NOTES ON THE CHARACTERISTIC VARIETY

BENJAMIN MCKAY

Abstract. These are my notes on chapter 5 of the book [1]. They were the subject of lectures I gave at Duke University in 1998.

1. Introduction

The characteristic directions of a solution to a system of partial differential equations are the directions in which the solution can be perturbed "most easily", with the partial differential equations "not noticing" or "paying the least attention". Motions in those directions are ignored at highest order. We will make this precise shortly. Intuitively, we would like to take advantage of these directions in trying to deform through a family of solutions; but also these directions may allow high frequency perturbation, so creasing or spiking.

Characteristic directions at each point form an algebraic variety. In the simplest case, where this variety is a linear subspace, it is the tangent space to a foliation of the solution. This is often a useful way to carve up solutions.

2. Calculating the Characteristic Variety of a Linear Pfaffian System

Given

$$d\begin{pmatrix}\theta^1\\ \vdots\\ \theta^{s_0}\end{pmatrix} = -\begin{pmatrix}\varpi^1_1 & \dots & \varpi^1_n\\ \vdots & \ddots & \vdots\\ \varpi^{s_0}_1 & \dots & \varpi^{s_0}_n\end{pmatrix}\wedge\begin{pmatrix}\omega^1\\ \vdots\\ \omega^n\end{pmatrix} \qquad \mathrm{mod}\ \theta^1,\dots,\theta^{s_0}$$

where

$$\theta^1,\dots,\theta^{s_0},\omega^1,\dots,\omega^n,\pi^1,\dots,\pi^t$$

is a coframing, and each ϖ^i_j is a multiple of the π^i, we can write down a basis for the linear relations among the ϖ^i_j. For example suppose

$$d\begin{pmatrix}\theta^1\\ \theta^2\\ \theta^3\\ \theta^4\end{pmatrix} = -\begin{pmatrix}0 & 0 & 0\\ \pi^1 & 0 & \pi^2\\ 0 & \pi^3 & \pi^4\\ \pi^2 & \pi^4 & \pi^5\end{pmatrix}\wedge\begin{pmatrix}\omega^1\\ \omega^2\\ \omega^3\end{pmatrix} \qquad \mathrm{mod}\ \theta^1,\dots,\theta^4$$

Date: July 18, 2001.

(as on page 18, equation (1) of [2]). Then the linear relations are

$$\begin{aligned} \varpi^1_1 &= 0 \\ \varpi^1_2 &= 0 \\ \varpi^1_3 &= 0 \\ \varpi^2_2 &= 0 \\ \varpi^3_1 &= 0 \\ \varpi^2_3 - \varpi^4_1 &= 0 \\ \varpi^3_3 - \varpi^4_2 &= 0 \end{aligned}$$

For each such relation, for instance the relation

$$\varpi^3_3 - \varpi^4_2 = 0$$

we write down a row vector with s_0 columns (in our example $s_0 = 4$), and translate each term in our linear relation as follows: $a\varpi^i_j$ tells us to place $a\xi_j$ in the i-th column. So

$$\varpi^3_3 - \varpi^4_2 = 0$$

gives

$$\begin{pmatrix} 0 & 0 & \xi_3 & -\xi_2 \end{pmatrix}$$

If we do this for all of our linear relations, we obtain a collection of row vectors, which we write down one after another to form a matrix:

$$\begin{pmatrix} \xi_1 & 0 & 0 & 0 \\ \xi_2 & 0 & 0 & 0 \\ \xi_3 & 0 & 0 & 0 \\ 0 & \xi_2 & 0 & 0 \\ 0 & 0 & \xi_1 & 0 \\ 0 & \xi_3 & 0 & -\xi_1 \\ 0 & 0 & \xi_3 & -\xi_2 \end{pmatrix}$$

This matrix is the *symbol matrix*. The characteristic variety of a linear Pffafian system can be defined to be the projectivization of the set of

$$\xi = (\xi_1, \xi_2, \dots, \xi_n)$$

for which the symbol matrix does not have full rank. In our example, we see if we take the determinants of 3×3 minors, they are

$$\xi_1^2\xi_2, \ \xi_1\xi_2^2, \ \xi_1\xi_2\xi_3$$

up to sign. In any case there will be a factor of $\xi_1\xi_2$. Dividing out that factor leaves ξ_1, ξ_2, ξ_3. But the ideal generated by ξ_1, ξ_2, ξ_3 is the whole ring of homogeneous polynomials. Therefore the factor ξ_1, ξ_2 cuts out the characteristic variety. The variety generated will be, when projectivized, a pair of lines

$$(\xi_1 = 0) \cup (\xi_2 = 0)$$

One obvious but crucial distinction has to be made: between the real and complex characteristic variety. The real points ξ satisfying the above condition belong to the real characteristic variety, but it may also be useful to study the complex

characteristic variety consisting of the complex ξ satisfying the above real algebraic equations. Either characteristic variety is defined by the same real equations.

If there are no Cauchy characteristics, then $\xi = (\xi_1, \dots \xi_n)$ belongs to the characteristic variety of an n dimensional integral element E precisely when the hyperplane

$$\xi_i \omega^i = 0$$

in E is characteristic, i.e. is contained in more than one integral element.

If the real characteristic variety consists of just a point, then we can use the integrability of characteristics to see that this provides a foliation of any integral n dimensional manifold by $n-1$ dimensional submanifolds. In the example case, we have the two projective lines $\xi_1 = 0$ and $\xi_2 = 0$, which are dual to the projective points

$$\omega^2 = \omega^3 = 0$$

and

$$\omega^1 = \omega^3 = 0$$

respectively. So on a solution 3 manifold, we will find that it is foliated by 2 families of characteristic curves.

Note that if we have a row of 0's in the tableau, say in the first row, as we did in the first row of our example, then it will lead to rows in the symbol matrix like:

$$\begin{pmatrix} \xi_1 & 0 & \dots & 0 \\ \xi_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \xi_n & 0 & \dots & 0 \end{pmatrix}$$

which have no effect on the resulting characteristic variety, since the rest of the symbol matrix will never use this first column. Consequently, we can just forget these rows. On the other extreme, if we have a row of the tableau which has $\pi's$ in it which are entirely independent of each other and of anything else in the tableau, say in the first row:

$$d\theta^1 = -\pi^1_1 \wedge \omega^1 - \dots - \pi^1_n \wedge \omega^n$$

with all π^1_j independent of each other, and appearing nowhere else in the tableau, then this contributes nothing to the symbol matrix, and thus nothing to the characteristic variety.

3. What does this have to do with classical notions?

Recall the concept of characteristic of a linear differential equation with constant coefficients

$$Pu = \sum_{|\alpha| \le d} c_\alpha \partial^\alpha u$$

of degree d, where $u : V \to W$ is a map between two finite dimensional real vector spaces, and the α varies over appropriate multiindicies, using some system of linear coordinates on V, and

$$c_\alpha \in \operatorname{Lin}(W, Z)$$

for Z some finite dimensional real vector space. Suppose that we pick a large number λ and a function $f : V \to \mathbb{R}$, and consider for $u : V \to W$ the expression

$$e^{-i\lambda f} \cdot P\left[e^{i\lambda f} u\right]$$

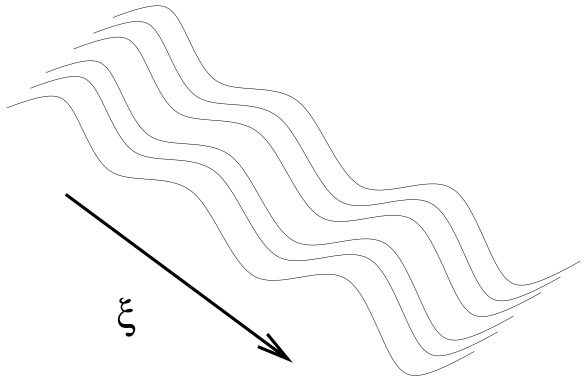

Figure 1. A wave and its covector

The idea is that $e^{i\lambda f}$ is a high frequency disturbance of u, and we want to see how P reacts to it, to highest order in λ. If $x^1, \dots, x^n$ are coordinates on V, then

$$e^{-i\lambda f} \cdot P\left[e^{i\lambda f} u\right] = (i\lambda)^d \sum_{|\alpha|=d} c_\alpha \prod_j \left(\frac{\partial f}{\partial x^j}\right)^{\alpha_j} u + O\left(|\lambda|^{d-1}\right)$$

(To convince yourself of this, try it out on $P = \frac{\partial}{\partial x^j}$ or $P = \frac{\partial^2}{\partial x^i x^j}$.) We call this leading order term the symbol:

$$\sigma_P(\xi) = \sum_{|\alpha|=d} c_\alpha \prod_j \xi_j^{\alpha_j} \in \operatorname{Lin}(W, Z)$$

for $\xi \in V^*$, so

$$e^{-i\lambda f} \cdot P\left[e^{i\lambda f} u\right] = (i\lambda)^d \sigma_P(df) u + O\left(|\lambda|^{d-1}\right)$$

We define the characteristic variety to be the real algebraic variety

$$\Xi = \{[\xi] \in \mathbb{P}V^* : \ker \sigma_P(\xi) \neq 0\}$$

The complex characteristic variety is

$$\Xi^{\mathbb{C}} = \{[\xi] \in \mathbb{CP}(V \otimes \mathbb{C}^*) : \ker \sigma_P(\xi) \neq 0\}$$

We can picture a covector $\xi \in V^*$ as a linear wave: a function constant in the directions perpendicular to ξ, with differential proportional to ξ (as in figure 1). A solution u of $Pu = 0$ can be multiplied by a wave $e^{i\lambda f}$, which has ripples travelling in the direction df, and this will still solve the equation to highest order precisely when

$$\sigma_P(df) \cdot u = 0$$

So for general choice of u it may not be possible to have nonconstant f, i.e. not every solution will admit such perturbations.

The first order equation

$$\sigma_P(df) \cdot u = 0$$

is called the *eikonal equation*. The word *eikonal* means having to do with images, since the term arises from geometric optics. In geometric optics, an example of such a function f is the "optical length" of a light ray passing from object to image.

3.1. **Example: the wave equation.** The wave operator

$$Pu = \frac{\partial^2 u}{\partial t^2} - c^2 \sum_j \frac{\partial^2 u}{\partial x_j^2}$$

(with c a constant) has symbol

$$\sigma_P(\xi) = \omega^2 - c^2 \sum_j k_j^2$$

where we write

$$\xi = (\omega, k)\,.$$

The characteristic variety is called the light cone. Every solution of the wave equation can be perturbed infinitesimally by any $e^{i\lambda f}$ whenever the eikonal equation

$$(\frac{\partial f}{\partial t})^2 = c^2 \sum_j (\frac{\partial f}{\partial x_j})^2$$

i.e. the differential is on the light cone. For example, the functions

$$f(t,x) = e^{i(\omega t - kx)}$$

solves the eikonal equation exactly, not just to leading order, as long as (ω, k) lies on the light cone. View from the perspective of space, these solutions look like travelling waves. But

3.2. **Example: a vector field.** For a vector field X the equation $\mathcal{L}_X u = 0$ has as characteristic variety the hyperplanes perpendicular to X. Thus solutions u are most easily perturbed by functions whose differential is perpendicular to the vector field, i.e. functions constant along the integral curves of the vector field. Of course these are actual perturbations of solutions, not just infinitesimal ones.

3.3. **Linear equations with variable coefficients.** Simply freeze the variable coefficients at a point. Then you can calculate the characteristic variety at that point as above.

3.4. **Vector bundles.** Given a differential operator P of order d

$$\begin{tikzcd} J^d V_0 \arrow[r, "P"] \arrow[dr] & V_1 \\ & M \end{tikzcd}$$

on vector bundles $V_0 \to M$ and $V_1 \to M$, (with J^d meaning the bundle of d jets) define $\sigma_P \in \Omega^0 \left(\operatorname{Sym}^k (T^* M) \otimes V_0^* \otimes V_1\right)$ by

$$e^{-i\lambda f} P \left[e^{i\lambda f}\right] = (i\lambda)^d \, \sigma_P\,(df) + O\left(|\lambda|^{d-1}\right)$$

for $f : M \to \mathbb{R}$ and $\lambda \in \mathbb{R}$.

3.5. **The tableau.** Let us see how to pass between the two definitions of characteristic variety. First, any system of differential equations can be written (by adding variables to represent derivatives) in terms of only first derivatives. Next, we can linearize equations about solutions. Finally, we can linearize about jets of solutions, to obtain linear constant coefficient equations, and we can try to approximate those by dropping their lower order terms. Therefore, we can approximate any system of partial differential equations by a first order linear system with constant coefficients, and no zero order terms. Consider a linear subspace $A \subset \operatorname{Lin}(V, W)$, a linear family of linear maps. Think of our differential equation as

$$u'(x) \in A$$

In effect we have found that this sort of equation can be used to approximate any differential equation, although of course approximating it poorly. We can define $Z = \operatorname{Lin}(V, W)/A$ and define the projection map

$$p : \operatorname{Lin}(V, W) \to Z, \qquad M \mapsto M + A$$

and then our differential operator will be

$$P[u(x)] = p(u'(x)) = u'(x) + A \in Z$$

Calculating the symbol, we find

$$\sigma_P(\xi) \cdot w = p(w \otimes \xi) = w \otimes \xi + A \in Z$$

for $w \in W$. Therefore $\sigma_P(\xi)$ is not injective precisely when the rank one linear map $w \otimes \xi$ belongs to A. Therefore characteristics $[\xi] \in \Xi$ correspond to covectors $\xi \in V^*$ for which there is a rank one linear map $w \otimes \xi$ with kernel ξ, satisfying our system of partial differential equations.

Take linear coordinates $x^1, \dots x^n$ on V and $y^1, \dots, y^m$ on W, and let p^μ_i be the coordinates on $\operatorname{Lin}(V, W)$, determined by

$$p^\mu_i(\phi) = dy^\mu \circ \phi \cdot \frac{\partial}{\partial x^i}$$

Define

$$\begin{aligned}
\omega^i &= dx^i \\
\pi^\mu_i &= dp^\mu_i \\
\theta^\mu &= dy^\mu - p^\mu_i \cdot dx^i
\end{aligned}$$

and then the equations $\theta = 0, d\theta = -\pi \wedge \omega$ with $\pi \in A$ and $\omega^1 \wedge \dots \wedge \omega^n \neq 0$ hold exactly on the graphs of solutions (locally) of the system of partial differential equations. In other words, we want to solve $\theta = d\theta + \pi \wedge \omega = 0$ and force π to satisfy the relations $p\pi = 0 \in Z$. So these $p\pi$ are the relations among the π, i.e. a basis of $Z = \operatorname{Lin}(V, W)/A$ describes the relations among the π. Following the algorithm we initially outlined, each relation among the $\pi' s$ gives a row of the symbol matrix, and a relation $a^i_j \pi^j_i$ gives the row

$$\begin{pmatrix} a^i_1 \xi_i & a^i_2 \xi_i & a^i_3 \xi_i & \dots & a^i_n \xi_i \end{pmatrix}$$

which hits some $w \in W$ to give

$$a^i_j \xi_i w^j = a^i_j (w \otimes \xi)^j_i$$

We see that the symbol matrix is the map

$$w \in W \mapsto p(w \otimes \xi) \in Z$$

having as rows the relations among the $\pi's$. Consequently the characteristic variety according to our algorithm consists in finding the ξ for which the symbol matrix has nonzero kernel, i.e. so that there are nonzero $w \in W$ with $\xi \otimes w \in A$, or in other words

$$[\xi] \in \Xi \text{ iff } \sigma_P(\xi) = 0$$

Therefore Cartan's algorithm recovers the classical story for linear equations with constant coefficients, and we hope it is clear that Cartan's algorithm is invariant under approximating nonlinear differential equations, stored in a linear Pffafian system, by their linearization about an integral element.

A system of partial differential equations is called *elliptic* if the real characteristic variety of its linearization about any jet is empty. Ellipticity thus corresponds to refusal to allow high frequency perturbations in any direction. Intuitively, a kind of smoothness and refusal to have small bumps or creases.

4. Symplectic aspects of the characteristic variety

Suppose that $\mathcal{I}$ is a differential ideal $\mathcal{I} \subset \Omega^* M$ on a manifold M, and that $\mathcal{I}$ has no Cauchy characteristics. Let $N \subset M$ be an integral manifold. All of this data can be merely C^1 I believe. Then the characteristic variety Ξ of $\mathcal{I}$ assigns an algebraic variety

$$\Xi_E \subset \mathbb{P}E^*$$

for each integral element E, defined as follows. We take for $E \subset T_x M$ any k dimensional vector subspace the polar space

$$H(E) = \{v \in T_x M : \omega_x(v, u_1, \ldots, u_k) = 0, \text{ for all } \omega \in \mathcal{I}, \ u_1, \ldots, u_k \in E\}$$

So then E is an integral element precisely if $E \subset H(E)$. Then we can say that for E an integral element

$$\Xi_E = \{[\xi] \in \mathbb{P}E^* : H(\xi = 0) \neq E\}$$

This gives rise on any integral manifold N to the induced characteristic variety

$$\Xi_x N := \Xi_{T_x N} \subset \mathbb{P}T_x^* N$$

This ΞN is a subset of $\mathbb{P}T^* N$, i.e. a set of hyperplanes. The set of nonzero covectors in $T^* N$ which project to the smooth points of ΞN I will write as $\tilde{\Xi} N$.

Theorem 1. *(O. Gabber) If $\mathcal{I}$ is involutive near N, and both $\mathcal{I}$ and N are real analytic, then $\tilde{\Xi} N$ is a coisotropic submanifold of $T^* N$, and is a fiber bundle over N.*

Recall that in a symplectic vector space V with symplectic form ω, the skew orthogonal of a vector subspace W is the space

$$W^{\angle} = \{v \in V \ : \omega(v, w) = 0, \text{ for all } w \in W\}$$

and a subspace is called *isotropic* if it is contained in its skew orthogonal, and *coisotropic* if contains its skew orthogonal, and *Lagrangian* if it is equal to its skew orthogonal. Similarly a submanifold of a symplectic manifold is isotropic if all of its tangent spaces are, etc.

The cotangent bundle $T^* M$ of a manifold M is equipped with the symplectic form ω defined by taking $\pi : T^* M \to M$ the obvious projection, and defining θ a 1-form on $T^* M$ by $\theta(\xi) = \xi \circ \pi'$ for $\xi \in T^* M$, and then letting $\omega = -d\theta$.

If we have coordinates $q^1, \dots q^n$ on an open set of M, they induce coordinates $q^1, \dots, q^n, p_1, \dots, p_n$ on T^*M by $q^j(\xi) = q^j \circ \pi(\xi)$ and

$$p_j(\xi) \cdot dq^j = \xi$$

Then $\theta = p \cdot dq$ and $\omega = dq \wedge dp$.

It follows from coisotropy that there is a canonical foliation of $\tilde{\Xi}N$ by isotropic submanifolds, whose dimension is k where the dimension of $\tilde{\Xi}N$ is $2n-k$. These are simply the maximal dimensional isotropic submanifolds of $\tilde{\Xi}N$, and their tangent spaces are the skew orthogonal spaces to the tangent spaces of $\tilde{\Xi}N$. Moreover it is elementary to see that the leaves project locally diffeomorphically to N, and that they strike each fiber of $\tilde{\Xi}N \to N$ transversely in discrete points. Also the leaf through a point $\xi \in \tilde{\Xi}N$ projects to the same submanifold of N as the leaf through $\lambda\xi$, for $0 \neq \lambda \in \mathbb{R}$. Because the characteristic variety is cut out by equations which are homogeneous polynomials, we find by Euler's formula that the canonical 1-form on the cotangent bundle T^*N is perpendicular to the isotropic leaves of the characteristic variety. We see this as follows: for $H(q,p)$ homogeneous of degree N in p, calculate that along the flow of Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}$$
$$\dot{p} = -\frac{\partial H}{\partial q}$$

we have

$$p \cdot dq = N \cdot H(q,p) \cdot dt$$

so on $H = 0$ we have $p \cdot dq$ perpendicular to the integral curves of H. Since the isotropic leaves of the characteristic variety are the leaves of the flows of all of the Hamiltonians whose vanishing cuts out the characteristic variety, the result is clear.

Another characterization of coisotropy is that the (real or complex valued) functions which vanish on a coisotropic submanifold are closed under Poisson bracket. Yet another is that the eikonal equation

$$[df] \in \mathbb{P}T^*M$$

is involutive in the sense of Cartan, for $f : M \to \mathbb{R}$, i.e. high frequency perturbations exist.

The case of the wave equation makes it apparent that the characteristic variety is a kind of light cone, and the isotropic leaves like paths of particles of light. Thus the construction of the characteristic variety reminds us of a wave to particle limit.

5. Contact geometry and the characteristic variety

It is more natural to view the characteristic variety as a submanifold of the contact manifold $\mathbb{P}T^*M$ of hyperplanes in the tangent spaces of M, for M an integral manifold. The symplectic perspective is artificial.

Given a manifold Z and a field of hyperplanes ξ on Z, we write the field locally as $\xi = \alpha^\perp$ where α is a 1-form. Then the rank of the field of hyperplanes at a point $z \in Z$ is r if

$$\alpha \wedge d\alpha^r \neq 0$$
$$\alpha \wedge d\alpha^{r+1} = 0$$

This occurs exactly when the null space of $d\alpha$ on our hyperplane $\xi(z)$ is of dimension $n-1-2r$, with $n = \dim Z$. It is easy to see that rank is independent of the choice of α. If the rank is full, i.e. $n = 2r+1$, then we call our hyperplane field a contact structure. A submanifold $X \subset Z$ is called nondegenerate if $TX \cap \xi$ is a hyperplane field on X.

The projectivized cotangent bundle $\mathbb{P}T^*M = T^*M/\mathbb{R}^\times$ (the bundle of hyperplanes in tangent spaces) of a manifold M is equipped with a contact structure ξ defined as follows: let $\pi : \mathbb{P}T^*M \to M$ be the obvious projection. Then

$$\xi_\Pi = \pi^{-1}\Pi \qquad \Pi \in \mathbb{P}T^*M$$

In local coordinates $q^1, \dots, q^n$ on an open set of M, we can build coordinates on $\mathbb{P}T^*M$ by setting $q^j = \pi^* q^j$ and taking homogeneous coordinates

$$[p_1, \dots, p_n]\,(\Pi)$$

defined by

$$p_j \, dq^j = 0 \text{ on } \Pi$$

Then our hyperplane field on $\mathbb{P}T^*M$ is $p\,dq = 0$. On the open set on which $p_1 \neq 0$, we can scale it to have $p_1 = 1$, and then our contact structure is

$$\xi = \left\{ dq^1 + \sum_{i>1} p_i \, dq^i = 0 \right\}.$$

Lemma 1 (Pfaff). *The rank of a hyperplane field ξ on a manifold Z is constant r iff there are coordinates z, q^i, p_i, x^μ locally on Z so that $\xi = \alpha^\perp$ with*

$$\alpha = dz + p_i \, dq^i$$

with $z \in \mathbb{R}, q \in \mathbb{R}^r, p \in \mathbb{R}^r, x \in \mathbb{R}^{n-(2r+1)}$.

The submanifolds on which $\alpha = d\alpha = 0$, cut out in Pfaff's coordinates by z, q, p constants, are called the *bicharacteristics* or *Cauchy characteristics.* If the bicharacteristic foliation is a fiber bundle, then we see from Pfaff's coordinates that its base space is a contact manifold. Bicharacteristics are precisely the submanifolds swept out by the symmetries of the hyperplane field which are tangent to the hyperplanes.

Think of the bicharacteristics by analogy with the case of the wave equation. (We will work out this case in detail below.) Each bicharacteristic is the response of a solution to a small, very localized (hence very high frequency) perturbation. For the wave equation, this is the emission of a single particle of light. It does not produce an entire expanding sphere of light, since it is localized in position and in momentum.

Given Z a contact manifold, with ξ its hyperplane field, and a submanifold $X \subset Z$, we can define $\xi_X = \xi \cap TX$, and define $X^\angle$ to be

$$X^\angle_z = \{ v \in \xi : d\alpha(v, w) = 0, \text{ for all } w \in \xi_X \}$$

and call X

$$\text{isotropic if } X^\angle \supset \xi_X$$
$$\text{coisotropic if } X^\angle \subset \xi_X$$
$$\text{rank 0 if } X^\angle = \xi_X$$

We find

$$2r + 1 = \dim X^{\angle} + \dim X$$

Bicharacteristics of X are submanifolds tangent to $X^{\angle}$ and to X. For generic X they won't be very large, but coisotropy forces them to be as large as possible.

A submanifold $X^r \subset Z^{2r+1}$ of a contact manifold is called *Legendre* if it is tangent to the hyperplane field. If $X \subset Y \subset Z$ with Y coisotropic, and X rank 0, we find that X is foliated by bicharacteristics, essentially because if we added the bicharacteristics to X we would not add any rank to X. We find moreover that if $\check{Y} = Y/\text{bichars}$ is the quotient space, and if it is a manifold, then a rank 0 submanifold $X \subset Y$ is precisely the preimage in Y of a rank 0 submanifold of $\bar{Y}$. Moreover rank 0 submanifolds in $\bar{Y}$ are simply arbitrary one parameter families of Legendre submanifolds.

Lemma 2. *Suppose that $X \subset T^*N - 0$ is a coisotropic submanifold, invariant under scaling the fibers, and $Y \subset \mathbb{P}T^*N$ is the quotient of X by scaling, and suppose that Y is a smooth manifold. Then X is coisotropic in the sense of symplectic geometry iff Y is in the sense of contact geometry.*

In the case of the characteristic variety, it is not hard to see that because the constraints on the characteristic variety are in the momentum variables (i.e. in the fibers of the contangent bundle) the bicharacteristics will actually strike the fibers of $\Xi \to N$ transversely, in discrete points. So they project diffeomorphically to submanifolds of N, and this is how one might think of them.

Now let us reconsider the problem of perturbing solutions of partial differential equations by high frequency waves. We want first to solve the equation

$$[df] \in \Xi \subset \mathbb{P}T^*N$$

for $f : N \to \mathbb{R}$ a function on an integral manifold. This is a system of partial differential equations on f, which are involutive in the sense of Cartan, so local solutions exist in the real analytic category. But we can do better: the graph of $[df]$ in $\mathbb{P}T^*N$ is a rank 0 submanifold of Ξ, a coisotropic manifold. Conversely, the equation $[df] \in \Xi$ has as local solutions exactly the rank 0 submanifolds on Ξ which satisfy the nondegeneracy hypothesis that they must project diffeomorphically down to N. The level sets of f correspond to the Legendre submanifolds of the graph of $[df]$. To construct $[df]$, we need to write down Legendre submanifolds of $\check{\Xi} = \Xi/\text{bichars}$.

5.1. **Example: the wave equation.** Consider the equation

$$\partial_t^2 u = c^2 \sum_i \partial_{x_i}^2 u$$

in coordinates (t, x) on $\mathbb{R}^{n+1}$, for $u : \mathbb{R}^{n+1} \to \mathbb{R}$. We take coordinates $(t, x, [\omega, k])$ on $\mathbb{P}T^*\mathbb{R}^{n+1} = \mathbb{R}^{n+1} \times \mathbb{P}^n$, with the square brackets around $[\omega, k]$ reminding us that these are homogeneous coordinates on projective space.

Note that a linear equation has characteristic variety independent of which solution we work on, while a nonlinear equation can have different characteristics on different solutions.

The characteristic variety is $\omega^2 = c^2|k|^2$. Let us fix our attention on the real points. At all real points of the characteristic variety, $\omega \neq 0$. So we can work in

affine coordinates in which $\omega = c$. The contact structure is $\xi = \{\alpha^{\perp}\}$ where

$$\alpha = \omega\, dt - k\, dx$$

and $d\alpha = d\omega \wedge dt - dk \wedge dx$. The funny minus sign in front of the k is really due to treating a covector (ω, σ) as a vector $(\omega, -k)$ using the Lorentz metric. The bicharacteristics are curves, because the dimension of a bicharacteristic is the codimension of the characteristic variety. The bicharacteristics are exactly the curves in the characteristic variety whose tangent vectors v satisfy $\alpha(v) = v \lrcorner\, d\alpha = 0$. To belong to the characteristic variety, they also satisfy $|k| = 1$ in our affine coordinates. We find that they are the curves on which $|k| = 1$, $dx = ck\, dt$, and $dk = 0$. We see that $dt \neq 0$ on all bicharacteristics, so we can integrate to find that the bicharacteristics are the straight lines

$$\begin{aligned} x &= \check{x} + c\check{k}(t - \check{t}) \\ k &= \check{k} \end{aligned}$$

for any constants $\check{t}, \check{x}, \check{k}$ giving our initial values. Picking a value for $\check{t}$, say $\check{t} = 0$, we can parameterize the space of bicharacteristics $\check{\Xi}$ by coordinates $(\check{x}, \check{k})$ with $|\check{k}| = 1$, i.e. by $\mathbb{R}^n \times S^{n-1}$. The map to bicharacteristics is

$$\begin{gathered} \Xi \to \check{\Xi} \\ (t, x, k) \mapsto (x - ckt, k) = (\check{x}, \check{k}). \end{gathered}$$

The 1-form $-\check{k}\, d\check{x}$ pulls back to

$$k\, d(x - ckt) = c\, dt - k\, dx = \alpha$$

the contact form restricted to Ξ. Therefore this 1-form $-\check{k}\, d\check{x}$ is perpendicular to the contact structure on the space of bicharacteristics. Essentially, the space of bicharacteristics is the projectivized cotangent bundle of $\mathbb{R}^n$, but not quite: it is the unit sphere bundle in the cotangent bundle.

The generic Legendre submanifold of $\check{\Xi}$ is therefore given by taking a hypersurface in $\mathbb{R}^n$ and forming its field of unit normal vectors (with either orientation). A rank 0 submanifold will emerge from a one parameter family of hypersurfaces in $\mathbb{R}^n$.

Let us reconsider more carefully how a perturbation of a solution of the wave equation corresponds to a one parameter of hypersurfaces in $\mathbb{R}^n$. Let us start with a function $f(t, x)$ satisfying $[df] \in \Xi$, i.e. a perturbation of a solution. The condition $[df] \in \Xi$ is

$$\left(\frac{\partial f}{\partial t}\right)^2 = c^2 \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2$$

or

$$\partial_t f = \pm c |\partial_x f|.$$

This implies that f is constant along the bicharacteristic curves

$$x(t) = \check{x} + c\check{k}t$$

where $\check{x} \in \mathbb{R}^n$ is arbitrary and

$$\check{k} = \pm \left.\frac{\partial_x f}{|\partial_x f|}\right|_{t=0, x=\check{x}}$$

Therefore, f is a function on the hypersurface $t = 0$, extended to $\mathbb{R}^{n+1}$ by making it constant on the bicharacteristics. The one parameter family of hypersurfaces is just the family of level sets of f on $\mathbb{R}^n$.

More generally, to form our perturbation of an integral manifold N of an exterior differential system, we will pick a submanifold of N, say Σ, whose codimension matches the codimension of the characteristic variety $\Xi \subset \mathbb{P}T^*N$. Then we pick on Σ any function $F : \Sigma \to \mathbb{R}$ and a choice of hyperplane $[\xi] \in \Xi$ from the characteristic variety, so that ξ restricts to the tangent spaces of Σ to be a multiple of dF, and so that the bicharacteristic through $[\xi]$ in Ξ projects to N to be transverse to Σ. This choice of $[\xi]$ maps Σ into Ξ. Then we can extend F to be constant on the bicharacteristics through Σ in Ξ. As long as the submanifold swept out in Ξ by those bicharacteristics projects down to N diffeomorphically, we can keep extending the function F. The resulting function on N will be our perturbation.

We need coisotropy to ensure that the bicharacteristics have large enough dimension to enable this recipe to work.

6. Complex characteristics and complex structures

Take N^{2n} a real manifold. An almost complex structure on N can be defined in two ways: first, as a vector bundle map

$$\begin{tikzcd} TN \arrow[rr, "J"] \arrow[dr] & & TN \arrow[dl] \\ & N & \end{tikzcd}$$

so that $J^2 = -1$. Secondly, we can take such J, consider its eigenvalues, and see that there must be a complex vector subbundle

$$V \subset TN \otimes \mathbb{C}$$

of complex rank n on which J acts as $\sqrt{-1}$. Moreover,

$$V \cap TN = 0$$

Conversely, given a complex subbundle

$$V \subset TN \otimes \mathbb{C}$$

of complex dimension n with

$$V \cap TN = 0$$

we can define J by setting J to act as $\sqrt{-1}$ on V and as $-\sqrt{-1}$ on $\bar{V}$ (the complex conjugates). Thus V determines J. Another way to describe an almost complex structure is to describe the subspace $W \subset T^*N \otimes \mathbb{C}$ of its $(1, 0)$ forms. These are the complex valued 1-forms ξ on TN which satisfy

$$\xi(Jv) = \sqrt{-1}\xi(v)$$

for all $v \in TN$. If we pick any subspace

$$W \subset T^*N \otimes \mathbb{C}$$

which has n complex dimensions, satisfying

$$W \cap T^*N = 0$$

then we can define J at a point $p \in N$ by picking any complex basis

$$\zeta_1, \dots, \zeta_n \in W_p$$

and writing

$$\zeta = (\zeta_1, \dots, \zeta_n) : T_p N \to \mathbb{C}^n$$

and letting J_p be the map

$$J_p = \zeta^{-1} \sqrt{-1} \zeta.$$

An almost complex structure J is called integrable if the associated subbundle V is locally spanned by the holomorphic vector fields associated to a complex structure on N. This complex structure is unique, if it exists.

Theorem 2 (Newlander–Nirenberg)**.** *An almost complex structure is integrable precisely when any pair of complex vector fields X, Y which are sections of $V \to N$ have bracket $[X, Y]$ also a section of $V \to N$.*

Corollary 1. *An almost complex structure*

$$W \subset T^*N \otimes \mathbb{C}$$

is integrable precisely when any local section

$$\zeta^1, \dots, \zeta^n$$

of W satisfies

$$d\zeta^\mu = 0 \mod \zeta^1, \dots, \zeta^n$$

Note that $\bar{\zeta}^1$ is not zero modulo ζ^1.

Proposition 1. *If N^{2n} is a C^1 manifold with a C^1 almost complex structure*

$$W \subset T^*N \otimes \mathbb{C}$$

then W is a coisotropic submanifold of $T^*N \otimes \mathbb{C}$ precisely if W is integrable as an almost complex structure.

Proof. Take ζ^μ spanning W at each point of some open set. Write these as

$$\zeta^\mu = \xi^\mu + \sqrt{-1}\eta^\mu$$

where ξ^μ, η^μ are real valued 1-forms. Then let

$$X_\mu, Y_\mu$$

be real vector fields dual to ξ^μ and η^μ. The bundle V is locally spanned by the complex vector fields

$$Z_\mu = X_\mu - \sqrt{-1} Y_\mu$$

while the complex conjugate is spanned by

$$Z_{\bar{\mu}} = X_\mu + \sqrt{-1} Y_\mu$$

Suppose

$$[Z_\mu, Z_\nu] = a^\sigma_{\mu\nu} Z_\sigma + a^{\bar{\sigma}}_{\mu\nu} Z_{\bar{\sigma}}$$

for some complex valued functions

$$a^\sigma_{\mu\nu}, a^{\bar{\sigma}}_{\mu\nu}$$

Integrability is precisely the condition

$$a^{\bar{\sigma}}_{\mu\nu} = 0$$

Similarly

$$[Z_{\bar\mu}, Z_{\bar\nu}] = a^{\sigma}_{\bar\mu\bar\nu} Z_\sigma + a^{\bar\sigma}_{\bar\mu\bar\nu} Z_{\bar\sigma}$$

and moreover we have the relations

$$\begin{aligned} a^{\sigma}_{\bar\mu\bar\nu} &= \overline{a^{\bar\sigma}_{\mu\nu}} \\ a^{\bar\sigma}_{\bar\mu\bar\nu} &= \overline{a^{\sigma}_{\mu\nu}} \end{aligned}$$

Define the functions

$$h_\mu, h_{\bar\mu} : T^*N \otimes \mathbb{C} \to \mathbb{C}$$

by

$$\begin{aligned} h_\mu(\xi) &= \xi(Z_\mu) \\ h_{\bar\mu}(\xi) &= \xi(Z_{\bar\mu}) \end{aligned}$$

Then the submanifold

$$W \subset T^*N \otimes \mathbb{C}$$

is cut out by the equations $h_{\bar\mu} = 0$. We calculate the Poisson brackets

$$\begin{aligned} \{h_{\bar\mu}, h_{\bar\nu}\} &= a^{\sigma}_{\bar\mu\bar\nu} h_\sigma + a^{\bar\sigma}_{\bar\mu\bar\nu} h_{\bar\sigma} \\ &= \overline{a^{\bar\sigma}_{\mu\nu}} h_\sigma + \overline{a^{\sigma}_{\mu\nu}} h_{\bar\sigma} \end{aligned}$$

so that closure under Poisson bracket implies integrability. Moreover one can easily see that integrability implies closure under Poisson bracket. ☐

Corollary 2. *If the complex characteristic variety of an involutive exterior differential system contains as a component a complex linear subspace of the same real dimension as the dimension of the integral elements, and this subspace has no real points, then the integral manifolds are canonically equipped with the local structure of complex manifolds. Oriented integral manifolds are complex manifolds.*

7. Other stuff

7.1. **Singular points.** The singular locus of the characteristic variety is not always involutive, so for instance if the characteristic variety is smooth except for an isolated singular point, the singular point is not necessarily going to determine a foliation of an integral manifold.

7.2. **Trees.** Given a submanifold in a symplectic manifold, we can take A to be the algebra of smooth functions vanishing on it. The submanifold is coisotropic precisely when the algebra is closed under Poisson bracket. Since all coisotropic submanifolds are locally symplectomorphic, this isn't interesting. But in the special case of the characteristic variety $\tilde\Xi \subset T^*N$ of an integral manifold of an exterior differential system, we can take A to be the algebra of functions vanishing on $\tilde\Xi$ which are homogeneous polynomials in the fibers of T^*N. Then these are also closed under Poisson bracket, again by coisotropy. We have also the derived series

$$A^{(1)} = \{A, A\}, \quad A^{(2)} = \left\{A^{(1)}, A^{(1)}\right\}, \quad \dots$$

and the lower central series

$$A_1 = \{A, A\}, \quad A_2 = \{A_1, A\}, \quad \dots$$

If A is generated by functions which are of degree d in the fibers of T^*N, then $A^{(k)}$ is generated by functions of degree $2^k d - 2^k + 1$ and A_k by functions of degree $(k+1)d - k$. These will define varieties containing the characteristic variety,

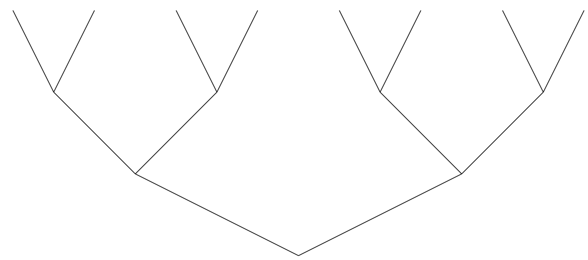

FIGURE 2. The tree for the derived series looks like this

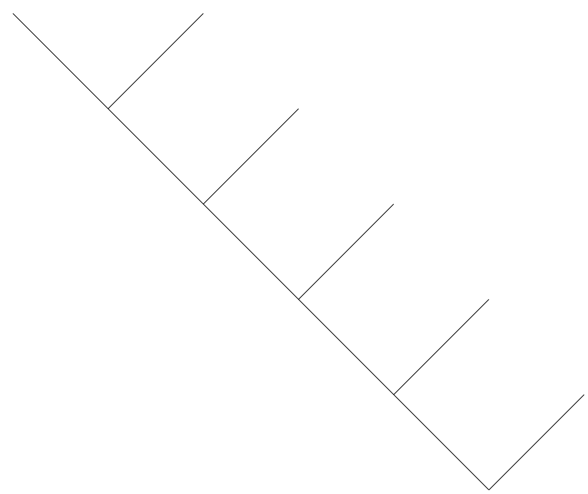

FIGURE 3. The tree for the lower central series looks like this

also coisotropic, with smaller dimensional isotropic submanifolds contained in the isotropic submanifolds of A, in a nested sequence.

Let $\mathfrak{g}$ be a Lie algebra. Suppose that we take any finite trivalent tree, with one root chosen, and interpret it as the subalgebra of $\mathfrak{g}$ given by the following recipe. First put a copy of $\mathfrak{g}$ on all of the roots except the chosen one. Now whenever faced by a vertex v with algebras $\mathfrak{h}_1, \mathfrak{h}_2$ at two of its neighbouring vertices, and nothing in v itself, we place at vertex v the algebra $[\mathfrak{h}_1, \mathfrak{h}_2]$ (i.e. the Lie subalgebra generated by the elements expressible as brackets of something from $\mathfrak{h}_1$ with something from $\mathfrak{h}_2$). Repeating this process, finitely many times, all vertices get algebras on them, and the chosen root now has an algebra from the trivalent vertex it is attached to. We thus obtain from any finite trivalent tree with chosen root a subalgebra of a given Lie algebra. This is probably in the literature somewhere. Call it a tree derivation of a Lie algebra. In my case, it gives rise to a collection of tree derivations of the characteristic variety, all of them coisotropic. For some overdetermined systems, this might be interesting.

References

[1] Bryant, R. L., Chern, S. S., Gardner, R. B., Goldschmidt, H., and Griffiths, P. A., *Exterior Differential Systems*, Springer-Verlag, New York, 1991.

[2] E. Cartan, *Sur les systèmes en involution d'équations aux dérivées partielles du second ordre a une fonction inconnue de trois variables independantes*, Bull. Soc. Math. France, t.39, 1911, pp. 352-443.

Duke University, Durham, North Carolina
E-mail address: mckay@math.duke.edu