Wiles received the Abel Prize from Crown Prince Haakon of Norway.

Editor's Note: *Christopher Skinner kindly accepted our invitation to put together this feature in honor of Andrew Wiles on the occasion of his receiving the 2016 Abel Prize.*

On May 24, 2016, Sir Andrew J. Wiles received the Abel Prize in a ceremony held in the Aula of the University of Oslo in Oslo, Norway. Wiles, who received the prize from H.R.H. Crown Prince Haakon at the award ceremony, was the fourteenth recipient of the 6 million NOK (about 750,000 USD) prize. A prize honoring the Norwegian mathematician Niels Henrik Abel was first proposed by the world-renowned mathematician Sophus Lie, also from Norway, and initially planned for the one-hundredth anniversary of Abel's birth in 1902, but the establishment of the Abel Prize had to wait another hundred years. The Abel Prize is administered by the Norwegian Academy of Science and Letters.

The Abel Prize was awarded to Wiles for "his stunning proof of Fermat's Last Theorem," which opened a new era in number theory. The citation of the Abel Committee, read by committee chair John Rognes on the occasion of the announcement of the 2016 Abel Prize, recounts the early history of Fermat's Last Theorem—the assertion that for any given integer $n \geq 3$, there are no integer solutions to $x^n + y^n = z^n$ with $xyz \neq 0$—and how it was eventually linked to the then-conjectural modularity of semistable elliptic curves and that it was this modularity that Andrew Wiles ultimately established in a proof both surprising and profound. It is especially appropriate that Wiles's groundbreaking work on elliptic curves was recognized by the awarding of the Abel Prize, elliptic curves being the natural domains of the elliptic functions introduced by Abel. As the citation for the 2016 Abel Prize concludes: "Few results have as rich a mathematical history and as dramatic a proof as Fermat's Last Theorem."

The awarding of the Abel Prize was followed by the Abel Lectures on the next day, May 25. In his lecture, Wiles explained how his proof of Fermat's Last Theorem exemplified the movement of number theory from the abelian to the nonabelian. Henri Darmon, in his lecture "Andrew Wiles's Marvelous Proof," described Wiles's work as "a centerpiece of the Langlands program" and explained its transformative impact, and Manjul Bhargava spoke about how Wiles's work has implications for the Birch–Swinnerton-Dyer Conjecture.

Photo Credits

Photos are courtesy of Audun Braastad.

The full text of the citation of the Abel Committee can be found at `www.abelprize.no/c54154/binfil/download.php?tid=67039`. An expanded version of Henri Darmon's Abel Lecture is included in this issue.

DOI: http://dx.doi.org/10.1090/noti1486

Interview with Abel Laureate Sir Andrew J. Wiles

Martin Raussen and Christian Skau

Andrew J. Wiles is the recipient of the 2016 Abel Prize of the Norwegian Academy of Science and Letters. The following interview was conducted by Martin Raussen and Christian Skau in Oslo on May 23, 2016, in conjunction with the Abel Prize celebration. This article originally appeared in the September 2016 issue of the *Newsletter of the European Mathematical Society*—see www.ems-ph.org/journals/newsletter/pdf/2016-09-101.pdf#page=31, pp. 29–38—and is reprinted here with permission of the EMS.

Raussen and Skau: *Professor Wiles, please accept our congratulations for having been selected as the Abel Prize Laureate for 2016. To be honest, the two of us expected this interview to take place several years ago!*

You are famed not only among mathematicians but also among the public at large for (and now we cite the Abel Committee): "the stunning proof of Fermat's Last Theorem, by way of the Modularity Conjecture for elliptic curves, opening a new era in number theory." This proof goes back to 1994, which means that you had to wait for more than twenty years before it earned you the Abel Prize. Nevertheless, you are the youngest Abel Prize Laureate so far. After you finished your proof of Fermat's Last Theorem you had to undergo a deluge of interviews, which makes our task difficult. How on earth are we to come up with questions that you have not answered several times before? Well, we will try our best.

Sir Andrew J. Wiles received the Abel Prize from Crown Prince Haakon of Norway.

Fermat's Last Theorem: A Historical Account

*We have to start at the very beginning, with a citation in Latin: "…*nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere," *which means:* "it is impossible to separate any power higher than the second into two like powers." *In modern mathematical jargon, this can be written: "The equation* $x^n + y^n = z^n$ *has no solution in natural numbers for n greater that two." And then it continues:* "cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet," *which means:* "I have discovered a truly marvellous proof of this, which this margin is too narrow to contain." *This remark was written in the year 1637 by the French lawyer and amateur mathematician Pierre de Fermat [1601–1665] in the margin of his copy of Diophantus'* Arithmetica. *He certainly did not expect that it would keep mathematicians, professionals, and amateurs alike busy for centuries trying to unearth the proof.*

Could you please give us a short account of some of the attempts towards proving Fermat's Last Theorem up until the time you embarked on your successful journey? Furthermore, why was such a simple-minded question so attractive, and why were attempts to prove it so productive in the development of number theory?

Wiles: The first serious attempt to solve it was presumably by Fermat himself. But, unfortunately, we know nothing about it except for what he explained about his proofs in the specific cases of $n = 3$ and $n = 4$.[1] That is, he showed that you can't have the sum of two cubes be another cube or the sum of two fourth powers being a

Martin Raussen is professor of mathematics at Aalborg University, Denmark. His e-mail address is raussen@math.aau.dk.

Christian Skau is professor of mathematics at the Norwegian University of Science and Technology, Trondheim, Norway. His e-mail address is csk@math.ntnu.no.

DOI: http://dx.doi.org/10.1090/noti1484

[1] *Strictly speaking, Euler was the first to spell out a complete proof in the case* $n = 3$.

fourth power. He did this by a beautiful method, which we call infinite descent. It was a new method of proof, or at least a new way of presenting proofs in arithmetic. He explained this method to his colleagues in letters, and he also wrote about it in his famous margin, which was big enough for some of it at least. After the marginal notes were published by Fermat's son after his father's death, it lay dormant for a while. Then it was picked up by Euler [1707–1783] and others who tried to find this truly marvellous proof. And they failed. It became quite dramatic in the mid-nineteenth century—various people thought they could solve it. There was a discussion concerning this in the French Academy: Lamé [1795–1870] claiming he was just about to prove it, and Cauchy [1789–1857] saying he thought he could too, and so on.

In fact, it transpired that the German mathematician Kummer [1810–1893] had already written a paper where he explained that the fundamental problem was what is known now as the fundamental theorem of arithmetic. In our normal number system, any number can be factorized in essentially one way into prime factors. Take a number like 12; it is 2 times 2 times 3. There is no other way of breaking it up. But in trying to solve the Fermat problem, you actually want to use systems of numbers where this uniqueness does not hold. Every attempt that was made to solve the Fermat problem had stalled because of this failure of unique factorization. Kummer analyzed this in incredible detail. He came up with the most beautiful results, and the end product was that he could solve it for many, many cases. For example, for $n \leq 100$ he solved it for all primes except for 37, 59, and 67. But he did not finally solve it. His method was based on the idea that Fermat had introduced—the method of infinite descent—but in these new number systems.

The new number systems he was using spawned algebraic number theory as we see it today. One tries to solve equations in these new systems of numbers instead of solving them with ordinary integers and rational numbers. Attempts in the style of Fermat carried on for a while but somewhat petered out in the twentieth century. No one came up with a fundamentally new idea. In the second half of the twentieth century, number theory moved on and considered other questions. Fermat's problem was all but forgotten by the professionals.

Then, in 1985, Gerhard Frey, a German mathematician, came up with a stunning new idea where he took a hypothetical solution to the Fermat problem and rewrote it so that it made what is called an elliptic curve. And he showed, or suggested, that this elliptic curve had very peculiar properties. He conjectured that you can't really have such an elliptic curve. Building on this a year later, an American mathematician, Kenneth Ribet, demonstrated, using this Frey curve, that any solution of Fermat would contradict another well-known conjecture called the Modularity Conjecture. This conjecture had been proposed in a weak form by Taniyama [1927–1958] and was refined by Shimura, but the first real evidence for it came from André Weil [1906–1998], who made it possible to check this precise form of the Modularity Conjecture in some detail. And a lot of evidence was amassed showing that this should certainly be true. So, at that point, mathematicians could see that: "Yes, Fermat is going to be true. Moreover, there has to be a proof of it." What happened was that the Modularity Conjecture was a problem that mathematics could not just put to one side and go on for another five hundred years. It was a roadblock right in the middle of modern mathematics. It was a very, very central problem. As for Fermat, you could just leave it aside and forget it almost forever. This Modularity Conjecture you could not forget. So, at the point when I heard that Ribet had done this, I knew that this problem could be solved and I was going to try.

It was a roadblock right in the middle of modern mathematics.

Raussen and Skau: *Concerning speculations about Fermat's claimed proof, do you think he had the same idea as Lamé, assuming, wrongly as it turned out, that the cyclotomic integers have unique factorization?*

Wiles: No, I don't think so, though the idea might be in there somewhere. It is very hard to understand. André Weil wrote about this. All the other problems Fermat considered had to do with curves that were of genus zero or genus one. And suddenly he is writing down a curve that has higher genus. How is he going to think about it? When I was trying this myself as a teenager, I put myself in Fermat's frame of mind because there was hardly anything else I could do. I was capable of understanding his mathematics from the seventeenth century but probably not much beyond that. It seemed to me that everything he did came down to something about quadratic forms, and I thought that might be a way of trying to think about it. Of course, I never succeeded, but there is nothing else that suggests Fermat fell into this trap with unique factorization. In fact, from the point of view of quadratic forms, he understood that sometimes there was unique factorization and sometimes there was not. So he understood that difference in his own context. I think it is unlikely that that was the mistake.

Raussen and Skau: *In the same book by André Weil that you referred to, titled* Number Theory: An approach through History from Hammurapi to Legendre*, it is mentioned that Fermat looked at the equation of a cube minus a square equal to* 2 $[x^3 - y^2 = 2]$ *and he showed that it has essentially only one solution, namely* $x = 3$ *and* $y = \pm 5$*. André Weil speculates that Fermat at the time looked at the ring* $Z[\sqrt{-2}]$*, which does have unique factorization.*

Wiles: Yes, he used unique factorization, but the way he did it was in terms of quadratic forms. And I think he also looked at quadratic forms corresponding to $Z[\sqrt{-6}]$ where there is not unique factorization. So I think he understood. It was my impression when I thought about it that he understood the difference.

A Mathematical Education

Raussen and Skau: *You were apparently already interested in mathematical puzzles as quite a young boy. Have you any thoughts about where this interest came from? Were you influenced by anyone in particular?*

Wiles: I just enjoyed mathematics when I was very young. At the age of ten, I was looking through library shelves devoted to mathematics. I would pull out books and at one point I pulled out a book of E. T. Bell [1883–1960] titled *The Last Problem*, which on its cover describes the Fermat equation, the Wolfskehl Prize, and the romantic history of the problem. I was completely captivated by it.

Raussen and Skau: *Were there other things that fascinated you in this book by Eric Temple Bell?*

Wiles: It is entirely about that one equation, really. And it is actually quite wordy. So there is less mathematics in some sense than you might think. I think it was more the equation. Then, when I found this equation, I looked for other elementary books on number theory and learned about congruences and solved congruences and so on, and looked at other things that Fermat did.

Raussen and Skau: *You did this work besides your ordinary schoolwork?*

Wiles: Yes, I don't think my schoolwork was too taxing from that point of view.

Raussen and Skau: *Was it already clear to you at that time that you had an extraordinary mathematical talent?*

Wiles: I certainly had a mathematical aptitude and obviously loved to do mathematics, but I don't think I felt that I was unique. In fact, I don't believe I was unique in the school I attended. There were others who had just as strong a claim to be future mathematicians, and some of them have become mathematicians, too.

Raussen and Skau: *Did you already plan to study mathematics and to embark on a mathematical career at that age?*

Wiles: No, I don't think I really understood you could spend your life doing mathematics. I think that only came later. But I certainly wanted to study it as long as I could. I'm sure that as far as my horizon extended, it involved mathematics.

Raussen and Skau: *You started to study mathematics as a student at Oxford in 1971. Can you tell us a little bit about how that worked out? Were there any particular teachers or any particular areas that were particularly important for you?*

Wiles: Before I went to college (actually in high school), one of my teachers had a PhD in number theory. He gave me a copy of Hardy and Wright's *An Introduction to the Theory of Numbers*, and I also found a copy of Davenport's *The Higher Arithmetic*. And these two books I found very, very inspiring in terms of number theory.

Raussen and Skau: *So you were on track before you started studying?*

Wiles: Yes, I was on track before. In fact, to some extent, I felt college was a distraction because I had to do all these other things: applied maths, logic, and so on, and I just wanted to do number theory. You were not allowed to do number theory in your first year. And you could not really get down to it before your third year.

Raussen and Skau: *But you were not interested in geometry, not as much as in algebra and number theory, anyway?*

Wiles: No, I was primarily interested in algebra and number theory. I was happy to learn these other things, but I really was most excited about number theory. My teachers arranged for me to take extra classes in number theory, but there was not that much to offer.

At one point, I decided that I should put all the years of Latin I had done at school to good use and try to read some of Fermat in the original, but I found that was actually too hard. Even if you translated the Latin, the way they wrote in those days wasn't in the algebraic symbols I was used to, so it was quite difficult.

Raussen and Skau: *It must have been a relief when you were done and came to Cambridge to start studying number theory for real, with John Coates as your supervisor.*

Wiles: That's right. I had a year, a preliminary year, in which I just studied a range of subjects and then I could do a special paper. John Coates was not yet at Cambridge, but I think he helped me, maybe over the summer. Anyway, that summer I met him and started working with him right away, and that was just wonderful. The transition from undergraduate work, where you were just reading and studying, to research—that was the real break for me. It was just wonderful.

Elliptic Curves

Raussen and Skau: *We assume it was John Coates who initiated your work on elliptic curves and Iwasawa theory?*

Wiles: Absolutely. He had some wonderful ideas and was generous to share them with me.

Raussen and Skau: *Did you tell John Coates that you were interested in the Fermat problem?*

Wiles: Perhaps I did. I don't remember. It is really true that there hadn't been any new ideas since the nineteenth century. People were trying to refine the old methods and, yes, there were refinements. But it didn't look like these refinements and the solution were going to converge. It was just too hard that way.

Raussen and Skau: *At the time you started to work with John Coates, you had no idea that these elliptic curves were going to be crucial for the solution of Fermat's Last Theorem?*

Wiles: No, it's a wonderful coincidence. The strange thing is that, in a way, the two things that are most prominent in Fermat that we remember today are his work on elliptic curves and his famous last theorem. For example, this equation you mentioned, $y^2 + 2 = x^3$, is an elliptic curve. And the two strands came together in the proof.

Raussen and Skau: *Could you explain what an elliptic curve is and why elliptic curves are of interest in number theory?*

Wiles: For a number theorist, the life of elliptic curves started with Fermat as equations of the form y^2 equals a cubic polynomial in x with rational coefficients. Then,

the problem is to find the rational solutions to such an equation. What Fermat noticed was the following. Sometimes you can start with one or even two rational solutions and use them to generate infinitely many others. And yet sometimes there are no solutions. This latter situation occurs, for example, in the case $n = 3$ of Fermat's Last Theorem, the equation being, in fact, an elliptic curve in disguise. Sometimes you can show there are no rational solutions. You could have infinitely many and you could have none. This was already apparent to Fermat.

In the early nineteenth century, one studied these equations in complex numbers. Abel [1802–1829] himself came in at this point and studied elliptic functions and related these to elliptic curves, implying that elliptic curves have a group structure. They were very well understood in terms of doubly periodic functions in the early nineteenth century. But that is what underlies the complex solutions, solutions to the equation in complex numbers.

The solutions to the equation in rational numbers were studied by Poincaré [1854–1912]. What's now known as the Mordell–Weil theorem was proved by Mordell [1888–1972] and then Weil in the 1920s, answering a question of Poincaré. In our setting, it says that the K-rational points on an elliptic curve over a number field K, in particular for K equal to the rationals, form a finitely generated abelian group. That is, from Fermat's language, you can start with a finite number of solutions and, using those, generate all the solutions by what he called the chord-and-tangent process.

By now you know the structure; it is a very beautiful algebraic structure, the structure of a group, but that does not actually help you find the solutions. So, no one really had any general methods for finding the solutions until the conjectures of the 1960s, which emerged from the Birch and Swinnerton-Dyer Conjecture. There are two aspects to it; one is somewhat analytic, and one is in terms of what is called the Tate–Shafarevich group. Basically, the Tate–Shafarevich group gives you the obstruction to an algorithm for finding the solutions. And the Birch and Swinnerton-Dyer Conjecture tells you that there is actually an analytic method for analyzing this so-called Tate–Shafarevich group. If you combine all this together, ultimately it should give you an algorithm for finding the solutions.

Birch and Swinnerton-Dyer, Tate–Shafarevich, Selmer…

Raussen and Skau: *You worked on the Birch and Swinnerton-Dyer Conjecture when you were a graduate student together with John Coates?*

Wiles: Yes, that is exactly what he proposed working on. We got the first result in certain special families of elliptic curves on this analytic link between the solutions and what is called the L-function of the elliptic curve.

Raussen and Skau: *These were curves admitting complex multiplication?*

Wiles: Exactly; these were the elliptic curves with complex multiplication.

Interviewed in Oslo in May, Wiles told Martin Raussen and Christian Skau that he set out to prove the Modularity Conjecture with no idea from what branch of mathematics the answer would come.

Raussen and Skau: *Was this the first general result concerning the Birch and Swinnerton-Dyer Conjecture?*

Wiles: It was the first one that treated a family of cases rather than individual cases. There was a lot of numerical data for individual cases, but this was the first infinite family of cases.

Raussen and Skau: *This was over the rational numbers?*

Wiles: Yes.

Raussen and Skau: *We should mention that the Birch and Swinnerton-Dyer Conjecture is one of the Clay Millennium Prize Problems, which would earn a person who solves it one million dollars.*

Wiles: That's right. I think it's appealing, partly because it has its roots in Fermat's work, just like the Fermat problem. It is another "elementary-to-state" problem concerned with equations—in this case of very low degree—which we can't master and which Fermat initiated. I think it is a very appealing problem.

Raussen and Skau: *Do you think it is within reach? In other words, do we have the necessary tools for somebody daring enough to attack it and succeed? Or do we have to wait for another three hundred years to see it solved?*

Wiles: I don't suppose it will take three hundred years, but I don't think it is the easiest of the Millennium Problems. I think we are still lacking something. Whether the tools are all here now, I am not sure. They may be. There are always these speculations with these really difficult problems; it may be that the tools simply aren't there. I don't believe that anyone in the nineteenth century could have solved Fermat's Last Theorem, certainly not in the way it was eventually solved. There was just too big a gap in mathematical history. You had to wait another hundred years for the right pieces to be in place. You can never be quite sure about these problems, whether they are accessible to your time. That is really what makes them so challenging; if you had the intuition for what can be done now and what can't be done now, you would be a long way towards a solution!

Raussen and Skau: *You mentioned the Tate–Shafarevich group and in that connection the Selmer group appears. Selmer [1920–2006] was a Norwegian mathematician, and it was Cassels [1922–2015] who was responsible for naming this group the Selmer group. Could you say a few words about the Selmer group and how it is related to the Tate–Shafarevich group, even if it's a little technical?*

Wiles: It is technical, but I can probably explain the basic idea of what the Selmer group is. What you are trying to do is to find the rational solutions on an elliptic curve. The method is to take the rational points on the elliptic curve—suppose you have got some—and you generate field extensions from these. So when I say generate extensions, I mean that you can take roots of those points on the elliptic curve. Just like taking the nth root of 5 or the cube root of 2. You can do the same thing on an elliptic curve; you can take the nth root of a point. These are all points which *added to themselves n times* give you the point you started with. They generate certain extensions of the number field you started with, in our case the rational number field $\mathbf{Q}$.

You can put a lot of restrictions on those extensions. And the Selmer group is basically the smallest set of extensions you can get putting on all the obvious restrictions.

Let me summarize this. You've got the group of points. They generate some extensions; that's too big—you don't want all extensions. You cut that down as much as you can using local criteria, using p-adic numbers; that's called the Selmer group. And the essential difference between the group generated by the points and the Selmer group is the Tate–Shafarevich group. So the Tate–Shafarevich group gives you the error term, if you like, in trying to get at the points via the Selmer group.

Raussen and Skau: *Selmer's paper, which Cassels refers to, studied the Diophantine equation* $3x^3 + 4y^3 + 5z^3 = 0$ *and similar ones. Selmer showed that it has just a trivial solution in the integers, while modulo n it has nontrivial solutions for all n. In particular, this curve has no rational points. Why did Cassels invoke Selmer's name in naming the group?*

Wiles: Yes, there are quite subtle relationships between these. What happens is you are actually looking at one elliptic curve, which in this case would be $x^3+y^3+60z^3 = 0$. That is an elliptic curve, in disguise, if you like, and the Tate–Shafarevich group involves looking at other ones like it, for example, $3x^3 + 4y^3 + 5z^3 = 0$, which is a genus one curve but which has no rational points. Its Jacobian is the original elliptic curve $x^3 + y^3 + 60z^3 = 0$. One way of describing the Tate–Shafarevich group is in terms of these curves that have genus one but don't have rational points. And by assembling these together you can make the Tate–Shafarevich group, and that is reflected in the Selmer group. It is too intricate to explain in words, but it is another point of view. I gave it in more arithmetic terminology in terms of extensions. The more geometric terminology was in terms of these twisted forms.

The Modularity Conjecture

Raussen and Skau: *What you proved in the end was a special case of what is now called the Modularity Conjecture. In order to explain this, one has to start with modular forms and how modular forms can be put in relation with elliptic curves. Could you give us some explanations?*

Wiles: Yes; we have described an elliptic curve (over the rationals) as an equation $y^2 = x^3 + ax + b$, where the a and b are assumed to be rational numbers. (There is also a condition that the discriminant should not vanish.) As I said, at the beginning of the nineteenth century you could describe the complex solutions to this equation. You could describe these very nicely in terms of the Weierstrass $\wp$-function, in terms of a special elliptic function. But what we want is actually a completely different uniformization of these elliptic curves which captures the fact that the a and b are rational numbers. It is a parametrization just for the rational elliptic curves. And because it captures the fact that it is defined over the rationals, it gives you a much better hold on solutions over the rationals than the elliptic functions do. The latter really only sees the complex structure.

And the place it comes from are modular forms or modular curves. To describe modular functions first: we are used to functions which satisfy the relation of being invariant under translation. Every time we write down a Fourier series, we have a function which is invariant under translation. Modular functions are ones which are invariant under the action of a much bigger group, usually a subgroup of $SL_2(\mathbb{Z})$. So, you would ask for a function $f(z)$ in one complex variable, usually on the upper-half-plane, which satisfies $f(z)$ is the same as $f((az+b)/(cz+d))$ (or, more generally, is that times a power of $cz+d$).

Modular functions hold the key to the arithmetic of elliptic curves.

These are called modular functions, and they were extensively studied in the nineteenth century. Surprisingly, they hold the key to the arithmetic of elliptic curves. Perhaps the simplest way to describe it is that because we have an action of $SL_2(\mathbb{Z})$ on the upper-half-plane H—by the action z goes to $(az+b)/(cz+d)$—we can look at the quotient H modulo this action. You can then give the quotient the structure of a curve. In fact, it naturally gets the structure of a curve over the rational numbers. If you take a subgroup of $SL_2(\mathbb{Z})$, or more precisely what is called a congruence subgroup, defined by the c value being divisible by N, then you call the curve a modular curve of level N. The Modularity Conjecture asserts that every elliptic curve over the rationals is actually a quotient of one of these modular curves for some integer N. It gives you a uniformization of elliptic curves by these other entities, these modular curves. On the face of it, it might seem we are losing because this is a high genus curve—it is more complicated. But it actually has a lot more structure because it is a moduli space.

Raussen and Skau: *And that is a very powerful tool?*

Wiles: That is a very powerful tool, yes. You have function theory, you have deformation theory, geometric methods, etc. You have a lot of tools to study it.

Raussen and Skau: *Taniyama, the young Japanese mathematician who first conjectured or suggested these connections; his conjecture was more vague, right?*

Wiles: His conjecture was more vague. He didn't pin it down to a function invariant under the modular group. I've forgotten exactly what he conjectured; it was invariant under some kind of group, but I forget exactly which group he was predicting. But it was not as precise as the congruence subgroups of the modular group. I think it was originally written in Japanese, so it was not circulated as widely as it might have been. I believe it was part of notes compiled after a conference in Japan.

Raussen and Skau: *It was an incredibly audacious conjecture at that time, wasn't it?*

Wiles: Apparently, yes.

Raussen and Skau: *But then it gradually caught the attention of other mathematicians. You told us already about Gerhard Frey, who came up with a conjecture relating Fermat's Last Theorem with the Modularity Conjecture.*

Wiles: That's right. Gerhard Frey showed that if you take a solution to the Fermat problem, say $a^p + b^p = c^p$, and you create the elliptic curve $y^2 = x(x - a^p)(x + b^p)$, then the discriminant of that curve would end up being a perfect pth power. And if you think about what that means assuming the Modularity Conjecture—you have to assume something a bit stronger as well (the so-called epsilon conjecture of Serre)—then it forces this elliptic curve to have the level N that I spoke about to be equal to one, and hence the associated congruence subgroup is equal to $SL_2(\mathbb{Z})$. But H modulo $SL_2(\mathbb{Z})$ is a curve of genus zero. It has no elliptic curve quotient, so it wasn't there after all, and hence there can't be a solution to the Fermat problem.

The Quest for a Proof

Raussen and Skau: *That was the point of departure for your own work, with crucial further ingredients due to Serre and Ribet making this connection clear. May we briefly summarize the story that then followed? It has been told by you many times, and it is the focus of a BBC documentary.*

You had moved to the United States, first to Harvard, then to Princeton University, becoming a professor there. When you heard of Ribet's result, you devoted all your research time to proving the Modularity Conjecture for semistable elliptic curves over the rationals. This work went on for seven years of really hard work in isolation. At the same time you were working as a professor in Princeton and you were raising small children.

A proof seems to be accomplished in 1993, and the development culminates in a series of three talks at the Isaac Newton Institute in Cambridge back in England, announcing your proof of Fermat's Last Theorem. You are celebrated by your mathematical peers. Even the world press takes an interest in your results, which happens very rarely for mathematical results.

But then, when your result is scrutinized by six referees for a highly prestigious journal, it turns out that there is a subtle gap in one of your arguments and you are sent back to the drawing board. After a while, you send for your former student, Richard Taylor, to come to Princeton to help you in your efforts. It takes a further ten months of hard and frustrating work; we think we do not exaggerate by calling it a heroic effort under enormous pressure. Then, in a sudden flash of insight, you realize that you can combine some of your previous attempts with new results to circumvent the problem that had caused the gap. This turns out to be what you need in order to get the part of the Modularity Conjecture that implies Fermat's Last Theorem. What a relief that must have been! Would you like to give a few comments on this dramatic story?

Wiles: With regard to my own work, when I became a professional mathematician working with Coates, I realized I really had to stop working on Fermat because it was time-consuming and I could see that in the last hundred years almost nothing had been done. And I saw others, even very distinguished mathematicians, had come to grief on it. When Frey came out with this result, I was a bit skeptical that the Serre part of the conjecture was going to be true, but when Ribet proved it, then, okay, this was it!

It is irresponsible to work on one problem to the exclusion of everything else.

And it was a long, hard struggle. In some sense, it is irresponsible to work on one problem to the exclusion of everything else, but this is the way I tend to do things. Whereas Fermat is very narrow (I mean, it is just this one equation, whose solution may or may not help with anything else), the setting of the Modular Conjecture was one of the big problems in number theory. It was a great thing to work on anyway, so it was just a tremendous opportunity.

When you are working on something like this, it takes many years to really build up the intuition to see the kinds of things you need and the kinds of things a solution will depend on. It's something like discarding everything you can't use and won't work till your mind is so focused that even making a mistake, you've seen enough that you'll find another way to the end.

Funnily enough, concerning the mistake in the argument that I originally gave, people have worked on that aspect of the argument and quite recently they have actually shown that you can produce arguments very like that. In fact, in every neighboring case, arguments similar to the original method seem to work, but there is this unique case that it doesn't seem to work for, and there is not yet any real explanation for it. So the same kind of argument I was trying to use, using Euler systems and so on, has been made to work in every surrounding case, but not the one I needed for Fermat. It's really extraordinary.

Raussen and Skau: *You once likened this quest for the proof of the Modularity Theorem to a journey through a dark, unexplored mansion. Could you elaborate?*

Wiles: I started off really in the dark. I had no prior insights of how the Modularity Conjecture might work or how you might approach it. One of the troubles with this problem—it's a little like the Riemann Hypothesis but perhaps even more so—was that you didn't even know what branch of mathematics the answer would be coming from.

To start with, there are three ways of formulating the problem: one is geometric, one is arithmetic, and one is analytic. And there were analysts—I would not understand their techniques at all well—who were trying to make headway on this problem.

I think I was a little lucky because my natural instinct was with the arithmetic approach and I went straight for the arithmetic route, but I could have been wrong. The only previously known cases where the Modularity Conjecture was known to hold were the cases of complex multiplication, and that proof is analytic, completely analytic.

Partly out of necessity, I suppose, and partly because that's what I knew, I went straight for an arithmetic approach. I found it very useful to think about it in a way that I had been studying in Iwasawa theory. With John Coates, I had applied Iwasawa theory to elliptic curves. When I went to Harvard, I learned about Barry Mazur's work, where he had been studying the geometry of modular curves using a lot of the modern machinery. There were certain ideas and techniques I could draw on from that. I realized after a while, I could actually use that to make a beginning—to find some kind of entry into the problem.

Raussen and Skau: *Before you started on the Modularity Conjecture, you published a joint paper with Barry Mazur, proving the main theorem of Iwasawa theory over the rationals. Can you please tell us what Iwasawa theory is all about?*

Wiles: Iwasawa theory grew out of the work of Kummer on cyclotomic fields and his approach to Fermat's Last Theorem. He studied the arithmetic, and in particular the ideal class groups, of prime cyclotomic fields. Iwasawa's idea was to consider the tower of cyclotomic fields obtained by taking all p-power roots of unity at once. The main theorem of Iwasawa theory proves a relation between the action of a generator of the Galois group on the p-primary class groups and the p-adic L-functions. It is analogous to the construction used in the study of curves over finite fields where the characteristic polynomial of Frobenius is related to the zeta function.

Raussen and Skau: *And these tools turned out to be useful when you started to work on the Modularity Conjecture?*

Wiles: They did; they gave me a starting point. It wasn't obvious at the time, but when I thought about it for a while, I realized that there might be a way to start from there.

Parallels to Abel's Work

Raussen and Skau: *We want to read you a quotation: "The ramparts are raised all around but, enclosed in its last redoubt, the problem defends itself desperately. Who will be the fortunate genius who will lead the assault upon it or force it to capitulate?"*

Wiles: It must have been E. T. Bell, I suppose. Is it?

Raussen and Skau: *No, it's not. It is actually a quote from the book* Histoire des Mathématiques *by Jean-Étienne Montucla [1725–1799], written in the late eighteenth century. It is really the first book ever written on the history of mathematics. The quotation refers to the solvability or unsolvability of the quintic equation by radicals. As you know, Abel [1802–1829] proved the unsolvability of the general quintic equation when he was twenty-one years old. He worked in complete isolation, mathematically speaking, here in Oslo. Abel was obsessed, or at least extremely attracted, to this problem. He also got a false start. He thought he could prove that one could actually solve the quintic by radicals. Then he discovered his mistake and he finally found the unsolvability proof. Well, this problem was, at that time, almost three hundred years old and very famous. If we move fast-forward two hundred years, the same quotation could be used about the Fermat problem, which was around three hundred fifty years old when you solved it. It is a very parallel story in many ways. Do you have any comments?*

Wiles: Yes. In some sense, I do feel that Abel, and then Galois [1811–1832], were marking a transition in algebra from these equations, which were solvable in some very simple way, to equations which can't be solved by radicals. But this is an algebraic break that came with the quintic. In some ways, the whole trend in number theory now is the transition from basically abelian and possibly solvable extensions to insolvable extensions. How do we do the arithmetic of insolvable extensions?

I believe the Modularity Conjecture was solved because we had moved on from this original abelian situation to a nonabelian situation, and we were developing tools, modularity and so on which are fundamentally nonabelian tools. (I should say, though, that the proof got away mostly with using the solvable situation, not because it was more natural but because we have not solved the relevant problems in the general nonsolvable case.)

It is the same transition in number theory that he was making in algebra, which provides the tools for solving this equation. So I think it is very parallel.

Raussen and Skau: *There is an ironic twist with Abel and the Fermat problem. When he was twenty-one years old, Abel came to Copenhagen to visit Professor Degen [1766–1825], who was the leading mathematician in Scandinavia at that time. Abel wrote a letter to his mentor in Oslo, Holmboe [1795–1850], stating three results about the Fermat equation without giving any proofs—one of them is not easy to prove, actually. This, of course, is just a curiosity today.*

But in the same letter, he gives vent to his frustration, intimating that he can't understand why he gets an equation of degree n^2 and not n when dividing the lemniscate arc in n equal pieces. It was only after returning to Oslo

that he discovered the double periodicity of the lemniscate integral and also of the general elliptic integral of the first kind.

If one thinks about it, what he did on the Fermat equation turned out to be just a curiosity. But what he achieved on elliptic functions, and implicitly on elliptic curves, turned out later to be a relevant tool for solving it. Of course, Abel had no idea that this would have anything to do with arithmetic. So this story tells us that mathematics sometimes develops in mysterious ways.

Wiles: It certainly does, yes.

Work Styles

Raussen and Skau: *May we ask for some comments about work styles of mathematicians in general and also about your own? Freeman Dyson, a famous physicist and mathematician at IAS in Princeton, said in his Einstein Lecture in 2008: "Some mathematicians are birds, others are frogs. Birds fly high in the air and survey broad vistas of mathematics out to the horizon. They delight in concepts that unify our thinking and bring together diverse problems from different parts of the landscape. Frogs live in the mud below and see only the flowers that grow nearby. They delight in the details of particular objects and they solve problems one at a time."*

Freeman Dyson didn't say that birds were better than frogs or the other way around. He considered himself a frog rather than a bird.

When we are looking at your work, for us it seems rather difficult to decide where to place you in his classification scheme: among the birds (those who create theories) or among the frogs (those who solve problems). What is your own perception?

Wiles: Well, I don't feel like either. I'm certainly not a bird—unifying different fields. I think of frogs as jumping a lot. I think I'm very, very focused. I don't know what the animal analogy is, but I think I'm not a frog in the sense of enjoying the nearby landscape. I'm very, very concentrated on the problem I happen to work on and I am very selective. And I find it very hard to even take my mind off it enough to look at any of the flowers around, so I don't think that either of the descriptions quite fit.

Raussen and Skau: *Based on your own experience, could you describe the interplay between hard, concentrated, and persevering work on the one side and, on the other side, these sudden flashes of insight that seemingly come out of nowhere, often appearing in a more relaxed setting. Your mind must have worked unconsciously on the problem at hand, right?*

Wiles: I think what you do is that you get to a situation where you know a theory so well, and maybe even more than one theory, so that you have seen every angle and tried a lot of different routes.

There is this tremendous amount of work in the preparatory stage where you have to understand all the details and maybe some examples—that is your essential launch pad. When you have developed all this, you let the mind relax and then at some point—maybe when you go away and do something else for a little bit—you come back and suddenly it is all clear. Why did you not think of that? This is something the mind does for you. It is the flash of insight.

I remember (this is a trivial example in a non-mathematical setting) someone once showed me some script—it was some Gothic script—and I couldn't make head nor tail of it. I was trying to understand a few letters, and I gave up. Then I came back half an hour later and I could read the whole thing. The mind somehow does this for you and we don't quite know how, but we do know what we have to do to set up the conditions where it will happen.

Raussen and Skau: *This is reminiscent of a story about Abel. While in Berlin, he shared an apartment with some Norwegian friends who were not mathematicians. One of his friends said that Abel typically woke up during the night, lit a candle, and wrote down ideas that he woke up with. Apparently his mind was working while asleep.*

Wiles: Yes, I do that, except I don't feel the need to write them down when I wake up with it because I know I will not forget it. But if I have an idea when I am about to go to sleep, I am terrified that I will not wake up with that idea, so then I have to write it down.

Raussen and Skau: *Are you thinking in terms of formulas or in terms of geometric pictures or what?*

Wiles: It is not really geometric. I think it is patterns and I think it is just parallels between situations I have seen elsewhere and the one I am facing now. In a perfect world, what is it all pointing to? What are the ingredients that ought to go into this proof? What am I not using that I still have in my pocket? Sometimes it is just desperation. I assemble every piece of evidence I have and that's all I've got. I have got to work with that and there is nothing else.

I often feel that doing mathematics is like being a squirrel.

I often feel that doing mathematics is like being a squirrel and there are some nuts at the top of a very tall tree. But there are several trees and you don't know which one. What you do is that you run up one and you think, no, it does not look good on this one, and you go down and up another one, and you spend your whole life just going up and down these trees. But you've only got up to thirty feet. Now, if someone told you the rest of the trees—it's not in them, you have only one tree left—then you would just keep going until you found it. In some sense, it is ruling out the wrong things—that is really crucial. And if you just believe in your intuition and your intuition is correct and you stick with your one tree, then you will find it.

Problems in Mathematics

Raussen and Skau: *Felix Klein [1849–1925] once said: "Mathematics develops as old results are being understood and illuminated by new methods and insights. Proportionally with a better and deeper understanding new problems*

naturally arise." And David Hilbert [1862–1943] stressed that "problems are the lifeblood of mathematics." Do you agree?

Wiles: I certainly agree with Hilbert, yes. Good problems are the lifeblood of mathematics. I think you can see this clearly in number theory in the second half of the last century. For me personally, there is obviously the Modularity Conjecture but also the whole Langlands program and the Birch and Swinnerton-Dyer Conjecture. These problems give you a very clear focus on what we should try to achieve. We also have the Weil Conjectures on curves and varieties over finite fields and the Mordell Conjecture and so on.

These problems somehow concentrate the mind and also simplify our goals in mathematics. Otherwise, we can get very, very spread out and not sure what's of value and what's not of value.

Raussen and Skau: *Do we have as good problems today as when Hilbert formulated his twenty-three problems in 1900?*

Wiles: I think so, yes.

Raussen and Skau: *Which one do you think is the most important problem today? And how does the Langlands program fit in?*

Wiles: Well, I think the Langlands program is the broadest spectrum of problems related to my field. I think that the Riemann Hypothesis is the single greatest problem from the areas I understand. It is sometimes hard to say exactly why that is, but I do believe that solving it would actually help solve some of these other problems. And then, of course, I have a very personal attachment to the Birch and Swinnerton-Dyer Conjecture.

Raussen and Skau: *Intuition can lead us astray sometimes. For example, Hilbert thought that the Riemann Hypothesis would be solved in his lifetime. There was another problem on his list, the seventh, that he never thought would be solved in his lifetime, but which was solved by Gelfond [1906–1968] in 1934. So our intuition can be wrong.*

Wiles: That is right. I'm not surprised that Hilbert felt that way. The Riemann Hypothesis has such a clear statement, and we have the analogue in the function field setting. We understand why it is true there and we feel we ought to be able to translate it. Of course, many people have tried and failed. But I would still expect it to be solved before the Birch and Swinnerton-Dyer Conjecture.

Investing in Mathematics

Raussen and Skau: *Let's hope we'll find out in our lifetimes!*

Classical mathematics has, roughly speaking, two sources: one of them coming from the physical sciences and the other one from—let's for simplicity call it number theoretical speculations, with number theory not associated to applications.

That has changed. For example, your own field of elliptic curves has been applied to cryptography and security. People are making money with elliptic curves nowadays! On the other hand, many sciences apart from physics really take advantage and profit from mathematical thinking and mathematical results. Progress in industry nowadays often depends on mathematical modelling and optimization methods. Science and industry propose challenges to the mathematical world.

In a sense, mathematics has become more applied than it ever was. One may ask whether this is a problem for pure mathematics. It appears that pure mathematics sometimes is put to the sidelines, at least from the point of view of the funding agencies. Do you perceive this as a serious problem?

Wiles: Well, I think in comparison with the past, one feels that mathematicians two, three hundred years ago were able to handle a much broader spectrum of mathematics, and a lot more of it touched applied mathematics than a typical pure mathematician would do nowadays. On the other hand, that might be because we only remember the very best and most versatile mathematicians from the past.

I think it is always going to be a problem if funding agencies are short-sighted. If they want to see a result in three years, then it is not going to work. It is hard to imagine a pure development and then the application all happening within three to five years. It is probably not going to happen.

On the other hand, I don't believe you can have a happily functioning applied maths world without the pure maths to back it up, providing the future and keeping them on the straight and narrow. So it would be very foolish not to invest in pure mathematics. It is a bit like only investing in energy resources that you can see now. You have to invest in the future; you have to invest in fusion power or solar power or these other things. You don't just use up what is there and then start worrying about it when it is gone. It is the same with mathematics; you can't just use up the pure mathematics we have now and then start worrying about it when you need a pure result to generate your applications.

Mathematical Awards

Raussen and Skau: *You have already won a lot of prizes as a result of your achievements, culminating in proving Fermat's Last Theorem. You have won the Rolf Schock Prize, given by the Swedish Academy; the Ostrowski Prize, which was given to you in Denmark; the Fermat Prize in France; the Wolf Prize in Israel; the Shaw Prize in Hong Kong (the prize that has been named the Nobel Prize of the East), and the list goes on, culminating with the Abel Prize tomorrow. May we ask you whether you enjoy these awards and the accompanying celebrations?*

Wiles: I certainly love them, I have to say. I think they are a celebration of mathematics. I think with something like Fermat, it is something people are happy to see in their lifetime. I would obviously be very happy to see the Riemann Hypothesis solved. It is just exciting to see how it finally gets resolved and just to understand the end of the story—because a lot of these stories we won't live to see the end of. Each time we do see the end of such a story, it is something we will naturally celebrate. For me, I learned about the Fermat problem from this book of E. T. Bell and about the Wolfskehl Prize attached to it. The Wolfskehl Prize was still there—only just, I may

say; I only had a few years left before the deadline for it expired.

Raussen and Skau: *This gives us the lead to talk a little about that prize. The Wolfskehl Prize was founded in 1906 by Paul Wolfskehl [1856–1906], who was a German physician with an interest in mathematics. He bequeathed one hundred thousand Reichmarks (equivalent to more than one million dollars in today's money) to the first person to prove Fermat's Last Theorem. The prize was, according to the testament, valid until 13 September 2007 and you received it in 1997. By then, due in part to the hyperinflation Germany suffered after World War I, the prize money had dwindled a lot.*

Wiles: For me, the amount of money was unimportant. It was the sentimental feeling attached to the Wolfskehl Prize that was important for me.

Graduate Students

Raussen and Skau: *You have had altogether twenty-one PhD students and you have attracted very gifted students. Some of them are really outstanding. One of them, Manjul Bhargava, won the Fields Medal in 2014. It must be a pleasure to be advisor to such students.*

Wiles: Yes, I don't want to take too much credit for it. In the case of Manjul, I suggested a problem to him, but after that I had nothing much more to do. He was coming up with these absolutely marvelous discoveries. In some sense, you get more credit if you have very gifted students, but the truth is that very gifted students don't really require that much help.

Raussen and Skau: *What is the typical way for you of interacting with graduate students?*

Wiles: Well, I think the hardest thing to learn as a graduate student is that afterwards you need to carry on with the rest of your professional life; it's hard to pick problems. And if you just assign a problem and they do it, in some sense that hasn't given them terribly much. Okay, they solved that problem, but the hard thing is then to have to go off and find other problems! So I prefer it if we come to a decision on the problem together.

I give them some initial idea and which area of mathematics to look at, having not quite focused on the problem. Then, as they start working and become experts, they can see a better way of pinning down what the right question is. And then they are part of the process of choosing the problem. I think that is a much better investment for their future. It doesn't always work out that way, and sometimes the problem you give them turns out to be the right thing. But usually it is not that way and usually it's a process to find the right problem.

Hobbies and Interests

Raussen and Skau: *We always end the Abel interviews by asking the laureate what they enjoy doing when they are not working with mathematics. What are your hobbies and interests outside mathematics?*

Wiles: Well, it varies at different times. When I was doing Fermat and being a father with young children, that combination was all-consuming.

I like to read and I like various kinds of literature: novels, some biographies—it is fairly balanced. I don't have any other focused obsessions. When I was in school, I played on chess teams and bridge teams, but when I started to do serious mathematics, I completely lost interest in those.

Raussen and Skau: *What about music; are you fond of music?*

Wiles: I go and listen to concerts, but I am not myself actively playing anything. I enjoy listening to music—classical, preferably.

Raussen and Skau: *Are you interested in other sciences apart from mathematics?*

Wiles: I would say somewhat. These are things I do to relax, so I don't like them to be too close to mathematics. If it is something like animal behavior or astrophysics or something from a qualitative point of view—I certainly enjoy learning about those—likewise about what machines are capable of, and many other kinds of popular science, but I'm not going to spend my time learning the details of string theory. I'm too focused to be willing to do that. Not that I would not be interested, but this is my choice.

Raussen and Skau: *We would like to thank you very much for this wonderful interview, first of all on behalf of the two of us but also on behalf of the Norwegian, the Danish, and the European Mathematical Societies. Thank you so much!*

Wiles: Thank you very much!

Photo Credits

Photo of Sir Andrew J. Wiles with Crown Prince Haakon is courtesy of Audun Braastad.

Photo of Sir Andrew J. Wiles with Martin Raussen, and Christian Skau is courtesy of Erik F. Baardsen, DNVA.

Sidebar 1. Abel Prize Winners

2016: Andrew Wiles
2015: John Forbes Nash Jr. and Louis Nirenberg
2014: Yakov Sinai
2013: Pierre Deligne
2012: Endre Szemerédi
2011: John Milnor
2010: John Tate
2009: Mikhail Leonidovich Gromov
2008: John G. Thompson and Jacques Tits
2007: S. R. Srinivasa Varadhan
2006: Lennart Carleson
2005: Peter Lax
2004: Michael Atiyah and Isadore Singer
2003: Jean-Pierre Serre

Sidebar 2. *Notices* articles on Wiles

July/August 1993: *Wiles Proves Taniyama's Conjecture; Fermat's Last Theorem Follows*, by Kenneth A. Ribet, math.berkeley.edu/~ribet/Articles/notices.pdf

October 1993: *Fermat Fest Draws a Crowd*, by Allyn Jackson

October 1994: *Another Step Toward Fermat*, by Allyn Jackson, www.ams.org/notices/199501/rubin.pdf

July 1995: *The Proof of Fermat's Last Theorem by R. Taylor and A. Wiles*, by Gerd Faltings, www.ams.org/notices/199507/faltings.pdf

July 1996: *Wiles Receives NAS Award in Mathematics*, by John Coates, www.ams.org/notices/199607/comm-wiles.pdf

January 1997: *Review of BBC's Horizon Program, "Fermat's Last Theorem*," reviewed by Andrew Granville, www.ams.org/notices/199701/comm-granville.pdf

March 1997: Announcement: 1997 Cole Prize, www.ams.org/notices/199703/comm-cole.pdf

November 1997: *Paul Wolfskehl and the Wolfskehl Prize*, by Klaus Barner, www.ams.org/notices/199710/barner.pdf

November 1997: *Book Review: Fermat's Enigma by Simon Singh*, reviewed by Allyn Jackson, www.ams.org/notices/199710/comm-fermat.pdf

December 1999: *Research News: A Proof of the Full Shimura-Taniyama-Weil Conjecture Is Announced*, by Henri Darmon, www.ams.org/notices/199911/comm-darmon.pdf

December 2001: *Theater Review: Fermat's Last Tango*, reviewed by Robert Osserman, www.ams.org/notices/200111/rev-osserman.pdf

September 2005: *Wiles Receives 2005 Shaw Prize*, by Allyn Jackson, www.ams.org/notices/200508/comm-shaw.pdf

June/July 2016: *Sir Andrew J. Wiles Awarded Abel Prize*, www.ams.org/publications/journals/notices/201606/rnoti-p608.pdf

Andrew Wiles's Marvelous Proof

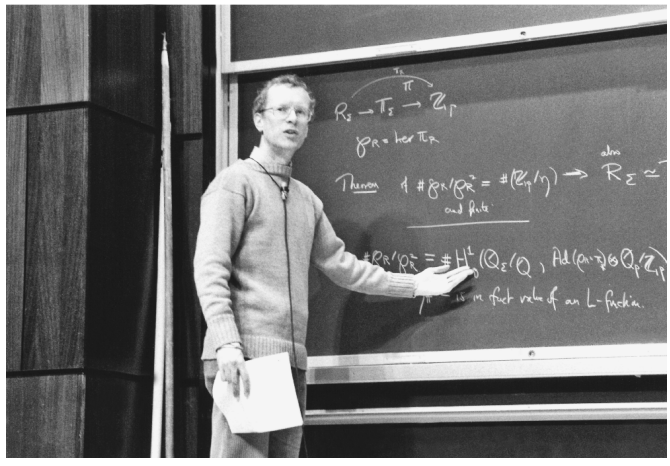
Henri Darmon

Fermat famously claimed to have discovered "a truly marvelous proof" of his Last Theorem, which the margin of his copy of Diophantus's *Arithmetica* was too narrow to contain. While this proof (if it ever existed) is lost to posterity, Andrew Wiles's marvelous proof has been public for over two decades and has now earned him the Abel Prize. According to the prize citation, Wiles merits this recognition "for his stunning proof of Fermat's Last Theorem by way of the modularity conjecture for semistable elliptic curves, opening a new era in number theory."

Few can remain insensitive to the allure of Fermat's Last Theorem, a riddle with roots in the mathematics of ancient Greece, simple enough to be understood and appreciated by a novice (like the ten-year-old Andrew Wiles browsing the shelves of his local public library), yet eluding the concerted efforts of the most brilliant minds for well over three centuries, becoming over its long history the object of lucrative awards like the Wolfskehl Prize and, more importantly, motivating a cascade of fundamental discoveries: Fermat's method of infinite descent, Kummer's theory of ideals, the ABC conjecture, Frey's approach to ternary diophantine equations, Serre's conjecture on mod p Galois representations,….

It is also a centerpiece of the "Langlands program," the imposing, ambitious edifice of results and conjectures which has come to dominate the number theorist's view of the world.

Even without its seemingly serendipitous connection to Fermat's Last Theorem, Wiles's modularity theorem is a fundamental statement about elliptic curves (as evidenced, for instance, by the key role it plays in the proof of Theorem 2 of Karl Rubin's contribution in this volume). It is also a centerpiece of the "Langlands program," the imposing, ambitious edifice of results and conjectures which has come to dominate the number theorist's view of the world. This program has been described as a "grand unified theory" of mathematics. Taking a Norwegian perspective, it connects the objects that occur in the works of Niels Hendrik Abel, such as elliptic curves and their associated Abelian integrals and Galois representations, with (frequently infinite-dimensional) linear representations of the continuous transformation groups whose study was pioneered by Sophus Lie. This report focuses on the role of Wiles's theorem and its "marvelous proof" in the Langlands program in order to justify the closing phrase in the prize citation: how Wiles's proof has opened "a new era in number theory" and continues to have a profound and lasting impact on mathematics.



Wiles giving his first lecture in Princeton about his approach to proving the Modularity Conjecture in early 1994.

Our "beginner's tour" of the Langlands program will only give a partial and undoubtedly biased glimpse of the full panorama, reflecting the author's shortcomings as well as the inherent limitations of a treatment aimed at a general readership. We will motivate the Langlands program by starting with a discussion of *diophantine equations*: for the purposes of this exposition, they are equations of the form

$$\mathcal{X}: \quad P(x_1, \dots, x_{n+1}) = 0, \tag{1}$$

where P is a polynomial in the variables $x_1, \dots, x_{n+1}$ with integer (or sometimes rational) coefficients. One can examine the set, denoted $\mathcal{X}(F)$, of solutions of (1) with coordinates in any ring F. As we shall see, the subject draws much of its fascination from the deep and subtle ways in which the behaviours of different solution sets

Henri Darmon is James McGill Professor of Mathematics at McGill University and a member of CICMA (Centre Interuniversitaire en Calcul Mathématique Algébrique) and CRM (Centre de Recherches Mathématiques). His e-mail address is darmon@math.mcgill.ca.

This report is a very slightly expanded transcript of the Abel Prize lecture delivered by the author on May 25, 2016, at the University of Oslo.

DOI: http://dx.doi.org/10.1090/noti1487

can resonate with each other, even if the sets $\mathcal{X}(\mathbb{Z})$ or $\mathcal{X}(\mathbb{Q})$ of integer and rational solutions are foremost in our minds. Examples of diophantine equations include Fermat's equation $x^d + y^d = z^d$, the Brahmagupta–Pell equation $x^2 - Dy^2 = 1$ with $D > 0$, as well as elliptic curve equations of the form $y^2 = x^3 + ax + b$, in which a and b are rational parameters, the solutions (x, y) with rational coordinates being the object of interest in the latter case.

It can be instructive to approach a diophantine equation by first studying its solutions over *simpler* rings, such as the complete fields of real or complex numbers. The set

$$\mathbb{Z}/n\mathbb{Z} := \{0, 1, \dots, n-1\} \tag{2}$$

of remainders after division by an integer $n \geq 2$, equipped with its natural laws of addition, subtraction, and multiplication, is another particularly simple collection of numbers of *finite cardinality*. If $n = p$ is *prime*, this ring is even a *field*: it comes equipped with an operation of division by nonzero elements, just like the more familiar collections of rational, real, or complex numbers. The fact that $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ is a field is an algebraic characterisation of the primes that forms the basis for most known efficient primality tests and factorisation algorithms. One of the great contributions of Evariste Galois, in addition to the eponymous theory which plays such a crucial role in Wiles's work, is his discovery of a field of cardinality p^r for any prime power p^r. This field, denoted $\mathbb{F}_{p^r}$ and sometimes referred to as the Galois field with p^r elements, is even *unique* up to isomorphism.

For a diophantine equation $\mathcal{X}$ as in (1), the most basic invariant of the set

$$\mathcal{X}(\mathbb{F}_{p^r}) := \left\{ \begin{array}{c} (x_1, \dots, x_{n+1}) \in \mathbb{F}_{p^r}^{n+1} \text{ such that} \\ P(x_1, \dots, x_{n+1}) = 0 \end{array} \right\} \tag{3}$$

of solutions over $\mathbb{F}_{p^r}$ is of course its *cardinality*

$$N_{p^r} := \#\mathcal{X}(\mathbb{F}_{p^r}). \tag{4}$$

What patterns (if any) are satisfied by the sequence

$$N_p, N_{p^2}, N_{p^3}, \dots, N_{p^r}, \dots ? \tag{5}$$

This sequence can be packaged into a generating series like

$$\sum_{r=1}^{\infty} N_{p^r} T^r \qquad \text{or} \qquad \sum_{r=1}^{\infty} \frac{N_{p^r}}{r} T^r. \tag{6}$$

For technical reasons it is best to consider the exponential of the latter:

$$\zeta_p(\mathcal{X}; T) := \exp\left(\sum_{r=1}^{\infty} \frac{N_{p^r}}{r} T^r \right). \tag{7}$$

This power series in T is known as the *zeta function* of $\mathcal{X}$ over $\mathbb{F}_p$. It has integer coefficients and enjoys the following remarkable properties:

(1) It is a *rational function* in T:

$$\zeta_p(\mathcal{X}; T) = \frac{Q(T)}{R(T)}, \tag{8}$$

where $Q(T)$ and $R(T)$ are polynomials in T whose degrees (for all but finitely many p) are *independent of* p and determined by the shape—the complex topology—of the set $\mathcal{X}(\mathbb{C})$ of complex solutions;

(2) the reciprocal roots of $Q(T)$ and $R(T)$ are complex numbers of absolute value $p^{i/2}$ with i an integer in the interval $0 \leq i \leq 2n$.

The first statement—the rationality of the zeta function, which was proved by Bernard Dwork in the early 1960s—is a key part of the Weil conjectures, whose formulation in the 1940s unleashed a revolution in arithmetic geometry, driving the development of étale cohomology by Grothendieck and his school. The second statement, which asserts that the complex function $\zeta_p(\mathcal{X}; p^{-s})$ has its roots on the real lines $\Re(s) = i/2$ with i as above, is known as the Riemann hypothesis for the zeta functions of diophantine equations over finite fields. It was proved by Pierre Deligne in 1974 and is one of the major achievements for which he was awarded the Abel Prize in 2013.

That the asymptotic behaviour of N_p can lead to deep insights into the behaviour of the associated diophantine equations is one of the key ideas behind the Birch and Swinnerton-Dyer conjecture. Understanding the patterns satisfied by the function

$$p \mapsto N_p \qquad \text{or} \qquad p \mapsto \zeta_p(\mathcal{X}; T) \tag{9}$$

as the prime p varies will also serve as our motivating question for the Langlands program.

It turns out to be fruitful to package the zeta functions over all the finite fields into a single function of a complex variable s by taking the infinite product

$$\zeta(\mathcal{X}; s) = \prod_p \zeta_p(\mathcal{X}; p^{-s}) \tag{10}$$

as p ranges over all the prime numbers. In the case of the simplest nontrivial diophantine equation $x = 0$, whose solution set over $\mathbb{F}_{p^r}$ consists of a single point, one has $N_{p^r} = 1$ for all p, and therefore

$$\zeta_p(x = 0; T) = \exp\left(\sum_{r \geq 1} \frac{T^r}{r} \right) = (1 - T)^{-1}. \tag{11}$$

It follows that

$$\zeta(x = 0; s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \tag{12}$$

$$= \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots\right) \tag{13}$$

$$= \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s). \tag{14}$$

The zeta function of even the humblest diophantine equation is thus a central object of mathematics: the celebrated Riemann zeta function, which is tied to some of the deepest questions concerning the distribution of prime numbers. In his great memoir of 1860, Riemann proved that, even though (13) and (14) only converge absolutely on the right half-plane $\Re(s) > 1$, the function $\zeta(s)$ extends to a meromorphic function of $s \in \mathbb{C}$ (with a single pole at $s = 1$) and possesses an elegant functional equation relating its values at s and $1 - s$. The zeta functions of linear equations $\mathcal{X}$ in $n + 1$ variables are just shifts of the Riemann zeta function, since N_{p^r} is equal to p^{nr}, and therefore $\zeta(\mathcal{X}; s) = \zeta(s - n)$.

Moving on to equations of degree two, the general quadratic equation in one variable is of the form $ax^2+bx+c=0$, and its behaviour is governed by its *discriminant*

$$\Delta := b^2 - 4ac. \tag{15}$$

This purely algebraic fact remains true over the finite fields, and for primes $p \nmid 2a\Delta$ one has

$$N_p = \begin{cases} 0 & \text{if } \Delta \text{ is a nonsquare modulo } p, \\ 2 & \text{if } \Delta \text{ is a square modulo } p. \end{cases} \tag{16}$$

A priori, the criterion for whether $N_p = 2$ or 0—whether the integer Δ is or is not a quadratic residue modulo p—seems like a subtle condition on the prime p. To get a better feeling for this condition, consider the example of the equation $x^2 - x - 1$, for which $\Delta = 5$. Calculating whether 5 is a square or not modulo p for the first few primes $p \le 101$ leads to the following list:

$$N_p = \begin{cases} 2 \text{ for } p = 11, 19, 29, 31, 41, 59, 61, 71, 79, 89, 101, \dots \\ 0 \text{ for } p = 2, 3, 7, 13, 17, 23, 37, 43, 47, 53, 67, 73, \dots \end{cases} \tag{17}$$

A clear pattern emerges from this experiment: whether $N_p = 0$ or 2 seems to depend only on the rightmost digit of p, i.e., on what the remainder of p is modulo 10. One is led to surmise that

$$N_p = \begin{cases} 2 & \text{if } p \equiv 1, 4 \pmod 5, \\ 0 & \text{if } p \equiv 2, 3 \pmod 5, \end{cases} \tag{18}$$

a formula that represents a dramatic improvement over (16), allowing a much more efficient calculation of N_p for example. The guess in (18) is in fact a consequence of Gauss's celebrated law of quadratic reciprocity:

Theorem (Quadratic reciprocity). *For any equation ax^2+bx+c, with $\Delta := b^2-4ac$, the value of the function $p \mapsto N_p$ (for $p \nmid a\Delta$) depends only on the residue class of p modulo 4Δ and hence is periodic with period length dividing $4|\Delta|$.*

The repeating pattern satisfied by the N_p's as p varies greatly facilitates the manipulation of the zeta functions of quadratic equations. For example, the zeta function of $\mathcal{X} : x^2 - x - 1 = 0$ is equal to

$$\zeta(\mathcal{X}; s) = \zeta(s) \times \left\{ \left(1 - \frac{1}{2^s} - \frac{1}{3^s} + \frac{1}{4^s}\right) + \left(\frac{1}{6^s} - \frac{1}{7^s} - \frac{1}{8^s} + \frac{1}{9^s}\right) + \left(1\frac{1}{11^s} - \frac{1}{12^s} - \frac{1}{13^s} + \frac{1}{14^s}\right) + \cdots \right\}. \tag{19}$$

The series that occurs on the right-hand side is a prototypical example of a *Dirichlet L-series*. These L-series, which are the key actors in the proof of Dirichlet's theorem on the infinitude of primes in arithmetic progressions, enjoy many of the same analytic properties as the Riemann zeta function: an analytic continuation to the entire complex plane and a functional equation relating their values at s and $1-s$. They are also expected to satisfy a Riemann hypothesis which generalises Riemann's original statement and is just as deep and elusive.

It is a (not completely trivial) fact that the zeta function of the general quadratic equation in n variables

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c = 0 \tag{20}$$

involves the same basic constituents, Dirichlet series, as in the one-variable case. This means that quadratic diophantine equations in any number of variables are well understood, at least as far as their zeta functions are concerned.

The plot thickens when equations of higher degree are considered. Consider for instance the cubic equation $x^3 - x - 1$ of discriminant $\Delta = -23$. For all $p \neq 23$, this cubic equation has no multiple roots over $\mathbb{F}_{p^r}$, and therefore $N_p = 0$, 1, or 3. A simple expression for N_p in this case is given by the following theorem of Hecke:

Theorem (Hecke). *The following hold for all primes $p \neq 23$:*

(1) *If p is not a square modulo 23, then $N_p = 1$.*
(2) *If p is a square modulo 23, then*

$$N_p = \begin{cases} 0 & \text{if } p = 2a^2 + ab + 3b^2, \\ 3 & \text{if } p = a^2 + ab + 6b^2, \end{cases} \tag{21}$$

for some $a, b \in \mathbb{Z}$.

Hecke's theorem implies that

$$\zeta(x^3 - x - 1; s) \quad = \quad \zeta(s) \times \sum_{n=1}^{\infty} a_n n^{-s}, \tag{22}$$

where the generating series

$$F(q) := \sum a_n q^n = q - q^2 - q^3 + q^6 + q^8 - q^{13} - q^{16} + q^{23} + \cdots \tag{23}$$

is given by the explicit formula

$$F(q) = \frac{1}{2}\left(\sum_{a,b \in \mathbb{Z}} q^{a^2+ab+6b^2} - q^{2a^2+ab+3b^2} \right). \tag{24}$$

The function $f(z) = F(e^{2\pi i z})$ that arises by setting $q = e^{2\pi i z}$ in (24) is a prototypical example of a *modular form*: namely, it satisfies the transformation rule

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d) f(z), \quad \begin{cases} a, b, c, d \in \mathbb{Z}, & ad - bc = 1 \\ 23 | c, & \left(\frac{a}{23}\right) = 1. \end{cases} \tag{25}$$

under so-called *modular substitutions* of the form $z \mapsto \frac{az+b}{cz+d}$. This property follows from the *Poisson summation formula* applied to the expression in (24). Thanks to (25), the zeta function of $\mathcal{X}$ can be manipulated with the same ease as the zeta functions of Riemann and Dirichlet. Indeed, Hecke showed that the L-series $\sum_{n=1}^{\infty} a_n n^{-s}$ attached to a modular form $\sum_{n=1}^{\infty} a_n e^{2\pi i n z}$ possesses very similar analytic properties, notably an analytic continuation and a Riemann-style functional equation.

The generating series $F(q)$ can also be expressed as an infinite product:

$$\frac{1}{2}\left(\sum_{a,b \in \mathbb{Z}} q^{a^2+ab+6b^2} - q^{2a^2+ab+3b^2} \right) = q \prod_{n=1}^{\infty} (1-q^n)(1-q^{23n}). \tag{26}$$

The first few terms of this power series identity can readily be verified numerically, but its proof is highly nonobvious and indirect. It exploits the circumstance that the space of holomorphic functions of z satisfying the transformation rules (25) together with suitable growth properties is a one-dimensional complex vector space which also contains

the infinite product above. Indeed, the latter is equal to $\eta(q)\eta(q^{23})$, where

$$\eta(q) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n) \tag{27}$$

is the Dedekind eta function whose logarithmic derivative (after viewing η as a function of z through the change of variables $q = e^{2\pi iz}$) is given by

$$\frac{\eta'(z)}{\eta(z)} = -\pi i\left(\frac{-1}{12} + 2\sum_{n=1}^{\infty}\Big(\sum_{d|n} d\Big)e^{2\pi inz}\right) \tag{28}$$

$$= \frac{i}{4\pi}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\frac{1}{(mz+n)^2}, \tag{29}$$

where the term attached to $(m,n) = (0,0)$ is excluded from the last sum. The Dedekind η-function is also connected to the generating series for the partition function $p(n)$ describing the number of ways in which n can be expressed as a sum of positive integers via the identity

$$\eta^{-1}(q) = q^{-1/24}\sum_{n=0}^{\infty} p(n)q^n, \tag{30}$$

which plays a starring role alongside Jeremy Irons and Dev Patel in a recent film about the life of Srinivasa Ramanujan.

"There are five elementary arithmetical operations: addition, subtraction, multiplication, division,…and modular forms."

Commenting on the "unreasonable effectiveness and ubiquity of modular forms," Martin Eichler once wrote, "There are five elementary arithmetical operations: addition, subtraction, multiplication, division,…and modular forms." Equations (26), (29), and (30) are just a few of the many wondrous identities which abound, like exotic strains of fragrant wild orchids, in what Roger Godement has called the "garden of modular delights."

The example above and many others of a similar type are described in Jean-Pierre Serre's delightful monograph [Se], touching on themes that were also covered in Serre's lecture at the inaugural Abel Prize ceremony in 2003.

Hecke was able to establish that all cubic polynomials in one variable are *modular*; i.e., the coefficients of their zeta functions obey patterns just like those of (24) and (25). Wiles's achievement was to extend this result to a large class of cubic diophantine equations in two variables over the rational numbers: the *elliptic curve* equations which can be brought to the form

$$y^2 = x^3 + ax + b \tag{31}$$

after a suitable change of variables and which are nonsingular, a condition equivalent to the assertion that the *discriminant* $\Delta := -16(4a^3 + 27b^2)$ is nonzero.

To illustrate Wiles's theorem with a concrete example, consider the equation

$$E : y^2 = x^3 - x, \tag{32}$$

of discriminant $\Delta = 64$. After setting

$$\zeta(E;s) = \zeta(s-1)\times(a_1 + a_2 2^{-s} + a_3 3^{-s} + a_4 4^{-s} + \cdots)^{-1}, \tag{33}$$

the associated generating series satisfies the following identities reminiscent of (24) and (26):

$$F(q) = \sum a_n q^n = q - 2q^5 - 3q^9 + 6q^{13} + 2q^{17} - q^{25} + \cdots \tag{34}$$

$$= \sum_{a,b} a\cdot q^{(a^2+b^2)} \tag{35}$$

$$= q\prod_{n=1}^{\infty}(1-q^{4n})^2(1-q^{8n})^2, \tag{36}$$

where the sum in (35) runs over the $(a,b)\in\mathbb{Z}^2$ for which the Gaussian integer $a+bi$ is congruent to 1 modulo $(1+i)^3$. (This identity follows from Deuring's study of zeta functions of elliptic curves *with complex multiplication* and may even have been known earlier.) Once again, the holomorphic function $f(z) := F(e^{2\pi iz})$ is a modular form satisfying the slightly different transformation rule

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 f(z),\ \begin{cases} a,b,c,d\in\mathbb{Z},\ ad-bc=1,\\ 32|c. \end{cases} \tag{37}$$

Note the exponent 2 that appears in this formula. Because of it, the function $f(z)$ is said to be a *modular form of weight* 2 and level 32. The modular forms of (25) attached to cubic equations in one variable are of weight 1, but otherwise the parallel of (35) and (36) with (24) and (26) is striking. The original conjecture of Shimura–Taniyama, and Weil asserts the same pattern for all elliptic curves:

Conjecture (Shimura–Taniyama–Weil). *Let E be any elliptic curve. Then*

$$\zeta(E;s) = \zeta(s-1)\times\left(\sum_{n=1}^{\infty} a_n n^{-s}\right)^{-1}, \tag{38}$$

where $f_E(z) := \sum a_n e^{2\pi inz}$ is a modular form of weight 2.

The conjecture was actually more precise and predicted that the level of f_E—i.e., the integer that appears in the transformation property for f_E as the integers 23 and 32 in (25) and (37) respectively—is equal to the *arithmetic conductor* of E. This conductor, which is divisible only by primes for which the equation defining E becomes singular modulo p, is a measure of the arithmetic complexity of E and can be calculated explicitly from an equation for E by an algorithm of Tate. An elliptic curve is said to be *semistable* if its arithmetic conductor is squarefree. This class of elliptic curves includes those of the form

$$y^2 = x(x-a)(x-b) \tag{39}$$

with $\gcd(a,b) = 1$ and $16|b$. The most famous elliptic curves in this class are those that ultimately do not exist:

Andrew Wiles, Henri Darmon, and Mirela Çiperiani in June 2016 at Harvard University during a conference in honor of Karl Rubin's sixtieth birthday.

the "Frey curves" $y^2 = x(x - a^p)(x + b^p)$ arising from putative solutions to Fermat's equation $a^p + b^p = c^p$, whose nonexistence had previously been established in a landmark article of Kenneth Ribet[1] under the assumption of their modularity. It is the proof of the Shimura–Taniyama–Weil conjecture for semistable elliptic curves that earned Andrew Wiles the Abel Prize:

Theorem (Wiles). *Let E be a semistable elliptic curve. Then E satisfies the Shimura–Taniyama–Weil conjecture.*

The semistability assumption in Wiles's theorem was later removed by Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor around 1999. (See, for instance, the account [Da] that appeared in the *Notices* at the time.)

As a prelude to describing some of the important ideas in its proof, one must first try to explain why Wiles's theorem occupies such a central position in mathematics. The Langlands program places it in a larger context by offering a vast generalisation of what it means for a diophantine equation to be "associated to a modular form." The key is to view modular forms attached to cubic equations or to elliptic curves as in (24) or (34) as vectors in certain irreducible infinite-dimensional representations of the locally compact topological group

$$\mathbf{GL}_2(\mathbb{A}_{\mathbb{Q}}) = \prod{}'_p \mathbf{GL}_2(\mathbb{Q}_p) \times \mathbf{GL}_2(\mathbb{R}) \tag{40}$$

(where $\prod'_p$ denotes a restricted direct product over all the prime numbers consisting of elements $(\gamma_p)_p$ for which the pth component γ_p belongs to the maximal compact subgroup $\mathbf{GL}_2(\mathbb{Z}_p)$ for all but finitely many p). The shift in emphasis from modular forms to the so-called *automorphic representations* which they span is decisive. Langlands showed how to attach an L-function to any irreducible automorphic representation of $G(\mathbb{A}_{\mathbb{Q}})$ for an arbitrary reductive algebraic group G, of which the matrix groups $\mathbf{GL}_n$ and more general algebraic groups of Lie type are prototypical examples. This greatly enlarges the notion of what it means to be "modular": a diophantine equation is now said to have this property if its zeta function can be expressed in terms of the Langlands L-functions attached to automorphic representations. One of the fundamental goals in the Langlands program is to establish further cases of the following conjecture:

Conjecture. *All diophantine equations are modular in the above sense.*

This conjecture can be viewed as a far-reaching generalisation of quadratic reciprocity and underlies the non-Abelian reciprocity laws that are at the heart of Andrew Wiles's achievement.

Before Wiles's proof, the following general classes of diophantine equations were known to be modular:

- Quadratic equations, by Gauss's law of quadratic reciprocity;
- Cubic equations in one variable, by the work of Hecke and Maass;
- Quartic equations in one variable.

This last case deserves further comment, since it has not been discussed previously and plays a primordial role in Wiles's proof. The modularity of quartic equations follows from the seminal work of Langlands and Tunnell. While it is beyond the scope of this survey to describe their methods, it must be emphasised that Langlands and Tunnell make essential use of the *solvability by radicals* of the general quartic equation, whose underlying symmetry group is contained in the permutation group S_4 on 4 letters. Solvable extensions are obtained from a succession of Abelian extensions, which fall within the purview of the class field theory developed in the nineteenth and first half of the twentieth century. On the other hand, the modularity of the general equation of degree > 4 in one variable, which cannot be solved by radicals, seemed to lie well beyond the scope of the techniques that were available in the "pre-Wiles era." The reader who perseveres to the end of this essay will be given a glimpse of how our knowledge of the modularity of the general quintic equation has progressed dramatically in the wake of Wiles's breakthrough.

Prior to Wiles's proof, modularity was also not known for any interesting general class of equations (of degree > 2, say) in more than one variable; in particular it had only been verified for finitely many elliptic curves over $\mathbb{Q}$ up to isomorphism over $\bar{\mathbb{Q}}$ (including the elliptic curves over $\mathbb{Q}$ with complex multiplication, of which the elliptic curve of (31) is an instance.) Wiles's modularity theorem confirmed the Langlands conjectures in the important test case of elliptic curves, which may seem like (and, in fact, are) very special diophantine equations, but have provided a fertile terrain for arithmetic investigations,

[1] *See the interview with Ribet as the new AMS president in this issue, page 229.*

both in theory and in applications (cryptography, coding theory…).

Wiles's proof is also important for having introduced a revolutionary new approach which has opened the floodgates for many further breakthroughs in the Langlands program.

Returning to the main theme of this report, Wiles's proof is also important for having introduced a revolutionary new approach which has opened the floodgates for many further breakthroughs in the Langlands program.

To expand on this point, we need to present another of the *dramatis personae* in Wiles's proof: *Galois representations*. Let $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ be the absolute Galois group of $\mathbb{Q}$, namely, the automorphism group of the field of all algebraic numbers. It is a profinite group, endowed with a natural topology for which the subgroups $\mathrm{Gal}(\bar{\mathbb{Q}}/L)$ with L ranging over the finite extensions of $\mathbb{Q}$ form a basis of open subgroups. Following the original point of view taken by Galois himself, the group $G_{\mathbb{Q}}$ acts naturally as permutations on the roots of polynomials with rational coefficients. Given a finite set S of primes, one may consider only the monic polynomials with integer coefficients whose discriminant is divisible only by primes $\ell \in S$ (eventually after a change of variables). The topological group $G_{\mathbb{Q}}$ operates on the roots of such polynomials through a quotient, denoted $G_{\mathbb{Q},S}$—the automorphism group of the maximal algebraic extension *unramified* outside S, which can be regarded as the symmetry group of all the zero-dimensional varieties over $\mathbb{Q}$ having "nonsingular reduction outside S."

In addition to the permutation representations of $G_{\mathbb{Q}}$ that were so essential in Galois's original formulation of his theory, it has become important to study the (continuous) *linear* representations

$$\varrho : G_{\mathbb{Q},S} \longrightarrow GL_n(L) \tag{41}$$

of this Galois group, where L is a complete field, such as the fields $\mathbb{R}$ or $\mathbb{C}$ of real or complex numbers, the finite field $\mathbb{F}_{\ell^r}$ equipped with the discrete topology, or a finite extension $L \subset \bar{\mathbb{Q}}_\ell$ of the field $\mathbb{Q}_\ell$ of ℓ-adic numbers.

Galois representations were an important theme in the work of Abel and remain central in modern times. Many illustrious mathematicians in the twentieth century have contributed to their study, including three former Abel Prize winners: Jean-Pierre Serre, John Tate, and Pierre Deligne. Working on Galois representations might seem to be a prerequisite for an algebraic number theorist to receive the Abel Prize!

Like diophantine equations, Galois representations also give rise to analogous zeta functions. More precisely, the group $G_{\mathbb{Q},S}$ contains, for each prime $p \notin S$, a distinguished element called the *Frobenius element* at p, denoted σ_p. Strictly speaking, this element is well defined only up to conjugacy in $G_{\mathbb{Q},S}$, but this is enough to make the arithmetic sequence

$$N_{p^r}(\varrho) := \mathrm{Trace}(\varrho(\sigma_p^r)) \tag{42}$$

well defined. The zeta function $\zeta(\varrho; s)$ packages the information from this sequence in exactly the same way as in the definition of $\zeta(\mathcal{X}; s)$.

For example, if $\mathcal{X}$ is attached to a polynomial P of degree d in one variable, the action of $G_{\mathbb{Q},S}$ on the roots of P gives rise to a d-dimensional permutation representation

$$\varrho_{\mathcal{X}} : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_d(\mathbb{Q}), \tag{43}$$

and $\zeta(\mathcal{X}, s) = \zeta(\varrho_{\mathcal{X}}, s)$. This connection goes far deeper, extending to diophantine equations in $n+1$ variables for general $n \geq 0$. The glorious insight at the origin of the Weil conjectures is that $\zeta(\mathcal{X}; s)$ can be expressed in terms of the zeta functions of Galois representations arising in the *étale cohomology* of $\mathcal{X}$, a cohomology theory with ℓ-adic coefficients which associates to $\mathcal{X}$ a collection

$$\left\{H^i_{\mathrm{et}}(\mathcal{X}_{/\bar{\mathbb{Q}}}, \mathbb{Q}_\ell)\right\}_{0 \leq i \leq 2n}$$

of finite-dimensional $\mathbb{Q}_\ell$-vector spaces endowed with a continuous linear action of $G_{\mathbb{Q},S}$. (Here S is the set of primes q consisting of ℓ and the primes for which the equation of $\mathcal{X}$ becomes singular after being reduced modulo q.) These representations generalise the representation $\varrho_{\mathcal{X}}$ of (43), insofar as the latter is realised by the action of $G_{\mathbb{Q},S}$ on $H^0_{\mathrm{et}}(\mathcal{X}_{\bar{\mathbb{Q}}}, \mathbb{Q}_\ell)$ after extending the coefficients from $\mathbb{Q}$ to $\mathbb{Q}_\ell$.

Theorem (Weil, Grothendieck,…). *If $\mathcal{X}$ is a diophantine equation having good reduction outside S, there exist Galois representations ϱ_1 and ϱ_2 of $G_{\mathbb{Q},S}$ for which*

$$\zeta(\mathcal{X}; s) = \zeta(\varrho_1; s)/\zeta(\varrho_2; s). \tag{44}$$

The representations ϱ_1 and ϱ_2 occur in $\oplus H^i_{\mathrm{et}}(\mathcal{X}_{/\bar{\mathbb{Q}}}, \mathbb{Q}_\ell)$, where the direct sum ranges over the odd and even values of $0 \leq i \leq 2n$ for ϱ_1 and ϱ_2 respectively. More canonically, there are always *irreducible* representations $\varrho_1, \ldots, \varrho_t$ of $G_{\mathbb{Q},s}$ and integers $d_1, \ldots d_t$ such that

$$\zeta(\mathcal{X}; s) = \prod_{i=1}^{t} \zeta(\varrho_i; s)^{d_i}, \tag{45}$$

arising from the decompositions of the (semisimplification of the) $H^i_{\mathrm{et}}(\mathcal{X}_{\bar{\mathbb{Q}}}, \mathbb{Q}_\ell)$ into a sum of irreducible representations. The $\zeta(\varrho_i, s)$ can be viewed as the "atomic constituents" of $\zeta(\mathcal{X}, s)$ and reveal much of the "hidden structure" in the underlying equation. The decomposition of $\zeta(\mathcal{X}; s)$ into a product of different $\zeta(\varrho_i; s)$ is not unlike the decomposition of a wave function into its simple harmonics.

A Galois representation is said to be *modular* if its zeta function can be expressed in terms of generating series attached to modular forms and automorphic representations and is said to be *geometric* if it can be realised in

an étale cohomology group of a diophantine equation as above. The "main conjecture of the Langlands program" can now be amended as follows:

Conjecture. *All geometric Galois representations of $G_{\mathbb{Q},S}$ are modular.*

Given a Galois representation

$$\varrho : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_n(\mathbb{Z}_\ell) \tag{46}$$

with ℓ-adic coefficients, one may consider the resulting mod ℓ representation

$$\bar{\varrho} : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_n(\mathbb{F}_\ell). \tag{47}$$

The passage from ϱ to $\bar{\varrho}$ amounts to replacing the quantities $N_{p^r}(\varrho) \in \mathbb{Z}_\ell$ as p^r ranges over all the prime powers with their mod ℓ reduction. Such a passage would seem rather contrived for the sequences $N_{p^r}(X)$—why study the solution counts of a diophantine equation over different finite fields, taken modulo ℓ?—if one did not know a priori that these counts arise from ℓ-adic Galois representations with coefficients in $\mathbb{Z}_\ell$. There is a corresponding notion of what it means for $\bar{\varrho}$ to be modular, namely, that the data of $N_{p^r}(\bar{\varrho})$ agrees, very loosely speaking, with the mod ℓ reduction of similar data arising from an automorphic representation. We can now state Wiles's celebrated *modularity lifting theorem*, which lies at the heart of his strategy:

Since then, "modularity lifting theorems" have proliferated, and their study, in ever more general and delicate settings, has spawned an industry.

Wiles's Modularity Lifting Theorem. *Let*

$$\varrho : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_2(\mathbb{Z}_\ell) \tag{48}$$

be an irreducible geometric Galois representation satisfying a few technical conditions (involving, for the most part, the restriction of ϱ to the subgroup $G_{\mathbb{Q}_\ell} = Gal(\bar{\mathbb{Q}}_\ell/\mathbb{Q}_\ell)$ of $G_{\mathbb{Q},S}$). If $\bar{\varrho}$ is modular and irreducible, then so is ϱ.

This stunning result was completely new at the time: nothing remotely like it had ever been proved before! Since then, "modularity lifting theorems" have proliferated, and their study, in ever more general and delicate settings, has spawned an industry and led to a plethora of fundamental advances in the Langlands program.

Let us first explain how Wiles himself parlays his original modularity lifting theorem into a proof of the Shimura–Taniyama-Weil conjecture for semistable elliptic curves. Given such an elliptic curve E, consider the groups

$$E[3^n] := \{P \in E(\bar{\mathbb{Q}}) \ : \ 3^n P = 0\}, \qquad T_3(E) := \lim_{\leftarrow} E[3^n], \tag{49}$$

the inverse limit being taken relative to the multiplication-by-3 maps. The groups $E[3^n]$ and $T_3(E)$ are free modules of rank 2 over $(\mathbb{Z}/3^n\mathbb{Z})$ and $\mathbb{Z}_3$ respectively and are endowed with continuous linear actions of $G_{\mathbb{Q},S}$, where S is a set of primes containing 3 and the primes that divide the conductor of E. One obtains the associated Galois representations:

$$\begin{aligned} &\bar{\varrho}_{E,3} : G_{\mathbb{Q},S} \longrightarrow \mathrm{Aut}(E[3]) \simeq \mathbf{GL}_2(\mathbb{F}_3), \\ &\varrho_{E,3} : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_2(\mathbb{Z}_3). \end{aligned} \tag{50}$$

The theorem of Langlands and Tunnell about the modularity of the general quartic equation leads to the conclusion that $\bar{\varrho}_{E,3}$ is modular. This rests on the happy circumstance that

$$\mathbf{GL}_2(\mathbb{F}_3)/\langle \pm 1 \rangle \simeq S_4, \tag{51}$$

and hence that $E[3]$ has essentially the same symmetry group as the general quartic equation! The isomorphism in (51) can be realised by considering the action of $\mathbf{GL}_2(\mathbb{F}_3)$ on the set $\{0, 1, 2, \infty\}$ of points on the projective line over $\mathbb{F}_3$.

If E is semistable, Wiles is able to check that both $\varrho_{E,3}$ and $\bar{\varrho}_{E,3}$ satisfy the conditions necessary to apply the Modularity Lifting Theorem, at least when $\bar{\varrho}_{E,3}$ is *irreducible*. It then follows that $\varrho_{E,3}$ is modular, and therefore so is E, since $\zeta(E;s)$ and $\zeta(\varrho_{E,3};s)$ are the same.

Note the key role played by the result of Langlands–Tunnell in the above strategy. It is a dramatic illustration fo the unity and historical continuity of mathematics that the solution in radicals of the general quartic equation, one of the great feats of the algebraists of the Italian Renaissance, is precisely what allowed Langlands, Tunnell, and Wiles to prove their modularity results more than five centuries later.

Having established the modularity of all semistable elliptic curves E for which $\bar{\varrho}_{E,3}$ is irreducible, Wiles disposes of the others by applying his lifting theorem to the prime $\ell = 5$ instead of $\ell = 3$. The Galois representation $\bar{\varrho}_{E,5}$ is always irreducible in this setting, because no elliptic curve over $\mathbb{Q}$ can have a rational subgroup of order 15. Nonetheless, the approach of exploiting $\ell = 5$ seems hopeless at first glance, because the Galois representation $E[5]$ is not known to be modular a priori, for much the same reason that the general quintic equation cannot be solved by radicals. (Indeed, the symmetry group $\mathbf{SL}_2(\mathbb{F}_5)$ is a double cover of the alternating group A_5 on 5 letters and thus is closely related to the symmetry group underlying the general quintic.) To establish the modularity of $E[5]$, Wiles constructs an auxiliary semistable elliptic curve E' satisfying

$$\bar{\varrho}_{E',5} = \bar{\varrho}_{E,5}, \qquad \bar{\varrho}_{E',3} \text{ is irreducible.} \tag{52}$$

It then follows from the argument in the previous paragraph that E' is modular, hence that $E'[5] = E[5]$ is modular as well, putting E within striking range of the modularity lifting theorem with $\ell = 5$. This lovely epilogue of Wiles's proof, which came to be known as the "3-5 switch," may have been viewed as an expedient trick at the time. But since then the prime switching argument has become firmly embedded in the subject, and many

variants of it have been exploited to spectacular effect in deriving new modularity results.

The modularity of elliptic curves was only the first in a series of spectacular applications.

Wiles's modularity lifting theorem reveals that "modularity is contagious" and can often be passed on to an ℓ-adic Galois representation from its mod ℓ reduction. It is this simple principle that accounts for the tremendous impact that the Modularity Lifting Theorem, and the many variants proved since then, continue to have on the subject. Indeed, the modularity of elliptic curves was only the first in a series of spectacular applications of the ideas introduced by Wiles, and since 1994 the subject has witnessed a real golden age, in which open problems that previously seemed completely out of reach have succumbed one by one.

Among these developments, let us mention:

- The two-dimensional Artin conjecture, first formulated in 1923, concerns the modularity of all odd, two-dimensional Galois representations

$$\varrho : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_2(\mathbb{C}). \tag{53}$$

  The image of such a ϱ modulo the scalar matrices is isomorphic either to a dihedral group, to A_4, to S_4, or to A_5. Thanks to the earlier work of Hecke, Langlands, and Tunnell, only the case of projective image A_5 remained to be disposed of. Many new cases of the two-dimensional Artin conjecture were proved in this setting by Kevin Buzzard, Mark Dickinson, Nick Shepherd-Barron, and Richard Taylor around 2003 using the modularity of all mod 5 Galois representations arising from elliptic curves as a starting point.
- Serre's conjecture, which was formulated in 1987, asserts the modularity of all odd, two-dimensional Galois representations

$$\varrho : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_2(\mathbb{F}_{p^r}), \tag{54}$$

  with coefficients in a finite field. This result was proved by Chandrasekhar Khare and Jean-Pierre Wintenberger in 2008 by a glorious extension of the "3-5 switching technique" in which essentially all the primes are used. (See Khare's report in this volume.) This result also implies the two-dimensional Artin conjecture in the general case.
- The two-dimensional Fontaine–Mazur conjecture concerning the modularity of odd, two-dimensional p-adic Galois representations

$$\varrho : G_{\mathbb{Q},S} \longrightarrow \mathbf{GL}_2(\bar{\mathbb{Q}}_p) \tag{55}$$

  satisfying certain technical conditions with respect to their restrictions to the Galois group of $\mathbb{Q}_p$. This theorem was proved in many cases as a consequence of work of Pierre Colmez, Matthew Emerton, and Mark Kisin.
- The Sato–Tate conjecture concerning the distribution of the numbers $N_p(E)$ for an elliptic curve E as the prime p varies, whose proof was known to follow from the modularity of all the symmetric power Galois representations attached to E, was proved in large part by Laurent Clozel, Michael Harris, Nick Shepherd-Barron, and Richard Taylor around 2006.
- One can also make sense of what it should mean for diophantine equations over more general number fields to be modular. The modularity of elliptic curves over all real quadratic fields has been proved very recently by Nuno Freitas, Bao Le Hung, and Samir Siksek by combining the ever more general and powerful modularity lifting theorems currently available with a careful diophantine study of the elliptic curves which could a priori fall outside the scope of these lifting theorems.
- Among the spectacular recent developments building on Wiles's ideas is the proof, by Laurent Clozel and Jack Thorne, of the modularity of certain symmetric powers of the Galois representations attached to holomorphic modular forms, which is described in Thorne's contribution to this volume.

These results are just a sampling of the transformative impact of modularity lifting theorems. The Langlands program remains a lively area, with many alluring mysteries yet to be explored. It is hard to predict where the next breakthroughs will come, but surely they will continue to capitalise on the rich legacy of Andrew Wiles's marvelous proof.

References

[Da] H. Darmon, A proof of the full Shimura-Taniyama-Weil conjecture is announced, *Notices of the AMS* **46** (1999), no. 11, 1397–1401. MR1723249

[Se] J-P. Serre, *Lectures on* $N_X(p)$, Chapman & Hall/CRC Research Notes in Mathematics, 11, CRC Press, Boca Raton, FL. MR2920749

Photo Credits

Photo of Wiles giving his first lecture is by C. J. Mozzochi, courtesy of the Simons Foundation.

Photo of the conference in honor of Karl Rubin's sixtieth birthday is courtesy of Kartik Prasanna.

ABOUT THE AUTHOR

Henri Darmon received the 2017 AMS Cole Prize in Number Theory and the 2017 CRM-Fields-PIMS Prize for his contributions to the arithmetic of elliptic curves and modular forms.

The Mathematical Works of Andrew Wiles

Christopher Skinner, with contributions from Karl Rubin, Barry Mazur, Mirela Çiperiani, Chandrashekhar Khare, and Jack Thorne

Sir Andrew J. Wiles was awarded the Abel Prize for 2016 for "his stunning proof of Fermat's Last Theorem by way of the modularity conjecture for semistable elliptic curves, opening a new era in number theory."[1] Andrew Wiles announced his proof of Fermat's Last Theorem in June of 1993 in a series of three lectures at a conference at the Isaac Newton Institute for Mathematical Sciences at the University of Cambridge. Overnight Wiles and his proof became an international media sensation, making headlines in papers around the world. The story of this proof—the subsequent discovery of a gap and its ultimate and beautiful completion in September of 1994—has entered into popular legend.[2] The surprising drama of the proof is told in the 1996 BBC Horizon documentary *Fermat's Last Theorem*, directed by Simon Singh, which ably conveys the human side of what is often seen as the distant and rarefied world of mathematical research.[3]

All this is well known. What is less well known, possibly even among number theorists, is that before his proof of Fermat's Last Theorem, Wiles had made significant contributions to two of the most important problems for late-twentieth-century number theory:

Andrew Wiles with his PhD advisor, John Coates (left); his first PhD student Karl Rubin (right); and Ken Ribet (second from right), whose work linking the Modularity Conjecture to Fermat's Last Theorem inspired Wiles to reengage with the problem that had fascinated him since childhood. This photo was taken at the Newton Institute in Cambridge, UK, during the conference at which Wiles first announced his proof of the Modularity Conjecture.

- Wiles proved, together with John Coates, the first theoretical evidence for the celebrated Birch–Swinnerton-Dyer Conjecture; this is now known as the Coates–Wiles Theorem.
- Wiles proved Iwasawa's Main Conjecture for $\mathbb{Q}$, in joint work with Barry Mazur, and for all totally real fields.

Each of these is a landmark result on its own and would be considered the highlight of a distinguished career. Two of the following contributions describe these works and their proofs. Karl Rubin, Wiles's first PhD student, writes about the Coates–Wiles Theorem. Barry Mazur, Wiles's collaborator on his first proof of Iwasawa's Main Conjecture for $\mathbb{Q}$, writes about Wiles's work on the Main Conjectures. Anyone seeking to learn about the context, significance, and ideas of Wiles's proof of Fermat's Last Theorem can do no better than to read Henri Darmon's Abel Prize lecture in this same issue of the *Notices*.

Christopher Skinner is professor of mathematics at Princeton University. His e-mail address is cmcls@math.princeton.edu.

[1] *Citation by the Abel Prize Committee of the Norwegian Academy of Science and Letters for the 2016 Abel Prize Laureate:* www.abelprize.no/c67107/binfil/download.php?tid=67059

[2] *No doubt many number theorists share my own experiences of striking up conversations with strangers, who, upon discovering that I am a mathematician and even a number theorist, ask about "that guy who solved that famous problem—the one who worked in his attic for seven years."*

[3] *I have watched this documentary many times with groups of mathematically talented high school students from around the world. Twenty years later it still inspires questions and conversation about what it means to do mathematical research or what a life spent doing mathematics can be.*

DOI: http://dx.doi.org/10.1090/noti1485

What are you working on now? This is a question that eminent mathematicians are frequently asked (or so I am reliably informed) and certainly one that Andrew Wiles has repeatedly faced in the years following his proof of Fermat's Last Theorem. In her contribution below, Mirela Çiperiani writes about her collaboration with Andrew Wiles, from the mid-2000s, on a very natural Diophantine question: does a genus one curve over $\mathbb{Q}$ have a rational point over a solvable extension of $\mathbb{Q}$? As Çiperiani explains, results and techniques arising from the proof of Fermat's Last Theorem also play a role in this work.

What distinguishes a great mathematical proof? There is, of course, no definitive answer to this question. Certainly proofs of famous or important open conjectures can lay claim to being great. By this measure, Andrew Wiles's proof of Fermat's Last Theorem is a truly great proof. But proofs that introduce new ideas or open doors to progress on problems that were previously viewed as out of reach also have their claim to greatness. As noted in the Abel Prize citation, Andew Wiles's proofs also achieve greatness by this second measure.

The new techniques and ideas that led to a successful proof of the modularity of semistable elliptic curves have been remarkably robust, also leading to the resolution of a host of problems within the circle of the Langlands Program: proofs of Serre's conjecture, the Artin conjecture for odd two-dimensional representations, the Sato–Tate conjecture, meromorphic continuation of Hasse–Weil L-functions of elliptic curves over totally real fields, modularity of all elliptic curves over real quadratic fields, and automorphy of small symmetric powers of modular forms, to list just a few.

Much of this progress has been achieved by the efforts of Richard Taylor, who was a PhD student of Wiles and later a coauthor of the paper that introduced one of the key ingredients[4] of the proof of modularity; by Wiles's PhD students; and by Taylor's collaborators and PhD students. But Wiles's work also inspired many others to push his ideas further. Among these is Chandrashekhar Khare, who writes below about some of the progress on modularity of two-dimensional Galois representations that followed Wiles's proof and of the use of this in Khare's own proof, with Jean-Pierre Wintenberger, of Serre's conjecture. In a related contribution, Jack Thorne describes how Wiles's original techniques evolved and were adapted to proving the automorphy of higher-dimensional Galois representations, leading to his proof, with Laurent Clozel, of the automorphy of some small symmetric powers of a holomorphic modular form.

But we should not lose sight of Wiles's earlier contributions in the glow of the successes arising from the proof of Fermat's Last Theorem. The proof of the Coates–Wiles Theorem and Wiles's proof of Iwasawa's Main Conjecture for totally real fields have inspired similar progress. For example, Kazuya Kato's (2004) spectacular success in constructing an Euler system for elliptic curves and then relating it to the special values of the Hasse–Weil L-function of the curve via an explicit reciprocity law can be seen as a vast generalization of the ideas in the proof of the Coates–Wiles theorem. My own work with Eric Urban (2014), which together with Kato's result proves much of the Main Conjecture in the Iwaswa theory of elliptic curves, is in large part the natural generalization to some unitary groups of Wiles's methods for proving the Iwasawa Main Conjecture for totally real fields. Andrew Wiles's ideas continue to inform and shape progress on some of the fundamental problems in algebraic number theory.

[4] *Now generally known as the Taylor–Wiles method.*

Karl Rubin

Wiles's Work on Elliptic Curves with Complex Multiplication

During my senior year at Princeton, 1975–76, there was a lot of buzz in the common room about something everyone referred to as "Coates–Wiles." I had a vague idea of what an elliptic curve was, but I doubt that I had any clear idea of what John Coates and Andrew Wiles had done. But I could tell that it was important, and I had a mental image of two very senior mathematicians who had made a great breakthrough.

I arrived at Harvard as a graduate student in 1976. Wiles arrived a year later, and I discovered that this "very senior" mathematician was scarcely older than I was and had been a graduate student at the time of this spectacular result. I became his student and had the unusual experience of attending my advisor's thesis defense.

Here's what the excitement was about. An elliptic curve E over the field $\mathbb{Q}$ of rational numbers is a curve defined by an equation $y^2 = x^3 + ax + b$, with $a, b \in \mathbb{Q}$ satisfying $4a^3 + 27b^2 \neq 0$. It is classical that the rational points $E(F)$ on E over any field F containing $\mathbb{Q}$ form an abelian group, and Mordell proved that $E(\mathbb{Q})$ is finitely generated. The *rank* of E is the dimension of the $\mathbb{Q}$-vector space $E(\mathbb{Q}) \otimes_{\mathbb{Z}} \mathbb{Q}$.

In the late 1950s, after extensive computations, Birch and Swinnerton-Dyer made the following conjecture.

Conjecture (Birch & Swinnerton-Dyer)**.** *For every elliptic curve E over $\mathbb{Q}$, we have*

$$\text{(BSD)} \qquad \operatorname{rank}(E(\mathbb{Q})) = \operatorname{ord}_{s=1} L(E,s).$$

Here $L(E,s)$ is the Hasse–Weil L-function of E, defined by an Euler product that converges on the complex half-plane $\Re(s) > 3/2$, and $\operatorname{ord}_{s=1} L(E,s)$ is its order of vanishing at $s = 1$. This conjecture is still unproved and is one of the Clay Millennium Problems.

Clearly, to make progress on this conjecture, or even for the statement to make sense, one needs to know that $L(E,s)$ has an analytic continuation at least to $s = 1$. This was already known, thanks to a theorem of Deuring, for elliptic curves with *complex multiplication*: we say that E has complex multiplication if the ring of endomorphisms $\operatorname{End}(E)$ is larger than $\mathbb{Z}$, in which case $\operatorname{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$ is an

Karl Rubin is Edward and Vivian Thorp Professor of Mathematics at UC Irvine. His e-mail address is krubin@math.uci.edu.

imaginary quadratic field. (The case $\mathrm{End}(E) = \mathbb{Z}$ is much more common.) For example, the elliptic curve $y^2 = x^3 + ax$ has complex multiplication because $(x, y) \mapsto (-x, iy)$ is an automorphism of E of order 4. Deuring proved that if E has complex multiplication, then $L(E, s)$ can be identified with the L-function attached to a Hecke character and hence has an analytic continuation to the entire complex plane. Further, in this case Damerell proved that there is an explicit $\Omega \in \mathbb{R}^\times$ such that $L(E, 1)/\Omega \in \mathbb{Z}$.

By the mid-1970s little was known about (BSD) beyond computational examples. That was how things stood when Coates and Wiles made their breakthrough:

Theorem 1 (Coates and Wiles, 1977 [2])**.** *Suppose E has complex multiplication. If $E(\mathbb{Q})$ is infinite, then $L(E, 1) = 0$.*

In other words, if $\mathrm{rank}(E(\mathbb{Q})) > 0$, then $\mathrm{ord}_{s=1} L(E, s) > 0$.

The key to the proof of Theorem 1 is the use of Robert's elliptic units to provide the crucial link between the algebraic and analytic sides of (BSD). Elliptic units are global units in abelian extensions of imaginary quadratic fields, defined by analytic functions, so they live in both the algebraic and analytic worlds.

Suppose E has complex multiplication, and let K be the imaginary quadratic field $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$. Suppose $p \geq 5$ is a prime where E has good reduction, and p factors as $p = \pi\bar{\pi} \in \mathrm{End}(E) \subset K$. For $n \geq 1$ let $E[\pi^n] \subset E(\bar{K})$ denote the kernel of the endomorphism π^n, let $K_n := K(E[\pi^n])$, and let

$$\Gamma_n := \mathrm{Gal}(K_n/K) \cong (\mathbb{Z}/p^n\mathbb{Z})^\times.$$

Since K_n is an abelian extension of K, there is a subgroup $C_n \subset K_n^\times$ of elliptic units. We also let U_n denote the units in the ring of integers of the completion of K_n at the unique prime above π.

Coates and Wiles constructed a "logarithmic derivative" homomorphism $\psi_n : U_n \to E[\pi^n]$ and computed that

$$\psi_n(C_n) = \frac{L(E, 1)}{\Omega} E[\pi^n].$$

Under a mild additional assumption on p, they showed that the group of Γ_n-equivariant homomorphisms $\mathrm{Hom}_{\Gamma_n}(U_n, E[\pi^n])$ is cyclic of order p^n, generated by ψ_n, and therefore

$$\mathrm{Hom}_{\Gamma_n}(U_n/C_n, E[\pi^n]) \cong \mathbb{Z}/(p^n, \tfrac{L(E,1)}{\Omega})\mathbb{Z}. \tag{1}$$

Now suppose $E(\mathbb{Q})$ is infinite. Fix a point $P \in E(\mathbb{Q})$ of infinite order, and for every positive integer n choose a point $Q_n \in E(\bar{K})$ such that $\pi^n(Q_n) = P$. The "Kummer map" that sends $\sigma \in \mathrm{Gal}(\bar{K}/K_n)$ to $\sigma(Q_n) - Q_n$ defines a homomorphism

$$\kappa_n \in \mathrm{Hom}_{\Gamma_n}(\mathrm{Gal}(\bar{K}/K_n), E[\pi^n]).$$

Using class field theory, κ_n induces a homomorphism

$$\tilde{\kappa}_n \in \mathrm{Hom}_{\Gamma_n}(U_n/C_n, E[\pi^n]),$$

and Coates and Wiles showed that there is an integer k, independent of n, such that $\tilde{\kappa}_n$ has order p^{n-k} for all $n \geq k$. Comparing this with (1) as n grows proves Theorem 1.

The ideas in the proof of Theorem 1, along with methods Wiles developed for his work on Iwasawa's Main Conjecture and on the modularity of elliptic curves (see the contributions by Barry Mazur and Henri Darmon, respectively, in this issue), have continued to play an important role in progress on the Birch and Swinnerton-Dyer conjecture.

- Kolyvagin (1990) recognized that elliptic units form what he calls an *Euler system*. (In fact, as one of very few known examples, elliptic units helped him to formulate the concept of an Euler system.) Combining the methods of Coates and Wiles with Kolyvagin's Euler system machinery led to my 1991 proof of Iwasawa's Main Conjecture for imaginary quadratic fields.
- Using a quite different Euler system of Heegner points, the combined results in the 1980s of Kolyvagin, Gross–Zagier, Bump–Friedberg–Hoffstein, and Murty–Murty showed that (BSD) holds if E is modular and $\mathrm{ord}_{s=1} L(E, s) \leq 1$.
- The work of Wiles on modularity [6], completed by Taylor–Wiles [4] and Breuil–Conrad–Diamond–Taylor [1], showed in the 1990s that every elliptic curve over $\mathbb{Q}$ is modular. Hence we have the following result for all elliptic curves over $\mathbb{Q}$, with or without complex multiplication.

Theorem 2. *If* $\mathrm{ord}_{s=1} L(E, s) \leq 1$*, then* $\mathrm{rank}(E(\mathbb{Q})) = \mathrm{ord}_{s=1} L(E, s)$.

Along with some Iwasawa-theoretic results due to Kato (2004) and Skinner and Urban (2014), this is currently the best result in the direction of the Birch and Swinnerton-Dyer conjecture for elliptic curves over $\mathbb{Q}$.

References

[1] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc.* **14** (2001), 843–939. MR1839918

[2] J. Coates and A. Wiles, On the conjecture of Birch and Swinnerton-Dyer, *Invent. Math.* **39** (1977), 223–251. MR0463176

[3] K. Rubin, The "main conjectures" of Iwasawa theory for imaginary quadratic fields, *Invent. Math.* **103** (1991), 25–68. MR1079839

[4] C. Skinner and E. Urban, The Iwasawa main conjectures for GL_2, *Invent. Math.* **195** (2014), 1–277. MR3148103

[5] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math.* **141** (1995), 553–572. MR1333036

[6] A. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math.* **141** (1995), 443–551. MR1333035

Barry Mazur

Andrew Wiles's Work on the Main Conjecture of Iwasawa Theory

What a joy it was to work with Andrew on the Main Conjecture of Iwasawa theory over $\mathbb{Q}$, and how great that Andrew went on to establish it, more generally, for totally real fields.

Barry Mazur is Gerhard Gade University Professor at Harvard University. His e-mail address is mazur@math.harvard.edu.

The basic idea behind the Main Conjecture might be thought of as having, as a starting place, the classical analytic formulas of number theory, such as Dirichlet's Class Number Theorem for quadratic imaginary fields:

$$\frac{L(\chi,1)}{2\pi} = \frac{h(-d)}{w\sqrt{d}},$$

where χ is the quadratic Dirichlet character cutting out the quadratic imaginary field $\mathbb{Q}(\sqrt{-d})$ of (integer) conductor $-d < 0$; $h(-d)$ is the order of the ideal class group of that field; w is the number of roots of unity in it; and $L(\chi, s)$ is the Dirichlet L-function. One of the striking aspects of this formula is that the left-hand side is *analytic*; the right-hand side is *arithmetic*.

The "Main Conjecture" is not a misnomer: for any prime number p the conjecture asserts a fundamental relationship that establishes a close tie between

- *the p-adic L-functions of a number field k*—these being p-adic analytic objects closely related to the classical (complex) L-functions of k,
- *the deep arithmetic of that number field*—namely, the p-primary parts of ideal class groups of certain abelian extensions of k.

Slightly more specifically, the Main Conjecture identifies the zeroes of p-adic L-functions of a number field k with the eigenvalues of an operator on a p-adic vector space constructed from the p-primary parts of ideal class groups of the abelian extensions of k alluded to above or constructed from some closely related arithmetic objects.[1]

This puts the conjecture somewhat in the spirit of a suggestion attributed to Hilbert and Pólya that a complex L-function of a number field might arise naturally as the characteristic series attached to a certain unbounded operator on a (naturally defined) Hilbert space. (Their suggestion, though, goes further by noting that if these operators were self-adjoint, this would relate well to the Riemann Hypothesis.) It also is in the spirit of the classical theory of L-functions attached to varieties over finite fields, for such an L-function can also be thought of as the characteristic polynomial of the Frobenius operator on an appropriate étale cohomology group.

To outline an example of the Main Conjecture, we'll discuss the relevant *group of operators*, *vector space*, *p-adic L-functions*, and *the method for the construction of the relevant arithmetic objects*.

The p-Cyclotomic Tower and the Fundamental Operator in the Main Conjecture

Let $n \geq 1$. Consider the finite field extension $\mathbb{Q}[\mu_n]/\mathbb{Q}$ obtained by adjoining the group of nth roots of unity,

$$\mu_n := \{e^{2\pi i a/n} \mid a = 0, 1, \ldots, n-1\} \subset C^*,$$

to the field $\mathbb{Q}$ of rational numbers. The group $(\mathbb{Z}/n\mathbb{Z})^*$ is canonically isomorphic to the group $Aut(\mu_n)$ of automorphisms of the cyclic group μ_n, this isomorphism being defined by sending $a \in (\mathbb{Z}/n\mathbb{Z})^*$ to the automorphism $\zeta \mapsto \zeta^a$ for any $\zeta \in \mu_n$. Any automorphism of μ_n extends uniquely to an automorphism of the field $\mathbb{Q}[\mu_n]$, giving us canonical isomorphisms:

$$(\mathbb{Z}/n\mathbb{Z})^* \xrightarrow{\simeq} Aut(\mu_n) \xrightarrow{\simeq} \mathrm{Gal}(\mathbb{Q}[\mu_n]/\mathbb{Q}).$$

Fixing a prime p (for simplicity, suppose $p > 2$) and letting n run through powers of p, form the **p-cyclotomic tower of (abelian Galois) extensions**

$$\mathbb{Q} \subset \mathbb{Q}[\mu_p] \subset \mathbb{Q}[\mu_{p^2}] \subset \mathbb{Q}[\mu_{p^3}] \subset \ldots$$

and put $\mathbb{Q}[\mu_{p^\infty}] := \bigcup_{\nu=1}^{\infty} \mathbb{Q}[\mu_{p^\nu}]$. We have that

$$\mathrm{Gal}(\mathbb{Q}[\mu_{p^\infty}]/\mathbb{Q}) = \lim_{\nu\to\infty} (\mathbb{Z}/p^\nu\mathbb{Z})^* = \mathbb{Z}_p^*,$$

the latter, $\mathbb{Z}_p^*$, being the profinite topological group of p-adic units, which decomposes as a product, $\mathbb{Z}_p^* = \mathbf{F}_p^* \times \{1+p\mathbb{Z}_p\}$, where $\mathbf{F}_p^* \simeq \mathrm{Gal}(\mathbb{Q}[\mu_p]/\mathbb{Q})$ is a cyclic group of order $p-1$, and the subgroup $\Gamma := \{1+p\mathbf{Z}_p\} \subset \mathbb{Z}_p^*$ is an infinite cyclic pro-p-group. A neat topological generator to choose for Γ is the p-adic unit $\gamma := (1+p) \in \Gamma \subset \mathbb{Z}_p^*$.

The field $\mathbb{Q}[\mu_{p^\infty}]$ is generated by two linearly disjoint subfields: $\mathbb{Q}[\mu_p]$ and a field, call it $\mathbb{Q}_\infty$, Galois over $\mathbb{Q}$ with Galois group

$$\mathrm{Gal}(\mathbb{Q}_\infty/\mathbb{Q}) = \Gamma := 1 + p\mathbf{Z}_p \subset \mathbf{Z}_p^*.$$

The topological group Γ is our "group of operators."

The Vector Space Containing Basic Arithmetic Data Related to Number Fields

Let $k \subset K$ be number fields, contained in C, linearly disjoint from $\mathbb{Q}_\infty$, with K totally real, and K/k a cyclic Galois extension. Put $K_\infty = K \cdot \mathbb{Q}_\infty \subset C$. So $\mathrm{Gal}(K_\infty/K) = \mathrm{Gal}(\mathbb{Q}_\infty/\mathbb{Q}) = \Gamma$.

Let L_∞ denote the maximal unramified pro-p abelian extension of K_∞. Let $X := \mathrm{Gal}(L_\infty/K_\infty)$, which, since it is a projective limit of p-abelian groups, we can view naturally as a $\mathbb{Z}_p$-module. Also, $\Gamma = \mathrm{Gal}(K_\infty/K)$ acts naturally (and $\mathbb{Z}_p$-linearly) on X. The action of an element in Γ on X is defined by lifting it to an element in $\mathrm{Gal}(L_\infty/K)$ and then noting that conjugation by that lifted element doesn't depend on the lifting and induces a well-defined automorphism of X.

One knows that $V := X \otimes_{\mathbb{Z}_p} \bar{\mathbb{Q}}_p$ is a finite-dimensional vector space. If $\chi : \mathrm{Gal}(K/k) \hookrightarrow C^*$ is an odd (faithful) character cutting out the field extension K/k, we will be considering the action of Γ on the χ-part of V, i.e., on $V^\chi := \{v \in V \mid g(v) = \chi(g) \cdot v\}$. We want to get as full an understanding of the V^χ as possible and specifically the eigenvalues of $1-\gamma$ on V^χ where $\gamma \in \Gamma$ is a topological generator.

L-Functions

By interpolating special values of the classical complex L-functions, Kubota and Leopoldt defined the p-adic L-functions over the field $k = \mathbb{Q}$, and subsequently Deligne and Ribet defined them over totally real fields k. Briefly, in the context above, let $\zeta_k(\sigma, s)$ denote the partial zeta-function of k associated to elements $\sigma \in \mathrm{Gal}(K/k)$. The special values $\zeta_k(\sigma, 1-n)$ are rational numbers, as proved by Klingen and Siegel. Let ψ be a one-dimensional character over k with values in $\overline{\mathbb{Q}_p}^*$ and $\chi_n := \psi^{-1}\omega^n$

[1] *For example, from abelian extensions of p-power degree unramified over those alluded-to extensions of k.*

where ω is the Teichmüller character. For any integer $n \geq 1$ one puts

$$L_p(1-n,\psi) = \sum_{\sigma \in \mathrm{Gal}(K/k)} \chi_n(\sigma)\zeta_k(\sigma, 1-n) \cdot \prod_{P\ |p} (1-\chi_n(P)N(P)^{1-n}),$$

where "$\prod_{P\ |p}$" is taken over all primes P of k lying above p, and $N(P)$ is the norm of the ideal P.

These special values, $\{1-n \mapsto L_p(1-n,\psi) \in \bar{\mathbb{Q}}_p\}$, interpolate to produce a p-adic analytic function $L_p(s,\psi)$ on $\mathbb{Z}_p$ if ψ is not trivial and on $\mathbb{Z}_p - \{1\}$ with a simple pole at $s=1$ when ψ is trivial. Moreover, there is a unique power series $G_\psi(T) \in \bar{\mathbb{Z}}_p[[T]]$ such that for the topological generator $\gamma \in \Gamma \subset \mathbb{Z}_p^*$ (viewed as an element of $\mathbb{Z}_p^*$) we have

$$L_p(s,\psi) = G_\psi(\gamma^s - 1).$$

The Main Conjecture, then, for χ odd and $p > 2$ identifies the characteristic polynomial of γ acting on V^χ as described above, with the Weierstrass polynomial of the power series $G_\psi(\gamma(1+T)^{-1} - 1)$ where $\psi = \chi^{-1}\omega$. That is, as Andrew proves [3], *the eigenvalues of γ acting on V^χ are identified with the zeroes of $L_p(s,\psi)$.*

Method

The essential issue in proving the Main Conjecture is to construct (by means of Galois representations attached to modular forms) as many abelian unramified extensions as would be predicted from the analytic side of the conjectured formula. A version of the classical analytic formula then allows one to conclude the conjecture. When the field k is $\mathbb{Q}$, we did this [1] by a two-step approach, starting from a marvelous idea of Ken Ribet relating divisibility of the p-adic L-function by p to a similar divisibility of the order of a specific ideal class group by p. (For a leisurely discussion of Ribet's method and connection with the earlier work of Herbrand, see [2].) Briefly, the p-adic L-function $L_p(s,\chi)$ (times an elementary nonzero factor) occurs as the constant term of a p-adic Eisenstein series (of p-adic weight determined by the value of s). Whenever the p-adic L-function vanishes, this Eisenstein series has constant term zero and can be shown, therefore, to be congruent modulo arbitrarily high powers of p to cuspidal eigenforms. Moreover, since these cuspidal modular eigenforms are congruent modulo a high power of p to Eisenstein series, their associated p-adic Galois representations are extremely well behaved modulo those powers of p and can be seen to cut out larger and larger unramified p-power abelian extensions of K_∞. The procedure employed by Wiles [3] for the totally real case is a good deal more delicate than in the case of $k = \mathbb{Q}$, in that one now uses the Galois representations of eigenforms on Hilbert–Blumenthal moduli spaces to construct the desired unramified abelian extensions. Andrew works systematically over the relevant weight space constructing the appropriate cuspidal Λ-adic Hilbert modular forms, using Hida's theory, and [3] proves much more.

References

[1] B. Mazur and A. Wiles, Class fields of abelian extensions of $\mathbb{Q}$, *Invent. Math.* **76** (1984), 179–330. MR742853

[2] B. Mazur, How can we construct abelian extensions of number fields?, *Bulletin of the A.M.S.* **48** (2011), 155–209. MR2774089

[3] A. Wiles, The Iwasawa conjecture for totally real fields, *Ann. of Math.* **131** (1990), 493–540. MR1053488

Mirela Çiperiani

Solvable Points on Genus One Curves

In the spring of 2002, as a graduate student at Princeton, I approached Andrew Wiles to ask about the possibility of working under his supervision. He suggested that I think about a problem that was thrillingly natural and beautiful: *Does every genus one curve, defined over the rational numbers $\mathbb{Q}$, have a point over some solvable extension of $\mathbb{Q}$?*

Solvable extensions of $\mathbb{Q}$—Galois extensions with solvable Galois group—are, concretely, fields contained in root extensions of $\mathbb{Q}$, i.e., in fields obtained by starting with $\mathbb{Q}$ and considering successive extensions of the form $F(\sqrt[n]{\alpha})/F$ for some $\alpha \in F$ and $n \in \mathbb{N}$. This relationship between solvability of the Galois group and iterated adjunction of roots is the Galois-theoretic criterion for solvability by radicals of a polynomial equation in one variable. Thus, as shown by Abel and Galois, equations in one variable with coefficients in $\mathbb{Q}$ and degree at least 5 need not be solvable by radicals. This connection to classical Galois theory is one of the appeals of the problem.

An obvious stumbling block is that, while this is evidently a diophantine equations problem, it is much less tangible than that of solving a quintic: we don't have a generic way of writing the equations that describe genus one curves defined over $\mathbb{Q}$. A priori, a genus one curve defined over $\mathbb{Q}$ is cut out by several homogeneous polynomial equations with rational coefficients. The famous Weierstrass cubic equations $y^2z = x^3 + axz^2 + bz^3$ (here $a, b \in \mathbb{Q}$ such that $x^3 + ax + b$ has no repeated roots) describe *elliptic* curves, i.e., genus one curves which *do* have a point over their field of definition, namely, $(0:1:0)$. However, each genus one curve is a torsor for an elliptic curve, its Jacobian; i.e., the Jacobian acts on the genus one curve, and over $\overline{\mathbb{Q}}$ that action becomes freely transitive. Conversely, every torsor for an elliptic curve E is a genus one curve with Jacobian E. In light of this correspondence, genus one curves over $\mathbb{Q}$ with fixed Jacobian E are parametrized by a (Galois) cohomology group, the Weil–Châtelet group $H^1(\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}), E(\bar{\mathbb{Q}}))$. So now we dispense with defining equations of genus one curves and work with elements of the Weil–Châtelet group.

In 2006, in joint work with Wiles, we showed that under certain restrictions, Wiles's original question has a positive answer, as stated momentarily. The removal of these restrictions is the subject of ongoing joint work.

Mirela Çiperiani is associate professor of mathematics at the University of Texas at Austin. Her e-mail address is mirela@math.utexas.edu.

Theorem ([ÇW]). *Let* C *be a genus one curve defined over* $\mathbb{Q}$*, with Jacobian* E*, such that*

(1) C *has a point defined over the* ℓ*-adics* $\mathbb{Q}_\ell$ *for all rational primes* ℓ*; and*
(2) *one of the following conditions holds:*
   (a) *the analytic rank of* $\mathrm{E}/\mathbb{Q}$ *is less than or equal to* 1 *or*
   (b) E *is semistable over* $\mathbb{Q}$.

Then C *has a point over a solvable extension of* $\mathbb{Q}$.

The analytic rank of an elliptic curve E is the order of vanishing of its L-series $L(\mathrm{E}, s)$ at $s = 1$. As for semistability, after lifting an elliptic curve from $\mathbb{Q}$ to $\mathbb{Z}$ and then reducing modulo a prime, it may remain nonsingular or it may acquire a node or cusp. Semistability disallows cusps.

I will now describe two fundamental ideas of Wiles which provide the frame of the proof. The first is referred to as the "unramified under ramified principle." Fix an elliptic curve E defined over $\mathbb{Q}$. Genus one curves C defined over $\mathbb{Q}$ with E as their Jacobian, and with points over $\mathbb{Q}_\ell$ for every rational prime ℓ, form a subgroup of the Weil–Châtelet group. (This subgroup contains the Tate–Shafarevich group, which consists of genus one curves that have both ℓ-adic and real points.) We in fact work with their preimages in the *Selmer group*. These Selmer classes become trivial after an extension of $\mathbb{Q}$ unramified at all but a finite set of primes ℓ. According to the unramified under ramified principle, if we find a finite set of primes $\mathcal{Q}$ obeying certain conditions and if we have sufficiently many cohomology classes which are ramified at primes in $\mathcal{Q}$, then under the cohomology group structure they will generate all the unramified classes.

We will apply this principle in the case where the genus one curve C satisfies conditions (1) and (2a) of our theorem by constructing ramified classes which correspond to genus one curves with points over some solvable extension of $\mathbb{Q}$. These ramified classes will then generate a group containing all genus one curves that have the same Jacobian as C and satisfy conditions (1) and (2a). Hence C has a solvable point. Thus the task is to choose the set $\mathcal{Q}$ and to construct sufficiently many ramified classes which have solvable points. This is achieved by using the cohomology classes constructed by Kolyvagin. The construction of these classes makes use of *Heegner points*. These points are defined on modular curves and pushed forward to the elliptic curve via its modular parametrization—whose existence is known by the work of Wiles extended by Breuil, Conrad, Diamond, and Taylor.

The unramified under ramified principle is sufficient to prove the theorem in the case when the analytic rank of $\mathrm{E}/\mathbb{Q}$ is less than or equal to 1. It is not sufficient to prove it in general, for the following two reasons:

(i) If the analytic rank of $\mathrm{E}/\mathbb{Q}$ is greater than 1, then the relevant Heegner points are trivial. Hence, in this case we cannot construct enough nontrivial Kolyvagin classes.
(ii) The set of primes $\mathcal{Q}$ depends on the order of the curve C viewed as an element of the Weil–Châtelet group of $\mathrm{E}/\mathbb{Q}$, and the existence of the sufficiently many cohomology classes ramified at primes in $\mathcal{Q}$ depends on the finiteness of all the p-primary components of the Tate–Shafarevich group of $\mathrm{E}/\mathbb{Q}$. This is only known for elliptic curves $\mathrm{E}/\mathbb{Q}$ of analytic rank less than or equal to 1 (by the modularity theorem, and the combined and celebrated work of Kolyvagin, Gross–Zagier, Bump–Friedberg–Hoffstein, and Murty–Murty).

When conditions (1) and (2b) hold, we can still find nontrivial Heegner points, using work of Cornut-Vatsal, by viewing C as a genus one curve over a nontrivial extension of $\mathbb{Q}$. We now attempt to apply the unramified under ramified principle. However, while this field extension enables us to construct ramified Kolyvagin classes, because of issue (ii) we are not able to see that we have sufficiently many of them.

It is in circumventing the potential existence of an infinite p-primary component of the Tate–Shafarevich group that we need the second idea. It is referred to as the "patching method." A similar method was used in the proof of the modularity theorem. Wiles's idea is the following. We construct as many ramified classes as we can for each in an infinite sequence of field extensions $\{F_n\}$. Observe that in order to do this we must choose unrelated sets of primes $\mathcal{Q}_n$ for the fields F_n, all with the same cardinality. We consider groups M_n of cohomology classes over F_n ramified at primes in $\mathcal{Q}_n$ (actually, M_n is viewed as a module over a ring related to F_n). There is no natural containment between these modules, but each of them contains all the classes over $\mathbb{Q}$ that we want to capture. However, our Kolyvagin classes generate a submodule $M_n' \subseteq M_n$, and for no n can we see that M_n' contains the desired classes. By considering their module structure (and ignoring their content), we construct injective maps $M_n \to M_{n+1}$. This gives rise to a module $\varinjlim M_n$ over an Iwasawa algebra. Miraculously, a structure theorem from Iwasawa theory shows us that each of our genus one curves C satisfying conditions (1) and (2b) lies in some M_n' for some n. Thus C has a solvable point.

I developed enormous admiration of Wiles for his generosity and modesty, and awe for his ability to see to the heart of the matter.

Working with Wiles has been a wonderful experience. At the beginning, when I barely knew what an elliptic curve was, I felt lucky but humbled and daunted to be entrusted with such a fantastic problem. I had to rapidly learn the background material needed to begin thinking about the problem and to understand Wiles's proposal to use the unramified under ramified principle. Later, at times when the problem seemed impossible, I was buoyed by Wiles's confidence that we could solve it. I developed

enormous admiration of Wiles for his generosity and modesty, and awe for his ability to see to the heart of the matter—feelings that stay with me today.

References

[ÇW] M. Çiperiani and A. Wiles, Solvable points on genus one curves, *Duke Math. J.* **142** (2008), 381–464. MR2412044

Chandrashekhar Khare

Modularity of GL_2 Galois Representations and the Work of Andrew Wiles

The work of Andrew Wiles on the modularity of elliptic curves provided some of the key ideas and techniques which led to the proof of Serre's modularity conjecture. Moreover, and equally importantly, it psychologically made it possible to imagine that there could be a strategy to prove results that before Wiles's work seemed completely inaccessible.

To retrace the path from Wiles's work [5] on modularity of elliptic curves to the proof of Serre's conjecture [2], we describe briefly Wiles's modularity lifting theorem and his strategy for proving it.

Let $G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ be the Galois group of an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. Wiles's modularity lifting theorem is about 2-dimensional p-adic Galois representations for a prime $p > 2$. These are certain homomorphisms

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p),$$

including those that arise from the action of $G_{\mathbb{Q}}$ on the torsion points of an elliptic curve defined over $\mathbb{Q}$. Wiles proved [5], [4] that under certain hypotheses on ρ and on its residual representation

$$\bar{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$$

(the reduction of ρ modulo p), the representation ρ is modular.

The hypotheses on the residual representation $\bar{\rho}$ in Wiles's theorem require that $\bar{\rho}$ be odd (the image of complex conjugation has eigenvalues $+1$ and -1), irreducible, and modular. For $\bar{\rho}$, being modular means that there is a cuspidal modular eigenform (an analytic function on the complex upper-half-plane with many symmetries)

$$f(\tau) = \sum_{n=1}^{\infty} a_n(f) e^{2\pi n \tau}$$

such that for all but finitely many primes ℓ, the Fourier coefficient $a_\ell(f)$ can be matched with the trace of $\bar{\rho}$ evaluated on a Frobenius element in $G_{\mathbb{Q}}$ for the prime ℓ. More precisely, the Fourier coefficients $a_n(f)$ are all algebraic integers (so belong to $\overline{\mathbb{Q}}$) and for some embedding $\iota : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$ of $\overline{\mathbb{Q}}$ into an algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$, for almost all primes ℓ the image of $\iota(a_\ell(f))$ in the residue field $\overline{\mathbb{F}}_p$ of $\overline{\mathbb{Q}}_p$ equals the trace of $\bar{\rho}$ on a Frobenius element for ℓ. Similarly, ρ is modular if there is an eigenform f and an embedding ι such that for almost all primes ℓ, $\iota(a_\ell(f))$ equals the trace of ρ on a Frobenius element for ℓ. Wiles's theorem "lifts modularity" in that it asserts that if $\bar{\rho}$ is modular, then the lift ρ of $\bar{\rho}$ is also modular.

Wiles's theorem allows for representations ρ where $\mathbb{Z}_p$ is replaced by more general p-adic rings. In particular, it allows for homomorphisms $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$; a continuity argument associates to such a ρ a residual representation $\bar{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$. Much earlier results of Shimura and Deligne attach to each eigenform f and embedding $\iota : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$ a Galois representation

$$\rho_{f,\iota} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p).$$

The spirit of Wiles's modularity lifting theorem is: if $\bar{\rho} \cong \bar{\rho}_{f,\iota}$ for some eigenform f, then $\rho \cong \rho_{g,\iota'}$ for some eigenform g (not necessarily the same as f). More about this theorem and its context can be found in Darmon's lecture in this issue and in the article by Thorne.

Wiles's strategy for proving his marvelous modularity lifting theorem can very roughly be paraphrased as follows. The kernel of the mod p reduction map $\mathrm{GL}_2(\mathbb{Z}_p) \to \mathrm{GL}_2(\mathbb{F}_p)$ is prosolvable (the inverse limit of solvable groups), which enables the use of duality theorems in Galois cohomology and congruences between modular forms to bootstrap the modularity property from $\bar{\rho}$ to ρ.

One spectacular application of Wiles's modularity lifting theorem was the proof of the modularity of semistable elliptic curves defined over $\mathbb{Q}$ (and hence Fermat's Last Theorem!). Elliptic curves are often encountered as the projective curves defined by an (affine) Weierstrass equation: $y^2 = x^3 + ax + b$ for constants a and b. The curve is defined over $\mathbb{Q}$ if $a, b \in \mathbb{Q}$. The points on an elliptic curve E form an abelian group, with the addition law defined by rational functions in x and y; the identity element of the group is the unique point at infinity (the point $(0 : 1 : 0)$ in projective coordinates). The torsion points $E[N]$ on E of (positive integer) order N form a group isomorphic to $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$. If the elliptic curve is defined over $\mathbb{Q}$, then the Galois group $G_{\mathbb{Q}}$ acts on $E[N]$ by its action on the coordinates of the points. The p-adic Tate module T_pE of E is then the inverse limit of the groups $E[p^n]$ of p-power torsion points: $T_pE = \varprojlim_n E[p^n] \cong \mathbb{Z}_p^2$. The action of $G_{\mathbb{Q}}$ on T_pE determines a p-adic Galois representation

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p).$$

The residual representation $\bar{\rho}_{E,p}$ of $\rho_{E,p}$ is just the representation of $G_{\mathbb{Q}}$ on the group of p-torsion points $E[p] \cong \mathbb{F}_p^2$. The modularity of an elliptic curve can be interpreted as the p-adic Galois representation $\rho_{E,p}$ being modular for some prime p (equivalently, all primes p).

The application of the modularity lifting theorem to modularity of elliptic curves over $\mathbb{Q}$ comes by taking $\rho = \rho_{E,p}$. For $p = 2$ or 3 the image of the representation $\rho_{E,p}$ is prosolvable (the inverse limit of finite solvable groups). This is one of the two places in his argument where Wiles uses lucky accidents which happen for small primes. He uses results of Langlands and Tunnell which imply that an odd irreducible representation $\bar{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_3)$ is modular. This uses the solvability of $\mathrm{GL}_2(\mathbb{F}_p)$ for $p = 3$ (indeed, $\mathrm{PGL}_2(\mathbb{F}_3)$ is isomorphic to S_4) and the fact that the map $\mathrm{GL}_2(\mathbb{Z}_p)$ to $\mathrm{GL}_2(\mathbb{F}_p)$ splits for $p = 3$; both these statements are false for $p > 3$!

Chandrashekhar Khare is professor of mathematics at UCLA. His e-mail address is shekhar@math.ucla.edu.

At this point, the modularity lifting theorem allows Wiles to conclude modularity of semistable E unless $\bar{\rho}_{E,3}$ is reducible. To deal with the case when $\rho_{E,3}$ is reducible, Wiles plays a "3-5" trick. He constructs another semistable elliptic curve E' over $\mathbb{Q}$, whose mod 3 representation surjects onto $\mathrm{GL}_2(\mathbb{F}_3)$ and such that the mod 5 representations arising from E and E' are isomorphic and irreducible. Here again the use of the small prime 5 is vital, as the moduli space he considers of elliptic curves with level 5 structure isomorphic to $E[5]$ turns out to be of genus zero, and in fact the projective line over $\mathbb{Q}$, and thus has many rational points. Then applying the modularity lifting theorem he deduces that $\rho_{E',3}$ is modular. This implies that $\rho_{E',5}$ is also modular, hence $\bar{\rho}_{E,5} = \bar{\rho}_{E',5}$ is modular, and then by another application of the theorem, that $\rho_{E,5}$, and therefore E, is modular.

Wiles's work was generalized by Breuil, Conrad, Diamond, and Taylor (2001) to prove the modularity of all elliptic curves over $\mathbb{Q}$. Recently, Freitas, Hung, and Siksek (2015) proved the modularity of elliptic curves defined over all real quadratic fields, using a "3-5-7" trick.

Before the work of Wiles there was no path from a p-adic Galois representation to a modular form. His completely new method of modularity lifting showed that if only a small quotient of a p-adic representation arose from a modular form, then the entire Galois representation did. This has turned out to be a very powerful method to prove modularity of Galois representations!

Conjectures of Serre and Fontaine–Mazur

Serre conjectured [3] that for any $\bar{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ that is continuous, odd, and irreducible, there is an eigenform f such that $\bar{\rho} \simeq \bar{\rho}_{f,\iota}$; i.e., $\bar{\rho}$ is modular.

J.-M. Fontaine and B. Mazur conjectured [1] that if $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ is continuous, odd, irreducible, unramified outside a finite set of primes, and potentially semistable with Hodge–Tate weights (a,b), say $a \le b$, then the cyclotomic twist $\rho(-a)$ is modular.

Wiles's modularity lifting results were in the direction of showing that the conjecture of Serre implies that of Fontaine–Mazur. These modularity lifting results were improved in various crucial ways by Diamond, Fujiwara, and Kisin, including generalizations with the base field $\mathbb{Q}$ replaced by a totally real field F. In an important development, Skinner and Wiles lifted the condition that $\bar{\rho}$ is irreducible in the case when the lift ρ is ordinary at p. All these developments were crucial in our later work on Serre's conjecture.

Potential Version of Serre's Conjecture

The automorphic descent results of Saito and Shintani, and Langlands are an important ingredient in the applications of modularity lifting theorems. These show that given a cyclic extension of totally real number fields K/F of prime degree and a cuspidal automorphic representation of π of $\mathrm{GL}_2(\mathbb{A}_K)$ which is a discrete series at the infinite places and invariant under $\sigma \in \mathrm{Gal}(K/F)$, then π is the base change of a cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_F)$. When combined with modularity lifting results, these descent results imply that modularity of $\bar{\rho}$ follows by showing that $\bar{\rho}|_{G_F}$ arises from a Hilbert modular form for a solvable, totally real extension $F/\mathbb{Q}$.

To study general representations $\bar{\rho}$ as in Serre's conjecture, Taylor considered moduli spaces over $\mathbb{Q}$ whose points over a number field K correspond to abelian varieties over K (with real multiplication) which give rise to $\bar{\rho}|_{G_K}$, and at an auxiliary place $\ell \neq p$ give rise to a mod ℓ dihedral representation. The latter are known to be modular by an old result of Hecke. Taylor used a theorem of Moret-Bailly to produce points of the moduli spaces he considered over totally real fields F (but which could not be guaranteed to be solvable over $\mathbb{Q}$). Then modularity lifting theorems for representations of G_F yield that $\bar{\rho}|_{G_F}$ arises from a Hilbert modular form over F. This may be regarded as a potential version of Serre's conjecture. Together with automorphic descent for cyclic prime degree extensions of totally real number fields, this led to the meromorphic continuation of the Hasse–Weil L-series attached to elliptic curves over all totally real fields.

To proceed along these lines to prove Serre's conjecture in the general case, it would be necessary either to find a general procedure to show existence of totally real solvable points on geometrically irreducible smooth projective varieties of general type over $\mathbb{Q}$ or to prove nonsolvable automorphic descent for Hilbert modular forms.

Proof of Serre's Conjecture

My proof with Wintenberger of Serre's conjecture [2] took a different path and used as a starting point results of Tate and Serre that proved Serre's conjecture for representations $\bar{\rho}$ of residue characteristic $p \le 3$ with limited ramification, with no a priori assumptions on the image of $\bar{\rho}$. Tate and Serre proved that any representation $\bar{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ unramified outside p and with $p \le 3$ is reducible. This is another instance of the magic of small primes!

Our proof measured the complexity of the representation $\bar{\rho}$ in terms of its ramification at p (Serre's weight) and the ramification away from p (its Artin conductor N). A double induction on (p,N) was used to reduce Serre's conjecture to the results of Tate and Serre. A principle of the proof is to use potential modularity to produce compatible systems of representations that lift $\bar{\rho}$ such as would exist if it were known that $\bar{\rho}$ were modular.

Our proof of Serre's conjecture owes its existence to the modularity lifting theorems initiated by Wiles, a tool to attack modularity which was as powerful as it was unexpected when it was introduced, and also to Wiles's prime switching trick in his proof of the modularity of elliptic curves.

References

[1] Jean-Marc Fontaine and Barry Mazur, Geometric Galois representations, *Elliptic Curves, Modular Forms, & Fermat's Last Theorem (Hong Kong, 1993)*, Ser. Number Theory, I, Internat. Press, Cambridge, MA, 1995, pp. 41–78. MR1363495

[2] Chandrashekhar Khare and Jean-Pierre Wintenberger, Serre's modularity conjecture (I), *Invent. Math.* **178** (2009), no. 3, 485–504. MR2551763
[3] Jean-Pierre Serre, Sur les représentations modulaires de degré 2 de Gal($\overline{\mathbf{Q}}/\mathbf{Q}$), *Duke Math. J.* **54** (1987), no. 1, 179–230. MR885783
[4] Richard Taylor and Andrew Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math. (2)* **141** (1995), no. 3, 553–572. MR1333036
[5] Andrew Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math. (2)* **141** (1995), no. 3, 443–551. MR1333035

Jack Thorne

Modularity of n-Dimensional Galois Representations

More than twenty years have passed since Andrew Wiles proved the first modularity lifting theorems and deduced Fermat's Last Theorem as a consequence. His theorems focussed on proving the modularity of certain 2-dimensional Galois representations. These were the homomorphisms

$$\varrho : G_{\mathbb{Q},S} \to \mathrm{GL}_2(\mathbb{Z}_\ell) \tag{1}$$

arising from elliptic curves E, $G_{\mathbb{Q},S}$ being the Galois group of the maximal extension of $\mathbb{Q}$ unramified outside some finite set of primes S. Despite this focus, the ideas of the proof have proved powerful and flexible enough that they are still a driving force of our understanding of much more general Galois representations today.

The first key hypothesis imposed on ϱ is the modularity of the residual representation

$$\overline{\varrho} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(\mathbb{F}_\ell) \tag{2}$$

that is the reduction of ϱ modulo ℓ. In other words, one assumes at the outset the existence of a cuspidal modular eigenform

$$f = \sum_{n\geq 1} a_n(f) q^n \tag{3}$$

which is matched with $\overline{\varrho}$, in the sense that for almost all primes p, the Fourier coefficient $a_p(f)$ is congruent modulo ℓ to the integer $a_p(E) = p + 1 - \#E(\mathbb{F}_p)$. This allows the introduction of a first key player, the Hecke algebra $\mathbb{T}_{\overline{\varrho},S}$, which acts faithfully on a space of cuspidal modular forms, all of whose Fourier coefficients agree modulo ℓ with those of f and which have level supported at the primes of S.

The second key player is the universal deformation ring of $\overline{\varrho}$, which we call $R_{\overline{\varrho},S}$. It is a complete local ring which is characterized by a universal property. For example, the set of homomorphisms $R_{\overline{\varrho},S} \to \mathbb{Z}_\ell$ is in bijection with the set of equivalence classes of lifts $\varrho' : G_{\mathbb{Q},S} \to \mathrm{GL}_2(\mathbb{Z}_\ell)$ of $\overline{\varrho}$ that one hopes to prove are modular.

These key players are related by a surjective ring homomorphism

$$R_{\overline{\varrho},S} \to \mathbb{T}_{\overline{\varrho},S}. \tag{4}$$

Jack Thorne is a Clay Research Fellow and Reader in Number Theory at the University of Cambridge. His e-mail address is thorne@dpmms.cam.ac.uk.

Passing to spectra, one can think of $\mathrm{Spec}\,\mathbb{T}_{\overline{\varrho},S}$ as a closed subspace of $\mathrm{Spec}\,R_{\overline{\varrho},S}$: it is the locus of the modular Galois representations inside the space of all Galois representations. In particular, the existence of this map expresses the *existence* of Galois representations attached to modular forms, itself a highly nontrivial fact. Diagrams such as (4) have achieved an iconic status in algebraic number theory.

The truth of the modularity lifting theorem is implied by the much more refined statement that the map (4) is an isomorphism. This goes some way towards explaining the importance of the universal deformation ring $R_{\overline{\varrho},S}$, which was first introduced by Mazur. A large part of Wiles's fundamental 1995 work is taken up with introducing the tools necessary to effectively study the map (4), putting him in a position to prove '$R = \mathbb{T}$' in many cases.

One tool that has turned out to be surprisingly versatile is the Taylor–Wiles method, introduced in the companion paper [3]. Roughly speaking, this is an effective machine to study (4) in the so-called minimal case where S is as small as possible. To pass from this case to the general case, Wiles introduced a numerical isomorphism criterion to compare the situation for varying sets S. The verification of this criterion then involves delicate calculations in Galois cohomology and with modular forms.

The first modularity lifting theorems for Galois representations of dimension $n > 2$ were proved by Clozel, Harris, and Taylor in a paper published in 2008, which was heavily influenced by an earlier unpublished manuscript of Harris and Taylor. This was made possible thanks to the construction of n-dimensional Galois representations attached to modular forms on unitary groups, itself initiated by Clozel and Kottwitz and then studied in great detail by Harris and Taylor in their proof of the local Langlands conjectures for GL_n, allowing one to write the n-dimensional analogue of the map $R_{\overline{\varrho},S} \to \mathbb{T}_{\overline{\varrho},S}$.

The Taylor–Wiles method was generalized by Clozel, Harris, and Taylor in order to prove a modularity lifting theorem in the minimal case. However, such a restriction on ramification makes these theorems very difficult to apply in interesting situations. The general case was treated using a generalization of the numerical criterion of Wiles, but only conditional on a conjecture (referred to colloquially as Ihara's lemma) that remains unproven today.

Kisin had earlier developed a generalization of the Taylor–Wiles method in his study of modularity for GL_2, in a work published in 2009. He enlarged the diagram (4) to a diagram

$$\widehat{\bigotimes}_{p\in S} R_{\overline{\varrho},p} \to R_{\overline{\varrho},S} \to \mathbb{T}_{\overline{\varrho},S}, \tag{5}$$

where each ring $R_{\overline{\varrho},p}$ is an object parameterizing deformations of the restriction of $\overline{\varrho}$ to a decomposition group at p (in other words, a local Galois group $D_p = \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$). In this point of view, the geometry of the rings $R_{\overline{\varrho},p}$ begins to play a key role. The Taylor–Wiles–Kisin method finally allows one to link the modularity of Galois representations $\varrho_1, \varrho_2 : G_{\mathbb{Q},S} \to \mathrm{GL}_n(\mathbb{Z}_\ell)$ which have the property that the representations $\varrho_1|_{D_p}, \varrho_2|_{D_p}$ determine points on

the *same* irreducible component of $\operatorname{Spec} R_{\overline{\varrho},p}$ for each prime $p \in S$.

Building on this, Taylor [2] made a detailed study of the irreducible components of certain 'local' deformation rings and introduced a very surprising trick that allowed him to circumvent completely the conjectural Ihara's lemma and prove the first modularity lifting theorems for GL_n without restriction on the permitted ramification of ϱ relative to $\overline{\varrho}$.

These theorems have had spectacular applications. Combined with the earlier work of Harris, Shepherd-Barron, and Taylor, they implied that the even-dimensional symmetric power Galois representations

$$\operatorname{Sym}^{n-1} \varrho : G_{\mathbb{Q},S} \to \mathrm{GL}_n(\mathbb{Z}_\ell) \tag{6}$$

are potentially modular: there exists a number field $F/\mathbb{Q}$ such that the restriction $\operatorname{Sym}^{n-1} \varrho|_{G_F}$ to the absolute Galois group of F is modular, in the sense of being associated to modular (or automorphic) forms on $\mathrm{GL}_{n,F}$. These ideas in turn led to the proof of the Sato–Tate conjecture for elliptic curves over $\mathbb{Q}$:

Theorem 1. *Let E be an elliptic curve over $\mathbb{Q}$ without complex multiplication. Then the quantities $a_p(E)/2\sqrt{p} \in [-1,1]$ are equidistributed as $p \to \infty$ with respect to the Sato–Tate measure*

$$\frac{2}{\pi}\sqrt{1-t^2}\,dt. \tag{7}$$

There are many fruitful directions that remain to be explored. For example, can one show that the symmetric powers (6) are modular and not just potentially modular? In joint work with Clozel [1], I showed that the answer to this question is affirmative for $n \le 9$. A major part of our proof is a generalization of an important modularity lifting theorem of Skinner and Wiles which applies to Galois representations ϱ for which the residual representation $\overline{\varrho}$ is *reducible*. This modularity lifting theorem also played a major part in the proof of Serre's conjecture.

The work of Wiles has had a transforming effect.

The proof of the Skinner–Wiles theorem employs many ideas which appear in Andrew Wiles's most famous works, such as p-adic families of modular forms and congruences between Eisenstein series and cuspidal modular forms, as well as many other ideas whose importance would become apparent only later: we mention in particular an emphasis on the geometry of Galois deformation rings. The work of Wiles has had a transforming effect on this corner of number theory, and his influence continues to be felt throughout the subject.

References

[1] Laurent Clozel and Jack A. Thorne, Level raising and symmetric power functoriality, II, *Ann. of Math. (2)* **181** (2015), no. 1, 303–359. MR3272927

[2] Richard Taylor, Automorphy for some l-adic lifts of automorphic mod l Galois representations. II, *Publ. Math. Inst. Hautes Études Sci.* No. 108 (2008), 183–239. MR2470688 (2010j:11085)

[3] Richard Taylor and Andrew Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math. (2)* **141** (1995), no. 3, 553–572. MR1333036 (96d:11072)

Photo Credits

Photo taken at the Newton Institute is courtesy of Ken Ribet.

Photo of Christopher Skinner is courtesy of William Crow/Princeton University.

Photo of Barry Mazur is courtesy of Jim Harrison.

Photo of Mirela Çiperiani is by C. J. Mozzochi, courtesy of the Simons Foundation.

Photo of Chandrashekhar Khare is by David Weisbart.

ABOUT THE AUTHORS

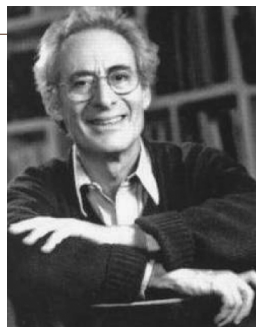
Christopher Skinner studied at Princeton with Andrew Wiles in the mid-1990s.

Christopher Skinner

Karl Rubin

Karl Rubin was Andrew Wiles's first PhD student.

Barry Mazur delights in the memories of the times Andrew and he had at Harvard, as colleagues.



Barry Mazur

Mirela Çiperiani

Mirela Çiperiani completed her PhD in 2006 under the supervision of Andrew Wiles.

Using techniques and ideas that grew out of Wiles's proof of Fermat's Last Theorem, **Chandrashekhar Khare** (with Jean-Pierre Wintenberger) proved Serre's Conjecture.

Chandrashekhar Khare

Jack Thorne

Jack Thorne studied at Harvard with Richard Taylor, himself a student of Andrew Wiles.