MATH 918 LECTURE NOTES, SPRING 2023

Lecture of January 24, 2023

1. Derivations

1.1. **Definition and first examples.** 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.

1.1.1. *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 $R' = A[X_\mu \mid \mu \in \Lambda \setminus \{\lambda\}]$. 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. $\square$

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

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

Lecture of January 26, 2023

Example 1.8. Other natural uses of derivatives actually take values in modules rather than the ring itself. Let's consider $R = \mathcal{C}^\infty(\mathbb{R}^3)$, the ring of infinitely differentiable real valued functions from $\mathbb{R}^3$ to $\mathbb{R}$, with pointwise operations. One has a notion of gradient ∇ of a function:

$$f(x, y, z) \mapsto \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix}.$$

The output is a vector of three functions in R, so this is a function $\nabla : R \to R^3$. It follows from calculus that this is an $\mathbb{R}$-linear derivation.

Similarly, one sometimes talks about the *total derivative* of a function $f \in \mathcal{C}^\infty(\mathbb{R}^3)$ as

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz.$$

This rule $f \mapsto df$ is a derivation from R to a free R-module with basis dx, dy, dz.

Example 1.9. Let's try out a slightly more interesting ring. Let's consider $R = \mathbb{C}[x, y]/(x^2 - y^3)$ and try out $\frac{d}{dx}$ on this ring. Of course, this is a quotient ring, so if this means anything, it means apply this rule to an equivalence class and take the class of the result. But this is a problem, since $0 = x^2 + y^3$ and $\frac{d}{dx}(0) = 0 \neq 2x = \frac{d}{dx}(x^2 + y^3)$. So this derivation doesn't even make sense, and in hindsight, perhaps it looks a little bit silly to try. But we can actually get by with something surprisingly similar. Let's write $\frac{d}{dx}|_{(0,0)}$ for the rule

$$\frac{d}{dx}|_{(0,0)}(f) = \frac{d}{dx}(f)(0, 0);$$

i.e., partial derivative with respect to x at the origin. It is in fact well-defined: we have

$$\begin{aligned}\frac{d}{dx}|_{(0,0)}(f + (x^2 + y^3)g) &= \frac{d}{dx}|_{(0,0)}(f) + \frac{d}{dx}|_{(0,0)}((x^2 + y^3)g) \\ &= \frac{d}{dx}|_{(0,0)}(f) + g|_{(0,0)}\frac{d}{dx}|_{(0,0)}(x^2 + y^3) + (x^2 + y^3)|_{(0,0)}\frac{d}{dx}|_{(0,0)}(g) = \frac{d}{dx}|_{(0,0)}(f)\end{aligned}$$

and, along the same lines as previous examples, is a $\mathbb{C}$-linear derivation to $\mathbb{C}$, viewed as a module via the rule $f \cdot c = f(0, 0)c$.

Example 1.10. Let's end with a boring example. For any A-algebra R and any R-module M, the zero map is an A-linear derivation from R to M.

1.2. **Properties of derivations.** Let's collect some basic properties of derivations. The first includes the fact that constants go to zero.

Proposition 1.11. *Let $\partial : R \to M$ be a derivation.*

(1) $\partial(0) = \partial(1) = 0$,
(2) $\partial(-r) = -\partial(r)$,
(3) *The kernel of ∂ is a subring of R,*
(4) *For $A \subseteq R$, ∂ is A-linear if and only if $A \subseteq \ker(\partial)$.*

Proof. (1) $\partial(0) = \partial(0 + 0) = \partial(0) + \partial(0)$, and $\partial(1) = \partial(1 \cdot 1) = 1\partial(1) + 1\partial(1) = \partial(1) + \partial(1)$; in each case we cancel.

(2) $0 = \partial(r - r) = \partial(r) + \partial(-r)$, and move $\partial(r)$ to the other side.
(3) If $\partial(r) = \partial(s) = 0$, then $\partial(r - s) = \partial(r) - \partial(s) = 0$ and $\partial(rs) = r\partial(s) + s\partial(r) = 0$.
(4) If $A \subseteq \ker(\partial)$, $a \in A$, and $r \in R$, then $\partial(ar) = a\partial(r) + r\partial(a) = a\partial(r)$, so ∂ is A-linear; conversely, if $\partial(ar) = a\partial(r)$ for all $a \in A$ and $r \in R$, then $r\partial(a) = 0$ for all $a \in A$ and $r \in R$, and in particular $\partial(a) = 1\partial(a) = 0$. □

Remark 1.12. It follows that every derivation of R into M is $\mathbb{Z}$-linear since every derivation is linear over its kernel, and its kernel is a subring.

There are lots of ways to make derivations out of other derivations.

Proposition 1.13. *Let $\alpha, \beta : R \to M$ be derivations over A, $t \in R$, and $\gamma : M \to N$ be an R-module homomorphism, and $\phi : S \to R$ an A-algebra homomorphism.*

(1) *$\alpha + \beta : R \to M$ is a derivation over A,*
(2) *$t\alpha : R \to M$ is a derivation over A,*
(3) *$\gamma \circ \alpha : R \to N$ is a derivation over A.*
(4) *$\alpha \circ \phi : S \to M$ is a derivation over A.*

Proof. In each case, the map under consideration is definitely A-linear, so we just need to check the product rule.

(1) $(\alpha + \beta)(rs) = \alpha(rs) + \beta(rs) = r\alpha(s) + s\alpha(r) + r\beta(s) + s\beta(r) = r(\alpha + \beta)(s) + s(\alpha + \beta)(r)$;
(2) $t\alpha(rs) = t(r\alpha(s) + s\alpha(r)) = r(t\alpha(s)) + s(t\alpha(r))$;
(3) $(\gamma \circ \alpha)(rs) = \gamma((r\alpha(s) + s\alpha(r)) = r\gamma \circ \alpha(s) + s\gamma \circ \alpha(r)$.
(4) $(\alpha \circ \phi)(rs) = \alpha(\phi(r)\phi(s)) = \phi(s)\alpha(\phi(r)) + \phi(r)\alpha(\phi(s)) = s(\alpha \circ \phi)(r) + r(\alpha \circ \phi)(s)$, where the last equality is just recalling that M is a module by restriction of scalars. □

Definition 1.14. Let R be a ring, and M be and R-module. We set $\mathrm{Der}_R(M)$ to be the *module of derivations* of R into M. If R is an A-algebra via $\phi : R \to M$, we set $\mathrm{Der}_{R|A}(M)$ or $\mathrm{Der}_\phi(M)$ to be the *module of A-linear derivations* of R into M.

These are R-modules as a consequence of the proposition above.

Example 1.15. If A is a ring and $R = A[x_1, \dots, x_n]$ is a polynomial ring over A, then for any $f_1, \dots, f_n \in R$,

$$\begin{aligned} \textstyle\sum_{i=1}^n f_i \frac{d}{dx_i} : \ & R \longrightarrow R \\ & r \longmapsto \textstyle\sum_{i=1}^n f_i \frac{dr}{dx_i} \end{aligned}$$

is an A-linear derivation on R.

If M is an R-module, then for any $m_1, \dots, m_n \in M$, the map

$$\begin{aligned} m_i \frac{d}{dx_i} : \ & R \longrightarrow M \\ & r \longmapsto \frac{dr}{dx_i} m_i \end{aligned}$$

is an A-linear derivation, since it is the composition of the derivation $R \xrightarrow{\frac{d}{dx_i}} R$ and the R-linear map $R \xrightarrow{m} M$; adding these, the map

$$\begin{aligned} \textstyle\sum_{i=1}^n m_i \frac{d}{dx_i} : \ & R \longrightarrow M \\ & r \longmapsto \textstyle\sum_{i=1}^n \frac{dr}{dx_i} m_i \end{aligned}$$

is an A-linear derivation.

Example 1.16. Let's jack this example up. Let A be a ring, $R = A[X_\lambda \mid \lambda \in \Lambda]$ a polynomial ring over A, and $\{f_\lambda \mid \lambda \in \Lambda\}$ a sequence of elements in bijection with the variables then the formal sum

$$\sum_{\lambda \in \Lambda} f_\lambda \frac{d}{dX_\lambda} : R \to R$$

given by $r \mapsto \sum_{\lambda\in\Lambda} f_\lambda \frac{dr}{dX_\lambda}$ gives a well-defined map, since any $r \in R$ involves at most finitely many variables, and hence $\frac{dr}{dX_\lambda} = 0$ for all but finitely many $\lambda \in \Lambda$. This map is an A-linear derivation. Indeed, A-linearity is striaghtforward. To check the product rule, take $r, s \in R$; between the two, they involve only finitely many variables, and for these elements, the formula for this derivation agrees with the rule for the finitely many variables involved. By the last example, the product rule holds.

Similarly, for any R-module M and Λ-tuple of elements of M, there is a derivation

$$\sum_{\lambda \in \Lambda} m_\lambda \frac{d}{dX_\lambda} : R \to M$$

given by $f \mapsto \sum_{\lambda\in\Lambda} \frac{d}{dX_\lambda} m_\lambda$.

We would like to compute modules of derivations in some examples. The following lemma will help us recognize when we're done.

Lemma 1.17. *Let R be an A-algebra and $\{f_\lambda \mid \lambda \in \Lambda\}$ be a generating set of R as an A-algebra. Let M be an R-module. Then any A-linear derivation on R is determined by the images of f_λ. That is, $\alpha, \beta : R \to M$ are A-linear derivations with $\alpha(f_\lambda) = \beta(f_\lambda)$ for all λ, then $\alpha = \beta$.*

Proof. We need to show that $\alpha(r) = \beta(r)$ for any $r \in R$. Any element of R can be written as a sum of monomial expressions in the $f's$; i.e., a sum of terms of the form $r = a f_{\lambda_1}^{\mu_1} \cdots f_{\lambda_n}^{\mu_n}$ with $a \in A$ so it suffices to show that α and β take e same value on such a monomial r. We proceed by induction on $k = \mu_1 + \cdots + \mu_n$. When $k = 0$, $r \in A$ so $\alpha(r) = 0 = \beta(r)$. For the inductive step, take $k > 0$, so WLOG $\mu_1 \neq 0$; then $r = r' f_{\lambda_1}$, and

$$\alpha(r' f_{\lambda_1}) = r' \alpha(f_{\lambda_1}) + f_{\lambda_1} \alpha(r')$$

and likewise for β. By the starting assumption, $\alpha(f_{\lambda_1}) = \beta(f_{\lambda_1})$ and by the induction hypothesis $\alpha(r') = \beta(r')$. The equality follows. $\square$

Theorem 1.18. *Let A be a ring and $R = A[X_\lambda \mid \lambda \in \Lambda]$ be a polynomial ring over A. For any R-module M, the map*

$$\begin{aligned} \textstyle\prod_\Lambda M &\xrightarrow{\mu} \mathrm{Der}_{R|A}(M) \\ (m_\lambda)_\lambda &\longmapsto \textstyle\sum m_\lambda \frac{d}{dX_\lambda} \end{aligned}$$

is an isomorphism.

Proof. Consider the map $\nu : \mathrm{Der}_{R|A}(M) \to \prod_\Lambda M$ given by $\alpha \mapsto (\alpha(X_\lambda))_{\lambda\in\Lambda}$. The previous lemma shows that ν is injective. On the other hand,

$$(\nu \circ \mu)(m_\lambda)_\lambda = ((\sum_\lambda m_\lambda \frac{d}{dX_\lambda})(X_\lambda))_\lambda = (m_\lambda)_\lambda.$$

Thus, μ is injective. Then μ must be an isomorphism. Indeed, $\nu = (\nu\mu)\nu = \nu(\mu\nu)$ and ν injective implies $\mu\nu$ is the identity as well. $\square$

Lecture of January 31, 2023

We can give a description the derivations on any ring now.

Proposition 1.19. *Let R be an A-algebra. Write $R = S/I$ with $S = A[X_\lambda \mid \lambda \in \Lambda]$ and $I = (f_\gamma \mid \gamma \in \Gamma)$. Let M be an R-module. Then every A-linear derivation ∂ from R to M can be written in the form*

$$\sum_{\lambda\in\Lambda} m_\lambda \overline{\tfrac{d}{dx_\lambda}}$$
$$r = [s] \mapsto \sum_\lambda \tfrac{d}{dx_\lambda}(s) m_\lambda$$

for some unique $(m_\lambda)_\lambda \in \prod_\Lambda M$. A tuple of elements $(m_\lambda)_\lambda$ induces a well-defined derivation from R to M if and only if the corresponding derivation $\widetilde{\partial} : S \to M$ has $\widetilde{\partial}(f_\gamma) = 0$ for all γ.

Proof. Let $\pi : S \to R$ be the quotient map. Given an A-linear derivation $\partial : R \to M$, there is an A-linear derivation $\pi \circ \partial : S \to M$ that can be written in the form above by the previous theorem, so any derivation has this form. Since derivations are addition, such a derivation is well-defined so long as $\widetilde{\partial}(I) = 0$. This certainly implies that $\widetilde{\partial}(f_\gamma) = 0$ for all γ; conversely, any element of I can be written as $\sum_i s_i f_i$ and $\widetilde{\partial}(\sum_i s_i f_i) = \sum_i s_i \widetilde{\partial}(f_i) + \sum_i f_i \widetilde{\partial}(s_i)$, and the first sum is zero by hypothesis and the second since M is an R-module which is necessarily killed by I. $\square$

Example 1.20. Let's find some $\mathbb{C}$-linear derivations on $R = \frac{\mathbb{C}[x,y]}{x^2+y^3}$ to itself. Any such derivation must be a map of the form $\partial = r_1 \overline{\frac{d}{dx}} + r_2 \overline{\frac{d}{dy}}$ where $\widetilde{\partial} = r_1 \frac{d}{dx} + r_2 \frac{d}{dy} : \mathbb{C}[x,y] \to R$ has $\widetilde{\partial}(x^2 + y^3) = 0$, or just as well $\partial' = r'_1 \frac{d}{dx} + r'_2 \frac{d}{dy} : \mathbb{C}[x,y] \to \mathbb{C}[x,y]$ has $\partial'(x^2 + y^3) \in (x^2 + y^3)$. Since $\partial'(x^2 + y^3) = 2x\partial'(x) + 3y^2\partial'(y)$, we must have $2xr'_1 + 3y^2 r'_2 \in (x^2 + y^3)$. Here are a couple:

$$3x\overline{\frac{d}{dx}} + 2y\overline{\frac{d}{dy}} \qquad \text{and} \qquad 3y^2\overline{\frac{d}{dx}} + 2x\overline{\frac{d}{dy}}.$$

Example 1.21. Let's look at something simpler: $\mathrm{Der}_{\mathbb{C}|\mathbb{R}}(\mathbb{C})$. We can write $\mathbb{C} = \mathbb{R}[x]/(x^2 + 1)$, so such a derivation is of the form $r\overline{\frac{d}{dx}}$ where $r\frac{d}{dx}(x^2 + 1) \in (x^2 + 1)$. Since $2x = \frac{d}{dx}(x^2 + 1)$ and $x^2 + 1$ are coprime in $\mathbb{R}[x]$, r must be a multiple of $x^2 + 1$, so the corresponding derivation must be the zero map. Thus, there are no $\mathbb{R}$-linear derivations on $\mathbb{C}$.

1.2.1. *Lie algebra structure on* $\mathrm{Der}_{R|A}(R)$. Even more than a module, there is extra structure on $\mathrm{Der}_{R|A}(R)$. Any two elements of $\mathrm{Der}_{R|A}(R)$ have the same source and target, so we can compose them. The result is essentially never a derivation though.

Example 1.22. In $\mathbb{C}[x]$,

$$\frac{d^2}{dx^2}(x \cdot x) = 2 \neq 0 = x\frac{d^2}{dx^2}(x) + x\frac{d^2}{dx^2}(x).$$

However:

Proposition 1.23. *Let R be an A-algebra, and $\alpha, \beta \in \mathrm{Der}_{R|A}(R)$. Then the map $\alpha \circ \beta - \beta \circ \alpha : R \to R$ is an A-linear derivation.*

Proof. A-linearity follows since we have linear combinations or compositions of A-linear maps. Given $r, s \in R$,

$$\begin{aligned}(\alpha\beta - \beta\alpha)(rs) &= \alpha(r\beta(s) + s\beta(r)) - \beta(r\alpha(s) + s\alpha(r)) \\ &= \alpha(r\beta(s)) + \alpha(s\beta(r)) - \beta(r\alpha(s)) - \beta(s\alpha(r)) \\ &= \alpha(r)\beta(s) + r\alpha\beta(s) + s\alpha\beta(r) + \alpha(s)\beta(r) - \beta(r)\alpha(s) - r\beta\alpha(s) - \alpha(r)\beta(s) - s\beta\alpha(r) \\ &= r\alpha\beta(s) + s\alpha\beta(r) - r\beta\alpha(s) - s\beta\alpha(r) \\ &= r(\alpha\beta - \beta\alpha)(s) + s(\alpha\beta - \beta\alpha)(r)\end{aligned}$$

□

We write $[\alpha, \beta] := \alpha \circ \beta - \beta \circ \alpha$ and call this the *commutator* of α and β. This operation isn't a product operation for a ring (we will see soon that it's not associative), but it gives the structure of a *Lie algebra*.

Definition 1.24. A *Lie algebra* over a ring A is an A-module M equipped with an operation $[-,-] : M \times M \to M$ such that, for all $l, m, n \in M$ and $a \in A$:

- $[l + m, n] = [l, n] + [m, n]$ and $[l, m + n] = [l, m] + [l, n]$,
- $[am, n] = a[m, n]$ and $[m, an] = a[m, n]$,
- $[m, m] = 0$,
- $[l, [m, n]] + [m, [n, l]] + [n, [l, m]] = 0$.

Example 1.25. If N is an A-module, then $E = \mathrm{End}_A(N)$ (the collection of A-linear *endomorphisms* of N) with bracket $[\alpha, \beta] := \alpha \circ \beta - \beta \circ \alpha$ is a Lie algebra over A. The first three conditions are straightforward. The third follows from associativity of composition: To avoid foiling all these out, note that each term after expanding is a triple involving l, m, n. The expression above is stable under the permutation $l \mapsto m \mapsto n \mapsto l$, so it suffices to check that the triples lmn and lnm appear a cancelling number of times. Indeed, lmn appears with $+1$ from the first and -1 from the second and lnm appears with -1 from the first and $+1$ from the second.

Proposition 1.26. *Let R be an A-algebra. The commutator operation endows $\mathrm{Der}_{R|A}(R)$ with the structure of a Lie algebra over A.*

Proof. $\mathrm{Der}_{R|A}(R)$ is a submodule of $\mathrm{End}_A(R)$ and the bracket operation is consistent with that on the Lie algebra $\mathrm{End}_A(R)$, so it suffices to note that it is closed under the bracket operation. □

1.3. **Derivations and ideals.**

Proposition 1.27. *Let R be a ring, and I an ideal. Let $\partial : R \to M$ be a derivation. Then $\partial(I^n) \subseteq I^{n-1}M$ for all $n \in \mathbb{N}$.*

Proof. We proceed by induction on n, with $n = 1$ trivial. Given $r \in I^n$, write $r = \sum_i a_i b_i$ with $a_i \in I^{n-1}$ and $b_i \in I$. Then

$$\partial(r) = \sum_i \partial(a_i b_i) = \sum_i a_i \partial(b_i) + \sum_i b_i \partial(a_i).$$

Clearly $a_i\partial(b_i) \in I^{n-1}M$, and by the induction hypothesis $\partial(a_i) \in I^{n-2}M$, so $b_i\partial(a_i) \in I^{n-1}M$. □

It follows that every A-linear derivation $\partial : R \to M$ gives rise, by restriction/quotient, to a well-defined A-linear map $\overline{\partial} : I^n/I^{n+1} \to I^{n-1}M/I^nM$, and in particular $\overline{\partial} : I/I^2 \to M/IM$.

Proposition 1.28. *Let R be an A-algebra, I an ideal, and M an R-module. If $IM = 0$, then there is an isomorphism*

$$\mathrm{Der}_{R|A}(M) \to \mathrm{Der}_{(R/I^2)|A}(M)$$

and a well-defined map

$$\mathrm{Der}_{R|A}(M) = \mathrm{Der}_{(R/I^2)|A}(M) \to \mathrm{Hom}_A(I/I^2, M)$$

induced by restriction.

Example 1.29. Consider $R = \mathbb{C}[x_1, \ldots, x_n]$ and $\mathfrak{m}$ maximal. We have the restriction map

$$\mathrm{Der}_{R|\mathbb{C}}(R/\mathfrak{m}) \xrightarrow{\text{res}} \mathrm{Hom}_{\mathbb{C}}(\mathfrak{m}/\mathfrak{m}^2, R/\mathfrak{m}) = (\mathfrak{m}/\mathfrak{m}^2)^*,$$

where $(-)^*$ denotes $\mathbb{C}$-linear dual. The map is an isomorphism! To see it, note that $\mathfrak{m} = (x_1 - a_1, \ldots, x_n - a_n)$ for some vector a. Write $\widetilde{x_i} = x_i - a_i$. After a change of coordinates, we can consider R as a polynomial ring in the $\widetilde{x_i}$'s. Then $\mathfrak{m}/\mathfrak{m}^2$ is a vector space with basis given by the classes of the $\widetilde{x_i}$'s. By our proposition on derivations on polynomial rings, for any n-tuple of elements in $R/\mathfrak{m} \cong \mathbb{C}$, there is a unique derivation sending the corresponding variables there. That's what it means for the restriction to be an isomorphism! Concretely, the map

$$\sum_i \lambda_i \widetilde{x_i}^* \mapsto \sum_i \lambda_i \frac{d}{dx_i}|_{x=a}$$

is an inverse.

Lecture of February 2, 2023

We actually don't need the extremely strong hypothesis of polynomial ring in the last example. Let's party hard and figure out when, for a module with $IM = 0$, the map

$$\mathrm{Der}_{(R/I^2)|A}(M) \to \mathrm{Hom}_A(I/I^2, M)$$

is surjective (i.e., every homomorphism from the "I-top" extends to a derivation). A reasonable starting point is to take M to be I/I^2, which is the part of the ring R/I^2 itself that is definitely killed by I.

Theorem 1.30. *Let R be an A-algebra and I an ideal. Then an A-linear map $\alpha : R/I^2 \to I/I^2$ is an A-linear derivation if and only if the map*

$$\begin{aligned} R/I^2 &\xrightarrow{1+\alpha} R/I^2 \\ r &\longmapsto r + \alpha(r) \end{aligned}$$

is an A-algebra homomorphism.

Proof. We observe that the map $1 + \alpha$ is a sum of A-module homomorphisms, and hence A-linear. We just need to check that the product rule for α lines up with $1 + \alpha$ respecting multiplication. If α is a derivation, then

$$\begin{aligned} (1+\alpha)(rs) = rs + \alpha(rs) = rs + r\alpha(s) + s\alpha(r) &= rs + r\alpha(s) + s\alpha(r) + \alpha(r)\alpha(s) \\ = (r + \alpha(r))(s + \alpha(s)) &= (1+\alpha)(r)(1+\alpha)(s) \end{aligned}$$

where we used that $\alpha(r), \alpha(s) \in I/I^2$ so their product is zero. Conversely, following the equalities above, we must have $\alpha(rs) = r\alpha(s) + s\alpha(r)$ for the products to agree. □

This theorem gives an interesting and useful new way to think of derivations: they are "perturbations" of the identity map.

It also allows us to unlock many derivations.

Proposition 1.31. *Let R be an A-algebra and I an ideal. Suppose that the quotient map $\pi : R/I^2 \to R/I$ has an A-algebra right inverse, i.e., there is some A-algebra map $\tau : R/I \to R/I^2$ such that $\pi \circ \tau$ is the* *identity on R/I. Then for every R-module M with $IM = 0$, the map*

$$\mathrm{Der}_{R|A}(M) \xrightarrow{\text{res}} \mathrm{Hom}_A(I/I^2, M)$$

is surjective.

Proof. Consider the ring homomorphism $\tau \circ \pi : R/I^2 \to R/I^2$. Set $\alpha : R/I^2 \to R/I^2$ by $\tau \circ \pi - 1$. We claim that the image of α is in I/I^2. Indeed, for $r \in R/I^2$ we have $\pi\alpha(r) = \pi\tau\pi(r) - \pi(r) = \pi(r) - \pi(r) = 0$, so $\alpha : R/I^2 \to I/I^2$ has image in I/I^2. But $1 + \alpha = \tau\pi$ is a ring homomorphism, so α is a derivation, and α as well. Additionally, if $a \in I/I^2$, then $\pi(a) = 0$, so $-\alpha(a) = (\tau \circ \pi - 1)(-a) = a$. Thus, given $\phi : I/I^2 \to M$ R-linear, $\phi \circ -\alpha : R/I^2 \to M$ is an A-linear derivation on R/I^2 with restriction to I/I^2 being just ϕ. □

Example 1.32. Let R be a finitely generated $\mathbb{C}$-algebra, and $\mathfrak{m}$ a maximal ideal. Then $\mathbb{C} \subseteq R/\mathfrak{m}^2$ and $R/\mathfrak{m} \cong \mathbb{C}$, so there is a right inverse of the quotient map $R/I^2 \to R/I$. Moreover, $R/\mathfrak{m}^2$ is generated by $\mathfrak{m}/\mathfrak{m}^2$ as a $\mathbb{C}$-algebra, since $R/\mathfrak{m}^2 \cong \mathbb{C} \oplus \mathfrak{m}/\mathfrak{m}^2$ (or many other reasons). It follows that the map

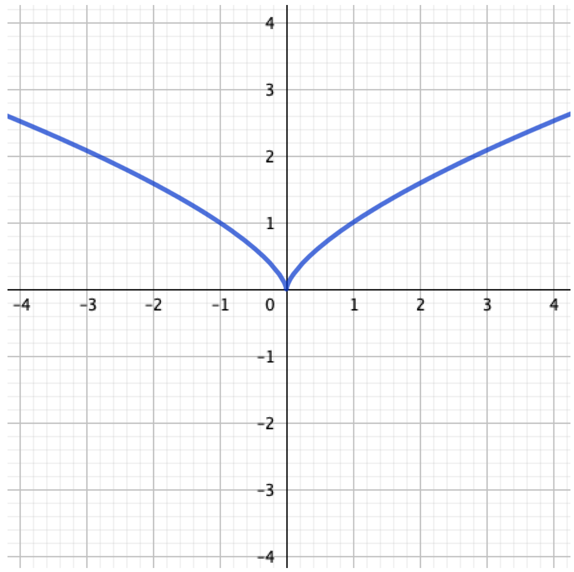
$$\mathrm{Der}_{R|\mathbb{C}}(R/\mathfrak{m}) \xrightarrow{\text{res}} (\mathfrak{m}/\mathfrak{m}^2)^*$$

is an isomorphism.

1.4. **Quick review of affine varieties.** Many of the constructions and questions we will consider will be motivated geometrically, and we will want to compare and contrast many of our main theorems with things we encounter in multivariable calculus, manifold theory, analysis, and other disciplines. We'll want to remember how to think of rings and ring homomorphisms geometrically. Over $\mathbb{C}$ (or an algebraically closed field) we have the following correspondence:

algebra	geometry
$\mathbb{C}[x_1, \dots, x_n]$	$\mathbb{C}^n$
reduced finitely-generated $\mathbb{C}$-algebra	variety
$R = \dfrac{\mathbb{C}[x_1, \dots, x_n]}{(f_1, \dots, f_m)} =: \mathbb{C}[X]$	$X :=$ solution set of $f_1 = \cdots = f_m = 0$
maximal ideal $\mathfrak{m}_a = (x_1 - a_1, \dots, x_n - a_n)$	point $a = (a_1, \dots, a_n)$
$r \in \mathbb{C}[X]$	polynomial function on X
going modulo $\mathfrak{m}_a$	evaluation at a
$\mathbb{C}$-algebra homomorphism $\mathbb{C}[X] \to \mathbb{C}[Y]$ $x_i \mapsto f_i((\underline{y}))$	morphism of varieties $Y \to X$ $b \mapsto (f_1(b), \dots, f_n(b))$

Example 1.33. Take $R = \mathbb{C}[x, y]/(x^2 - y^3)$. Geometrically, this corresponds to the solution set of $x^2 = y^3$ in 2-space. We can only draw the "real" picture, and we'll have to live with that.



Note the corner at $(0,0)$; we will see later that this has something to do with our unexpected derivation in the example above.

Example 1.34. This business about maps of varieties going the wrong way is a bit disorienting. Let's try a couple of examples of this.

- Given a (radical) ideal $I \subset S = \mathbb{C}[x_1, \ldots, x_n]$, the quotient map $S \to S/I$ is given by sending $x_i \mapsto x_i$, so the corresponding map of varieties $V(I) \to \mathbb{C}^n$ is just the inclusion map.
- Consider $\mathbb{C}[x,y]/(x^2 - y^3) \cong \mathbb{C}[t^2, t^3]$ (via $x \mapsto t^3, y \mapsto t^2$) and take the inclusion of rings $\mathbb{C}[t^2, t^3] \subseteq \mathbb{C}[t]$. Under the composition $x \mapsto t^3, y \mapsto t^2$ in $\mathbb{C}[t]$, and corresponding map of varieties goes from $\mathbb{C} \mapsto V(x^2 - y^3)$ and sends $b \mapsto (b^3, b^2)$.

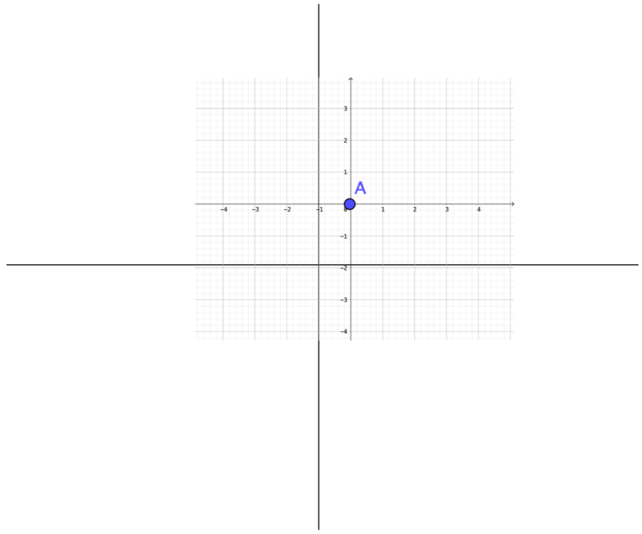
One important thing that is *not* included in this correspondence is the usual *Euclidean* topology on $\mathbb{C}^n$ or a subset $X \subseteq \mathbb{C}^n$ with an open basis given by $B_\varepsilon(a) = \{x \mid |x - a| < \varepsilon\}$ with the usual norm $|\cdot|$. We have the *Zariski* topology in which the closed sets are subvarieties, but this has no knowledge of what things are close in the Euclidean sense.

The magic making this all work out so nicely is the Nullstellensatz, which guarantees that maximal ideals of $\mathbb{C}[X]$ all correspond to points of X. In general, we just take the (instead of maximal) prime ideals to be our points and work from there.

algebra	"geometry"
ring	prime spectrum
R	$\mathrm{Spec}(R) = \{\mathfrak{p} \mid \mathfrak{p} \subset R \text{ prime ideal}\}$
prime ideal	point
maximal ideal	closed point
ring homomorphism $R \to S$	continuous map $Y \to X$
ϕ	$\mathfrak{q} \mapsto \phi^{-1}(\mathfrak{q})$

Whereas the correspondence between varieties and reduced f.g. $\mathbb{C}$-algebras was bijective above, the correspondence between rings and their spectra as topological spaces is far from: in particular, every field K has $\mathrm{Spec}(K)$ a singleton.

1.4.1. *Tangent spaces of varieties.* Let's get to the bottom of this corner business while we're at it. Let's define the *tangent space* of an affine variety X at a point a, $T_a(X)$. For starters, the tangent space of affine space $\mathbb{C}^n$ at a point a will be the vector space $\mathbb{C}^n$, thought of as centered at a.



We can recenter our coordinates there as $\widetilde{x_j} := x_j - a_j$. Now, given a variety $X = V(f_1, \ldots, f_m)$, for each f_i we look at its *linear part* near a: we can take its Taylor expansion at a

$$f_i = f_i(a) + \sum_j \frac{d}{dx_j}|_{x=a}(f_i)(x_j - a_j) + \text{higher order terms} .$$

Since $a \in X$, $f_i(a) = 0$, and we have

$$f_i = \sum_j \frac{d}{dx_j}|_{x=a}(f_i)\widetilde{x_j} + \text{higher order terms} ,$$

so the linear part of f is given by the linear functional $\nabla(f_i)|_{x=a} \cdot \widetilde{x}$. Then we take $T_a(X)$ to be the linear subspace of $T_a(\mathbb{C}^n)$ cut out by the linear equations $\nabla(f_1)|_{x=a}v = \nabla(f_m)|_{x=a}v = 0$. In particular, $T_a(\mathbb{C}^n)$ is the kernel of the *Jacobian matrix*

$$J(f_1, \ldots, f_m)|_{x=a} = \begin{bmatrix} \frac{d}{dx_1}|_{x=a}(f_1) & \cdots & \frac{d}{dx_n}|_{x=a}(f_1) \\ \vdots & \ddots & \vdots \\ \frac{d}{dx_1}|_{x=a}(f_m) & \cdots & \frac{d}{dx_n}|_{x=a}(f_m) \end{bmatrix},$$

whose rows are the gradient vectors.