p-adic differential equations
18.787, Kiran S. Kedlaya, MIT, fall 2007
Introduction and motivation

In this unit, we list the various ways that p-adic differential equations occur in modern number theory. We then explain in a bit of detail one example; it is in fact of the original examples of Dwork that prompted the creation of the whole theory.

Since this unit is intended as an overview, it involves sweeping statements with proofs not included. Starting with the next unit, things will be done more carefully.

To keep things better organized, I am creating a central references file "references.pdf" on the web site. However, I will attach to each unit some afternotes, and some exercises; the latter are optional unless you need a genuine course grade (e.g., if you are an undergraduate).

1 Where are the p-adic differential equations?

Under what circumstances does the consideration of a differential equation involving p-adic numbers, rather than real or complex numbers, occur in number theory? Let me count the ways; note that I may not get to discuss all of these in much detail during the semester, though I'll try to fit some of them in at the end.

- The original circumstance, which will be described later in this lecture, was Dwork's work on the variation of zeta functions of algebraic varieties over finite fields. Roughly speaking, solving certain p-adic differential equations can give rise to explicit formulas for number of points on varieties over finite fields.

  In contrast to methods involving étale cohomology, methods for studying zeta functions based on p-adic analysis (including also the next item) lend themselves well to numerical computation. Interest in computing zeta functions for varieties where straightforward point-counting is not an option (e.g., curves over tremendously large prime fields) has been driven by potential applications in computer science, the principal example being cryptography based on elliptic or hyperelliptic curves. (There may also be some interesting applications generated by coding theory, but this remains to be seen.)

- Dwork's work suggested, but did not immediately lead to, a proper analogue of étale cohomology based on p-adic analytic techniques. Such an analogue was eventually developed by Berthelot (based on work of Monsky and Washnitzer, and also ideas of Grothendieck); it is called *rigid cohomology* (see the unit notes for the origin of the word "rigid"). It is not yet a fully functional analogue of étale cohomology, particularly because there are still open problems related to the construction of a good category of coefficients. These problems are rather closely related to questions concerning p-adic differential equations, and in fact some of the results presented in this course have been (or will be) used for this purpose.

- The subject of *p-adic Hodge theory* aims to do for the cohomology of varieties over p-adic fields what ordinary Hodge theory does for the cohomology of varieties over $\mathbb{C}$, namely abstract away the variety and enable a better understanding of the cohomology of the variety as an object in its own right. In the p-adic case, the cohomology in question is often étale cohomology, which carries the structure of a Galois representation.

  The study of such representations, as pioneered by Fontaine, involves a number of exotic auxiliary rings ("rings of p-adic periods") which serve their intended purposes but are otherwise a bit mysterious. More recently, the work of Berger has connected much of the theory to the study of p-adic differential equations; notably, a key result that was originally intended for use in p-adic cohomology (the "p-adic local monodromy theorem") turned out to imply an important conjecture about Galois representations (Fontaine's conjecture on potential semistability).

- There are some interesting analogies between properties of differential equations over $\mathbb{C}$ with meromorphic singularities, and wildly ramified Galois representations of p-adic fields. At some level, this is suggested by the parallel formulation of the Langlands conjectures in the number field and function field cases. One can use p-adic differential equations to interpolate between the two situations, by associating differential equations to Galois representations (as in the previous item) and then using "differential invariants" (irregularity) to recover "Galois invariants" (Artin and Swan conductor).

  For representations of the étale fundamental group of a variety over a field of positive characteristic of dimension greater than 1, it is quite a tough problem to construct meaningful numerical invariants from the Galois point of view. Recent work of Abbes and Saito attempts to do this, but the resulting quantities are quite difficult to calculate. One can alternatively use p-adic differential equations to define invariants which are somewhat easier to deal with for some purposes; for instance, one can define a "differential Swan conductor" which is guaranteed to be an integer, whereas one does not know this for the Abbes-Saito conductor. (In fact, the two are expected to coincide; see Liang Xiao's thesis project.)

2 Zeta functions of varieties

Let us now say something more detailed about the first item in the previous list. For this, I must recall properties of zeta functions of algebraic varieties.

For λ in some field K, let E_λ be the elliptic curve over K defined by the equation

$$E_\lambda : y^2 = x(x-1)(x-\lambda)$$

in the projective plane. Remember that there is one point $O = [0:1:0]$ at infinity, and that there is a natural group law on $E_\lambda(K)$ under which O is the origin, and three points add to zero if and only if they are collinear (or better, if they are the three intersections of E_λ with some line; this correctly allows for degenerate cases).

Theorem 1 (Hasse)**.** *Suppose λ belongs to a finite field $\mathbb{F}_q$. Then $\#E_\lambda(\mathbb{F}_q) = q + 1 - a_q(\lambda)$ where $|a_q(\lambda)| \leq 2\sqrt{q}$.*

Proof. See [Sil91, Theorem V.1.1]. □

Hasse's theorem was later vastly generalized as follows, originally as a set of conjectures by Weil. (Despite no longer being conjectural, these are still commonly referred to as the *Weil conjectures*.) For X an algebraic variety over $\mathbb{F}_q$, the *zeta function* of X is defined as the formal power series

$$\zeta_X(T) = \exp\left(\sum_{n=1}^{\infty} \frac{T^n}{n} \#X(\mathbb{F}_{q^n})\right);$$

another way to write it, which makes it look more like zeta functions you've seen before, is

$$\zeta_X(T) = \prod_x (1 - T^{\deg(x)})^{-1},$$

where x runs over Galois orbits of $X(\overline{\mathbb{F}_q})$, and deg is the size of the orbit. (If you prefer algebro-geometric terminology: x runs over closed points of X, and deg is the degree over $\mathbb{F}_q$.) For $X = E_\lambda$, one checks (exercise) that

$$\zeta_X(T) = \frac{1 - a_q(\lambda)T + qT^2}{(1-T)(1-qT)}.$$

Theorem 2 (Dwork, Grothendieck, Deligne, et al)**.** *Let X be an algebraic variety over $\mathbb{F}_q$. Then $\zeta_X(T)$ represents a rational function of T. Moreover, if X is smooth and proper of dimension d, we can write*

$$\zeta_X(T) = \frac{P_1(T) \cdots P_{2d-1}(T)}{P_0(T) \cdots P_{2d}(T)},$$

where each $P_i(T)$ has integer coefficients, satisfies $P_i(0) = 1$, and has all roots in $\mathbb{C}$ on the circle $|T| = q^{-i/2}$.

Proof. The proof of this theorem is a sufficiently massive undertaking that even a reference is not reasonable here; instead, we give [Har77, Appendix C] as a metareference. □

It is worth pointing out that the first complete proof uses the fact that you can interpret

$$\#X(\mathbb{F}_{q^n}) = \sum_i (-1)^i \operatorname{Trace}(F^n, H^i_{\text{et}}(X, \mathbb{Q}_\ell)),$$

where for any prime $\ell \neq p$, $H^i_{\text{et}}(X, \mathbb{Q}_\ell)$ is the i-th étale cohomology group of X with coefficients in $\mathbb{Q}_\ell$.

3 Why p-adic differential equations?

All well and good, but there are several downsides of the interpretation in terms of étale cohomology. One important one is that étale cohomology is not explicitly computable; for instance, it is not straightforward to describe étale cohomology to a computer well enough that the computer can make calculations. (The main problem is that while one can write down étale cocycles, it is very hard to tell whether or not a cocycle is a coboundary.)

Another important downside is that you don't get extremely good information about what happens to ζ_X when you vary X. This is where p-adic differential equations enter the picture. It was observed by Dwork that when you have a family of algebraic varieties defined over $\mathbb{Q}$, the same differential equations appear when you study variation of complex periods, and when you study variation of zeta functions over $\mathbb{F}_p$.

Here is an explicit example due to Dwork. Recall that the *hypergeometric series*

$$F(a,b;c;z) = \sum_{i=0}^{\infty} \frac{a(a+1)\cdots(a+i)b(b+1)\cdots(b+i)}{c(c+1)\cdots(c+i)i!} z^i$$

satisfies the *hypergeometric differential equation*

$$z(1-z)y'' + (c-(a+b+1)z)y' - aby = 0.$$

Set in particular

$$\alpha(z) = F(1/2, 1/2; 1; z);$$

over $\mathbb{C}$, α is related to an elliptic integral, for instance, by the formula

$$\alpha(\lambda) = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1-\lambda \sin^2 \theta}} \qquad (0 < \lambda < 1).$$

(You can extend this to complex λ by being careful about branch cuts.) This elliptic integral can be viewed as a period integral for the curve E_λ, i.e., you're integrating some meromorphic form on E_λ around some loop (homology class).

Let $p \neq 2$ be an odd prime. We now try to interpret $\alpha(z)$ as a function of a p-adic variable rather than a complex variable. Beware that this means z can take *any* value in a field with a norm extending the p-adic norm on $\mathbb{Q}$, not just $\mathbb{Q}_p$ itself. (For the moment, you can imagine z running over a completed algebraic closure of $\mathbb{Q}_p$.)

Lemma 3. *The series $\alpha(z)$ converges p-adically for $|z| < 1$.*

Proof. Straightforward. □

Dwork discovered that a closely related function admits "analytic continuation". To explain what the result says, we define the *Igusa polynomial*

$$H(z) = \sum_{i=0}^{(p-1)/2} \binom{(p-1)/2}{i}^2 z^i.$$

Modulo p, the roots of $H(z)$ are the values of $\lambda \in \overline{\mathbb{F}_p}$ (which actually all belong to $\mathbb{F}_{p^2}$, for reasons we will not discuss) for which E_λ is a supersingular elliptic curve, i.e., $a_q(\lambda) \equiv 0 \pmod{p}$.

Dwork's analytic continuation result then is the following.

Proposition 4 (Dwork)**.** *There exists a series $\xi(z) = \sum_j P_i(z)/H(z)^i$ converging uniformly for $|z| \leq 1$ and $|H(z)| = 1$, with each $P_i(z) \in \mathbb{Q}_p[z]$, such that*

$$\xi(z) = (-1)^{(p-1)/2} \frac{\alpha(z)}{\alpha(z^p)} \qquad (|z| < 1).$$

Proof. See [vdP86, §7]. □

Note that ξ itself satisfies a differential equation, which I won't write out just yet. We will see it again later.

For $\lambda \in \mathbb{F}_q$, let $\mathbb{Z}_q$ be the unramified extension of $\mathbb{Z}_p$ with residue field $\mathbb{F}_q$. Let $[\lambda]$ be the unique q-th root of 1 in $\mathbb{Z}_q$ congruent to λ mod p (the *Teichmüller lift* of λ).

Theorem 5 (Dwork)**.** *If $q = p^a$ and $\lambda \in \mathbb{F}_q$ is not a root of $H(z)$, then*

$$T^2 - a_q(\lambda)T + q = (T - u)(T - q/u),$$

where

$$u = \xi([\lambda])\xi([\lambda]^p) \cdots \xi([\lambda]^{p^{a-1}}).$$

That is, the quantity u is the "unit root" of the polynomial $T^2 - a_q(\lambda)T + q$ occurring (up to reversal) in the zeta function.

Proof. See [vdP86, §7]. □

4 A word of caution

Before we embark on the study of p-adic ordinary differential equations, a cautionary note is in order, concerning the rather innocuous-looking differential equation $y' = y$. Over $\mathbb{R}$ or $\mathbb{C}$, this equation is nonsingular everywhere, and its solutions $y = ce^x$ are defined everywhere.

Over a p-adic field, things are quite different. As a power series around $x = 0$, we have

$$y = c \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

and the denominators hurt us rather than helping. In fact, the series only converges for $|x| < p^{-1/(p-1)}$ (assuming that we are normalizing $|p| = p^{-1}$). For comparison, note that the logarithm series

$$\log \frac{1}{1-x} = \sum_{n=1}^{\infty} \frac{x^n}{n}$$

converges for $|x| < 1$.

The conclusion to be taken away is that there is no fundamental theorem of ordinary differential equations over the p-adics! In fact, the hypergeometric differential equation in the previous example was somewhat special; the fact that it had a solution in a disc where it had no singularities was not a foregone conclusion. One of Dwork's discoveries is that this typically happens for differential equations that "come from geometry", such as the Picard-Fuchs equations that arise from integrals of algebraic functions (e.g., elliptic integrals). Another is that one can quantify rather well the obstruction to solving a p-adic differential equation in a nonsingular disc, using similar techniques to those used to study obstructions to solving complex differential equations in singular discs.

5 Notes

I alluded above to the notion of an analytic function, defined as a uniform limit of rational functions with poles prescribed to certain regions. To keep down the background required for the course, I will stick throughout to this approach of defining everything in terms of rings, and not making any attempt to introduce analytic geometry over a nonarchimedean field. However, it must be noted that it is much better in the long run to build this theory in terms of nonarchimedean analytic geometry; for example, it is pretty hopeless to deal with partial differential equations without doing so.

That said, there are several ways to develop a theory of analytic spaces over a nonarchimedean field. The traditional method is Tate's theory of rigid analytic spaces, so-called because one develops everything "rigidly" by imitating the theory of schemes in algebraic geometry, but using rings of convergent power series instead of polynomials. The canonical foundational reference for rigid geometry is the book of Bosch, Güntzer, and Remmert [BGR84], but novices may find the text of Fresnel and van der Put [FvdP04] or the lecture notes of Bosch [Bos05] more approachable. A more recent method, which in some ways is more robust, is Berkovich's theory of nonarchimedean analytic spaces (commonly called *Berkovich spaces*), as introduced in [Ber90] and further developed in [Ber93]. For both points of view, see also the lecture notes of Conrad [Con07].

6 Exercises

Exercises are purely optional unless you are an undergraduate, in which case please see me at once if you have not done so yet.

1. Explain why the Dwork-Grothendieck-Deligne et al theorem implies Hasse's theorem (this includes verifying the formula for the zeta function of E_λ).

2. Check that the usual formula
$$\liminf_{n\to\infty} |a_n|^{-1/n}$$

for the radius of convergence of the Taylor series $\sum_{n=0}^{\infty} a_n x^n$ still works over a nonarchimedean field. (That is, the series converges inside that radius, and diverges outside.)

3. Check that the exponential series has radius of convergence $p^{-1/(p-1)}$.

4. Show that a power series which converges for $|x| \leq 1$ may have an integral which only converges for $|x| < 1$, but that its derivative still converges for $|x| \leq 1$. This is backwards from the archimedean situation.