Differentiation and Division

Rohan Hitchcock

24 August 2022

Let k be a commutative ring and R a k-algebra (soon we will assume k is a field of characteristic zero and R is a polynomial ring). Let $t = (t_1, \cdots, t_n)$ be a sequence of elements in R

Last talk we constructed a strong deformation retract involving the Koszul complex of t. This relied on the existence of certain k-linear maps $\partial_{t_1}, \dots, \partial_{t_n} : R \to R$, which together we referred to as a *system of t-derivatives*. In this talk we will show that such maps exist under certain assumptions on t, and that they can be computed algorithmically. These assumptions are satisfied when t is the sequence of partial derivatives of a potential (i.e. in the setting relevant to $\mathcal{LG}$). Recall that a system of t-derivatives is defined as follows. Let $k[t]$ denote the k-algebra generated by $1, t_1, \dots, t_n$.

Definition 1. Given $t = (t_1, \dots, t_n)$, *system of t-derivatives* are k-linear maps $\partial_{t_i} : R \to R$, $i = 1, \dots, n$ which satisfy the following properties:

(1) Every $f \in R$ can be written uniquely in the form

$$f = \sum_{u \in \mathbb{N}^n} r_u t^u$$

where each $r_u \in \bigcap_i \ker(\partial_{t_i})$ and finitely many $r_u \neq 0$.

(2) $\partial_{t_i}(t^v) = v_i t^{v - e_i}$ for all $v \in \mathbb{N}^n$ (where we understand that $0 t_j^{-1} = 0$).

(3) For $f \in k[t]$ and $r \in \bigcap_i \ker(\partial_{t_i})$ we have $\partial_{t_i}(rf) = r\partial_{t_i}(f)$.

Now suppose that $R = k[x] = k[x_1, \dots, x_m]$. When $n = m$ and $t = x = (x_1, \dots, x_n)$ our construction will recover the usual partial derivative maps $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$. In other words our goal is the generalise the notion of taking partial derivatives with respect to the sequence $x = (x_1, \dots, x_n)$ to taking partial derivatives with respect to other sequences in $k[x]$.

This generalisation is motivated by the observation that taking derivatives of polynomials is related to polynomial division. Consider the one variable case $k[x] = k[x_1]$ and let $f \in k[x]$. For another formal variable y one can show that in the polynomial ring $k[x, y]$ we have

$$f(x) = \sum_{p=1}^{\infty} \frac{1}{p!} f^{(p)}(y)(x - y)^p$$

analogously to the analytic Taylor's Theorem, where $f^{(p)} = \frac{d^p}{dx^p}(f)$ is the p^{th} partial derivative of f. Notice that the right-hand-side is a polynomial since eventually $f^{(m)} = 0$. Rearranging this we have

$$f(x) - f(y) = f'(y)(x - y) + (x - y)^2 \sum_{p=2}^{\infty} \frac{1}{p!} f^{(p)}(y)(x - y)^{p-2}$$

or in other words, $f'(y)$ is the remainder of $f(x) - f(y)$ divided by $(x-y)^2$. A similar result can be shown in the multivariate case.

1 Iterated Euclidean division

Our first task is to prove a corollary of the division algorithm in $k[x]$. Let k be a field. We begin by recalling some concepts related to polynomial division in the multivariate polynomial ring $k[x]$ following the conventions of [CLO15, Chapter 2]. A *monomial* in $k[x]$ is any polynomial in the set $\{x^u\}_{u\in\mathbb{N}^n}$. A *monomial ordering* on $k[x]$ is a well-founded, total order relation $<$ on $\mathbb{N}^n$ with the property that $u < v \implies u + w < v + w$. A typical example of a monomial ordering is the lexicographic ordering on $\mathbb{N}^n$, and others are given in [CLO15, Section 2.2].

Let $f = \sum_{u\in\mathbb{N}^n} c_u x^u \in k[x]$ where $c_u \in k$ and finitely many $c_u \neq 0$. Any $c_u x^u$ for which $c_u \neq 0$ is called a *term* of f. Given a monomial ordering on $k[x]$, if $f \neq 0$ we define the *multi-degree* of f as

$$\operatorname{multideg}(f) = \max\{u \in \mathbb{N}^n \mid c_u \neq 0\}$$

where the maximum is taken with respect to the monomial ordering. Setting $u^* = \operatorname{multideg}(f)$ we define the *leading term* of f with respect to the given monomial ordering to be $\operatorname{LT}(f) = c_{u^*} x^{u^*}$. The coefficient c_{u^*} is called the *leading coefficient* and is denoted $\operatorname{LC}(f)$. Consider the division algorithm on $k[x]$ given in [CLO15, Theorem 2.3.3]. We recall its properties:

Theorem 2 (Division Algorithm)**.** *Let $f, g_1, \ldots, g_m \in k[x]$ and suppose we have a monomial ordering on $k[x]$ for which $\operatorname{LC}(g_1), \ldots, \operatorname{LC}(g_m)$ are invertible in k. Given such a monomial ordering the division algorithm produces $r, q_1, \ldots, q_m \in k[x]$ which satisfy*

(1)
$$f = r + \sum_{i=1}^{m} q_i g_i$$

(2) None of the terms of r are divisible by any of the $\operatorname{LT}(g_i)$, $i = 1, \ldots, m$.

(3) For all $i = 1, \ldots, m$ with $q_i \neq 0$ we have $\operatorname{multideg}(f) \geq \operatorname{multideg}(q_i g_i)$.

In [CLO15] the division algorithm theorem is stated under the assumption that k is a field, but by inspecting the algorithm and proof given in [CLO15] one can observe that this hypothesis is only needed so the $\operatorname{LC}(g_i)$ can be inverted.

We show that the division algorithm can be iterated to "completely divide" the coefficients of the divisors $g_1, \ldots, g_m$ of Theorem 2. The properties of this iterated division algorithm are stated and proved in Theorem 4 and the algorithm itself is given in Algorithm 1.0.3.

Algorithm 1.0.3 Iterated Division Algorithm

Require: A polynomial $f \in k[x]$, non-constant polynomials $g_1, \dots, g_m \in k[x]$, and a monomial ordering on $k[x]$ such that $\mathrm{LC}(g_1), \dots, \mathrm{LC}(g_m)$ are invertible in k.

procedure ITERATEDDIVISION$(f, g_1, \dots, g_m)$
  $Q \leftarrow \{(\vec{0}, f)\}$ $\triangleright$ $\vec{0} = (0, \dots, 0) \in \mathbb{N}^m$
  $R \leftarrow \emptyset$
  while $Q \neq \emptyset$ **do**
    $Q_{\text{new}} \leftarrow \emptyset$
    for all $(u, q) \in Q$ **do**
      Apply the division algorithm to obtain $r, p_1, \dots, p_m \in k[x]$ satisfying

$$q = r + \sum_{i=1}^{m} p_i g_i$$

along with the other conditions in Theorem 2.
      $Q_{\text{new}} \leftarrow Q_{\text{new}} \cup \{(u + e_i, p_i) \mid i = 1, \dots, m \text{ where } p_i \neq 0\}$
      $R \leftarrow \{(u, r)\} \cup R$
    $Q \leftarrow$ COLLECTTERMS(Q_{new})
  return R
function COLLECTTERMS(Q)
  $Q_{\text{collected}} \leftarrow \emptyset$
  for all u where $(u, p) \in Q$ for some p **do**
    Let $p_1, \dots, p_s$ be all the polynomials such that $(u, p_i) \in Q$
    if $\sum_{i=1}^{s} p_i \neq 0$ **then**
      $Q_{\text{collected}} \leftarrow \{(u, \sum_{i=1}^{s} p_i)\} \cup Q_{\text{collected}}$
  return $Q_{\text{collected}}$

Theorem 4. *Let $f \in k[x]$ be a polynomial and $g_1, \ldots, g_m \in k[x]$ be non-constant polynomials. Suppose we have a monomial ordering on $k[x]$ such that $\mathrm{LC}(g_1), \ldots, \mathrm{LC}(g_m)$ are invertible in k. Given this monomial ordering,* $\textsc{iteratedDivision}(f, g_1, \ldots, g_m)$ *in Algorithm 1.0.3 computes an expression of the form*

$$f = \sum_{u \in \mathbb{N}^n} r_u g^u$$

where $g^u = g_1^{u_1} \cdots g_m^{u_m}$, all but finitely many of the $r_u = 0$ and for each $r_u \neq 0$, all terms of r_u are not divisible by any of the $\mathrm{LT}(g_i)$, $i = 1, \ldots, m$.

Proof. Let R_N and Q_N be the values of R and Q respectively in Algorithm 1.0.3 at the end of the N^{th} repetition of the loop on line 4, where we start counting from $N = 0$. Let $i(Q_N)$ and $i(R_N)$ be the indices arising in Q_N and R_N respectively, so $u \in i(Q_N)$ if and only if $(u, q) \in Q_N$ for some $q \in k[x]$ and likewise for $i(R_N)$.

We begin by showing that if the algorithm terminates then we obtain an expression of the stated form. First note that every index $u \in \mathbb{N}^n$ appears in Q_N and R_N at most once. That is, $|i(Q_N)| = |Q_N|$ and $|i(R_N)| = |R_N|$. Hence we can define

$$r_{u,N} = \begin{cases} 0 & u \notin i(R_N) \\ r & \text{where } (u, r) \in R_N \end{cases} \qquad \text{and} \qquad q_{u,N} = \begin{cases} 0 & u \notin i(Q_N) \\ q & \text{where } (u, q) \in Q_N \end{cases}$$

We aim to show that for all N we have

$$f = \sum_{u \in \mathbb{N}^m} r_{u,N} g^u + \sum_{u \in \mathbb{N}^m} q_{u,N} g^u$$

where $g^u = g_1^{u_1} \cdots g_m^{u_m}$. We proceed by induction on N. The base case is clear if we define R_{-1} and Q_{-1} to be the initial values of R and Q defined prior to line 4. Now consider the inductive case. For $u \in \mathbb{N}^m$ let $|u| = \sum_{i=1}^m u_i$. Note that if $r_{u,N-1} \neq 0$ then $|u| \leq N - 1$ and if $q_{u,N-1} \neq 0$ then $|u| \geq N$. We also have that if $r_{u,N-1} \neq 0$ then $r_{u,N} = r_{u,N-1}$ since we do not remove elements from R. Then we have

$$\begin{aligned} f &= \sum_{u:|u|<N} r_{u,N-1} g^u + \sum_{u:|u|\geq N} q_{u,N-1} g^u \\ &= \sum_{u:|u|<N} r_{u,N} g^u + \sum_{u:|u|\geq N} \left(r_{u,N} + \sum_{i=1}^m p_{u,i} g_i \right) g^u \\ &= \sum_u r_{u,N} g^u + \sum_u \sum_{i=1}^n p_{u,i} g^{u+e_i} \\ &= \sum_u r_{u,N} g^u + \sum_u q_{u,N} g^u \end{aligned}$$

where $p_{u,1}, \ldots, p_{u,m}$ are obtained by applying the division algorithm to $q_{u,N-1}$ as on line 7. Since the algorithm terminates when $Q = \emptyset$ and all $r_u \neq 0$ satisfy the required property this proves we have an expression of the desired form on termination.

It remains to prove that the algorithm terminates. We abuse notation and write $q \in Q_N$ to mean $(u, q) \in Q_N$ for some $u \in \mathbb{N}^m$. Now define

$$b_N = \max\{\mathrm{multideg}(q) \mid q \in Q_N\}$$

where the maximum is taken with respect to the chosen monomial ordering. Consider $q \in Q_{N-1}$ and let $p_1, \ldots, p_m$ be the polynomials computed from q on line 7. By Theorem 2 we have that $\text{multideg}(q) \geq \text{multideg}(p_i g_i)$. By hypothesis g_i is not a constant polynomial so this implies $\text{multideg}(q) > \text{multideg}(p_i)$ and in particular $b_{N-1} > \text{multideg}(p_i)$. Now, the elements of Q_N consist of sums of the various $p_1, \ldots, p_m$ generated on line 7. Since for any $s + t \neq 0$ we have $\text{multideg}(s + t) \leq \max\{\text{multideg}(s), \text{multideg}(t)\}$ [CLO15, Lemma 2.2.8] it follows that for any $q' \in Q_N$ that $\text{multideg}(q') < b_{N-1}$. Therefore $b_N < b_{N-1}$ and since monomial orderings are well-founded the algorithm terminates. $\square$

2 Differentiating with respect to a sequence of polynomials

Fix a monomial ordering $>_x$ on $k[x]$. We extend this to a monomial ordering on $k[x, y] = k[x_1, \ldots, x_m, y_1, \ldots, y_m]$ as follows. Let $(a, b), (a', b') \in \mathbb{N}^m \times \mathbb{N}^m$ where $a = (a_1, \ldots, a_m) \in \mathbb{N}^m$ and likewise for b, a' and b'. We define

$$(a, b) >_{x,y} (a', b') \equiv a >_x a' \text{ or } (a = a' \text{ and } b >_x b')$$

That is $>_{x,y}$ is the lexicographic ordering on $\mathbb{N}^m \times \mathbb{N}^m$ given by considering $>_x$ on each factor. This is clearly a monomial ordering on $k[x, y]$ which agrees with the monomial order on $k[x]$ when restricted to monomials involving only x-variables, and for which $x_i > y_j$ for all $i, j = 1, \ldots, m$. In particular $\text{LT}_x(f(x)) = \text{LT}_{x,y}(f(x) + f(y))$ for all $f \in k[x]$. From now on we dispense with distinguishing between $>_x$ and $>_{x,y}$ and simply use $>$ and LT to refer to both monomial orderings.

Now suppose $t = (t_1, \ldots, t_n)$ satisfies the following assumption:

Assumption 5. *If we have an expression of the form*

$$\sum_{r_u} t^u = 0$$

such that finitely many $r_u \neq 0$ and if $r_u \neq 0$ then no term of r_u is divisible by any of the $\text{LT}(T_i)$*, then $r_u = 0$ for all $u \in \mathbb{N}^n$*

This assumption means that expressions with coefficients satisfying the above conditions are unique. This assumption is satisfied when k is a field and t is quasi-regular, so holds in the context we need for $\mathcal{LG}$.

We next define $T_i = t_i(x) - t_i(y)$. If t satisfies the above property then so will T, and so we have the following result.

Lemma 6. *Any $F \in k[x, y]$ can be written uniquely in the form*

$$F = \sum_{u \in \mathbb{N}^n} r_u T^u$$

where $T^u = T_1^{u_1} \cdots T_n^{u_n}$, we have finitely many $r_u \neq 0$ and if $r_u \neq 0$ then no term of r_u is divisible by any of the $\text{LT}(T_i)$.

Proof. That such an expression exists follows from Theorem 4 and only uses that k is a field. Since T is quasi-regular such an expression is unique by the above assumption. $\square$

Given $f \in k[x]$ write

$$f(x) - f(y) = \sum_{u \in \mathbb{N}^n} r_u T^u$$

where the $r_u \in k[x,y]$ are the unique polynomials satisfying the conditions in Lemma 6. For each $u \in \mathbb{N}^n$ define a map $\rho_u : k[x] \to k[x,y]$ by setting $\rho_u(f) = r_u$. We now prove some facts about these maps. For $u, v \in \mathbb{N}^n$ define $u! = u_1!u_2! \cdots u_n!$ and

$$\binom{v}{u} = \begin{cases} 0 & \text{if any } v_i - u_i < 0 \\ \frac{v!}{u!(v-u)!} & \text{otherwise} \end{cases}$$

Lemma 7. *ρ_u is k-linear.*

Proof. Let $f, g \in k[x]$. Then we can write

$$(f+g)(x) - (f+g)(y) = \sum_{u \in \mathbb{N}^n} (\rho_u(f) + \rho_u(g)) T^u$$

Now, if $\rho_u(f) + \rho_u(g) \neq 0$ then no term of $\rho_u(f) + \rho_u(g)$ is divisible by any of the $\mathrm{LT}(T_i)$. Hence the right-hand-side satisfies the conditions in Lemma 6 and so by uniqueness $\rho_u(f+g) = \rho_u(f) + \rho_u(g)$. Likewise $\rho_u(cf) = c\rho_u(f)$ for $c \in k$. □

Lemma 8. *$\rho_u(t^v) = \binom{v}{u} t^{v-u}(y)$ for all $v \in \mathbb{N}^n$ and $u \neq 0$.*

Proof. It suffices to prove that

$$t^v(x) = \sum_u \binom{v}{u} t^{v-u}(y) T^u \tag{1}$$

Indeed, having shown (1) holds we have

$$t^v(x) - t^v(y) = \sum_{u \neq 0} \binom{v}{u} t^{v-u}(y) T^u$$

where we note that no term of $t^{v-u}(y)$ is divisible by any of the $\mathrm{LT}(T_i) = \mathrm{LT}(t_i(x))$.

We proceed by induction on $|v| = \sum_i v_i$. If $v = 0$ then both sides of (1) are equal to 1. Now suppose that $|v| \geq 1$. Let i be such that $v_i > 0$. Then, using the induction hypothesis, we have

$$\begin{aligned} t^v(x) &= t_i(x) t^{v-e_i}(x) \\ &= t_i(x) \sum_u \binom{v-e_i}{u} t^{v-e_i-u}(y) T^u \\ &= (t_i(y) + T_i) \sum_u \binom{v-e_i}{u} t^{v-e_i-u}(y) T^u \\ &= \sum_u \binom{v-e_i}{u} t^{v-u}(y) T^u + \sum_u \binom{v-e_i}{u} t^{v-e_i-u}(y) T^{u+e_i} \\ &= \sum_u \binom{v-e_i}{u} t^{v-u}(y) T^u + \sum_{u \neq 0} \binom{v-e_i}{u} t^{v-u}(y) T^u \\ &= t^v(y) + \sum_{u \neq 0} \left(\binom{v-e_i}{u} + \binom{v-e_i}{u-e_i} \right) t^{v-u}(y) T^u \\ &= t^v(y) + \sum_{u \neq 0} \binom{v}{u} t^{v-u}(y) T^u \\ &= \sum_u \binom{v}{u} t^{v-u}(y) T^u \end{aligned}$$

which proves the claim. □

Lemma 9. *Let $f \in k[t]$ and $r \in k[x]$ be such that no term of r is divisible by any of the* $\mathrm{LT}(t_i)$. *Then for $u \neq 0$ we have $\rho_u(rf) = r(x)\rho_u(f)$.*

Proof. It suffices to prove this for $f = t^v$ for $v \in \mathbb{N}^n$. Using Lemma 8 we have

$$\begin{aligned} r(x)t^v(x) - r(y)t^v(y) &= r(x)t^v(x) - r(x)t^v(y) + r(x)t^v(y) - r(y)t^v(y) \\ &= r(x)(t^v(x) - t^v(y)) + (r(x) - r(y))t^v(y) \\ &= r(x)\sum_{u\neq 0}\binom{v}{u}t^{v-u}(y)T^u + (r(x) - r(y))t^v(y) \\ &= (r(x) - r(y))t^v(y) + \sum_{u\neq 0}\binom{v}{u}r(x)t^{v-u}(y)T^u \end{aligned}$$

Notice that $\mathrm{LT}(t_j) = \mathrm{LT}(T_j)$ does not divide any term of $(r(x)-r(y))t^v(y)$ or $\binom{v}{u}r(x)t^{v-u}(y)$ for all $j = 1,\ldots,n$ and $u \in \mathbb{N}^n$. Hence by Lemma 6 this proves the claim. $\square$

Now let $e_i \in \mathbb{N}^n$ have a 1 in the i^{th} coordinate and 0 elsewhere and let $\varphi : k[x,y] \to k[x]$ be the k-algebra morphism identifying x and y. For each t_i we define a map $\partial_{t_i} : k[x] \to k[x]$ by setting $\partial_{t_i}(f) = \varphi\rho_{e_i}(f)$.

Proposition 10. *The maps $\partial_{t_i},\ldots,\partial_{t_i} : k[x] \to k[x]$ form a system of t-derivatives as defined in Definition 1.*

Proof. We need to show that $\partial_{t_1},\ldots,\partial_{t_n}$ are k-linear and satisfy

(1) Every $f \in k[x]$ can be written uniquely in the form

$$f = \sum_{u\in\mathbb{N}^n} r_u t^u$$

where $r_u \in \bigcap_i \ker(\partial_{t_i})$.

(2) $\partial_{t_i}(t^v) = v_i t^{v-e_i}$ for all $v \in \mathbb{N}^n$ (where we understand that $0t_j^{-1} = 0$).

(3) For $f \in k[t]$ and $r \in \bigcap_i \ker(\partial_{t_i})$ we have $\partial_{t_i}(rf) = r\partial_{t_i}(f)$.

That $\partial_{t_1},\ldots,\partial_{t_n}$ are k-linear, and properties (2) and (3) follow directly from Lemma 7, Lemma 8 and Lemma 9 respectively. For (1) note that we can write any $f \in k[x]$ in the form

$$f(x) = \sum_u r_u(x)t^u(x)$$

where if $r_u \neq 0$ then no term of r_u is divisible by any of the $\mathrm{LT}(t_i)$. This expression exists by Theorem 4 and is unique by the assumption on t, and note that $\rho_{e_i}(r_u) = 0$ for all u by Lemma 6. $\square$

Proposition 10 is the main result of this section, but before continuing we note some other properties of the maps $\partial_{t_1},\ldots,\partial_{t_n}$ defined in Proposition 10. As noted previously, since $\partial_{t_1},\ldots,\partial_{t_n}$ is a system of t-derivatives we have $\partial_{t_i}\partial_{t_j} = \partial_{t_j}\partial_{t_i}$ for all i,j. Hence for $a \in \mathbb{N}^n$ we define $\partial_t^a = \partial_{t_1}^{a_1}\cdots\partial_{t_n}^{a_n}$. The next result is analogous to Taylor's Theorem.

Proposition 11. $\partial_t^a = a!\varphi\rho_a$ *for all $a \neq 0$.*

Proof. Let $f \in k[x]$ and write

$$f(x) = \sum_u r_u(x) t^u(x)$$

where finitely many $r_u \neq 0$ and if $r_u \neq 0$ then no term of r_u is divisible by any of the $\mathrm{LT}(t_i)$. By Lemma 8 we have $\rho_a(t^u) = \binom{u}{a} t^{u-a}(y)$ and so

$$\begin{aligned}
\partial_t^a(f) &= \sum_u a! \binom{u}{a} r_u(x) t^{u-a}(x) \\
&= a! \sum_u r_u(x) \varphi \rho_a(t^u) \\
&= a! \sum_u \varphi \rho_a(r_u t^u) \\
&= a! \varphi \rho_a(f)
\end{aligned}$$

where we have that $r_u(x)\rho_a(t^u) = \rho_a(r_u t^u)$ by Lemma 9. □

Let $f \in k[x]$. Clearly one way to compute $\partial_{t_i}(f)$ is to use Algorithm 1.0.3 to compute an expression for f of the form

$$f(x) = \sum_u r_u(x) t^u(x)$$

where finitely many $r_u \neq 0$ and if $r_u \neq 0$ then no term of r_u is divisible by any of the $\mathrm{LT}(t_i)$. We then have

$$\partial_{t_i}(f) = \sum_u r_u(x) u_i t^{u-e_i}(x)$$

This approach needs many calls to the division algorithm as the whole expansion of $f(x)$ in terms of $t_1(x), \ldots, t_n(x)$ must be computed. A more efficient approach which only calls the division algorithm twice is given in Algorithm 2.0.12, in which $\partial_{t_j}(f) = \textsc{differentiate}(f, j, t_1, \ldots, t_n)$.

Algorithm 2.0.12 Computing ∂_{t_j}

1: **procedure** $\textsc{differentiate}(f, j, t_1, \ldots, t_n)$

2: Use the division algorithm in $k[x, y]$ to obtain $r(x, y), q_1(x, y), \ldots, q_n(x, y)$ satisfying

$$f(x) - f(y) = r(x, y) + \sum_{i=1}^{n} q_i(x, y)(t_i(x) - t_i(y))$$

along with the other conditions in Theorem 2.

3: Use the division algorithm in $k[x, y]$ to obtain $r'(x, y), p_1(x, y), \ldots, p_n(x, y)$ satisfying

$$q_j(x, y) = r'(x, y) + \sum_{i=1}^{n} p_i(x, y)(t_i(x) - t_i(y))$$

4: **return** $\varphi(r'(x, y))$

References

[CLO15] David A. Cox, John Little and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Cham: Springer International Publishing, 2015.