L'histoire de la "modularity conjecture"

J.-P. Serre

À la suite de la publication de la lettre de Serge Lang dans le courrier des lecteurs du numéro 90 de la Gazette[1], *Jean-Pierre Serre nous fait parvenir le texte d'une lettre qu'il a écrite récemment à D. Goss sur ce sujet.*

Lettre à David Goss, 30 mars 2000

Dear Goss,

You are right : the history of the "modularity conjecture" has been somewhat distorted recently, and it would be good to put the record straight. Let me try.
As you well know, the main actors have been Taniyama, Shimura and Weil. What they have published on it is as follows :

1. Taniyama. At the Tokyo-Nikko conference (1955), the organizers asked for a list of open questions. This list was typed and distributed to all participants (but it was not included in the Conference volume). Taniyama contributed several such questions. One of them (problem n° 12) is about elliptic curves ; you may find it (in Japanese) in his Collected Papers, p. 167, and (in English) in mine, Vol. III, p. 399 (see below). It is clear that Taniyama had been influenced by the results of Eichler of 1954 (which had also made a deep impression on Weil). His conjecture was, more or less, that Eichler's construction gives the zeta functions of all the elliptic curves over **Q**. Unfortunately, he chose to state it over an arbitrary algebraic number field ; this made him invoke a "field of automorphic functions" which does not make sense in such a general setting. Still, it was a brilliant insight.

2. Shimura. He clarified Eichler's theory by using the action of the Hecke algebra on the Jacobian of the modular curve (1958) ; this allowed him to split the Jacobian, up to isogeny. He obtained the "Eichler-Shimura" relation (for the reduction mod p of T_p) for large enough (but unspecified) primes p. It was Igusa (1959) who proved the important fact that this holds for every p not dividing the level.
As for the modularity conjecture, Shimura published nothing on it. He did not mention it (not even as a "problem") in his 1971 book, nor in any of the many papers on modular forms he wrote between 1955 and 1985. The most he did was to ask a few people (verbally, only) whether they believed in it or not. An explicit statement in print would have been more useful ; maybe he felt he did not have enough evidence to do so.

3. Weil. In his paper on "Funktionalgleichungen" (Coll. Papers, [1967a]), he mentions the conjecture, tongue-in-cheek, as an "exercise for the interested reader", without quoting Eichler or Taniyama (as he could have). He adds two decisive ingredients :
a) A characterization of modular forms by functional equations of Hecke type for the corresponding L functions, and their twists by Dirichlet characters. A remarkable aspect of his theorem is the way the constant of the functional equation depends on the twisting character. This has been the starting point of what is now called "converse theorems" in Langlands theory.
b) He suggests that, not only every elliptic curve over **Q** should be modular, but its "level" (in the modular sense) should coincide with its "conductor" (defined in terms of the local Néron models, say).
Part b) was a beautiful new idea ; it was

[1] Pages 46–52.

not in Taniyama, nor in Shimura (as Shimura himself wrote to me after Weil's paper had appeared). 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 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.

I was not the only one to be convinced : people such as Birch, Tate, Swinnerton-Dyer, Mazur, ... felt the same way ; moreover, a lot of numerical evidence was soon collected by Swinnerton-Dyer and others. Of course, there were several loose ends which needed tying up, but this was done within a few years :

– the Galois-representation definition of the conductor, and of the gamma factors of the functional equation (Ogg, Tate, myself) ;

– the newform theory of Atkin-Lehner (1970) ;

– the fact that an elliptic curve is determined, up to isogeny, by its ℓ-adic representation (for any given ℓ). This was harder. I did it in my McGill lectures (1967) when the j-invariant is not an algebraic integer (a case which turned out later to be sufficient for Wiles), but I could not do it in general ; it had to wait until 1983, when Faltings proved the general Tate conjecture.

This period (end of the '60 s and beginning of the '70 s) was a very exciting one for people working on modular forms, elliptic curves and the like. To wit :

– Langlands's theory (especially his 1967 Yale notes), with its relations with motives ;

– Deligne's construction (1968) of the ℓ-adic representations associated with modular forms of weight $\geqslant 2$, confirming a conjecture I had made the year before ;

– the theory of modular forms mod p (Swinnerton-Dyer, 1970), which I applied to define p-adic modular forms and to construct the p-adic zeta function of an arbitrary totally real number field (Antwerp, 1972) ;

– Shimura's correspondence between modular forms of half-integral weight and those of integral weight (1972) : a surprising, and beautiful, application of Weil's "converse theorem" ;

– the crowning part (1973) : Deligne's proof of Weil's conjecture for varieties over finite fields, and, as a consequence, the proof of the Ramanujan-Petersson conjecture.

Quite a list, don't you think ?

Note that, during the ten years following Weil's paper, the modularity conjecture was called "Weil's conjecture", and Taniyama's original insight was all but forgotten. Around 1976, 1 bought a copy of Taniyama's Collected Papers, and I noticed that "problem n° 12" was included there in Japanese, but not in English. To make it more widely available, I reproduced its 1955 English version in my 1977 paper on l-adic representations ; a fitting place, since the notion of system of l-adic representations is due to Taniyama ! From then on, I started saying "Taniyama-Weil conjecture" instead of "Weil conjecture" : it seemed natural to me that the credit be divided between the two of them. Little did I know that I was thus starting a bitter controversy. In the '90 s, Lang took the matter to heart (as he often does) and launched a big campaign, in order to have Weil's name removed and the conjecture called "Taniyama-Shimura", which I find strange in view of Shimura's record (or absence of record, see above). I still feel that "Taniyama-Weil" is more accurate. Maybe your suggestion of "modularity conjecture" is even better ? Anyway, one should not take such terminology quarrels too seriously. As Weil was fond to say, "Pell's equation" is not due to Pell, and Klein did not do much on the "Kleinian functions" of Poincaré ...

Best wishes
J.-P. Serre

PS : Lang's paper in the 1995 Notices describes a would-be discussion between Shimura and myself, at the Institute, in 1962-64 (sic). You ask whether this discussion actually happened. The answer will surprise you : I don't know! I have no memory of it. However, it is perfectly possible that Shimura said once "... don't you believe that every elliptic curve is modular ?" and that I replied something like "... why should it be so ?". I know very well that memory erases what is not important. If he had given me even one little piece of evidence, I would have been impressed and I would not have forgotten. (The discussion with Weil, on the other hand, was memorable ; the evidence was there.)

PPS : You would probably be interested by the letters I exchanged with Tate, Swinnerton-Dyer, Shimura,. . . , between 1966 and 1968, on the modularity conjecture. No controversy then : just mathematics.

[Extrait de "Wolf Prize in Mathematics", vol. 2, World Sci. Publ. Co., Singapore, 2001, pages 537–539]