# ABSTRACT DERIVATIONS

## A STUDY IN CONNECTION WITH HOMOMORPHISMS, ALGEBRAIC SEPARABILITY, AND PURELY INSEPARABLE ALGEBRAIC FIELD EXTENSIONS OF EXPONENT ONE



by

ALISTER O'BRIEN CAMPBELL

A Thesis presented to the University of Manitoba in partial fulfilment of the requirements for the degree of Master of Science

February, 1968

# INTRODUCTION

Throughout our considerations we shall presuppose a knowledge of the basic concepts of modern abstract algebra. In particular, we shall presuppose a knowledge of vector spaces, rings and ideals, elementary Galois theory of field extensions, the construction (intrinsic) of a tensor product of algebras as in, for example, Zariski and Samuel [12]. Unless otherwise stated, we shall assume that rings (subrings) are commutative with identity $1 \neq 0$, and that algebras are associative with identity $1 \neq 0$.

The first section deals with some properties of derivations and gives their connections with extensions of algebra homomorphisms and separable and inseparable algebraic extensions. In the second section we introduce the notion of p-dependence and give further discussions on derivation algebras. In both these sections, there are worked exercises from Jacobson [9], some of which lead to results due to Baer and Hochschild. We also derive an analogue to the normal basis theorem. In the third section we derive a Galois type correspondence between subfields $\Phi$ of a given field P which is purely inseparable of exponent one over $\Phi$, $[P:\Phi] < \infty$, and derivation algebras which are finite dimensional over $\Phi$. In the fourth section we introduce the notion of higher derivations (of finite rank) and examine briefly higher derivations of purely inseparable fields P over $\Phi$. We include the case where P is a tensor product of simple extensions.

Finally, and without in any way making him responsible for the

contents of this work, I should like to take this opportunity to publicly thank Dr. K. W. Armstrong most of all for the criticisms he made and the encouragement I received during the preparation of this thesis.

## SECTION I

Definition 1.1. A non-associative (= not necessarily associative) algebra $\mathfrak{A}$ over a field $\Phi$, usually denoted by $\mathfrak{A}/\Phi$, is a vector space over $\Phi$ in which a product $xy \in \mathfrak{A}$ is defined for $x$, $y$ in $\mathfrak{A}$ such that

(1) $(x_1 + x_2)y = x_1y + x_2y$, $x(y_1 + y_2) = xy_1 + xy_2$

(2) $\alpha(xy) = (\alpha x)y = x(\alpha y)$, $\alpha \in \Phi$

An algebra $\mathfrak{A}$ is called associative if its multiplication satisfies the associative law

$$(xy)z = x(yz), \quad x, y, z \text{ in } \mathfrak{A}.$$

We recall that a sub-algebra of $\mathfrak{A}$ is a subspace of $\mathfrak{A}$ which is also a subring. Observe that $\Phi[x_1, \ldots, x_n]$, the ring of polynomials in the $n$ indeterminates $x_1, \ldots, x_n$ with coefficients in $\Phi$ is an algebra (commutative) over $\Phi$. For this reason, $\Phi[x_1, \ldots, x_n]$ and $\Phi(x_1, \ldots, x_n)$, the field of rational functions of $\Phi[x_1, \ldots, x_n]$ are frequently referred to as algebras over $\Phi$.

Definition 1.2. A non-associative algebra $\mathfrak{A}$ is called a Lie algebra if its multiplication satisfies the Lie conditions

(3) $x^2 = 0$, $(xy)z + (yz)x + (zx)y = 0$.

The second condition is the so-called Lie-Jacobi identity.

Definition 1.3. If $\mathfrak{A}/\Phi$ is a sub-algebra of an algebra $\mathfrak{B}/\Phi$, a derivation $D$ of $\mathfrak{A}/\Phi$ into $\mathfrak{B}/\Phi$ is a mapping of $\mathfrak{A}/\Phi$ into $\mathfrak{B}/\Phi$ such that for $x, y$ in $\mathfrak{A}$, $\alpha$ in $\Phi$,

(4) $\quad (x + y)D = xD + yD, \quad (x\alpha)D = (xD)\alpha$

(5) $\quad (xy)D = (xD)y + x(yD)$.

Condition (4) states that D is linear. If $\mathfrak{U} = \mathfrak{B}$, then we speak of a derivation in $\mathfrak{B}$ or a derivation of $\mathfrak{B}$ into itself. The mapping of the polynomial algebra $\Phi[x]$ into itself given by $f(x) \to f'(x)$ the formal derivative of $f(x)$ is clearly an example of a derivation in $\Phi[x]$.

Let $\mathrm{Der}_{\Phi}(\mathfrak{U},\mathfrak{B})$ denote the set of derivations of $\mathfrak{U}/\Phi$ into $\mathfrak{B}/\Phi$. Then $D \in \mathrm{Der}_{\Phi}(\mathfrak{U},\mathfrak{B})$ is a linear transformation of $\mathfrak{U}/\Phi$ into $\mathfrak{B}/\Phi$ satisfying the special condition (5). If D, $D_1$, $D_2$ are linear transformations of $\mathfrak{U}/\Phi$ into $\mathfrak{B}/\Phi$, $x \in \mathfrak{U}$, $\alpha \in \Phi$, define $D_1 \pm D_2$, $D\alpha$, $D_1 D_2$ respectively by $x(D_1 \pm D_2) = xD_1 \pm xD_2$, $x(D\alpha) = (xD)\alpha$, and $x(D_1 D_2) = (xD_1)(D_2)$. Thus $D_1 \pm D_2$, $D\alpha$, and $D_1 D_2$, are linear transformations of $\mathfrak{U}/\Phi$ into $\mathfrak{B}/\Phi$. In particular, if $D_1, D_2 \in \mathrm{Der}_{\Phi}(\mathfrak{U},\mathfrak{B})$, $x,y \in \mathfrak{U}$, $\alpha \in \Phi$, we then have

$$\begin{aligned}
(xy)(D_1 \pm D_2) &= (xy)D_1 \pm (xy)D_2 \\
&= (xD_1)y + x(yD_1) \pm (xD_2)y + x(yD_2) \\
&= (xD_1 \pm xD_2)y + x(yD_1 \pm yD_2) \\
&= \{x(D_1 \pm D_2)\}y + x\{y(D_1 \pm D_2)\}
\end{aligned}$$

and

$$\begin{aligned}
(xy)D\alpha &= \{(xy)D\}\alpha = \{(xD)y + x(yD)\}\alpha \\
&= \{(xD)y\}\alpha + \{x(yD)\}\alpha \\
&= \{(xD)\alpha\}y + x\{(yD)\alpha\} \\
&= \{x(D\alpha)\}y + x\{y(D\alpha)\} .
\end{aligned}$$

This shows that $D_1 \pm D_2$, $D\alpha$ belong to $\mathrm{Der}_{\Phi}(\mathfrak{U},\mathfrak{B})$.

Remark 1.1. Take $\mathfrak{U} = \mathfrak{B}$. Then we observe that $D\alpha$, $D_1 \pm D_2$ are derivations in $\mathfrak{B}$. However, it should not be inferred that $D_1 D_2$ is also a derivation in $\mathfrak{B}$. Indeed, it is clear that

$$\begin{aligned}(xy)(D_1 D_2) &= \{(xy)D_1\}D_2 = \{(xD_1)y + x(yD_1)\}D_2 \\ &= \{(xD_1)y\}D_2 + \{x(yD_1)\}D_2 \\ &= x(D_1 D_2)y + (xD_1)(yD_2) + (xD_2)(yD_1) + x\{y(D_1 D_2)\}\end{aligned}$$

Hence $(xy)(D_1 D_2) \neq \{x(D_1 D_2)\}y + x\{y(D_1 D_2)\}$ for all x,y in $\mathfrak{B}$. In view of this remark, we can say no more than Der $(\mathfrak{B},\mathfrak{B})$ = Der $(\mathfrak{B})$ is a subspace of $\mathfrak{L}(\mathfrak{B},\mathfrak{B}) = \mathfrak{L}(\mathfrak{B})$ the space of linear transformations in $\mathfrak{B}$.

Definition 1.4. Let $\mathfrak{U}$ be an associative algebra. If $D, D_1, D_2, \ldots$ belong to Der $\mathfrak{U}$, then the Lie product or additive commutator of $D_1$ and $D_2$ is given by $[D_1, D_2] = D_1 D_2 - D_2 D_1$. It is clear that $[D_1, D_2]$ belongs to $\mathfrak{L}(\mathfrak{U})$. We next observe that the following relation is satisfied

(6) $\quad [D,D] = 0; \quad \big[[D_1,D_2], D_3\big] + \big[[D_2,D_3], D_1\big] + \big[[D_3,D_1], D_2\big] = 0$

The first part of the relation (6) is evident. The second part follows immediately since

$$\begin{aligned}\big[[D_1,D_2], D_3\big] &= [(D_1 D_2 - D_2 D_1), D_3] \\ &= (D_1 D_2 - D_2 D_1)D_3 - D_3(D_1 D_2 - D_2 D_1) \\ &= D_1 D_2 D_3 - D_2 D_1 D_3 - D_3 D_1 D_2 + D_3 D_2 D_1 \\ \big[[D_2,D_3], D_1\big] &= (D_2 D_3 - D_3 D_2)D_1 - D_1(D_2 D_3 - D_3 D_2) \\ &= D_2 D_3 D_1 - D_3 D_2 D_1 - D_1 D_2 D_3 + D_1 D_3 D_2 \\ \big[[D_3,D_1], D_2\big] &= (D_3 D_1 - D_1 D_3)D_2 - D_2(D_3 D_1 - D_1 D_3) \\ &= D_3 D_1 D_2 - D_1 D_3 D_2 - D_2 D_3 D_1 + D_2 D_1 D_3 .\end{aligned}$$

We next observe that

$$\begin{aligned}
[D_1 + D_2 , D_3] &= (D_1 + D_2)D_3 - D_3(D_1 + D_2) \\
&= D_1 D_3 + D_2 D_3 - D_3 D_1 - D_3 D_2 \\
&= D_1 D_3 - D_3 D_1 + D_2 D_3 - D_3 D_2 \\
&= [D_1 , D_3] + [D_2 , D_3] \qquad ,
\end{aligned}$$

$$\begin{aligned}
[D_1 , D_2 + D_3] &= D_1 (D_2 + D_3) - (D_2 + D_3) D_1 \\
&= D_1 D_2 + D_1 D_3 - D_2 D_1 - D_3 D_1 \\
&= [D_1 , D_2] + [D_1 , D_3]
\end{aligned}$$

and $$\begin{aligned}
\alpha[D_1 , D_2] &= \alpha(D_1 D_2) - \alpha(D_2 D_1) \\
&= (\alpha D_1) D_2 - D_2 (\alpha D_1) = [\alpha D_1 , D_2] \ .
\end{aligned}$$

Since $\alpha(D_1 D_2) - \alpha(D_2 D_1) = D_1 (\alpha D_2) - (\alpha D_2) D_1 = [D_1 , \alpha D_2]$ ,

we must therefore have

$$\alpha[D_1 , D_2] = [\alpha D_1 , D_2] = [D_1 , \alpha D_2] \text{ for any } \alpha \in \Phi \ .$$

We have already shown (cf. Remark 1.1) that

$$(xy)\, D_1 D_2 = \{x(D_1 D_2)\}y + (xD_1)(yD_2) + (xD_2)(yD_1) + x\{y(D_1 D_2)\}.$$

Since this relation is clearly symmetrical in $D_1$ and $D_2$ ,

$$(xy) D_2 D_1 = \{x(D_2 D_1)\}y + (xD_2)(yD_1) + (xD_1)(yD_2) + x\{y(D_2 D_1)\}.$$

Clearly $x\{y(D_1 D_2)\} - x\{y(D_2 D_1)\} = x\{y(D_1 D_2) - y(D_2 D_1)\}$

and $\{x(D_1 D_2)\}y - \{x(D_2 D_1)\}y = \{x(D_1 D_2 - D_2 D_1)\}y$ .

Therefore $(xy)[D_1 , D_2] = (x[D_1 , D_2])y + x(y[D_1 , D_2])$.

Hence $D$, $D_1$ , $D_2$ belong to $\mathrm{Der}_\Phi(\mathfrak{A})$ and $\alpha \in \Phi$ together imply that $D_1 \pm D_2$ , $D\alpha$, $[D_1 , D_2]$ belong to $\mathrm{Der}_\Phi(\mathfrak{A})$. These observations lead to the following definition:

Definition 1.5. Der $\mathfrak{A}$ the set of derivations in the algebra $\mathfrak{A}$ is called the Lie algebra of derivations or simply the derivation algebra of $\mathfrak{A}$.

Let D be a derivation in $\mathfrak{A}$ and $x,y \in \mathfrak{A}$. Then induction on k gives the Leibniz rule

$$(7) \qquad (xy)D^k = (xD^k)y + \sum_{i=1}^{k-1} \binom{k}{i}(xD^i)(yD^{k-i}) + x(yD^k)$$

where $\binom{k}{i}$ is the usual binomial coefficient[1] $\dfrac{k(k-1)\ldots(k-i+1)}{1.2\ldots i}$

Proof. Take $D^0 = 1$. Then (7) holds for k = 1 by definition. We may point out that if $D_1 = D_2$ in Remark 1.1, the formula (7) holds for k = 2. Let us now assume that (7) holds for all $k \leq n$. Since $(xy)D^{n+1} = \{(xy)D\}D^n = \{(xD)y + x(yD)\}D^n$, we now have

$$\begin{aligned}(xy)D^{n+1} &= \{(xD)y\}D^n + \{x(yD)\}D^n \\ &= (xD^{n+1})y + \sum_{i=1}^{n-1} \binom{n}{i}(xD^{i+1})(yD^{n-i}) + (xD)(yD^n) \\ &\quad + (xD^n)(yD) + \sum_{j=1}^{n-1} \binom{n}{j}(xD^j)(yD^{n+1-j}) + x(yD^{n+1}) \\ &= (xD^{n+1})y + \sum_{k=1}^{n} (xD^h)(yD^{n+1-h})\left(\binom{n}{h} + \binom{n}{h-1}\right) + x(yD^{n+1}) \ .\end{aligned}$$

Since $\binom{n}{h} + \binom{n}{h-1} = \binom{n+1}{h}$, we now have

---

[1] This coefficient will be assumed to be a rational integer.

$$(xy)D^{n+1} = (xD^{n+1})y + \sum_{i=1}^{n} \binom{n+1}{i}(xD^i)(yD^{n+1-i}) + x(yD^{n+1}) \ .$$

Thus (7) holds for all rational integers $k = 1,2, \ldots$

If the characteristic of $\Phi$ is 0 we can divide (7) by $k!$ and obtain the relation

$$(8) \qquad (xy)\frac{D^k}{k!} = \frac{(xD^k)y}{k!} + \sum_{i=1}^{k-1} \left(\frac{xD^i}{i!}\right)\left(\frac{yD^{k-i}}{(k-i)!}\right) + \frac{x(yD^k)}{k!}$$

$$= \sum_{i=0}^{k} \left(\frac{xD^i}{i!}\right)\left(\frac{yD^{k-i}}{(k-i)!}\right)$$

We shall now give a direct connection between derivations and automorphisms. Let $\mathfrak{A}$ be the polynomial algebra $\Phi[x]$ where $\Phi$ is a field of characteristic 0. Let a derivation D in $\mathfrak{A}$ be defined by $f(x)D = f'(x)$ the formal derivative of $f(x) \in \Phi[x]$. Consider the series

$$G = \exp D = 1 + \frac{D}{1!} + \frac{D^2}{2!} + \ldots \ .$$

If $f(x)$ is of degree n, then $f(x)\ D^{n+1} = 0$. Hence the series $f(x)G$ converges. We assert that $f(x)G = f(x + 1)$.

Let us set $f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$ .

Then $f(x)G = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$

$$+ a_0\binom{n}{1}x^{n-1} + a_1\binom{n-1}{1}x^{n-2} + a_2\binom{n-2}{1}x^{n-3} + \ldots + a_{n-2}\binom{2}{1}x + a_{n-1}$$

$$+ a_0\binom{n}{2}x^{n-2} + a_1\binom{n-1}{2}x^{n-3} + a_2\binom{n-2}{2}x^{n-4} + \ldots + a_{n-2}$$

$$+ a_0\binom{n}{3}x^{n-3} + a_1\binom{n-1}{3}x^{n-4} + \ldots \ .$$

$$\ldots\ldots\ldots + a_0$$

Rearranging the terms, we now have

$$f(x)G = a_0 x^n + a_0\binom{n}{1}x^{n-1} + a_0\binom{n}{2}x^{n-2} + \ldots + a_0$$

$$+ a_1 x^{n-1} + a_1\binom{n-1}{1}x^{n-2} + \ldots + a_1 + \ldots$$

$$\ldots\ldots\ldots + a_n$$

$$= a_0 (x + 1)^n + a_1 (x + 1)^{n-1} + \ldots + a_n = f(x + 1) \in \Phi[x]$$

This shows that $f(x - 1)G = f(x)$, whence G is onto.

Since the map D is linear, it is clear that the map G is also linear. Consequently, in order to show that $\Big(f(x)\ h(x)\Big)G = f(x)G \cdot h(x)G$, we need only verify this statement for $f(x) = x^r$ and $h(x) = x^s$, $0 \leq r,s$. We have seen that $f(x)G = (x + 1)^r$ and $h(x)G = (x + 1)^s$. Therefore $\Big(f(x)\ h(x)\Big)G = x^{r+s}G = (x + 1)^{r+s} = f(x)G \cdot h(x)G$. Finally, we assert that the kernel of the map G is the zero polynomial. Let $h(x) = b_m x^m + \ldots + b_1 x + b_0$ be an arbitrary non-zero polynomial of degree m. Then $b_m \neq 0$. Hence $h(x - 1) = b_m(x - 1)^m + \ldots$ $+ b_1 (x - 1) + b_0$ is also a non-zero polynomial of degree m. This shows that the kernel of the map G is the zero polynomial. We have thus shown that exp D is an automorphism of $\Phi[x]$.

Definition 1.6. Let $\Phi$ be a field of characteristic p (= 0 or otherwise). A <u>restricted Lie algebra</u> of characteristic p is an algebra $\mathfrak{R}_p$ over $\Phi$ in which the multiplication [x,y] satisfies

$$[x,y] = -[y,x]$$

$$\Big[[x,y],z\Big] + \Big[[y,z],\ x\Big] + \Big[[z,x],y\Big] = 0 \quad ,$$

and for every y in $\mathfrak{R}_p$ there exists an element called $y^p$ such that

$$\Big[\ldots\ldots\big[\overbrace{[x,y],y],\ldots\ldots y}^{p \text{ times}}\big]\Big] = [x,y^p] \text{ for any } x \in \mathfrak{R}.$$

A restricted subalgebra $\mathfrak{R}_p'$ of $\mathfrak{R}_p$ is a subalgebra containing $y^p$ for every y in $\mathfrak{R}_p'$. Similarly we define a restricted ideal, etc. It should not be inferred that this element $y^p$ is necessarily an ordinary p-th power since multiplication is not necessarily associative. If $p \neq 2$, $[x,x] = 0$ for all $x \in \mathfrak{R}_p$ and $\mathfrak{R}_p$ is then a Lie algebra.

Exercise 1.1. Let $\mathfrak{U}$ be an (associative) algebra over $\Phi$ and $d \in \mathfrak{U}$. Verify that the mapping $I_d : a \to [a,d] = ad - da$ is a derivation in $\mathfrak{U}$. Such a mapping is called an inner derivation in $\mathfrak{U}$. Prove that

$I_{\alpha_1 d_1 + \alpha_2 d_2} = \alpha_1 I_{d_1} + \alpha_2 I_{d_2}$ , $\alpha_i \in \Phi$ , and

$I_{[d_1,d_2]} = [I_{d_1}, I_{d_2}]$. Show that if $\Phi$ is of characteristic $p \neq 0$, then $I_{d^p} = (I_d)^p$ .

Proof. Let $a,b \in \mathfrak{U}$ , $a \neq 0$. Then we have

$$\begin{aligned}(ab)I_d &= (ab)d - d(ab)\\ &= abd - adb + adb - dab\\ &= a(bd - db) + (ad - da)b\\ &= a(b\ I_d) + (a\ I_d)b\end{aligned}$$

Since $I_d$ is evidently linear, $I_d$ is a derivation in $\mathfrak{U}$.

Now $a\, I_{\alpha_1 d_1 + \alpha_2 d_2} = a(\alpha_1 d_1 + \alpha_2 d_2) - (\alpha_1 d_1 + \alpha_2 d_2)a$

$= a\,\alpha_1 d_1 - \alpha_1 d_1 a + a\,\alpha_2 d_2 - \alpha_2 d_2 a$

$= a(I_{\alpha_1 d_1} + I_{\alpha_2 d_2})$ .

Since $a\, I_{\alpha d} = a(\alpha d) - (\alpha d)a$

$= \alpha(ad) - \alpha(da) = \alpha(ad - da)$

implies that $I_{\alpha d} = \alpha\, I_d$ , we must have

$$I_{\alpha_1 d_1 + \alpha_2 d_2} = \alpha_1 I_{d_1} + \alpha_2 I_{d_2} \quad .$$

Next, $a\, I_{[d_1, d_2]} = a[d_1, d_2] - [d_1, d_2]a$

$= ad_1 d_2 - ad_2 d_1 - d_1 d_2 a + d_2 d_1 a$

$= ad_1 d_2 - d_1 ad_2 - d_2 ad_1 + d_2 d_1 a$

$+ d_2 ad_1 - ad_2 d_1 + d_1 ad_2 - d_1 d_2 a$

$= (ad_1 - d_1 a)d_2 - d_2(ad_1 - d_1 a)$

$- (ad_2 - d_2 a)d_1 + d_1(ad_2 - d_2 a)$

$= (a\, I_{d_1})I_{d_2} - (a\, I_{d_2})I_{d_1}$

$= a(I_{d_1} I_{d_2} - I_{d_2} I_{d_1}) = a[I_{d_1}, I_{d_2}]$ .

Hence $I_{[d_1, d_2]} = [I_{d_1}, I_{d_2}]$

It is clear that $a\, I_d{}^p = [a, d^p]$ and that

$a(I_d)^p = \Big[ \ldots\ldots [\overbrace{[a,d],d, \ldots\ldots d}^{p \text{ times}}] \Big]$. Let $d_R$ denote the mapping $a \to ad$ and $d_L$ denote the mapping $a \to da$. Clearly

$I_d = d_R - d_L$ and $a(d_R d_L) = (ad_R)d_L = d(ad) = (da)d = (ad_L)d_R$. We note that $a(I_d)^2 = [(ad - da), d] = ad^2 - 2dad + d^2 a$.

Let us assume that

$$(8') \quad a(I_d)^k = ad^k + \sum_{i=1}^{k-1} \binom{k}{i}(-d)^i ad^{k-i} + (-d)^k a \text{ for all } 1 \leq k \leq n.$$

Then

$$a(I_d)^{n+1} = \left(ad^n + \sum_{i=1}^{n-1} \binom{n}{i}(-d)^i ad^{n-i} + (-d)^n a\right)d$$

$$- d\left(ad^n + \sum_{i=1}^{n-1} \binom{n}{i}(-d)^i ad^{n-i} + (-d)^n a\right)$$

$$= ad^{n+1} + \sum_{j=1}^{n} \left(\binom{n}{j} + \binom{n}{j-1}\right)(-d)^j ad^{n+1-j} + (-d)^{n+1} a$$

$$= ad^{n+1} + \sum_{j=1}^{n} \binom{n+1}{j}(-d)^j ad^{n+1-j} + (-d)^{n+1} a$$

This shows that (8′) holds for all k = 1,2, . . . In particular, if k = p, $\binom{k}{i} = 0$ for each i = 1, 2, . . . , p - 1. Hence $a(I_d)^p = ad^p + (-d)^p a$. We know that p = 2 or p is odd. If p = 2, $y^2 = -y^2$ for all $y \in \mathfrak{A}$. If p is odd, $(-d)^p = -d^p$. Hence $(I_d)^p = I_{d^p}$ or equivalently

$$\Big[\ldots\ldots\big[[a,d],\overbrace{d,\ldots\ldots d}^{p \text{ times}}\big]\Big] = [a,d^p].$$

We may here comment that if $D, D_1$ are elements of $\mathrm{Der}_{\Phi}(\mathfrak{A})$, $\mathfrak{A}$ an algebra over $\Phi$ with characteristic $p \neq 0$, the above method shows that

$$(9) \quad \Big[\ldots\ldots\big[[D_1,D],\overbrace{D,\ldots\ldots D}^{p \text{ times}}\big]\Big] = [D_1,D^p]$$

In this case ($p \neq 0$), Leibniz's rule (7) reduces to

$$(xy)D^p = (xD^p)y + x(yD^p)$$

and so implies that $D^p \in \mathrm{Der}_\Phi(\mathfrak{U})$. Hence $\mathrm{Der}_\Phi(\mathfrak{U})$ is a restricted Lie algebra of characteristic $p \neq 0$.

Remark 1.2. Following upon the results of Exercise 1.1, it can be shown that $\mathfrak{I}(\mathfrak{U})$ the set of inner derivations in $\mathfrak{U}$ is a restricted right ideal in Der $(\mathfrak{U})$. It remains to show that $I_{dD} = [I_d, D]$ is in $\mathfrak{I}(\mathfrak{U})$, where $D \in$ Der $(\mathfrak{U})$, $d \in \mathfrak{U}$.

Proof. We have seen that $I_d = d_R - d_L$. By definition, $(da)D = (dD)a + d(aD)$. Put otherwise, $(da)D - d(aD) = (dD)a$. This can be written in operator form as $[d_L, D] = (dD)_L$. Similarly, $(ad)D - (aD)d = a(dD)$ can be written as $[d_R, D] = (dD)_R$. Hence we obtain the relation

$$(dD)_R - (dD)_L = (d_R D - Dd_R) - (d_L D - Dd_L)$$

$$= (d_R - d_L)D - D(d_R - d_L) = [(d_R - d_L), D]$$

Since $(dD)_R - (dD)_L = I_{dD}$, we now have $I_{dD} = [I_d, D]$.

Remark 1.3. Let c be an element of the centre of $\mathfrak{B}$ (i.e., $cx = xc$ for all $x$ in $\mathfrak{B}$) and $D \in \mathrm{Der}_\Phi(\mathfrak{U}, \mathfrak{B})$. Let $c_R$ denote the mapping $x \to xc$ in $\mathfrak{B}$. Then $\mathrm{Der}_\Phi(\mathfrak{U}, \mathfrak{B})$ is closed under right multiplication by $c_R$.

It is clear that $(x + y)D\ c_R = \Big((x + y)D\Big)c_R = (xD + yD)\, c_R$

$$= x\ D\ c_R + y\ D\ c_R.$$

If $\alpha \in \Phi$, $(x\alpha)\ D\ c_R = \Big((x\alpha)D\Big)\ c_R = \Big((xD)\alpha\Big)\ c$

$$= \Big((xD)c\Big)\ \alpha = (x\ D\ c_R)\ \alpha$$

Hence $D\ c_R$ is linear. If y is an element of $\mathfrak{U}$, we also have

$$(xy)D\, c_R = \Big((xD)y + x(yD)\Big)c_R = (xD)yc + x(yD)c$$

$$= (xD)cy + x(yD)c = (x\ D\ c_R)y + x(y\ D\ c_R)\ .$$

Therefore $D\ c_R$ is an element of $\mathrm{Der}_{\Phi}(\mathfrak{A},\mathfrak{B})$.

This result may be specialized for the case $\mathfrak{A} = \mathfrak{B}$ where $\mathfrak{B}$ is a field P over $\Phi$. Moreover, without specifying which field is the base field of P, we observe that Der (P) is closed under right multiplication by elements $\rho_R$, $\rho \in P$.

In order to give another connection between derivations and homomorphisms, let us construct the so-called algebra of dual numbers. Recall that if $(x^2)$ is the principal ideal generated by $x^2$ over the base field $\Phi$, then $1 + (x^2)$ and $x + (x^2)$ form a basis for the algebra $\Phi[x] / (x^2)$ over $\Phi$. Let us denote the coset $x + (x^2)$ in $\Phi[x] / (x^2)$ by t and set $\mathfrak{C} = \Phi[x] / (x^2)$. Then $\mathfrak{C}$ is an associative algebra with basis (1,t) over $\Phi$ and the multiplication rule $t^2 = 0$. If $\mathfrak{B}$ is an arbitrary (associative) algebra over $\Phi$, form the algebra (= Kronecker or tensor product) $\mathfrak{B} \otimes \mathfrak{C}$ over $\Phi$ (see, e.g. Zariski and Samuel [12], pp. 182-183). Here, multiplication is defined by

$$(b_1 \otimes u_1)(b_2 \otimes u_2) = b_1 b_2 \otimes u_1 u_2\ ,\ b_i \in \mathfrak{B},\ u_i \in \mathfrak{C}\ .$$

In particular, $(b \otimes 1)(1 \otimes t) = b \otimes t$ and $(1 \otimes b)(t \otimes 1) = t \otimes b$. Since $b \otimes 1 = 1 \otimes b$ and $1 \otimes t = t \otimes 1$, we must therefore have $b \otimes t = t \otimes b$. If we identify $\mathfrak{B}$ with the subalgebra of the elements $b \otimes 1$, $b \in \mathfrak{B}$ and identify $\mathfrak{C}$ with the subalgebra of the elements $1 \otimes u$, $u \in \mathfrak{C}$, then the elements of $\mathfrak{B} \otimes \mathfrak{C}$ can be uniquely written as $b_1 + b_2 t$, $b_i \in \mathfrak{B}$. This follows readily from the fact that any element

of $\mathfrak{C}$ can be uniquely written as $\alpha_1 + \alpha_2 t$, $\alpha_i \in \Phi$. In $\mathfrak{B} \otimes \mathfrak{C}$ we now have $bt = tb$ and the multiplication rule

$$(10) \qquad (b_1 + b_2 t)(b_3 + b_4 t) = b_1 b_3 + (b_1 b_4 + b_2 b_3)t \ , \quad b_i \in \mathfrak{B} \ .$$

The algebra $\mathfrak{B} \otimes \mathfrak{C}$ is called the algebra of dual numbers over $\mathfrak{B}$. This construction shows that if $\mathfrak{B}$ is an arbitrary (associative) algebra, then $\mathfrak{B}$ is indeed a subalgebra of an (associative) algebra $\mathfrak{T}$ in which there is an element $t$ such that $t^2 = 0$, $bt = tb$ for all $b \in \mathfrak{B}$, and every element $u \in \mathfrak{T}$ can be uniquely written as $b_1 + b_2 t$, $b_i \in \mathfrak{B}$.

Let $D$ be a derivation of $\mathfrak{A}$ into $\mathfrak{B}$. Define a mapping $s = s(D)$ of $\mathfrak{A}$ into $\mathfrak{B} \otimes \mathfrak{C}$ by

$$(11) \qquad a \to a^s \equiv a + (aD)t$$

Then $(a + b)^s = (a + b) + \Big((a + b)D\Big)t, \quad a,b \in \mathfrak{A}$

$= (a + b) + (aD + bD)t = a + (aD)t + b + (bD)t$

$= a^s + b^s \ ,$

and $(a\alpha)^s = a\alpha + \Big((a\alpha)D\Big)t \qquad , \ \alpha \in \Phi \ ,$

$= a\alpha + \Big((aD)\alpha\Big)t = a\alpha + (aD)t\alpha$

$= \Big(a + (aD)t\Big)\alpha = a^s\alpha$ . Hence the mapping $s$ is linear. Furthermore, we have

$$a^s b^s = \Big(a + (aD)t\Big)\Big(b + (bD)t\Big) = ab + a(bD)t + (aD)tb$$

$$= ab + \Big(a(bD) + (aD)b\Big)t = ab + \Big((ab)D\Big)t = (ab)^s \ .$$

Hence the mapping $s$ is a homomorphism of $\mathfrak{A}$ into $\mathfrak{B} \otimes \mathfrak{C}$. Let us now consider a mapping $\pi$ of $\mathfrak{B} \otimes \mathfrak{C}$ into $\mathfrak{B}$ given by

$$(a + bt) \to (a + bt)^\pi = a, \qquad a,b \in \mathfrak{B} \ .$$

Then $(a + bt)^\pi (c + dt)^\pi = ac = \Big((a + bt)(c + dt)\Big)^\pi$ by rule (10). It

is clear that the mapping $\pi$ is linear. Hence the mapping $\pi$ is a homomorphism of $\mathfrak{B} \otimes \mathfrak{C}$ into $\mathfrak{B}$ which is the identity on $\mathfrak{B}$. In particular, if $a \in \mathfrak{A}$ and the mapping s is defined as in (11), then $a^{s\pi} = \Big(a + (aD)t\Big)^{\pi} = a$. This shows that the mapping s is one-to-one (1 - 1), since $a_1{}^{s} = a_2{}^{s}$ would imply that $a_1{}^{s\pi} = a_2{}^{s\pi}$, that is $a_1 = a_2$ .

Conversely, let s be any homomorphism of $\mathfrak{A}$ into $\mathfrak{B} \otimes \mathfrak{C}$ such that $a^{s\pi} = a$, $a \in \mathfrak{A}$ . Then we have $a^{s} = a + bt$, $a \in \mathfrak{B}$, $b \in \mathfrak{B}$ . The uniqueness of the form $a + bt$ implies that b is uniquely determined by a. Hence, we have the mapping $D: a \to b$ and we may write $a^{s} = a + (aD)t$. We shall now prove that D is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$ .

Proof. Since the mapping s is linear, $(a + c)^{s} = a^{s} + c^{s}$.

Therefore $(a + c) + \Big((a + c)D\Big)t = a^{s} + c^{s} = \Big(a + (aD)t\Big) + \Big(c + (cD)t\Big)$

$$= (a + c) + \Big((aD) + (cD)\Big)t .$$

Hence $(a + c)D = aD + cD$. We also have $(a\alpha)^{s} = a^{s}\alpha$.

Therefore $a\alpha + \Big((a\alpha)D\Big)t = a^{s}\alpha = \Big(a + (aD)t\Big)\alpha$ , $\alpha \in \Phi$

$$= a\alpha + \Big((aD)\alpha\Big)t .$$

We have therefore shown that D is linear. Since $a^{s}c^{s} = (ac)^{s}$ for all $a, c, \in \mathfrak{A}$ ,

$$\Big(a + (aD)t\Big)\Big(c + (cD)t\Big) = ac + \Big((aD)c\Big)t + \Big(a(cD)\Big)t$$

$$= (ac)^{s} = ac + \Big((ac)D\Big)t .$$

Hence $(ac)D = (aD)c + a(cD)$. This completes the proof that D is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$.

We can therefore state the following result (cf. Jacobson [9],

p. 169).

Theorem 1.A. If $\mathfrak{A}$ is a subalgebra of $\mathfrak{B}$ and D is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$, then $s : a \to a + (aD)t$ is an isomorphism of $\mathfrak{A}$ into the algebra of dual numbers $\mathfrak{B} \otimes \mathfrak{C}$ over $\mathfrak{B}$ such that $a^{s\pi} = a$. Conversely, any homomorphism of $\mathfrak{A}$ into $\mathfrak{B} \otimes \mathfrak{C}$ satisfying this condition has the form

$a \to a + (aD)t$ where D is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$.

Here the author [9] gives two consequences of this connection between derivations and isomorphisms. First, if two derivations coincide on a set X of generators of $\mathfrak{A}$, these derivations are identical. Secondly, if s is a homomorphism of $\mathfrak{A}$ into $\mathfrak{B} \otimes \mathfrak{C}$ such that $x^{s\pi} = x$ for $x \in X$, then $a^{s\pi} = a$ for all $a \in \mathfrak{A}$. Hence s defines a derivation D in the manner indicated.

Exercise 1.2. Let $\mathfrak{A}$ be a subalgebra of an algebra $\mathfrak{B}$. Verify that the mapping D of $\mathfrak{A}$ into $\mathfrak{B}$ is a derivation if and only if the mapping

$$s : a \to \begin{bmatrix} a & aD \\ 0 & a \end{bmatrix}$$

of $\mathfrak{A}$ into the matrix algebra $\mathfrak{B}_2$ of $2 \times 2$ matrices over $\mathfrak{B}$ is an isomorphism.

Proof. Suppose that D is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$ and $a,b \in \mathfrak{B}$. We then have

$$(a + b)^s = \begin{bmatrix} (a + b) & (a + b)D \\ 0 & a + b \end{bmatrix} = \begin{bmatrix} a + b & aD + bD \\ 0 & a + b \end{bmatrix}$$

$$= \begin{bmatrix} a & aD \\ 0 & a \end{bmatrix} + \begin{bmatrix} b & bD \\ 0 & b \end{bmatrix} = a^s + b^s ,$$

$$(a\alpha)^s = \begin{bmatrix} a\alpha & (a\alpha)D \\ 0 & a\alpha \end{bmatrix} = \begin{bmatrix} a\alpha & (aD)\alpha \\ 0 & a\alpha \end{bmatrix}$$

$$= \begin{bmatrix} a & aD \\ 0 & a \end{bmatrix} \begin{bmatrix} \alpha \cdot 1 & 0 \\ 0 & \alpha \cdot 1 \end{bmatrix} = a^s \cdot \alpha ,$$

and $$(ab)^s = \begin{bmatrix} ab & (ab)D \\ 0 & ab \end{bmatrix} = \begin{bmatrix} ab & (aD)b + a(bD) \\ 0 & ab \end{bmatrix}$$

$$= \begin{bmatrix} a & aD \\ 0 & a \end{bmatrix} \begin{bmatrix} b & bD \\ 0 & b \end{bmatrix} = a^s b^s$$

We also have

$$\begin{bmatrix} a & aD \\ 0 & a \end{bmatrix} = \begin{bmatrix} b & bD \\ 0 & b \end{bmatrix}$$

if and only if $a = b$. Hence the mapping $s$ is an isomorphism of $\mathfrak{A}$ into $\mathfrak{B}_2$.

Conversely, suppose that $s$ is an isomorphism. It is clear that the linearity of $s$ implies that $D$ is linear. Since we have

$$(ab)^s = \begin{bmatrix} ab & (ab)D \\ 0 & ab \end{bmatrix} = a^s b^s$$

$$= \begin{bmatrix} a & aD \\ 0 & a \end{bmatrix} \begin{bmatrix} b & bD \\ 0 & b \end{bmatrix} = \begin{bmatrix} ab & (aD)b + a(bD) \\ 0 & ab \end{bmatrix} ,$$

we must therefore have $(ab)D = (aD)b + a(bD)$. We have thus proved that $D$ is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$.

Definition 1.7. An element c of a subalgebra $\mathfrak{U}$ of an algebra $\mathfrak{B}$ whose image under a derivation D of $\mathfrak{U}$ into $\mathfrak{B}$ is zero is called a D-constant.

Remark 1.4. The relation (11) implies that an element $c \in \mathfrak{U}$ is a D-constant if and only if $c^s = c$ for the isomorphism $s = s(D)$.

Remark 1.5. The set of D-constants form a subalgebra of $\mathfrak{U}$ with identity $1 \neq 0$.

Proof. It is clear that $1^2 = 1$ implies that $1^2D = (1D)1 + 1(1D) = 1D$. Hence $1D = 0$ for every derivation D of $\mathfrak{U}$ into $\mathfrak{B}$. Let $a,b \in \mathfrak{U}$ be D-constants and $\alpha \in \Phi$. Then $(a \pm b)D = aD \pm bD = 0$, $(a\alpha)D = (aD)\alpha = 0$ and $(ab)D = (aD)b + a(bD) = 0$.

Remark 1.6. If $\mathfrak{B}$ is commutative and $\Phi$ is of characteristic p, then every p-th power in $\mathfrak{U}$ is a D-constant.

Proof. By definition, $a^2 = a(aD) + (aD)a = 2a(aD)$ for all $a \in \mathfrak{U}$. It is clear that $a^3D = (a^2D)a + a^2(aD) = 2a(aD)a + a^2(aD) = 3a^2(aD)$. Let us assume that $a^kD = ka^{k-1}(aD)$ for all $k \leq n$. Then

$$a^{n+1}D = (a^nD)a + a^n(aD) = na^{n-1}(aD)a + a^n(aD)$$
$$= na^n(aD) + a^n(aD) = (n + 1)\, a^n(aD).$$

It now follows by induction that

$$(12) \qquad a^kD = k\, a^{k-1}(aD), \qquad k = 1,2, \ldots .$$

Take $k = p$ and conclude that $a^p$ is a D-constant.

Remark 1.7. If $\mathfrak{U} = P$ is a field over $\Phi$, then the set of D-constants of P form a subfield $\Gamma$ of P which contains $\Phi$. Moreover,

the trivial derivation $D = 0$ is the only derivation on $\Phi$ .

Proof. Since $\alpha D = (1 \cdot \alpha)D = (1D)\alpha = 0$ for all $\alpha \in \Phi$ , $\Phi \subseteq \Gamma$. In view of Remark 1.5, it remains only to show that for all $a \neq 0$ in $\Gamma$, $a^{-1}$ is also in $\Gamma$. This is clear since

$$0 = 1D = (a\, a^{-1})D = (aD)a^{-1} + a(a^{-1}D)$$

implies that

$$(13) \qquad a^{-1}D = -(aD)a^{-2} \text{ , for all } a \in \mathfrak{U} .$$

Hence $a \in \Gamma$ , $a \neq 0$, implies that $a^{-1} \in \Gamma$ . Since $\alpha D = 0$ for all $\alpha \in \Phi$ and for all $D$ in $\mathrm{Der}_{\Phi}(P)$, $D = 0$ is the only derivation on $\Phi$.

Exercise 1.3. Let $D$ be a derivation in $P/\Phi$, $\Gamma$ the subfield of $D$-constants of $P$ over $\Phi$. Prove that the elements $\rho_1, \rho_2, \ldots, \rho_m$ of $P$ are linearly independent over $\Gamma$ if and only if the so-called Wronskian determinant

$$\Delta = \begin{vmatrix} \rho_1 & \rho_2 & \cdots & \rho_r & \cdots & \rho_m \\ \rho_1 D & \rho_2 D & \cdots & \rho_r D & \cdots & \rho_m D \\ \vdots & \vdots & & \vdots & & \vdots \\ \rho_1 D^{r-1} & \rho_2 D^{r-1} & \cdots & \rho_r D^{r-1} & \cdots & \rho_m D^{r-1} \\ \vdots & \vdots & & \vdots & & \vdots \\ \rho_1 D^{m-1} & \rho_2 D^{m-1} & \cdots & \rho_r D^{m-1} & \cdots & \rho_m D^{m-1} \end{vmatrix} = 0$$

In order to prove that this condition holds, we shall use the following lemma.[1]

---

[1] See, e.g., Scott, R. F. and Mathews, G. B., [11], pp. 36, 62-63.

Lemma 1.0. Let $a_{tr}$ denote the element in the t-th row (column) and r-th column (row) of an m × m determinant $(a_{ij})$, $A_{tr}$ denote the minor corresponding to $a_{tr}$, and $\Delta$ denote the value of the determinant. If $\delta_{rs}$ denotes the Kronecker delta ($\delta_{rs} = 1$, if $r = s$, and $\delta_{rs} = 0$, if $r \neq s$), then

$$(14) \qquad \sum_{t} a_{tr} A_{ts} = \delta_{rs} \Delta .$$

Secondly, if

$$M = \begin{vmatrix} A_{ik} & A_{is} & \cdots & A_{if} \\ A_{rk} & A_{rs} & \cdots & A_{rf} \\ \vdots & \vdots & & \vdots \\ A_{tk} & A_{ts} & \cdots & A_{tf} \end{vmatrix}$$

is an h × h determinant and $\Delta^{*}$ is the complementary (m - h) × (m - h) determinant of

$$\Delta^{**} = \begin{vmatrix} a_{ik} & a_{is} & \cdots & a_{if} \\ a_{rk} & a_{rs} & \cdots & a_{rf} \\ \vdots & \vdots & & \vdots \\ a_{tk} & a_{ts} & \cdots & a_{tf} \end{vmatrix}$$

formed from $\Delta$ by deleting the h rows and h columns which contain the elements $\Delta^{**}$, then we have the identity

$$(15) \qquad \Delta^{h} \Delta^{*} = \Delta M$$

We next remark that $\Delta$ can be regarded as a polynomial in the $m^2$ variables $a_{ij}$, $1 \leq i, j \leq m$. Since, for example, the coefficient of

$a_{11}a_{22} \ldots a_{mm}$ is $+1$, $\Delta$ is a polynomial which is not identically zero. Hence the relation (15) is an identity in which each member is a polynomial in the $m^2$ variables $a_{ij}$, $1 \leq i, j \leq m$, with $\Delta \not\equiv 0$. Therefore, we obtain the relation

$$(15') \quad M = \Delta^{h-1}\Delta^* \quad .$$

Proof (Exercise 1.3). We recall that $\rho_1, \rho_2, \ldots \rho_m$ are linearly dependent over $\Gamma$ if and only if there exist $c_1, c_2, \ldots c_m$ in $\Gamma$, not all $c_i = 0$, such that $c_1\rho_1 + c_2\rho_2 + \ldots + c_m\rho_m = 0$. Let us assume that this condition holds. Write $\Delta$ briefly as $(1, D, \ldots, D^{m-1})|\rho_1, \rho_2, \ldots \rho_m|$ and denote the minor corresponding to $\rho_r D^{m-1}$ by the Wronskian $\Delta_r$. If each $\rho_i \in \Gamma$, there is nothing to prove since $\rho_i D^k = 0$ for $k = 1, 2, \ldots, m-1$. We shall assume that not all $\rho_i \in \Gamma$. We next observe that $(c_i\rho_i)D = c_i(\rho_i D)$, $(c_i\rho_i)D^2 = \Big(c_i(\rho_i D)\Big)D = c_i(\rho_i D^2)$. If we assume that $(c_i\rho_i)D^k = c_i(\rho_i D^k)$ for all $k \leq n$, then $(c_i\rho_i)D^{n+1} = \Big((c_i\rho_i)D^n\Big)D = \Big(c_i(\rho_i D^n)\Big)D = c_i(\rho_i D^{n+1})$ implies that $(c_i\rho_i)D^k = c_i(\rho_i D^k)$ for all $k = 1, 2, \ldots$. Hence

$$\begin{aligned}
&c_1\rho_1 + c_2\rho_2 + \ldots + c_m\rho_m = 0 \\
&c_1(\rho_1 D) + c_2(\rho_2 D) + \ldots + c_m(\rho_m D) = 0 \\
&\quad\vdots \\
&c_1(\rho_1 D^{m-1}) + c_2(\rho_2 D^{m-1}) + \ldots + c_m(\rho_m D^{m-1}) = 0
\end{aligned}$$

is a system of $m$ linear equations which has non-trivial solutions in the $c_i$. This shows that $\Delta = 0$.

Conversely, assume that $\Delta = 0$ but one of the Wronskians, say $\Delta_1$, does not vanish. Let us write

$$\Delta_r = \begin{vmatrix} \rho_1 & \cdots & \rho_{r-1} & \rho_{r+1} & \cdots & \rho_m \\ \rho_1 D & \cdots & \rho_{r-1} D & \rho_{r+1} D & \cdots & \rho_m D \\ \vdots & & & & & \\ \rho_1 D^{m-2} & \cdots & \rho_{r-1} D^{m-2} & \rho_{r+1} D^{m-2} & \cdots & \rho_m D^{m-2} \end{vmatrix}$$

as $(1, D, \ldots, D^{m-2}) \left| \rho_1 \rho_2 \ldots \hat{\rho}_r \ldots \rho_m \right|$ . It follows directly from the rule for differentiating a determinant that $\Delta_r D = (1, D, \ldots, D^{m-3}, D^{m-1}) \left| \rho_1 \rho_2 \ldots \hat{\rho}_r \ldots \rho_m \right|$ which is the minor corresponding to $\rho_r D^{m-2}$. Similarly, $\Delta_s D$ is the minor corresponding to $\rho_s D^{m-2}$. From (15′), we see that $\Delta = 0$ implies that

$$(\Delta_r D)\Delta_s - \Delta_r(\Delta_s D)$$

$$= \begin{vmatrix} \Delta_r D & \Delta_s D \\ \Delta_r & \Delta_s \end{vmatrix} = 0 \quad , \quad r,s = 1,2, \ldots, m$$

In particular, $(\Delta_r D)\Delta_1 - \Delta_r(\Delta_1 D) = 0$. We now have

$$\begin{aligned} (\Delta_r \Delta_1^{-1})D &= (\Delta_r D)\Delta_1^{-1} + \Delta_r(\Delta_1^{-1} D) \\ &= (\Delta_r D)\Delta_1^{-1} - \Delta_r(\Delta_1 D)\Delta_1^{-2} \quad , \text{ by (13)}, \\ &= \Big((\Delta_r D)\Delta_1 - \Delta_r(\Delta_1 D)\Big)\Delta_1^{-2} = 0 \ . \end{aligned}$$

Hence $\Delta_r \Delta_1^{-1} = c_r \in \Gamma$ . By (14) above, $\Delta_1 \rho_1 + \Delta_2 \rho_2 + \ldots + \Delta_m \rho_m = 0$. This proves that $\rho_1 + c_2 \rho_2 + \ldots + c_m \rho_m = 0$, which is of the required form.

Finally, we observe that if one of the Wronskians, say $\Delta_1 = 0$ , we can start with $\Delta_1$ as the leading Wronskian and arrive at a particular relation of the form $\rho_2' + c_3 \rho_3' + \ldots + c_m \rho_m' = 0, \quad \rho_i' \in P$.

We now recall two results on extensions of homomorphisms and give reference to corresponding results on extensions of derivations.

(i) Let $\mathfrak{U}$ be a subring (with 1) of a field P, M be a subset of non-zero elements of $\mathfrak{U}$ containing 1 and closed under multiplication, $\mathfrak{U}_M$ the subring of P generated by $\mathfrak{U}$ and the inverses of the elements of M. ($\mathfrak{U}_M = \{ab^{-1}, a \in \mathfrak{U}, b \in M\}$). Let s be a homomorphism of $\mathfrak{U}$ into a field P' such that $\beta^s \neq 0$ for every $\beta \in M$. Then s has a unique extension to a homomorphism S of $\mathfrak{U}_M$ into P'. Moreover, S is an isomsrphism if and only if s is an isomorphism. (cf. Jacobson [9], pp. 2-3).

The corresponding result on derivations is given by the following theorem (Jacobson [9], p. 170).

Theorem 1.B. Let P be a field over $\Phi$, $\mathfrak{U}$ a subalgebra of $P/\Phi$ (containing 1), M a multiplicatively closed subset of non-zero elements of $\mathfrak{U}$ containing 1, and let $\mathfrak{U}_M$ be the subalgebra of P of elements of the form $ab^{-1}$, $a \in \mathfrak{U}$, $b \in M$. Let D be a derivation of $\mathfrak{U}$ into P. Then D can be extended in one and only one way to a derivation of $\mathfrak{U}_M$ into P.

Remark. Here we observe that the isomorphism $a \to a + (aD)t$ of $\mathfrak{U}$ into $P \otimes \mathbb{C}$ the algebra of dual numbers over P has a unique extension to an isomorphism s of $\mathfrak{U}_M$ into $P \otimes \mathbb{C}$ given by

$$(ab^{-1})^s = ab^{-1} + \Big((aD)b^{-1} - ab^{-2}(bD)\Big)t$$

Since $(ab^{-1})D = (aD)b^{-1} - ab^{-2}(bD)$, we can write

$(ab^{-1})^s = ab^{-1} + \Big((ab^{-1})D\Big)t$. It can easily be shown (cf. discussion leading to Theorem 1.A) that the mapping $\theta$ associated with

s of $\mathfrak{U}_M$ into P and given by

(16) $\theta : ab^{-1} \to (aD)b^{-1} - ab^{-2}(bD) \equiv (ab^{-1})D$

is indeed a derivation of $\mathfrak{U}_M$ into P.

Remark. Let $\mathfrak{U}$ be a subalgebra over $\Phi$ of a field $P/\Phi$, D a derivation of $\mathfrak{U}$ into P, $\xi_1, \xi_2, \ldots, \xi_m$ be elements of P. Let $f(x_1, x_2, \ldots, x_m)$ be a polynomial in $\mathfrak{U}[x_1, x_2, \ldots, x_m]$, $x_1, x_2, \ldots, x_m$, indeterminates. If $f^D(x_1, x_2, \ldots, x_m)$ is the polynomial obtained from f by replacing the coefficients by their images under D, then $\theta : f \to f^D(\xi_1, \xi_2, \ldots, \xi_m)$ is a derivation of $\mathfrak{U}[x_1, x_2, \ldots, x_m]$ into P. Secondly, if we define the partial derivative of

$$f = \sum a_{k_1 \ldots k_m} x_1^{k_1} \ldots x_m^{k_m}$$

relative to $x_i$ as

$$\frac{\partial f}{\partial x_i} = \sum k_i \, a_{k_1 \ldots k_m} x_1^{k_1} \ldots x_i^{k_i - 1} \ldots x_m^{k_m},$$

and denote its value at $(\xi_1, \xi_2, \ldots, \xi_m)$ by $\left(\frac{\partial f}{\partial x_i}\right)_{x_j = \xi_j}$,

then $\pi : f \to \left(\frac{\partial f}{\partial x_i}\right)_{x_j = \xi_j}$ is also a derivation of $\mathfrak{U}[x_1, x_2, \ldots, x_m]$ into P.

Proof. In $\mathfrak{U}[x_1, x_2, \ldots x_m]$ addition and multiplication by elements of $\Phi$ are defined coefficientwise in the usual manner. Hence it is clear that the mapping $\theta$ is linear over $\Phi$. It is therefore sufficient to show that $\theta : fg \to f^D g + f g^D$ holds for monomials

$f, g \in \mathfrak{U}[x_1, x_2, \ldots x_m]$. Let $f = \alpha\, x_1^{k_1} \ldots x_m^{k_m}$ and $g = \beta\, x_1^{\ell_1} \ldots x_m^{\ell_m}$, $\alpha\, \beta \in \mathfrak{U}$, $0 \le k_i, \ell_j < p$. We then have

$$f^D g + f g^D = (\alpha D)\, \xi_1^{k_1} \ldots \xi_m^{k_m} (\beta \xi_1^{\ell_1} \ldots \xi_m^{\ell_m}) + \alpha\, \xi_1^{k_1} \ldots \xi_m^{k_m} \left((\beta D)\, \xi_1^{\ell_1} \ldots \xi_m^{\ell_m}\right)$$

$$= \left((\alpha D)\beta + \alpha(\beta D)\right) \xi_1^{r_1} \ldots \xi_m^{r_m}, \quad r_i = k_i + \ell_i ,$$

$$= (fg)^D .$$

This shows that $\theta : fg \to f^D g + f g^D$.

In the second case we have

$$\pi : fg \to \alpha\, \beta\, (k_i + \ell_i)\, \xi_1^{r_1} \ldots \xi_i^{r_i - 1} \ldots \xi_m^{r_m}$$

$$= \alpha\beta\, k_i \xi_1^{r_1} \ldots \xi_i^{r_i - 1} \ldots \xi_m^{r_m} + \alpha\beta\, \ell_i \xi^{r_1} \ldots \xi_i^{r_i - 1} \ldots \xi_m^{r_m}$$

$$= \alpha\, k_i \xi_1^{k_1} \ldots \xi_i^{k_i - 1} \ldots \xi_m^{k_m}\, \beta\, \xi_1^{\ell_1} \ldots \xi_i^{\ell_i} \ldots \xi_m^{\ell_m}$$

$$+ \alpha\, \xi_1^{k_1} \ldots \xi_i^{k_i} \ldots \xi_m^{k_m}\, \beta\, \ell_i \xi_1^{\ell_1} \ldots \xi_i^{\ell_i - 1} \ldots \xi_m^{\ell_m}$$

$$= \left(\frac{\partial f}{\partial x_i}\right)_{x_j = \xi_j} g + f \left(\frac{\partial g}{\partial x_i}\right)_{x_j = \xi_j}$$

Since $\pi$ is clearly linear, $\pi$ is also a derivation.

Finally, we remark that the mapping

$$f \to f^D + \left(\frac{\partial f}{\partial x_i}\right)_{x_j = \xi_j} \eta_i \qquad , \quad \eta_i \in P,$$

is a derivation of $\mathfrak{U}[x_1, \ldots, x_m]$ into P follows immediately from the fact that the set of derivations of $\mathfrak{U}[x_1, \ldots, x_m]$ into P is closed under addition and right multiplication by elements of F.

A second result on homomorphisms is given by

(ii) Let $\mathfrak{B}$ and $\mathfrak{B}'$ be commutative rings, $\mathfrak{A}$ a subring of $\mathfrak{B}'$, and s a homomorphism of $\mathfrak{A}$ into $\mathfrak{B}$. Let X be a set of generators of the ideal $\mathfrak{R}$ of polynomials f in $\mathfrak{A}[x_1, x_2, \ldots, x_r]$, $x_i$ indeterminates, such that $f(t_1, t_2, \ldots t_r) = 0$, $\{t_1, t_2, \ldots t_r\} \subseteq \mathfrak{B}'$. Then there exists a homomorphism S of $\mathfrak{A}[t_1, t_2, \ldots, t_r]$ into $\mathfrak{B}$ such that $a^S = a^s$, $a \in \mathfrak{A}$, and $t_i^S = u_i$, $1 \leq i \leq r$, if and only if $g^s(u_1, \ldots, u_r) = 0$ for every $g \in X$. (Here $g^s$ is obtained from g by replacing the coefficients of $g(x_1, \ldots x_r)$ by their images under s.) If S exists, then it is unique (cf. Jacobson [9], pp. 5-6).

The corresponding result on derivations is given by the following theorem (cf. Jacobson [9], pp. 170-172).

Theorem 1.C. Let $\mathfrak{A}$ be a subalgebra over $\Phi$ of a field $P/\Phi$ and let $\xi_1, \xi_2, \ldots \xi_m, \eta_1, \eta_2, \ldots \eta_m$ be elements of P, D a derivation of $\mathfrak{A}$ into P. Let $\mathfrak{R}$ be the ideal of polynomials $f(x_1, \ldots, x_m)$ in $\mathfrak{A}[x_1, \ldots, x_m]$ such that $f(\xi_1, \ldots \xi_m) = 0$ and let X be any set of generators for $\mathfrak{R}$. Then D can be extended to a derivation D of $\mathfrak{A}[\xi_1, \ldots, \xi_m]$ into P such that $\xi_i D = \eta_i$, $i = 1, 2, \ldots, m$, if and only if

$$(17) \qquad g^D(\xi_1, \ldots \xi_m) + \sum_{i=1}^{m} \left(\frac{\partial g}{\partial x_i}\right)_{x_j=\xi_j} \eta_i = 0$$

for every $g \in X$. If this extension exists, then it is unique.

Remark. The condition (17) can be replaced by

$$(18) \qquad f^D(\xi_1, \ldots, \xi_m) + \sum_{i=1}^{m} \left(\frac{\partial f}{\partial x_i}\right)_{x_j=\xi_j} \eta_i = 0$$

for any $f \in \mathfrak{R}$.

Proof. Let $f_k$, $k = 1, 2, \ldots$, be arbitrary polynomials in $\mathfrak{U}[x_1, \ldots, x_m]$. Then $f = \sum_{k=1}^{r} f_k g_k$, $g_k \in X$. Since D is a derivation of $\mathfrak{U}[x_1, \ldots, x_m]$ into P and every $g_k \in X$ satisfies $g_k(\xi_1, \ldots, \xi_m) = 0$,

$$f^D = \sum_{k=1}^{r} (f_k{}^D g_k + f_k g_k{}^D) = \sum_{k=1}^{r} f_k g_k{}^D .$$

Similarly, $\left(\frac{\partial f}{\partial x_i}\right)_{x_j=\xi_j} = \sum_{k=1}^{r} f_k \left(\frac{\partial g_k}{\partial x_i}\right)_{x_j=\xi_j}$, whence

$\sum_i \left(\frac{\partial f}{\partial x_i}\right)_{x_j=\xi_j} \eta_i = \sum_i \sum_{k=1}^{r} f_k \left(\frac{\partial g_k}{\partial x_i}\right)_{x_j=\xi_j} \eta_i$. We must now have

$$f^D + \sum_i \left(\frac{\partial f}{\partial x_i}\right)_{x_j=\xi_j} \eta_i = f_1 g_1{}^D + f_1 \sum_i \left(\frac{\partial g_1}{\partial x_i}\right)_{x_j=\xi_j} \eta_i + \cdots$$

$$+ f_r g_r{}^D + f_r \sum_i \left(\frac{\partial g_r}{\partial x_i}\right)_{x_j=\xi_j} \eta_i$$

$$= r \cdot 0 = 0 , \text{ by (17).}$$

For the remainder of this section we shall be interested mainly in derivations in a field $P/\Phi$. Suppose that E is a subfield of $P/\Phi$ and D is a derivation of $E/\Phi$ into $P/\Phi$. Let $\xi$ be an element of P. If $\xi$ is transcendental over E, then, as a special case of Theorem 1.C with $\mathfrak{R} = 0$, D can be extended to $E[\xi]$ so that $\xi \to \eta$, any chosen element of P. We may now apply Theorem 1.B and extend D on $E[\xi]$ uniquely to the subfield $E(\xi)$ of P of elements of the form $fg^{-1}$ where $f, g \in E[\xi]$, $g \neq 0$, so that $\xi \to \eta$. Hence we have the following remark.

Remark 1.8. Let E be a subfield of $P/\Phi$ and D a derivation of $E/\Phi$ into $P/\Phi$. Let $\xi \in P$ be transcendental over E and $\eta$ be any element in P. Then D can be extended to $E[\xi]$ so that $\xi \to \eta$. Moreover, D can be extended to the subfield $E(\xi)$ of P such that $\xi \to \eta$.

Next assume that $\xi$ is algebraic over E so that $E[\xi] = E(\xi)$ and let $f(x)$ be the minimal polynomial of $\xi$ over E. Then it is well known (see, e.g. Jacobson [7], p. 100) that the ideal $\mathfrak{R}$ in $E[x]$ of the polynomials $h(x)$ such that $h(\xi) = 0$ is the principal ideal $\big(f(x)\big)$. Hence Theorem 1.C shows that D can be extended to $E(\xi)$ such that $\xi \to \eta$ if and only if

$$f^D(\xi) + f'(\xi)\eta = 0 \tag{19}$$

where $f'(x)$ is the formal derivative of $f(x)$. Recall that an element $\xi$ is separable over E if its minimal polynomial $f(x) \in E[x]$ has no repeated roots in its splitting field (i.e., the field in which it factors linearly). Let us first assume that $\xi$ is separable. Then $f'(\xi) \neq 0$ and (19) gives a unique value of $\eta = -f^D(\xi) \cdot f'(\xi)^{-1}$. In particular, if $D = 0$ on E then $\xi \to 0$ and, in this case, $D = 0$ is the only extension of D to $E(\xi)$. We can now state the following

Remark 1.9. If $E(\xi)$ is separable algebraic over E, then a derivation of $E/\Phi$ into $P/\Phi$ can be uniquely extended to a derivation of $E(\xi)/\Phi$ into $P/\Phi$. If $D = 0$ on E, then $D = 0$ is the only extension of D to $E(\xi)$.

Secondly, we suppose that $\xi$ is inseparable (= not separable) over E. Then $f'(\xi) = 0$ and D can be extended to $E(\xi)$ so that $\xi \to \eta$,

any chosen element of P if and only if $f^D(\xi) = 0$. Let the minimal polynomial for $\xi$ be $f(x) = x^n + a_1 x^{n-1} + \ldots + a_n$, $a_i \in E$. Then $f^D(\xi) = (a_1 D)\xi^{n-1} + \ldots + (a_n D)$. Since the minimal polynomial for $\xi$ is of degree n, $f^D(\xi) = 0$ if and only if each $a_i D = 0$. This proves the following:

Remark 1.10. If $\xi$ is inseparable algebraic over E, then a derivation D of E into P can be extended to $E(\xi)$ so that $\xi \to \eta$, any element of P, if and only if each coefficient of the minimal polynomial of $\xi$ is a D-constant.

In general, it may be pointed out that the condition (19) for the extendibility of D to $E(\xi)$, $\xi$ algebraic over E, is connected to a previous result on homomorphisms given by

(iii) Let $\mathfrak{U}$ and $\mathfrak{B}$ be commutative rings, $\Phi$ a subfield of $\mathfrak{B}$, t an element of $\mathfrak{B}$ which is algebra over $\Phi$, and s an isomorphism of $\Phi$ into $\mathfrak{U}$. Then s can be extended to a homomorphism S of $\Phi[t]$ into $\mathfrak{U}$ so that $t^S = u$, if and only if $f^s(u) = 0$ for the minimal polynomial f(x) of t over $\Phi$. When the extension exists it is unique (cf. Jacobson [9], p. 6).

Here we observe that in the case of extendibility of D to E(t), uniqueness is achieved only for t separable (algebraic) over E.

Now let $P = \Phi(\xi_1, \xi_2, \ldots, \xi_m)$ be a finitely generated extension field of $\Phi$. Let $\mathfrak{R}$ be the ideal in $\Phi[x_1, \ldots, x_m]$ of polynomials $f(x_1, \ldots, x_m)$ such that $f(\xi_1, \ldots, \xi_m) = 0$ and let X be a basis for $\mathfrak{R}$. If D is a derivation of the polynomial algebra

$\Phi[\xi_1, \ldots, \xi_m]$ over $\Phi$ into P over $\Phi$, then D has a unique extension to P over $\Phi$ (cf. Theorem 1.B). We have seen (cf. Remark 1.7) that $D = 0$ is the only derivation on $\Phi$. Theorem 1.C applied to $D = 0$ on $\Phi$ shows that there exists a derivation D of $\Phi[\xi_1, \ldots, \xi_m]$ over $\Phi$ into P over $\Phi$, and hence $\Phi(\xi_1, \ldots, \xi_m) = P$ into itself, such that $\xi_i D = \eta_i$, $i = 1, 2, \ldots, m$, if and only if

$$(20) \qquad \sum_i \left(\frac{\partial g}{\partial x_i}\right)_{x_j=\xi_j} \eta_i = 0$$

for every $g \in X$. By the relation (18), we can replace (20) by

$$(21) \qquad \sum_i \left(\frac{\partial f}{\partial x_i}\right)_{x_j=\xi_j} \eta_i = 0$$

for any $f \in \mathfrak{R}$. Moreover, we have the following criterion.

Lemma 1.1. Let $P = \Phi(\xi_1, \ldots \xi_r, \ldots \xi_m)$ be a finitely generated extension field of $\Phi$. Then 0 is the only derivation of P into itself if and only if P is separable algebraic over $\Phi$.

Proof. Assume that P is separable algebraic over $\Phi$. We have seen (cf. Remark 1.7) that 0 is the only derivation on $\Phi$. By Remark 1.9, if $\xi_1$ is separable algebraic over $\Phi$, then $D = 0$ is the only derivation on $\Phi(\xi_1)$. In the same way, if $\xi_2 \in P$, $\xi_2 \notin \Phi(\xi_1)$, then $\xi_2$ is separable algebraic over $\Phi$ and so over $\Phi(\xi_1)$, and we now have $D = 0$ is the only derivation on $\Phi(\xi_1, \xi_2)$. We can repeat this process for a finite number (=m) of times and thus obtain that $D = 0$ is the only derivation of $\Phi(\xi_1, \ldots, \xi_m) = P$ into itself. Conversely, let $r < m$ be the largest integer such that P is not separable algebraic

over $\Phi(\xi_1, \circ\ \circ\ \circ, \xi_r) = \Phi'$. Then $\Phi'(\xi_{r+1})$ is either transcendental over $\Phi'$or is inseparable over $\Phi'$ by the transitivity of algebraic separability. In either case we have observed that the zero derivation on $\Phi'$ can be extended to a non-zero derivation D on $\Phi'(\xi_{r+1})$ such that $\xi_{r+1}D = \eta$, an arbitrary element of P. Since, by assumption, P is separable over $\Phi(\xi_1, \circ\ \circ\ \circ, \xi_{r+1})$ we can extend this derivation to a non-zero derivation of P into itself (cf. Remark 1.9). Hence, 0 is the only derivation if P is separable algebraic over $\Phi$.

Exercise 1.4. Let $P = \Phi(\xi_1, \ldots, \xi_m)$. Show that P is separable algebraic if and only if there exist m polynomials $g_1(x_1, \circ\ \circ\ x_m), \ldots, g_m(x_1, \circ\ \circ\ \circ, x_m)$ in $\Phi[x_1, \ldots, x_m]$ such that $0 = g_1(\xi_1, \ldots, \xi_m) = \ldots = g_m(\xi_1, \circ\ \circ\ \circ, \xi_m)$ and the Jacobian

$$J = \det\left(\left(\frac{\partial g_i}{\partial x_j}\right)_{x_k=\xi_k}\right) \neq 0 .$$

Proof. By Lemma 1.1, P is separable algebraic over $\Phi$ if and only if D = 0 is the only derivation of P into itself. By (21), this is equivalent to

$$\sum_{j=1}^{m}\left(\frac{\partial g_i}{\partial x_j}\right)_{x_k=\xi_k}(\xi_j D) = 0, \quad i = 1, 2, \circ\ \circ\ \circ, m,$$

where this system of m homogeneous linear equations has only trivial solutions $(\xi_j D) = 0$ for all $j = 1, 2, \ldots, m$. Such solutions are possible if and only if $J \neq 0$.

We have seen that the system $\mathrm{Der}_{\Phi}(P,P) = \mathrm{Der}_{\Phi}(P)$ of derivations in a field $P/\Phi$ is a Lie algebra of linear transformations which is restricted if the characteristic of $\Phi$ is $p \neq 0$, and that $\mathrm{Der}_{\Phi}(P)$ is

closed under right multiplication by $\rho_R$, $\rho \in P$. If $D \in Der(P)$, then define $D\rho_R = D\rho$. From a discussion given in Jacobson [9], pp. 19-20, it is seen that $\mathfrak{L}(P,P) = \mathfrak{L}(P)$, the system of linear transformations in P, with $A\rho_R = A\rho$, $A \in \mathfrak{L}(P)$, $\rho \in P$, is a right vector space over P. Hence, $Der_{\bar{\Phi}}(P)$ is a subspace of the right vector space $\mathfrak{L}_{\bar{\Phi}}(P)$ over P. Let us denote the dimensionality of $Der_{\bar{\Phi}}(P)$ over P by $[Der_{\bar{\Phi}}(P):P]_R$. Then another connection between derivations and separable algebraic extensions is given by the following theorem.

Theorem 1.D. If $P = \bar{\Phi}(\xi_1, \ldots, \xi_m)$, then $[Der_{\bar{\Phi}}(P):P]_R$ is the smallest rational integer s such that there exists a subset $S = \{\xi_{i_1}, \ldots, \xi_{i_s}\}$ of $\{\xi_1, \ldots, \xi_m\}$ such that P is separable algebraic over $\bar{\Phi}(S)$. (Jacobson [9], pp. 178-179)

Proof: Consider the mapping $D \overset{\theta}{\to} (\xi_1 D, \ldots, \xi_m D)$ of $Der_{\bar{\Phi}}(P)$ into $P^{(m)}$, the right vector space of m-tuples $(\rho_1, \ldots, \rho_m)$, $\rho_j \in P$. In $P^{(m)}$, equality, addition, and scalar multiplication is defined coordinate-wise in the usual manner. Let D, $D_1$ be elements of $Der_{\bar{\Phi}}(P)$. Then $(\xi_1 D, \ldots, \xi_m D) + (\xi_1 D_1, \ldots, \xi_m D_1)$ $= \left(\xi_1(D + D_1), \ldots, \xi_m(D + D_1)\right)$. Hence $D^\theta + D_1^\theta = (D + D_1)^\theta$. We also have $(\xi_1 D\rho, \ldots, \xi_m D\rho) = (\xi_1 D, \ldots, \xi_m D)\rho$. Hence $(D\rho)^\theta = D^\theta\rho$. We now suppose that $\xi_i D = 0$, $i = 1, 2, \ldots, m$. In this case $D = 0$ since the $\xi_i$ generate P over $\bar{\Phi}$. This shows that the kernel of this mapping is 0. This mapping is therefore P-linear and one-one. Let $(D_1, \ldots, D_s)$ be a right basis of $Der_{\bar{\Phi}}(P)$ over P. Then $s \leq m$ and the image of $Der_{\bar{\Phi}}(P)$ in $P^{(m)}$ has the basis $(\xi_1 D_j, \ldots, \xi_m D_j)$, $1 \leq j \leq s$. We now see that the rank of the

$s \times m$ matrix $(\xi_i D_j)$ is $s$, so we can re-order the $\xi_i$'s so that $\det(\xi_i D_j) \neq 0$, $1 \leq i, j \leq s$. We now set $E = \Phi(\xi_1, \ldots, \xi_s)$ and let D be a derivation of $P/E$ into itself. Then D is also an element of $\mathrm{Der}_\Phi(P)$ and so

$$D = \sum_{j=1}^{s} D_j \rho_j , \quad \rho_j \in P.$$

Since $\xi_1, \ldots, \xi_s$ belong to E, we also have

$$0 = \xi_i D = \sum_{j=1}^{s} (\xi_i D_j)\rho_j \quad \text{for } i = 1, 2, \ldots, s.$$

Since $\det(\xi_i D_j) \neq 0$, then every $\rho_j = 0$ and hence $D = 0$. Therefore, by Lemma 1.1, P is separable algebraic over $E = \Phi(\xi_1, \ldots, \xi_s)$. Next suppose that there is a subset $\{\xi_1, \ldots, \xi_t\}$ of the $\xi_i$'s (re-ordered appropriately) such that P is separable algebraic over $\Phi(\xi_1, \ldots, \xi_t)$. We now use these $\xi$'s to map $\mathrm{Der}_\Phi(P)$ into $P^{(t)}$ by means of $D \to (\xi_1 D, \ldots, \xi_t D)$. It is clear that this map is P-linear. If $\xi_1 D = 0 = \ldots = \xi_t D$, then D maps $\Phi(\xi_1, \ldots \xi_t)$ into 0 and so D is a derivation of $P/\Phi(\xi_1, \ldots, \xi_t)$ into itself. Since P is separable algebraic over $\Phi(\xi_1, \ldots, \xi_t)$, $D = 0$. This map is therefore P-linear and one-one. Hence $s = [\mathrm{Der}_\Phi(P):P]_R \leq t$.

## SECTION 2

Definition 2.1. If the only elements of an algebraic extension $P/\Phi$ which are separable are the elements of $\Phi$, then P is said to be purely inseparable over $\Phi$. Similarly, an element $\rho$ is said to be inseparable over $\Phi$ if $\Phi(\rho)/\Phi$ is purely inseparable.

Remark 2.1. An element $\rho \in P/\Phi$ is purely inseparable over $\Phi$ if and only if $\rho$ is a root of the polynomial $x^{\rho^e} - a$, $a \in \Phi$, e some non-negative rational integer (cf. Jacobson [9], p. 48). It is clear that if $e = 0$, $\rho \in \Phi$.

Definition 2.2. We shall call e the exponent of the purely inseparable element $\rho$. If there exists a maximum k for the exponents of the elements of P, we say that P is of exponent k over $\Phi$; otherwise, the exponent of $P/\Phi$ is infinite.

Let $P/\Phi$ be purely inseparable of exponent $e = 1$. If $\rho, \sigma \in P$, then $\rho^p = a \in \Phi$ and $\sigma^p = b \in \Phi$. We recall that $(\rho \pm \sigma)^p = \rho^p \pm \sigma^p$. Hence $(\rho \pm \sigma)^p = a \pm b \in \Phi$. We also see that if $\rho \neq 0$, $(\rho^{-1})^p = a^{-1} \in \Phi$. Therefore the p-th powers of P form a subfield $P^p$ of P over $\Phi$. It now follows that P is purely inseparable of exponent $\leq 1$ over $\Phi$ if and only if $P^p \subseteq \Phi$. Hence we may also say that any field P is purely inseparable of exponent $\leq 1$ over $\Phi(P^p)$.

Remark 2.2. If $P/\Phi$ is a field of characteristic $p \neq 0$ and $E/\Phi$ is a subfield of $P/\Phi$, then D is a derivation of $E/\Phi$ into $P/\Phi$ if and only if D is a derivation of $E/\Phi(E^p)$ into $P/\Phi(E^p)$.

Proof. We recall that if D is a derivation of $E/\Phi$ into $P/\Phi$, then $a^pD = pa^{p-1}(aD) = 0$ for all $a \in E$, and that the set of D-constants

forms a subfield $\Gamma$ of E over $\Phi$. Hence $\Gamma \supseteq \Phi(E^p)$. Moreover, if $c \in \Gamma$, $a \in E$, then $(ac)D = a(cD) + (aD)c = (aD)c$. This shows that D is also a derivation of $E/\Gamma$ into $P/\Gamma$. Since $\Gamma \supseteq \Phi(E^p)$, it follows that a derivation of $E/\Phi$ into $P/\Phi$ is also a derivation of $E/\Phi(E^p)$ into $P/\Phi(E^p)$ and the converse is immediate.

Since E is purely inseparable of exponent $\leq 1$ over $\Phi(E^p)$, in discussing derivations of $E/\Phi$ into $P/\Phi$ we can consider only those subfields $E/\Phi$ which are purely inseparable of exponent $\leq 1$.

Remark 2.3. If $\Phi$ is a field of characteristic $p \neq 0$, then $x^p - a$, $a \in \Phi$ is irreducible in $\Phi[x]$ unless $a = b^p$, $b \in \Phi$, in which case $x^p - a = (x - b)^p$ in $\Phi[x]$.

Proof. Suppose $x^p - a$ factors linearly in K[x], K a field over $\Phi$. If b is a root of $x^p - a$, then $x^p - a = x^p - b^p = (x - b)^p$ in K[x]. Assume that $x^p - a = f(x)\, g(x)$ in $\Phi[x]$ where the degree of f(x) is k, $1 \leq k < p$. Then in K[x] we must have $f(x) = (x - b)^k$ since factorization in K[x] is unique. But we have

$$(x - b)^k = x^k - kbx + \ldots + (-b)^k.$$

This implies that $kb \in \Phi$ and consequently $b \in \Phi$. In this case we now have $x^p - a = (x - b)^p$ in $\Phi[x]$.

Remark 2.4. Let $P = \Phi(\xi_1, \xi_2, \ldots \xi_m)$ be a finitely generated extension field of $\Phi$ with characteristic $p \neq 0$, and $\xi_i \in P$, $\xi_i \notin \Phi(\xi_1, \ldots, \xi_{i-1})$, satisfy the minimal polynomial $x^p - a_i$, $a_i \in \Phi$. Then $[P : \Phi] = p^m < \infty$.

Proof. We recall that since $x^p - a_1$ is irreducible in $\Phi[x]$ (see Remark 2.3), $[\Phi(\xi_1) : \Phi] = p$. Now $\xi_2 \in P$, $\xi_2 \notin \Phi(\xi_1)$ implies that $x^p - a_2$ is irreducible in $\Phi(\xi_1)[x]$. For otherwise, $a_2 = \beta^p$, $\beta \in \Phi(\xi_1)$ and hence $\xi_2 = \beta \in \Phi(\xi_1)$ which is impossible. It now follows, a fortiori, that $x^p - a_2$ is also irreducible in $\Phi[x]$. Hence

$$[\Phi(\xi_1,\xi_2) : \Phi] = [\Phi(\xi_1, \xi_2) : \Phi(\xi_1)][\Phi(\xi_1) : \Phi] = p^2$$

(See, e.g., Herstein [3], p. 168). This process terminates after m applications and we now have $[P:\Phi] = p^m < \infty$.

It will be seen later that the theory of derivations in the case of characteristic $p \neq 0$ is connected with the study of purely inseparable extensions of exponent $e = 1$. In order to show this, we shall introduce a special kind of dependence relation.

Definition 2.3. An element $\rho \in P$ is said to be p-dependent in P over $\Phi$ on the subset S of P if $\rho \in \Phi^*(S)$ where $\Phi^* = \Phi(P^p)$. We indicate this relation by $\rho <_p S$ .

Accordingly, we call a finite subset S of P p-independent if $\sigma \not<_p S - \{\sigma\}$ for all $\sigma$ in S, and say that any arbitrary subset S of an arbitrary field P is p-independent if every finite subset F of S is p-independent.

Definition 2.4. A p-independent subset B of P over $\Phi$ which is such that every element of P is p-dependent on B is called a <u>p-basis</u> for P over $\Phi$. In this case, $P = \Phi^*(B)$.

We next recall the following theorem for an arbitrary set P and

a set of four axioms for a generalized dependence relation.

I If $\sigma \in S$, then $\sigma < S$ (= $\sigma$ depends on S).

II If $\rho < S$, then $\rho < F$ for some finite subset F of S.

III If $\rho < S$ and every $\sigma$ in S satisfies $\sigma < T$, then $\rho < T$.

IV If $\rho < S$ and $\rho \not< S - \{\sigma\}$ where $\sigma \in S$, then $\sigma < \left(S - \{\sigma\} \cup \{\rho\}\right)$.

Basis Theorem. The set P has a basis. Moreover, any two bases have the same cardinal number (Jacobson [9], pp. 154-155).

We next show that p-dependence satisfies the conditions of the given axioms.

First, if $\sigma \in S$, then $\sigma \in \Phi^*(S)$ or equivalently $\sigma <_p S$. Secondly, since $S = \cup F$, the union of all finite subsets F of S, $\rho \in \Phi^*(S)$ implies that $\rho \in \Phi^*(F)$ for some finite subset F of S. Hence, if $\rho <_p S$, then $\rho <_p F$. Next, if $\rho \in \Phi^*(S)$ and every $\sigma$ in S satisfies $\sigma \in \Phi^*(T)$, then $\rho \in \Phi^*(T)$. That is, if $\rho <_p S$ and every $\sigma$ in S satisfies $\sigma <_p T$, then $\rho <_p T$. We next suppose that $\rho \in \Phi^*(S)$, $\rho \notin \Phi^*(S - \{\sigma\})$ where $\sigma \in S$, and write $S - \{\sigma\} = T$. It is now clear that $\Phi^*(T)$ is a proper subfield of $\Phi^*(T,\sigma) = \Phi^*(T,\rho,\sigma)$ and of $\Phi^*(T,\rho)$. It is also clear that $\sigma \in \Phi^*(T,\sigma)$, $\sigma \notin \Phi^*(T)$ and that $\rho \in \Phi^*(T,\rho)$, $\rho \notin \Phi^*(T)$ by assumption. From the discussion with regards to Remark 2.4, we now have

$$[\Phi^*(T,\sigma):\Phi^*(T)] = [\Phi^*(T,\rho,\sigma):\Phi^*(T)] = p$$
$$= [\Phi^*(T,\rho):\Phi^*(T)]$$

We assert that $\sigma \in \Phi^*(T,\rho)$. For otherwise $\sigma \in \Phi^*(T,\rho,\sigma)$ and $\sigma \notin \Phi^*(T,\rho)$ would imply that $[\Phi^*(T,\rho,\sigma): \Phi^*(T,\rho)] = p$, which is impossible since $[\Phi^*(T,\rho,\sigma): \Phi^*(T)] = [\Phi^*(T,\rho): \Phi^*(T)] = p$. Hence,

if $\rho <_p S$ and $\rho \not<_p S - \{\sigma\}$ for some $\sigma$ in S, then $\sigma <_p (S - \{\sigma\} \cup \{\rho\})$. We can now apply the Basis Theorem and state the following:

Let P be a field of characteristic $p \neq 0$, P purely inseparable (algebraic) over $\Phi$ of exponent one. Then P has a p-basis. Moreover, any two p-bases have the same cardinal number.

If $F = \{\rho_1, \rho_2, \ldots, \rho_m\}$ is a p-independent set, then $\rho_i \notin \Phi^*(\rho_1, \rho_2, \ldots, \rho_{i-1})$ and $\rho_i^p = \alpha_i \in \Phi^*$. As we have seen earlier, we now have $[\Phi^*(\rho_1, \rho_2, \ldots, \rho_m): \Phi^*] = p^m$. Hence, the $p^m$ elements

$$\rho_1^{k_1}\rho_2^{k_2} \ldots \rho_m^{k_m}, \quad 0 \leq k_i < p, \tag{22}$$

form a basis for $\Phi^*(\rho_1, \ldots, \rho_m)$ over $\Phi^*$. Conversely, we suppose that the $p^m$ elements given by (22) form a basis for $\Phi^*(\rho_1, \ldots, \rho_m)$ over $\Phi^*$. Then $[\Phi^*(\rho_1, \ldots, \rho_m): \Phi^*] = p^m$. If, for example, $\rho_{i_o} \in \Phi^*(\rho_1, \ldots, \rho_{i_o-1})$, then we have $[\Phi^*(\rho_1, \ldots, \rho_m): \Phi^*] = [\Phi^*(\rho_1, \ldots, \rho_m): \Phi^*(\rho_1, \ldots, \rho_{i_o})][\Phi^*(\rho_1, \ldots, \rho_{i_o-1}): \Phi^*] \leq p^{m-i_o} \cdot p^{i_o-1} = p^{m-1}$ which is a contradiction. Therefore the $\rho_i$ are p-independent.

Remark 2.5. A maximal p-independent set B is necessarily a p-basis.

Proof. We first assert that if $\eta \in P$ satisfies $\eta \not<_p B$, then $B \cup \{\eta\}$ is p-independent. Otherwise we must have an $\xi \in B$ such that $\xi <_p (B \cup \{\eta\} - \{\xi\})$. Since $\xi \not<_p (B - \{\xi\}) = (B \cup \{\eta\} - \{\xi\} - \{\eta\})$, then Axiom IV implies that $\eta <_p B = (B - \{\xi\} \cup \{\xi\})$ which contradicts

the hypothesis of our assertion. This shows that our assertion holds. In particular, if B is maximal, we must therefore have $\rho \in P$ satisfies $\rho <_p B$ for all $\rho \in P$. Hence B is a p-basis.

Theorem 2.A. Let P be an arbitrary field of characteristic $p \neq 0$, $\Phi$ a subfield and E an intermediate field. Let B be a p-basis of E over $\Phi$. Let $\delta$ be an arbitrary mapping of B into P. Then there exists one and only one derivation D of $E/\Phi$ into $P/\Phi$ such that $\beta D = \delta(\beta)$ for every $\beta \in B$.

Proof. We have seen in Remark 2.2 that we can assume that E is purely inseparable of exponent $\leq 1$ over $\Phi$. Let us assume that $E \supset \Phi$ (that is, E properly contains $\Phi$) which means that B is non-empty and the exponent of $E/\Phi$ is one. Let $\beta \in B$ and set $B_\beta = B - \{\beta\}$. Then $\beta \notin \Phi(B_\beta)$ and its minimal polynomial over $\Phi(B_\beta)$ is $x^p - b$. We have seen that 0 is the only derivation on $\Phi(B_\beta)/\Phi(B_\beta)$. Hence, by Remark 1.10, there exists a derivation $D_\beta$ (relative to $\beta$) of $E = \Phi(B_\beta, \beta)$ over $\Phi(B_\beta)$ into P over $\Phi(B_\beta)$ such that $\beta D_\beta = \delta(\beta)$. Note that $\beta' D_\beta = 0$, $\beta' \neq \beta$, $\beta' \in B$. Since $\Phi(B_\beta) \supseteq \Phi$, a derivation D of E over $\Phi(B_\beta)$ into P over $\Phi(B_\beta)$ is also a derivation D of $E/\Phi$ into $P/\Phi$. Let F be a finite subset $\{\beta_1, \ldots, \beta_r\}$ of B, then $D_F = D_{\beta_1} + \ldots + D_{\beta_r}$ is a derivation of $E/\Phi$ into $P/\Phi$ such that $\beta_i D_F = \delta(\beta_i)$, and $\beta' D_F = 0$ if $\beta' \in B - \{\beta_1, \ldots, \beta_r\}$, $i = 1,2, \ldots, r$.

If G is any finite subset of B containing F, then the restriction of $D_G$ to $\Phi(F)$ coincides with the restriction of $D_F$ to $\Phi(F)$. If $\xi$ is an arbitrary element of E, we can choose a finite subset F of B such that $\xi \in \Phi(F)$ and map $\xi \to \xi D_F$. Since the $p^r$ elements $\beta_1^{k_1} \ldots \beta_r^{k_r}$,

$0 \leq k_i < p$, form a basis for $\Phi(F)$ over $\Phi = \Phi(F^p)$, the form

$$\xi = \sum a_{ik_1 \ldots k_r} \beta_1^{k_1} \ldots \beta_r^{k_r} \quad ,$$

$a_{ik_1 \ldots k_r} \in \Phi$, is unique for $\xi \in \Phi(F)$. It is clear that there exists a smallest finite subset $F_o$ of B such that $\xi \in \Phi(F_o)$. Hence $\xi D_F = \xi D_{F_o}$ is the same for any finite subset F of B such that $\xi \in \Phi(F)$. Therefore the mapping $D: \xi \to \xi D_F$ is single valued. Clearly this mapping is a derivation of $E/\Phi$ into $P/\Phi$ since $D_F$ is a derivation of $E/\Phi$ into $P/\Phi$. Moreover, $\beta D = \beta D_F = \delta(\beta)$ for every $\beta \in B$. Since $E = \Phi(B)$, D is unique.

Corollary 1. If E has a finite p-basis B, then

$$[\mathrm{Der}_{\Phi}(E,P): P]_R = |B|.$$

Proof. Let $\Delta(B,P)$ be the set of mappings of B into P which we can consider as a right vector space over P by defining $(\delta_1 + \delta_2)(\beta) = \delta_1(\beta) + \delta_2(\beta)$ and $(\delta\rho)(\beta) = \delta(\beta)\rho$, $\delta, \delta_1, \delta_2 \in \Delta(B,P)$, $\beta \in B$, $\rho \in P$. We now map $\mathrm{Der}_{\Phi}(E,P)$ into $\Delta(B,P)$ by sending $D \in \mathrm{Der}_{\Phi}(E,P)$ into its restriction $\delta$ to B. Let $D, D_1, D_2 \in \mathrm{Der}_{\Phi}(E,P)$. Then $\beta(D_1 + D_2) = \beta D_1 + \beta D_2 = \delta_1(\beta) + \delta_2(\beta) = (\delta_1 + \delta_2)(\beta)$. We also have $\beta(D\rho) = (\beta D)\rho = \delta(\beta)\rho = (\delta\rho)(\beta)$. Hence this map is P-linear. It is clear that the kernel of this map is $D = 0$. Moreover, the theorem shows that this map is onto. This map is therefore an isomorphism. If $B = \{\beta_1, \ldots, \beta_m\}$ is finite, then the m-mappings $\delta_1, \ldots, \delta_m$ such that $\delta_i(\beta_j) = \delta_{ij}$ (the Kronecker delta) form a basis for $[\Delta(B,P): P]_R$. This follows immediately since $\det\left(\delta_i(\beta_j)\right) = 1 \neq 0$ implies that the system

$$(\delta_1\rho_1 + \ldots + \delta_m\rho_m)(\beta_j) = 0, \quad 1 \leq j \leq m, \text{ of m linear}$$

homogeneous equations

$$\delta_1(\beta_1)\rho_1 + \cdots + \delta_m(\beta_1)\rho_m = 0$$
$$\vdots$$
$$\delta_1(\beta_m)\rho_m + \cdots + \delta_m(\beta_m)\rho_m = 0$$

has only trivial solutions $0 = \rho_1 = \cdots = \rho_m$. This means that the $\delta_i$ are right linearly independent over P. It is also clear that the only mapping that sends B into $0 \in P$ is the zero mapping (= the restriction of $D = 0$ on $\Phi(B)$ to B). Moreover, if $\rho_k$ is an arbitrary element of P, $(\delta_i\rho_k)(\beta_j) = \delta_{ij}\rho_k$. Hence any element of $\Delta(B,P)$ has the form $\delta_1\rho_1 + \cdots + \delta_m\rho_m$, $\rho_i \in P$. This means that $(\delta_1, \cdots, \delta_m)$ is a right basis for $\Delta(B,P)$. Therefore $[\Delta(B,P): P]_R = [\mathrm{Der}_{\bar{\Phi}}(E,P):P]_R = |B|$.

Corollary 2. Every derivation of $E/\bar{\Phi}$ into $P/\bar{\Phi}$ can be extended to a derivation of $P/\bar{\Phi}$ if and only if the elements of any p-basis B of $E/\bar{\Phi}$ are p-independent in $P/\bar{\Phi}$.

Proof. Suppose the elements of any p-basis B of $E/\bar{\Phi}$ are p-independent in $P/\bar{\Phi}$. Then B can be imbedded in a p-basis C of $P/\bar{\Phi}$. If D is a derivation of $E/\bar{\Phi}$ into $P/\bar{\Phi}$, then the restriction $\delta_B$ of D to B can be extended to a mapping $\delta_C$ (= the restriction of D to C) of C into P. By the theorem we now have a unique derivation $D'$ of $\Phi(C) = P$ over $\bar{\Phi}$ into itself such that $\gamma D' = \delta_C(\gamma)$ for all $\gamma \in C$. Since $\beta D' = \delta_C(\beta) = \delta_B(\beta) = \beta D$ for every $\beta \in B$, $D'$ is an extension of D. Conversely, suppose B is not p-independent in $P/\bar{\Phi}$ and let $\beta$ be an element of B which is p-dependent in P on $B_\beta = B - \{\beta\}$. If $D''$ is any derivation in P such that $\mu D'' = 0$ for all $\mu \in B_\beta$, since every p-th power

of P is a $D''$-constant and $\beta \in \Phi(P^p, B_\beta)$, $\beta D'' = 0$. We have seen in the proof of the theorem that there is a derivation $D_\beta$ of $E/\Phi$ into $P/\Phi$ such that $\beta D_\beta \neq 0$ and $\mu D_\beta = 0$, $\mu \in B_\beta$. Clearly, such a derivation cannot be extended to P since $\beta D'' = 0$ and $\beta D_\beta \neq 0$ is impossible.

We shall next take $\Phi$ to be the prime field $\Phi_0$ which can be identified with the field of rational integers modulo p. We have seen that for all rational integers x, $x^p \equiv x \pmod{p}$ (Fermat's theorem). Hence we may write $\Phi_0(E^p) = E^p$. We shall agree to refer to a derivation of $E/\Phi_0$ into $P/\Phi_0$ as a derivation of E into P. Our object is to show that the condition given in Corollary 2 is equivalent to separability, in the general sense, of P/E. We recall that an extension field A of a field P is called the algebraic closure (up to an isomorphism) of P if: (1) A is algebraic over P, and (2) every polynomial $f(x) \in A[x]$ of positive degree can be written as a product of linear factors in A[x]. We take $E^{p^{-1}} = \{\gamma \in A : \gamma^p \in E\}$.

Definition 2.5. Let X, Y be two subspaces of a vector space S over a field E. Then X and Y are said to be linearly disjoint over E if the following condition is satisfied: whenever $x_1, \ldots, x_n$ are elements of X which are linearly independent over E and $y_1, \ldots, y_m$ are elements of Y which are linearly independent over E, then the mn products $x_i y_j$ are also linearly independent over E.

Remark 2.6. The following property is equivalent to linear disjointness: whenever $x_1, x_2, \ldots, x_n$ are elements of X which are linearly independent over E then these elements $x_1, x_2, \ldots, x_n$ are linearly independent over Y.

Proof. Assume that X and Y are linearly disjoint over E. Let $x_1, x_2, \ldots, x_n$ be elements of X which are linearly independent over E and suppose that

$$\sum_{i=1}^{n} \alpha_i x_i = 0, \quad \alpha_i \in Y .$$

If $\{y_i : i \in I\}$ is a basis for Y over E, then $\alpha_i = \sum_j c_{ij} y_j$, $c_{ij} \in E$. We now have $0 = \sum_j c_{1j} y_j x_1 + \ldots + \sum_j c_{nj} y_j x_n = \sum_{i,j} c_{ij} x_i y_j$. Since X and Y are assumed to be linearly disjoint over E, we must have the $c_{ij} = 0$ and hence the $\alpha_i = 0$, showing that the condition is satisfied. Conversely, suppose that the condition is satisfied. Let the two sets $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_m\}$ be as in Definition 2.5. Assume that $\sum_{i,j} c_{ij} x_i y_j = 0$, $c_{ij} \in E$. Since the condition is satisfied, the the $x_i$ are linearly independent over Y. Hence

$$\sum_j c_{1j} y_j = 0 = \sum_j c_{2j} y_j = \ldots = \sum_j c_{nj} y_j .$$

Since the $y_j$ are linearly independent over E, this means that every $c_{ij} = 0$. Put otherwise, X and Y are linearly disjoint over E.

Definition 2.6. A field P is separable (not necessarily algebraic) over a field E of characteristic $p \neq 0$ if P and $E^{p^{-1}}$ are linearly disjoint over E.

We have seen (cf. Jacobson [9], p. 163) that if P is an algebraic extension of E (possibly infinite dimensional), then P is separable over E if and only if P is linearly disjoint to $E^{p^{-1}}$ over E. Hence, the

following criterion is applicable in the case where P is separable algebraic over E.

Theorem 2.b. Every derivation of E into P can be extended to a derivation in P if and only if P is separable over E.

Proof. Assume that every derivation of E into P can be extended to a derivation in P. Then by Corollary 2 of Theorem 2.A, every p-basis of E over $\Phi_0$ is p-independent in P. We must show that whenever the elements $x_1, \ldots, x_n$ of P are linearly independent over E, these elements $x_i$ are also linearly independent over $E^{p^{-1}}$, or equivalently, that $x_1^p, \ldots, x_n^p$ are linearly independent over E. Let $x_1, \ldots, x_n \in P$ and $\alpha_1, \ldots, \alpha_n$ be elements of E not all zero such that $\sum \alpha_i x_i^p = 0$. If B is a p-basis for E, then $\alpha_i = \sum \gamma_{ik_1 \ldots k_r} \beta_1^{k_1} \cdots \beta_r^{k_r}$ where $0 \le k_i < p$, $\beta_j \in B$, $\gamma_{ik_1 \ldots k_r} \in \Phi_0(E^p) = E^p$. We now have $\sum \delta_{k_1 \ldots k_r} \beta_1^{k_1} \cdots \beta_r^{k_r} = 0$ where $\delta_{k_1 \ldots k_r} = \sum_i \gamma_{ik_1 \ldots k_r} x_i^p \in P^p$. We can write $\gamma_{ik_1 \ldots k_r} = (\eta_{ik_1 \ldots k_r})^p$, $\eta_{ik_1 \ldots k_r} \in E$. Since the $\beta$'s are p-independent in $P/\Phi_0$, we have $0 = \delta_{k_1 \ldots k_r} = \sum_i (\eta_{ik_1 \ldots k_r})^p x_i^p$. Hence $\sum_i \eta_{ik_1 \ldots k_r} x_i = 0$. Since the $\alpha_i$ are not all zero, not all $\eta_{ik_1 \ldots k_r}$ are zero. We have thus proved that the $x_i$ are linearly independent over E implies that $x_i^p$ are linearly independent over E, which is, by Remark 2.6, equivalent to separability of P/E. Conversely,

assume that P is separable over E. Let $\{\beta_1, \ldots, \beta_r\}$ be an arbitrary set of elements of E which is p-dependent in $P/\Phi_o$. Then we have

$$0 = \sum \gamma_{k_1 \ldots k_r} \beta_1^{k_1} \ldots \beta_r^{k_r}, \quad \gamma_{k_1 \ldots k_r} = (\eta_{k_1 \ldots k_r})^p \in P^p,$$

$0 \leq k_i < p$, not all $\gamma_{k_1 \ldots k_r} = 0$. Let $\{x_t\}$ be a basis for P over E and write $\eta_{k_1 \ldots k_r} = \sum_{i=1}^{n} \lambda_{ik_1 \ldots k_r} x_i$, the $\lambda$'s in E. Then we have

$$0 = \sum \gamma_{k_1 \ldots k_r} \beta_1^{k_1} \ldots \beta_r^{k_r} = \sum \Big( \sum_i \lambda_{ik_1 \ldots k_r}^{p} x_i^p \Big) \beta_1^{k_1} \ldots \beta_r^{k_r}.$$

That is, $0 = \sum \mu_i x_i^p$ where $\mu_i = \sum \lambda_{ik_1 \ldots k_r}^{p} \beta_1^{k_1} \ldots \beta_r^{k_r}$.

By Remark 2.6, every $\mu_i = 0$ since P is separable over E. Hence the $\beta$'s are p-dependent in $E/\Phi_o$. Since the $\beta$'s were arbitrarily chosen, it follows that any p-basis of $E/\Phi_o$ is p-independent in $P/\Phi_o$.

Remark 2.7. Let $P/\Phi$ be an extension field of $\Phi$ with characteristic p, P purely inseparable of exponent one over $\Phi$. Then $[P:\Phi] = p^m < \infty$ if and only if $\{\rho_1, \rho_2, \ldots, \rho_m\}$ is a p-basis for P over $\Phi$.

Proof. Suppose that $\{\rho_1, \ldots, \rho_m\}$ is a p-basis for P over $\Phi$, then Remark 2.4 shows that $[P:\Phi] = p^m$. Conversely, assume that $[P:\Phi] = p^m < \infty$. Clearly $\{\rho_1, \ldots, \rho_n\}$ is a p-basis for P over $\Phi$, $n \neq m$, would also imply that $[P:\Phi] = p^n \neq p^m$. This shows that $[P:\Phi] = p^m$ implies that $\{\rho_1, \ldots \rho_m\}$ is a p-basis for P over $\Phi$.

Exercise 2.1 (Baer). P is purely inseparable of exponent one over $\Phi$ and $[P:\Phi] = p^m < \infty$. Show that there exists a derivation D of $P/\Phi$

such that the only D-constants are the elements of $\Phi$.

Proof. In view of Remark 2.7, we can take $P = \Phi(B)$ where $B = \{x_1, \ldots x_m\}$ is a p-basis for P over $\Phi$. A derivation D of $P/\Phi$ is then completely determined by its values at the $x_i$ (cf. Theorem 2.A). Choose $x_iD = \lambda_i x_i$, $\lambda_i \in \Phi$ and let M be any monomial of the form $x_1^{k_1} \ldots x_m^{k_m}$, $0 \leq k_i < p$. Then $(M)D = k_1 x_1^{k_1 - 1}\lambda_1 x_1 x_2^{k_2} \ldots x_m^{k_m}$ $+ \ldots + k_m x_1^{k_1} \ldots x_m^{k_m - 1}\lambda_m x_m = (k_1\lambda_1 + \ldots + k_m\lambda_m)M$. It is clear that $M = 1$ implies that $(M)D = 0$. Let us show that we can choose the $\lambda_i$ so that $k_1\lambda_1 + \ldots + k_m\lambda_m = 0$ if and only if each $k_i = 0$. Since the $x_i$ are linearly independent over $\Phi$, it is clear that the $x_i^p$ are linearly independent over $\Phi^p$ and hence over the prime field $\Phi_0$ $(= \Phi_0^p$, using Fermat's theorem). Clearly it is sufficient to take $\lambda_i = x_i^p$. Then the only monomial which belongs to the constant field of D is the trivial monomial $1 \in \Phi$. We have seen that if $M_i$ is a monomial of the form $x_1^{k_1} \ldots x_m^{k_m}$, $0 \leq k_i < p$, we can write $(M_i)D = \mu_i M_i$, $0 \neq \mu_i \in P^p \subseteq \Phi$. Hence $(M_i)D^n = \mu_i^n M_i$. We observe that for $M_i \neq M_j$ we obtain $\mu_i \neq \mu_j$. Suppose there exist $\alpha_i \in \Phi = \Phi(P^p)$ such that $\left(\sum_{i=1}^{n} \alpha_i M_i\right) D = 0$. Then we obtain a system

$$\begin{aligned} \alpha_1\mu_1 M_1 + \ldots + \alpha_n\mu_n M_n &= 0 \\ &\vdots \\ \alpha_1\mu_1^n M_1 + \ldots + \alpha_n\mu_n^n M_n &= 0 \end{aligned}$$

of n linear homogeneous equations in the $\alpha_i$. Since the determinant

$$\begin{vmatrix} \mu_1 M_1 & \mu_2 M_2 & \cdots & \mu_n M_n \\ \mu_1^2 M_1 & \mu_2^2 M_2 & \cdots & \mu_n^2 M_n \\ \vdots & \vdots & & \vdots \\ \mu_1^n M_1 & \mu_2^n M_n & \cdots & \mu_n^n M_n \end{vmatrix}$$

$$= \mu_1 \cdots \mu_n M_1 \cdots M_n \begin{vmatrix} 1 & \cdots & 1 \\ \mu_1 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_1^{n-1} & & \mu_n^{n-1} \end{vmatrix}$$

$= \mu_1 \cdots \mu_n M_1 \cdots M_n \prod_{i>j} (\mu_i - \mu_j) \neq 0$ for distinct $M_i$ and hence distinct $\mu_i$, then we must have every $\alpha_i = 0$. Therefore the constant field of D is precisely $\Phi$.

Exercise 2.2. Let D be a non-zero derivation in a field P of characteristic $p \neq 0$ over $\Phi$. Show that $1\cdot\rho_0 + D\rho_1 + \cdots + D^{p-1}\rho_{p-1}$, $\rho_i \in P$, is a derivation only if every $\rho_i = 0$, $i \neq 1$. Show that if $\rho \in P$, then

(23) $(D\rho)^k = D^k\rho^k + D(\rho E^{k-1}) + \sum_{j=2}^{k-1} D^j\rho_j$, where $\rho_j \in P$ and $E = D\rho$ $(\equiv D\rho_R)$. Hence prove the following formula due to Hochschild:

$$E^p = (D\rho)^p = D^p\rho^p + D(\rho E^{p-1}) .$$

Proof. We have seen that $1, D, D^2, \ldots, D^{p-1}$ are elements of $\mathfrak{L}_\Phi(P)$. Write $T = 1\cdot\rho_0 + D\rho_1 + \cdots + D^{p-1}\rho_{p-1}$. If $T = 0$, then $yT = 0$ for all $y \in P$. In particular, let $y_0, y_1, \ldots, y_{p-1}$ be

p linearly independent (over $\Phi$) elements of P. Then the Wronskian $(1, D, \ldots, D^{p-1})\left|y_0 y_1 \ldots y_{p-1}\right| \neq 0$. This means that the p linear homogeneous equations $y_i T = 0$, $i = 0, 1, \ldots, p-1$, have only trivial solutions $\rho_i = 0$, $i = 0, 1, 2, \ldots, p-1$. Hence the operators $1, D, D^2, \ldots, D^{p-1}$ are linearly independent over P.

We assert that T is a derivation in P only if every $\rho_i = 0$, $i \neq 1$. We have seen that $D\rho$ $(\equiv D\rho_R)$ is a derivation in P. If $p = 2$, T becomes $1\cdot\rho_0 + D\rho_1$ and for $x,y \in P$ we have, on the one hand, $(xy)T = (xy)\rho_0 + (xD\rho_1)y + x(yD\rho_1)$, and one the other

$(xT)y + x(yT) = (x\rho_0 + xD\rho_1)y + x(y\rho_0 + yD\rho_1)$

$= (x\rho_0)y + x(y\rho_0) + (xD\rho_1)y + x(yD\rho_1)$

$= 2(xy)\rho_0 + (xD\rho_1)y + x(yD\rho_1)$. Therefore T is a derivation only if $\rho_0 = 0$, in which case $T = D\rho_1$. We shall now assume that $p > 2$. We have seen that

$$(xy)D^k = x(yD^k) + \sum_{i=1}^{k-1} \binom{k}{i} (xD^i)(yD^{k-i}) + (xD^k)y \text{ for } k = 1, 2, \ldots$$

and that $D^p$ is a derivation in P. Now

$$(xT)y = (x\rho_0)y + \left((xD)\rho_1\right)y + \left((xD^2)\rho_2\right)y + \ldots + \left((xD^{p-1})\rho_{p-1}\right)y$$

$$x(yT) = x(y\rho_0) + x\left((yD)\rho_1\right) + x\left((yD^2)\rho_2\right) + \ldots + x\left((yD^{p-1})\rho_{p-1}\right)$$

$$x(yT) + (xT)y = 2(xy)\rho_0 + (xy)D\rho_1 + (xD^2\rho_2)y + x(yD^2\rho_2) +$$

$$\ldots + (xD^{p-1}\rho_{p-1})y + x(yD^{p-1}\rho_{p-1}).$$

We also have the relation

$$(xy)T = (xy)\rho_0 + (xy)D\rho_1 + (xy)D^2\rho_2 + \ldots + (xy)D^{p-1}\rho_{p-1} .$$

If T is a derivation in P we must now have $\rho_o = 0$ and

$$(xy)D^k\rho_k = x(yD^k\rho_k) + (xD^k\rho_k)y, \quad k = 2,3, \ldots, p-1.$$

This means that $\sum_{i=1}^{k-1} \binom{k}{i} (xD^i)(yD^{k-i})\rho_k = 0$. Put otherwise,

$\sum_{i=1}^{k-1} \binom{k}{i} D^i(yD^{k-i})\rho_k = 0$. Since $k < p$, it is clear that p does not divide any $\binom{k}{i}$. Moreover, we have just shown that $1, D, D^2, \ldots, D^{p-1}$ are right linearly independent over P. Therefore if $\sum_{i=1}^{k-1} \binom{k}{i} D^i(yD^{k-i})\rho_k = 0$ for all $y \in P$, we must have $\rho_k = 0$ for each $k = 2,3, \ldots, p-1$. This means that if T is a derivation it must reduce to the form $E = D\rho$, $\rho \in P$.

We next show that the relation (23) holds for $k = 1,2, \ldots$. If $k = 1$, we write $D\rho = D\rho_R$. If $k = 2$, we have

$$(D\rho)^2 = (D\rho)(D\rho) = D^2\rho^2 + D(\rho D)\rho$$
$$= D^2\rho^2 + D\left(\rho(D\rho)\right) = D^2\rho^2 + D(\rho E).$$

We next assume that $(D\rho)^n = D^n\rho^n + D(\rho E^{n-1}) + \sum_{i=2}^{n-1} D^i\rho_i$ for each $n < k$, where $\rho_i \in P$ and $E = D\rho$. Since $\rho^k D = \rho^{k-1}(\rho D)k$, we now have

$$(D\rho)^{n+1} = (D^n\rho^n)(D\rho) + \left(D(\rho E^{n-1})D\rho\right) + \sum_{i=2}^{n-1} (D^i\rho_i)(D\rho)$$
$$= D^{n+1}\rho^{n+1} + D^n\left(\rho^{n-1}(\rho D)\rho n\right) + D(\rho E^n)$$
$$+ D^2\rho(\rho E^{n-1}) + \sum_{i=2}^{n-1} \left(D^{i+1}\rho_i\rho + D^i\rho_i(D\rho)\right)$$

$$= D^{n+1}\rho^{n+1} + D(\rho E^n) + \sum_{j=2}^{n} D^j\rho_j \ , \quad \rho_j \in P.$$

This proves the general result given by (23).

We recall that Der (P) is closed under right multiplication by elements of P and under p-th powers. Therefore $E^p$, $D^p\rho^p$, $D(\rho E^{p-1})$ are in Der (P). We also have

$$E^p - D^p\rho^p = D(\rho E^{p-1}) + \sum_{i=2}^{p-1} D^i\rho_i \ , \quad \rho_i \in P.$$

Since $E^p - D^p\rho^p$ is in Der (P) we apply our preceding result and conclude that each $\rho_i = 0$, $i = 2,3, \ldots, p-1$. Hence

$$E^p = (D\rho)^p = D^p\rho^p + D(\rho E^{p-1}).$$

Exercise 2.3. Let $P = \Phi(\rho_1, \ldots, \rho_m)$, $\Phi$ of characteristic $p \neq 0$, $\rho_i^p = \beta_i \in \Phi$, $[P:\Phi] = p^m < \infty$. Let D be a derivation in $P/\Phi$ such that $\Phi$ is the subfield of D-constants (see Exercise 2.1). Show that the minimal polynomial of D as a linear transformation in P over $\Phi$ is a so-called p-polynomial of the form

$$(24) \quad X^{p^m} + \beta_1 x^{p^{m-1}} + \ldots + \beta_m x \ , \quad \beta_i \in \Phi.$$

Show that every element in the algebra $\mathfrak{L}_{\Phi}(P)$ of linear transformations in P over $\Phi$ can be written in one and only one way in the form

$$(25) \quad 1\sigma_0 + D\sigma_1 + D^2\rho_2 + \ldots + D^{N-1}, \quad N = p^m, \ \sigma_i \in P.$$

Proof. Let us first consider the differential equation $(yD^m)k_0 + (yD^{m-1})k_1 + \ldots + yk_m = 0$, $k_i \in P$ and not all $k_i = 0$. Suppose that this equation has more than m solutions in P and let $y_1, y_2, \ldots, y_{m+1}$ be m + 1 solutions. Consider the m + 1 linear

equations

$$y_i(D^m k_0 + D^{m-1}k_1 + \ldots + 1\cdot k_m) = 0$$

and employ the notation of Exercise 1.3. We obtain non-trivial solutions for the $k_i$ if and only if the Wronskian $(1,D,D^2, \ldots, D^m)|y_1,y_2, \ldots, y_{m+1}| = 0$. By Exercise 1.3, this Wronskian vanishes if and only if $y_1,y_2, \ldots, y_{m+1}$ are linearly dependent over $\Phi$. Hence there exist at most m linearly independent solutions of the given differential equation.

We next recall that for $p \neq 0$, $\mathrm{Der}_\Phi(P)$ is closed under p-th powers. Hence $D, D^p, D^{p^2}, \ldots$ are derivations in P over $\Phi$. Since $\mathrm{Der}_\Phi(P)$ is closed under right multiplication by elements of P, we now have

$$Da_0 + D^p a_1 + \ldots + D^{p^{m-1}} a_{m-1}, \quad a_i \in P$$

is a derivation in P over $\Phi$. More generally,

$$T = 1\cdot\sigma_0 + D\sigma_1 + D^2\sigma_2 + \ldots + D^{N-1}\sigma_{N-1}, \quad N = p^m, \quad \sigma_i \in P,$$

is a linear transformation in P over $\Phi$. We have seen in Remark 2.7 that $\{\rho_1, \ldots, \rho_m\}$ is a p-basis for P over $\Phi$. Hence the $p^m$ monomials $\rho_1^{k_1} \ldots \rho_m^{k_m}$, $0 \leq k_j < p$, form a basis for P over $\Phi(P^p) = \Phi$. Choose D such that $\rho_i D = \lambda_i \rho_i$, $\lambda_i = \rho_i^p \in \Phi$. Then, as in Exercise 2.1, if $M_j$ is any monomial of the above form, $M_j D = \mu_j M_j$, $\mu_j \in P^p \subseteq \Phi$, $\mu_j = 0$ if and only if $M_j = 1$, and $\mu_i \neq \mu_j$ whenever $i \neq j$. If the transformation $T = 0$, then $\xi T = 0$ for all $\xi \in P$. In particular, as $\xi$ takes on the values of each $M_j$, we obtain the N linear equations $(M_j)T = 0$, $j = 0, 1, 2, \ldots, N - 1$. As in Exercise 2.1, we note

that $M_j D^r = \mu_j^r M_j$. Hence we may write the above N equations as

$$(M_j \cdot 1)\sigma_0 + (\mu_j M_j)\sigma_1 + \ldots + (\mu_j^{N-1} M_j)\sigma_{N-1} = 0,$$

$$j = 0,1,2, \ldots, N-1.$$

Since $\mu_j \neq \mu_i$ whenever $M_j \neq \mu_i$, the determinant

$$\begin{vmatrix} M_0 & \mu_0 M_0 & \cdots & \mu_0^{N-1} M_0 \\ M_1 & \mu_1 M_1 & \cdots & \mu_1^{N-1} M_1 \\ \vdots & \vdots & & \\ M_{N-1} & \mu_{N-1} M_{N-1} & \cdots & \mu_{N-1}^{N-1} M_{N-1} \end{vmatrix}$$

$$= \prod_{k=0}^{N-1} M_k \prod_{i>j} (\mu_i - \mu_j) \neq 0$$ for distinct $M_k$. This means that each $\sigma_i = 0$, $i = 0, 1, \ldots, N-1$, whence $1, D, D^2, \ldots, D^{N-1}$ are right linearly independent over P. Therefore $1, D, D^2, \ldots, D^{N-1}$ is a basis for $\mathfrak{L}_\Phi(P)$ over P (cf. Jacobson [9], p. 20). Hence every element of $\mathfrak{L}_\Phi(P)$ can be written uniquely in the form (25).

Since the set $\{1, D, D^2, \ldots, D^{N-1}\}$ consists of right linearly independent elements over P, therefore the set $\{D, D^p, \ldots, D^{p^{m-1}}\} \subseteq \{1, D, D^2, \ldots, D^{N-1}\}$ also consists of right linearly independent elements over P. Corollary 1 of Theorem 2.A implies that $[\mathrm{Der}_\Phi(P):P]_R = m$. Therefore $\{D, D^p, \ldots, D^{p^{m-1}}\}$ is indeed a basis for $\mathrm{Der}_\Phi(P)$ over P. We shall call D a <u>generator</u> of $\mathrm{Der}_\Phi(P)$. Since $D^{p^m}$ is derivation in P over $\Phi$,

$$D^{p^m} = D^{p^{m-1}} b_{m-1} + D^{p^{m-2}} b_{m-2} + \ldots + Db_0$$

for fixed $b_i \in P$. Observe that

$$[D^k b_k, D] = (D^k b_k)D - D(D^k b_k)$$

$$= (D^k D)b_k + D^k(b_k D) - (D D^k)b_k = D^k(b_k D).$$

Therefore $[D^{p^m}, D] = 0 = D^{p^{m-1}}(b_{m-1}D) + \ldots + D(b_0 D)$ and since $D, D^p, \ldots, D^{p^{m-1}}$ are right linearly independent over P, each $(b_i D) = 0$. Hence each $b_i \in \Phi$. Write $b_i = -\beta_i$ and conclude that the minimal polynomial for D is as given by (24).

We shall now derive an analogue of the normal basis theorem for separable normal extensions. We recall (cf. Jacobson [9], pp. 40-41) that a field $P/\Phi$ is finite dimensional Galois over $\Phi$ if and only if P is a splitting field over $\Phi$ of a separable polynomial $f(x) \in \Phi[x]$. In the latter case we call $P/\Phi$ a <u>normal extension</u>. Let $P/\Phi$ be finite dimensional Galois over $\Phi$ with Galois group $G = \{s_1, s_2, \ldots, s_n\}$. If $\rho \in P$, we call the images $\rho^{s_i}$ under $s_i \in G$ the conjugates of $\rho$ in $P/\Phi$. If $[P:\Phi] = n$ and $\{\rho^{s_i} : s_i \in G\}$ is a set of linearly independent elements, then the set $\{\rho^{s_1}, \rho^{s_2}, \ldots, \rho^{s_n}\}$ forms a basis for P over $\Phi$. Such a basis is called a <u>normal basis</u> for the Galois extension. We call an extension field $P/\Phi$ cyclic if it is finite dimensional Galois and its Galois group G is cyclic. We next recall the following result for finite base fields $\Phi$ :

Any cyclic extension $P/\Phi$ has a normal basis over $\Phi$ (cf. Jacobson [9], p. 61).

In particular, we observe that, if s is a generator of G, there exists an element $\rho \in P$ such that $\{\rho, \rho^s, \ldots, \rho^{s^{n-1}}\}$ is a normal basis for $P/\Phi$.

Let $P = \Phi(\rho_1, \rho_2, \ldots, \rho_m)$, $\Phi$ of characteristic $p \neq 0$, $\rho_i^p = \gamma_i \in \Phi$, $[P:\Phi] = p^m < \infty$, and D be a derivation in P over $\Phi$. We can assume (see Exercise 2.3) that $\Phi$ is the field of D-constants, whence D is a generator of $\mathrm{Der}_{\Phi}(P)$ and satisfies

$$D^{p^m} = D^{p^{m-1}}\beta_1 + D^{p^{m-2}}\beta_2 + \ldots + D\beta_m , \quad \beta_i \in \Phi .$$

In this case, $x^{p^m} - x^{p^{m-1}}\beta_1 - x^{p^{m-2}}\beta_2 - \ldots - x\beta_m$ is in fact the minimal polynomial of D over $\Phi$ (see Exercise 2.3).

Suppose $(y_1, y_2, \ldots, y_N)$, $N = p^m$, is a basis for P over $\Phi$ and $(y_1D, y_2D, \ldots, y_ND)$ is an ordered set of vectors in P defined by

$$y_iD = \sum_{j=1}^{N} a_{ji}y_j , \quad a_{ji} \in \Phi , \quad i = 1, 2, \ldots, N.$$

As elements of P, each $y_iD$ is represented in a unique manner as linear combinations of $y_1, y_2, \ldots, y_N$. Therefore the matrix $A = (a_{ij})$ is uniquely determined by the basis $(y_1, y_2, \ldots, y_N)$ and the ordered set $(y_1D, y_2D, \ldots, y_ND)$. We observe that $(y_1D, \ldots, y_ND) = (y_1, \ldots, y_N)A$. Let us call A the <u>matrix of D relative to the basis</u> $(y_1, \ldots, y_N)$. If $f(\lambda)$ is the characteristic function $|A - \lambda I|$, then by the Hamilton-Cayley theorem, $f(D) = 0$. This shows that the minimal polynomial of D over $\Phi$ divides $f(\lambda)$. Since $f(\lambda)$ is also of degree $p^m = N$, we must have $f(\lambda) = \lambda^{p^m} - \lambda^{p^{m-1}}\beta_1 - \ldots - \lambda\beta_m$. This shows that the characteristic function and the minimal equation of A are identical. We now have the matrix A is similar to the so-called <u>companion matrix</u> of $f(\lambda)$ given by

$$B = \begin{Vmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & & & \beta_m \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \beta_1 \\ & & & \ddots & \vdots \\ & & & 1 & 0 \end{Vmatrix}$$

with 1's in sub-diagonal and 0's elsewhere except the entries in the N-th column corresponding to the coefficients of powers of $\lambda$ in the characteristic function $|B - \lambda I| = f(\lambda)$. Let $(z, z_2, \ldots, z_N)$ be a basis for P over $\Phi$ such that B is the matrix of D relative to this basis. We observe that $zD = z_2$, $z_2D = z_3$, $\ldots$, $z_{N-1}D = z_N$. Hence P has a basis of the form $(z, zD, zD^2, \ldots, zD^{p^m-1})$.

## SECTION 3

Let P be a field extension which is purely inseparable of exponent one over $\Phi$ with characteristic $p \neq 0$. As before, we let $\mathrm{Der}_{\Phi}(P)$ denote the set of derivations of P over $\Phi$. We recall that $\mathrm{Der}_{\Phi}(P)$ is a restricted Lie algebra of linear transformations in P over $\Phi$ and that $\mathrm{Der}_{\Phi}(P)$ is a right vector space over P relative to right multiplication $D\rho = D\rho_R$, $\rho \in P$. In this section we shall consider derivations in P as a ring $\mathcal{E}(P)$ of endomorphisms[1] D of the additive group (P, +) with the condition that $(\rho\sigma)D = (\rho D)\sigma + \rho(\sigma D)$, $\rho, \sigma \in P$.

We now suppose that $\mathfrak{D}$ is a set of derivations in P with the following closure properties: (1) $\mathfrak{D}$ is closed under addition. (2) $\mathfrak{D}$ is closed under Lie commutation $[D_1, D_2]$. (3) $\mathfrak{D}$ is closed under p-th powers. (4) $\mathfrak{D}$ is closed under right multiplication by elements $\rho_R$, $\rho \in P$. These four conditions imply that $\mathfrak{D}$ is a subspace of the right vector space of endomorphisms of (P, +).

Definition 3.1. Any set $\mathfrak{D}$ of endomorphisms of (P, +) which satisfy conditions (1) to (4) will be called a restricted P-Lie algebra of endomorphisms of (P, +).

Remark 3.1. It should not be inferred that $\mathfrak{D}$ is an algebra over P as a base field. To this end, we shall show that the relationship

$$[D_1, D_2]\rho = [D_1\rho, D_2] = [D_1, D_2\rho]$$

---

[1] cf. Jacobson [7], pp. 78-80.

does not hold for all $\rho \in P$.

Proof. $[D_1, D_2]\rho = (D_1D_2)\rho - (D_2D_1)\rho = D_1(D_2\rho) - D_2(D_1\rho)$

$[D_1\rho, D_2] = (D_1\rho)D_2 - D_2(D_1\rho) = (D_1D_2)\rho - D_1(\rho D_2) - D_2(D_1\rho)$

$[D_1, D_2\rho] = D_1D_2\rho - D_2\rho D_1 = D_1(D_2\rho) - D_2(D_1\rho) - D_2(\rho D_1)$

It is now clear that this relationship holds only if $D_1(\rho D_2) = D_2(\rho D_1) = 0$. In this case $\rho \in C(\mathfrak{D})$ the constant field of $\mathfrak{D}$ which is properly contained in P.

Theorem 3.A. (Jacobson). Let P be a field of characteristic $p \neq 0$ and let $\mathfrak{D}$ be a restricted P-Lie algebra of derivations in P such that $[\mathfrak{D}:P]_R = m < \infty$. Then: (1) if $\Phi$ is the subfield of $\mathfrak{D}$-constants, then P is purely inseparable of exponent $\leq 1$ over $\Phi$ and $[P:\Phi] = p^m$; (2) if D is any derivation in P over $\Phi$, then $D \in \mathfrak{D}$; (3) if $(D_1, \ldots, D_m)$ is any right basis for $\mathfrak{D}$ over P, then the set of monomials $D_1^{k_1} \ldots D_m^{k_m}$, $0 \leq k_i < p$, $D_i^0 = 1$, is a right basis for the ring $\mathfrak{L}_\Phi(P)$ of linear transformations of P over $\Phi$ considered as a right vector space over P (Jacobson [9], pp. 186-188).

We next observe that the author [9] gives the following proposition as an exercise.

Let $P = \Phi(\rho_1, \ldots, \rho_m)$, $\Phi$ of characteristic $p \neq 0$, $\rho_i^p = \beta_i \in \Phi$, $[P:\Phi] = p^m$. Let $\mathrm{Der}_\Phi(P)$ be the set of derivations in $P/\Phi$, $\mathfrak{F}$ be a subspace of the right vector space $\mathrm{Der}_\Phi(P)$ over P which is closed under p-th powers. Prove that $\mathfrak{F}$ is also closed under commutation so $\mathfrak{F}$ satisfies all the conditions of the above theorem (cf. Exercise 4, p. 190).

The possibility that one may delete the assumption that $\mathfrak{D}$ be a Lie subring of $\mathrm{Der}_{\Phi}(P)$ is answered by M. Gerstenhaber [2], p. 561. The author [2] points out that if we define a restricted subspace of Der (P) to be a subset which is a vector space over P and which is closed under p-th powers, then one may make the following claim.

Claim. If $\mathfrak{D}$ is a finite dimensional restricted subspace of Der (P) and if $\Phi$ is the subfield of $\mathfrak{D}$-constants, then $\mathfrak{D} = \mathrm{Der}_{\Phi}(P)$.

The author [2] then remarks that it follows a posteriori that $\mathfrak{D}$ must be a Lie subring of Der (P).

Let P be a field of characteristic $p \neq 0$ and Der (P) denote the set of derivations of P into itself. Given a derivation D in Der (P), denote the constant field of D by $\Gamma$. We recall that $P^p \subseteq \Gamma$ for all D in Der (P) and that if $a \in P$, $a \notin \Gamma$, then $[\Gamma(a): \Gamma] = p$. Suppose that $a \in P$ satisfies $aD \neq 0$. Then setting $D' = Da(aD)^{-1}$, we have $aD' = (aD)a(aD)^{-1} = a$, since P is a field. Let us denote the constant field of $D'$ by $\Gamma'$ and choose an element $b \in P$, $b \neq a$. Then $bD' = (bD)a(aD)^{-1}$ vanishes if and only if $bD = 0$. Hence $\Gamma = \Gamma'$.

Next, suppose that for some D in Der (P) and a,b in P we have $aD = \lambda a$, $bD = \mu b$, $\lambda, \mu$ in $\Gamma$. Then

$$(ab)D = (aD)b + a(bD) = \lambda ab + a\mu b = (\lambda + \mu)ab$$

since P is a field. We also have

$$1D = 0 = (a\, a^{-1})D$$

$$= (\lambda a)a^{-1} + a(a^{-1}D). \quad \text{Hence } a^{-1}D = -\lambda a^{-1}.$$

It is now clear that if $a_1 D = \lambda a_1$ and $a_2 D = \lambda a_2$, $\lambda \in \Gamma$, then

$$\left(\frac{a_1}{a_2}\right) D = \lambda a_1 a_2^{-1} - a_1 \lambda a_2^{-1} = 0. \quad \text{That is, } \frac{a_1}{a_2} \in \Gamma.$$

Therefore the set of all a in P such that $aD = \lambda a$, $\lambda \in \Gamma$, is either reduced to the zero element or is one-dimensional over $\Gamma$.

Lemma 3.1. Suppose D is in Der (P) and $a \neq 0 \in P$ such that $aD = a$. Set $D' = D^p - D$. Then $\Gamma' = \Gamma(a)$.

Proof. Since $aD = a$ implies $aD^p = a$, it follows that $aD' = 0$. Therefore $\Gamma' \supseteq \Gamma(a)$. It remains to prove that $\Gamma' \subseteq \Gamma(a)$.

Set $f(t) = t^p - t$. Since the formal derivative $f'(t)$ of $f(t)$ does not vanish, $f(t)$ has distinct roots $0,1, \ldots, p-1$ in $Z_p$ the Galois field of p elements. Therefore $f(t) = t(t-1) \ldots (t-p+1)$ over $Z_p$. Define polynomials $f_i(t)$, $i = 0,1, \ldots, p-1$, of degree $p-1$ in $t$ by the relation $f_i(t) = f(t)(t-i)^{-1}$. It is clear that $f_i(i) \neq 0$ and $f_i(j) = 0$ if $i \neq j$. Suppose that $\sum_i c_i f_i(t) = 0$, $c_i \in Z_p$. Then $\sum_i c_i f_i(j) = 0$, $j = 0,1, \ldots, p-1$. Since $\det\left(f_i(j)\right) = 1 \neq 0$, each $c_i = 0$. Put otherwise, the polynomials $f_i(t)$ are linearly independent over $Z_p$. In particular, we have $1 = \sum_i c_i f_i(t)$ for suitable $c_i \in Z_p$. Since we are viewing derivations in P as elements of the ring $\mathcal{E}(P)$, we can write $D^p - D = D(D-1) \ldots (D-p+1)$ in $\mathcal{E}(P)$. If $b \in \Gamma'$, $b(D^p - D) = 0 = b\left(D(D-1) \ldots (D-p+1)\right)$. This can be written as $bf_i(D)(D-i) = 0$. Setting $bf_i(D) = b_i$, we obtain $b_i(D-i) = 0$,

$i = 0,1, \ldots, p - 1$. It is also clear that $a^i(D - i) = a^iD - a^ii = ia^{i-1}(aD) - ia^i = 0$. Hence, for $a \neq 0 \in P$,

$$(b_i/a^i)(D - i) = b_i(D - i)a^{-i} - b_ia^{-2i}\left(a^i(D - i)\right) = 0.$$

This shows that

$$\begin{aligned}(b_i/a^i)(D - i) &= (b_iD)a^{-i} - ib_ia^{-i} - b_ia^{-2i}(a^iD - ia^i)\\ &= (b_iD)a^{-i} - b_ia^{-2i}(a^iD)\\ &= (b_i/a^i)D = 0\end{aligned}$$

Therefore $b_i/a^i \in \Gamma$ and so $b_i \in \Gamma(a)$, $i = 0,1, \ldots, p - 1$. However, $\sum f_i(D)c_i = 1$ for suitable $c_i \in Z_p$. Hence $\sum c_ib_i = b$, thus showing that $b \in \Gamma(a)$. This proves that $\Gamma' \subseteq \Gamma(a)$.

Remark 3.2. If $\{\beta_1, \ldots, \beta_m\} \subseteq P$ and $D_1, \ldots, D_m$ in $\mathrm{Der}_\Phi(P)$ are such that, for $1 \leq i, j \leq m$, $\beta_iD_j = \delta_{ij}$ the Kronecker delta, then $\beta_1, \ldots, \beta_m$ are p-independent and $D_1, \ldots, D_m$ are right linearly independent over P. Moreover, if $[P:\Phi] = p^m$, then $(D_1, \ldots, D_m)$ is a right basis for $\mathrm{Der}_\Phi(P)$ over P, and $(\beta_1, \ldots, \beta_m)$ is a p-basis for P over $\Phi$.

Proof. Let us write $B = \{\beta_1, \ldots, \beta_m\}$ and denote the set $\{\beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots \beta_m\}$ by $B_{\beta_i}$. If $\beta_1, \ldots, \beta_m$ are not p-independent, then $\beta_i \in \Phi(B_{\beta_i})$. It is clear that for all $\rho \in \Phi(B_{\beta_i})$, $\rho D_i = 0$, since $\beta'D_i = 0$ for each $\beta' \in B_{\beta_i}$. This implies (in particular) that $\beta_iD_i = 0$ which is contrary to our hypothesis that $\beta_iD_i = 1$. Hence $\beta_1, \ldots, \beta_m$ are p-independent.

Secondly, let us consider the derivation given by

$E = D_1\sigma_1 + \ldots + D_m\sigma_m$, $\sigma_i \in P$. If $E = 0$, then $yE = 0$ for all $y \in P$. In particular, $\beta_i E = 0$ for each $i = 1,2, \ldots, m$. This implies that $(\beta_i D_i)\sigma_i = 0$ for each $i = 1,2, \ldots, m$. Hence each $\sigma_i = 0$ and, consequently, the $D_i$ are right linearly independent over P.

We next assume that $[P:\Phi] = p^m$. It follows directly from Remark 2.7 that $(\beta_1, \ldots, \beta_m)$ is a p-basis for P over $\Phi$. By Corollary 1 of Theorem 2.A $[\mathrm{Der}_{\Phi}(P): P]_R = m$. Hence $(D_1, \ldots, D_m)$ is a right basis for $\mathrm{Der}_{\Phi}(P)$ over P.

Given D in Der (P), let ((D)) denote the smallest restricted subspace of Der (P) containing D. We have seen that Der (P) is closed under addition, under right multiplication $D\rho = D\rho_R$, $\rho \in P$, and under p-th powers. Hence ((D)) is the set of all derivations of the form $D\rho_0 + D^p\rho_1 + \ldots$. follows immediately from the fact (cf. Exercise 2.2) that

(26) $(D\rho)^p = D^p\sigma_1 + D\sigma_2$, for suitable $\sigma_1, \sigma_2 \in P$.

Lemma 3.2. If $D \in \mathrm{Der}\ (P)$ and $\beta_1, \ldots, \beta_m$ are p-independent over $\Gamma$ the constant field of D, then there exist $D_1, \ldots D_m$ in ((D)) such that $\beta_i D_j = \delta_{ij}$. (These $D_j$ are right linearly independent over P, by Remark 3.2).

Proof. It is clear that, for example, $\beta_2 D \neq 0$. Otherwise, we would have $\beta_2 \in \Gamma$, whence $\beta_2 \in \Gamma(\beta_1, \beta_3, \ldots, \beta_m)$, thus contradicting our hypothesis that $\beta_1, \ldots, \beta_m$ are p-independent over $\Gamma$. Hence $\beta_2 D$ has inverse $(\beta_2 D)^{-1}$ in P. Set $(\beta_2 D)^{-1}\beta_2 = a \in P$. Then $\beta_2(Da) = \beta_2$. Next, write $(Da)^p - Da = D'$ which clearly belongs to

$((D))$. By Lemma 3.1, we see that the constant field of $D'$ is $\Gamma(\beta_2)$. Since $\beta_3 \notin \Gamma(\beta_2)$, we can construct $D''$ in $((D))$ (see relation (26)) such that the constant field of $D''$ is $\Gamma(\beta_2, \beta_3)$. We can repeat this process for a finite number (= m - 1) of times until we obtain $D^*$ in $((D))$ with constant field $\Gamma(\beta_2, \beta_3, \ldots, \beta_m)$. Since $\beta_1 \notin \Gamma(\beta_2, \beta_3, \ldots, \beta_m)$, we must have $\beta_1 D^* = \lambda_1 \neq 0$ in P. Setting $D^*\lambda_1^{-1} = D_1$, we obtain $\beta_1 D_1 = 1$. It is clear that, for $\rho \neq 0$ in P, the constant field of $D \in \text{Der}(P)$ coincides with the constant field of $D\rho$. This follows immediately since $cD\rho = (cD)\rho = 0$ if and only if $cD = 0$. We have therefore constructed $D_1 \in ((D))$ such that $\beta_i D_i = 0$, $i \neq 1$. We can repeat the above construction for a finite number (= m) of times by changing the choice of the $\beta$ and so obtain derivations $D_j$ satisfying the given condition.

Corollary. If $\Gamma$ is the constant field of $D \in \text{Der}(P)$, then $[P:\Gamma]$ is finite if and only if $[((D)) : P]_R$ is finite, and in that case $((D))$ is the set of all derivations vanishing on $\Gamma$. In particular, $((D))$ is then a Lie subring of $\text{Der}(P)$.

Proof. Since $P^p \subseteq \Gamma$, if $[P:\Gamma] = [P:\Gamma(P^p)]$ is infinite, then the dimension of a p-basis of P over $\Gamma$ is infinite. By Lemma 3.2, we then conclude that $[((D)) : P]_R$ is infinite. On the other hand, if $[P:\Gamma] = p^n$ and $(\beta_1, \ldots, \beta_n)$ is a p-basis for P over $\Gamma$, then $[\text{Der}_\Gamma(P):P] = n$ and the $D_1, \ldots, D_n$ of Lemma 3.2 form a right basis for $\text{Der}_\Gamma(P)$ the Lie ring of derivations of P over $\Gamma$. We now have $((D))$ contains a basis over P for $\text{Der}_\Gamma(P)$ and $((D)) \subseteq \text{Der}_\Gamma(P)$. Therefore $((D)) = \text{Der}_\Gamma(P)$.

Given $D,D'$ in Der (P), let $((D,D'))$ denote the smallest restricted subspace of Der (P) containing both $D,D'$. Denote the constant fields of $D,D'$ by $\Gamma,\Gamma'$ respectively. If $D''$ is an element of $((D,D'))$, denote its constant field by $\Gamma''$.

Lemma 3.3. Given $D,D'$ in Der (P), let $\beta_1, \beta_2, \ldots, \beta_m$ be elements of P which are p-independent over $\Gamma$, and let $y \in \Gamma$ satisfy $yD' \neq 0$. Then there exists an element $D''$ in $((D,D'))$ such that $\beta_1, \ldots, \beta_m, y$ are p-independent over $\Gamma''$. Further, if $[((D,D')) : P]_R$ is finite, there exists $D''$ in $((D,D'))$ such that $\Gamma'' = \Gamma \cap \Gamma'$ and $((D'')) = ((D,D'))$.

Proof. By Lemma 3.2, there exist $D_1, D_2, \ldots, D_m$ in $((D))$ such that $\beta_j D_i = \delta_{ij}$. Since each $D_i$ is a right linear combination over P of p-th powers of D, we must have $yD_i = 0$, for each $i = 1,2, \ldots, m$. Let us consider the derivation $E = \left(D' - (D_1\sigma_1 + \ldots + D_m\sigma_m)\right)\rho$, $\rho, \sigma_i \in P$. If $yE = 1$, we must have $(yD')\rho = 1$, that is, $\rho = (yD')^{-1}$. If $\beta_j E = 0$, $j = 1, 2, \ldots, m$, we must also have $(\beta_j D' - \sigma_j)\rho = 0$. Since $\rho \neq 0$, $\sigma_j = \beta_j D'$. Let us then define E in $((D,D'))$ by

$$E = \left(D' - (D_1\sigma_1 + \ldots + D_m\sigma_m)\right)(yD')^{-1}$$

where $\sigma_j = \beta_j D'$ and so $yE = 1$ and $\beta_j E = 0$. We now have a set $\{\beta_1, \ldots \beta_m, y\} \subseteq P$ and derivations $D_1, \ldots, D_m, E$ in $((D,D'))$ such that $\beta_j E = 0$, $yE = 1$ and $yD_i = 0$, $\beta_j D_i = \delta_{ij}$. Therefore, by Remark 3.2, $\beta_1, \beta_2, \ldots, \beta_m, y$ are p-independent over K the constant field of $((D,D'))$ and $D_1, \ldots, D_m, E$ are right P-independent. Hence $D'' = \sum_i D_i\lambda_i + E\mu$, $\mu,\lambda_i \in P$ is an element of $((D,D'))$. Since E is a

right P-linear combination of $D'$, $D_1$, . . . , $D_m$, we can write

$D'' = \sum_i D_i\alpha_i + D'\alpha$ , $\alpha, \alpha_i \in P$. By relation (26),

$(D'')^p = \sum_i (D_i\alpha_i)^p + (D'\alpha)^p = \sum_i D_i^p \gamma_i + D_i\rho_i + D'^p\gamma + D'\rho$, for

suitable $\gamma_i$, $\gamma$, $\rho_i$, $\rho$ in P. Since every element of $((D''))$ has the form $D''v_0 + (D'')^p v_1 + \ldots$ , it is now clear that $((D'')) \subseteq ((D,D'))$.

By the corollary of Lemma 3.2, $\Gamma''$ is the constant field of $((D''))$. It is clear that $\Gamma'' \subseteq K$ the constant field of $((D,D'))$. Since $\beta_1, \ldots, \beta_m$, y are p-independent over K, it now follows that $\beta_1, \ldots, \beta_m$, y are p-independent over $\Gamma''$.

Since, by assumption, $((D,D'))$ has finite dimension over P, there exists $D''$ in $((D,D'))$ such that the dimension of $(D''))$ over P is maximal. We now claim that $\Gamma'' = \Gamma \cap \Gamma'$. It is clear that if $a \in \Gamma \cap \Gamma'$, then $aD' = 0$ and $aD_i = 0$, $i = 1, 2, \ldots, m$. Therefore $aD'' = 0$, whence $\Gamma \cap \Gamma' \subseteq \Gamma''$. If $\Gamma'' \neq \Gamma \cap \Gamma'$, there is an element $y_1$ in $\Gamma''$ which is not in $\Gamma \cap \Gamma'$. We must now have $y_1 D \neq 0$ or $y_1 D' \neq 0$. Without loss of generality, assume that $y_1 D' \neq 0$ and let $x_1, \ldots x_n$ be a p-basis for P over $\Gamma''$. Therefore $[P:\Gamma''] = p^n$ and, consequently, there exist $D_1, \ldots, D_n$ in $((D''))$ such that $(D_1, \ldots, D_n)$ is a right basis for $((D''))$ over P. We can now construct a new $D''$, call it $E''$, such that $x_1, \ldots, x_n, y_1$ are p-independent over $\Gamma^{(3)}$ the constant field of $E''$. Hence $[P:\Gamma^{(3)}] \geq p^{n+1}$ and the dimension of $((E''))$ over $P \geq n + 1$. This contradicts the maximality of $((D''))$. We must therefore have $\Gamma'' = \Gamma \cap \Gamma'$.

By the Corollary of Lemma 3.2, $((D''))$ is the set of all derivations vanishing on $\Gamma''$. Since $((D,D')) \subseteq \mathrm{Der}_{\Gamma \cap \Gamma'}(P)$, $((D,D'))$ vanishes on $\Gamma \cap \Gamma'$. Thus $((D,D'))$ vanishes on $\Gamma''$, whence $((D,D')) \subseteq ((D''))$. We have seen that $((D'')) \subseteq ((D,D'))$. Therefore $((D'')) = ((D,D'))$.

Corollary. Let $\mathfrak{D}$ be a finite-dimensional restricted subspace of Der (P) and $\Phi$ its constant field. Then there exists D in $\mathfrak{D}$ such that the constant field of D is $\Phi$, whence $((D)) = \mathrm{Der}_{\Phi}(P)$.

Proof. Let $D_1, \ldots, D_m$ be elements of Der (P) whose constant fields are $\Gamma^{(1)}, \ldots, \Gamma^{(m)}$ respectively. We shall assume that $((D_1, \ldots, D_i))$ is finite dimensional for each $i = 1, 2, \ldots, m$. In particular we shall write $\mathfrak{D} = ((D_1, \ldots, D_m))$. We have seen in Lemma 3.3 that there exists an element $E_2$ in $((D_1, D_2))$ such that $((E_2)) = ((D_1, D_2))$ and $\Gamma^{(1)} \cap \Gamma^{(2)} = K^{(2)}$ the constant field of $E_2$. Now $((D_1, D_2, D_3))$ is finite dimensional implies that there exists an element $E_3$ in $((E_2, D_3)) = ((D_1, D_2, D_3))$ such that $((E_3)) = ((D_1, D_2, D_3))$ and $\Gamma^{(3)} \cap K^{(2)} = K^{(3)}$ the constant field of $E_3$. Continuing in this way, we can construct $D = E_m$ in $((E_{m-1}, D_m)) = \mathfrak{D}$ such that $((D)) = \mathfrak{D}$ and $\Gamma^{(1)} \cap \ldots \cap \Gamma^{(m)} = K^{(m)}$ the constant field of D. By hypothesis, the constant field of $\mathfrak{D}$ is $\Phi$. Since $((D)) = \mathfrak{D}$, we must now have $K^{(m)} = \Phi$. By the corollary of Lemma 3.2, $((D))$ is the set of all derivations vanishing on $\Phi$, whence $((D)) = \mathrm{Der}_{\Phi}(P)$.

Theorem. Let P be a field of characteristic $p \neq 0$, $\mathfrak{D}$ a restricted subspace of Der (P) and $\Phi$ be the constant field of $\mathfrak{D}$. Then

$[P:\Phi]$ is finite only if $[\mathfrak{D}:P]_R$ is finite and in this case $\mathfrak{D} = \mathrm{Der}_{\Phi}(P)$. Moreover, $\mathfrak{D}$ is then a Lie subring of Der (P), there is an element D in $\mathfrak{D}$ such that the constant field of D is $\Phi$, and for any such D we have $\mathfrak{D} = ((D))$.

Proof. We have seen in the corollary of Lemma 3.3 that the constant field of D is $\Phi$, that $((D)) = \mathrm{Der}_{\Phi}(P)$ and $((D)) = \mathfrak{D}$. Since $((D))$ is finite dimensional over P, by the corollary of Lemma 3.2, $[P:\Phi]$ is finite and $((D)) = \mathfrak{D} = \mathrm{Der}_{\Phi}(P)$ a Lie subring of Der (P).

Let $\Phi$ be a subfield of P with characteristic $p \neq 0$, $\mathfrak{D}$ a finite dimensional restricted subspace of Der (P). We shall denote the constant fields of $\mathfrak{D}$, $\mathrm{Der}_{\Phi}(P)$ by $C(\mathfrak{D})$ and $C\left(\mathrm{Der}_{\Phi}(P)\right)$ respectively, and indicate correspondence between subfields of P and restricted subspaces of Der (P) by $\rightarrow$. If $\Phi$ is the constant field of $\mathfrak{D}$, from the theorem we obtain the correspondence

$$\Phi \rightarrow \mathrm{Der}_{\Phi}(P) \rightarrow C\left(\mathrm{Der}_{\Phi}(P)\right) = C(\mathfrak{D}) = \Phi,$$

or

$$\mathfrak{D} \rightarrow C(\mathfrak{D}) \rightarrow \mathrm{Der}_{C(\mathfrak{D})}(P) = \mathrm{Der}_{\Phi}(P) = \mathfrak{D}.$$

This is the type of Galois correspondence which we set out to establish.

## SECTION 4

Definition 4.1. Let $\mathfrak{A}$ be a subalgebra of an algebra $\mathfrak{B}$ over $\Phi$. Then a sequence of mappings $D^{(m)} = \{D_0=1, D_1, \ldots, D_m\}$ of $\mathfrak{A}$ into $\mathfrak{B}$ is called a <u>higher derivation of rank</u> m of $\mathfrak{A}$ into $\mathfrak{B}$ if every $D_i$ is $\Phi$-linear and

(27) $(ab)D_j = \sum_{i=0}^{j} (aD_i)(bD_{j-i})$, $j = 0, 1, \ldots, m$ holds for every $a, b \in \mathfrak{A}$. A <u>higher derivation of infinite rank</u> is an infinite sequence $\{D_0=1, D_1, \ldots\}$ of linear mappings of $\mathfrak{A}$ into $\mathfrak{B}$ such that (27) holds for all $j = 0,1,2, \ldots$.

It is clear that if $\{D_0, D_1, \ldots\}$ is a higher derivation of infinite rank, then the section $\{D_0, D_1, \ldots, D_m\}$ is a higher derivation of rank m and any section $\{D_0, D_1, \ldots, D_q\}$, $q \leq m$, of the higher derivation $\{D_0, D_1, \ldots, D_m\}$ is also a higher derivation. If we set $j = 1$ in (27), we obtain

$$(ab)D_1 = (aD_0)(bD_1) + (aD_1)(bD_0) = a(bD_1) + (aD_1)b.$$

Hence the mapping $D_1$ is a derivation of $\mathfrak{A}$ into $\mathfrak{B}$ since every $D_i$ is assumed to be $\Phi$-linear. We say that $D^{(m)}$ is <u>proper</u> if $D_1 \neq 0$.

Let $\mathfrak{A} = \mathfrak{B} = \Phi[x]$ where x is transcendental and let $D_i$ be the linear mapping in $\mathfrak{A}$ whose effect on the basis $(1, x, x^2, \ldots)$ is given by

(28) $x^m D_i = \binom{m}{i} x^{m-i}$

where $\binom{m}{i}$ is the usual binomial coefficient, $\binom{m}{i} = 0$ if $i > m$. Then

we must have

$$x^{m+n}D_j = \binom{m+n}{j} x^{m+n-j}$$

and

$$(x^m D_i)(x^n D_{j-i}) = \binom{m}{i} x^{m-i} \binom{n}{j-i} x^{n-j+i}$$

$$= \binom{m}{i}\binom{n}{j-i} x^{m+n-j}$$

We next make the following claim:

$$\sum_{i=0}^{j} \binom{m}{i}\binom{n}{j-i} = \binom{m+n}{j}$$ for arbitrary rational positive integers m, n, j with $j \leq m + n$.

Proof. We have agreed that $\binom{m+n}{j} = 0$ for $j > m + n$.

We have seen that $(1 + x)^m = \sum_{r=0}^{m} \binom{m}{r} x^r$, x an indeterminate. Hence

$$(1 + x)^m(1 + x)^n = \sum_{r=0}^{m} \binom{m}{r} x^r \sum_{s=0}^{n} \binom{n}{s} x^s$$

$$= \sum_{j=0}^{m+n} \sum_{i=0}^{j} \binom{m}{i}\binom{n}{j-i} t^j$$

We also have $(1 + x)^m(1 + x)^n = (1 + x)^{m+n} = \sum_{j=0}^{m+n} \binom{m+n}{j} t^j$ .

Therefore $\sum_{i=0}^{j} \binom{m}{i}\binom{n}{j-i} = \binom{m+n}{j}$ .

We can now state that

$$\sum_{i=0}^{j} (x^m D_i)(x^n D_{j-i}) = \sum_{i=0}^{j} \binom{m}{i}\binom{n}{j-i} x^{m+n-j}$$

$$= x^{m+n-j} \sum_{i=0}^{j} \binom{m}{i} \binom{n}{j-i}$$

$$= \binom{m+n}{j} x^{m+n-j} = x^{m+n} D_j .$$

Hence $\{1, D_1, D_2, \ldots\}$ is a higher derivation of infinite rank in $\Phi[x]$.

If $\Phi$ is of characteristic 0, then (28) shows that $i!x^m D_i = m(m-1) \ldots (m-i+1)\, x^{m-i}$ $(i! = 1.2.3 \ldots i)$. Set $f(x)D_1 = f'(x)$ the formal derivative of $f(x) \in \Phi[x]$. Then $i!D_i = D_1^i$ or, equivalently, $D_i = \frac{1}{i!} D_1^i$. More generally, if $a, b \in \mathfrak{A}/\Phi$, $\Phi$ has characteristic 0, then the relation (8) gives

$$(ab)\ \frac{D^j}{j!} = \sum_{i=0}^{j} \left(\frac{aD^i}{i!}\right) \left(\frac{bD^{j-i}}{(j-i)!}\right) .$$

Hence, if $D_1$ is a derivation in $\mathfrak{A}$ and $D_i = \frac{1}{i!} D_1^i$, then $\{1, D_1, D_2, \ldots\}$ is a higher derivation of infinite rank in $\mathfrak{A}$.

In Section 1 we discussed a connection between derivations D of $\mathfrak{A}/\Phi$ into $\mathfrak{B}/\Phi$ and homomorphisms $s = s(D)$ of $\mathfrak{A}$ into $\mathfrak{B} \otimes \mathbb{C}$ where $\mathbb{C}$ is an algebra with basis (1,t) over $\Phi$ with the multiplication rule $t^2 = 0$. In the case of higher derivations we shall introduce an algebra $\mathbb{C}^{(m)}$ with basis $(1, t, \ldots, t^m)$ over $\Phi$ such that $t^{m+1} = 0$. If $(x^{m+1})$ is the principal ideal generated by $x^{m+1}$ over $\Phi$, then we can identify $\mathbb{C}^{(m)}$ with $\Phi[x]/(x^{m+1})$. Construct the algebra $\mathfrak{B} \otimes \mathbb{C}^{(m)} = \mathfrak{B}^{(m)}$. If $D^{(m)} = \{1, D_1, \ldots, D_m\}$ is a higher derivation of rank m of $\mathfrak{A}$ into $\mathfrak{B}$, introduce the mapping $s = s(D^m)$ of $\mathfrak{A}$ into $\mathfrak{B}^{(m)}$ defined by

(29) $\quad a \to a^s \equiv a + (aD_1)t + \ldots + (aD_m)t^m.$

Since each $D_i$ is $\Phi$-linear, s is also linear. We have

$$a^s b^s = \sum_{i=0}^{m} (aD_i)t^i \sum_{k=0}^{m} (bD_k)t^k = \sum_{j=0}^{m} \sum_{i=0}^{j} (aD_i)(bD_{j-i})t^j$$

$$= \sum_{j=0}^{m} (ab)D_j t^j = (ab)^s .$$

This shows that s is a homomorphism of $\mathfrak{A}$ into $\mathfrak{B}^{(m)}$. We next introduce the mapping $\pi : a_0 + a_1 t + \ldots + a_m t^m \to a_0$ , $a_i \in \mathfrak{B}$. This mapping is clearly a homomorphism of $\mathfrak{B}^{(m)}$ into $\mathfrak{B}$ which is the identity on $\mathfrak{B}$. It is also clear that $a^{s\pi} = a$ for every $a \in \mathfrak{A}$ and that this requirement guarantees that s is 1 - 1.

Conversely, let s be any homomorphism of $\mathfrak{A}$ into $\mathfrak{B}^{(m)}$ such that $a^{s\pi} = a$, for all $a \in \mathfrak{A}$. Then we write $a^s = a + (aD_1)t + \ldots + (aD_m)t^m$. We now claim that $\{1, D_1, \ldots, D_m\}$ is a higher derivation of $\mathfrak{A}$ into $\mathfrak{B}$. This follows immediately since s is linear implies that every $D_i$ is linear and we have

$$a^s b^s = \sum_{i=0}^{m} (aD_i)t^i \sum_{k=0}^{m} (bD_k)t^k$$

$$= \sum_{j=0}^{m} \sum_{i=0}^{j} (aD_i)(bD_{j-i})t^j$$

$$= (ab)^s = \sum_{j=0}^{m} (ab)D_j t^j ,$$

implies the relation (27).

Similar considerations apply to higher derivations of infinite rank. The place of $\mathfrak{B}^{(m)}$ is now taken by the algebra $\mathfrak{B}[[t]]$ of power series

(30) $a_0 + a_1 t + a_2 t^2 + \cdots$

where the $a_i \in \mathfrak{B}$ (cf. Vol. I, p. 95).[1]

Definition 4.2. If $\{D_i\}$ is a higher derivation of rank m (possibly infinite) of $\mathfrak{A}$ into $\mathfrak{B}$, an element $c \in \mathfrak{A}$ is a constant relative to $\{D_i\}$ if $c D_i = 0$ for all $i \geq 1$.

We shall often refer to constants relative to $\{D_i\}$ simply as $\{D_i\}$-constants. We observe that if c is a $\{D_i\}$-constant, $c^s = c$ (and conversely) for the homomorphism associated with $\{D_i\}$. Let $c_1, c_2 \in \mathfrak{A}/\Phi$, be $\{D_i\}$-constants. Then

$$(c_1 \pm c_2)^s = c_1^s \pm c_2^s = c_1 \pm c_2, \quad (c_1 c_2)^s = c_1^s c_2^s = c_1 c_2,$$

and, if $\alpha \in \Phi$, $\left(c_1(\alpha c_2)\right)^s = c_1^s (\alpha c_2)^s = c_1(\alpha c_2)$. Moreover, since $1^s = 1$, the set of $\{D_i\}$-constants form a subalgebra (with identity) of the algebra $\mathfrak{A}/\Phi$. In particular, if $\mathfrak{A} = P$ is a field and $b \neq 0$ in P, then $(cb^{-1})^s = c^s(b^{-1})^s = c^s(b^s)^{-1} = cb^{-1}$. Therefore the set of $\{D_i\}$-constants form a subfield $\Gamma$ of P. It is clear that $\Gamma \supseteq \Phi$ since $\alpha D_i = (1\cdot\alpha)D_i = 1\, D_i\alpha = 0$ for all $\alpha \in \Phi$.

We suppose now that $P/\Phi$ is a field of characteristic $p \neq 0$. Let E be a subfield of $P/\Phi$ and let $D^{(m)} = \{1, D_1, \ldots, D_m\}$ be a higher derivation of rank m of $E/\Phi$ into $P/\Phi$. In general, if $0 = D_1 = D_2 = \cdots = D_{q-1}$, but $D_q \neq 0$, we shall say that $D^{(m)}$ is of

---

[1] Jacobson [9], p. 193.

order q. In this case, the associated homomorphism $s = s(D^{(m)})$ of $E/\Phi$ into $P^{(m)}$ has the form

$$(31) \quad \zeta^s = \zeta + (\zeta D_q)t^q + (\zeta D_{q+1})t^{q+1} + \ldots + (\zeta D_m)t^m \quad ,$$

where $\zeta D_q \neq 0$ for some $\zeta$ in E. We shall use this to prove the following theorem (cf. Jacobson [9], p. 194).

Theorem 4.A. Let $P/\Phi$ be a field of characteristic $p \neq 0$. Let E be a subfield of $P/\Phi$, $D^{(m)}$ a higher derivation of rank m and order q of $E/\Phi$ into $P/\Phi$. Let $\Gamma$ be the subfield of $D^m$-constants of E and let $p^e$ be the smallest power of p such that $p^e > \frac{m}{q}$. Then E is purely inseparable of exponent e over $\Gamma$.

Proof. We have to show that $\zeta^{p^e} \in \Gamma$ for every $\zeta \in E$ and that there exists some $\zeta \in E$ such that $\zeta^{p^{e-1}} \notin \Gamma$. The first part is clear from (31) since

$$\begin{aligned}(\zeta^{p^e})^s = (\zeta^s)^{p^e} &= \left(\zeta + (\zeta D_q)t^q + \ldots + (\zeta D_m)t^m\right)^{p^e} \\ &= \zeta^{p^e} + (\zeta D_q)^{p^e} t^{qp^e} + \ldots + (\zeta D_m)^{p^e} t^{mp^e} \\ &= \zeta^{p^e} \quad , \quad \text{since } t^{qp^e} = \ldots = t^{mp^e} = 0 \text{ follows}\end{aligned}$$

immediately from the fact that $q\,p^e > m$ and $t^k = 0$ for all $k > m$. This shows that $\zeta^{p^e} \in \Gamma$. Let us now choose an element $\zeta \in E$ which is not a $D^m$-constant. In this case, $\zeta D_q \neq 0$ and

$$\begin{aligned}(\zeta^{p^{e-1}})^s &= \zeta^{p^{e-1}} + (\zeta D_q)^{p^{e-1}} t^{qp^{e-1}} + \ldots \\ &\neq \zeta^{p^{e-1}} \quad \text{since } q\,p^{e-1} \leq m. \text{ Hence } \zeta^{p^{e-1}} \notin \Gamma.\end{aligned}$$

We shall now consider a purely inseparable extension field

$P = \Phi(\xi)$ where $x^{p^e} - \alpha$ is the minimal polynomial of $\xi$ over $\Phi$. Let $\{D_i\}$ be the higher derivation in the polynomial algebra $\Phi[x]$ defined by (28) and let $D^{(p^e-1)} = \{1, D_1, \ldots, D_{p^e-1}\}$ be the higher derivation of rank $p^e - 1$ which is a section of this higher derivation $\{D_i\}$. Since each $D_j \in \{D_i\}$ is linear, $[x^{p^e} - \alpha]D_j = x^{p^e}D_j - \alpha D_j = 0 - 0 = 0$ for $1 \leq j \leq p^e - 1$. We have seen that $P = \Phi(\xi)$ can be identified with $\Phi[x]/(x^{p^e} - \alpha)$ where $(x^{p^e} - \alpha)$ is the principal ideal generated by $x^{p^e} - \alpha$ and that $1 + (x^{p^e} - \alpha), \ldots, x^{p^e-1} + (x^{p^e} - \alpha)$ then form a basis over $\Phi$ for $\Phi(\xi)$. Moreover, if we write $x^{p^e} - \alpha = h(x)$ and $f(x) \in \Phi[x]$ is arbitrarily chosen, then the defining relations (27) show that

$$\Big(f(x)\ h(x)\Big)\ D_j = \sum_{i=0}^{j} \Big(f(x)\ D_i\Big)\Big(h(x)\ D_{j-i}\Big)$$

$$= \Big(f(x)\ D_j\Big)\ h(x)\ .$$

This implies that this principal ideal is mapped into itself by every $D_j$. Hence every $D_j$ induces a linear mapping, which we denote again by $D_j$, in $P = \Phi(\xi)$. The conditions in $\Phi[x]$ for $D_j$ go over to the same conditions (27) for $D_j$ in $\Phi(\xi)$ since

$$x^{m+n}D_j = \sum_{i=0}^{j} (x^m D_i)(x^n D_{j-i})$$

holds in $\Phi[x]$ and each $D_j$ is linear. Hence we obtain a higher derivation $D^{(p^e-1)}$ in $\Phi(\xi)$ such that

$$(32) \qquad \xi^m D_i = \binom{m}{i} \xi^{m-i}\ , \qquad m = 0, 1, \ldots, p^e - 1\ .$$

We shall now show that the subfield $\Gamma$ of $\{D_i\}$-constants for

$D^{(p^e-1)}$ is precisely $\Phi$. Suppose that $\Phi$ is properly contained in $\Gamma$ (i.e., $\Phi \subset \Gamma$). Then the minimal polynomial of $\xi$ over $\Gamma$ is $x^{p^f} - \beta$, $\beta \in \Gamma$, $f < e$. In this case, $\xi^{p^f} \in \Gamma$. On the other hand, the definition (32) gives $\xi^{p^f} D_{p^f} = 1 \neq 0$. This is a contradiction since $\xi^{p^f} \in \Gamma$ implies that $\xi^{p^f} D_{p^f} = 0$. Hence $\Phi = \Gamma$, since we already have $\Phi \subseteq \Gamma$.

We next assume that P is a purely inseparable extension of $\Phi$ given by $P = P_1 \otimes P_2 \otimes \ldots \otimes P_r$ the tensor product of simple extensions $P_i = \Phi(\xi_i)$, $\xi_i^{p^e} = \alpha_i \in \Phi$. Then $P = \Phi(\xi_1, \ldots, \xi_r)$ and the monomials $\xi_1^{k_1} \ldots \xi_r^{k_r}$, $0 \leq k_i < p^{e_i}$, form a basis for P over $\Phi$. If we set $\Phi_i = \Phi(\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_r)$, then $P = \Phi_i(\xi_i)$. By the above argument, there exists a higher derivation of finite rank in P whose constants are the elements of $\Phi_i$. This statement holds for $i = 1, 2, \ldots, r$. Hence it is clear that $\bigcap_{i=1}^{r} \Phi_i = \Phi$ is the subfield of P of elements which are constants relative to all the higher derivations of finite rank in P over $\Phi$.

Definition 4.3. A higher derivation in $\mathfrak{A}$ of infinite rank is called <u>iterative</u> if $D_i D_j = \binom{i+j}{i} D_{i+j}$, and a higher derivation $D^{(m)} = \{D_0 = 1, D_1, \ldots, D_m\}$ of finite rank m is called <u>iterative</u> if (1) $D_i D_j = \binom{i+j}{j} D_{i+j}$ for $i + j \leq m$, and (2) $D_i D_j = 0$ if $i + j > m$.

Remark 4.1. The higher derivations defined by (28) and (32) are iterative.

Proof. Since the mappings $D_i$ defined by (28) constitute a higher derivation $\{1, D_1, D_2, \ldots\}$ of infinite rank in $\Phi[x]$, we have to show that $D_i D_j = \binom{i+j}{i} D_{i+j}$. This follows readily since

$$x^m D_i D_j = \left(\binom{m}{i} x^{m-i}\right) D_j = \binom{m}{i}\binom{m-i}{j} x^{m-i-j}$$

$$= \frac{m!}{i!(m-i)!}\,\frac{(m-i)!}{j!(m-i-j)!}\, x^{m-i-j}$$

$$= \frac{(i+j)!}{i!j!}\,\frac{m!}{(i+j)!\,(m-i-j)!}\, x^{m-i-j} = \binom{i+j}{i} x^m D_{i+j}$$

Secondly, the mappings $D_i$ defined by (32) constitute a higher derivation $D^{(p^e-1)}$ of rank $p^e - 1$ over $\Phi(\xi)$. It is clear that $\xi^m D_i D_j = \binom{m}{i}\left(\xi^{m-i} D_j\right) = \binom{i+j}{i} \xi^m D_{i+j}$, as above. This implies that $D_i D_j = \binom{i+j}{i} D_{i+j}$, $i + j \le m \le p^e - 1$. Since the fields of constants for $D^{(p^e-1)}$ is $\Phi$, we must have $D_i D_j = 0$ for $i + j > p^e - 1$.

This proves that the higher derivations defined by (28) and (32) are iterative.

BIBLIOGRAPHY


1. CHAUNDY, T. The Differential Calculus, Oxford, 1935.

2. GERSTENHABER, M. On The Galois Theory of Inseparable Extensions, Bull. Amer. Math. Soc. vol. 70, 1964, pp. 561-566.

3. HERSTEIN, I. N. Topics in Algebra, Blaisdell, 1964.

4. HOCHSCHILD, G. Double Vector Spaces over Division Rings, Amer. Jour. Math. vol. 71, 1949, pp. 443-460.

5. Simple Algebras with Purely Inseparable Splitting Fields of Exponent One, Trans. Amer. Math. Soc. vol. 79, 1955, pp. 477-489.

6. JACOBSON, N. Abstract Derivations and Lie Algebras, Trans. Amer. Soc. vol. 42, 1937, pp. 206-224.

7. Lectures in Abstract Algebra, vol. 1 - Basic Concepts, Van Nostrand, 1951.

8. Lectures in Abstract Algebra, vol. 11 - Linear Algebra, Van Nostrand, 1953.

9. Lectures in Abstract Algebra, vol. 111 - Theory of Fields and Galois Theory, Van Nostrand, 1964.

10. Lie Algebras, John Wiley, 1962.

11. SCOTT, R. F., and MATTHEWS, G. B. The Theory of Determinants and Their Applications, 2nd. edition, revised, Cambridge, 1904.

12. ZARISKI, O., and SAMUEL, P., Commutative Algebra, vol. 1, Van Nostrand, 1958.