Section 5.5  More about Roots

We return to the general exposition. Let $F$ be any field and, as usual, let $F[x]$ be the ring of polynomials in $x$ over $F$.

**Definition:** If

$$f(x) = \alpha_0 x^n + \alpha_1 x^{n-1} + \ldots + \alpha_i x^{n-i} + \ldots + \alpha_{n-1} x + \alpha_n$$

in $F[x]$, then the *derivative* of $f(x)$, written as $f'(x)$, is the polynomial

$$f'(x) = n\alpha_0 x^{n-1} + (n-1)\alpha_1 x^{n-2} + \ldots + (n-i)\alpha_i x^{n-i-1} + \ldots + \alpha_{n-1}$$

in $F[x]$.

To make this definition or to prove the basic formal properties of the derivative, as applied to polynomials, does not require the concept of a limit. However, since the field $F$ is arbitrary, we might expect some strange things to happen.

At the end of Section 3.2, we defined what is meant by characteristic of a field. Let us recall it now. A field $F$ is said to be of characteristic 0 if $ma \neq 0$ for $a \neq 0$ in $F$ and $m > 0$, an integer. If $ma = 0$ for some $m > 0$ and some $a \neq 0 \in F$, then $F$ is said to be of finite characteristic. In this second case, the characteristic of $F$ is defined to be the smallest positive integer $p$ such that $pa = 0$ for all $a \in F$. It turned out that if $F$ is of finite characteristic then its characteristic $p$ is a prime number.

We return to the question of the derivative. Let $F$ be a field of characteristic $p \neq 0$. In this case, the derivative of the polynomial $x^p$ is $px^{p-1} = 0$. Thus the usual result from the calculus that a polynomial whose derivative is 0 must be a constant no longer need hold true. However, if the characteristic of $F$ is 0 and if $f'(x) = 0$ for $f(x) \in F[x]$, it is indeed true that $f(x) = \alpha \in F$ (see Problem 1). Even when the characteristic of $F$ is $p \neq 0$, we can still describe the polynomials with zero derivative; if $f'(x) = 0$, then $f(x)$ is a polynomial in $x^p$ (see Problem 2).

We now prove the analogs of the formal rules of differentiation that we know so well.

**Lemma 5.5.1:** For any $f(x), g(x) \in F[x]$ and any $\alpha \in F$,

1. $(f(x) + g(x))' = f'(x) + g'(x)$.
2. $(\alpha f(x))' = \alpha f'(x)$.
3. $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.

**Proof:** The proofs of parts 1 and 2 are extremely easy and are left as exercises. To prove part 3, note that from parts 1 and 2 it is enough to prove it in the highly special case $f(x) = x^i$ and $g(x) = x^j$ where both $i$ and $j$ are positive. But then $f(x)g(x) = x^{i+j}$, whence $(f(x)g(x))' = (i+j)x^{i+j-1}$; however,

$$f'(x)g(x) = ix^{i-1}x^j = ix^{i+j-1} \quad \text{and} \quad f(x)g'(x) = jx^ix^{j-1} = jx^{i+j-1};$$

consequently,

$$f'(x)g(x) + f(x)g'(x) = (i+j)x^{i+j-1} = (f(x)g(x))'.$$

\[\square\]
Recall that in elementary calculus the equivalence is shown between the existence of a multiple root of a function and the simultaneous vanishing of the function and its derivative at a given point. Even in our setting, where $F$ is an arbitrary field, such an interrelation exists.

**Lemma 5.5.2:** The polynomial $f(x) \in F[x]$ has a multiple root if and only if $f(x)$ and $f'(x)$ have a nontrivial (that is, of positive degree) common factor.

**Proof:** Before proving the lemma proper, a related remark is in order, namely, if $f(x)$ and $g(x)$ in $F[x]$ have a nontrivial common factor in $K[x]$, for $K$ an extension of $F$, then they have a nontrivial common factor in $F[x]$. For, were they relatively prime as elements in $F[x]$, then we would be able to find two polynomials $a(x)$ and $b(x)$ in $F[x]$ such that

$$a(x)f(x) + b(x)g(x) = 1.$$  

Since this relation also holds for those elements viewed as elements of $K[x]$, in $K[x]$ they would have to be relatively prime.

Now to the lemma itself. From the remark just made, we may assume, without loss of generality, that the roots of $f(x)$ all lie in $F$ (otherwise extend $F$ to $K$, the splitting field of $f(x)$). If $f(x)$ has a multiple root $\alpha$, then $f(x) = (x - \alpha)^mq(x)$, where $m > 1$. However, as is easily computed, $((x - \alpha)^m)' = m(x - \alpha)^{m-1}$, whence, by Lemma 5.5.1,

$$f'(x) = (x - \alpha)^mq'(x) + m(x - \alpha)^{m-1}q(x) = (x - \alpha)r(x),$$

since $m > 1$. But this says that $f(x)$ and $f'(x)$ have the common factor $x - \alpha$, thereby proving the lemma in one direction.

On the other hand, if $f(x)$ has no multiple root then

$$f(x) = (x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n)$$

where the $\alpha$’s are all distinct (we are supposing $f(x)$ to be monic). But then

$$f'(x) = \sum_{i=1}^{n}(x - \alpha_1)\ldots(\overline{x - \alpha_i})\ldots(x - \alpha_n)$$

where the $\wedge$ denotes the term is omitted. We claim no root of $f(x)$ is a root of $f'(x)$, for

$$f'(\alpha_i) = \prod_{j \neq i}(\alpha_i - \alpha_j) \neq 0,$$

since the roots are all distinct. However, if $f(x)$ and $f'(x)$ have a nontrivial common factor, they have a common root, namely, any root of this common factor. The net result is that $f(x)$ and $f'(x)$ have no nontrivial common factor, and so the lemma has been proved in the other direction. $\blacksquare$

**Corollary 1:** If $f(x) \in F[x]$ is irreducible, then

1. If the characteristic of $F$ is 0, $f(x)$ has no multiple roots.

2. If the characteristic of $F$ is $p \neq 0$, $f(x)$ has a multiple root only if it is of the form $f(x) = g(x^p)$. 

2
Proof. Since $f(x)$ is irreducible, its only factors in $F[x]$ are 1 and $f(x)$. If $f(x)$ has a multiple root, then $f(x)$ and $f'(x)$ have a nontrivial common factor by the lemma, hence $f(x) \mid f'(x)$. However, since the degree of $f'(x)$ is less than that of $f(x)$, the only possible way that this can happen is for $f'(x)$ to be 0. In characteristic 0 this implies that $f(x)$ is a constant, which has no roots; in characteristic $p \neq 0$, this forces $f(x) = g(x^p)$. □

We shall return in a moment to discuss the implications of Corollary 1 more fully. But first, for later use in Chapter 7 in our treatment of finite fields, we prove the rather special

**COROLLARY 2**: If $F$ is a field of characteristic $p \neq 0$, then the polynomial $x^p - x \in F[x]$, for $n \geq 1$, has distinct roots.

**Proof.** The derivative of $x^p - x$ is $p^n x^{p-1} - 1 = -1$, since $F$ is of characteristic $p$. Therefore, $x^p - x$ and its derivative are certainly relatively prime, which, by the lemma, implies that $x^p - x$ has no multiple roots. □

Corollary 1 does not rule out the possibility that in characteristic $p \neq 0$ an irreducible polynomial might have multiple roots. To clinch matters, we exhibit an example where this actually happens. Let $F_0$ be a field of characteristic 2 and let $F = F_0(x)$ be the field of rational functions in $x$ over $F_0$. We claim that the polynomial $t^2 - x \in F[t]$ is irreducible over $F$ and that its roots are equal. To prove irreducibility we must show that there is no rational function in $F_0(x)$ whose square is $x$; this is the content of Problem 4. To see that $t^2 - x$ has a multiple root, notice that its derivative (the derivative is with respect to $t$; for $x$, being in $F$, is considered as a constant) is $2t = 0$. Of course, the analogous example works for any prime characteristic.

Now that the possibility has been seen to be an actuality, it points out a sharp difference between the case of characteristic 0 and that of characteristic $p$. The presence of irreducible polynomials with multiple roots in the latter case leads to many interesting, but at the same time complicating, subtleties. These require a more elaborate and sophisticated treatment which we prefer to avoid at this stage of the game. Therefore, we make the flat assumption for the rest of this chapter that all fields occurring in the text material proper are fields of characteristic 0.

**DEFINITION:** The extension $K$ of $F$ is a simple extension of $F$ if $K = F(\alpha)$ for some $\alpha$ in $K$.

In characteristic 0 (or in properly conditioned extensions in characteristic $p \neq 0$; see Problem 14) all finite extensions are realizable as simple extensions. This result is

**THEOREM 5.5.1**: If $F$ is of characteristic 0 and if $a, b$, are algebraic over $F$, then there exists an element $c \in F(a, b)$ such that $F(a, b) = F(c)$.

**Proof.** Let $f(x)$ and $g(x)$, of degrees $m$ and $n$, be the irreducible polynomials over $F$ satisfied by $a$ and $b$, respectively. Let $K$ be an extension of $F$ in which both $f(x)$ and $g(x)$ split completely. Since the characteristic of $F$ is 0, all the roots of $f(x)$ are distinct, as are all those of $g(x)$. Let the roots of $f(x)$ be $a = a_1, a_2, \ldots, a_m$ and those of $g(x)$, $b = b_1, b_2, \ldots, b_n$.

If $j \neq 1$, then $b_j \neq b_1 = b$, hence the equation $a_i + \lambda b_j = a_1 + \lambda b_1 = a + \lambda b$ has only one solution $\lambda$ in $K$, namely,

$$\lambda = \frac{a_i - a}{b - b_j}.$$

Since $F$ is of characteristic 0 it has an infinite number of elements, so we can find an element
\[ \gamma \in F \text{ such that } a_i + \gamma b_j \neq a + \gamma b \text{ for all } i \text{ and for all } j \neq 1. \text{ Let } c = a + \gamma b; \text{ our contention is that } F(c) = F(a, b). \text{ Since } c \in F(a, b), \text{ we certainly do have that } F(c) \subset F(a, b). \text{ We will now show that both } a \text{ and } b \text{ are in } F(c) \text{ from which it will follow that } F(a, b) \subset F(c). \]

Now \( b \) satisfies the polynomial \( g(x) \) over \( F \), hence satisfies \( g(x) \) considered as a polynomial over \( K = F(c) \). Moreover, if \( h(x) = f(c - \gamma x) \) then \( h(x) \in K[x] \) and

\[
h(b) = f(c - \gamma b) = f(a) = 0,
\]

since \( a = c - \gamma b \). Thus in some extension of \( K \), \( h(x) \) and \( g(x) \) have \( x - b \) as a common factor. We assert that \( x - b \) is in fact their greatest common divisor. For, if \( b_j \neq b \) is another root of \( g(x) \), then \( h(b_j) = f(c - \gamma b_j) \neq 0 \), since by our choice of \( \gamma \), \( c - \gamma b_j \) for \( j \neq 1 \) avoids all roots \( a_i \) of \( f(x) \). Also, since \( (x - b)^2 \nmid g(x) \), \( (x - b)^2 \) cannot divide the greatest common divisor of \( h(x) \) and \( g(x) \). Thus \( x - b \) is the greatest common divisor of \( h(x) \) and \( g(x) \) over some extension of \( K \). But then they have a nontrivial greatest common divisor over \( K \), which must be a divisor of \( x - b \). Since the degree of \( x - b \) is 1, we see that the greatest common divisor of \( g(x) \) and \( h(x) \) in \( K[x] \) is exactly \( x - b \). Thus \( x - b \in K[x] \), whence \( b \in K \); remembering that \( K = F(c) \), we obtain that \( b \in F(c) \). Since \( a = c - \gamma b \), and since \( b, c \in F(c) \), \( \gamma \in F \subset F(c) \), we get that \( a \in F(c) \), whence \( F(a, b) \subset F(c) \). The two opposite containing relations combine to yield \( F(a, b) = F(c) \). \( \blacksquare \)

A simple induction argument extends the result from 2 elements to any finite number, that is, if \( \alpha_1, \ldots, \alpha_n \) are algebraic over \( F \), then there is an element \( c \in F(\alpha_1, \ldots, \alpha_n) \) such that \( F(c) = F(\alpha_1, \ldots, \alpha_n) \). Thus the

COROLLARY: Any finite extension of a field of characteristic 0 is a simple extension.