Recall: • A (radical/prime) differential ideal I of a differential ring (R, δ) is a (radical/prime) ideal I of R with $\delta(I) \subset I$.

• Let I be a proper radical differential ideal of R. Then

$$I = \cap_{I \subset P \text{ prime}} P.$$

• If R is a Ritt algebra (i.e., $\mathbb{Q} \subset R$), then $\{S\} = \sqrt{[S]}$ for any $S \subset R$, and a maximal differential ideal of R is always prime.

• The differential polynomial ring $K\{y_1, \ldots, y_n\} := K[\delta^k y_j : k \in \mathbb{N}; j = 1, \ldots, n]$.

In this chapter, we shall prove the Ritt-Raudenbush basis theorem, which is the differential analog of the Hilbert basis theorem for the differential polynomial ring. Hilbert's basis theorem states that the polynomial ring $k[x_1, \ldots, x_n]$ over a field k is Noetherian. That is, every ideal of $k[x_1, \ldots, x_n]$ is finitely generated (every ascending chain of ideals in $k[x_1, \ldots, x_n]$ is finite.) One might hope that the ACC condition holds for differential ideals in $K\{y_1, \ldots, y_n\}$. However, this is not true and *we do not have an exact analog of Hilbert's basis theorem for differential ideals.*

Non-example: Consider $\mathbb{Q}\{y\}$ and write $\delta(y) = y', \delta^2(y) = y'', \delta^3(y) = y''', \delta^k(y) = y^{(k)} (k \geq 4)$. Then the differential ideal $\mathcal{J}$ generated by

$$y^2, (y')^2, (y'')^2, \ldots, (y^{(k)})^2, \ldots$$

is not finitely differentially generated.

Proof. Suppose that $\mathcal{J}$ is finitely differentially generated. Then there exists $h \in \mathbb{N}$ such that

$$\mathcal{J} = [y^2, (y')^2, (y'')^2, \ldots, (y^{(h)})^2].$$

We will show that

$$(y^{(h+1)})^2 \notin [y^2, (y')^2, (y'')^2, \ldots, (y^{(h)})^2]$$

which will yield a contradiction.

Suppose the contrary. Obseve that when we consider $y, y', \ldots$ as usual variables over K, then all the $\delta^m((y^{(k)})^2)$ are homogenous of degree 2. Also, if we define the weight of $y^{(i)}y^{(j)}$ to be $i + j$, then the p-th derivative of $y^{(i)}y^{(j)}$ will be isobaric, with each term of weight $i + j + p$. Now if $(y^{(h+1)})^2 = \sum_{i=0}^{h} \sum_{j=0}^{m} A_{i,j} \delta^j((y^{(i)})^2)$, by the degree property, $A_{i,j} \in \mathbb{Q}$. Again considering the weights of the various forms, we find that

$$(y^{(h+1)})^2 = C_1 \delta^{2(h+1)}(y^2) + C_2 \delta^{2h}((y')^2) + \cdots + C_h \delta^2((y^{(h)})^2)$$

with $C_i \in \mathbb{Q}$. Now $\delta^{2(h+1)}(y^2)$ contains a term $y^{(2h+2)}y$ and none of the other derivatives yield such a term. We conclude that $C_1 = 0$. Continuing, we find each $C_i = 0$. This proves our statement. □

Obviously the differential ideal $\mathcal{J}$ in the above example is not radica. So, we can still hope that $K\{y_1, \ldots, y_n\}$ satisfies the ascending chain condition on radical differential ideals. That this is true is the content of the Ritt-Raudenbush basis theorem.

As a preparation, we first introduce characteristic set method, which is the main computational tool in differential algebra and also could provide some theoretical insights. For example, the use of characteristic sets makes the proof of the Ritt-Raudenbush basis theorem (Theorem 2.3.3) more transparent. The idea behind characteristic sets is similar to the notion of Gröbner basis.

2.2 Differential characteristic sets

Motivated Example (Ideal membership problem):

① In $\mathbb{Q}[x]$, every ideal is of the form $I = (f)$ for some $f \in \mathbb{Q}[x]$. By the Euclidean division algorithm, $g = qf + r$ with $r = \mathrm{rem}(g, f)$. Then $g \in I \Leftrightarrow r = 0$.

② In $\mathbb{Q}[x_1, \ldots, x_n]$, given an ideal $I = (f_1, \ldots, f_s) \subseteq \mathbb{Q}[x_1, \ldots, x_n]$, we use Gröbner basis to test whether $g \in I$.

③ How about the differential ideal membership problem? (differential characteristic sets)

Let (K, δ) be a differential field of characteristic zero. The differential polynomial ring $K\{Y\} \triangleq K\{y_1, \ldots, y_n\}$ in the differential variables $Y = \{y_1, \ldots, y_n\}$ can be viewed as a polynomial ring in the algebraic variables $\Theta(Y) \triangleq \{\delta^i(y_j) \mid i \in \mathbb{N}, j = 1, \ldots, n\}$. (i.e., $K\{Y\} = K[\Theta(Y)]$)

A **differential ranking** on $\Theta(Y)$ is a total ordering on $\Theta(Y)$ satisfying

(1) $u < \delta(u)$ for all $u \in \Theta(Y)$ and

(2) if $u, v \in \Theta(Y)$ with $u < v$, then $\delta(u) < \delta(v)$.

Example:

- The set $\Theta(y) = \{\delta^i(y) : i \in \mathbb{N}\}$ has a unique ranking $y < \delta(y) < \delta^2(y) < \delta^3(y) < \cdots$.
- Two important rankings on $\Theta(Y)$ are the following:
  1) Elimination ranking: $y_i > y_j \Rightarrow \delta^k(y_i) > \delta^l(y_j)$ for any $k, l \in \mathbb{N}$.
  2) Orderly ranking: $k > l \Rightarrow \delta^k(y_i) > \delta^l(y_j)$ for all $i, j \in \mathbb{N}$.

Lemma 2.2.1. *Every ranking is a well-ordering (i.e., every nonempty subset of $\Theta(Y)$ has a least element).*

Proof. Let $U \subseteq \Theta(Y)$ and $U \neq \emptyset$. For each $j \in \{1, \ldots, n\}$, if $\exists\, i \in \mathbb{N}$ s.t. $\delta^i(y_j) \in U$, then set $k_j = \min\{i \mid \delta^i(y_j) \in U\}$ and set $u_j = \delta^{k_j}(y_j)$. Then the least element of U is the least element in the finite set of u_j's. □

Until the end of this subsection, we assume a ranking $\mathscr{R}$ is fixed. And by convention, $1 < \delta^i(y_j)$.

Definition 2.2.2. *Let $f \in K\{y_1, \ldots, y_n\} \backslash K$. The* **leader** *of f is the largest element of $\Theta(Y)$ with respect to $\mathscr{R}$ which appears effectively in f, denoted by u_f or $\mathrm{ld}(f)$. By the two conditions in the definition of ranking, for each $i \in \mathbb{N}$, $\mathrm{ld}(\delta^i(f)) = \delta^i(\mathrm{ld}(f))$. We write f as a univariate polynomial of u_f, then $f = I_d(u_f)^d + I_{d-1}(u_f)^{d-1} + \cdots + I_1 u_f + I_0$, where I_i is free of u_f and $d = \deg(f, u_f)$. The leading coefficient I_d is called the* **initial** *of f and denoted by I_f. The pair $\mathrm{rk}(f) := (u_f, d)$ is called the* **rank** *of f.*

Example: Let $f = (y')^2 - 4y \in \mathbb{Q}\{y\}$. Then $u_f = \mathrm{ld}(f) = y'$ and $\mathrm{I}_f = 1$. Apple δ to f, then we have $\delta(f) = 2y'y'' - 4y'$. So we get $u_{\delta(f)} = y'' = \delta(u_f)$ and $\mathrm{I}_{\delta(f)} = 2y' = \frac{\partial f}{\partial y'}$.

Note that in the above example, $\deg(\delta(f), u_{\delta(f)}) = 1$ and $\mathrm{I}_{\delta(f)} = \frac{\partial f}{\partial u_f}$.

Definition 2.2.3. *Let $f \in K\{y_1, \ldots, y_n\} \backslash K$. We call $\frac{\partial f}{\partial u_f}$ the* **separant** *of f, denoted by S_f.*

Remark:

1) $f = \sum_{i=0}^{d} I_i {u_f}^i \Longrightarrow \delta(f) = \sum_{i=1}^{d} I_i \delta({u_f}^i) + \sum_{i=0}^{d} \delta(I_i){u_f}^i = (\sum_{i=1}^{d} I_i \cdot i \cdot {u_f}^{i-1})\delta(u_f) + \sum_{i=0}^{d} \delta(I_i){u_f}^i = \mathrm{S}_f \cdot \delta(u_f) + \sum_{i=0}^{d} \delta(I_i){u_f}^i$.

   Note that $u_{\delta(f)} = \delta(u_f)$, $\mathrm{I}_{\delta(f)} = \mathrm{S}_f$ and $\deg(\delta(f), u_{\delta(f)}) = 1$. ($\mathrm{char}(K) = 0$)

   Also, for $k > 0$, $\delta^k(f) = \mathrm{S}_f \cdot \delta^k(u_f) +$ tail polynomial involving derivatives less than $\delta^k(u_f)$.

   So $u_{\delta^k(f)} = \delta^k(u_f)$, $\mathrm{I}_{\delta^k(f)} = \mathrm{S}_f$, $\deg(\delta^k(f), u_{\delta^k(f)}) = 1$.

   ((K, δ) is a δ-field, c is algebraic over $K \Rightarrow$ there is a unique way to make $(K(c), \delta)$ a δ-field.)

2) By convention, for $f \in K\backslash\{0\}, u_f = 1$.

Definition 2.2.4. *Let $f, g \in K\{Y\}$, we say that f is* **partially reduced** *with respect to g if none of the proper derivatives of u_g (i.e., $\delta^i(u_g)$ with $i > 0$) appears effectively in f.*

Example:

1) Let $f = y^2, g = y + 1$. Since $u_g = y$ and none of the proper derivatives of y appears in f, f is partially reduced with respect to g.

2) Let $f = 2y\delta(y)^2 + y$ and $g = y + 1$. Since $\delta(u_g) = \delta(y)$ appears in the first term of f, f isn't partially reduced with respect to g.

Definition 2.2.5. *We say f is* **reduced** *with respect to g if*

1) f is partially reduced with respect to g, and

2) $\deg(f, u_g) < \deg(g, u_g)$.

Definition 2.2.6. *A subset $\mathcal{A} \subseteq K\{y_1, \ldots, y_n\}$ is called an* **autoreduced set** *if any element of $\mathcal{A}$ is reduced with respect to any other element of $\mathcal{A}$.*

Remark: If an autoreduced set $\mathcal{A}$ contains an element $A \in K\backslash\{0\}$, then $\mathcal{A} = \{A\}$.

Lemma 2.2.7. *Every autoreduced set of $K\{y_1, \ldots, y_n\}$ is finite.*

Proof. Let $\mathcal{A}$ be an autoreduced set. For each $i = 1, \ldots, n$, there exists at most one differential polynomial $A \in \mathcal{A}$ such that $\mathrm{ld}(A) = \delta^k(y_i)$ for some $k \in \mathbb{N}$, for two differential polynomials A_1, A_2 with $\mathrm{ld}(A_j) = \delta^{k_j}(y_i)$ couldn't be reduced with respect to each other. Thus $|A| \leq n$. □

Remark: The partial differential analogues of Lemma 2.2.1 and Lemma 2.2.7 are also valid and can be proved by Dickson's lemma.

Definition 2.2.8. *Let $f, g \in K\{y_1, \ldots, y_n\}\backslash K$. We say f has lower rank than g if* $\mathrm{rk}(f) <_{lex} \mathrm{rk}(g)$[2]*, denoted by $f < g$. By convention, each element of $K\backslash\{0\}$ has lower rank than elements of $K\{Y\}\backslash K$.*

Notation: We use $f \leq g$ to denote either $f < g$ or f and g have the same rank. Note that " $\leq$ " is a pre-order among $K\{y_1, \ldots, y_n\}$, that is, a binary relation that is reflexive and transitive.

In the following, we write an autoreduced set in the order of increasing rank, i.e., $\mathcal{A} = A_1, \ldots, A_p$ with $\mathrm{rk}(A_1) <_{lex} \mathrm{rk}(A_2) <_{lex} \cdots <_{lex} \mathrm{rk}(A_p)$. In the following, we shall define an ordering on autoreduced sets.

Let $\mathcal{A} = A_1, \ldots, A_p$ and $\mathcal{B} = B_1, \ldots, B_q$ be two autoreduced sets. We say $\mathcal{A}$ **has lower rank than** $\mathcal{B}$ and write $\mathcal{A} < \mathcal{B}$ if either

[2] $<_{lex}$ is a well-ordering of $\Theta(Y) \times \mathbb{N}^*$.

1) $\exists\, k\, (\leq \min\{p, q\})$ such that $\forall\, i < k, \mathrm{rk}(A_i) = \mathrm{rk}(B_i)$ and $A_k < B_k$, or

2) $p > q$ and for each $i \leq q, \mathrm{rk}(A_i) = \mathrm{rk}(B_i)$.

If neither $\mathcal{A} < \mathcal{B}$ nor $\mathcal{B} < \mathcal{A}$, we say $\mathcal{A}$ and $\mathcal{B}$ are of the same rank. Clealry, $\mathcal{A}$ and $\mathcal{B}$ have the same rank if and only if $p = q$ and $\forall\, i \leq p, \mathrm{rk}(A_i) = \mathrm{rk}(B_i)$. Say $\mathcal{A} \leq \mathcal{B}$ iff $\mathcal{A} < \mathcal{B}$ or $\mathcal{A}$ and $\mathcal{B}$ have the same rank. (" $\leq$ " is a pre-order.)

Example: Consider $K\{y_1, y_2\}$ and take the orderly ranking with $y_1 < y_2$. Let $\mathcal{A} = \{A_1 = (y_2')^2 + 1, A_2 = y_1'' + y_2\}$, $\mathcal{B} = \{B_1 = y_2' + 2\}$ and $\mathcal{C} = \{C_1 = (y_2')^2 + 2\}$. Since $\mathrm{rk}(A_1) > \mathrm{rk}(B_1), \mathcal{B} < \mathcal{A}$. Since $\mathrm{rk}(A_1) = \mathrm{rk}(C_1)$ and $|\mathcal{A}| > |\mathcal{C}|, \mathcal{A} < \mathcal{C}$.

Proposition 2.2.9. *Any nonempty set of autoreduced sets in $K\{Y\} = K\{y_1, \ldots, y_n\}$ contains an autoreduced set of lowest rank.*

Proof. Let U be any nonempty set of autoreduced sets of $K\{Y\}$. Define by induction a sequence of subsets of U as follows: $U_0 \triangleq U$, and for $i > 0$, we define $U_i = \{\mathcal{A} \in U_{i-1} \mid \mathrm{card}(\mathcal{A}) \geq i$, the i-th element of $\mathcal{A}$ is of lowest rank$\}$. Then $U_0 \supseteq U_1 \supseteq \cdots$. By Lemma 2.2.7, $\exists\, i \in \mathbb{N}$ (actually $i \leq n$ in our ordinary differential case) such that $U_i \neq \emptyset$ and $U_{i+1} = \emptyset$. Actually, any element of U_i is an autoreduced set in U of lowest rank. □

Definition 2.2.10. *Let $I \subseteq K\{Y\}$ be a differential ideal. An autoreduced set of lowest rank contained in I is called a* **characteristic set** *of I (with respect to the given ranking).*

Remark: By convention, $\emptyset$ and $\{a\}$ with $a \in K^*$ are autoreduced sets. (Here, $\mathrm{rk}(a) = (1, 1)$.)

We start to introduce *pesudo-division* of differential polynomials:

Lemma 2.2.11. *Let $\mathcal{A} = A_1, \ldots, A_p$ be an autoreduced set in $K\{Y\}$ and $F \in K\{Y\}$. Then there exist $\tilde{F} \in K\{Y\}$ and $t_i \in \mathbb{N}$ satisfying*

1) $\tilde{F}$ is partially reduced with respect to $\mathcal{A}$ (i.e., $\tilde{F}$ is partially reduced w.r.t. each A_i),

2) the rank of $\tilde{F}$ is not higher than that of F,

3) $\prod_{i=1}^{p} \mathrm{S}_{A_i}^{t_i} F \equiv \tilde{F} \bmod [\mathcal{A}]$.

More precisely, $\prod_{i=1}^{p} \mathrm{S}_{A_i}^{t_i} F - \tilde{F}$ can be expressed as a linear combination of derivatives $\theta(A_i)$ with coefficients in $K\{Y\}$ such that $\theta(u_{A_i}) \leq u_F$.

Proof. If F is partially reduced with respect to $\mathcal{A}$, then set $\tilde{F} = F$ and $t_i = 0\, (i \leq p)$. Otherwise, F contains a proper derivative $\delta^k(u_{A_i})$ of the leader of some A_i. Let v_F be the maximal one among all such derivatives. We shall prove the lemma by induction on v_F. Suppose for all $G \in K\{Y\}$ that doesn't involve a proper derivative of any u_{A_i} of rank $\geq v_F$, the corresponding $\tilde{G}$ and natural numbers are defined satisfying the desired properties. There exists a unique $A \in \mathcal{A}$ such that $v_F = \delta^k(u_A)$ for some $k > 0$. If $A = \sum_{i=0}^{d} I_i {u_A}^i$, then

$$\delta^k(A) = \mathrm{S}_A \delta^k(u_A) + T \text{ with } T \text{ having lower rank than } \delta^k(u_A) = v_F.$$

Denoting $l = \deg(F, v_F)$ and write F as $F = \sum_{i=0}^{l} J_i {v_F}^i$ where $J_0, \ldots, J_l$ don't involve proper derivatives of any u_{A_i} of rank $\geq v_F$. Then we have

$$\mathrm{S}_A{}^l F = \sum_{i=0}^{l} J_i \mathrm{S}_A{}^{l-i} (\mathrm{S}_A v_F)^i \equiv \sum_{i=0}^{l} J_i \mathrm{S}_A{}^{l-i} (-T)^i \bmod (\delta^k(A)).$$

Clearly, $G = \sum_{i=0}^{l} J_i \mathrm{S}_A{}^{l-i}(-T)^i$ doesn't involve proper derivatives of any u_{A_i} of rank $\geq v_F$. By the induction hypothesis, $\exists\ \tilde{G}$ partially reduced with respect to $\mathcal{A}$ and $k_i \in \mathbb{N}$ such that $\prod_{i=1}^{p} \mathrm{S}_{A_i}^{k_i} G \equiv \tilde{G} \bmod [\mathcal{A}]$. Now it suffices to set $\tilde{F} = \tilde{G}$, $t_i = \begin{cases} k_i, & A_i \neq A \\ k_i + l, & A_i = A \end{cases}$. □

Remark: $\tilde{F}$ constructed by the process in the proof is called the *partial remainder* of F w.r.t $\mathcal{A}$.

Recall the *pseudo reduction algorithm* in commutative algebra:
Let D be an integral domain and v an indeterminate over D. Let $F, A \in D[v]$ be of respective degrees d_F, d_A. Suppose $A = I_{d_A} v^{d_A} + \cdots + I_1 v + I_0 \neq 0$ with $I_i \in D$. Let $e = \max\{d_F - d_A + 1, 0\}$. Then we can compute unique $Q, R \in D[v]$ s.t. $I_{d_A}^e F = QA + R$ and $\deg(R, v) < \deg(A, v)$.

Theorem 2.2.12. *Let $\mathcal{A} = A_1, \ldots, A_p$ be an autoreduced set in $K\{y_1, \ldots, y_n\}$. If $F \in K\{y_1, \ldots, y_n\}$, then $\exists$ a δ-polynomial F_0 (called the* **differential remainder** *of F w.r.t. $\mathcal{A}$) and $r_i, t_i \in \mathbb{N}$ such that*

1) *F_0 is reduced w.r.t $\mathcal{A}$,*
2) *The rank of F_0 is no higher than the rank of F,*
3) $\prod_{i=1}^{p} \mathrm{S}_{A_i}^{t_i} \mathrm{I}_{A_i}^{r_i} F \equiv F_0 \bmod [\mathcal{A}]$.

Proof. Let $\tilde{F}$ be the partial remainder of F with respect to $\mathcal{A}$ and $\prod_{i=1}^{p} \mathrm{S}_{A_i}^{t_i} F \equiv \tilde{F} \bmod [\mathcal{A}]$. Let $r_p = \max\{0, \deg(F, u_{A_p}) - \deg(A_p, u_{A_p}) + 1\}$. Then $\exists\ F_{p-1} \in K\{Y\}$ partially reduced with respect to $\mathcal{A}$ and reduced with respect to A_p such that $\mathrm{I}_{A_p}^{r_p} \tilde{F} \equiv F_{p-1} \bmod (A_p)$. If $p = 1$, then we are done. Otherwise, we can find r_{p-1} and $F_{p-2} \in K\{Y\}$ partially reduced with respect to $\mathcal{A}$ and reduced with respect to A_{p-1}, A_p s.t. $\mathrm{I}_{A_{p-1}}^{r_{p-1}} \mathrm{I}_{A_p}^{r_p} \tilde{F} \equiv F_{p-2} \bmod (A_{p-1}, A_p)$ and is not higher than $\tilde{F}$. Continuing in this way, we get F_0 satisfying the desired properties.

□

Remark: The reduction procedures above could be summarized in an algorithm, called the *Ritt-Kolchin algorithm* to compute the δ-remainder of a δ-polynomial F with respect to an autoreduced set $\mathcal{A}$. Denote F_0 above by $\delta\text{-rem}(F, \mathcal{A})$, or $F \underset{\mathcal{A}}{\to} F_0$.

Example: Consider $K\{y_1, y_2\}$ and fix the orderly ranking with $y_1 > y_2$.

(1) Let $f = y_1$ and $\mathcal{A} = A_1 = y_2 y_1$. Here $f \underset{\mathcal{A}}{\to} 0$, and $\mathrm{I}_{A_1} f - 0 \in [\mathcal{A}]$.

(2) Let $f = y_1' + 1$ and $\mathcal{A} = A_1 = y_2 y_1^2$. $u_{A_1} = y_1$ and $\mathrm{S}_{A_1} = 2y_2 y_1$. Clearly, f is not partially reduced with respect to $\mathcal{A}$. Note that $\delta(A_1) = 2y_2 y_1 y_1' + y_2' y_1^2$. The partial remainder of f with respect to $\mathcal{A}$ is $2y_2 y_1 - y_2' y_1^2 = \tilde{f}$ and $\mathrm{S}_{A_1} f - A_1' = \tilde{f}$. Since

$$\mathrm{I}_{A_1} \tilde{f} - \mathrm{I}_{\tilde{f}} A_1 = y_2(2y_2 y_1 - y_2' y_1^2) - (-y_2') y_2 y_1^2 = 2y_2^2 y_1$$

is reduced with respect to $\mathcal{A}$, $f \underset{\mathcal{A}}{\to} 2y_2^2 y_1$ and $\mathrm{I}_{A_1} \mathrm{S}_{A_1} f - 2y_2^2 y_1 = -y_2' A_1 + \mathrm{I}_{A_1} A_1' \in [\mathcal{A}]$.