

4.5 The Formal Derivative

The polynomials of this chapter are coefficient sequences, and Theorem 4.4.2 showed how to evaluate them and how to substitute into them. One more operation from school survives the passage to an arbitrary commutative ring: the derivative. This section defines it by its coefficients, proves the sum, product and power rules with the move of Lemma 4.2.3, and asks when a derivative can be 0. Over a field the answer depends on the characteristic of Chapter 2, and over $\mathbb{F}_p$ it is not the answer of calculus.

At school a double root shows itself in the graph. The parabola $y = (x - 1)^2$ touches the axis at 1 instead of crossing it, and its slope there is 0: the derivative $2(x - 1)$ vanishes at 1 as well. Over $\mathbb{F}_p$ there is no graph and no slope; the field has p elements and no limits. What survives is the formula: the derivative of ax^n is nax^{n-1} . We keep the formula and drop the limit. Chapter 5 uses the result to detect repeated roots; this section builds the tool.

Throughout, R is a commutative ring (Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Definition 2.3.1). Polynomials are written f, g, h , with coefficients f_k, g_k, h_k , because the letter p is kept for a prime. A dot with an integer on its left is a **multiple**, $n \cdot a = a + \cdots + a$ with n terms (the “Multiples” paragraph of *Number Systems: From $\mathbb{N}$ to $\mathbb{R}$* , Section 2.3, collected in Lemma 4.1.1 (d)); the product of two ring elements is written by juxtaposition.

4.5.1 The derivative by coefficients

At school, $(3x^4 + 5x^2)' = 12x^3 + 10x$. Each term moves down one place, and its coefficient is multiplied by the exponent it had. In $R[x]$ “multiplied by the exponent” has a meaning without any division or limit: the exponent n is a natural number, and $n \cdot a$ is the n -th multiple of a . That recurring computation deserves a name.

The formal derivative moves every coefficient down one place and takes it as many times as its old exponent.

Definition 4.5.1. Formal Derivative

Let R be a commutative ring and $f \in R[x]$. The **formal derivative** of f is the sequence $f': \mathbb{N} \rightarrow R$ given by

$$(f')_k := (k + 1) \cdot f_{k+1} \quad \text{for every } k \in \mathbb{N},$$

where $(k + 1) \cdot f_{k+1}$ is the $(k + 1)$ -th multiple of f_{k+1} . The **higher derivatives** are defined by

$$f^{(0)} := f, \quad f^{(k+1)} := (f^{(k)})' \quad (k \in \mathbb{N}),$$

and we write f'' for $f^{(2)}$.

In words: the coefficient of x^k in f' is read off from the coefficient of x^{k+1} in f , added to itself $k + 1$ times. The definition uses only the addition of R , so it makes sense, word for word, over $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Z}/6\mathbb{Z}$ and $\mathbb{F}_p$.

Well-definedness. Two things need checking: that f' is again a polynomial, and that the higher derivatives exist.

- $f' \in R[x]$. First, $m \cdot 0 = 0$ for every $m \in \mathbb{N}$: by Lemma 4.1.1 (d), $m \cdot 0 = (m \cdot 1)0$, and this is 0 by Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Proposition 2.3.2 (b). Now choose $N \in \mathbb{N}$ with $f_k = 0$ for every $k > N$ (Definition 4.1.1). For $k \geq N$ we have $k + 1 > N$,

so $f_{k+1} = 0$ and $(f')_k = (k+1) \cdot 0 = 0$. So f' has only finitely many non-zero terms, and $f' \in R[x]$. Moreover $(f')_k = 0$ for $k > N-1$, so for $N \geq 1$, Proposition 4.2.1 writes f' as an honest sum:

$$f' = \sum_{k=0}^{N-1} ((k+1) \cdot f_{k+1})x^k = 1 \cdot f_1 + (2 \cdot f_2)x + \cdots + (N \cdot f_N)x^{N-1}.$$

- **Higher derivatives.** Since $u \mapsto u'$ is a function $R[x] \rightarrow R[x]$, Sets, Relations, and Cardinality, Theorem 2.3.1, with $X = R[x]$, starting value f and step $(k, u) \mapsto u'$, gives exactly one sequence $k \mapsto f^{(k)}$ with the two displayed properties.

Monomials. Let $a \in R$ and $n \in \mathbb{N}$. The monomial ax^n has a in position n and 0 elsewhere (Definition 4.1.1), so $(ax^n)_{k+1}$ is a when $k+1 = n$ and 0 otherwise. Hence $((ax^n)')_k$ is $n \cdot a$ when $k = n-1$, and $(k+1) \cdot 0 = 0$ otherwise:

$$(ax^n)' = (n \cdot a)x^{n-1} \quad (n \geq 1), \quad (ax^0)' = 0. \quad (\text{M})$$

The case $n = 0$ is separate because x^{-1} has no meaning there: $-1 \notin \mathbb{N}$. Every constant has derivative 0.

The element $n \cdot 1$. By Lemma 4.1.1 (d), $n \cdot a = (n \cdot 1)a$. So the exponent enters the derivative only through one element of R , the multiple $n \cdot 1$, and **that element can be 0**. In $\mathbb{F}_2$, $2 \cdot 1 = 1 + 1 = 0$, so $(x^2)' = (2 \cdot 1)x = 0$. This is the whole difference between the derivative here and the derivative of calculus, and the rest of the section keeps track of it.

Why this definition. The formula is the school formula with the limit removed. Exercise C1 asks you to show that it is the only additive map $R[x] \rightarrow R[x]$ that sends constants to 0, sends x to 1 and obeys the product rule. A definition by limits is not available, since $\mathbb{F}_p$ and $\mathbb{Z}/6\mathbb{Z}$ have no distances.

Warning. The coefficient of x^k in f' is $(k+1) \cdot f_{k+1}$, not $k \cdot f_k$: the derivative shifts before it multiplies. For $f = x^2$ over $\mathbb{Q}$, the coefficient of x^1 in $f' = 2x$ is $2 \cdot f_2 = 2$, while $1 \cdot f_1 = 0$. Reading the formula without the shift gives $\sum_k (k \cdot f_k)x^k$, which is xf' , not f' .

Example. Formal derivatives in five rings

Compute the following derivatives.

- (i) $(x^4 - 2x^3 + 2x - 1)'$ over $\mathbb{Q}$.
- (ii) a' , x' and $(x - c)'$ over any commutative ring R , for $a, c \in R$.
- (iii) $(x^3 + x)'$ over $\mathbb{F}_3$, and $(x^2)'$ over $\mathbb{F}_2$.
- (iv) $([2]x^3 + x)'$ over $\mathbb{Z}/6\mathbb{Z}$.
- (v) $(x^n)'$ over $\mathbb{Z}$, for $n \geq 1$.
- (vi) The higher derivatives $(x^4)'$, $(x^4)''$ and $(x^4)'''$ over $\mathbb{Q}$ and over $\mathbb{F}_3$.

Solution. In each case, by (M) and Lemma 4.1.1 (d), the derivative of ax^n with $n \geq 1$ is $((n \cdot 1)a)x^{n-1}$, and the derivative of a sum is computed coefficient by coefficient from Definition 4.5.1.

- (i) The coefficients are $f_0 = -1$, $f_1 = 2$, $f_2 = 0$, $f_3 = -2$, $f_4 = 1$, and $f_k = 0$ for $k \geq 5$. So $(f')_0 = 1 \cdot 2 = 2$, $(f')_1 = 2 \cdot 0 = 0$, $(f')_2 = 3 \cdot (-2) = -6$, $(f')_3 = 4 \cdot 1 = 4$, and $(f')_k = 0$

for $k \geq 4$. In $\mathbb{Q}$ the multiple $n \cdot a$ is the product na : $n \cdot a = (n \cdot 1)a$ by Lemma 4.1.1 (d), and $n \cdot 1$ is the rational number n , by induction on n from $(n+1) \cdot 1 = n \cdot 1 + 1$, since the embeddings $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$ preserve sums and 1 (Number Systems: From $\backslash(\backslash n \mathbb{N} \backslash)$ to $\backslash(\backslash n \mathbb{R} \backslash)$, Proposition 2.2.3, Number Systems: From $\backslash(\backslash n \mathbb{N} \backslash)$ to $\backslash(\backslash n \mathbb{R} \backslash)$, Proposition 2.4.4). Hence

$$(x^4 - 2x^3 + 2x - 1)' = 4x^3 - 6x^2 + 2,$$

as at school.

- (ii) $a = ax^0$, so $a' = 0$ by (M). $x = 1x^1$ (Definition 4.1.1), so $x' = (1 \cdot 1)x^0 = 1$, since $1 \cdot 1 = 0 \cdot 1 + 1 = 1$ by the recursion of multiples. The polynomial $x - c$ has $-c$ in position 0, 1 in position 1 and 0 elsewhere, so $((x-c)')_0 = 1 \cdot 1 = 1$, and $((x-c)')_k = (k+1) \cdot 0 = 0$ for $k \geq 1$: $(x-c)' = 1$.
- (iii) Over $\mathbb{F}_3$, $(x^3 + x)' = (3 \cdot 1)x^2 + 1 = 1$, since $3 \cdot 1 = 1 + 1 + 1 = 0$ in $\mathbb{F}_3$ (Number Systems: From $\backslash(\backslash n \mathbb{N} \backslash)$ to $\backslash(\backslash n \mathbb{R} \backslash)$, Theorem 3.4.3 (d): $p \cdot [1] = [p] = [0]$ in $\mathbb{Z}/p\mathbb{Z}$). Over $\mathbb{F}_2$, $(x^2)' = (2 \cdot 1)x = 0$, since $1 + 1 = 0$ in $\mathbb{F}_2$ (Example (Arithmetic on $\mathbb{Z}/n\mathbb{Z}$ and the Field $\mathbb{F}_2$)).
- (iv) $([2]x^3 + x)' = (3 \cdot [2])x^2 + 1 = [6]x^2 + 1 = 1$, since $3 \cdot [2] = (3 \cdot [1])[2] = [3][2] = [6] = [0]$, by Lemma 4.1.1 (d), Number Systems: From $\backslash(\backslash n \mathbb{N} \backslash)$ to $\backslash(\backslash n \mathbb{R} \backslash)$, Theorem 3.4.3 (d), and the product of classes computed on representatives (Number Systems: From $\backslash(\backslash n \mathbb{N} \backslash)$ to $\backslash(\backslash n \mathbb{R} \backslash)$, Theorem 3.4.1 (b)). The polynomial has degree 3 and its derivative has degree 0: the degree dropped by 3, not by 1.
- (v) $(x^n)' = (n \cdot 1)x^{n-1} = nx^{n-1}$, since in $\mathbb{Z}$ the multiple $n \cdot 1$ is the integer n (the “Multiples” paragraph cited above).
- (vi) Apply (M) three times, with the rule $m \cdot (j \cdot 1) = (mj) \cdot 1$ of Lemma 4.1.1 (d). Over any commutative ring,

$$(x^4)' = (4 \cdot 1)x^3, \quad (x^4)'' = (3 \cdot (4 \cdot 1))x^2 = (12 \cdot 1)x^2, \quad (x^4)''' = (2 \cdot (12 \cdot 1))x = (24 \cdot 1)x.$$

Over $\mathbb{Q}$, as in (i), these are $4x^3$, $12x^2$ and $24x$. Over $\mathbb{F}_3$, $3 \cdot 1 = 0$, so $4 \cdot 1 = 3 \cdot 1 + 1 \cdot 1 = 1$ and $12 \cdot 1 = 4 \cdot (3 \cdot 1) = 4 \cdot 0 = 0$, by Lemma 4.1.1 (d) and $m \cdot 0 = 0$. So $(x^4)' = x^3$, but already $(x^4)'' = 0$, and then $(x^4)''' = 0' = 0$ by (M).

Examples (iii) and (iv) are the new phenomena. Over $\mathbb{F}_2$ a non-constant polynomial has derivative 0; over $\mathbb{Z}/6\mathbb{Z}$ the coefficient of x^2 in the derivative, $3 \cdot [2]$, is $[0]$ although neither $3 \cdot [1] = [3]$ nor $[2]$ is. Both come from the element $n \cdot 1$ above.

Quick check. Compute $(x^3 + x^2 + x + 1)'$ over $\mathbb{F}_2$.

Solution. $(3 \cdot 1)x^2 + (2 \cdot 1)x + 1 = x^2 + 1$, since $3 \cdot 1 = 1 + 1 + 1 = 1$ and $2 \cdot 1 = 0$ in $\mathbb{F}_2$.

4.5.2 The rules of differentiation

Calculus proves the sum, product and power rules with limits. Here they must come from the coefficients. The sum rule is one line per coefficient. The product rule is where a direct attack hurts: the k -th coefficient of $(fg)'$ is $(k+1) \cdot \sum_{i=0}^{k+1} f_i g_{k+1-i}$, and matching it with $f'g + fg'$ means

splitting each term between two sums with shifted indices. The move of Section 4.2, **check on monomials, extend by additivity**, avoids that: both sides of the product rule are additive in f and in g , so it is enough to compare them on monomials, where (M) does the work.

Theorem 4.5.1. Rules for the Formal Derivative

Let R be a commutative ring, $f, g \in R[x]$, $a, c \in R$ and $m \in \mathbb{Z}_{>0}$.

- (a) (Linearity) $(f + g)' = f' + g'$, and $(af)' = af'$ for the constant a ; moreover $a' = 0$.
- (b) (Product rule) $(fg)' = f'g + fg'$.
- (c) (Power rule) $(g^m)' = (m \cdot 1)g^{m-1}g'$. In particular $((x - c)^m)' = (m \cdot 1)(x - c)^{m-1}$.

Idea. ① Linearity is one coefficient at a time: the multiple $(k + 1) \cdot b$ is additive in b and commutes with multiplication by a constant. ② For the product rule, both $(fg)'$ and $f'g + fg'$ are additive in each argument, by ①; so Lemma 4.2.3 reduces the rule to $f = ax^i$, $g = bx^j$. There (M) and Lemma 4.2.2 (b) turn both sides into one monomial, and the coefficients agree because $(i + j) \cdot (ab) = i \cdot (ab) + j \cdot (ab)$. ③ The power rule is the product rule applied $m - 1$ times, by induction.

Proof. Equality of polynomials is equality of all coefficients (Definition 4.1.1), so it is enough to compare coefficients in (a); (b) and (c) are then built on (a).

- (a) Let $k \in \mathbb{N}$. Since addition in $R[x]$ is coefficientwise (Definition 4.2.1),

$$\begin{aligned} ((f + g)')_k &= (k + 1) \cdot (f_{k+1} + g_{k+1}) \\ &= (k + 1) \cdot f_{k+1} + (k + 1) \cdot g_{k+1} = (f' + g')_k, \end{aligned}$$

where the second equality is the rule $n \cdot (b + d) = n \cdot b + n \cdot d$ of Lemma 4.1.1 (d). For the constant a , $(af)_{k+1} = af_{k+1}$ by Theorem 4.2.4 (d), so

$$\begin{aligned} ((af)')_k &= (k + 1) \cdot (f_{k+1}a) \\ &= ((k + 1) \cdot f_{k+1})a = a(f')_k = (af')_k, \end{aligned}$$

using (F6) of R in the first and third equalities, the rule $(n \cdot b)d = n \cdot (bd)$ of Lemma 4.1.1 (d) in the second, and Theorem 4.2.4 (d) in the last. Finally $a' = 0$ is (M) with $n = 0$.

- (b) Define $\Phi, \Psi: R[x] \times R[x] \rightarrow R[x]$ by

$$\Phi(f, g) := (fg)', \quad \Psi(f, g) := f'g + fg'.$$

By Theorem 4.2.4 (a), $R[x]$ is a commutative ring, so (F1)–(F7), (F9) and Number Systems: From $\setminus(\setminus n \mathbb{N} \setminus)$ to $\setminus(\setminus n \mathbb{R} \setminus)$, Proposition 2.3.2 hold in it.

Additivity. Let $f, h, g \in R[x]$. By the distributive law (F9) of $R[x]$ and (a),

$$\Phi(f + h, g) = (fg + hg)' = (fg)' + (hg)' = \Phi(f, g) + \Phi(h, g).$$

By (a), (F9) and a regrouping of four terms by (F1) and (F2),

$$\begin{aligned} \Psi(f + h, g) &= (f' + h')g + (f + h)g' \\ &= (f'g + fg') + (h'g + hg') = \Psi(f, g) + \Psi(h, g). \end{aligned}$$

Both maps are symmetric: $\Phi(f, g) = \Phi(g, f)$ because $fg = gf$, and $\Psi(f, g) = \Psi(g, f)$ because $f'g + fg' = gf' + g'f = g'f + gf'$, by (F6) and (F2) of $R[x]$. So each is additive in its second argument as well.

Monomials. Let $u = ax^i$ and $v = bx^j$, with $a, b \in R$ and $i, j \in \mathbb{N}$. By Lemma 4.2.2 (b), $uv = (ab)x^{i+j}$.

Case $i \geq 1$ and $j \geq 1$. Then $i + j \geq 1$, and (M) gives

$$\Phi(u, v) = ((i + j) \cdot (ab))x^{i+j-1}.$$

On the other side, (M) gives $u' = (i \cdot a)x^{i-1}$ and $v' = (j \cdot b)x^{j-1}$, and Lemma 4.2.2 (b), with $(i - 1) + j = i + (j - 1) = i + j - 1$ in $\mathbb{N}$, gives

$$\Psi(u, v) = ((i \cdot a)b)x^{i+j-1} + (a(j \cdot b))x^{i+j-1}.$$

Now $(i \cdot a)b = i \cdot (ab)$, and $a(j \cdot b) = (j \cdot b)a = j \cdot (ba) = j \cdot (ab)$, by Lemma 4.1.1 (d) and (F6) of R . Two monomials in the same position add coefficientwise (Definition 4.2.1), and $i \cdot (ab) + j \cdot (ab) = (i + j) \cdot (ab)$ (Lemma 4.1.1 (d)). So $\Psi(u, v) = \Phi(u, v)$.

Case $i = 0$. Then $u = a$ is a constant, $u' = 0$ by (a), and $uv = (ab)x^j$. So $\Psi(u, v) = 0 \cdot v + av' = av'$, by Number Systems: From $\setminus(\setminus \mathbb{N} \setminus)$ to $\setminus(\setminus \mathbb{N} \setminus)$, Proposition 2.3.2 (b) and (F3) of $R[x]$. If $j = 0$, then $v' = 0$ and uv is a constant, so both sides are 0 (by (a) and Number Systems: From $\setminus(\setminus \mathbb{N} \setminus)$ to $\setminus(\setminus \mathbb{N} \setminus)$, Proposition 2.3.2 (b)). If $j \geq 1$, then by (M) and Lemma 4.2.2 (b) with $m = 0$,

$$av' = (a(j \cdot b))x^{j-1} = (j \cdot (ab))x^{j-1} = ((ab)x^j)' = \Phi(u, v),$$

where the middle step is the computation of the previous case.

Case $j = 0, i \geq 1$. By the symmetry of Φ and Ψ this is the case $i = 0$ with u and v exchanged.

So Φ and Ψ agree whenever both arguments are monomials, and both are additive in each argument. By Lemma 4.2.3, with two arguments and $T = R[x]$, $\Phi = \Psi$. This is the product rule.

- (c) First, $(f - g)' = f' - g'$: by (a), $(f - g)' + g' = ((f - g) + g)' = f'$, and adding $-g'$ to both sides gives the claim, by (F2)–(F4) of $R[x]$. We prove the power rule by induction on $m \geq 1$ (Logic and Proof, Axiom 4.2.1, with base 1). For $m = 1$: $g^0 = 1$ (Lemma 4.1.1 (c)) and $1 \cdot 1 = 1$, so the right side is $1 \cdot 1 \cdot g' = g'$. Suppose $(g^m)' = (m \cdot 1)g^{m-1}g'$ for some $m \geq 1$. Since $g^{m+1} = g^m g$, the product rule (b) gives

$$\begin{aligned} (g^{m+1})' &= (g^m)'g + g^m g' = (m \cdot 1)g^{m-1}g'g + g^m g' \\ &= (m \cdot 1)g^m g' + 1 \cdot g^m g' = ((m + 1) \cdot 1)g^m g'. \end{aligned}$$

The third equality uses (F5), (F6), $g^{m-1}g = g^m$ (the power recursion, since $(m - 1) + 1 = m$) and (F7); the last uses (F9), the fact that constants add as in R (Theorem 4.2.4 (b)), and the recursion $(m + 1) \cdot 1 = m \cdot 1 + 1$ of multiples. This completes the induction.

For the particular case, $(x - c)' = x' - c' = 1 - 0 = 1$, by the difference rule and [Example \(Formal derivatives in five rings\)](#) (ii). The power rule with $g = x - c$ gives $((x - c)^m)' = (m \cdot 1)(x - c)^{m-1} \cdot 1$, and $\cdot 1$ may be dropped by (F7). This proves the theorem. ■

The theorem is the calculus toolkit with every limit removed, and it holds over every commutative ring. The power rule shows the element $m \cdot 1$ once more: the derivative of $(x - c)^m$ is

$(x - c)^{m-1}$ times the constant $m \cdot 1$, and over $\mathbb{F}_p$ that constant is 0 for $m = p$. So over $\mathbb{F}_p$ the polynomial $(x - c)^p$ has derivative 0, although it is not constant; the next subsection says exactly when this happens.

Remark. Over $\mathbb{Q}$, and over the real numbers, f' is the derivative of the polynomial function $c \mapsto f(c)$ in the sense of calculus. This book never uses the agreement, and nothing in it depends on it; it explains only why the formulas look familiar.

Quick check. Over $\mathbb{F}_3$, compute $((x + 1)^3)'$ twice: by the power rule, and by expanding $(x + 1)^3$ first.

Solution. By Theorem 4.5.1 (c), $((x + 1)^3)' = (3 \cdot 1)(x + 1)^2 = 0$, since $3 \cdot 1 = 0$ in $\mathbb{F}_3$. Expanding, $(x + 1)^3 = x^3 + 3x^2 + 3x + 1 = x^3 + 1$ in $\mathbb{F}_3[x]$, and $(x^3 + 1)' = (3 \cdot 1)x^2 = 0$. The two answers agree.

4.5.3 When the derivative vanishes

In calculus, a function with derivative 0 on an interval is constant. Over $\mathbb{F}_2$ the polynomial x^2 is not constant, yet $(x^2)' = 0$ (Example (Formal derivatives in five rings) (iii)). Which polynomials have derivative 0? Over a field the answer depends only on the characteristic. Recall Definition 2.5.1: $\text{char } F$ is the least positive integer n with $n \cdot 1 = 0$ in F , and 0 if there is none; by Theorem 2.5.3 it is 0 or a prime.

Proposition 4.5.2. When the Formal Derivative Is Zero

Let F be a field (Definition 2.1.1) and $f \in F[x]$.

- (a) If $\text{char } F = 0$, then $f' = 0$ **if and only if** f is constant. Moreover, $\deg f' = \deg f - 1$ whenever $\deg f \geq 1$.
- (b) If $\text{char } F = p > 0$, then p is prime, and the following are equivalent: (i) $f' = 0$; (ii) $f_k = 0$ for every $k \in \mathbb{N}$ **not** divisible by p ; (iii) $f = g \circ x^p$ for some $g \in F[x]$ (Definition 4.4.3).

Idea. Everything rests on one identity: for $k \geq 1$, the coefficient of x^{k-1} in f' is $k \cdot f_k = (k \cdot 1)f_k$. In a field a product is 0 only when a factor is. So $f' = 0$ forces $f_k = 0$ exactly at the k with $k \cdot 1 \neq 0$: in characteristic 0 that is every $k \geq 1$, and in characteristic p it is every k not divisible by p . A polynomial whose only non-zero coefficients sit at multiples of p is a polynomial in x^p .

Proof. For every $k \geq 1$, Definition 4.5.1 at the index $k - 1$ and Lemma 4.1.1 (d) give

$$(f')_{k-1} = k \cdot f_k = (k \cdot 1)f_k. \quad (*)$$

As k runs through the integers ≥ 1 , the index $k - 1$ runs through $\mathbb{N}$, so $f' = 0$ **if and only if** $(k \cdot 1)f_k = 0$ for every $k \geq 1$.

- (a) Let $\text{char } F = 0$. Then $k \cdot 1 \neq 0$ for every $k \geq 1$, by Definition 2.5.1. Suppose $f' = 0$, and let $k \geq 1$. By (*), $(k \cdot 1)f_k = 0$ with $k \cdot 1 \neq 0$, so $f_k = 0$ by Theorem 2.1.1 (f). Thus f and the constant f_0 agree in every position, and $f = f_0$ (Definition 4.1.1). Conversely, a constant has derivative 0 by Theorem 4.5.1 (a).

Now let $n := \deg f \geq 1$, so that $f_n \neq 0$ and $f_k = 0$ for every $k > n$ (Definition 4.3.1). By (*), $(f')_{n-1} = (n \cdot 1)f_n$, a product of two non-zero elements of F , which is non-zero by Theorem 2.1.1 (f). For $k \geq n$, $(f')_k = (k + 1) \cdot f_{k+1} = (k + 1) \cdot 0 = 0$, as in the well-definedness paragraph after Definition 4.5.1. So $n - 1$ is the largest index of a non-zero coefficient of f' : $\deg f' = n - 1$.

- (b) Let $\text{char } F = p > 0$. Then p is prime by Theorem 2.5.3.

By Proposition 2.5.2 (a), with $c = p$: for $k \in \mathbb{N}$, $k \cdot 1 = 0$ **if and only if** $p \mid k$.

(i) $\Rightarrow$ (ii). Let $f' = 0$ and let $k \in \mathbb{N}$ with $p \nmid k$. Then $k \neq 0$, since $p \mid 0$; so $k \geq 1$, and $(k \cdot 1)f_k = 0$ by (*). By Proposition 2.5.2 (a), $k \cdot 1 \neq 0$, so $f_k = 0$ by Theorem 2.1.1 (f).

(ii) $\Rightarrow$ (i). Let $k \geq 1$. If $p \nmid k$, then $f_k = 0$ by (ii); if $p \mid k$, then $k \cdot 1 = 0$ by Proposition 2.5.2 (a). Either way $(k \cdot 1)f_k = 0$ (Number Systems: From $\backslash(\backslash\mathbb{N}\backslash)$ to $\backslash(\backslash\mathbb{N}\mathbb{R}\backslash)$, Proposition 2.3.2 (b)), so $f' = 0$ by (*).

(ii) $\Rightarrow$ (iii). Choose $N \in \mathbb{N}$ with $f_k = 0$ for every $k > N$, and define $g: \mathbb{N} \rightarrow F$ by $g_j := f_{pj}$. For $j > N$ we have $pj \geq j > N$, so $g_j = 0$, and $g \in F[x]$. By Definition 4.4.3, then the power law $(x^p)^j = x^{pj}$ in the commutative ring $F[x]$ (Lemma 4.1.1 (c)) and Theorem 4.2.4 (c),

$$g \circ x^p = \sum_{j=0}^N g_j (x^p)^j = \sum_{j=0}^N f_{pj} x^{pj}.$$

We compare the coefficients of both sides with those of f at an index $k \in \mathbb{N}$. Addition in $F[x]$ is coefficientwise (Definition 4.2.1), so the k -th coefficient of a finite sum of polynomials is the finite sum of their k -th coefficients (induction on the number of terms, through the sum recursion of Lemma 4.1.1 (a)). The monomial $f_{pj}x^{pj}$ has k -th coefficient f_k if $pj = k$, and 0 otherwise.

- If $p \nmid k$, no j has $pj = k$, so every term is 0 at k , and $f_k = 0$ by (ii).
- If $p \mid k$, then $k = pj_0$ for some $j_0 \in \mathbb{Z}$, and $j_0 \geq 0$: if $j_0 < 0$, then $pj_0 < 0 \cdot p = 0$ by Number Systems: From $\backslash(\backslash\mathbb{N}\backslash)$ to $\backslash(\backslash\mathbb{N}\mathbb{R}\backslash)$, Proposition 2.3.5 (c), since $p > 0$, against $k \geq 0$. This j_0 is unique (if $pj = pj'$, then $j = j'$ by Number Systems: From $\backslash(\backslash\mathbb{N}\backslash)$ to $\backslash(\backslash\mathbb{N}\mathbb{R}\backslash)$, Proposition 2.3.5 (f), since $p \neq 0$). If $j_0 \leq N$, the sum of the k -th coefficients has the single possibly non-zero term $f_{pj_0} = f_k$, so it equals f_k by ($\Sigma 3$) of Lemma 4.1.1 (b). If $j_0 > N$, every term is 0 at k , and $f_k = 0$ because $k = pj_0 \geq j_0 > N$.

In every case the k -th coefficients agree, so $g \circ x^p = f$.

(iii) $\Rightarrow$ (ii). Let $f = g \circ x^p$, and choose N with $g_j = 0$ for $j > N$. The computation above gives $f = \sum_{j=0}^N g_j x^{pj}$, and by the first bullet, the coefficient of this sum at any k with $p \nmid k$ is 0. This proves the proposition. ■

So in characteristic p the polynomials with derivative 0 are exactly the polynomials in x^p . Over $\mathbb{F}_2$, $x^2 + 1 = (x + 1) \circ x^2$; over $\mathbb{F}_3$, $x^6 + 2x^3 + 1 = (x^2 + 2x + 1) \circ x^3$, and indeed its derivative

is $(6 \cdot 1)x^5 + (6 \cdot 1)x^2 = 0$, since $6 \cdot 1 = 0$ in $\mathbb{F}_3$. Part (a) is the case where the school statement survives, and its proof spends the hypothesis at exactly one place: $k \cdot 1 \neq 0$.

Warning. “ $f' = 0$ means f is constant” and “differentiating lowers the degree by one” both fail in characteristic p . Over $\mathbb{F}_p$, $(x^p)' = (p \cdot 1)x^{p-1} = 0$, although x^p has degree p ; and $(x^p - x)' = (p \cdot 1)x^{p-1} - 1 = -1$ has degree 0, not $p - 1$. Tests for repeated roots that use the derivative must therefore say which characteristic they assume. Chapter 5 concludes from $(x^p - x)' = -1$ that $x^p - x$ has no repeated roots in $\mathbb{F}_p$.

Quick check. Compute $(x^5 + x)'$ over $\mathbb{F}_5$.

Solution. $(x^5 + x)' = (5 \cdot 1)x^4 + 1 = 1$, since $5 \cdot 1 = 0$ in $\mathbb{F}_5$.

4.5.4 Exercises

4.5.4.1 A. Check your understanding

Exercise. A1

- Compute $(x^5 - x)'$ over $\mathbb{Q}$ and over $\mathbb{F}_5$.
- True or false: if $f' = 0$, then f is constant. Answer over $\mathbb{Q}$ and over $\mathbb{F}_2$, with a reason or a witness.
- Let $f \in (\mathbb{Z}/6\mathbb{Z})[x]$ have degree 3 and leading coefficient $[2]$. What can $\deg f'$ be? Give a witness for each possibility.
- In the proof of the product rule ([Theorem 4.5.1 \(b\)](#)), why is the case $i = 0$ treated separately?

Solution.

- Over $\mathbb{Q}$: $(x^5 - x)' = 5x^4 - 1$, by (M), with $5 \cdot 1 = 5$. Over $\mathbb{F}_5$: $(x^5 - x)' = (5 \cdot 1)x^4 - 1 = -1$, since $5 \cdot 1 = 0$ in $\mathbb{F}_5$; in the elements $0, \dots, 4$ of $\mathbb{F}_5$, $-1 = 4$.
- Over $\mathbb{Q}$: **true**. $\mathbb{Q}$ is a field (Number Systems: From $\mathbb{N}$ to $\mathbb{R}$), Theorem 2.4.3) of characteristic 0 (Example (The characteristics of the standard fields)); so [Proposition 4.5.2 \(a\)](#) applies. Over $\mathbb{F}_2$: **false**. The witness is x^2 , which is not constant, with $(x^2)' = (2 \cdot 1)x = 0$ ([Example \(Formal derivatives in five rings\)](#) (iii)).
- $\deg f' \in \{1, 0, -\infty\}$, and each occurs. Write $f = [2]x^3 + f_2x^2 + f_1x + f_0$. Then

$$f' = (3 \cdot [2])x^2 + (2 \cdot f_2)x + f_1 = (2 \cdot f_2)x + f_1,$$

since $3 \cdot [2] = [6] = [0]$. So $\deg f' \leq 1$ (Definition 4.3.1). Witnesses: $[2]x^3 + x^2$ has derivative $[2]x$, of degree 1; $[2]x^3 + x$ has derivative 1, of degree 0; $[2]x^3$ has derivative 0, of degree $-\infty$. The school rule “degree drops by one” fails because $3 \cdot [2] = [0]$: the leading coefficient of f' would be $(3 \cdot 1) \text{lc}(f)$, and here that product is 0.

- For $i = 0$ the monomial $u = ax^0$ is a constant, and (M) gives $u' = 0$, not $(0 \cdot a)x^{-1}$: the

exponent $i - 1 = -1$ is not in $\mathbb{N}$, so x^{i-1} has no meaning. The general computation, which writes $u' = (i \cdot a)x^{i-1}$, therefore cannot be used, and the case is checked directly.

4.5.4.2 B. Practice

Exercise. B1: The chain rule

Let R be a commutative ring and $f, g \in R[x]$. Prove that

$$(f \circ g)' = (f' \circ g) g'.$$

Check the formula on $f = x^2$, $g = x^2 + 1$ over $\mathbb{Z}$.

Solution. Fix g , and define $\Phi, \Psi: R[x] \rightarrow R[x]$ by

$$\Phi(f) := (f \circ g)', \quad \Psi(f) := (f' \circ g) g'.$$

The move is the one of the product rule: check on monomials, extend by additivity.

Additivity. By Definition 4.4.3 and Theorem 4.4.2 (a), $u \mapsto u \circ g$ is a ring homomorphism $R[x] \rightarrow R[x]$ that fixes the constants and sends x to g ; in particular $(f + h) \circ g = f \circ g + h \circ g$. With Theorem 4.5.1 (a),

$$\Phi(f + h) = (f \circ g + h \circ g)' = \Phi(f) + \Phi(h),$$

and, using also (F9) of $R[x]$,

$$\Psi(f + h) = (f' \circ g + h' \circ g) g' = \Psi(f) + \Psi(h).$$

Monomials. Let $f = ax^n$ with $a \in R$, $n \in \mathbb{N}$. By Theorem 4.2.4 (c), ax^n is the product of the constant a and the n -th power of x ; a ring homomorphism preserves products and powers (Definition 4.4.2), so $(ax^n) \circ g = ag^n$.

If $n = 0$, then $f = a$ and $f \circ g = a$, so $\Phi(f) = 0$ by Theorem 4.5.1 (a). Also $f' = 0$, and $0 \circ g = 0$ because a ring homomorphism sends 0 to 0 (Definition 4.4.2); so $\Psi(f) = 0 \cdot g' = 0$.

If $n \geq 1$, then by Theorem 4.5.1 (a) and (c),

$$\Phi(f) = (ag^n)' = a(g^n)' = a(n \cdot 1)g^{n-1}g'.$$

On the other side $f' = (n \cdot a)x^{n-1}$ by (M), so $f' \circ g = (n \cdot a)g^{n-1}$ and $\Psi(f) = (n \cdot a)g^{n-1}g'$. The constants agree: $a(n \cdot 1) = (n \cdot 1)a = n \cdot a$, by (F6) and Lemma 4.1.1 (d), and constants multiply as in R (Theorem 4.2.4 (b)). So $\Phi(f) = \Psi(f)$.

By Lemma 4.2.3 with one argument and $T = R[x]$, $\Phi = \Psi$, which is the chain rule.

Check. Over $\mathbb{Z}$, $f \circ g = (x^2 + 1)^2 = x^4 + 2x^2 + 1$, whose derivative is $4x^3 + 4x$. On the other side, $f' = 2x$, $f' \circ g = 2(x^2 + 1)$ and $g' = 2x$, so $(f' \circ g)g' = 4x^3 + 4x$.

Exercise. B2: Higher derivatives

For $n, k \in \mathbb{N}$ write $P(n, k)$ for the product $n(n-1) \cdots (n-k+1)$ of k factors, defined by $P(n, 0) := 1$ and $P(n, k+1) := P(n, k)(n-k)$; it is a natural number for $k \leq n$.

- (a) Let R be a commutative ring. Show that $(x^n)^{(k)} = (P(n, k) \cdot 1)x^{n-k}$ for $k \leq n$, and $(x^n)^{(k)} = 0$ for $k > n$.
- (b) Let p be a prime. Show that $f^{(p)} = 0$ for every $f \in \mathbb{F}_p[x]$.

Solution.

- (a) Induction on $k \leq n$. For $k = 0$, $(x^n)^{(0)} = x^n = (1 \cdot 1)x^n$. Let $k < n$ and suppose $(x^n)^{(k)} = (P(n, k) \cdot 1)x^{n-k}$. Since $n-k \geq 1$, (M) gives

$$\begin{aligned} (x^n)^{(k+1)} &= ((n-k) \cdot (P(n, k) \cdot 1))x^{n-k-1} \\ &= (P(n, k+1) \cdot 1)x^{n-(k+1)}, \end{aligned}$$

by the rule $m \cdot (j \cdot b) = (mj) \cdot b$ of Lemma 4.1.1 (d) and $(n-k)P(n, k) = P(n, k+1)$ in $\mathbb{N}$. This proves the formula for $k \leq n$. At $k = n$ the polynomial $(x^n)^{(n)}$ is a constant, so $(x^n)^{(n+1)} = 0$ by (M), and the derivative of 0 is 0; by induction, $(x^n)^{(k)} = 0$ for every $k > n$.

- (b) *The coefficients of $f^{(k)}$.* Let R be any commutative ring and $f \in R[x]$. We show by induction on k that

$$(f^{(k)})_j = P(j+k, k) \cdot f_{j+k} \quad \text{for every } j \in \mathbb{N}.$$

For $k = 0$ both sides are f_j . If it holds for k , then by Definition 4.5.1 and Lemma 4.1.1 (d),

$$\begin{aligned} (f^{(k+1)})_j &= (j+1) \cdot (P(j+k+1, k) \cdot f_{j+k+1}) \\ &= (P(j+k+1, k)(j+1)) \cdot f_{j+k+1} = P(j+k+1, k+1) \cdot f_{j+k+1}, \end{aligned}$$

since the factor added by the recursion is $(j+k+1) - k = j+1$.

The case $k = p$ over $\mathbb{F}_p$. The number $P(j+p, p) = (j+p)(j+p-1) \cdots (j+1)$ is a product of the p consecutive integers $j+1, \dots, j+p$. Divide $j = qp+r$ with $0 \leq r < p$ (Number Systems: From $\backslash(\mathbb{N}\backslash)$ to $\backslash(\mathbb{N}\mathbb{R}\backslash)$, Theorem 3.1.2); then $j + (p-r) = (q+1)p$, and $1 \leq p-r \leq p$, so one of the factors is divisible by p . A product with a factor divisible by p is divisible by p (Number Systems: From $\backslash(\mathbb{N}\backslash)$ to $\backslash(\mathbb{N}\mathbb{R}\backslash)$, Proposition 3.1.1 (c), one factor at a time). So $P(j+p, p) = Mp$ with $M \in \mathbb{Z}$, and

$$P(j+p, p) \cdot f_{j+p} = M \cdot (p \cdot f_{j+p}) = M \cdot ((p \cdot 1)f_{j+p}) = M \cdot 0 = 0,$$

by Lemma 4.1.1 (d), which allows integer multiples, $p \cdot 1 = 0$ in $\mathbb{F}_p$ (Number Systems: From $\backslash(\mathbb{N}\backslash)$ to $\backslash(\mathbb{N}\mathbb{R}\backslash)$, Theorem 3.4.3 (d)), Number Systems: From $\backslash(\mathbb{N}\backslash)$ to $\backslash(\mathbb{N}\mathbb{R}\backslash)$, Proposition 2.3.2 (b) and $M \cdot 0 = 0$. So every coefficient of $f^{(p)}$ is 0. For instance, over $\mathbb{F}_3$, $(x^4)''' = (P(4, 3) \cdot 1)x = (24 \cdot 1)x = 0$.

4.5.4.3 C. Going deeper

Exercise. C1: Derivations

Let R be a commutative ring, and let $\delta: R[x] \rightarrow R[x]$ be a map such that, for all $f, g \in R[x]$ and every constant a ,

$$\delta(f + g) = \delta(f) + \delta(g), \quad \delta(a) = 0, \quad \delta(fg) = \delta(f)g + f\delta(g).$$

Prove that $\delta(f) = f' \delta(x)$ for every $f \in R[x]$. Deduce that $f \mapsto f'$ is the only such map with $\delta(x) = 1$.

Solution. Both sides are additive in f . δ is additive by hypothesis, and $f \mapsto f' \delta(x)$ is additive by [Theorem 4.5.1](#) (a) and (F9) of $R[x]$ ([Theorem 4.2.4](#) (a)).

Powers of x . We show $\delta(x^n) = (n \cdot 1)x^{n-1}\delta(x)$ for every $n \geq 1$, by induction. For $n = 1$, $(1 \cdot 1)x^0\delta(x) = \delta(x)$. If it holds for n , then, since $x^{n+1} = x^n x$,

$$\begin{aligned} \delta(x^{n+1}) &= \delta(x^n)x + x^n\delta(x) = (n \cdot 1)x^{n-1}\delta(x)x + x^n\delta(x) \\ &= (n \cdot 1 + 1)x^n\delta(x) = ((n + 1) \cdot 1)x^n\delta(x), \end{aligned}$$

by (F5), (F6), (F7), (F9) of $R[x]$, $x^{n-1}x = x^n$, and the recursion of multiples. Also $\delta(x^0) = \delta(1) = 0$, since 1 is a constant.

Monomials. Let $a \in R$ and $n \in \mathbb{N}$. By [Theorem 4.2.4](#) (c), ax^n is the product of the constant a and x^n , so

$$\delta(ax^n) = \delta(a)x^n + a\delta(x^n) = a\delta(x^n).$$

If $n = 0$, this is $0 = (ax^0)'\delta(x)$, by (M). If $n \geq 1$, it is $a(n \cdot 1)x^{n-1}\delta(x) = (n \cdot a)x^{n-1}\delta(x) = (ax^n)'\delta(x)$, by (F6), [Lemma 4.1.1](#) (d) and (M).

Conclusion. The two additive maps δ and $f \mapsto f' \delta(x)$ agree on every monomial, so they are equal by [Lemma 4.2.3](#) (one argument, $T = R[x]$).

Uniqueness. The formal derivative satisfies the three conditions, by [Theorem 4.5.1](#) (a) and (b), and $x' = 1$ ([Example \(Formal derivatives in five rings\)](#) (ii)). If δ satisfies them with $\delta(x) = 1$, then $\delta(f) = f' \cdot 1 = f'$ for every f .

Marking scheme. (10 points) - 2: both sides additive in f , with the additivity of $f \mapsto f' \delta(x)$ justified. - 2: $\delta(ax^n) = a\delta(x^n)$ from the product condition and $\delta(a) = 0$. - 3: $\delta(x^n) = (n \cdot 1)x^{n-1}\delta(x)$ by induction, with the case $n = 0$ handled separately. - 2: comparison with $(ax^n)'$, using $a(n \cdot 1) = n \cdot a$. - 1: conclusion by extending from monomials, and the uniqueness statement.

A solution that shows instead that $K := \{f \in R[x] : \delta(f) = f' \delta(x)\}$ is all of $R[x]$ is marked on the same points: closure of K under $+$ earns the first 2; the constants and x in K earn the next 2; closure of K under products, by [Theorem 4.5.1](#) (b), earns the 3; $K = R[x]$ earns the 2; the uniqueness statement earns the last 1.

Fatal errors (0 for the step): checking only on x^n , with no account of the coefficient a (that is where $\delta(a) = 0$ is needed); writing x^{-1} in the case $n = 0$.

Exercise. C2: Find the error

Here is an argument.

Claim. If $f \in \mathbb{Q}[x]$ and $f' = 0$, then f is constant.

Proof. (1) For $k \geq 1$, the coefficient of x^{k-1} in f' is $k \cdot f_k = (k \cdot 1)f_k$, by the definition of f' and the rules for multiples. (2) Since $f' = 0$, $(k \cdot 1)f_k = 0$ for every $k \geq 1$. (3) For $k \geq 1$ we have $k \neq 0$, so $k \cdot 1 \neq 0$; since $\mathbb{Q}$ has no zero divisors, $f_k = 0$ for every $k \geq 1$. (4) So $f = f_0 x^0$ is constant.

Is the claim true? Does the argument prove it? Name the line that fails and the reason, give a counterexample to the illegal step, and repair the argument.

Solution. **The claim is true:** it is Proposition 4.5.2 (a) for $F = \mathbb{Q}$. **The argument, as written, does not prove it.**

Lines (1), (2) and (4) are correct: (1) is Definition 4.5.1 with Lemma 4.1.1 (d), (2) reads $f' = 0$ coefficient by coefficient, and (4) compares coefficients (Definition 4.1.1). The second half of line (3) is correct too: $\mathbb{Q}$ is a field (Number Systems: From $\mathbb{N}$ to $\mathbb{R}$), Theorem 2.4.3), so it has no zero divisors (Theorem 2.1.1 (f)).

Line (3) fails, in its first half. “ $k \neq 0$ in $\mathbb{Z}$, so $k \cdot 1 \neq 0$ in the ring” is a general “fact” that is neither proved nor true, and the line uses nothing about $\mathbb{Q}$ to support it: entry (N9) of Logic and Proof, Example (What is not allowed). *Counterexample to the step:* in $\mathbb{F}_2$, $k = 2$ gives $2 \cdot 1 = 1 + 1 = 0$. The same four lines, read over $\mathbb{F}_2$, would “prove” that x^2 is constant, since $(x^2)' = 0$ (Example (Formal derivatives in five rings) (iii)). There is no counterexample to the conclusion, which is true.

Repair. What line (3) needs is that $\mathbb{Q}$ has characteristic 0 (Definition 2.5.1): for $k \geq 1$, $k \cdot 1$ is the rational number k (induction on k , since the embeddings $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$ preserve sums and 1, Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Proposition 2.2.3 and Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Proposition 2.4.4), and $k \neq 0$. With that sentence in place of “so”, the argument is the proof of Proposition 4.5.2 (a), whose hypothesis is exactly the characteristic.

Marking scheme. (10 points) - 2: says the claim is true and that the argument, as written, does not prove it. - 3: names the first half of line (3) and the reason: $k \neq 0$ does not give $k \cdot 1 \neq 0$ in a general ring, (N9), or an equivalent statement. - 2: a counterexample to the step, such as $2 \cdot 1 = 0$ in $\mathbb{F}_2$, or the “proof” that x^2 is constant over $\mathbb{F}_2$. - 3: repairs the step by proving $k \cdot 1 \neq 0$ in $\mathbb{Q}$ for every $k \geq 1$, that is, $\text{char } \mathbb{Q} = 0$. Two routes earn full marks: $k \cdot 1 = k$ via the embeddings, as in the solution; or $k \cdot 1 > 0$ by Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Proposition 2.3.5 (h), since $\mathbb{Q}$ is an ordered field (Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Theorem 2.5.1) with $1 \neq 0$. Fatal errors (0 for the step): a counterexample to the conclusion (there is none); blaming the no-zero-divisors half of line (3), which is correct.

4.5.4.4 R. Review

Exercise. R1

Let R be a commutative ring and $E: R[x] \rightarrow R[x]$, $E(f) := f + f'$. Show that E is an injective homomorphism of the group $(R[x], +)$ to itself. If $1 \neq 0$ in R , show that $E(fg) = E(f)E(g)$ fails for some f, g .

Solution. The tools are the notion of a group homomorphism and its kernel, Definition 3.3.1, written additively (Definition 3.1.1), and Proposition 3.3.3.

The group. By Theorem 4.2.4 (a), $R[x]$ satisfies (F1)–(F4), so $(R[x], +)$ is an abelian group, with identity the zero polynomial and inverse $-f$.

A homomorphism. By Theorem 4.5.1 (a) and (F1), (F2) of $R[x]$,

$$E(f + g) = (f + g) + (f' + g') = (f + f') + (g + g') = E(f) + E(g),$$

which is the homomorphism condition in additive notation.

The kernel is $\{0\}$. Let $f \neq 0$, with $n := \deg f$, so $f_n \neq 0$ and $f_{n+1} = 0$ (Definition 4.3.1). By Definition 4.5.1, $(f')_n = (n+1) \cdot f_{n+1} = (n+1) \cdot 0 = 0$, as in the well-definedness paragraph after it. Addition is coefficientwise (Definition 4.2.1), so $E(f)_n = f_n + 0 = f_n \neq 0$, and $E(f) \neq 0$. Also $0 = 0x^0$, so $0' = 0$ by (M), and $E(0) = 0 + 0 = 0$ by (F3) of $R[x]$. So $\ker E = \{0\}$, and E is injective by Proposition 3.3.3.

Not multiplicative. Take $f = g = x$. By Example (Formal derivatives in five rings) (ii), $x' = 1$, so $E(x) = x + 1$, and $E(x)E(x) = (x + 1)^2 = x^2 + 2x + 1$ by Number Systems: From $\setminus(\setminus n\mathbb{N}\setminus)$ to $\setminus(\setminus n\mathbb{R}\setminus)$, Proposition 2.3.2 (d). On the other side, (M) gives $(x^2)' = (2 \cdot 1)x = 2x$, so $E(x^2) = x^2 + 2x$. The two differ in position 0, where they have 1 and 0; since $1 \neq 0$, $E(x \cdot x) \neq E(x)E(x)$. So only the group structure of $R[x]$ is respected, and E is not a ring homomorphism.

Exercise. R2

For $f \in \mathbb{R}[x]$ and $z \in \mathbb{C}$, let $f(z)$ denote the value obtained by substituting z for x , where a real coefficient a is read as the complex number a . Show that $f(\bar{z}) = \overline{f(z)}$. Deduce that if $z_0 \in \mathbb{C}$ is a root of f that is not real, then $\bar{z}_0$ is a different root of f .

Solution. The tool is the uniqueness in Theorem 4.4.2 (b), applied to the field $\mathbb{C}$ of Chapter 1.

The setting. $\mathbb{C}$ is a field, and $\mathbb{R} \subseteq \mathbb{C}$ is closed under its operations (Theorem 1.2.4); so $\mathbb{C}$ and $\mathbb{R}$ are commutative rings and $\mathbb{R}[x]$ is defined. Let $\iota: \mathbb{R} \rightarrow \mathbb{C}$ be the map of Proposition 1.1.1; it preserves $+$, $\cdot$ and 1 , so it is a ring homomorphism (Definition 4.4.2). For $s \in \mathbb{C}$, let $\Phi_s: \mathbb{R}[x] \rightarrow \mathbb{C}$ be the homomorphism of Theorem 4.4.2 with $\varphi = \iota$ and this s ; then $f(z)$ means $\Phi_z(f)$.

Conjugation is a ring homomorphism. By Theorem 1.2.2 (a) and (b), $\overline{u + v} = \bar{u} + \bar{v}$ and $\overline{uv} = \bar{u}\bar{v}$, and $\bar{1} = 1$, since $1 = 1 + 0i$ and $\bar{1} = 1 - 0i = 1$ (Definition 1.2.1).

Two homomorphisms. Let $A(f) := \overline{\Phi_z(f)}$ and $B(f) := \Phi_{\bar{z}}(f)$. Then A is a composite of ring homomorphisms, hence one (Definition 4.4.2), and B is one by Theorem 4.4.2 (a). On a constant $a \in \mathbb{R}$, $A(a) = \overline{\iota(a)} = \iota(a)$, since $a = a + 0i$ has conjugate $a - 0i = a$; and $B(a) = \iota(a)$. On x , $A(x) = \bar{z} = B(x)$. By the uniqueness in Theorem 4.4.2 (b), $A = B$: $f(z) = f(\bar{z})$ for every $f \in \mathbb{R}[x]$.

Roots. If $f(z_0) = 0$, then $f(\bar{z}_0) = \bar{0} = 0$. If $z_0 = a + bi$ is not real, then $b \neq 0$, so $\bar{z}_0 = a - bi \neq a + bi$: equality would give $-b = b$ by comparing imaginary parts (Definition 1.1.1), so $2b = 0$, and $b = 0$ because $2 = 2 \cdot 1 \neq 0$ in the ordered field $\mathbb{R}$ (Number Systems: From

$\backslash(\backslash n N \backslash)$ to $\backslash(\backslash n R \backslash)$, Proposition 2.3.5 (h)) and $\mathbb{R}$ has no zero divisors (Number Systems: From $\backslash(\backslash n N \backslash)$ to $\backslash(\backslash n R \backslash)$, Proposition 2.3.5 (f)). So the non-real roots of a real polynomial come in pairs $z_0 \neq \bar{z}_0$.

The identity and the first deduction are the statement of Proposition 1.6.4, and the distinctness of $\bar{z}_0$ is the paragraph after it; Chapter 1 proved both by a direct computation with the sum $\sum_k a_k z^k$. Here the computation is replaced by the uniqueness of a substitution homomorphism. Polynomials with rational coefficients are a special case, since $\mathbb{Q}$ sits inside $\mathbb{R}$.

Exercise. R3

Show that $\mathbb{Z}[x]$ and $\mathbb{F}_2[x]$ are countable, and that the set of all sequences $\mathbb{N} \rightarrow \mathbb{F}_2$ is uncountable. Which clause of the definition of a polynomial does the difference rest on?

Solution. The tools are the countability results of *Sets, Relations, and Cardinality*, Chapter 3, read on $\mathbb{N} = \omega$ through Number Systems: From $\backslash(\backslash n N \backslash)$ to $\backslash(\backslash n R \backslash)$, Corollary 1.7.3.

$\mathbb{Z}$ and $\mathbb{F}_2$ are countable. $\mathbb{N} \times \mathbb{N}$ is countable (Sets, Relations, and Cardinality, Theorem 3.2.6). The class map $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim$ is surjective (Sets, Relations, and Cardinality, Definition 2.5.4), so $\mathbb{Z}$ is countable by Sets, Relations, and Cardinality, Proposition 3.2.4; this is the count of Number Systems: From $\backslash(\backslash n N \backslash)$ to $\backslash(\backslash n R \backslash)$, Exercise (R2). The map $\mathbb{N} \rightarrow \mathbb{F}_2$, $n \mapsto [n]$, is surjective, since every class is $[0]$ or $[1]$ (Number Systems: From $\backslash(\backslash n N \backslash)$ to $\backslash(\backslash n R \backslash)$, Theorem 3.4.3 (b)); so $\mathbb{F}_2$ is countable (Sets, Relations, and Cardinality, Definition 3.2.1).

Polynomials of bounded length. Let R be $\mathbb{Z}$ or $\mathbb{F}_2$, and for $N \in \mathbb{N}$ let $A_N := \{f \in R[x] : f_k = 0 \text{ for every } k > N\}$. By Exercise (B1: Polynomials of bounded length as tuples), A_N is in bijection with R^{N+1} . For $N \geq 1$, R^{N+1} is countable by Sets, Relations, and Cardinality, Corollary 3.2.7 (b); for $N = 0$ it is R itself. So each A_N is countable, as the image of a countable set under a bijection (Sets, Relations, and Cardinality, Proposition 3.2.4).

The union. Every polynomial lies in some A_N (Definition 4.1.1), so $R[x] = \bigcup_{N \in \mathbb{N}} A_N$, a countable union of countable sets, which is countable by Sets, Relations, and Cardinality, Theorem 3.3.3 (a). (That theorem uses the axiom of countable choice to pick a list of each A_N .)

All sequences. The set $\{0, 1\}^{\mathbb{N}}$ is uncountable (Sets, Relations, and Cardinality, Theorem 3.4.1). The map $\{0, 1\}^{\mathbb{N}} \rightarrow \mathbb{F}_2^{\mathbb{N}}$, $0 \mapsto [0]$, $1 \mapsto [1]$, is a bijection; composing a sequence with its inverse turns every sequence $\mathbb{N} \rightarrow \mathbb{F}_2$ into a sequence of zeros and ones, and this is a surjection onto $\{0, 1\}^{\mathbb{N}}$. If the sequences $\mathbb{N} \rightarrow \mathbb{F}_2$ were countable, so would be $\{0, 1\}^{\mathbb{N}}$, by Sets, Relations, and Cardinality, Proposition 3.2.4. So they are uncountable.

The clause. The difference is the clause **only finitely many** of Definition 4.1.1: a polynomial is a sequence that is 0 from some point on, so it is determined by a finite tuple, and there are countably many finite tuples. A general sequence carries infinitely many independent choices. The same count, with the polynomials written as integer tuples, is the proof of Sets, Relations, and Cardinality, Theorem 3.3.4.

Exercise. R4

Find all $f \in \mathbb{Q}[x]$ with $f \circ x^2 = f^2$.

Solution. Answer: $f = 0$, and $f = x^n$ for $n \in \mathbb{N}$.

They work. $0 \circ x^2 = 0 = 0^2$, and $x^n \circ x^2 = (x^2)^n = x^{2n} = (x^n)^2$ by Definition 4.4.3 and the power laws (Lemma 4.1.1 (c)).

The tool is **look at the top coefficient**, then at the next non-zero one. Let $f \neq 0$ satisfy $f \circ x^2 = f^2$, with $n := \deg f$ and $a := \text{lc}(f)$. As in the proof of Proposition 4.5.2 (b), $f \circ x^2 = \sum_{j=0}^n f_j x^{2j}$, whose coefficient at $2j$ is f_j and whose coefficients at odd positions are 0. So its coefficient at $2n$ is a , and none beyond $2n$ is non-zero. By Theorem 4.3.2 (b), since $\mathbb{Q}$ has no zero divisors, f^2 has degree $2n$ and leading coefficient a^2 . Hence $a^2 = a$, so $a(a-1) = 0$, and $a = 1$ since $a \neq 0$ (Theorem 2.1.1 (f)).

Now let $r := f - x^n$, so $f = x^n + r$ and $\deg r < n$ (Theorem 4.3.1, since the coefficients at n cancel). Suppose $r \neq 0$, and let $k := \deg r < n$. By Number Systems: From $\mathbb{N}$ to $\mathbb{R}$, Proposition 2.3.2 (d) in the commutative ring $\mathbb{Q}[x]$ (Theorem 4.2.4 (a)),

$$f^2 = x^{2n} + 2x^n r + r^2.$$

Its coefficient at $n+k$ is $2r_k$: x^{2n} contributes 0, since $n+k < 2n$; $2x^n r$ contributes $2r_k$, since multiplying by x^n shifts by n (Lemma 4.2.2 (a)); and r^2 contributes 0, since $\deg r^2 \leq 2k < n+k$ (Theorem 4.3.2 (a)). On the other side, the coefficient of $f \circ x^2$ at $n+k$ is 0: if $n+k$ is odd, it is an odd position; if $n+k = 2j$, then $k < j < n$, so $f_j = r_j = 0$, because $j < n$ and $j > \deg r$. Hence $2r_k = 0$ in $\mathbb{Q}$, and $r_k = 0$ since $2 \neq 0$: a contradiction with $r_k = \text{lc}(r) \neq 0$. So $r = 0$ and $f = x^n$.

Why the look-alikes fail. Evaluating at points gives $f(c^2) = f(c)^2$ for every $c \in \mathbb{Q}$, a statement about polynomial functions; turning it back into a statement about coefficients needs the fact that a polynomial over an infinite field is determined by its values, which this chapter does not prove (Proposition 4.4.4 (d) shows that over a finite ring with $1 \neq 0$ it is false). Differentiating with the chain rule of Exercise B1 gives $2x(f' \circ x^2) = 2ff'$, an equation of the same kind and no easier.

Over $\mathbb{F}_2$ every f works. There $(u+v)^2 = u^2 + v^2$, since $2uv = 0$, and $c^2 = c$ for $c \in \mathbb{F}_2$; so $f^2 = \sum_k f_k^2 x^{2k} = \sum_k f_k x^{2k} = f \circ x^2$. The step that fails is “ $2r_k = 0$ implies $r_k = 0$ ”: in $\mathbb{F}_2$, $2 = 0$.