3 Derivatives and multiple roots

We begin by recalling the definition of a repeated root.

Definition 3.1. Let $F$ be a field and let $\alpha \in F$. Then there is a unique integer $m \geq 0$ such that $(x - \alpha)^m$ divides $f$ but $(x - \alpha)^{m+1}$ does not divide $f$. We define this integer $m$ to be the multiplicity of the root $\alpha$ in $f$. Note that, by the correspondence between roots of a polynomial and its linear factors, $\alpha$ has multiplicity 0 in $f$, i.e. $m = 0$ above, $\iff f(\alpha) \neq 0$. More generally, if $\alpha$ has multiplicity $m$ in $f$, then $f = (x - \alpha)^mg$ with $g(\alpha) \neq 0$, and conversely.

If $\alpha$ has multiplicity 1 in $f$, we call $\alpha$ a simple root of $f$. If $\alpha$ has multiplicity $m \geq 2$ in $f$, then we call $\alpha$ a multiple root or repeated root of $f$.

We would like to find conditions when a nonconstant polynomial does, or does not have a multiple root in $F$ or in some extension field $E$ of $F$. To do so, we introduce the formal derivative:

Definition 3.2. Let $F$ be a field. Define the function $D: F[x] \to F[x]$ by the formula

$$D(\sum_{i=0}^{n} a_i x^i) = \sum_{i=1}^{n} ia_i x^{i-1}.$$ 

Here the notation $ia_i$ means the ring element $i \cdot a_i = a_i + \cdots + a_i$, with the convention that $0a_0 = 0$. We usually write $D(f)$ as $Df$. Note that either $Df = 0$ or $\deg Df \leq \deg f - 1$.

Clearly, the function $D$ is compatible with field extension, in the sense that, if $F \leq E$, then we have $D: F[x] \to F[x]$ and $D: E[x] \to E[x]$, and given $f \in F[x]$, $Df$ is the same whether we view $f$ as an element of $F[x]$ or of $E[x]$. Also, an easy calculation shows that:
Proposition 3.3. \( D: F[x] \to F[x] \) is \( F \)-linear. \( \square \)

This result is equivalent to the sum rule: for all \( f, g \in F[x] \), \( D(f + g) = Df + Dg \) as well as the constant multiple rule: for all \( f \in F[x] \) and \( c \in F \), \( D(cf) = cDf \). Once we know that \( D \) is \( F \)-linear, it is specified by the fact \( D(1) = 0 \) and, that, for all \( i > 0 \), \( Dx^i = ix^{i-1} \). Also, viewing \( D \) as a homomorphism of abelian groups, we can try to compute

\[
\text{Ker} \, D = \{ f \in F[x] : Df = 0 \}.
\]

Our expectation from calculus is that a function whose derivative is 0 is a constant. But if \( \text{char} \, F = p > 0 \), something strange happens:

**Proposition 3.4.** If \( \text{Ker} \, D = \{ f \in F[x] : Df = 0 \} \), then

\[
\text{Ker} \, D = \begin{cases} F, & \text{if } \text{char} \, F = 0; \\ F[x^p], & \text{if } \text{char} \, F = p > 0. \end{cases}
\]

Here \( F[x^p] = \{ \sum_{i=0}^{n} a_i x^{ip} : a_i \in F \} \) is the subring of all polynomials in \( x^p \).

**Proof.** Clearly, \( f = \sum_{i=0}^{n} a_i x^i \) is in \( \text{Ker} \, D \iff \) for every \( i \) such that the coefficient \( a_i \) is nonzero, the monomial \( ix^{i-1} = 0 \). In case \( \text{char} \, F = 0 \), this is only possible if \( i = 0 \), in other words \( f \in F \) is a constant polynomial. In case \( \text{char} \, F = p > 0 \), this happens exactly when \( p|i \) for every \( i \) such that \( a_i \neq 0 \). This is equivalent to saying that \( f \) is a polynomial in \( x^p \). \( \square \)

As is well-known in calculus, \( D \) is not a ring homomorphism. In other words, the derivative of a product of two polynomials is not in general the product of the derivatives. Instead we have:

**Proposition 3.5** (The product rule). For all \( f, g \in F[x] \),

\[
D(f \cdot g) = Df \cdot g + f \cdot Dg.
\]

**Proof.** If \( f = x^a \) and \( g = x^b \), then we can verify this directly:

\[
D(x^a x^b) = D(x^{a+b}) = (a + b)x^{a+b-1}; \\
(Dx^a)x^b + x^a(Dx^b) = ax^{a-1}x^b + bx^a x^{b-1} = (a + b)x^{a+b-1}.
\]

The general case follows from this by writing \( f \) and \( g \) as sums of monomials and expanding (but is a little messy to write down). Another approach using formal difference quotients is in the HW. \( \square \)
If $R$ is a ring, a function $d: R \to R$ which is an additive homomorphism (i.e. $d(r + s) = d(r) + d(s)$ for all $r, s \in R$) satisfying $d(rs) = d(r)s + rd(s)$ for all $r, s \in R$ is called a derivation of $R$. Thus, $D$ is a derivation of $F[x]$.

As a corollary of the product rule, we obtain:

**Corollary 3.6 (The power rule).** For all $f \in F[x]$ and $n \in \mathbb{N}$, 

$$D(f)^n = n(f)^{n-1}Df.$$ 

**Proof.** This is an easy induction using the product rule and starting with the case $n = 1$ (or 0). $\square$

The connection between derivatives and multiple roots is as follows:

**Lemma 3.7.** Let $f \in F[x]$ be a nonconstant polynomial and let $E$ be an extension field of $F$. Then $\alpha \in E$ is a multiple root of $f$ $\iff$ $f(\alpha) = Df(\alpha) = 0$.

**Proof.** Write $f = (x - \alpha)^mg$ with $m$ equal to the multiplicity of $\alpha$ in $f$ and $g \in F[x]$ a polynomial such that $g(\alpha) \neq 0$. If $m = 0$, then $f(\alpha) = g(\alpha) \neq 0$. Otherwise, 

$$Df = m(x - \alpha)^{m-1}g + (x - \alpha)^mDg.$$ 

If $m = 1$, then $Df(\alpha) = g(\alpha) \neq 0$. If $m \geq 2$, then $f(\alpha) = Df(\alpha) = 0$. Thus we see that $\alpha \in F$ is a multiple root of $f$ $\iff$ $m \geq 2$ $\iff$ $f(\alpha) = Df(\alpha) = 0$. $\square$

In practice, an (unknown) root of $f$ will only exist in some (unknown) extension field $E$ of $F$. We would like to have a criterion for when a polynomial $f$ has some multiple root $\alpha$ in some extension field $E$ of $F$, without having to know what $E$ and $\alpha$ are explicitly. In order to find such a criterion, we begin with the following lemma, which says essentially that divisibility, greatest common divisors, and relative primality are unchanged after passing to extension fields.

**Lemma 3.8.** Let $E$ be an extension field of a field $F$, and let $f, g \in F[x], \text{not both } 0$.

(i) $f | g$ in $F[x] \iff f | g$ in $E[x]$.

(ii) The polynomial $d \in F[x]$ is a gcd of $f, g$ in $F[x] \iff d$ is a gcd of $f, g$ in $E[x]$.

(iii) The polynomials $f, g$ are relatively prime in $F[x] \iff f, g$ are relatively prime in $E[x]$. 

15
Proof. (i): \(\Rightarrow\) : obvious. \(\Leftarrow\) : We can assume that \(f \neq 0\), since otherwise \(f|g\) (in either \(F[x]\) or \(E[x]\)) \(\iff\) \(g = 0\). Suppose that \(f|g\) in \(E[x]\), i.e. that \(g = fh\) for some \(h \in E[x]\). We must show that \(h \in F[x]\).

By long division with remainder in \(F[x]\), there exist \(q,r \in F[x]\) with either \(r = 0\) or \(\deg r < \deg f\), such that \(g = fq + r\). Now, in \(E[x]\), we have both \(g = fh\) and \(g = fq + r\). By uniqueness of long division with remainder in \(E[x]\), we must have \(h = q\) (and \(r = 0\)). In particular, \(h = q \in F[x]\), as claimed.

(ii): \(\Rightarrow\) : Let \(d \in F[x]\) be a gcd of \(f, g\) in \(F[x]\). Then, by (i), since \(d|f\), \(d|g\) in \(F[x]\), \(d[f], d[g]\) in \(E[x]\). Moreover, there exist \(a, b \in F[x]\) such that \(d = af + bg\). Now suppose that \(e \in E[x]\) and that \(e|f, e|g\) in \(E[x]\). Then \(e|af + bg = d\). It follows that \(d\) satisfies the properties of being a gcd in \(E[x]\). \(\Leftarrow\) : Let \(d \in F[x]\) be a gcd of \(f, g\) in \(E[x]\). Then \(d|f, d|g\) in \(E[x]\), hence by (i) \(d|f, d|g\) in \(F[x]\). Suppose that \(e \in F[x]\) and that \(e|f, e|g\) in \(F[x]\). Then \(e|f, e|g\) in \(E[x]\). Hence \(e|d\) in \(E[x]\). Since both \(e, d \in F[x]\), it again follows by (i) that \(e|d\) in \(F[x]\). Thus \(d\) is a gcd of \(f, g\) in \(F[x]\).

(iii): The polynomials \(f, g\) are relatively prime in \(F[x]\) \(\iff\) \(1 \in F[x]\) is a gcd of \(f\) and \(g\) in \(F[x]\) \(\iff\) \(1 \in F[x]\) is a gcd of \(f\) and \(g\) in \(F[x]\), by (ii), \(\iff\) \(f, g\) are relatively prime in \(E[x]\).

\(\Box\)

Corollary 3.9. Let \(f \in F[x]\) be a nonconstant polynomial. Then there exists an extension field \(E\) of \(F\) and a multiple root of \(f\) in \(E\) \(\iff\) \(f\) and \(Df\) are not relatively prime in \(F[x]\).

Proof. \(\Rightarrow\) : If \(E\) and \(\alpha\) exist, then, by Lemma 3.7, \(f\) and \(Df\) have a common factor \(x - \alpha\) in \(E[x]\) and hence are not relatively prime. Thus by Lemma 3.8 \(f\) and \(Df\) are not relatively prime in \(F[x]\).

\(\Leftarrow\) : Suppose that \(f\) and \(Df\) are not relatively prime in \(F[x]\), and let \(g\) be a common nonconstant factor of \(f\) and \(Df\). There exists an extension field \(E\) of \(F\) and an \(\alpha \in E\) which is a root of \(g\). Then \(\alpha\) is a common root of \(f\) and \(Df\), and hence a multiple root of \(f\). \(\Box\)

We now apply the above to an irreducible polynomial \(f \in F[x]\).

Corollary 3.10. Let \(f \in F[x]\) be an irreducible polynomial. Then there exists an extension field \(E\) of \(F\) and a multiple root of \(f\) in \(E\) \(\iff\) \(Df = 0\).

Proof. \(\Rightarrow\) : By the previous corollary, if there exists an extension field \(E\) of \(F\) and a multiple root of \(f\) in \(E\), then \(f\) and \(Df\) are not relatively prime in \(F[x]\). In this case, since \(f\) is irreducible, it must be that \(f\) divides \(Df\). Hence, if \(Df \neq 0\), then \(\deg Df \geq \deg f\). But we have seen that either \(\deg Df < \deg f\) or \(Df = 0\). Thus, we must have \(Df = 0\).
Clearly, if $Df = 0$, then $f$ is a gcd of $f$ and $Df$, hence $f$ and $Df$ are not relatively prime in $F[x]$. \qed

**Corollary 3.11.** Let $F$ be a field of characteristic 0 and let $f \in F[x]$ be an irreducible polynomial. Then there does not exist an extension field $E$ of $F$ and a multiple root of $f$ in $E$. In particular, if $E$ is an extension field of $F$ such that $f$ factors into linear factors in $E$, say

$$f = c(x - \alpha_1) \cdots (x - \alpha_n),$$

then the $\alpha_i$ are distinct, i.e. for $i \neq j$, $\alpha_i \neq \alpha_j$. \qed

If $\text{char} F = p > 0$, then it is possible for an irreducible polynomial $f \in F[x]$ to have a multiple root in some extension field, but it takes a little effort to produce such examples. For example, it is not possible to find such an example for a finite field. The basic example arises as follows: consider the field $F_p(t)$, where $t$ is an indeterminate (here we could replace $F_p$ by any field of characteristic $p$). Then $t$ is not a $p^{\text{th}}$ power in $F_p(t)$, and in fact one can show that the polynomial $x^p - t$ is irreducible in $F_p(t)[x]$. Let $E$ be an extension field of $F_p(t)$ which contains a root $\alpha$ of $x^p - t$, so that by definition $\alpha^p = t$. Then

$$x^p - t = x^p - \alpha^p = (x - \alpha)^p,$$

because we are in characteristic $p$. Thus $\alpha$ is a multiple root of $x^p - t$, of multiplicity $p$.

The key property of the field $F_p(t)$ which made the above example work was that $t$ was not a $p^{\text{th}}$ power in $F_p(t)$. More generally, define a field $F$ of characteristic $p$ to be **perfect** if every element of $F$ is a $p^{\text{th}}$ power, or equivalently if the Frobenius homomorphism $\sigma_p : F \to F$ is surjective. For example, we shall show below that a finite field is perfect. An algebraically closed field is also perfect. We also declare every field of characteristic zero to be perfect. By a problem on HW, if $F$ is a perfect field and $f \in F[x]$ is an irreducible polynomial, then there does not exist an extension field $E$ of $F$ and a multiple root of $f$ in $E$.\[17]