CAMBRIDGE

Cambridge University Press & Assessment
978-1-009-12334-1 — p-adic Differential Equations
Kiran S. Kedlaya
Excerpt
More Information

0
Introductory remarks

The theory of ordinary differential equations is a fundamental instrument of *continuous* mathematics, in which the central objects of study are functions involving real numbers. It is not immediately apparent that this theory has anything useful to say about *discrete* mathematics in general, or number theory in particular.

In this book we consider ordinary differential equations in which the role of the real numbers is instead played by the field of p-adic numbers, for some prime number p. The p-adics form a number system with enough formal similarities to the real numbers to permit meaningful analogues of notions from calculus, such as continuity and differentiability. However, the p-adics incorporate data from arithmetic in a fundamental way; two numbers are p-adically close together if their difference is divisible by a large power of p.

In this chapter, we first survey some ways in which p-adic differential equations appear in number theory. We then focus on an example of Dwork, in which the p-adic behavior of Gauss's hypergeometric differential equation relates to the manifestly number-theoretic topic of the number of points on an elliptic curve over a finite field.

Since this chapter is meant only as an introduction, it is full of statements for which we give references instead of proofs. This practice is not typical of the rest of this book, except for the forward-looking discussions in the appendices. On a related note, the reader new to p-adic numbers should postpone this chapter's exercises until after reading Part I.

0.1 Why p-adic differential equations?

Although the very existence of a highly developed theory of p-adic ordinary differential equations is not entirely well known even within number theory, the subject is actually almost 50 years old. Here are some circumstances, past and present, in which it arises; some of these will be taken up again in the appendices.

Variation of zeta functions (see Appendix A). The original circumstance in which p-adic differential equations appeared in number theory 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 p-adic cohomology) lend themselves well to numerical computation. Interest in computing zeta functions for varieties for which direct point-counting is not an option (e.g., curves over tremendously large finite fields) has been driven by applications in computer science, the principal example being cryptography based on elliptic or hyperelliptic curves.

p-adic cohomology (see Appendix B). 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 by synthesizing work of Monsky and Washnitzer with ideas of Grothendieck); it is called *rigid cohomology* (see the chapter notes for the origin of the term "rigid"). The development of rigid cohomology has lagged somewhat behind that of étale cohomology, partly due to the emergence of some thorny problems related to the construction of a good category of coefficients. These problems, which have only recently been resolved, are rather closely related to questions concerning p-adic differential equations; in fact, some of the results presented in this book have been used to address these problems.

p-adic Hodge theory (see Appendix C). 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, to provide a better understanding of the cohomology of a variety in its own right, independently of the geometry of the variety. 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 of Fontaine on the potential semistability of Galois representations.

Ramification theory (see Chapter 19). There are some interesting analogies between properties of differential equations over $\mathbb{C}$ with meromorphic singularities and properties of 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 (such as irregularity) to recover Galois invariants (such as Artin and Swan conductors).

For representations of the étale fundamental group of a variety over a field of positive characteristic of dimension greater than 1, it is difficult to construct meaningful Galois-theoretic numerical invariants. Recent work of Abbes and Saito [1, 2] provides satisfactory definitions, but the resulting quantities are quite difficult to calculate. One can alternatively use p-adic differential equations to define invariants which can be somewhat easier to deal with; for instance, one can define a *differential Swan conductor* which is guaranteed to be an integer [238], whereas this is not clear for the Abbes–Saito conductor. One can then equate the two conductors, deducing integrality for the Abbes–Saito conductor; this has been carried out by Chiarellotto and Pulita [86] for one-dimensional representations, and by L. Xiao [411] in the general case.

0.2 Zeta functions of varieties

For the rest of this introduction, we return to Dwork's original example showing the role of p-adic differential equations and their solutions in number theory. This example refers to elliptic curves, for which see Silverman's book [373] for background.

Definition 0.2.1. For λ in some field K, let E_λ be the elliptic curve over K defined by the equation

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

in the projective plane. Remember that there is one point $O = [0:1:0]$ at infinity. There is a natural commutative group law on $E_\lambda(K)$ with identity element O, characterized by the property that three points add to zero if and only if they are collinear. (It is better to say that three points add to zero if they are the three intersections of E_λ with some line, as this correctly permits degenerate cases. For instance, if two of the points coincide, the line must be the tangent to E_λ at that point.)

For elliptic curves over finite fields, one has the following result of Hasse, which generalizes some observations made by Gauss and others.

Theorem 0.2.2 (Hasse)**.** *Suppose λ belongs to a finite field $\mathbb{F}_q$. If we write $\#E_\lambda(\mathbb{F}_q) = q + 1 - a_q(\lambda)$, then $|a_q(\lambda)| \leq 2\sqrt{q}$.*

Proof See [373, 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*.)

Definition 0.2.3. 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 like more familiar examples of zeta functions, is

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

where x runs over Galois orbits of $X(\overline{\mathbb{F}}_q)$, and $\deg(x)$ denote the size of the orbit x. (If you prefer algebro-geometric terminology, you may run x over closed points of the scheme X, in which case $\deg(x)$ denotes the degree of the residue field of x over $\mathbb{F}_q$.)

Example 0.2.4. For $X = E_\lambda$, one can verify that

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

using properties of the Tate module of E_λ; see [373, Theorem V.2.2].

The statement of the Weil conjectures is the following theorem.

Theorem 0.2.5 (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 [196, Appendix C] as a metareference. (Another useful exposition is [330]; see also the chapter notes.) □

Remark 0.2.6. It is worth pointing out that the first complete proof of Theorem 0.2.5 used the fact that for any prime $\ell \neq p$, one has

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

where $H^i_{\mathrm{et}}(X, \mathbb{Q}_\ell)$ is the i-th étale cohomology group of X (or rather, the base change of X to $\overline{\mathbb{F}}_q$) with coefficients in $\mathbb{Q}_\ell$. This is an instance of the *Lefschetz trace formula* in étale cohomology.

0.3 Zeta functions and p-adic differential equations

Remark 0.3.1. The interpretation of Theorem 0.2.5 in terms of étale cohomology (Remark 0.2.6) is all well and good, but there are several downsides. 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 any given cocycle is a coboundary.)

Another important downside is that étale cohomology does not yield 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 on one hand when you study variation of complex periods, and on the other hand when you study variation of zeta functions over $\mathbb{F}_p$.

Here is an explicit example due to Dwork.

Definition 0.3.2. Recall that the *hypergeometric series*

$$F(a, b; c; z) = \sum_{i=0}^{\infty} \frac{a(a+1)\cdots(a+i-1)b(b+1)\cdots(b+i-1)}{c(c+1)\cdots(c+i-1)i!} z^i \tag{0.3.2.1}$$

Cambridge University Press & Assessment
978-1-009-12334-1 — p-adic Differential Equations
Kiran S. Kedlaya
Excerpt
More Information

satisfies the *hypergeometric differential equation*

$$z(1-z)y'' + (c - (a+b+1)z)y' - aby = 0. \tag{0.3.2.2}$$

Set

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

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

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

(One can extend this formula to complex λ, but this requires some care with branch cuts.) This elliptic integral can be viewed as a period integral for the curve E_λ, i.e., one is integrating some meromorphic differential form on E_λ around some loop (or more properly, around some homology class).

Let p 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 that 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 0.3.3. *The series $\alpha(z)$ converges p-adically for $|z| < 1$.*

Proof Exercise. □

Dwork discovered that a closely related function admits a sort of analytic continuation.

Definition 0.3.4. 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 $\lambda \in \overline{\mathbb{F}}_p$ for which E_λ is a *supersingular* elliptic curve, i.e., for which $a_q(\lambda) \equiv 0 \pmod{p}$. (In fact, the roots of $H(z)$ all belong to $\mathbb{F}_{p^2}$, by a theorem of Deuring; see [373, Theorem V.3.1].)

Dwork's analytic continuation result is the following.

Theorem 0.3.5 (Dwork)**.** *There exists a series $\xi(z) = \sum_{i=0}^\infty P_i(z)/H(z)^i$, with each $P_i(z) \in \mathbb{Q}_p[z]$, converging uniformly for those z satisfying $|z| \le 1$ and $|H(z)| = 1$ and such that*

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

Proof See [402, §7]. □

Remark 0.3.6. Note that ξ itself satisfies a differential equation derived from the hypergeometric equation. We will see such equations again once we introduce the notion of a Frobenius structure on a differential equation, in Chapter 17.

In terms of the function ξ, we can compute zeta functions in the Legendre family as follows.

Definition 0.3.7. Let $\mathbb{Z}_q$ be the unique unramified extension of $\mathbb{Z}_p$ with residue field $\mathbb{F}_q$. For $\lambda \in \mathbb{F}_q$, let $[\lambda]$ be the unique q-th root of 1 in $\mathbb{Z}_q$ congruent to λ mod p. (See the notes for Chapter 14 for more discussion of this construction.)

Theorem 0.3.8 (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" (meaning the root of valuation 0) of the polynomial $T^2 - a_q(\lambda)T + q$ occurring (up to reversal) in the zeta function.

Proof See [402, §7]. □

0.4 A word of caution

Example 0.4.1. 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 $x = 0$,

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

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 in such a way that $|p| = p^{-1}$). For comparison, note that the logarithm series

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

converges for $|x| < 1$.

Remark 0.4.2. The conclusion to be taken away from the previous example 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 *Picard–Fuchs equations* which arise from integrals of algebraic functions (e.g., elliptic integrals). Another of Dwork's discoveries is that one can quantify 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. We will carry this out later in the book.

Notes

For detailed notes on the topics discussed in §0.1, see the notes for the chapters referenced.

We again mention [196, Appendix C] and [330] as starting points for further reading about the Weil conjectures. See also [261].

The notion of an analytic function in terms of a uniform limit of rational functions with poles prescribed to certain regions is the original such notion, introduced by Krasner. For this book, we will restrict our consideration of p-adic analysis to working with complete rings in this fashion, without attempting to introduce any notion of nonarchimedean analytic geometry. However, it must be noted that it is much better in the long run to work in terms of 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 [69], but novices may find the text of Fresnel and van der Put [171] or the lecture notes of Bosch [68] more approachable. Two more recent methods, which in some ways are more robust, is Berkovich's theory of nonarchimedean analytic spaces (commonly called *Berkovich spaces*), as introduced in [52] and further developed in [53]; and Huber's theory of *adic spaces*, as introduced in [211] and further developed in [212]. For all three points of view, see also the lecture notes of Conrad [111]. The Berkovich approach will manifest at a very superficial level in Part VII.

Dwork's original analysis of the Legendre family of elliptic curves via the associated hypergeometric equation (which expands upon earlier work of Tate) appears in [140, §8]. The treatment in [402] is more overtly related to p-adic cohomology.

The family of hypergeometric equations with $a, b, c \in \mathbb{Q} \cap \mathbb{Z}_p$ is rich enough that one could devote an entire book to the study of its p-adic properties. Indeed, Dwork did exactly this; the result is [145].

It is possible to resurrect partially the fundamental theorem of ordinary differential equations in the p-adic setting. The best possible results in that direction seem to be those of Priess-Crampe and Ribenboim [341]. One consequence of their work is that a differential equation over $\mathbb{Q}_p$ has a solution if and only if it has a sufficiently good approximate solution; this amounts to a differential version of Hensel's lemma. We too will need noncommutative forms of Hensel's lemma; see Theorem 2.2.2.

Christol [92] has given an interesting retrospective on some of the key ideas of Dwork, including generic points, the transfer principle, and Frobenius structures, which resonate throughout this book.

Exercises

0.1 Prove directly from the definition that the series $F(a, b; c; z)$ converges p-adically for $|z| < 1$ whenever a, b, c are rational numbers with denominators not divisible by p. This implies Lemma 0.3.3.

0.2 Using the fact that $\alpha(z)$ satisfies the hypergeometric equation, write down a nontrivial differential equation with coefficients in $\mathbb{Q}(z)$ satisfied by the function $\xi(z)$.

0.3 Check that the usual formula

$$\liminf_{n\to\infty} |a_n|^{-1/n}$$

for the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n z^n$ still works over a nonarchimedean field. That is, the series converges when $|z|$ is less than this radius and diverges when $|z|$ is greater than this radius.

0.4 Show that in the previous exercise, just like in the archimedean case, a power series over a nonarchimedean field can either converge or diverge at a value of z for which $|z|$ equals the radius of convergence.

0.5 Check that (as claimed in Example 0.4.1) under the normalization

$|p| = p^{-1}$, the exponential series $\exp(z)$ over $\mathbb{Q}_p$ has radius of convergence $p^{-1/(p-1)}$, while the logarithm series $\log(1 - z)$ has radius of convergence 1.

0.6 Show that over $\mathbb{Q}_p$, a power series in z which converges for $|z| \leq 1$ may have an antiderivative which only converges for $|z| < 1$, but its derivative still converges for $|z| \leq 1$. This is the reverse of what happens over an archimedean field.