Lecture 15: April 1

The Cauchy-Kovalevskaya theorem. Last time, we showed that the solution to the initial value problem

$$\frac{du}{dt} = F(u), \quad u(0) = 0,$$

is real-analytic near $t = 0$, provided that this is true for the function F. I also showed you Cauchy's proof, using the "method of majorants". Today, we are going to generalize this result to partial differential equations. We work on $\mathbb{R}^n$, with coordinates $x_1, \dots, x_n$, and consider a partial differential equation of the form

$$Pu = \sum_{|\alpha| \leq k} f_\alpha \partial^\alpha u = 0,$$

where each f_α is a real-analytic function in a neighborhood of the origin, say. (And $\partial_j = \partial/\partial x_j$, as usual.) In other words, P is a linear differential operator of order k with real-analytic coefficients. We will specify the initial conditions on the hyperplane $x_n = 0$, which is a copy of $\mathbb{R}^{n-1}$. They are

$$u\big|_{\mathbb{R}^{n-1}} = g_0, \quad \partial_n u\big|_{\mathbb{R}^{n-1}} = g_1, \quad \dots, \quad \partial_n^{k-1} u\big|_{\mathbb{R}^{n-1}} = g_{k-1},$$

where $g_0, g_1, \dots, g_{k-1}$ are real-analytic in a neighborhood of the origin in $\mathbb{R}^{n-1}$. From this data, we can of course compute all partial derivatives of u of order at most $k-1$ on $\mathbb{R}^{n-1}$; indeed, if $\alpha \in \mathbb{N}^n$ is a multi-index, then

$$\partial^\alpha u\big|_{\mathbb{R}^{n-1}} = \partial_1^{\alpha_1} \cdots \partial_{n-1}^{\alpha_{n-1}} g_{\alpha_n}, \tag{15.1}$$

provided that $\alpha_n \leq k-1$.

The goal is to show that the solution u is real-analytic near the origin. For that to be true, the Taylor series of u at the origin needs to be determined by the equation $Pu = 0$ plus the initial conditions, and so we had better be able to compute *all* partial derivatives of u at the origin. Since we can always differentiate along $\mathbb{R}^{n-1}$, the real question is how to find

$$\partial_n^j u\big|_{\mathbb{R}^{n-1}}$$

for $j \geq k$. Clearly, this information has to come from $Pu = 0$. Since P has order k, we can rewrite $Pu = 0$ as

$$f_{(0,\dots,0,k)} \cdot \partial_n^k u = - \sum_{\alpha_n \leq k-1} f_\alpha \partial^\alpha u,$$

and in view of (15.1), we can solve this for $\partial_n^k u\big|_{\mathbb{R}^{n-1}}$ if and only if the restriction of the coefficient function $f_{(0,\dots,0,k)}$ to $\mathbb{R}^{n-1}$ is everywhere nonzero. (If we only care about what happens at the origin, then the condition is that $f_{(0,\dots,0,k)}$ should be nonzero at the origin.) If that is the case, we can of course divide through by $f_{(0,\dots,0,k)}$ and arrange that ∂_n^k appears with coefficient 1.

Definition 15.2. We say that P is *non-characteristic* with respect to the hypersurface $x_n = 0$ if the coefficient function $f_{(0,\dots,0,k)}$ is everywhere nonzero on $\mathbb{R}^{n-1}$.

Assuming that P is non-characteristic (and $f_{(0,\dots,0,k)} = 1$), we can rewrite the equation $Pu = 0$ in the form

$$\partial_n^k u = Qu,$$

where Q is a differential operator of order k in which ∂_n^k does not appear. We can now use this equation recursively, together with (15.1), to compute $\partial^\alpha u\big|_{\mathbb{R}^{n-1}}$ for every $\alpha \in \mathbb{N}^n$. In particular, assuming that P is non-characteristic, the equation $Pu = 0$ together with the initial conditions on $\mathbb{R}^{n-1}$ give enough information to compute the Taylor series for u at the origin. We can now state the PDE version of the Cauchy-Kovalevskaya theorem.

Theorem 15.3 (Cauchy-Kovalevskaya)**.** *Let P be a linear partial differential operator of order k whose coefficients are real-analytic near the origin in $\mathbb{R}^n$. If P is non-characteristic with respect to $x_n = 0$, then the boundary-value problem*

$$Pu = 0, \quad u\big|_{\mathbb{R}^{n-1}} = g_0, \quad \partial_n u\big|_{\mathbb{R}^{n-1}} = g_1, \quad \ldots, \quad \partial_n^{k-1} u\big|_{\mathbb{R}^{n-1}} = g_{k-1},$$

has a unique real-analytic solution u near the origin in $\mathbb{R}^n$, for every choice of functions $g_0, g_1, \ldots, g_{k-1}$ real-analytic near the origin in $\mathbb{R}^{n-1}$.

Example 15.4. Here is an example to show that the solution can fail to be real-analytic if P is "characteristic". This example is due to Kovalevskaya herself. Consider the heat equation $\partial_t u = \partial_x^2 u$ in $\mathbb{R}^2$, with coordinates (x, t). Since the equation is first-order in t, we only need a single initial condition $u(x, 0) = g(x)$. Note that the operator $P = \partial_t - \partial_x^2$ is characteristic with respect to $t = 0$, because it has order 2, but no term involving ∂_t^2. Here is a heuristic reason why we cannot expect u to be real-analytic in general. From the equation, we get

$$\partial_t^n u = \partial_x^{2n} u,$$

and at $(x, t) = (0, 0)$, this evaluates to $g^{(2n)}(0)$. If the Taylor series of g at the origin has a finite radius of convergence, then

$$|g^{(2n)}(0)| \geq C \frac{(2n)!}{r^{2n}}$$

for some $C, r > 0$. But this means that the function $h(t) = u(0, t)$ cannot be real-analytic in t: indeed, from the above, we deduce that

$$|h^{(n)}(0)| \geq C \frac{(2n)!}{r^{2n}},$$

and since $(2n)!$ grows so much faster than $n!$, the Taylor series of $h(t)$ has radius of convergence equal to zero. For an actual example, take $g(x) = 1/(x^2 + 1)$.

Now let me give an outline of the proof of Theorem 15.3. As explained above, we can rewrite the equation $Pu = 0$ in the form

$$\partial_n^k u = Qu,$$

where Q is a differential operator of order k with real-analytic coefficients, such that Q has order at most $k-1$ in ∂_n. Moreover, we can subtract a suitable real-analytic function from u to arrange that $g_0 = g_1 = \ldots = g_{k-1} = 0$. We now rewrite the problem as a system of first-order PDE for $N = \binom{n+k-1}{n} + 1$ unknown functions $u_1, \ldots, u_N$. These functions are the $N - 1$ partial derivatives $\partial^\alpha u$ for $|\alpha| \leq k - 1$, and the auxiliary $u_N = x_n$. In vector notation, the system takes the form

$$\frac{\partial \underline{u}}{\partial x_n} = \sum_{j=1}^{n-1} B_j(x_1, \ldots, x_{n-1}) \frac{\partial \underline{u}}{\partial x_j} + B_0(x_1, \ldots, x_{n-1}) \underline{u}, \tag{15.5}$$

where $\underline{u} = (u_1, \ldots, u_N)$, and where the coefficient matrices $B_0, \ldots, B_{n-1}$ are derived from Q, hence real-analytic near the origin. Note that we threw in the function $u_N = x_n$ in order to make the coefficients be independent of x_n; of course, the corresponding equation is simply $\partial u_N / \partial x_n = 1$. The initial condition is that $\underline{u}$ is the zero vector for $x_n = 0$.

Now one can again use the method of majorants to prove that $\underline{u}$ is real-analytic near the origin in $\mathbb{R}^n$. From (15.5), all partial derivatives of $\underline{u}$ at the origin are given by (very complicated) universal polynomials with nonnegative integer coefficients in the partial derivatives of $B_0, \ldots, B_{n-1}$ at the origin. Using the fact that the coefficient matrices are real-analytic near the origin, one can again write down simple majorants for each of them, and then explicitely solve the resulting system of first-order PDE to show that its solution $\underline{v}$, and hence also $\underline{u}$, is real-analytic near the origin.

Non-characteristic $\mathscr{D}$-modules. Here is a geometric interpretation for the condition that P is non-characteristic with respect to $x_n = 0$. If $P = \sum_\alpha f_\alpha \partial^\alpha$ has order k as above, then its principal symbol

$$\sigma_k(P) = \sum_{|\alpha|=k} f_\alpha(x_1, \dots, x_n) \cdot \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$$

is a homogeneous polynomial of degree k in the variables $\xi_1, \dots, \xi_n$. We said that P is non-characteristic iff $f_{(0,\dots,0,k)}(x_1, \dots, x_{n-1}, 0) \neq 0$ for every $x_1, \dots, x_{n-1}$. Another way of saying this is that if we set $x_n = 0$ and assign arbitrary values to the variables $x_1, \dots, x_{n-1}, \xi_1, \dots, \xi_{n-1}$, then $\sigma_k(P)$, considered as a polynomial in the remaining variable ξ_n, always has degree exactly k. The geometric meaning of this condition is as follows. We have the usual maps between the cotangent bundles $T^*\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ and $T^*\mathbb{R}^{n-1} = \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$:

$$\begin{array}{ccc} \mathbb{R}^{n-1} \times_{\mathbb{R}^n} T^*\mathbb{R}^n & \xrightarrow{di} & T^*\mathbb{R}^{n-1} \\ \Big\downarrow{\scriptstyle p_2} & & \\ T^*\mathbb{R}^n & & \end{array}$$

Using $x_1, \dots, x_n, \xi_1, \dots, \xi_n$ as coordinates on $T^*\mathbb{R}^n$, the maps are just

$$\begin{aligned} p_2(x_1, \dots, x_{n-1}, \xi_1, \dots, \xi_n) &= (x_1, \dots, x_{n-1}, 0, \xi_1, \dots, \xi_n) \\ di(x_1, \dots, x_{n-1}, \xi_1, \dots, \xi_n) &= (x_1, \dots, x_{n-1}, \xi_1, \dots, \xi_{n-1}). \end{aligned}$$

Consider the subset $\operatorname{Ch}(P) \subseteq T^*\mathbb{R}^n$ defined by the equation $\sigma_k(P) = 0$. Setting $x_n = 0$ and prescribing values for $x_1, \dots, x_{n-1}, \xi_1, \dots, \xi_{n-1}$ amounts to looking at the fibers of $p_2^{-1} \operatorname{Ch}(P)$ over $T^*\mathbb{R}^{n-1}$, and so P is non-characteristic exactly when the projection from $p_2^{-1} \operatorname{Ch}(P)$ to $T^*\mathbb{R}^{n-1}$ is a finite morphism of degree k. If we observe that $\operatorname{Ch}(P)$ is the characteristic variety of the $\mathscr{D}$-module $A_n(\mathbb{R})/A_n(\mathbb{R})P$, this finiteness condition makes sense for arbitrary coherent $\mathscr{D}$-modules.

Let me now give the general definition. Suppose that $f \colon X \to Y$ is a morphism between two nonsingular algebraic varieties. Here is the diagram of the induced morphisms between cotangent bundles:

$$\begin{array}{ccc} X \times_Y T^*Y & \xrightarrow{df} & T^*X \\ \Big\downarrow{\scriptstyle p_2} & & \\ T^*Y & & \end{array}$$

Definition 15.6. Let $\mathcal{M}$ be a coherent left $\mathscr{D}_Y$-module. We say that $\mathcal{M}$ is *non-characteristic* with respect to $f \colon X \to Y$ if the morphism

$$df \colon p_2^{-1} \operatorname{Ch}(\mathcal{M}) \to T^*X$$

is finite over its image.

Example 15.7. Consider the closed embedding $i \colon \mathbb{A}_k^{n-1} \hookrightarrow \mathbb{A}_k^n$, defined by $x_n = 0$. Our earlier discusion shows that if $P \in A_n$ is nonzero, then the left A_n-module A_n/A_nP is non-characteristic with respect to i if and only if the differential operator P is non-characteristic with respect to $x_n = 0$ in the classical sense.

Example 15.8. If $f \colon X \to Y$ is a smooth morphism, then every coherent $\mathscr{D}_Y$-module is non-characteristic with respect to f. Indeed, smoothness means that we have a short exact sequence

$$0 \to f^*\Omega^1_{Y/k} \to \Omega^1_{X/k} \to \Omega^1_{X/Y} \to 0,$$

with $\Omega^1_{X/Y}$ locally free of rank $\dim X - \dim Y$. But this says that

$$df\colon X \times_Y T^*Y \to T^*X$$

is a closed embedding (of codimension $\dim X - \dim Y$), and so $p_2^{-1}\operatorname{Ch}(\mathcal{M})$ is trivially finite over its image in T^*X.

In the following example, we compute the pullback of an A_n-module of the form A_n/A_nP to the hypersurface $x_n = 0$, in the case where P is non-characteristic.

Example 15.9. Consider the left A_n-module $M = A_n/A_nP$, where $P \in A_n$ is a nonzero differential operator of order $r \geq 0$. Suppose that M is non-characteristic with respect to the closed embedding $i\colon \mathbb{A}^{n-1} \hookrightarrow \mathbb{A}^n$ defined by the equation $x_n = 0$. We claim that, in this case, the pullback i^*M is not only coherent, but actually a free A_n-module of rank r. The definition of the pullback gives

$$i^*M = k[x_1, \dots, x_{n-1}] \otimes_{k[x_1,\dots,x_n]} M \cong A_n/(x_nA_n + A_nP), \tag{15.10}$$

where the right-hand side is a left A_{n-1}-module in the obvious way. We have a morphism of left A_{n-1}-modules

$$\varphi\colon A_{n-1}^{\oplus r} \to A_n/(x_nA_n + A_nP)$$
$$(Q_0, Q_1, \dots, Q_{r-1}) \mapsto Q_0 + Q_1\partial_n + \cdots + Q_{r-1}\partial_n^{r-1}.$$

We will show that φ is an isomorphism. Let us first argue that ∂_n^r is in the image. We can write our differential operator $P \in A_n$ uniquely in the form

$$P = f\partial_n^r - P_{r-1}\partial_n^{r-1} - \cdots - P_1\partial_n - P_0,$$

where $f \in k[x_1, \dots, x_n]$ and where $P_0, \dots, P_{r-1} \in A_n$ do not involve ∂_n. The fact that P is non-characteristic means that f is nowhere vanishing on $\mathbb{A}^{n-1}$; after rescaling, we can assume that $f = 1 - x_ng$. Writing $P_j = Q_n + x_nR_j$, with $Q_j \in A_{n-1}$, we get

$$\partial_n^r = \sum_{j=0}^{r-1} Q_j\partial_n^j + x_n\Big(g\partial_n^r + \sum_{j=0}^{r-1} R_j\partial_n^j\Big) + P, \tag{15.11}$$

and so ∂_n^r belongs to the image of φ. Using the relation in (15.11) repeatedly, we see that this is true for all powers of ∂_n, and so φ is surjective.

It remains to prove that φ is injective. This is equivalent to saying that if

$$Q_0 + Q_1\partial_n + \cdots + Q_{r-1}\partial_n^{r-1} = x_nS + TP$$

for some $Q_0, \dots, Q_{r-1} \in A_{n-1}$ and $S, T \in A_n$, then actually $Q_0 = \cdots = Q_{r-1} = 0$. We can write $T = x_nT_0 + T_1$, in such a way that x_n does not appear in T_1; since $x_nS + TP = x_n(S + T_0) + T_1P$, we can therefore assume without loss of generality that T does not involve x_n. Now suppose, for the sake of contradiction, that $T \neq 0$. On the right-hand side of the equation, ∂_n^r appears with a nonzero coefficient: indeed, P contains $(1 - x_ng)\partial_n^r$, and since T does not involve x_n, it is not possible to cancel this term against anything from x_nS. But this clearly contradicts the fact that ∂_n^r does not appear on the left-hand side of the equation. The conclusion is that $T = 0$; and then also $Q_0 = \cdots = Q_{r-1} = 0$, because the right-hand side is divisible by x_n, whereas the left-hand side does not involve x_n.

The preceding example, together with the Cauchy-Kovalevskaya theorem, sheds some light on what the pullback of $\mathscr{D}$-modules has to do with differential equations.

Example 15.12. Continuing with the previous example, let us take $k = \mathbb{R}$. Set $\mathcal{M} = \mathscr{D}_{\mathbb{R}^n}/\mathscr{D}_{\mathbb{R}^n}P$. Let us denote by $\mathscr{R}_{\mathbb{R}^n}$ the sheaf of real-analytic functions on $\mathbb{R}^n$; it is a left $\mathscr{D}_{\mathbb{R}^n}$-module in the obvious way. Recall from Lecture 1 that real-analytic solutions to the equation $Pu = 0$ on an open subset $U \subseteq \mathbb{R}^n$ correspond naturally to morphisms of left $\mathscr{D}_{\mathbb{R}^n}$-modules $\mathcal{M} \to \mathscr{R}_{\mathbb{R}^n}$ over U; here the morphism takes the generator $1 \in \Gamma(U, \mathscr{D}_{\mathbb{R}^n})$ to the corresponding function $u \in \Gamma(U, \mathscr{R}_{\mathbb{R}^n})$.

In this notation, the Cauchy-Kovalevskaya theorem says that if $V \subseteq \mathbb{R}^{n-1}$ is an open subset, and $g_0, g_1, \dots, g_{r-1} \in \Gamma(V, \mathscr{R}_{\mathbb{R}^{n-1}})$ are arbitrary real-analytic functions on V, there is an open subset $U \subseteq \mathbb{R}^n$ with $U \cap \mathbb{R}^{n-1} = V$, and a real-analytic function $u \in \Gamma(U, \mathscr{R}_{\mathbb{R}^n})$, such that $Pu = 0$ and

$$\partial_n^j u\big|_{\mathbb{R}^{n-1}} = g_j \quad \text{for } j = 0, 1, \dots, r-1.$$

By what we have just said, u may be viewed as a section of the sheaf

$$i^{-1}\mathcal{H}om_{\mathscr{D}_{\mathbb{R}^n}}\big(\mathcal{M}, \mathscr{R}_{\mathbb{R}^n}\big)$$

on the open subset V. Now we have a natural morphism of sheaves

$$i^{-1}\mathcal{H}om_{\mathscr{D}_{\mathbb{R}^n}}\big(\mathcal{M}, \mathscr{R}_{\mathbb{R}^n}\big) \to \mathcal{H}om_{\mathscr{D}_{\mathbb{R}^{n-1}}}\big(i^*\mathcal{M}, i^*\mathscr{R}_{\mathbb{R}^n}\big) \to \mathcal{H}om_{\mathscr{D}_{\mathbb{R}^{n-1}}}\big(i^*\mathcal{M}, \mathscr{R}_{\mathbb{R}^{n-1}}\big);$$

it works by applying the pullback functor i^* to a morphism of left $\mathscr{D}_{\mathbb{R}^n}$-modules $\mathcal{M} \to \mathscr{R}_{\mathbb{R}^n}$, and then composing with the restriction morphism $i^*\mathscr{R}_{\mathbb{R}^n} \to \mathscr{R}_{\mathbb{R}^{n-1}}$. The preceding example shows that $i^*\mathcal{M}$ is a free $\mathscr{D}_{\mathbb{R}^{n-1}}$-module of rank r, generated by the images of $1, \partial_n, \dots, \partial_n^{r-1}$. Thus

$$\mathcal{H}om_{\mathscr{D}_{\mathbb{R}^{n-1}}}\big(i^*\mathcal{M}, \mathscr{R}_{\mathbb{R}^{n-1}}\big) \cong \mathscr{R}_{\mathbb{R}^{n-1}}^{\oplus r},$$

and one checks that the resulting morphism

$$i^{-1}\mathcal{H}om_{\mathscr{D}_{\mathbb{R}^n}}\big(\mathcal{M}, \mathscr{R}_{\mathbb{R}^n}\big) \to \mathscr{R}_{\mathbb{R}^{n-1}}^{\oplus r}$$

takes u to its boundary values

$$u\big|_{\mathbb{R}^{n-1}}, \partial_n u\big|_{\mathbb{R}^{n-1}}, \cdots, \partial_n^{r-1} u\big|_{\mathbb{R}^{n-1}}.$$

This means that we can interpret the Cauchy-Kovalevskaya theorem, in more fancy language, as the statement that the morphism

$$i^{-1}\mathcal{H}om_{\mathscr{D}_{\mathbb{R}^n}}\big(\mathcal{M}, \mathscr{R}_{\mathbb{R}^n}\big) \to \mathcal{H}om_{\mathscr{D}_{\mathbb{R}^{n-1}}}\big(i^*\mathcal{M}, \mathscr{R}_{\mathbb{R}^{n-1}}\big)$$

is an isomorphism of sheaves on $\mathbb{R}^{n-1}$. This tells us that the $\mathscr{D}$-module pullback $i^*\mathcal{M}$ has to do with the *boundary conditions* for the partial differential equation $Pu = 0$; the fact that $i^*\mathcal{M}$ is free of rank r means that we can specify r independent real-analytic functions as boundary conditions.

Non-characteristic pullback. Our next goal is to show that if $f \colon X \to Y$ is a morphism between nonsingular algebraic varieties, and if $\mathcal{M}$ is a coherent left $\mathscr{D}_Y$-module that is non-characteristic with respect to f, then the pullback $f^*\mathcal{M}$ is coherent over $\mathscr{D}_X$. To simplify the analysis, we are going to factor f through its graph. Let us see how this factorization interacts with being non-characteristic.

Suppose for a moment that we have an arbitrary factorization

$$\begin{tikzcd} X \arrow[r, "g"] \arrow[rr, bend right, "f"'] & Z \arrow[r, "h"] & Y \end{tikzcd}$$

with Z nonsingular. We can then draw the following big diagram of induced morphisms between cotangent bundles:

$$\begin{tikzcd} X \times_Y T^*Y \arrow[r, "q"] \arrow[d, "g\times\mathrm{id}"] \arrow[rr, bend left, "df"] & X \times_Z T^*Z \arrow[r, "dg"] \arrow[d, "p_2"] & T^*Z \\ Z \times_Y T^*Y \arrow[r, "dh"] \arrow[d, "p_2"] & T^*Z & \\ T^*Y & & \end{tikzcd}$$

If $h\colon Z \to Y$ is a smooth morphism, then dh is a closed embedding, and so its base change along $g\colon X \to Z$, which is denoted by q in the diagram above, is also a closed embedding. Since $df = dg \circ q$, we see that the subset $p_2^{-1}\operatorname{Ch}(\mathcal{M})$ of $X \times_Y T^*Y$ is finite over T^*Z if and only if its image under q is finite over T^*Z. This observation can be used to reduce the study of non-characteristic pullback to two special cases: smooth morphisms and closed embeddings.

Exercises.

Exercise 15.1. On $\mathbb{R}^n$, we use coordinates $x_1, \dots, x_n$. Let $\mathcal{M} = \mathscr{D}_{\mathbb{R}^n}/\mathscr{D}_{\mathbb{R}^n}P$, where P is a differential operator of order r that is non-characteristic with respect to $x_n = 0$. Show that the morphism

$$i^{-1}\mathcal{H}om_{\mathscr{D}_{\mathbb{R}^n}}\big(\mathcal{M}, \mathscr{R}_{\mathbb{R}^n}\big) \to \mathscr{R}_{\mathbb{R}^{n-1}}^{\oplus r}$$

in Example 15.12 takes a real-analytic solution to the equation $Pu = 0$ to the r-vector of its normal derivatives

$$u\big|_{\mathbb{R}^{n-1}}, \partial_n u\big|_{\mathbb{R}^{n-1}}, \cdots, \partial_n^{r-1} u\big|_{\mathbb{R}^{n-1}}.$$