Pure and Applied Mathematics Quarterly
Volume 19, Number 1, 3–6, 2023

Working with Don

Benedict H. Gross

I wrote six joint papers with Don Zagier. Two are reports of some computation, and one is a Comptes Rendus note. The other three papers represent my best mathematical work.

I first heard about Don when I was in graduate school and read his paper with Hirzebruch on the intersection of curves on a Hilbert modular surface. When Ken Ribet pointed Don out to me, at a conference in Corvallis in the summer of 1977, I was incredulous. Here was a senior mathematician, planning to come to Harvard that fall as a visiting professor, who was a year younger than I was. Life is so unfair! I didn't get to meet Don at the conference – he spent most his time computing special values of the L-functions associated to symmetric powers of the modular form Δ, giving Deligne convincing evidence for his general conjectures. But we became friends when he arrived at Harvard that fall.

I was in the process of putting together my thesis, on elliptic curves E with complex multiplication and their associated Hecke characters χ. In this case, the special values of the L-function of $\mathrm{Sym}^n(E)$ satisfy Deligne's conjecture, by Damerell's theorem for the L-functions of powers of χ. We decided to look at the rational numbers that came up, especially for the L-functions of the odd powers χ^{2n-1} at the point $s = n$ in the center of the critical strip. I had arranged a visit to IBM's research center at Yorktown that fall, and had ample computational time to calculate these rational numbers. The denominator contained only small primes, which Don immediately recognized as $(n-1)!$ (he mentioned that it was a good idea to memorize the factorizations of standard quantities, like factorials and Bernoulli numbers). After we removed this term, the numerator was always twice a square! That seemed worth publishing in a short note, as it provided evidence for higher Tate-Shafarevich groups.

Don arranged for me to visit Maryland for their special year in number theory in the spring of 1978, and came up to Boston for my thesis defense in May. My advisor John Tate asked me to describe a conjecture I had made with Deligne, that the order of vanishing of the L-functions of motives of odd weight in the center of their critical strip remained constant under conjugation of the coefficients of the Dirichlet series. I asked John whether he wanted me to

Received November 12, 2022.

explain this for abelian varieties with endomorphisms, or for general motives. At which point Don interjected “Why don’t you explain it for elliptic curves with complex multiplication, so that you know what you are talking about?” I was suitably humbled. When Don returned to Bonn I purchased his car – an Audi on its last legs. It cost $75 and lasted for four months.

After the special year in number theory, Maryland appointed Don as a chaired Professor, and he visited for a semester each year. That was lucky for me – I came down to visit for a week in September of 1982. I had been corresponding with Bryan Birch about Heegner points on the modular curve $X_0(N)$, and had formulated a general conjecture on their heights, related to the first derivatives of Rankin L-functions in the center of the critical strip. Don knew more about the special values of Rankin L-functions than anyone else, so I challenged him to compute the first derivative.

Even to an amateur like me, it was clear that this computation was daunting – we had to take the product of an Eisenstein series of weight one and level ND with the theta function of a binary quadratic form of discriminant D, then trace this form of weight 2 from level ND to level N, then differentiate the result with respect to the parameter s in the Eisenstein series and set $s = 1$, then calculate the holomorphic projection of the corresponding non-holomorphic derivative. The derivative of the Rankin L-function is then the inner product of this projection with a cusp form of weight 2. Even Don was reluctant to begin. When I pointed out that a proof of my conjecture would have implications for the conjecture of Birch and Swinnerton-Dyer, he was unmoved. The implications for the class numbers of imaginary quadratic fields certainly interested him, but he continued to demur. As a last resort, I turned to Steve Kudla, who was attending our discussion, and remarked “Don is just afraid that he can’t compute it.” That did the trick.

The week that followed was spent at Don’s apartment, with sporadic breaks for food and sleep. He began the computation by saying that we would work at level $N = 1$ and weight $2k \geq 4$. Since my conjecture was only for weight $2k = 2$, and even in that weight there were no cusp forms of level $N = 1$, I tried to object. But Don insisted, saying that we could keep k as a variable in the computation and try to set $k = 1$ at the very end. He pulled out a yellow note pad, drew a vertical line down the center, and began to work. Suffice it to say that I have never seen so much mathematics pass by so quickly. Don kept the computation on the right side of the page, running compatibility tests on the left to detect the few errors that crept in. All the fine points of analysis, such as the convergence of infinite series and integrals, were ignored in the quest for some kind of answer. Five days into this madness, Don produced a formula for the Fourier coefficients of the holomorphic projection, each containing at least 20 terms, most of which were wildly divergent when we dared to set $k = 1$. In his formula for the first Fourier coefficient there was one term that was a finite sum:

$$\sum_{n=1}^{|D|} r_A(|D| - n) \sum_{d|n} \epsilon_A(d, n) \log(n/d^2).$$

Here A is the class of the binary form of discriminant D, $r_A(m)$ is the number of ideals in that class of norm m, and $\epsilon_A(d, n)$ is a sign. This expression can be rewritten as a sum over primes p

$$\sum_{p \leq |D|} m_A(p) \log(p).$$

Don pointed to the sum and asked me, in an exasperated manner, what the integers $m_A(p)$ were supposed to be. I guessed that when the class A is non-trivial, $m_A(p)$ should be the power of p dividing the norm of the difference $j(\tau) - j(\tau_A)$, where $j(z)$ is the modular invariant of level 1, τ is a point of discriminant D in the upper half-plane, and τ_A is another point of discriminant D, translated from the first by the class A. When A is trivial, $j(\tau) - j(\tau_A) = 0$, so Don differentiated to get

$$q\frac{d}{dq} j(z)/\eta(z)^4 = j(z)^{2/3}(j(z) - 1728)^{1/2}.$$

In this case, $m_A(p)$ should be the power of p dividing the norm of $j(\tau)^{2/3}(j(\tau) - 1728)^{1/2}$. In particular, this conjecture implies that all the primes dividing this norm are less than or equal to $|D|$.

This all seemed far-fetched, as nothing in Don's computation suggested any relation with singular moduli $j(\tau)$. But at least it was something that we could check! The algebraic integers $j(\tau)$ had been computed for discriminants with small class numbers, and were tabulated in the literature. Since by then it was 3 o'clock in the morning, I suggested that we go to sleep and look up these values in the math library the next morning. (There were no tables on the Internet in 1982.) If my guess turned out to be hopelessly off the mark, we could bring this nonsense to a stop. After these somewhat dispiriting remarks, I went to bed.

I should have stayed up. Although Don was completely exhausted and had only a small hand-held computer, he set about calculating as many singular moduli as he could, and in each case checking my guess. I woke up to find his living room covered with paper. It took me a while to figure out what was going on. At the top of each page was a discriminant. This was followed by a calculation of its singular moduli. At the bottom of each page was a check mark indicating that the powers of p in their factorization matched Don's formula. The last page was in large print, and said “Wake me up immediately!” Which I did.

That morning we were both in a state of mathematical euphoria. We didn't know whether Don's calculation could be made rigorous in weight $2k = 2$, or whether it would work in level $N > 1$. Also, I had no idea how to compute the heights of Heegner points on the modular curve $X_0(N)$, which would somehow generalize the prime factorization of singular moduli. But we knew that we were onto something important.

That night we went to a faculty dinner and sat with Steve Kudla. Don excitedly recounted his computation and we asked Steve if he had ever seen sums of the form $\sum_{a+b=n} R_A(a)R_B(b)$ which involved the product of representation numbers of two binary quadratic forms of the same discriminant. Steve remarked that these were representation numbers of the sum $A+B$ of the two forms, which is a quadratic form in four variables with square discriminant – the norm form on a quaternion algebra. I knew very little about quaternion algebras at the time, but I did know that the endomorphism ring of a supersingular elliptic curve is a maximal order in the rational quaternion algebra ramified at ∞ and p. Writing its norm form as the sum of two binary forms of the same discriminant is equivalent to embedding a quadratic order into the quaternion order. You get such an embedding (of endomorphism rings) when you have an elliptic curve with complex multiplication in characteristic zero which has supersingular reduction modulo p. And that turned out to be the right framework for the height computation.

I never had a week like that again in my life.

Benedict H. Gross
Department of Mathematics
Harvard University
Cambridge, MA 02138
USA
E-mail: gross@math.harvard.edu