**$p$-adic differential equations**
**18.787, Kiran S. Kedlaya, MIT, fall 2007**
**Formalism of differential algebra**

In this lecture, we set up some formalism for dealing with differential equations. These can be used for the start of an axiomatic treatment of *differential algebra*, but I will only introduce the minimum for my needs.

Convention: My rings are commutative and unital unless otherwise specified. When I say "noncommutative ring", I really mean "not necessarily commutative ring".

# 1 Differential modules

A *differential ring* is a ring $R$ equipped with a derivation $d : R \to R$, i.e., an additive map satisfying the Leibniz rule

$$d(ab) = ad(b) + bd(a) \qquad (a, b \in R).$$

We expressly allow $d = 0$ unless otherwise specified; this will come in handy in some situations. A differential ring which is also a domain, field, etc., will be called a *differential domain, field*, etc.

A *differential module* over a differential ring $(R, d)$ is a module $M$ equipped with an additive map $D : M \to M$ satisfying

$$D(am) = aD(m) + d(a)m;$$

such a $D$ will also be called a *differential operator* on $M$ relative to $d$. For example, $(R, d)$ is a differential module over itself; any differential module isomorphic to a direct sum of copies of $(R, d)$ is said to be *trivial*. (If we refer to "the" trivial differential module, though, we mean $(R, d)$ itself.) A *differential ideal* of $R$ is a differential submodule of $R$ itself, i.e., an ideal stable under $d$.

The kernel of the derivation $d$ on $R$ is a subring of $R$; if $R$ is a field, then $\ker(d)$ is a subfield. We call this the *constant subring/subfield*. For $(M, D)$ a differential module, an element of $\ker(D)$ is said to be *horizontal*. (This terminology makes sense if you consider *connections* in differential geometry, where the differential operator is measuring the extent to which a section of a vector bundle deviates from some prescribed "horizontal" direction identifying points on one fibre with points on its neighbors.)

For $(M, D)$ a differential module, define

$$H^0(M) = \ker(D), \qquad H^1(M) = \operatorname{coker}(D) = M/D(M).$$

The latter computes Yoneda extensions; see Lemma 1 below.

Convention: I will sometimes refer to the pair $(R, d)$ as a differential ring, but in some cases I will call $R$ itself a differential ring when $d$ is either evident from context or not relevant. Similarly for differential modules.

## 2 Differential modules and differential systems

Let $R$ be a differential ring, and let $M$ be a finite free differential module of rank $n$ over $R$. Let $e_1, \ldots, e_n$ be a basis of $M$. Then for any $v \in M$, we can write $v = v_1 e_1 + \cdots + v_n e_n$ for some $v_1, \ldots, v_n \in R$, and then compute

$$D(v) = v_1 D(e_1) + \cdots + v_n D(e_n) + d(v_1)e_1 + \cdots + d(v_n)e_n.$$

If we define the $n \times n$ matrix $N$ over $R$ by the formula

$$D(e_j) = \sum_{i=1}^{n} D_{ij} e_i,$$

we then have

$$D(v) = \sum_{i=1}^{n} \left( d(v_i) + \sum_j N_{ij} v_j \right) e_i.$$

That is, if we identify $v$ with the column vector $\mathbf{v} = [v_1, \ldots, v_n]$, then

$$D(\mathbf{v}) = N\mathbf{v} + d(\mathbf{v}).$$

Conversely, it is clear that given the underlying finite free $R$-module, any differential module structure is given by such an equation.

In other words, differential modules are a coordinate-free version of differential systems. If you are a geometer, you may wish to go further and think of *differential bundles*, i.e., vector bundles equipped with a differential operator. A differential operator on a vector bundle is usually called a *connection*.

## 3 Operations on differential modules

Differential modules over a fixed differential ring $R$ form a category in which the morphisms (or *homomorphisms*) from $M_1$ to $M_2$ are additive maps $f : M_1 \to M_2$ satisfying $D(f(m)) = f(D(m))$. This category admits certain functors corresponding to some familiar functors on the category of modules over an ordinary ring. (Beware that in the following notations, the subscripted $R$ in the above notations will often be suppressed when it is unambiguous.)

Given two differential modules $M_1, M_2$, the tensor product $M_1 \otimes_R M_2$ in the category of rings may be viewed as a differential module via the formula

$$D(m_1 \otimes m_2) = D(m_1) \otimes m_2 + m_1 \otimes D(m_2).$$

Similarly, the exterior power $\wedge_R^n M$ may be viewed as a differential module via the formula

$$D(m_1 \wedge \cdots \wedge m_n) = \sum_{i=1}^{n} m_1 \wedge \cdots \wedge m_{i-1} \wedge D(m_i) \wedge m_{i+1} \wedge \cdots \wedge m_n;$$

likewise for the symmetric power $\mathrm{Sym}^n_R M$. The module of $R$-homomorphisms $\mathrm{Hom}_R(M_1, M_2)$ may be viewed as a differential module via the formula

$$D(f)(m) = D(f(m)) - f(D(m));$$

the homomorphisms from $M_1$ to $M_2$ as differential modules are precisely the horizontal elements of $\mathrm{Hom}_R(M_1, M_2)$. If $M_2 \cong R$ is trivial, we write $R^\vee$ for $\mathrm{Hom}_R(M_1, R)$ and call it the *dual* of $M_1$; if $M_1$ is finite projective (which is the same as finite locally free if $R$ is a noetherian ring), then $\mathrm{Hom}_R(M_1, M_2) \cong M_1^\vee \otimes M_2$ and the natural map $M_1 \to (M_1^\vee)^\vee$ is an isomorphism.

**Lemma 1.** *Let $M, N$ be differential modules with $M$ finite projective. Then the group $H^1(M^\vee \otimes N)$ is canonically isomorphic to the Yoneda extension group* $\mathrm{Ext}(M, N)$.

*Proof.* The group $\mathrm{Ext}(M, N)$ consists of equivalence classes of exact sequences $0 \to N \to P \to M \to 0$ under the relation that this sequence is equivalent to a second sequence $0 \to N \to P' \to M \to 0$ if there is an isomorphism $P \cong P'$ that induces the identity maps on $M$ and $N$. The addition is to take two such sequences and return the *Baer sum* $0 \to N \to (P \oplus P')/\Delta \to M \to 0$, where $\Delta = \{(n, -n) : n \in N\}$. The identity element is the split sequence $0 \to N \to M \oplus N \to M \to 0$. The inverse of a sequence $0 \to N \to P \to M \to 0$ is the sequence $0 \to N \to P \to M \to 0$ with the map $N \to P$ negated.

Given an extension $0 \to N \to P \to M \to 0$, tensor with $M^\vee$ to get $0 \to M^\vee \otimes N \to M^\vee \otimes P \to M^\vee \otimes P \to 0$, and apply the connecting homomorphism $H^0(M^\vee \otimes M) \to H^1(M^\vee \otimes N)$ from the snake lemma to the trace (the element of $M^\vee \otimes M$ corresponding to the identity map in $\mathrm{Hom}(M, M)$) to get an element of $H^1(M^\vee \otimes N)$. This is the desired map $\mathrm{Ext}(M, N) \to H^1(M^\vee \otimes N)$. To construct its inverse, given an element $H^1(M^\vee \otimes N)$ represented by $x \in M^\vee \otimes N$, form the sequence

$$0 \to N \to \frac{M \oplus N}{(m, \langle m, x \rangle)} \to M \to 0$$

where $\langle \cdot, \cdot \rangle$ represents the natural map $M \times (M^\vee \otimes N) \to N$. □

# 4 Cyclic vectors

Let $R$ be a differential ring, and let $M$ be a finite free differential module of rank $n$ over $R$. A *cyclic vector* for $M$ is an element $m \in M$ such that $m, D(m), \ldots, D^{n-1}(m)$ form a basis of $M$.

**Theorem 2** (Cyclic vector theorem)**.** *Let $R$ be a differential field of characteristic zero with nonzero derivation. Then every finite differential module over $R$ has a cyclic vector.*

For a comment on characteristic $p$, see the exercises.

*Proof.* This is a folklore result, that is, it is old enough that giving a proper attribution is difficult. Many proofs are possible; here is the proof from [DGS, Theorem III.4.2].

We start by normalizing the derivation. For $u \in R^\times$, given one differential module $(M, D)$ over $(R, d)$, we get another differential module $(M, uD)$ over $(R, ud)$, and $m$ is a cyclic vector for one if and only if it is a cyclic vector for the other (because the image of $m$ under $(uD)^j$ is in the span of $u, D(u), \dots, D^j(u)$). We may thus assume (thanks to the assumption that the derivation is nontrivial) that there exists an element $x \in R$ such that $d(x) = x$.

Let $M$ be a differential module of dimension $n$, and choose $m \in M$ so that the dimension $\mu$ of the span of $m, D(m), \dots$ is as large as possible. We derive a contradiction under the hypothesis $\mu < n$.

For $z \in M$ and $\lambda \in \mathbb{Q}$, we now have

$$(m + \lambda z) \wedge D(m + \lambda z) \wedge \cdots + D^\mu(m + \lambda z) = 0$$

in the exterior power $\wedge^{\mu+1} M$. If we write this expression as a polynomial in $\lambda$, it vanishes for infinitely many values, so must be identically zero. Hence each coefficient must vanish separately, including the coefficient of $\lambda^1$, which is

$$\sum_{i=0}^{\mu} m \wedge \cdots \wedge D^{i-1}(m) \wedge D^i(z) \wedge D^{i+1}(m) \cdots \wedge D^\mu(m). \tag{1}$$

Pick $s \in \mathbb{Z}$, substitute $x^s z$ for $z$ in (1), divide by $x^s$, and set equal to zero. We get

$$\sum_{i=0}^{\mu} s^i \Lambda_i(m, z) = 0 \qquad (s \in \mathbb{Z}) \tag{2}$$

for

$$\Lambda_i(m, z) = \sum_{j=0}^{\mu-i} \binom{i+j}{i} m \wedge \cdots \wedge D^{i+j-1}(m) \wedge D^j(z) \wedge D^{i+j+1}(m) \wedge \cdots \wedge D^\mu(m).$$

Again because we are in characteristic zero, we may conclude that (2), viewed as a polynomial in $s$, has all coefficients equal to zero; that is, $\Lambda_i(m, z) = 0$ for all $m, z \in M$.

We now take $i = \mu$ to obtain

$$(m \wedge \cdots \wedge D^{\mu-1}(m)) \wedge z = 0 \qquad (m, z \in M);$$

since $\mu < n$, we may use this to deduce

$$m \wedge \cdots \wedge D^{\mu-1}(m) = 0 \qquad (m \in M).$$

But that means that the dimension of the span of $m, D(m), \dots$ is always at most $\mu - 1$, contradicting the definition of $\mu$. □

If $R$ is not a field, then one obstruction to having a cyclic vector is that $M$ itself might not be a finite free $R$-module. But even if it is, there is no reason to expect in general that cyclic vectors exist; this will create complications for us later.

# 5 Differential polynomials

Let $(R, d)$ be a differential ring. The *ring of twisted polynomials* $R\{T\}$ over $R$ in the variable $T$ is the additive group

$$R \oplus (R \cdot T) \oplus (R \cdot T^2) \oplus \cdots,$$

with noncommuting multiplication given by the formula

$$\left(\sum_{i=0}^{\infty} a_i T^i\right)\left(\sum_{j=0}^{\infty} b_j T^j\right) = \sum_{i,j=0}^{\infty} \sum_{h=0}^{j} \binom{j}{h} a_i d^h(b_j) T^{i+j-h}.$$

In other words, you impose the relation

$$Ta = aT + d(a) \qquad (a \in R)$$

and check that you get a sensible noncommutative ring.

We define the *degree* of a twisted polynomial in the usual way, as the exponent of the largest power of $T$ with a nonzero coefficient. (Pick your favorite convention for the degree of the zero polynomial.)

**Proposition 3** (Ore)**.** *For $R$ a differential field, the ring $R\{T\}$ admits a left division algorithm. That is, if $f, g \in R\{T\}$ and $g \neq 0$, then there exist unique $q, r \in R\{T\}$ with $\deg(r) < \deg(g)$ and $f = gq + r$. (There is also a right division algorithm.)*

*Proof.* Exercise. □

Using the Euclidean algorithm, this yields the following consequence as in the untwisted case.

**Theorem 4** (Ore)**.** *Let $R$ be a differential field. Then $R\{T\}$ is both left principal and right principal; that is, any left ideal (resp. right ideal) has the form $R\{T\}f$ (resp. $fR\{T\}$) for some $f \in R\{T\}$.*

Note that the opposite ring to $R\{T\}$, i.e., the ring with left and right reversed, is again a twisted polynomial ring, but for the derivation $-d$. Given $f \in R\{T\}$, we define the *formal adjoint* of $f$ as the element $f$ in the opposite ring. This operation looks a bit less formal if you also push the coefficients over to the other side, giving what we will call the *adjoint form* of $f$. For instance, the adjoint form of $T^3 + aT^2 + bT + c$ is

$$T^3 + T^2 a + T(b - 2d(a)) + d(d(a)) - d(b) + c.$$

The twisted polynomial ring is rigged up precisely so that for any differential module $M$ over $R$, we get an action of $R\{T\}$ on $M$ under which $T$ acts like $D$. In particular, $R\{T\}$ acts on $R$ itself with $T$ acting like $d$. In fact, the category of differential modules over $R$ is equivalent to the category of left $R\{T\}$-modules. Moreover, if $M$ is a differential module, any cyclic vector $m \in M$ corresponds to an isomorphism $M \cong R\{T\}/R\{T\}f$ for some monic twisted polynomial $f$, where the isomorphism carries $m$ to the class of 1. (You might want to think of $f$ as a sort of "characteristic polynomial" for $M$, except that it depends strongly on the choice of the cyclic vector.)

# 6 Differential equations

You may have been wondering when differential equations will appear, those supposedly being the objects of study of this course. If so, your wait is over.

A *differential equation of order* $n$ over the differential ring $(R, d)$ is an equation of the form

$$(a_n d^n + \cdots + a_1 d + a_0)y = b,$$

with $a_0, \ldots, a_n, b \in R$, and $y$ indeterminate. We say the equation is *homogeneous* if $b = 0$ and *inhomogeneous* otherwise.

Using our setup, we may write this equation as $f(d)y = b$ for some $f \in R\{T\}$. Similarly, we may view systems of differential equations as being equations of the form $f(D)y = b$ where $b$ lives in some differential module $(M, D)$. By the usual method (of introducing extra variables corresponding to derivatives of $y$), we can convert any differential system into a first-order system $Dy = b$. We can also convert an inhomogeneous system into a homogeneous one by adding an extra variable, with the understanding that we would like the value of that last variable to be 1 in order to get back a solution of the original equation.

Here is a more explicit relationship between adjoint polynomials and solving differential equations. Say you start with the cyclic differential module $M \cong R\{T\}/R\{T\}f$ and you want to find a horizontal element. That means that you want to find some $g \in R\{T\}$ such that $Tg \in R\{T\}f$; we may as well assume that $\deg(g) < \deg(f)$. Then by comparing degrees, we see that in fact $Tg = rf$ for some $r \in R$. Write $f$ in adjoint form as $f_0 + Tf_1 + \cdots + T^n$; then

$$rf \equiv rf_0 - d(r)f_1 + d^2(r)f_2 - \cdots \pm d^n(f) \mod TR\{T\}.$$

In this manner, finding a horizontal element becomes equivalent to solving a differential equation.

# 7 Cyclic vectors: a mixed blessing

The reader may at this point be wondering why so many points of view are necessary, since the cyclic vector theorem can be used to transform any differential module into a differential equation, and ultimately differential equations are the things one writes down and wants to solve. Permit me to interject here a countervailing opinion.

In ordinary linear algebra (or in other words, when considering differential modules for the trivial derivation), one can pass freely between linear transformations on a vector space and square matrices if one is willing to choose a basis. The merits of doing this depend on the situation, so it is valuable to have both the matricial and coordinate-free viewpoints well in hand. One can then pass to the characteristic polynomial, but not all information is retained (one loses information about nilpotency), and even information that in principle is retained is sometimes not so conveniently accessed. In short, no one would seriously argue that one can dispense with studying matrices because of the existence of the characteristic polynomial.

The situation is not so different in the differential case. The difference between a differential module and a differential system is merely the choice of a basis, and again it is valuable to have both points of view in mind. However, the cyclic vector theorem may seduce one into thinking that collapsing a differential system into a differential polynomial is an operation without drawbacks, and this is far from the case. For instance, determining whether two differential polynomials correspond to the same differential system is not straightforward.

More seriously for our purposes, the cyclic vector theorem only applies over a differential field. Many differential modules are more naturally defined over some ring which is not a field, e.g., those coming from geometry which should be defined over some sort of ring of functions on some sort of geometric space. Working with differential modules instead of differential polynomials has a tremendously clarifying effect over rings.

We find it unfortunate that much of the literature on complex ordinary differential equations, and nearly all of the literature on $p$-adic ordinary differential equations, is mired in the language of differential polynomials. By instead switching between differential modules and differential polynomials as appropriate, we will be able to demonstrate strategies that lead to a more systematic development of the $p$-adic theory.

# 8 Taylor series

Let $R$ be a *topological* differential ring, i.e., a ring equipped with a topology and a derivation such that all operations are continuous. Assume also that $R$ is a $\mathbb{Q}$-algebra. Let $M$ be a topological differential module over $R$, i.e., a differential module such that all operations are continuous.

For $r \in R$ and $m \in M$, We define the *Taylor series* $T(r, m)$ as the infinite sum

$$\sum_{i=0}^{\infty} \frac{r^i}{i!} D^i(m)$$

whenever the sum converges absolutely (i.e., all rearrangements converge to the same value). This map is additive whenever possible: if $m_1, m_2 \in M$, then

$$T(r, m_1) + T(r, m_2) = T(r, m_1 + m_2)$$

whenever all three terms make sense. Also, the map $T(r, \cdot) : R \to R$ is a ring homomorphism whenever it is defined. That is, if $s_1, s_2 \in R$, then (by the Leibniz rule)

$$T(r, s_1)T(r, s_2) = T(r, s_1 s_2)$$

whenever all three terms make sense. More generally, if $s \in R$, $m \in M$, then

$$T(r, s)T(r, m) = T(r, sm)$$

whenever all three terms make sense. Loosely put, the map $T(r, \cdot)$ on $M$ is semilinear for the map $T(r, \cdot)$ on $R$.

# 9 Notes

The subject of differential algebra is rather well-developed; a classic treatment is the book of Ritt [Rit50]. As in abstract algebra in general, development of differential algebra was partly driven by *differential Galois theory*, i.e., the study of when solutions of differential equations can be expressed in terms of solutions to ostensibly simpler differential equations. A good introduction to the latter is [SvdP03].

Twisted polynomials were introduced by Ore [Ore33]. They are actually somewhat more general than we have discussed; for instance, one can also twist by an endomorphism $\tau : R \to R$ by imposing the relation $Ta = \tau(a)T$. (This enters the realm of the analogue of differential algebra called *difference algebra*.) Moreover, one can twist by both an endomorphism and a derivation if they are compatible in an appropriate way, and one can even study differential/difference Galois theory in this setting. A unifying framework for doing so, which is also suitable for considering multiple derivations and automorphisms, is given by André [And01].

# 10 Exercises

1. Prove that if $M$ is a locally free differential module over $R$ of rank 1, then $M^\vee \otimes M$ is trivial (as a differential module).

2. Check that in characteristic $p > 0$, the cyclic vector theorem holds for modules of rank less than $p$, but may fail for modules of rank $p$.

3. Give a counterexample to the cyclic vector theorem for a differential field of characteristic zero with trivial derivation.

4. Verify that $R\{T\}$ is indeed a noncommutative ring; the content in this is to check associativity of multiplication.

5. Prove the division algorithm (Proposition 3).