological quantum field theories of Donaldson-Jones-Seiberg-Witten, and the gauge-theoretic origin of geometric Langlands) occupy, one each, the three chapters of the third part of this book.

What the stories have in common is the pattern Atiyah identified and Moore named. In each case, physicists working on theories without full experimental confirmation derived mathematical statements that the mathematicians, working strictly within their own discipline, had not derived and could not easily have derived. In each case, when the mathematicians went back to check, the physicists' statements turned out to be correct. In each case, the physicists' arguments, however compelling, did not qualify as mathematical proofs by the mathematicians' standards; they relied on objects like path integrals whose rigorous definition had not yet been established. And in each case, the mathematicians eventually supplied proofs on their own terms, proofs that depended on techniques and ideas often quite different from those the physicists had used. The physicists had seen something the mathematicians had then to see for themselves.

The phenomenon poses an epistemological puzzle, and to some a scandal. Mathematical truths are supposed to be eternal and a priori. They are supposed to be discoverable by sitting at a desk and reasoning carefully, without reference to any contingent feature of the natural world. For most of the history of mathematics, that is how they were in fact discovered. Yet here is a body of theorems whose discovery depended on reasoning about a specific and not experimentally certified theory of the physical universe. If one lived in a universe with different physical laws (if, for instance, string theory were not the approximately correct description of high-energy physics that some physicists take it to be), would the theorems cease to be true? Surely not. The mathematicians who verified the theorems by their own methods did not assume any particular physics. The theorems are true in every possible universe. And yet the path by which they were found ran through a theory of this universe. Somehow, the structure of the physical world turned out to know things about the structure of mathematics that mathematicians, working only with mathematical tools, did not know.

Stated carefully, this is a situation whose philosophical interpretation is not obvious. Mathematical knowledge is either a priori or it is not; the traditional answer has been that it is. If the traditional answer is correct, the Candelas prediction should not have been possible to make by physical reasoning alone. If the Candelas prediction was possible to make that way, and the evidence is that it was, then the traditional answer is at best incomplete. Either mathematics is not quite a priori in the way philosophers have thought, or the physical world is doing something we do not yet have a name for.

The remainder of this book is an attempt to understand what the reversal means. The chapter that follows takes a short historical detour. It argues that the reversal, though it looks unprecedented, is in fact only the current phase of a much older intermittent intimacy between mathematics and physics, whose divorce in the nineteenth century was the exception rather than the rule. Over the long run, what Atiyah described in 2002 is not a novelty. It is a return. It remains, for all that, a situation in need of philosophical explanation, because the terms on which the two disciplines have rejoined in the late twentieth century are not the terms on which they were joined in any previous period. The mathematics that arrives today from physics is different mathematics, in kind as well as in content, from what Newton and Euler produced. It is the mathematics of gauge fields, of Calabi-Yau manifolds, of quantum cohomology, of geometric Langlands. These are domains that no mathematician working before 1970 would have expected physics to teach.

3

Golden Age, Divorce, Remarriage

The phenomenon Atiyah described in 2002 looks unprecedented only if one looks from 1950. Seen from 1700, it is not unprecedented at all. Seen from 1700, it is a return. The separation between mathematics and theoretical physics that seemed, to a mathematician of the 1960s, to be the normal condition of the two disciplines, was in fact a historical anomaly of roughly a century's duration. Before 1850, the disciplines had not been separate; they had not been two disciplines at all. What we have witnessed since the 1970s is the end of the interregnum and the resumption, in a modified form, of the long prior state of affairs.

The prior state of affairs had a name. *Natural Philosophy* was the term used throughout Europe, into the early nineteenth century, for the integrated study of the mathematical structures governing the natural world. The title page of Newton's 1687 masterwork reads *Philosophiæ Naturalis Principia Mathematica*: *Mathematical Principles of Natural Philosophy*. The book is simultaneously the foundation of classical mechanics and a turning point in the development of the calculus; it treats motion, gravitation, fluids, tides, and the orbits of comets, and it does so by developing mathematical techniques in geometric dress that would remain the working language of physicists and mathematicians alike for a century and a half. Newton did not distinguish between his mathematical results and his physical results. Neither did anyone else; the distinction would not have made sense to him.

The list of names who operated within this framework is, from a twentieth-century vantage, startling. Leonhard Euler, the most prolific mathematician in history, devoted as much of his career to problems of continuum mechanics, optics, and astronomy as to the number theory and analysis for which he is now best remembered in mathematics courses. The calculus of variations, which Euler and Joseph-Louis Lagrange developed through the eighteenth century, was simultaneously a mathematical theory of extremal problems and a physical principle governing the motion of all classical systems: the Euler-Lagrange equations are the laws of mechanics restated in the language of functionals. Lagrange, whose *Mécanique Analytique* of 1788 reformulated Newton's mechanics without a single diagram, considered himself, according to his own testimony, as much a mathematician as a physicist, and did not think his reformulation belonged to one discipline more than to the other. Pierre-Simon Laplace worked indifferently on celestial mechanics, on probability theory, and on partial differential equations; his *Traité de Mécanique Céleste* and his *Théorie Analytique des Probabilités* are companion volumes written by the same mind using the same tools. Carl Friedrich Gauss, in the first decades of the nineteenth century, conducted geodetic surveys of the Kingdom of Hanover, developed the theory of electromagnetism with Wilhelm Weber, proved the fundamental theorem of algebra, and wrote the *Disquisitiones Arithmeticae*, which remains a founding text of modern number theory, without seeing these activities as different in nature. He was, for himself and for his contemporaries, practising a single unified discipline.

The unity lasted roughly from the publication of the *Principia* to around 1850. It was not perfect, and it had tensions; there had always been mathematicians whose interests were more abstract than physical, and physicists whose interests were more empirical than theoretical. But the institutional and conceptual machinery treated the two activities as continuous. A single trained scientist could produce original results in both. The academies and journals received both kinds of work without distinguishing them.

The separation began, for reasons that were in part conceptual and in part sociological, in the middle of the nineteenth century. Several currents combined. The first was the internal pressure for rigour inside mathematics itself. Bernard Bolzano in 1817, Augustin-Louis Cauchy in the 1820s, and Karl Weierstrass in the 1850s and 1860s each in turn pushed mathematicians to stop relying on the geometric and physical intuitions that had been good enough for Euler. The notion of a limit had to be given a definition that did not appeal to the motion of a point along a curve. The notion of a continuous function had to be given a definition that did not appeal to the tracing of a graph with a pencil. By the time Weierstrass was lecturing at Berlin, the standards of what counted as a mathematical argument had changed. Arguments acceptable to Newton and Euler no longer were.

The second current was the appearance of mathematical objects that had no evident physical interpretation. The non-Euclidean geometries discovered independently by Nikolai Lobachevsky in the 1820s and János Bolyai in the 1830s described spaces in which Euclid's parallel postulate failed, and for which the physicists of the time saw no application. Riemann's 1854 habilitation lecture, on the hypotheses that lie at the foundation of geometry, generalised geometry to spaces of any dimension and any curvature, treating geometry as a mathematical subject whose physical relevance was an open question to be decided, if at all, by some future experiment. Évariste Galois's theory of equations, worked out in the nights before his duel in May 1832 and published only after his death, concerned permutations of roots and had no evident connection to any physical problem. Mathematics was beginning to generate objects that mathematicians valued for their own sake, and whose physical significance, if any, would have to be discovered later, or might never be discovered at all.

The third current was the increasing experimentalism of physics. The laboratory physics of the nineteenth century, descending from Faraday, Oersted, Joule, and Maxwell on the electromagnetic side, and from Carnot, Clausius, and Kelvin on the thermodynamic side, required skills and dispositions that the geometer at his desk did not cultivate. The physicist had to handle galvanometers, thermocouples, cathode-ray tubes, heat engines. His working relation to mathematics was that of a consumer of ready-made tools. A physicist in 1890 would use differential equations, but he did not typically think of himself as generating new mathematics; the job of a physicist was to discover the laws of nature, not to advance the abstract structure of the calculus.

These three currents, over the course of half a century, produced a separation that by 1900 was institutional. Universities had split their mathematical and physical chairs. Journals catered to one audience or the other. A mathematician who attended a physics colloquium, or the reverse, did so as a visitor. The separation would reach its philosophical extreme in the work of Nicolas Bourbaki, the collective pseudonym adopted in 1935 by a group of young French mathematicians, most of them graduates of the École Normale Supérieure, who set themselves the task of reconstructing mathematics on foundations purified of every extraneous reference. Bourbaki's *Éléments de Mathématique* would, over the decades that followed, present set theory, algebra, topology, integration, and Lie theory with only rare and local recourse to physical or geometric motivation (the chapters on Lie groups being the most visible exception). The programme, if not quite as austere as its caricature holds, was emphatically self-contained. It was also extraordinarily influential; its pedagogical descendants dominated mathematical instruction, in France and well beyond, for a generation.

By the middle of the twentieth century, the separation was such that a working mathematician and a working theoretical physicist could pass their professional lives in the same building without encountering each other's ideas. The language had diverged. The problems had diverged. A paper published in *Annals of Mathematics* and a paper published in the *Physical Review* inhabited intellectual worlds whose commerce had become attenuated to the point of being occasional.

The condition was noticed, and lamented, by a figure uniquely placed to see both sides. Freeman Dyson, English-born, trained as a mathematician at Cambridge, had spent his adult career as a theoretical physicist at the Institute for Advanced Study in Princeton. In 1972 he was invited to give the Josiah Willard Gibbs Lecture, an annual address to the American Mathematical Society. He took the occasion to speak, under the title "Missed Opportunities," about exactly the separation described above. The lecture has become, in the decades since, one of the most cited texts in the history of twentieth-century mathematical physics. Its central sentence is this: "The marriage between mathematics and physics, which was so enormously fruitful in past centuries, has recently ended in divorce." Dyson did not mean the sentence rhetorically. He went on to list several missed opportunities, instances where, in his view, mathematicians had failed to notice or pursue problems that theoretical physicists had brought to their attention. The Yang-Mills equations, the Dirac equation, quantum field theory itself: each, according to Dyson, contained mathematical structure of the first importance that the mathematicians were ignoring because it had not reached them in a purely mathematical form.

The 1972 lecture was prophetic, and it was prophetic in a way that Dyson himself did not quite anticipate. In the decade that followed its delivery, the trend began to reverse. The turn had overlapping origins. In 1978, Michael Atiyah, Vladimir Drinfeld, Nigel Hitchin, and Yuri Manin constructed all the finite-action solutions of the Yang-Mills equations on four-dimensional Euclidean space, resolving a major problem of gauge theory by methods of algebraic geometry; the resulting construction was recognised immediately by physicists and mathematicians alike as belonging fully to both disciplines. In 1982, Simon Donaldson, a student of Atiyah at Oxford, used the moduli space of Yang-Mills instantons to construct invariants of four-dimensional manifolds, and to