1.9 06/28/24 - Derivations, Intersection Multiplicity

Today, we'll prove the Jacobi Criterion (Theorem 1.8.8), and start talking about intersection multiplicity for two curves.

1.9.1 Derivations and the Jacobi Criterion

We want to first discuss an algebraic way to differentiate things, for which we introduce derivations.

Definition 1.9.1. Let k be a field, and R be a ring containing k. A k-derivation on R is a k-linear map $D : R \to R$ satisfying the Leibniz rule, i.e. a map $D : R \to R$ such that

(a) for all $a, b \in k$ and $f, g \in R$, we have $D(af + bg) = a \cdot D(f) + b \cdot D(g)$, and
(b) for all $f, g \in R$, we have $D(fg) = D(f) \cdot g + f \cdot D(g)$.

The set of all k-derivations of R is denoted by $\mathrm{Der}_k(R)$.

Remark 1.9.2. The definition works also if k is any ring–then R can be any k-algebra, i.e. a ring with a homomorphism $\rho : k \to R$. Note that $\mathrm{Der}_k(R)$ is an R-module and a k-Lie algebra.[21]

Note that if $D \in \mathrm{Der}_k(R)$, then $D(c) = 0$ for all $c \in k$. This follows from

$$D(1) = D(1^2) = D(1) \cdot 1 + 1 \cdot D(1) = 2\,D(1),$$

so that $D(1) = 0$ and $D(c) = c \cdot D(1) = 0$. Therefore, a k-derivation on R captures the notion of differentiating elements of R, where elements of k function as "constants".

Example 1.9.3. If $R = k[x]$, then the operation $\sum_{i\geq 0} a_i x^i \mapsto \sum_{i\geq 1} i a_i x^{i-1}$ is a k-derivation on R, denoted ∂_x or $\partial/\partial x$. Note that if $c \in R$ is any element, then the operation $f \mapsto c \cdot \partial_x f$ is also a derivation of R. More generally, if $R = k[x_1, x_2, \dots, x_n]$, then the operations ∂_{x_j} are all derivations on R, and hence so are $\sum_{j=1}^n c_j \partial_{x_j}$. In fact, these are all the k-derivations of R.

Theorem 1.9.4. Let k be a field, and let $R = k[x_1, \dots, x_n]$ be the polynomial ring over R in $n \geq 1$ variables $x_1, \dots, x_n$. Then

$$\mathrm{Der}_k(R) = \bigoplus_{j=1}^n R \cdot \partial_{x_j}.$$

In other words, given any $c_1, c_2, \dots, c_n \in R$, there is a unique k-derivation $D : R \to R$ such that $D(x_j) = c_j$ for each $j = 1, \dots, n$.

Proof. It follows from the Leibniz rule that if $D : R \to R$ is any derivation and $f \in R$, then

$$D(f) = \sum_{j=1}^n \partial_{x_j}(f) \cdot D(x_j).$$

Therefore, a k-derivation D of R is determined by $D(x_j)$ for $j = 1, \dots, n$, showing uniqueness. Conversely, if $c_1, \dots, c_n$ are given, taking $D = \sum_{j=1}^n c_j \partial_{x_j}$ works, showing existence. ■

[21] As usual, if you don't know what this means, you can ignore it. If you do, what is the Lie algebra structure on $\mathrm{Der}_k(R)$?

It is now possible to derive algebraically the multivariable chain rule for polynomials. Let's do a special case–the only case we will need–to illustrate the process.

Lemma 1.9.5. Let $\phi : \mathbb{A}^2_k(x', y') \to \mathbb{A}^2_k(x, y)$ be an affine change of coordinates of the form $(x, y) = \phi(x', y') = (ax' + by' + p, cx' + dy' + q)$, where $a, b, c, d, p, q \in k$ satisfy $ad - bc \neq 0$. If $\phi^* : k[x, y] \to k[x', y']$ denotes the associated ring homomorphism, then for any $f \in k[x, y]$, we have

$$\partial_{x'}(\phi^* f) = a \cdot \phi^*(\partial_x f) + c \cdot \phi^*(\partial_y f) \text{ and}$$
$$\partial_{y'}(\phi^* f) = b \cdot \phi^*(\partial_x f) + d \cdot \phi^*(\partial_y f).$$

In particular, given any $Q \in \mathbb{A}^2_k$ and $f \in k[x, y]$, we have

$$\partial_x f|_Q = \partial_y f|_Q = 0 \Leftrightarrow \partial_{x'}(\phi^* f)|_{\phi^{-1}(Q)} = \partial_{y'}(\phi^* f)|_{\phi^{-1}(Q)} = 0.$$

The more traditional way to express the change of variables formula from Lemma 1.9.5 is to write

$$\frac{\partial f}{\partial x'} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial x'} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial x'} \text{ and}$$
$$\frac{\partial f}{\partial y'} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial y'} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial y'},$$

written which way, this formula is valid for other types of changes of coordinates as well.

Proof. We'll show the first identity; the proof of the second is similar. Since ϕ^* is a ring isomorphism, in light of Theorem 1.9.4, it suffices to show that the map

$$D : k[x, y] \to k[x, y] \text{ defined by } D(f) = (\phi^*)^{-1} \partial_{x'}(\phi^* f)$$

is a k-derivation, and that $D(x) = a$ and $D(y) = c$. This last part is easy: indeed,

$$D(x) = (\phi^*)^{-1} \partial_{x'}(\phi^* x) = (\phi^*)^{-1} \partial_{x'}(ax' + by' + p) = (\phi^*)^{-1} a = a,$$

and similarly $D(y) = c$. To check that this D is a derivation, note that condition (a) in Definition 1.9.1 is clear because ϕ^*, $\partial_{x'}$ and $(\phi^*)^{-1}$ are all k-linear, and condition (b) follows from the check that for all $f, g \in k[x, y]$ we have

$$\begin{aligned} D(fg) &= (\phi^*)^{-1} \partial_{x'}(\phi^*(fg)) \\ &= (\phi^*)^{-1} \partial_{x'}(\phi^* f \cdot \phi^* g) \\ &= (\phi^*)^{-1} \left[\partial_{x'}(\phi^* f) \cdot \phi^* g + \phi^* f \cdot \partial_{x'}(\phi^* g) \right] \\ &= \left((\phi^*)^{-1} \partial_{x'}(\phi^* f) \right) \cdot (\phi^*)^{-1} \phi^* g + (\phi^*)^{-1} \phi^* f \cdot \left((\phi^*)^{-1} \partial_{x'}(\phi^* g) \right) \\ &= D(f) \cdot g + f \cdot D(g). \end{aligned}$$

The second statement follows from the first by the same linear algebra as before, since again $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has nonzero determinant, i.e. is an invertible matrix. ∎

We are now ready to prove the Jacobi criterion, which we restate here for convenience.

Theorem 1.8.8 (Affine Jacobi Criterion)**.** Suppose we are given a curve $C \subset \mathbb{A}^2_k$ and a point $P = (p, q) \in \mathbb{A}^2_k$. Let $f \in k[x, y]$ be a minimal polynomial for C. Then

(a) $P \in C$ iff $f|_P := f(p, q) = 0$, and in this case
(b) P is a singular point of C iff

$$\left.\frac{\partial f}{\partial x}\right|_P = \left.\frac{\partial f}{\partial y}\right|_P = 0.$$

(c) If $P \in C$ is a smooth point, then the tangent line $\mathrm{T}_P C$ is defined by the vanishing of

$$\left.\frac{\partial f}{\partial x}\right|_P (x - p) + \left.\frac{\partial f}{\partial y}\right|_P (y - q) \in k[x, y].$$

Proof. The statement in (a) is clear. First , let's prove (b) and (c) for $P = O = (0, 0)$. If we write $f = f_1 + f_2 + \cdots + f_d$, where $d = \deg f$ and each f_j is homogeneous of degree j (note $P \in C$ is equivalent to $f_0 = 0$), then

$$f_1 = \lambda x + \mu y$$

for some $\lambda, \mu \in k$. Then

$$\partial_x f = \lambda + \partial_x f_2 + \cdots + \partial_x f_d,$$

and for each $j \geq 2$, we have $\partial_x f_j|_P = 0$, whence $\partial_x f|_P = \lambda$. Similarly, $\partial_y f|_P = \mu$. Therefore,

$$m_P(C) \geq 2 \Leftrightarrow f_1 = 0 \Leftrightarrow \lambda = \mu = 0 \Leftrightarrow \partial_x(f)|_P = \partial_y(f)|_P = 0.$$

Since

$$f_1 = \partial_x f|_P \cdot (x - 0) + \partial_y f|_P \cdot (y - 0),$$

the result of (c) is also clear. In general, let $\phi : \mathbb{A}^2_k \to \mathbb{A}^2_k$ be an affine change of coordinates such that $\phi(O) = P$. It is easy to see then that $\phi^* f$ is a minimal polynomial for $\phi^{-1} C$, and so we have

$$\begin{aligned} m_P(C) \geq 2 &\Leftrightarrow m_O(\phi^{-1}C) \geq 2 \\ &\Leftrightarrow \partial_{x'}(\phi^* f)|_O = \partial_{y'}(\phi^* f)|_O = 0 \\ &\Leftrightarrow \partial_x f|_P = \partial_y f|_P = 0 \end{aligned}$$

as needed, where in the last step we have used Lemma 1.9.5. The proof of (c) is similar, but can be simplified even more by noting that it suffices to consider a change of coordinates of the simple form $(x, y) = \phi(x', y') = (x' + p, y' + q)$; the details are left to the reader. ■

From this criterion, we can derive many important results. Here are a couple.

Theorem 1.9.6. A plane curve is singular at the points of intersection of its components. In particular, an affine curve is smooth iff its components are both individually smooth and pairwise disjoint.

Proof. Let f, g be two distinct irreducibles, and suppose $C = C_f \cup C_g = C_{fg}$; the general case is similar. By Theorem 1.8.8, it suffices to show that if $P \in C_f \cap C_g$, then $\partial_x(fg)|_P = \partial_y(fg)|_P = 0$, but this is clear because, for instance, we have

$$\partial_x(fg)|_P = \partial_x(f)|_P \cdot g|_P + f|_P \cdot \partial_x(g)|_P = 0$$

because $f|_P = g|_P = 0$. ■

Recall now our base assumption that k is algebraically closed.

Theorem 1.9.7. If $C \subset \mathbb{A}^2_k$ is any curve, then C has only finitely many singular points.

Proof. Let $C = C_1 \cup C_2 \cup \cdots \cup C_n$ be the irreducible decomposition of C (Theorem 1.7.10). For each $1 \le i < j \le n$, the intersection $C_i \cap C_j$ is finite by Theorem 1.7.11; therefore, it suffices to show the result for an irreducible C. Let $f \in k[x, y]$ be a minimal polynomial for C; then f is irreducible by Corollary 1.6.13(a). By Theorem 1.8.8, it suffices to show that the system of polynomial equations

$$f = \partial_x f = \partial_y f = 0$$

has only finitely many solutions in $\mathbb{A}^2_k$.[22] First suppose that $\partial_x f \neq 0$ (i.e. as a polynomial in $k[x, y]$). Since $\deg \partial_x f < \deg f$, it follows that either $\partial_x f$ is a nonzero constant (in which case C is smooth, and we are done), or that f and $\partial_x f$ are relatively prime (since f is prime and $\partial_x f$ cannot be a nonzero polynomial multiple of f for degree reasons), in which case we are done by Theorem 1.6.6. Similarly, if $\partial_y f \neq 0$, we are done.

This finishes the proof when $\operatorname{ch} k = 0$, because if $\operatorname{ch} k = 0$ and $f \in k[x, y]$ is any nonconstant polynomial, then one of $\partial_x f$ and $\partial_y f$ is nonzero. Unfortunately, when $\operatorname{ch} k = p > 0$, there are nonconstant $f \in k[x, y]$ such that $\partial_x f = \partial_y f = 0$, such as $f = x^p + y^p$. We will show that this cannot happen if f is irreducible: we will show that even if $\operatorname{ch} k = p > 0$, as long as $f \in k[x, y]$ is irreducible, then one of $\partial_x f$ and $\partial_y f$ is nonzero. Indeed, suppose not. Then to say that $\partial_x f = 0$ means that if we write $f = \sum_{i,j} a_{i,j} x^i y^j$, then $a_{i,j} = 0$ unless $p \mid i$. Similarly, $\partial_y f = 0$ implies that $a_{i,j} = 0$ unless $p \mid j$. Therefore, we conclude that

$$f = \sum_{i,j \ge 0} a_{pi,pj} x^{pi} x^{pj}.$$

Since k is algebraically closed, for each $i, j \ge 0$, we can find a p^{th} root $\alpha_{i,j} \in k$ of $a_{pi,pj}$, i.e. an element such that $\alpha_{i,j}^p = a_{pi,pj}$. Then, since we are in characteristic p,

$$f = \sum_{i,j \ge 0} \alpha_{i,j}^p x^{pi} y^{pj} = \left(\sum_{i,j \ge 0} \alpha_{i,j} x^i y^j \right)^p = g^p,$$

where $g := \sum_{i,j \ge 0} \alpha_{i,j} x^i y^j$, contradicting irreducibility of f. This completes the proof when $k = \overline{k}$; in general, we can reduce to this case by Theorem 1.4.5 as before. ■

Example 1.9.8. For any field k, consider the circle C defined by $f(x, y) := x^2 + y^2 - 1 \in k[x, y]$. This has partial derivatives

$$\partial_x f = 2x \text{ and } \partial_y f = 2y.$$

When $\operatorname{ch} k \neq 2$, it follows that this system $f = \partial_x f = \partial_y f = 0$ has no solutions, so that C is smooth. When $\operatorname{ch} k = 2$, it seems that $\partial_x f = \partial_y f = 0$, so that any point on C should be singular–why does this not contradict Theorem 1.9.7? Well, if we are to follow the proof of Theorem 1.9.7, we will observe that when $\operatorname{ch} k = 2$, in fact, we have that

$$f(x, y) = (x + y + 1)^2 \in k[x, y],$$

so that f is not reduced. In this case, the curve C is just a line with minimal polynomial $g(x, y) = x + y + 1$, which is also smooth. This example shows that when applying the Jacobi Criterion (Theorem 1.8.8), it is crucial to use a minimal polynomial for your curve. Another way to think about this is: a "curve" defined by a nonreduced polynomial is singular everywhere. This can be made precise using the language of schemes; we won't discuss this in this course.

[22] Here, I'm being a little sloppy about the distinction between polynomials and polynomial functions–given that we're in week 3, I'll presume you know what I mean and how to make this rigorous.

1.9.2 Intersection Multiplicity

Given curves $C, D \subset \mathbb{A}_k^2$ and a point $P \in C \cap D$, we want to make precise what we mean by the intersection multiplicity of C and D at P. Again, whatever this notion means, it should be invariant under affine (or even other kinds of) changes of coordinates, and as we observed in the previous sections, it is helpful to have this notion already for polynomials and not just curves–after all, we want to capture nonreduced behavior.

The goal, therefore, is to find a function

$$i : (k[x,y] \smallsetminus \{0\}) \times (k[x,y] \smallsetminus \{0\}) \times \mathbb{A}_k^2 \to \mathbb{Z}_{\geq 0} \cup \{\infty\}, \quad (f,g,P) \mapsto i_P(f,g)$$

that satisfies some reasonable properties. What properties should we have? Here I list a few.

(1) (Symmetry) $i_P(f,g) = i_P(g,f)$ for all f, g, P.
(2) (Finiteness for Proper Intersection) $i_P(f,g) = \infty$ iff f and g have a common component through P, i.e. there is a $q \in k[x,y]$ such that $q \mid f$ and $q \mid g$ and $q|_P = 0$.
(3) (Non-Intersection) $i_P(f,g) = 0$ iff $P \notin C_f \cap C_g$, i.e. either $f|_P \neq 0$ or $g|_P \neq 0$.
(4) (Additivity) $i_P(f_1 f_2, g) = i_P(f_1, g) + i_P(f_2, g)$ for all $f_1, f_2, g \in k[x,y] \smallsetminus \{0\}$ and $P \in \mathbb{A}_k^2$.
(5) (Coordinate Ring Dependence) $i_P(f,g) = i_P(f, g + hf)$ for all $f, g, h \in k[x,y] \smallsetminus \{0\}$.
(6) (Invariance under ACOCs) If $\phi : \mathbb{A}_k^2 \to \mathbb{A}_k^2$ is an ACOC, then $i_P(f,g) = i_{\phi^{-1}(P)}(\phi^*(f), \phi^*(g))$.
(7) (Normalization) For $P = O = (0,0)$, we have $i_O(x,y) = 1$.

The amazing result is, then, that these properties characterize intersection multiplicity uniquely.

Theorem 1.9.9. There is a unique function i satisfying (1)-(7) above.

We'll sketch a proof next time; today, let's work out a few examples this time. Firstly, by (3) and (4), scaling f or g by nonzero scalars does not change the intersection multiplicity.

Example 1.9.10. If $f = y^2 - x^2(x+1)$ and $g = x$ and $P = (0,0)$, then

$$i_P(y^2 - x^2(x+1), x) = i_P(y^2, x) = 2i_P(x,y) = 2.$$

If $g = y - tx$ for $t \in k$, then

$$\begin{aligned} i_P(y^2 - x^2(x+1), y - tx) &= i_P(y - tx, y^2 - x^2(x+1) - (y+tx)(y-tx)) \\ &= i_P(y - tx, x^2(-x + t^2 - 1)) \\ &= 2i_P(y - tx, x) + i_P(y - tx, -x + t^2 - 1) \\ &= 2 + \begin{cases} 1, & \text{if } t^2 - 1 = 0, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

This confirms our intuition that each line through P intersects the curve C_f at least twice, with even higher multiplicity (at most three) iff it is tangent to C_f at P.

Example 1.9.11. If C is a smooth curve with tangent line $L = \mathrm{T}_P C$ at $P \in C$ such that $C \neq L$, and f and ℓ are minimal polynomials for C and L, then $i_P(f, \ell) \geq 2$. Indeed, we can choose a suitable coordinate system so that $P = (0,0)$ and $\ell = y$; then $f_0 = 0$ and $f_1 = y$, whence

$$i_P(f, \ell) = i_P(y + (f - y), y) = i_P(f - y, y) \geq 2,$$

where in the last step we have used that $f - y$ is nonzero and homogeneous of degree at least 2. (How does this result follow?)

Example 1.9.12. Let $p(x) \in k[x]$ be a nonconstant polynomial of x alone, and let $f := y - p(x)$ and $g = y$. Then P is a point of intersection of the curves C_f (i.e. the graph of p) and C_g (i.e. the x-axis) iff $P = (\alpha, 0)$ for some root α of p. To compute the intersection multiplicity at this point, we factor $p(x) = (x - \alpha)^m q(x)$ for some integer $m \geq 1$ and $q(x) \in k[x]$ with $q(\alpha) \neq 0$, and then note that

$$i_P(f, g) = i_P((x-\alpha)^m q(x), y) = m \cdot i_P(x - \alpha, y) + i_P(q(x), y) = m \cdot 1 + 0 = m.$$

Therefore, the intersection multiplicity of f and g at P is exactly the multiplicity $m_\alpha(p)$ of α as a root of $p(x)$. In particular, we have

$$\sum_{P \in C_f \cap C_g} i_P(f, g) = \sum_{\alpha : p(\alpha) = 0} m_p(\alpha) = \deg p = (\deg f)(\deg g).$$

This is one simple manifestation of Bézout's Theorem, which we will soon get to. When $p(x) = 0$, every point on the x-axis is a point of infinite multiplicity, while if $p(x) = c$ is a nonzero constant, then there are no points of intersection, althought $(\deg f)(\deg g) = 1$; this is because the lines C_f and C_g are parallel (i.e. meet "at infinity"). We will soon develop tools to make this more precise.