

Virtues of Priority

MICHAEL HARRIS

In memory of Serge Lang

Acknowledgement. — This article was originally written in response to an invitation by a group of philosophers as part of “a proposal for a special issue of the philosophy journal *Synthèse* on virtues and mathematics”. The author thanks the editors of *Annals of Mathematics and Philosophy* for allowing me to rework the text for publication in the new journal, and especially to the referees for their suggestions for improving the exposition.

This work was partially supported by NSF Grant DMS-1701651. This work was also supported by the National Science Foundation under Grant No. DMS-1440140 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2019 semester.

§ — Introduction.

In the *Republic* Plato has Socrates insist to Glaucon that geometry, and by extension mathematics as a whole, “is the knowledge of what always is, and not of what at some particular time comes into being and passes away”⁽¹⁾; Glaucon readily agrees. Even the most dogmatic of contemporary platonists must nevertheless acknowledge that the “knowledge” to which Socrates refers first does come into being at some particular time, and to some particular person or group of persons. Socrates and Glaucon may have been surprised had they known how aggressively many of today’s mathematicians would defend their claims to *priority* in the development of new

⁽¹⁾Translation in G. E. R. Lloyd, “Mathematics and Narrative: An Aristotelian Perspective”, in A. Doxiadis and B. Mazur, eds., *Circles Disturbed*, (Princeton University Press, 2012), p. 389, adapted from *Republic* 527ab, ed. Paul Shorey, Loeb Classical Library 276 (Cambridge, MA: Harvard University Press, 1935)

mathematical knowledge. We, on the other hand, are not surprised; we know the importance today's Academy attaches to priority in assigning the marks of prestige that constitute a mathematician's standing in the community.

The arXiv includes a timestamp with every submission, going so far as to indicate time of day as well as date of submission. Since what is posted on arXiv only counts as "knowledge" when it has been accepted by a journal for publication, journals provide additional timestamps for the timeless truths their pages contain. Publications in the prestigious *Annals of Mathematics*, for example, list four "Milestones": date of submission, date of revision, date of acceptance, and date of publication.

These dates provide objective markers for adjudicating potential priority disputes, when it comes to mathematical *texts*. Priority in the creation of mathematical *ideas* is much less clear-cut. In this article I will consider a priority dispute concerning a famous *conjecture*. A conjecture is something more than an *idea*: it states a proposition that aspires to the status of a *theorem*. But it can be, and in the present instance is, considerably less than a *text*; if and when it first appears as a formal conjecture with a precise date, it may well already have been circulating informally for some time as something different from but recognizably on the path to its ultimate shape.

It is perhaps because of this liminal character of the *Modularity Conjecture* that the dispute over the priority for its creation has never been resolved. I will not attempt to resolve the dispute here. Instead, I will examine the arguments advanced in favor of each of the three attributions of priority — to Yutaka Taniyama, Goro Shimura, and André Weil. On my reading, each of the attributions is justified by the invocation of a different mathematical *virtue*. Each of these virtues, in turn, can be understood as loosely reflecting the priorities of a different school of philosophy: cartesianism in the case of Taniyama, realism or phenomenology for Shimura (and his champion Serge Lang), and Popperian falsificationism for Weil (and his champion Jean-Pierre Serre). My conclusion will be that the priority dispute cannot be settled independently, but only on the basis of the relative importance one is willing to assign to each of these virtues.

§ — Originality and other virtues.

If hiring committees are arbiters of mathematical virtue, then letters of recommendation should give a good sense of the virtues most appreciated by mathematicians. You will not see “proves true theorems” among them. That’s merely part of the job description, and drawing attention to it would be analogous to saying an electrician won’t burn your house down, or a banker won’t steal from your account. I don’t know how electricians or bankers recommend themselves to one another, but I have read a lot of letters for jobs and prizes in mathematics, and their language is revealing in its repetitiveness. Words like “innovative” or “original” are good, “influential” or “transformative” are better, and “breakthrough” or “decisive” carry more weight than “one of the best”. Best, of course, is “the best”, but it is only convincing when accompanied by some evidence of innovation or influence or decisiveness.

When we try to answer the questions: what is being innovated or decided? who is being influenced? — we conclude that the virtues highlighted in reference letters point to mathematics as an undertaking relative to and within a community. This is hardly surprising, because those who write and read these letters do so in their capacity as representative members of this very community. Or I should say: members of overlapping communities, because the virtues of a branch of mathematics whose aims are defined by precisely formulated conjectures (like much of my own field of algebraic number theory) are very different from the virtues of an area that grows largely by exploring new phenomena in the hope of discovering simple underlying principles (like much of the field of dynamical systems). The interpretation of letters is left to the experts who apply not only their knowledge of their field’s problems but also their familiarity with its value system. The same vocabulary (original, influential, breakthrough) can thus apply to developments whose character differs radically from one field to another. Some of the five mathematicians who in 2014 were the first laureats of Silicon Valley’s *Breakthrough Prizes* were honored for solving longstanding unsolved problems, while others defined new classes of objects or raised original questions.

“Originality” is a virtue recognized across disciplinary boundaries. The ability to understand and explain the details of another mathematician’s difficult proof is undoubtedly admired as a virtue, but it’s almost never enough to get you a job. For that you have to

have brought something *new* to your field. This sounds so banal to be unworthy of attention, but the philosophical problem is glaringly obvious: how do we determine what is and is not *new*? Here the word *new* is used as shorthand for the more complex notion of “bringing something new to your field”. Bare newness is not enough. You can take 99 theorems in your field and string them together into a single sentence connected by “and”s but that’s not *new* even if no one has ever thought to do it before. As I already hinted, most true theorems are not *new* in the virtuous sense I share with my fellow members of the hiring or prize committee.

Surely an element of taste is involved. It would be easiest to leave the clarification of the virtue of newness to specialists in mathematical aesthetics, who are in as difficult a position as any philosophers of mathematics. Specifically, they face the difficulty common to all philosophers who deal with collective characteristics of communities: how to distinguish the abstract (and presumably permanent) principle they wish to emphasize from the (presumably transitory) idiosyncracies of individual members of the community? Or more crudely, is a mathematical contribution *innovative* because it is claimed to be so by influential or powerful members of the community — the kind who are asked to write letters of recommendation? Are reference letters just politics by other means?

These considerations can be left in the background when it has somehow come to be accepted that answering a certain question — solving a famous conjecture, for example — meets the community’s standards of importance. No member in good standing of the community of geometers doubted that Grigori Perelman’s claim to have solved the Poincaré conjecture was *new* in the relevant sense, because (a) the problem had long been considered central to the field; (b) no one had solved it before Perelman; and (c) several teams of experts worked through the proof and vouched for its correctness.

It is already more difficult to characterize the newness of a proof of a theorem that was not long-awaited. Judgments in such cases are obtained as the result of a process of implicit negotiation among those recognized as leaders within the community. On the basis of the standards invoked in such negotiations it should be possible to form a reasonably coherent (or perhaps chaotic) picture of the community’s virtues.

Still more problematic are the mathematical ideas that express themselves in the form of *conjectures*. It is at least conventionally

believed that the *goal* of mathematics is to produce proofs of theorems.⁽²⁾ So faced with such a proof, one at least knows where the negotiation starts. The place of conjectures is much more ambiguous: they can guide the thoughts of mathematicians for generations, or even for centuries, but in principle a conjecture that turns out to be mistaken can also vanish without a trace.⁽³⁾

§ — The Modularity Conjecture: Weil’s 1967 paper.

In 1967 André Weil published an 8-page article entitled “Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen” — “On the determination of Dirichlet series by means of functional equations”. The title was an intentional allusion to a 1936 paper of Erich Hecke — “Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung”: only one functional equation is used in Hecke’s title and in his paper for determination, whereas the result of Weil’s paper requires a collection of functional equations. The main theorem of the paper, now known as *Weil’s converse theorem*, provides conditions that guarantee that a Dirichlet series — an infinite series of the form $\sum_n a_n n^{-s}$, where s is a complex variable — arises by a well-known procedure from a *modular form* of a specified type. I will return to this last point later.

Weil must have considered his result respectable, because it was his contribution to a special issue of the journal *Mathematische Annalen* in honor of the 70th birthday of Carl Ludwig Siegel, a mathematician whom Weil greatly admired. The last page or so of his brief text was devoted to “reflections on the zeta functions of elliptic curves that perhaps also merit some attention” (*Überlegungen über Zetafunktionen elliptischer Kurven anknüpfen, die vielleicht auch einige Aufmerksamkeit verdienen*). These reflections came very close to, but stopped short of proposing the conjecture, now solved, whose name has since been a matter of controversy:

Conjecture I. The zeta function of any elliptic curve with rational coefficients coincides with the zeta function of a modular form of the type specified in the main theorem of Weil’s paper.

⁽²⁾However, William Thurston famously wrote that “The product of mathematics is clarity and understanding. Not theorems, by themselves”. <https://mathoverflow.net/posts/44213/revisions>.

⁽³⁾The canonical reference for the role of conjecture within mathematical practice is still Barry Mazur’s article on the subject: Mazur, B., Conjecture. *Synthese*, vol. 111, no. 2, 1997, pp. 197–210. JSTOR, www.jstor.org/stable/20117628.

I choose to formulate this sentence as a conjecture for the sake of clarity, and to set the conjecture in boldface type, but Weil presented his “reflections” in normal type. Moreover, instead of coming out and stating a conjecture, he explained that this is sometimes the case — here he cited a result of Goro Shimura — but added that “whether things always behave that way” appears for the moment to be problematic (*scheint im Moment noch problematisch zu sein*) and may be recommended to the interested reader as an exercise (*mag dem interessierten Leser als Übungsaufgabe empfohlen werden*).

These last sentences came at the end of a paragraph that began with Weil’s very strong hints in favor of believing Conjecture I: “On certain theoretical grounds one may with reasonable certainty suspect”⁽⁴⁾ that something is true that entails the claim, in view of the main result of the paper. This “suspicion” — in a rather stronger form, mentioned in that same paragraph of Weil’s paper — soon was circulating under the name of the *Weil Conjecture*. I will have to provide some background in order to explain the nature of these hints, but for the moment let me focus on the superficial aspect of the situation. Weil has drawn his readers’ attention to a series of steps, some of them proved in his paper, some “reasonably certain”, leading to the suspicion and, with another step that is “of course obvious to expect” (*natürlich naheliegend zu erwarten*), to the stronger form that will be explained in a moment.

§ — **Lightning introduction to elliptic curves, modular forms, and their Dirichlet series.**

But first I had better review the mathematical notions invoked in the “suspicion”. An *elliptic curve with rational coefficients* is a geometric object E with number theoretic meaning: it is the set of solutions to a polynomial equation in two variables of the form

$$y^2 = x^3 + ax + b \quad (\star)$$

where a and b are rational numbers (with the property that the polynomial on the right-hand side of the equation has no multiple roots). A modular form is a different kind of geometric object

⁽⁴⁾In the original German: “Aus gewissen theoretischen Gründen darf man mit ziemlicher Sicherheit vermuten”. The German admits more than one translation. The online dictionary <https://www.linguee.com/german-english/translation/> provides sentences in which *mit ziemlicher Sicherheit* means “almost certainly”. Moreover, *Vermutung* is also the first and best-known German equivalent of the English word *conjecture*, which leads one to wonder whether Weil was being entirely sincere when in 1979 he argued against conjectures in general.

with a very different number-theoretic meaning. To explain this, we consider the set of complex numbers $\mathfrak{H} = \{z = x + iy, y > 0\}$ —the upper half-plane in the complex plane. The 2×2 matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with determinant 1 form a group $SL(2, \mathbb{Z})$ that acts on \mathfrak{H} by the formula

$$z \mapsto \frac{az + b}{cz + d}$$

Whenever Γ is a subgroup of finite index in $SL(2, \mathbb{Z})$, the quotient $\Gamma \backslash \mathfrak{H}$ can be identified with a compact Riemann surface with some points missing; adding the points yields the *modular curve* $X(\Gamma)$. A modular form of weight 2 and level Γ is a complex analytic function $f(z)$ on \mathfrak{H} that is not quite invariant under Γ ; rather the differential form $f(z)dz$ is invariant under Γ .

There is a procedure A for associating a Dirichlet series to such an f , and there is also a procedure B for associating a Dirichlet series — Weil denotes it Λ — to an elliptic curve E . Weil proves that any Dirichlet series that satisfies a property (A3) — the “functional equations” (plural) of Weil’s title — arises from procedure A — with specified *level subgroup* Γ — and alludes to “certain theoretical grounds” for believing that the Dirichlet series arising from procedure B also satisfy property (A3). This motivates the original suspicion, which I now present in full:

Aus gewissen theoretischen Gründen darf man mit ziemlicher Sicherheit vermuten, dass Λ eine Funktionalgleichung besitzt... Weiter lässt sich vermuten [that the functional equations of condition (A3) also hold upon twist by characters].

On the other hand, E can also be viewed as a Riemann surface, and there is a stronger form of the conjecture:

Conjecture II. Any elliptic curve E with rational coefficients admits a modular parametrization: a non-trivial map of Riemann surfaces from $X(\Gamma)$ to E , again with specified Γ .

For nearly 20 years, such a map was called a *Weil uniformization*, in conferences and seminars (including many that I attended) as well as in a number of influential papers, reflecting the consensus that Weil had formulated an idea whose implications included the existence of such a map. But then things got more complicated.

§ — One conjecture, many names.

Actually, things were already more complicated within a few years of the publication of Weil's papers, because it turned out that Yutaka Taniyama had listed something very like Conjectures I and II as items 12 and 13 of a collection of 36 problems. This list was published in Japanese in his collected works, but it was handed out in English at the 1955 Tokyo-Nikko conference on number theory, attended by Weil, Shimura, and Jean-Pierre Serre,⁽⁵⁾ who also has a part to play in this story. Great things were expected of Taniyama, but he committed suicide a few years after 1955 and is now mainly remembered for the theory of complex multiplication that he developed with Shimura, and for his part in formulating the boldface statement Conjecture I, whose official name had become the *Taniyama-Weil Conjecture*, no later than 1977. By then it was generally understood that the statement would follow from the *Langlands program*, a vast and ambitious network of conjectures and theorems that linked *automorphic forms* — generalizations of modular forms — to generalizations of the number-theoretic structures derived from elliptic curves. And it was generally believed that the Langlands program was a distant fantasy, in large part because no one had any idea how to begin to prove the (then-so-called) Taniyama-Weil Conjecture — which I will henceforth call the *Modularity Conjecture*.

Positions hardened about halfway between the announcement by Frey and Ribet that Fermat's Last Theorem would be a consequence of the Modularity Conjecture, and the announcement by Wiles of his proof (later completed in a paper of Taylor-Wiles) of enough cases of the Modularity Conjecture to derive Fermat's Last Theorem. In February 1988, before Wiles announced his result, Joseph Oesterlé had published a Bourbaki seminar talk whose first theorem, attributed to Ken Ribet, asserts that "*La conjecture de Taniyama-Weil implique le théorème de Fermat*". No paper by Shimura was cited in Oesterlé's bibliography. Ribet himself, writing in the United States two years later, published a paper in France with the title "From the Taniyama-Shimura conjecture to Fermat's last theorem", in which his reference to Weil's 1967 paper draws attention to what he considered to be Weil's contribution to the story.

⁽⁵⁾I am following Serge Lang's account in "Some history of the Shimura-Taniyama Conjecture", *Notices of the AMS*, 42, November 1995, 1301–1307.

Finally, A. Weil proved in [35] that an elliptic curve E over \mathbb{Q} is modular provided that its L -function $L(E, s)$ has the analytic properties one expects of it.

Mathematicians' choice of terminology for the conjecture largely broke down along national lines, and remains contested more than 20 years after a complete proof of the Modularity Conjecture was published by Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor.⁽⁶⁾ Public disputes over priority were acrimonious on occasion, private disputes much more often. My own discomfort with the situation, and my dissatisfaction with the quality of the arguments put forward in support of one position or the other, led me to seek clarity and consolation in philosophy. Thus, by a circuitous route, I find myself writing an article for a philosophical journal on the topic that revived my interest in philosophy in the first place.

§ — One conjecture, many virtues: comparing the cases for Taniyama, Shimura, and Weil.

In fact, although the difference of opinion had all the characteristics of a vulgar personality conflict — it was common knowledge that relations between Serre and Shimura were difficult, and the pages Shimura devotes to Serre in his memoirs are so gratuitously nasty that one wonders why Springer-Verlag agreed to publish them — arguments were advanced in support of each of the attributions. It is these arguments that deserve to be evaluated for their implicit or explicit assumptions about the virtues attributed to priority.

⁽⁶⁾Wikipedia is an international battleground. The English, German, French, Russian, and Arabic pages are entitled “The Modularity Theorem” but disagree on its former terminology as conjecture: in English and German it was the “Taniyama-Shimura Conjecture”, in Russian the “Taniyama-Shimura-Weil”, and in French either “Taniyama-Weil”, “Shimura-Taniyama-Weil”, or “Shimura-Taniyama”. The Arabic page mentions Taniyama-Shimura-Weil “and other names”. The Spanish, Italian, Swedish, Hebrew, Catalan — and, as far as I can tell, the Chinese — pages are entitled “The Taniyama-Shimura Theorem”; the Portuguese page calls it the “Shimura-Taniyama-Weil Theorem”; and the Dutch page is “The Shimura-Taniyama Theorem”. Wikipedia pages on mathematical topics are generally maintained by mathematicians, who, as far as I know, do not report to their national authorities. The proof by Breuil et al. is in the article “On the modularity of elliptic curves over \mathbb{Q} : wild 3-adic exercises”, *Journal of the American Mathematical Society*, 14 (4): 843–939 (2001).

1. The case for Taniyama. By all accounts, Taniyama was the first to have expressed an expectation along the lines of the conjecture, in items 12 and 13 of his list of problems at the Tokyo-Nikko conference. It is also agreed that his version was not literally correct as stated.⁽⁷⁾ I copy the English translation of Taniyama's questions from Lang's article in the *AMS Notices*:

12. Let C be an elliptic curve defined over an algebraic number field k , and $L_C(s)$ denote the L -function of C over k . Namely,

$$\zeta_C(s) = \frac{\zeta_k(s)\zeta_k(s-1)}{L_C(s)}$$

is the zeta function of C over k . If a conjecture of Hasse is true for $\zeta_C(s)$, then the Fourier series obtained from $L_C(s)$ by the inverse Mellin transformation must be an automorphic form of dimension -2 , of some special type (cf. Hecke). If so, it is very plausible that this form is an elliptic differential of the field of that automorphic functions. The problem is to ask if it is possible to prove Hasse's conjecture for C , by going back this considerations, and by finding a suitable automorphic form from which $L_C(s)$ may be obtained.

13. Concerning the above problem, our new problem is to characterize the field of elliptic modular functions of "Stufe" N , and especially, to decompose the Jacobian variety J of this function field into simple factors, in the sense of isogeneity.

It is well known, that, in case $N = q$ is a prime number, satisfying $q \equiv 3 \pmod{4}$, J contains elliptic curves with complex multiplication. Is this true for general N ?

2. The case for Shimura. (a) The case for Shimura's priority was made by Lang in his 1995 article: Lang established that Shimura had actually stated a version of the conjecture in the presence of Serre and Weil, at a specific place (the Institute for Advanced Study in Princeton) and during a specific time interval ("most likely in

⁽⁷⁾This opinion is shared by Lang, Serre, and Shimura. In a letter to Lang, Shimura wrote that Taniyama "was not completely careful, and if someone had pointed out this, he would have agreed that the problem would have to be revised accordingly".

1964”, according to a private message⁽⁸⁾ Shimura sent me in 1998). Weil mentions this conversation in his 1979 comments on his 1967 article, included in his collected works.⁽⁹⁾ The chronology is important for Lang, since the conversation definitely took place before Weil published his result.

When I first read Weil’s answer about “one and the other set being denumerable”, I characterized it as “stupid”. I have since also called it inane. But actually, Weil’s answer gives further evidence that he did not think of the conjecture himself. Indeed, as a result of his conversations with Serre and Weil, Shimura was directly responsible for changing the prevailing psychology about elliptic curves over \mathbb{Q} [my emphasis; we will be returning to psychology].⁽¹⁰⁾

(b) Shimura added two arguments that need to be considered separately. In the first place, there is the important theorem that he first published in 1971 in his textbook *Introduction to the Arithmetic Theory of Automorphic Functions*: namely, that

Theorem (Shimura). To a modular form of weight 2 for $\Gamma_0(N)$ with rational Fourier coefficients one can associate an elliptic curve with rational coefficients.

This is the evidence Weil cited in 1967 (as a *Mitteilung* from Shimura) and about which he asks “whether things always behave that way”.

Twenty years ago, Shimura was attaching a great deal of importance to being credited with the proof of this theorem. He was unhappy with Wiles’s discussion of this and a related point in one of his articles on Fermat’s Last Theorem:

“Strangely Wiles says ‘Any such elliptic curve has the property... has an analytic continuation ...’ without saying that the result is due to me. This is another defect. Does he think that it is trivial?”⁽¹¹⁾

⁽⁸⁾Shimura, private emails to the author, May 6-12, 1998.

⁽⁹⁾Weil, *Oeuvres scientifiques*, Vol III, p. 450.

⁽¹⁰⁾Lang, p. 1304.

⁽¹¹⁾Shimura, op. cit., May 11, 1998. In his letter to Serge Lang, February 14, 1999, appended to this article, he refers to “one more annoying item” in Wiles’s article: “He attributes that statement to nobody. I wonder if he considers it trivial. You may call it a non-reference by Wiles”.

The letter to Lang mainly complains that Wiles failed to appreciate the importance of the boldfaced Theorem, and only mentions the result about analytic continuation in the final paragraph. This observation could well serve as the starting point for an analysis of the virtues inherent in a new theorem; but I don’t have enough information to reconstruct Shimura’s thoughts on the matter.

(c) Shimura's third argument is of a different nature again. In his remarks on receiving the American Mathematical Society's Steele Prize for Lifetime Achievement, Shimura explained his motivation:

*...at first I thought that ... curves obtained from a division quaternion algebra B over \mathbb{Q} might not be modular... but I realized that no nonmodular \mathbb{Q} -rational elliptic curves could be obtained for the following reason: Eichler had shown [that] the Euler products on B are already included in those obtained from elliptic modular forms [2]. The Tate conjecture on this was explicitly stated much later, but the idea was known to many people, and so **it was natural for me to think** that two elliptic curves with the same zeta function are isogenous. This fact concerning B , in addition to the results I had about the zeta functions of modular curves, may have been the strongest reason for my stating the conjecture that every \mathbb{Q} -rational elliptic curve is modular.⁽¹²⁾ [my emphasis, to be explained later]*

3. The case for Weil. Of course Weil gets credit for writing and publishing the paper that includes his converse theorem, an insight at the basis of an entire research program. A vast generalization of the converse theorem by James Cogdell and Ilya Piatetski-Shapiro has been indispensable in 20 years of progress on the Langlands program.

Weil may have remembered the conjecture on elliptic curves from that 1964 conversation with Shimura. This doesn't strengthen the case for attaching his name to the conjecture. But Weil's paper included a major advance, already mentioned above, on the previous versions of the conjecture: he suggested that an elliptic curve E is a quotient of the Jacobian of the modular curve with level subgroup $\Gamma = \Gamma_0(N)$, where N is an explicit integer, the conductor of E , defined in terms of the arithmetic of E . This makes the conjecture considerably more precise: there is an equivalence between two sets of objects that matches a specific numerical property; moreover for a given N , the set of such objects is finite. Serre wrote that this was a *beautiful new idea ; it was not in Taniyama, nor in Shimura (as Shimura himself wrote to me after Weil's paper had appeared)*". But he then added an explicit value judgment:

*Its importance comes from the fact that it made the conjecture **checkable numerically** (while Taniyama's statement was not). I remember vividly when Weil explained it to me, in the summer of 1966, in some Quartier Latin coffee house. Now things really began to make sense. Why no elliptic*

⁽¹²⁾ Shimura, *Notices of the AMS*, November 1996, p. 1345.

curve with conductor 1 (i.e. good reduction everywhere)? Because the modular curve $X_0(1)$ of level 1 has genus 0, that's why! I went home and checked a few examples of curves with low conductor : I did not know any with conductor < 11 nor with conductor 16? No surprise, since $X_0(N)$ has genus 0 for such values of N , etc. Within a few hours, I was convinced that the conjecture was true.⁽¹³⁾

The virtue Serre identified in Weil's formulation — that it lends itself to numerical verification — has been taken up by defenders of the claim for Weil's priority; most of the defenders I know happen to be French.

§ — “Conceptual evidence”.

We should also hear what Andrew Wiles himself has to say about the origin of the modularity conjecture. Wiles may have been thinking of the virtue mentioned by Serre in the first page of his paper on Fermat's Last Theorem:

A well-known conjecture which grew out of the work of Shimura and Taniyama in the 1950's and 1960's asserts that every elliptic curve over \mathbb{Q} is modular. However, it only became widely known through its publication in a paper of Weil in 1967..., in which, moreover, Weil gave conceptual evidence for the conjecture.

At the time, though, I understood that the “conceptual evidence” was the derivation of modularity from the characterization of modular forms given by Weil's converse theorem, together with the Hasse conjecture, in Weil's version (A3), and the Tate conjecture, mentioned below.

It's important to emphasize that all three protagonists — Taniyama, Shimura, and Weil — justified their speculation in roughly the same terms. Taniyama's 12th problem includes the following sentence:

If a conjecture of Hasse is true ... then the Fourier series obtained from $L_C(s)$ by the inverse Mellin transformation must be an automorphic form of dimension -2 , of some special type (cf. Hecke).

Shimura never published any version of his conjecture, but Lang reported in 1995 that

⁽¹³⁾Jean-Pierre Serre, “L'histoire de la ‘modularity conjecture’”, *Gazette des mathématiciens*, **91**, Janvier 2002, 55–57.

The rationale for Shimura's conjecture was precisely the conjectured functional equation (Hasse), along the lines indicated in Taniyama's problem 12, suitably corrected. Shimura's bolder insight was that the ordinary modular functions ... suffice to uniformize elliptic curves defined over the rationals.

Shimura confirmed this a few years later in a private communication.⁽¹⁴⁾ And Hasse's conjecture constituted the "certain theoretical grounds" [*gewisse theoretische Gründe*] to which Weil alluded (elliptically) in his paper. However, Hasse's conjecture alone would have indicated that the function $L_C(s)$ satisfied some functional equation, but not necessarily the condition (A2) that had been shown by Hecke to characterize modular forms. So Weil actually took his speculations to the next step: "*Weiter lässt sich vermuten*",⁽¹⁵⁾ as indicated above, that $L_C(s)$ satisfies Weil's condition (A3), which Serre thought added the necessary precision to make the speculation of Taniyama and Shimura into a checkable and therefore full-fledged conjecture.

§ — Discussion: Naming Virtues.

1. Taniyama's clear and distinct idea. Taniyama's reasoning follows a pattern that is familiar to working mathematicians but that to my knowledge has not been taken up by philosophers. "If a conjecture of Hasse is true", he says, then something else must follow, because of Hecke's well-known characterization of modular forms.

⁽¹⁴⁾Shimura, unlike Wiles, did not see Weil's paper as a conceptual breakthrough. In his e-mail to me dated May 11, 1998, he wrote "'Conceptual evidence' is rather misleading. After all Taniyama and I were familiar with the results of Hecke on the characterization of an automorphic form by the functional equation of its Mellin transform, and our ideas were partly based on that correspondence. I always viewed Weil's result as its nice refinement. But "conceptual"? Really? I don't get it".

The word "conceptual" clearly continued to bother him, because in the letter to Serge Lang cited in footnote 10, he complained vigorously about Wiles's failure to understand his (Shimura's) contribution. Shimura sent me a copy of this letter and I have included a reproduction of the letter at the end of this article. The letter to Lang makes it clear, however, that Shimura was especially annoyed that Wiles and others failed to give him credit for the boldface Theorem cited above, under point (b) of the case for Shimura, and which he calls a "geometric construction".

⁽¹⁵⁾Hard to translate. On www.linguee.de one finds "it looks as if", "it can be assumed that", and "suggests that" among the translations, depending on the context.

Taniyama's 12th problem can be taken to be a sketch of a proof of a theorem:

Taniyama's "theorem". *If Hasse's conjecture is true, then the L-function of any elliptic curve over a number field k is an automorphic form of weight 2 of some special type.*

In more detail: (i) Hasse's conjecture implies that $L_C(s)$ satisfies a functional equation, and (ii) Hecke's theorem then implies that $L_C(s)$ is the L-function of an automorphic form. Step (ii) is false for at least two reasons: (a) Hecke's theorem only applies when k is the field of rational numbers, and (b) it only applies when the conductor N is 1, which is never the case (a point Serre mentioned in the text cited above). So the virtue of Taniyama's contribution seems to involve the sketch of an incorrect proof of an imprecisely stated and undoubtedly false theorem (see (a)). The mistakes do not enhance the contribution, of course; the imprecise and partially incorrect reasoning served as the basis of Taniyama's imaginative leap, making him the first to intuit an unexpected connection between elliptic curves and modular forms.

The imaginative leap was not totally unprecedented. Dirichlet series with Euler products and functional equations arise in number theory in two ways. One has to do with Galois theory, the other involves constructions in complex analysis that can be interpreted as some sort of harmonic analysis on topological groups.⁽¹⁶⁾ Class field theory, the major accomplishment of algebraic number theory in the first half of the 20th century, proves that both ways give rise to the same family of Dirichlet series when the structures involved (the Galois groups and the topological groups) are both abelian. Taniyama can be credited with the insight that something analogous is going on when the groups are not abelian. To the best of my knowledge, this sort of insight is considered contingent and thus not suitable as a topic for philosophical analysis. It is undoubtedly the most prized of all the mathematical virtues, however, and for that reason alone deserves the attention of anyone who is interested in mathematical practice. For example, the mathematical community would see no virtue in the publication of a new

⁽¹⁶⁾This perspective was already visible in Erich Hecke's theory of L-functions of Grössencharaktere, and was made systematic in John Tate's influential thesis of 1950. Kenkichi Iwasawa had independently discovered an essentially identical method slightly earlier. Although Iwasawa never published his results and never wrote them up in detail in English, there has never been a priority dispute.

speculative intuition every week on the hope that one of them will turn out to be right. A new conjecture not accompanied by a compelling motivation — what we might call a *clear and distinct idea* — will win no credit for its author. In this respect the reasoning that accompanied Taniyama’s imaginative leap qualifies as *cartesian* in Ian Hacking’s sense: any reader familiar with both Hasse’s conjecture and Hecke’s theorem will immediately see the pertinence of Taniyama’s 12th problem.

The apparent importance of the cartesian character of Taniyama’s insight has an apparently paradoxical implication that deserves to be stressed: there is more virtue in Taniyama’s incorrect formulation of his 12th problem, accompanied as it was by steps (i) and (ii) above, than there would have on the counterfactual assumption that he had presented with no justification whatsoever the fully-formed and correct modularity conjecture, as later verified by Breuil-Conrad-Diamond-Taylor. Something about this paradox feels wrong, but I am not sure what it is.

Of course, mathematicians generally show more appreciation for a conjecture when it is accompanied by a proof, however uncartesian it might be, than when it is merely accompanied by a clear and distinct idea. Nevertheless, although Hacking thought cartesian proofs are rare, I would argue that cartesian insights are fundamental to mathematical communication and progress. Everyone remembers that Caesar crossed the Rubicon; only eyewitnesses and specialized historians remember the details of what happened when he got to the other side.

Before turning to the other two claimants to priority, I remind the reader that the 13th problem on Taniyama’s list was a version of Conjecture II. I have alluded indirectly to the possession of the “clear and distinct idea” underlying this intuition on the part of each of the three protagonists. In Taniyama’s case, it is implicit in the words “Concerning the above problem” that form the transition between his problems 12 and 13, together with his use of the word “isogeneity”. This is a basic concept in the theory of *abelian varieties*, a class of algebraic varieties to which elliptic curves belong. An abelian variety is a projective algebraic variety which is also an algebraic group — its group structure is given by algebraic maps. To an algebraic curve is canonically assigned an abelian variety called its *Jacobian*; the Jacobian of the modular curve $X(\Gamma)$ is denoted $J(\Gamma)$. The theory of abelian varieties is one of the best understood parts of algebraic geometry. In particular, every abelian variety is

a product of “simple factors” but only “up to isogeny”. An isogeny between abelian varieties is an isomorphism up to finite error: it is a surjective homomorphism of algebraic groups with finite kernel. Taniyama’s problem 13 says the Jacobian $J(\Gamma)$ for Γ of “stufe” N “contains” a certain elliptic curve; by “contains” he means that it is one of its simple factors up to isogeny.

We have seen that an L-function can be attached to any elliptic curve, but in fact Hasse’s conjecture (or the Hasse-Weil conjecture) concerns the L-function attached to any algebraic variety. In particular, one can attach an L-function to an abelian variety. It has been known for a long time that two isogenous abelian varieties have the same L-function. Recall what Shimura wrote in 1996: *it was natural for me to think that two elliptic curves with the same zeta function are isogenous*. For want of a better account of the *nature* to which Shimura is referring here, let’s assume that this “idea”, of which he claimed that it was “known to many people”, was “clear and distinct” in their minds. It was this idea to which Weil alluded when we wrote that it was “of course obvious — *natürlich naheliegend zu erwarten*”. Here is the full citation:

“It is of course obvious to expect that under these conditions C' [the elliptic curve factor of the Jacobian attached to the modular form by Shimura] is isogenous to C [the original elliptic curve that on “certain theoretical considerations” should be modular]; this is in fact confirmed in some cases”. (*Es ist natürlich naheliegend zu erwarten, da unter diesen Umständen C' mit C isogen ist; das bestätigt sich tatsächlich in einigen Fällen.*)

With this in mind, problem 13 can be read in two ways. First, is it true that every $J(\Gamma)$ of “stufe” N has elliptic curves among their simple factors? This is the most obvious reading of the text, but it is a strange question, because there are many modular curves whose Jacobians do not involve any elliptic curves. The question was given a precise answer by Shimura’s analysis of the Jacobians of modular curves in his 1971 textbook.

The alternative reading is that every elliptic curve with rational coefficients is isogenous to a factor of the Jacobian of some modular curve. This is now known to be true, as a consequence of two very substantial theorems: the theorem of Breuil-Conrad-Diamond-Taylor on the modularity of elliptic curves, and the 1983 theorem of Gerd Faltings that includes the *Tate conjecture* that

Shimura had mentioned: that two abelian varieties over a number field whose L-functions coincide are necessarily isogenous.⁽¹⁷⁾ This is the expectation that Shimura found “natural” and Weil found “obvious”, once Conjecture I has been established. While one needs a good deal of expertise and sophistication in order to perceive Conjecture II as an obvious or natural consequence of Conjecture I and Tate’s Conjecture, the choice of words suggests that neither Weil nor Shimura is claiming priority for this particular aspect of the conjecture. So I will henceforth restrict my attention to Conjecture I.

2. Lang’s realism, Shimura’s phenomenology. Lang’s case for Shimura’s priority is primarily forensic in nature: his detective work places Weil and Serre in conversations with Shimura on this topic during a precise period in the early 1960s. Weil and Serre were skeptical when Shimura asserted that he “believed [that an elliptic] curve [over \mathbb{Q}] should always be a quotient of the Jacobian of a modular curve” (both Conjectures I and II, in other words). Since Weil was not only a skeptic at the time but even responded in a way Lang characterized as “stupid” and “inane”, how could anyone assign him any credit for the conjecture?

If Lang’s approach can be assigned to any philosophy, I would tend to call it *empiricist* and *realist*: to the extent that there is a fact of the matter (and Lang clearly believes there is), it can be established by gathering evidence. But Lang made a comment that calls this assignment into question: as we have seen, he claimed that “Shimura was directly responsible for changing the prevailing psychology about elliptic curves over \mathbb{Q} ”. We can agree that something changed about the way number theorists were thinking about elliptic curves over \mathbb{Q} without committing ourselves to Lang’s use of the

⁽¹⁷⁾ At some point in the 1980s I had the disconcerting experience of being asked by Shimura whether or not it was true that Faltings had proved this theorem. Unfortunately, I don’t remember whether it was before or after Faltings had joined Shimura in the Princeton mathematics department, nor whether it was before or after the work of Frey and Ribet, around 1985-86, that showed how Fermat’s Last Theorem would follow from the Modularity Conjecture, in view of Faltings’s theorem. Either way, Shimura was surrounded by people who were far better qualified than I to answer his question.

Tate’s conjectures on algebraic cycles, including the one about isogenies of abelian varieties that Faltings proved, were made public in the 1960s. They are fundamental to our thinking about the Galois representations attached to the l-adic cohomology of algebraic varieties, but I am not familiar with their prehistory and cannot comment on why Taniyama, Shimura, and Weil found it “natural to think” that it was valid for elliptic curves.

word “psychology”, nor do we have to accept his judgment that Shimura had direct responsibility for the change by speaking with Weil and Serre. Let’s say that we could ascribe to the community of number theorists the *disposition* to think about elliptic curves over \mathbb{Q} in certain ways in 1960, and a *markedly different disposition* in 1970. Then we can say that Lang counts among the prized virtues of mathematical practice *the ability to alter dispositions* with regard to important topics, and that this is one of the main virtues he identifies in Shimura’s role in proposing the modularity conjecture. This doesn’t suffice as a characterization of Shimura’s contribution, however. Other distinguished number theorists, including Barry Mazur, Yuri Manin, and especially John Tate, as well as Serre himself, were busily altering dispositions with regard to elliptic curves over \mathbb{Q} in the decade following the publication of Weil’s article, by giving lectures and writing survey articles⁽¹⁸⁾ promoting a very precise and detailed international research program that continues to frame thinking about elliptic curves even now. One presumes that they did so with dispositions altered as a direct result of Shimura’s intervention with Serre and Weil a few years earlier — and, I should add, with Tate (“I told it to Tate around the same time, and I remember that he was impressed”).⁽¹⁹⁾

Lang is no longer among us, and we don’t know what he had in mind when he refers to Shimura’s “changing the prevailing psychology”. My guess is that he would like his readers to believe that Weil’s conversation with Shimura inspired him to think about the questions that led to his 1967 paper, inviting us to interpret his failure to mention either the contents of the conversation or Taniyama’s 1955 questions as a deliberate attempt to claim the virtue of the original idea of Taniyama and Shimura for himself. That interpersonal conflict does not lend itself to philosophical analysis, whereas “dispositions to think about” various mathematical objects can be legitimately understood as targets of virtuous intervention within the community of mathematical practitioners. The dispositions, and the interventions to change them, can thus be incorporated in a philosophy of mathematical practice, whether or not Shimura’s private conversations were as effective as Lang claimed. At the very least, this way of looking at things raises interesting questions about

⁽¹⁸⁾Especially J. Tate, The arithmetic of elliptic curves, *Invent. Math.*, **23** (1974), 179–206, which is where I and everyone else I knew developed our dispositions at the time.

⁽¹⁹⁾Shimura, letter to the author, May 11, 1998.

the relative virtue of private and public interventions within the community — and we should not underestimate the importance of the letters of recommendation with which this essay began in creating dispositions in the first place.

Shimura's own arguments (b) and (c) make no reference, explicit or otherwise, to the community of mathematical practitioners. To put them in perspective, it's helpful to rewind to the early 1950s, when Hasse's conjecture was known in very few cases but was the basis of a very active research program. Weil, Taniyama, and Shimura were reporting on their contributions to this program as early as the 1955 Tokyo-Nikko Conference, but their results (like those of Max Deuring, also present at the conference) were based on what was already known about abelian class field theory. The first non-abelian example had been treated in a 1954 article by Martin Eichler. Shimura spent much of the next 20 years generalizing Eichler's observation and creating what is now called (following Pierre Deligne) the theory of *Shimura varieties*, of which the simplest (but by no means simple) case is the one treated in his 1971 textbook. In that book, and in a series of papers dating back to 1958, Shimura developed a systematic method for verifying new non-abelian cases of Hasse's conjecture for elliptic curves, in the process providing evidence for both Conjectures I and II. Lang puts it this way:

Shimura himself in the late fifties and sixties extended Eichler's results and proved that elliptic curves which are modular have zeta functions which have an analytic continuation. (Cf. the three papers [Sh 58], [Sh 61], and [Sh 67].)⁽²⁰⁾

Thirty years earlier, this had already been mentioned in a survey article of J.W.S. Cassels that was largely responsible for creating the "prevailing psychology" about elliptic curves in the 1960s:

Hasse's conjecture has also been verified for a few C without complex multiplication which arise in the theory of modular functions [Shimura (1958a, 1961a, 1964a)]⁽²¹⁾

What should we call the virtue, exemplified by this evidence, on which Shimura bases his claim (b) to priority? Wikipedia suggests one possible name:

⁽²⁰⁾ Lang, p. 1303.

⁽²¹⁾ Cassels, J. W. S., Diophantine equations with special reference to elliptic curves. *J. London Math. Soc.*, **41** (1966) p. 280.

Shimura's approach, later presented in his monograph, was largely **phenomenological**, pursuing the widest generalizations of the reciprocity law formulation of complex multiplication theory. [my emphasis]⁽²²⁾

Wikipedia's history page informs me that on March 31, 2007, Charles Matthews inserted the word that I have set in boldface, together with its philosophical overtones. The word is not used in Husserl's sense but rather as it is understood in philosophy of science. Shimura developed the theory in order to establish the facts of the matter.⁽²³⁾ In the course of this development, he made the observation reported above in argument (c): that the phenomena (elliptic curves) observed in his study of the Shimura varieties attached to quaternion division algebras provided additional evidence for Conjectures I and II.

Thus Shimura claimed the virtue that Matthews calls *phenomenological*. As he himself acknowledged, it built upon Taniyama's original *cartesian* insight, but from the standpoint of philosophical analysis it is conveniently different. Nevertheless, I don't want to leave Shimura without pointing out that mathematical practice sees an unbridgeable gulf between *providing* (*phenomenological*) *evidence* for a conjecture and *proving* the conjecture. This is obvious, of course, and since no one is claiming that any of the triad (Taniyama, Shimura, Weil) came close to proving the conjecture, its relevance to identifying the virtues of a conjecture may not be clear. This may best be understood *a posteriori*, in the light of the Langlands program for automorphic forms, specifically the program Langlands outlined in 1979 for the study of Shimura varieties. During the 1960s, Alexander Grothendieck and his school had developed the theory of étale cohomology, which made it possible to attach Dirichlet series to automorphic forms in both of the ways mentioned previously. Étale cohomology provided the connection with Galois theory, while the relation to *representation theory* — what we previously called harmonic analysis on topological groups — gave a variety of constructions of Dirichlet series. The generalization of Conjecture I (no uniform generalization is possible for Conjecture II) asserts roughly that

A. Galois-theoretic Dirichlet series are representation-theoretic Dirichlet series

⁽²²⁾Wikipedia, "Shimura variety", https://en.wikipedia.org/wiki/Shimura_variety.

⁽²³⁾A Wikipedia editor named Arcfrk emphasized (in 2010) the contrast with Deligne's definition, which served "to isolate the abstract features that played a role in Shimura's theory".

Specialists call this the “Galois-to-automorphic” direction of Langlands’s conjectures; current work on this direction largely derives from the methods Wiles developed in his proof of Fermat’s Last Theorem. The “automorphic-to-Galois” direction, naturally, asserts

B. *Representation-theoretic Dirichlet series are Galois-theoretic Dirichlet series*

Shimura’s phenomenological efforts established **B** for the class of modular forms that is relevant to elliptic curves with rational coefficients, and provided the starting point for all subsequent work on **B**, which by now is established in considerable (though by no means complete) generality. But it’s not immediately obvious why even a complete proof of **B** should help to establish a disposition to believe **A**, as it apparently did for Shimura.

Shimura actually did prove **A** — Conjecture I for the class of *elliptic curves with complex multiplication* — in a paper published in 1973.⁽²⁴⁾ Even though this class of elliptic curves has especially favorable properties (the problem in this case is very close to being abelian), phenomenological results of this kind typically do count as significant evidence for **A** within the framework of the Langlands program, since it is generally believed that we will have to wait several generations, if not centuries, for a complete proof of **A**.

3. Weil’s falsificationism, on Serre’s reading. Unlike some of his Bourbaki colleagues, Serre has never displayed an interest in philosophy for its own sake, but the virtue he discerns in Weil’s contribution to the conjecture is the easiest to translate into familiar philosophical terms. If we replace the word “checkable” with its synonym “verifiable”, then Weil’s virtue, according to Serre, is that he has endowed the Modularity Conjecture with a *meaning*, following the *logical positivist* dictum that (in the words of Schlick) “The meaning of a proposition is the method of its verification”.⁽²⁵⁾ In practice, though, *falsification* of the Modularity Conjecture by finding a counter-example would have carried more weight than its verification in special cases, if for no other reason than that

⁽²⁴⁾Goro Shimura, On the factors of the jacobian variety of a modular function field, *J. Math. Soc. Japan*, 25 (1973), 523–544. This theorem can be proved using Weil’s converse theorem, but this was not the method Shimura chose, although his paper on modular forms of half-integral weight, published the same year, was based on a highly original application of Weil’s 1967 paper.

⁽²⁵⁾Moritz Schlick, “Meaning and verification”, *The Philosophical Review*, 45, No. 4 (Jul., 1936), 339–369.

a proof by (philosophical rather than mathematical) induction would require infinitely many verifications, which is impossible, whereas a single counter-example suffices for disproof. Thus the virtue Serre highlights may just as well be *popperian*. Either way, by insisting on the virtue of making a conjecture “checkable”, Serre seems to be in agreement with Mach, when he asserted that “where neither confirmation nor refutation is possible, science is not concerned”.⁽²⁶⁾

The clarity of Serre’s argument simplifies the philosopher’s task; he has identified a virtue in conjecture-making that is both unambiguous and easy to map onto familiar concerns in the philosophy of science. I would go further and say that Serre has rendered a service to philosophy: while both the verificationism of the logical positivists and Popper’s falsificationism are known to be vulnerable to critique, on the grounds that there is more to the world than its logical *Aufbau*, Serre’s “checkability” may well be a viable criterion for the meaningfulness of a conjecture. At any rate, it has not been subject to philosophical analysis.

Most importantly, perhaps, Serre has shown some consistency in following his own dictum: Serre formulated his famous Conjecture on modular representations of Galois groups in 1973 but only published it in 1987, after, as he put it in his thanks,

Jean-François Mestre ...a réussi à programmer et vérifier un nombre d'exemples suffisant pour me convaincre que la conjecture méritait d'être prise au sérieux.⁽²⁷⁾

The examples occupy 11 of the 52 pages of his manuscript.

§ — Conclusion: the “prevailing psychology”.

Lang indirectly hinted at a model of the *mathematical subject* by assigning to Shimura direct responsibility “for changing the prevailing psychology” by privately conveying his belief in the truth of what we are calling the Modularity Conjecture to two or

⁽²⁶⁾Mach, Ernst, 1883, *Die Mechanik in ihrer Entwicklung*, Leipzig: Brockhaus, 9th ed. transl. by T.J. McCormick, *The Science of Mechanics*, Chicago: Open Court, 1960, p. 587.

⁽²⁷⁾Jean-Pierre Serre, Sur les représentations modulaires de degré 2 de $Gal(\bar{\mathbb{Q}}/\mathbb{Q})$, *Duke Mathematical Journal*, 54 (1): (1987) p. 180. Its proof some 20 years later, by Chandrashekhara Khare and Jean-Pierre Wintenberger, made extensive use of the methods introduced by Wiles, Taylor, and others in proving the Modularity Conjecture, to which Serre’s conjecture is very closely related.

three individuals. This subject is structured hierarchically, so that impressions acquired by those (like Weil, Serre, or Tate) at the pinnacle of the hierarchy have the power of detectably altering dispositions throughout the subject. Lang doesn't explain the means of this alteration, so we don't know the relative weight in the process of private exchanges, seminar presentations, or publications — to mention a few of the mechanisms with which most members of the community would be familiar.

But even granting that the dispositions of a collective subject can be altered as a result of a few brief conversations involving a group of exceptionally influential individuals, the question remains: why should there have been a “psychology about elliptic curves over \mathbb{Q} ” in the first place? Whether or not there is a virtue ethics for mathematics, in which every conceivable topic merits its own psychology, in practice the attention of the mathematical subject is rather narrowly focused. Why were elliptic curves so much on the minds of number theorists in the 1960s?

It's easier to understand this attention when we remember that number theory at its core is concerned with understanding *Diophantine equations* — polynomial equations in several variables with integer coefficients— and more specifically with characterizing their set of solutions in integers, or in rational numbers. Equation (★) that we encountered at the beginning of the article, the one whose solution set is an elliptic curve, is a Diophantine equation. It is among the simplest possible classes of Diophantine equations, whose theory has been studied intensively since the 19th century (if not longer), and yet the nature of its solution set remains an outstanding open question. Much was known by the 1960s, however. It was proved by L. E. J. Mordell nearly 100 years ago that its set of rational solutions forms a finitely generated abelian group, a result generalized in André Weil's thesis and now known as the *Mordell-Weil Theorem*. This landmark in number theory, arguably the deepest result known at the time about Diophantine equations, posed a challenge to number theorists: how many generators does the group have? And how can we tell whether the group is infinite or finite? The challenge was more acute in view of the special status of elliptic curves among all algebraic curves⁽²⁸⁾ with rational coefficients. Every such curve has a genus, a non-negative integer; elliptic curves are the curves of genus 1. The rational solutions of

⁽²⁸⁾For our present purposes, an algebraic curve is defined by an equation in two variables, like (★).

curves of genus 0 — like the circle with equation $x^2 + y^2 = 1$ — are well understood (rational points on the circle are essentially the same as right triangles with integer sides). The *Mordell Conjecture* — proved by Faltings in 1983 — is the statement that curves of genus 2 have only finitely many solutions. For elliptic curves, simple as they are, there is no elementary way to determine whether the (rational) solution set is finite or infinite.

But in the early 1960s, when Shimura was privately conversing in Princeton with Weil, Serre, and Tate, over in England Bryan Birch and H. P. F. Swinnerton-Dyer were busy formulating and testing a *conjectural answer* to the question. The *Birch-Swinnerton-Dyer (BSD) Conjecture* — the appropriateness of its name has never been called into question! — asserts that the number of generators of the group of rational points of the elliptic curve C can be determined by examining its associated Dirichlet series — Taniyama’s $L_C(s)$, Weil’s Λ — but only if the Dirichlet series had (at least some of) the properties conjectured by Hasse.

The (eminently checkable) BSD Conjecture, which was widely known⁽²⁹⁾ by the time Weil’s article was published, gave such an appealing answer to the questions raised by the Mordell-Weil Theorem that one can easily understand why number theorists wanted the Conjecture to be meaningful — and thus for Hasse’s conjectures to be true for elliptic curves, and thus for Taniyama’s “theorem” to be true. With the help of the BSD Conjecture, the Modularity Conjecture thus becomes an object of number theorists’ collective desire — a symptom of our mass psychology, as Lang intuited.⁽³⁰⁾

⁽²⁹⁾Thanks to the 1966 article by Cassels cited in note 20, and of course to lectures by Birch and Swinnerton-Dyer themselves, especially the latter’s chapter in the book *Algebraic Number Theory* edited by Cassels and A. Fröhlich (London: Academic Press (1967)) which for 50 years has been required reading for every student of number theory.

⁽³⁰⁾The importance of the BSD Conjecture in promoting the Modularity Conjecture is beyond the scope of this article, but two landmark articles written in the decades following the publication of Weil’s paper give a sense of the excitement it generated. The introduction to “Arithmetic of Weil Curves”, by Mazur and Swinnerton-Dyer, published in 1974 (but available several years earlier) explicitly identifies the “aim of Part I of this paper”: to “examine... structures” that are “associated to the Weil parametrization and the bearing they have (via the conjectures of Birch and Swinnerton-Dyer) on the arithmetic” of elliptic curves over \mathbb{Q} . Part II as well as Part I became instantly applicable to all elliptic curves over \mathbb{Q} with the proof of the Modularity Conjecture. In 1983 Benedict Gross and Don Zagier announced their proof of their famous formula the first that explicitly constructs infinitely

Along with the Modularity Conjecture's implications for Fermat's Last Theorem, spectacularly realized by Wiles, this helps to explain why the controversy over the Conjecture's origin stirred up so much interest. On its own, however, it cannot settle the controversy. The virtues examined in this article — Taniyama's cartesian insight, Lang's realism, Shimura's phenomenology, and Weil's falsificationism — are by no means mutually exclusive. But nor does any of these virtues subsume any of the others. I would be delighted if my consideration of the virtues of the respective attributions helps to resolve the priority dispute in the minds of some of my readers. It has not settled the matter in my mind in one direction or another. For that reason I'm not sorry to see my colleagues have grown more relaxed about these attributions in recent years. I hope I have demonstrated that there are legitimate reasons to attach the names of one, or two, or three authors to the conjecture. As for myself, I prefer the version with no authors.

many rational points on an elliptic curve using information about its Dirichlet series. Since the Modularity Conjecture was not yet established, their article (published in 1986, before the controversy over priority broke out) explains that the results apply to "strong Weil curves" and refers to the Mazur-Swinnerton-Dyer article for the terminology. B. Mazur and H. P. F. Swinnerton-Dyer, Arithmetic of Weil Curves, *Inventiones Math.*, **25** (1974) 1–61; B. Gross and D. Zagier, Heegner points and derivatives of L-series, *Inventiones Math.* **84** (1986) 225–320.

Princeton University

Department of Mathematics
 Fine Hall—Washington Road
 Princeton, New Jersey 08544-1000

February 14, 1999

Professor Serge Lang
 Department of Mathematics
 Yale University
 New Haven, Connecticut 06520-8283

Dear Serge:

This may not be your concern, but in my case the problem is the attitude of those people who make inaccurate or insufficient statements about the conjecture. What Weil did poses no longer any problem. Let me give examples.

In his paper *Modular elliptic curves and Fermat's Last Theorem* Andrew Wiles wrote "Weil gave conceptual evidence for the conjecture." I have been puzzled by this statement, and so recently I asked him about its meaning. His explanation was as follows: Given a \mathbf{Q} -rational elliptic curve C , assume that the twists of the zeta function of C satisfy the functional equations of suitable types; then Weil's result gives a modular form f . Therefore C is modular, that is, my conjecture follows from the assumption concerning the functional equations combined with Weil's result.

I pointed out to him that the argument would produce a modular form, but it wouldn't say that C is a modular elliptic curve. He conceded that it didn't give any geometric construction of a modular elliptic curve, but added "Doesn't one call C modular if such an f exists?" He contradicts himself, because, the first line of his paper says "An elliptic curve over \mathbf{Q} is said to be modular if it has a finite covering by a modular curve of the form $X_0(N)$." Therefore, if he really wishes to speak of *conceptual evidence* in the above sense, then he has to say something more. I would advise him to read Weil's 1967-paper again. Weil was more accurate than Wiles about this, since he at least mentioned my work which associated a modular elliptic curve E to such an f . Besides, even after finding E , we have to invoke the result of Faltings, which was of course established many years later. I was careful about this point in my 1971-Nagoya paper concerning elliptic curves with complex multiplication.

There is one more annoying item in the paper of Wiles: After the definition of modularity, he says "Any such elliptic curve has the property that its Hasse-Weil zeta function has an analytic continuation and satisfies a functional equation of the standard type." He attributes that statement to nobody. I wonder if he considers it trivial. You may call it a non-reference by Wiles.

These are examples of what I mean by inaccuracies and insufficiencies. Worse, there are also some people who accept or even copy that statement about *conceptual evidence*.

Sincerely yours,


 Goro Shimura

1

Michael Harris,
 Department of Mathematics, Columbia University,
 New York, NY 10027, USA

★
 ★ ★