MATH 918 LECTURE NOTES, SPRING 2023

Lecture of January 24, 2023

1. Derivations

Our goal will be to consider derivatives algebraically.

The usual notion of derivative of a function is a rule that turns certain real-valued or complex-valued functions into other real-valued or complex-valued functions as follows: at a given point x, we take

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}.$$

This certainly gives us derivative functions on some rings, for example, the ring of infinitely-differentiable functions on $\mathbb{R}$:

$$\mathcal{C}^\infty(\mathbb{R}) \xrightarrow{\frac{d}{dx}} \mathcal{C}^\infty(\mathbb{R})$$

or the ring of *entire functions*, i.e., *holomorphic*, a.k.a. complex-differentiable, functions on the complex plane:

$$\mathrm{Holo}(\mathbb{C}) \xrightarrow{\frac{d}{dx}} \mathrm{Holo}(\mathbb{C}).$$

Neither of these is the sort of ring that we usually consider in commutative algebra. In particular, neither is Noetherian.

Using our familiar rules of differentiation, we might recall that the derivative of a polynomial is a polynomial, and the derivative of a rational function is a rational function. So, we get derivatives on much more manageable rings:

$$\mathbb{R}[x] \xrightarrow{\frac{d}{dx}} \mathbb{R}[x], \quad \mathbb{R}(x) \xrightarrow{\frac{d}{dx}} \mathbb{R}(x), \quad \mathbb{C}[x] \xrightarrow{\frac{d}{dx}} \mathbb{C}[x], \quad \mathbb{C}(x) \xrightarrow{\frac{d}{dx}} \mathbb{C}(x).$$

To unlock some of the applications of derivatives, we would like to be able to do this as much as possible over arbitrary rings. We might be optimistic about doing this for arbitrary polynomial rings at least, given the examples above. To do it, we certainly must get rid of this limit approach, since moving around in fields like $\mathbb{Q}$ or $\mathbb{F}_p$ we certainly will miss out on lots of limits. Of course, when we actually compute the derivative of a real or complex polynomial, we don't consider the limit definition anymore, but instead use rules of derivative. Namely, we have a sum rule, a scalar rule, a product rule, a quotient rule, and a power rule, and knowing all of these, we easily and limitlessly compute derivatives of any polynomial or rational function over $\mathbb{R}$ or $\mathbb{C}$. Since the quotient rule and power rule (mostly) follow from the product rule, we will hone in on the first three for our definition of algebraic notion of derivative.

So, our first approximation of the definition of *derivation*, our notion of derivative, is a function ∂ from a ring R to itself that satisfies a sum rule, a scalar rule, and a product rule:

- $\partial(r + s) = \partial(r) + \partial(s)$ for all $r, s \in R$,
- $\partial(cr) = c\partial(r)$ for all $r \in R$ and c "constant???",
- $\partial(rs) = r\partial(s) + s\partial(r)$ for all $r, s \in R$.

There is something we must change ("constant???") and something else less clear we can/should change. Let's be openminded. If R is a ring, let's let our constants be any reasonable set of elements of R: any subring A of R. But let's be even more openminded. Look at the right-hand sides above. To make sense of them we have to be able to add our outputs together and multiply them by ring elements, but we don't have to multiply them with each other. They don't have to live in R—they just have to live in an R-module.

Definition 1.1. Let R be a ring and M be an R-module. A *derivation* from R to M is a function $\partial : R \to M$ such that

- $\partial(r+s) = \partial(r) + \partial(s)$ for all $r, s \in R$,
- $\partial(rs) = r\partial(s) + s\partial(r)$ for all $r, s \in R$.

If R is an A-algebra, then ∂ is a *derivation over* A or an *A-linear derivation* if in addition

- $\partial(ar) = a\partial(r)$ for all $a \in A$ and $r \in R$.

Remark 1.2. Recall that R is an A-algebra means that R is equipped with a ring homomorphism $\phi : A \to R$. In this case, every R-module is also an A-module by restriction of scalars: $am := \phi(a)m$; i.e., for A to act on M, just view elements of A as elements of R via ϕ and do the same action. This is what's going on in the right-hand side above. We'll circle back to restriction of scalars soon.

1.0.1. *Examples of derivations.* Let's consider some examples of derivations to buy into this notion.

First, let's construct the "usual derivative" for a polynomial or a power series ring, and show it is a derivation.

Definition 1.3. Let A be a ring and $R = A[x]$ a polynomial ring. We define $\frac{d}{dx} : R \to R$ by the rule

$$\frac{d}{dx}(\sum_{j=0}^{d} a_j x^j) = \sum_{j=1}^{d} j a_j x^{j-1}.$$

Similarly, for a power series ring, $R = A[\![x]\!]$, we define $\frac{d}{dx} : R \to R$ by the rule

$$\frac{d}{dx}(\sum_{j=0}^{\infty} a_j x^j) = \sum_{j=1}^{\infty} j a_j x^{j-1}.$$

Lemma 1.4. *The functions $\frac{d}{dx} : A[x] \to A[x]$ and $\frac{d}{dx} : A[\![x]\!] \to A[\![x]\!]$ are A-linear derivations.*

Proof. In either case, we have a well-defined function returning an object of the same type. The formulas are the same in both cases, just allowing infinite formal sums for power series, so we'll deal with both simultaneously.

Take $r = \sum_{j=0} a_j x^j$, $s = \sum_{j=0} b_j x^j$, and c with $a_j, b_j, c \in A$. Then

$$\frac{d}{dx}(r+s) = \frac{d}{dx}(\sum_{j=0}(a_j+b_j)x^j) = \sum_{j=1} j(a_j+b_j)x^{j-1} = \frac{d}{dx}(r) + \frac{d}{dx}(s),$$

$$\frac{d}{dx}(cr) = \frac{d}{dx}(\sum_{j=0}(ca_j)x^j) = \sum_{j=1} j(ca_j)x^{j-1} = c\frac{d}{dx}(r),$$

and

$$\begin{aligned} r\frac{d}{dx}(s) + s\frac{d}{dx}(r) &= (\sum_{i=0} a_i x^i)(\sum_{j=1} j b_j x^{j-1}) + (\sum_{j=0} b_j x^j)(\sum_{i=1} i a_i x^{i-1}) \\ &= \sum_{k=1}\sum_{i+j=k} (a_i j b_j) x^{i+j-1} + \sum_{k=1}\sum_{i+j=k} (i a_i b_j) x^{i+j-1} \\ &= \sum_{k=1}\sum_{i+j=k} k a_i b_j x^{i+j-1} \\ &= \frac{d}{dx}(\sum_{k=0}\sum_{i+j=k} (a_i b_j) x^k) = \frac{d}{dx}(rs). \quad \square \end{aligned}$$

Note that we could have written the formula above as $\frac{d}{dx}(\sum_{j=0}^{d} a_j x^j) = \sum_{j=1}^{d} j a_j x^{j-1}$ as well: it looks like we have something illegal when $j = 0$, but the coefficient of zero tells us to ignore it.

Proposition 1.5. *Let A be a ring, $\{X_\lambda \mid \lambda \in \Lambda\}$, and $R = A[X_\lambda \mid \lambda \in \Lambda]$ be a polynomial ring. Then the partial derivatives $\frac{d}{dX_\lambda}$ given by the rule*

$$\frac{d}{dx_\lambda}(\sum_\alpha a_\alpha X^\alpha) = \sum_\alpha \alpha_\lambda a_\alpha X^{\alpha - e_\lambda}$$

where $\alpha \in \mathbb{N}^\Lambda$ is an exponent tuple and e_λ is the unit vector in the λ coordinate, are A-linear derivations. Similarly for the power series ring $R = A[\![X_\lambda \mid \lambda \in \Lambda]\!]$.

Proof. Consider R as $R'[X_\lambda]$, with $R' = A[X_\mu \mid \mu \in \Lambda \smallsetminus \{\lambda\}]$. Then $\frac{d}{dX_\lambda}$ is just the "usual derivative" in this polynomial ring over R', so it is an R'-linear derivation of R. But since $A \subseteq R'$, this is an A-linear derivation as well. □

So we can differentiate over any polynomial ring now, e.g., over $R = \mathbb{F}_2[x]$. Let's not neglect our original derivatives.

Example 1.6. The standard derivatives

$$\mathcal{C}^\infty(\mathbb{R}) \xrightarrow{\frac{d}{dx}} \mathcal{C}^\infty(\mathbb{R})$$

and

$$\mathrm{Holo}(\mathbb{C}) \xrightarrow{\frac{d}{dx}} \mathrm{Holo}(\mathbb{C})$$

are $\mathbb{R}$-linear and $\mathbb{C}$-linear derivations, respectively.

We haven't seen examples where we take derivations into "actual" modules yet. It turns out that this is a natural thing to do. In fact, examples like this appear in calculus before derivations back into the ring!

Example 1.7. Let's return to old-fashioned derivatives of $\mathbb{C}^\infty$ functions. Before we get derivatives of functions as functions, we start with the notion of derivative at a point, which should just be a number. Let's try to realize "derivative at $x = x_0$" for some real number x_0, which we'll write as $\frac{d}{dx}|_{x=x_0}$, as a derivation on $\mathcal{C}^\infty(\mathbb{R})$. The target should be $\mathbb{R}$:

$$\frac{d}{dx}|_{x=x_0} : \mathcal{C}^\infty(\mathbb{R}) \to \mathbb{R},$$

so we need to view $\mathbb{R}$ as a $\mathcal{C}^\infty(\mathbb{R})$-module. A very $x = x_0$ flavored way of doing so is by the rule

$$f \cdot c = f(x_0)c.$$

Another useful way of thinking about this module structure is as the quotient $\mathcal{C}^\infty(\mathbb{R})/\mathfrak{m}_{x_0}$, where $\mathfrak{m}_{x_0}$ is the maximal ideal consisting of functions with $f(x_0) = 0$. Indeed, the evaluation at 0 map

$$\mathrm{ev}_{x_0}\mathcal{C}^\infty(\mathbb{R}) \to \mathbb{R}$$

has kernel $\mathfrak{m}_{x_0}$ by definition, and if $\mathbb{R}$ has the module structure given above, this map is $\mathcal{C}^\infty(\mathbb{R})$-linear: if $f \in \mathcal{C}^\infty(\mathbb{R})$ and $c \in \mathbb{R}$, then $\mathrm{ev}_{x_0}(fc) = f(x_0)c = f \cdot c$. Of course, if x_0 changed, we would get a different module structure.

Back to our derivative. Take $f, g \in \mathcal{C}^\infty(\mathbb{R})$ and $c \in \mathbb{R}$. Note that this c is an element of $\mathbb{R} \subseteq \mathcal{C}^\infty(\mathbb{R})$ as opposed to $\mathbb{R} \cong \mathcal{C}^\infty(\mathbb{R})/\mathfrak{m}_{x_0}$. Then

$$\frac{d}{dx}|_{x=x_0}\ (f+g) = \frac{d}{dx}|_{x=x_0}\ f + \frac{d}{dx}|_{x=x_0}\ g$$

$$\frac{d}{dx}|_{x=x_0}\ cf = c\frac{d}{dx}|_{x=x_0}\ f$$

and by the product rule

$$\frac{d}{dx}|_{x=x_0}\ (fg) = f(x_0)(\frac{d}{dx}|_{x=x_0}\ g) + g(x_0)(\frac{d}{dx}|_{x=x_0}\ f) = f\cdot(\frac{d}{dx}|_{x=x_0}\ g) + g\cdot(\frac{d}{dx}|_{x=x_0}\ f).$$