

WHEREFORE COTANGENT SPACE?

MICHAEL VANVALKENBURGH

At some point we all must have asked ourselves: Why is “cotangent space” (or, in physics, “position–momentum space”) the correct setting for PDE? After all, in high school physics, momentum is just mass times velocity, and when we draw pictures of arrows it’s easier to think of them as vectors rather than the more mysterious covectors. So, in particular, why don’t we study PDE in the setting of tangent space? The point of this note is to answer that question.

Say we are given a linear partial differential operator

$$P = \sum_{\alpha} a_{\alpha}(x) D_x^{\alpha},$$

where $D_{x_j} := \frac{1}{i} \frac{\partial}{\partial x_j}$. If P were to have constant coefficients, $a_{\alpha}(x) \equiv a_{\alpha}$, then the Fourier transform \mathcal{F} would conjugate P into a simple multiplication operator:

$$((\mathcal{F} \circ P)u)(\xi) = p(\xi) \mathcal{F}u(\xi),$$

where the function

$$p(\xi) = \sum_{\alpha} a_{\alpha} \xi^{\alpha}$$

is the so-called “symbol” of the operator. Hence, on the Fourier transform side, the operator P is easy to understand. When P has variable coefficients, the concept of “symbol” is more subtle. Since x and D_x do not commute, one can turn a symbol $p(x, \xi)$ into a PDO in more than one way, depending on how one wants to order the x ’s and D_x ’s.

Now the question is: Why do we consider $p(x, \xi)$ to be a function on cotangent space, $T^*(\mathbb{R}^n)$? Since P looks like a function of partial derivatives, and since tangent vectors are commonly identified with partial derivatives, one might guess that the symbol is a function on tangent space (or, more accurately, on the tangent bundle). However, this is not the case. But if we make our reasoning more precise, we can guess the right answer. A function on the tangent bundle doesn’t really act on (x, μ) , where μ is a tangent vector. What the function does is act on (x, μ) , where μ are the *coefficients* of a tangent vector with respect to the basis $\{\frac{\partial}{\partial x_j}\}_{j=1}^n$ adapted to the choice of coordinates x . And what a function on the cotangent bundle does is act on (x, ξ) , where ξ are the coefficients of a cotangent vector with respect to the basis $\{dx_j\}_{j=1}^n$ adapted to the coordinates x . So why do we think of a symbol $p(x, \xi)$ as having ξ be the coefficients of cotangent vectors?

Date: March 7, 2009.

Whether or not the coordinates $(x, \xi) \in \mathbb{R}^{2n}$ are in $T\mathbb{R}^n$ or $T^*\mathbb{R}^n$ depends on how the ξ 's change when one changes the coordinates x by a diffeomorphism

$$\kappa : x \mapsto \kappa(x) = y.$$

By the chain rule of elementary calculus, tangent coordinates (x, μ) change as follows:

$$\begin{aligned} \sum_{j=1}^n \mu_j \frac{\partial}{\partial x_j} &= \sum_{j,k=1}^n \mu_j \frac{\partial y_k}{\partial x_j} \frac{\partial}{\partial y_k} \\ &= \sum_{k=1}^n \nu_k \frac{\partial}{\partial y_k}, \end{aligned}$$

where

$$\nu = \left(\frac{\partial \kappa}{\partial x} \right) \mu.$$

Here we write

$$\left(\frac{\partial \kappa}{\partial x} \right) = \begin{pmatrix} \frac{\partial \kappa_1}{\partial x_1} & \cdots & \frac{\partial \kappa_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \kappa_n}{\partial x_1} & \cdots & \frac{\partial \kappa_n}{\partial x_n} \end{pmatrix}.$$

That is, tangent coordinates change as

$$(x, \mu) \mapsto \left(\kappa(x), \left(\frac{\partial \kappa}{\partial x} \right) \mu \right) = (y, \nu).$$

On the other hand, cotangent coordinates change as follows:

$$\begin{aligned} \sum_{j=1}^n \xi_j dx_j &= \sum_{j,k=1}^n \xi_j \frac{\partial x_j}{\partial y_k} dy_k \\ &= \sum_{k=1}^n \eta_k dy_k, \end{aligned}$$

where $\eta_k = \sum_{j=1}^n \frac{\partial x_j}{\partial y_k} \xi_j$. That is, with the notation above,

$$(x, \xi) \mapsto \left(\kappa(x), \left(\left(\frac{\partial \kappa}{\partial x} \right)^{-1} \right)^T \xi \right) = (y, \eta).$$

Now we can check a very simple example of a symbol, to at least rule out one possibility. We take the one dimensional operator $P = D_x$ with symbol $p(\xi) = \xi$, and we consider the change of variables $y = \kappa(x) = 2x$. Then

$$D_x = 2D_y,$$

so the change of variables κ acts on the symbols of the operators as

$$(x, \xi) \mapsto (y, \eta) = \left(2x, \frac{1}{2}\xi \right) = \left(\kappa(x), \left(\left(\frac{\partial \kappa}{\partial x} \right)^{-1} \right)^T \xi \right).$$

So at least we know that the symbol $p(x, \xi)$ does NOT live on the tangent bundle.

The general case takes place in the theory of pseudodifferential operators. We will just sketch the result, somewhat imprecisely; for details one may consult any book on pseudodifferential operators, for example, the books of Grigis and Sjöstrand. Let $\kappa : X \rightarrow \tilde{X}$ be a diffeomorphism between two open sets in \mathbb{R}^n , and let $\tilde{A} \in \Psi^m(\tilde{X})$ (i.e. be a pseudodifferential operator of order m on \tilde{X} ; note that I haven't even defined this). Let κ act on $u \in C^\infty(\tilde{X})$ in the obvious way: $\kappa^*u := u \circ \kappa \in C^\infty(X)$. Then $A = \kappa^* \circ \tilde{A} \circ \kappa^{*-1} \in \Psi^m(X)$, and we have the following relation between the principal symbols:

$$a(x, \xi) = \tilde{a} \left(\kappa(x), \left(\left(\frac{\partial \kappa}{\partial x} \right)^{-1} \right)^T \xi \right).$$

This shows that the (principal) symbol is invariantly defined on the cotangent bundle, which is the ultimate reason why cotangent space is so important in PDE. Moreover, there is a rich correspondence between PDE and classical mechanics in the cotangent bundle (Hamiltonian mechanics). For one thing, the above result generalizes in a way. One may check that κ^* is a simple example of a Fourier integral operator (this is Exercise 1.13 in Grigis and Sjöstrand's book). If one replaces κ^* by a more general (but not completely general) Fourier integral operator, one gets the same type of transformation of the symbol, but with the lift of κ replaced by a more general symplectomorphism.

Let us conclude by briefly discussing some physics language that appears in the theory of pseudodifferential operators. Often ξ is called the momentum variable. For one thing, in Hamiltonian mechanics, the cotangent bundle has the natural interpretation of the position-momentum space. (Lagrangian mechanics, on the other hand, is classical mechanics in the tangent bundle, having the natural interpretation of the position-velocity space.) One may then ask about the role of the Fourier transform, which we alluded to in the beginning of this note. Since x has units of position, since ξ has units of mass times velocity, and since we can only take the exponential of a dimensionless quantity (in the definition of the Fourier transform), what happened to the missing units? The quantity $x\xi$ has units of mass \times velocity squared \times units of time. So we come out with units of energy \times units of time (the units of a quantity called *action*). This is precisely where Planck's constant comes in: \hbar has dimensions of energy \times time, which is why in quantum mechanics and in semiclassical analysis the Fourier transform is of the form

$$\mathcal{F}u(\xi) = \int e^{-ix\xi/\hbar} u(x) dx.$$

If one is using the Fourier transform in signal analysis, when the Fourier transform takes functions of time to functions of frequency, there is no need to correct the units, since frequency has units of $(\text{time})^{-1}$.

UCLA DEPARTMENT OF MATHEMATICS, LOS ANGELES, CA 90095-1555, USA
E-mail address: `mvanvalk@ucla.edu`