THE GENESIS AND GESTATION OF FUNCTORIALITY

ROBERT P. LANGLANDS

Algebraic Numbers. Early Greeks. Quadratic irrationals, their existence, their nature: $\sqrt{2}, \sqrt{3}, \sqrt{5}, \dots$ appear in Theaetetus (one of Plato's Socratic dialogues)—before Plato simple surds, after Plato more elaborate theory. Found in the later books of Euclid's elements, the 10th and 12th. Combinations of simple surds, such as $a\sqrt{2} + b\sqrt{3}$, a and b rational. Proof of linear independence, thus first signs of Galois theory. Also the appearance of such irrationalities as lengths in regular polygons and regular polyhedra.

Renaissance Period (14th–15th centuries). Algebraic mysteries of cubic and quartic irrationalities—solution by extraction of roots.

18th Century. Lagrange and Vandermonde. Equations of general degree, symmetric functions of the roots, functions of the roots perhaps leading to solution by extraction of roots. Thus a kind of inchoate Galois theory.

Gauss. Elementary construction of regular heptadecagon—17 sides. Really an analysis of the resolution of the cyclotomic equation

$$X^n - 1 = 0$$

or

$$X^{n-1} + X^{n-2} + \cdots + 1 = 0.$$

Most important case is n a prime. Solve by radicals—rather by extraction of roots of low degree. It is not always clear what the earlier authors had in mind. It is however absolutely clear what Gauss is doing: building a sequence of functions of the roots $\zeta, \zeta^2, \dots, \zeta^{n-1}$ that give numbers that can be obtained by a successive extraction of roots of the lowest possible degree. For $n = 5$,

$$\zeta, \zeta^2, \zeta^3, \zeta^4 \longrightarrow \quad \zeta + \zeta^4, \zeta^2 + \zeta^3,$$

which are quadratic surds, and the equations

$$\zeta + \zeta^2 + \zeta^3 + \zeta^4 = -1, \quad (\zeta + \zeta^4)^2 = \zeta^2 + 2 + \zeta^3 = 1 - \zeta - \zeta^4$$

imply that $z = \zeta + \zeta^4$ satisfies $z^2 + z - 1 = 0$, $z = -\frac{1}{2} \pm \frac{\sqrt{5}}{2}$. As a consequence $z = -\frac{1}{2} + \frac{\sqrt{5}}{2}$. The number ζ itself is then found by the extraction of a further square root.

$$\zeta^2 - z\zeta + 1 = 0.$$

General theory is that of periods. First step is always extraction of a square. If a is a primitive root modulo n, then

$$\zeta + \zeta^{a^2} + \zeta^{a^4} + \cdots + \zeta^{a^{n-3}} \quad \text{and} \quad \zeta^a + \zeta^{a^3} + \cdots + \zeta^{a^{n-2}}$$

are the two roots of a quadratic equation. This is all found in Gauss's Treatise *Disquisitiones.*

TIFR, Mumbai, February 2005.

18th century again. Legendre's introduction of the law of quadratic reciprocity as a conjecture that he was unable to prove.

p quad. res. / not quad. res. modulo $q \leftrightarrow q$ quad. res. / not quad. res. modulo p.

Gauss again. He proved the conjecture in Disq. by an elementary method, but the problem is implicit in the theory of the cyclotomic equation. I use modern concepts for brevity. The equation $x^2 - p$ has a root modulo q say $a^2 \equiv p \pmod q$. Then q factors in the field of $\sqrt{p}$

$$(a - \sqrt{p})(a + \sqrt{p}) = a^2 - p$$

divisible by q. But we have seen that quadratic fields are contained in the cyclotomic fields $\mathbf{Q}(\zeta)$. When does q factor in $\mathbf{Q}(\zeta)$? If

$$X^n - 1 = 0$$

has a root modulo q. Take $n = p$. It has a root if $p|(q-1)$ because of Fermat's theorem $a^{q-1} \equiv 1 \pmod p$ if $(q,p) = 1$. Thus factorization depends on q modulo p. $a^p - 1 \equiv 0 \pmod q$ implies

$$q|(a - \zeta)(a - \zeta^2) \cdots (a - \zeta^{p-1})$$

Quadratic Forms. In the same complex of ideas it is natural to treat the theory of representations by quadratic forms initiated by Fermat, Euler, Legendre. The representability of q by a quadratic form with given discriminant is related to the equation

$$g = \mathfrak{q}\bar{\mathfrak{q}}$$

in the field $\mathbf{Q}(\sqrt{p})$.

Conclusion. Galois theory, reciprocity laws, quadratic forms, thus orthogonal groups belong to same complex of ideas and these ideas have given concrete expression in Gauss's magnificent treatise which has been translated into several languages and is of immediate access even to those with limited mathematical experience.

Gauss's successors in nineteenth and early twentieth century.

- Kummer (1810–1893)
- Galois (1811–1832).

Kummer. Arithmetic of cyclotomic fields, of extensions of these fields. Field is $\mathbf{Q}(\zeta)$, ζ satisfies $\zeta^\ell = 1$, ℓ a prime. Extensions of $\mathbf{Q}(\zeta)$ obtained by adjoining a root of $Y^\ell = \alpha$. Kummer analyzed these fields in the spirit described above and, again in the spirit described above, proved reciprocity laws for higher powers.

Kummer's work was followed by the general theory of number fields, by Dirichlet, by Dedekind and finally by the theory of abelian fields, Kronecker, Weber, Hilbert, Takagi, Artin. This theory, which establishes a decisive connection between abelian extensions of number fields on one hand and their ideal class groups on the other hand, is known as *class field theory* and should be regarded as arising by the evolution I have described from the simpler, but nevertheless deep, theory to be found in Euclid.

Class-field theory classifies and constructs all abelian extensions of number fields. So the problem arose after the completion of the theory to classify and construct (in some sense) <u>all</u> finite extensions. In 1956 at the bicentennial conference of Princeton University, Artin suggested that perhaps all we could know and all we needed to know in general was implicit in our knowledge of abelian extensions, so that there was in fact little left to do, although it was not clear what it might be.

Another development with origins in Gauss and therefore Legendre.

Quadratic forms $\to$ arithmetic theory of algebraic groups

Gauss $\to$ Eisenstein $\to$ Dirichlet $\to$ Hermite, H. J. S Smith, Minkowski and later analytic theory of Hardy-Littlewood.

All of this subsumed and developed by Siegel in work extending over a life-time. To him we owe I believe more than to any other mathematician the present overwhelming importance of algebraic groups in number theory—but of course not to him alone. There were other analytic developments:

Euler products and analytic continuation of L-functions. Riemann (perhaps along with rather than after Gauss the second major mathematician-philosopher of the nineteenth century). The analytic theory of L-functions was developed by Hecke and then Maaß. There was also of course an important although, in some respects, somewhat isolated development: Mordell's proof of part of Ramanujan's conjecture on the τ-function. I briefly recall the contributions of Hecke and Maaß that are pertinent here.

Holomorphic modular forms. Consider a function f holomorphic in upper half-plane

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z), \qquad \text{Weight} = k$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \pmod N$$

$$f(z) = \sum_{n=0}^{\infty} a(n) e^{2\pi i n z}$$

Cusp forms.

$$f(z) = \sum_{n=1}^{\infty} a(n) e^{2\pi i n z}$$

$$f \to \zeta_f(s) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s}$$

There is an analytic continuation and functional equation:

$$\zeta_f(k-s) = \gamma(s)\zeta_f(s).$$

$\gamma(s)$ is expressible in terms of Γ-functions.

Also if f is an eigenfunction of Hecke algebra then $\zeta_f(s)$ has an Euler product,

$$\zeta_f(s) \sim \prod_p \frac{1}{\left(1 - \frac{\alpha_p}{p^s}\right)\left(1 - \frac{\beta_p}{p^s}\right)}$$

$$\alpha_p \beta_p = p^{k-1}$$

Nonholomorphic Forms. Now f is infinitely differentiable in upper half-plane, say $f = f(x,y)$. It is an eigenfunction of the Laplacian

$$y^2\left\{\frac{d^2 f}{dx^2} + \frac{d^2 f}{dy^2}\right\} = \lambda f$$

There is a similar theory but with different Γ-factors. This is a contribution of Maaß.

Note. Ramanujan τ-function is defined by a holomorphic modular form of weight 12 and level 1.

$$\begin{aligned} g(x) &= x\big\{(1+x)(1-x^2)\cdots\big\}^{24} \\ &= \sum_{n=1}^{\infty} \tau(n)x^n \end{aligned}$$

Ramanujan conjectured that if n and n' are relatively prime then $\tau(nn') = \tau(n)\tau(n')$. This is a consequence of Hecke's formula but was in fact proved by Mordell some years before Hecke's work appeared and by methods that anticipated those of Hecke.

This was roughly the state in the late 50s of the subject I am describing, except that I have omitted the contributions of Selberg. I shall return to them but only after recalling my own circumstances at the time.

I took a BA from the University of British Columbia in 1957, an MA in 1958, and then moved on to Yale, where I had at first planned to work on partial differential equations. I was also captivated by the book of Hille and Phillips on analytic semigroups, so that without paying much attention to the matter, I had a thesis largely in parabolic equations and analytic semigroups before the first year was up. So my second year as a student at Yale, the last, was entirely free.

Indeed my first year was also pretty free and I spent most of my time in the library or with a few books of my own. I describe some I remember studying.

a) **Burnside on finite groups**—although I remember idly dreaming of solving the problem on simple groups of odd order solved indeed a short time thereafter by Feit-Thompson, I don't think I ever mastered much of Burnside.
b) The first edition of **Zygmund** on Fourier series—I knew the book inside and out at the time.
c) **Stone's** work on the spectral decomposition of self-adjoint operators in Hilbert space. Although I never understood this material really well, it did later stand me in good stead.
d) I did come also, for some reason or other, to understand not only the basis of function theory—from **Knopp's** books—but also something about holomorphic functions of several complex variables that I was very quickly able to exploit.

S. Gal and A. Selberg. Fortunately for me something else happened during my last year at Yale. S. Gal, a Hungarian, had fled Hungary after the failed revolution of 1956, and sponsored by Selberg, whose wife was Hungarian, had spent a year at the IAS before coming to Yale. He was fascinated by the ideas developed by Selberg in the first of his TIFR papers.

As is well known, A. Selberg published two papers in the proceedings of two different TIFR conferences. The first, strongly influenced by Maaß was on the **spectral problems** arising from Maaß's construction. The second, strongly influenced by Siegel, was on the rigidity of discrete subgroups of Lie groups of large rank. They were both highly original and very influential papers. It was the first that Gal was trying to understand and he began with Hecke's ideas!

While attending Gal's lecture I also looked at Selberg's papers on my own and combining what was there with some things about holomorphic functions of several variables was able to prove some theorems about the analytic construction of what are now called Eisenstein series. I confess that I was rather more interested in learning new things than in proving them and did not attach much significance to these results.

In 1960 because E. Nelson liked my work on analytic semigroups, perhaps partly because it was related to some work he himself has done on analytic vectors, I was appointed an instructor at Princeton University.

Princeton. At Princeton several things quickly happened that would be important for my mathematical development but exactly when and in what order I can no longer say. I list them.

a) Invited to speak in a seminar and having nothing better to report, I discussed the results on some simple Eisenstein series. **Bochner** who was present and who was a great fan of Dirichlet series was very excited, and from then on did everything possible to promote my career.

b) I had, during a general conversation at Yale, heard one of the professors speak of class field theory as a subject too arcane to be of any interest to mathematicians at large. My curiosity was piqued. Fortunately for me, although Artin had left Princeton to return to Hamburg, A. Brumer and M. Rosen, who had come to Princeton to study with him, decided to have a seminar on class field theory. When I saw this announced I immediately decided to take part. In addition to Brumer, Rosen, and me there was only one other participant, perhaps J. d'Arti. Brumer, who did most of the lecturing, and Rosen were experienced. I was not. I had certainly studied Northcott's book on ideal theory before leaving Vancouver, perhaps even persuaded myself that I understood it. Perhaps I had also, as I believe, already gone through Weyl's book on algebraic number theory, getting something out of it, but probably not a great deal. Anyhow, I plagued poor Brumer with silly questions often taxing his patience beyond its limits. Rosen was more tolerant of my impertinence.

c) Bochner pointed out my existence to Selberg and he invited me over to speak with him at the Institute. I have known Selberg for more than 40 years. We are on cordial terms and our offices have been essentially adjacent for more than 20 years. This is nevertheless the only mathematical conversation I ever had with him. It was a revelation. I had never talked in detail about mathematics with Bochner. So Selberg was the first powerful analyst I had seen up close.

   I recall that one of the results announced in the first Tata paper was a general proof of the analytical continuation of the Eisenstein series for discrete subgroups of $\mathrm{SL}(2, \mathbf{R})$ with quotient of finite volume. Maaß and Roelcke had searched in vain for such a proof. The ideas involved are just those of spectral theory for second-order self-adjoint equations on a half-line, but I had never really seen these before and certainly not in the hands of a master. It was a defining experience. I went away with a reprint of his paper and began to study it carefully, especially the trace formula.

d) Sometime in those first months or years in Princeton, I acquired and began to read the various Paris seminars of Cartan, Godement and others that had been inspired by the work of Siegel, Hecke, Selberg and by the work of the pioneers of representation theory, Gelfand and coauthors, Bargmann and Harish-Chandra.

e) Bochner urged me to give a graduate course in class field theory. Although I still knew almost nothing about the subject, had only two weeks to prepare, was very young, and scared stiff, I had no choice but to yield. I owe eternal gratitude to the two students who stayed to the end, D. Reich and R. Fuller. I don't suppose that any of us really understood the subject, even at the end, but it was enormously useful to me to go through the motions.

f) D. Lowdenslager, who died prematurely not long afterwards, observed to me when I told him that I was trying to use the trace formula of Selberg to compute the dimension of the space of holomorphic automorphic forms for higher-dimensional groups that it was generally felt that the work of Harish-Chandra was pertinent. So I began to study Harish-Chandra's work and, at least at that time, it was not possible to study one of his papers without studying all. Anyhow, on studying his papers I came to understand that the integrals from the trace formula were given by characters of the discrete series and could calculate the dimension.

g) I read Gelfand's address to the ICM in Stockholm, finally understood correctly the notion of a cusp form in general. Since, as I observed, I had some passive experience with the spectral theory of self-adjoint operators and with holomorphic forms of several variables, several months with my nose to the grindstone and a refusal to be discouraged by temporary setbacks—for the proof presented a good number of unexpected obstacles—gave me in the spring of 1964 a complete proof of analytic continuation. I was exhausted and, moreover, quite dissatisfied with the account of the proof but with no energy and no desire to revise the exposition. If **Harish-Chandra** had not taken time from his own researches to work through and present, at the Institute for Advanced Study, at least a part of my paper—that pertaining to Eisenstein series associated to cusp forms—no one may have taken me seriously. To Bochner and Harish-Chandra I owe an enormous amount.

Berkeley. In the fall of 1964, I went to Berkeley for a year. Although, as I appreciate in retrospect, it was not an entirely unsuccessful year—there were general results on the volumes of fundamental domains and a conjecture on the geometrical realization of discrete series that bore some fruit later—I did not have the feeling that things were working out. The Eisenstein series in hand, I tried to develop the general trace formula, but did not succeed. I ran a seminar together with P. Griffiths on abelian varieties, but in the end he did much more with the material than I.

I grew discouraged and 1965–66 when I was back in Princeton was at first even worse. I had several projects—more or less vague—with which I was trying to do something.

(i) Trace formula
(ii) Extension of the Hecke theory to groups other than GL(2).
(iii) A nonabelian class field theory.

(ii) and (iii) are of course related, yet different. All three projects were coming to nothing. I began to think of throwing it all up.

There were in the 1960s still residues of the romantic notions of British imperialism, embodied in memories of figures like Gertrude Bell or T. E Lawrence. So one could dream, even with a wife and four small children, of escaping into the life and language of some exotic land and beginning anew. I did; my wife, more generous than wise, did not discourage me. We made plans to spend a year, at first several years, but the department in Ankara only agreed to a provisional one-year appointment. The specific choice of Turkey was the result of accidental factors.

The decision had been taken by the summer of 1966 and all ambitious projects dropped. I took up again the study of Russian abandoned for many years, and continued I suppose to teach mathematics, perhaps even to learn it, but with no urgency.

As far as I recall, I began idly, simply to fill the time, to calculate the constant term of the Eisenstein series associated to maximal parabolics of split groups. I had no goal in mind, just nothing better to do.

Then I noticed that the constant terms were Euler products and that they had a uniform expression in terms of representations of the group defined by the dual Cartan matrix. That was probably an insight that came slowly. It was certainly there at the time of the Yale lectures. Indeed, before posting my letter to Weil on Casselman's UBC site, I verified some dates with the help of external evidence. The Yale lectures were given in April 1967, and the letter to Weil with the functoriality conjectures in their original form was sent in January 1967. The idea itself must have come during the vacation period of Christmas/New Year 1966–67. So I was apparently quite reticent during the Yale lectures.

I end by recalling the main idea of functionality. If I replace s by $s+k-1$ in the Hecke form then

$$\alpha_p \to \widetilde{\alpha}_p = \frac{\alpha_p}{p^{k-1}}, \quad \beta_p \to \widetilde{\beta}_p = \frac{\beta_p}{p^{k-1}}, \quad \widetilde{\alpha}_p\widetilde{\beta}_p = 1,$$

and the functional equation is between s and $1-s$. The matrix

$$\begin{pmatrix} \widetilde{\alpha}_p & \\ & \widetilde{\beta}_p \end{pmatrix} = \gamma_p$$

may be treated as a conjugacy class in $\mathrm{GL}(2,\mathbf{C})$ and the Hecke form may be written as

$$\prod_p \frac{1}{\det\Big(1-\frac{\gamma_p}{p^s}\Big)}$$

More generally, although Hecke did not do so, we may replace γ_p by $\rho(\gamma_p)$ where ρ is any finite-dimensional representation of $\mathrm{GL}(2,\mathbf{C})$

$$L(s,\pi,\rho) = \prod \frac{1}{\det\Big(1-\frac{\rho(\gamma_p)}{p^s}\Big)}$$

Here $\gamma_p = \gamma_p(f) = \gamma_p(\pi)$ where the automorphic form, an eigenform of the Hecke operators, is replaced in the notation by the representation it determines. For any reductive group G over F and any automorphic representation π of $G(\mathbf{A}_F)$, the theory of spherical functions allows us to define, for almost all p, $\gamma_p(\pi) = \gamma(\pi_p)$ as a conjugacy class in a finite-dimensional complex group associated to G, its L-group. The calculations on Eisenstein series I described give Euler products can then all be expressed as

$$\prod \frac{1}{\det\left(1 - \frac{\rho(\gamma_p)}{p^s}\right)}$$

They suggest that this function can be analytically continued and has a functional equation of the usual kind. This question posed, a way to answer it suggests itself. Tamagawa had on some occasion that could not have been too long before December 1966, but I am not sure, delivered a lecture in the auditorium of the old Fine Hall in which he discussed the standard L-function associated to automorphic representation on forms on the multiplicative group of a division algebra. I had no trouble believing that his method would also work for $\mathrm{GL}(n)$, as indeed it does, as later shown by Godement-Jacquet. Then, in analogy with the Artin reciprocity law, all we would need to do to show the analytic continuation of $L(s,\pi,p)$ is to establish the existence of an automorphic representation of $\mathrm{GL}(n)$, $n = \dim\rho$, such that

$$\left\{\rho\left(\gamma_p(\Pi)\right)\right\} = \left\{\gamma_p(\Pi)\right\}$$

for almost all p. It is a small step—at least conceptually—from this possibility to the possibility of functoriality in general.

Since the meeting with Weil that gave rise to the letter took place early in January and, according to my recollection, took place shortly after I began to appreciate these possibilities, the idea probably came to me, as I said, during the Christmas vacation.

Although the date the idea came is forgotten, I still have a vivid recollection of the place. In the old Fine Hall at Princeton University, there was a small seminar room on the ground floor directly to the east of the entrance. The building itself, I recall, was in Gothic style with leaded casement windows. I was looking through them into the ivy and the pines and across to the fence surrounding the gardens of the residence of the president when I realized that the conjecture I was in the course of formulating implied on taking $G = \{1\}$ the Artin conjecture. It was certainly one of the major moments in my mathematical career.

Compiled on April 18, 2025.