University of Minnesota, School of Mathematics

Math 4281: Introduction to Modern Algebra, Spring 2019: Midterm 3

Darij Grinberg

February 25, 2021

1 Exercise 1: Nonunital rings and local unities

1.1 Problem

A *nonunital ring* is defined in the same way as we defined a ring, except that we don't require it to be endowed with an element 1 (and, correspondingly, we omit the "Neutrality of one" axiom). This does not mean that a nonunital ring must not contain an element 1 that would satisfy the "Neutrality of one" axiom; it simply means that such an element is not required (and not considered part of the ring structure). So, formally speaking, a nonunital ring is a 4-tuple $(\mathbb{K}, +, \cdot, 0)$ (while a ring in the usual sense is a 5-tuple $(\mathbb{K}, +, \cdot, 0, 1)$) that satisfies all the ring axioms except for "Neutrality of one".

Thus, every ring becomes a nonunital ring if we forget its unity (i.e., if $(\mathbb{K}, +, \cdot, 0, 1)$ is a ring, then $(\mathbb{K}, +, \cdot, 0)$ is a nonunital ring). But there are other examples as well: For instance, if $n \in \mathbb{Z}$ is arbitrary, then $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\} = \{\text{all multiples of } n\}$ is a nonunital ring (when endowed with the usual $+$, $\cdot$ and 0).

An element z of a nonunital ring $\mathbb{K}$ is said to be a *unity* of $\mathbb{K}$ if every $a \in \mathbb{K}$ satisfies $az = za = a$. In other words, an element z of a nonunital ring $\mathbb{K}$ is said to be a *unity* of $\mathbb{K}$ if equipping $\mathbb{K}$ with the unity z results in a ring (in the usual sense of this word).

Prove the following:

(a) If $n \in \mathbb{Z}$, then the nonunital ring $n\mathbb{Z}$ has a unity if and only if $n \in \{1, 0, -1\}$.

(b) Any nonunital ring has **at most one** unity.

Now, let $\mathbb{K}$ be a nonunital ring. As usual, we write $+$ and $\cdot$ for its two operations, and 0 for its zero.

Let $z \in \mathbb{K}$. Define a subset U_z of $\mathbb{K}$ by

$$U_z = \{r \in \mathbb{K} \mid rz = zr = r\}.$$

(c) Prove that $0 \in U_z$, and that every $a, b \in U_z$ satisfy $a + b \in U_z$ and $ab \in U_z$.

Thus, we can turn U_z into a nonunital ring by endowing U_z with the binary operations $+$ and $\cdot$ (inherited from $\mathbb{K}$) and the element 0. Consider this nonunital ring U_z.

(d) Assume that $z^2 = z$. Prove that z is a unity of the nonunital ring U_z.

[**Hint:** In **(b)**, what would the product of two unities be?]

1.2 Solution sketch

(a) Let $n \in \mathbb{Z}$. We must prove that the nonunital ring $n\mathbb{Z}$ has a unity if and only if $n \in \{1, 0, -1\}$.

This is an "if and only if" statement. We are going to prove its "$\Longrightarrow$" and "$\Longleftarrow$" directions separately:

$\Longrightarrow$: Assume that the nonunital ring $n\mathbb{Z}$ has a unity. We must prove that $n \in \{1, 0, -1\}$.

Assume the contrary. Thus, $n \notin \{1, 0, -1\}$. Hence, $|n| > 1$ (since n is an integer) and $n \neq 0$.

Clearly, the integer n is a multiple of n (since $n = n \cdot 1$). Thus, $n \in n\mathbb{Z}$.

We have assumed that the the nonunital ring $n\mathbb{Z}$ has a unity. Let us denote this unity by u. Then, every $a \in n\mathbb{Z}$ satisfies $au = ua = a$ (by the definition of a unity). Applying this to $a = n$, we obtain $nu = un = n$ (since $n \in n\mathbb{Z}$). We can divide both sides of the equality $nu = n$ by n (since $n \neq 0$), and thus obtain $u = 1$. Thus, $1 = u \in n\mathbb{Z}$ (since u is a unity of $n\mathbb{Z}$); in other words, 1 is a multiple of n (by the definition of $n\mathbb{Z}$). In other words, $n \mid 1$. Thus, Proposition 2.2.3 **(b)** in the class notes (applied to $a = n$ and $b = 1$) yields $|n| \leq |1|$ (since $1 \neq 0$). Thus, $|n| \leq 1 = 1$. This contradicts $|n| > 1$. This contradiction shows that our assumption was false. Thus, the "$\Longrightarrow$" direction of part **(a)** of this exercise is proven.

$\Longleftarrow$: Assume that $n \in \{1, 0, -1\}$. We must prove that the nonunital ring $n\mathbb{Z}$ has a unity.

We have $n \in \{1, 0, -1\}$. Thus, we are in one of the following three cases:

Case 1: We have $n = 1$.

Case 2: We have $n = 0$.

Case 3: We have $n = -1$.

Let us first consider Case 1. In this case, we have $n = 1$. Hence,

$$n\mathbb{Z} = 1\mathbb{Z} = \{\text{all multiples of } 1\} = \{\text{all integers}\} \qquad \text{(since all integers are multiples of 1)}$$
$$= \mathbb{Z}.$$

But the nonunital ring $\mathbb{Z}$ has a unity (namely, 1), since $\mathbb{Z}$ equipped with the unity 1 is a ring. In other words, the nonunital ring $n\mathbb{Z}$ has a unity (since $n\mathbb{Z} = \mathbb{Z}$). Thus, we have proven in Case 1 that the nonunital ring $n\mathbb{Z}$ has a unity.

Let us next consider Case 2. In this case, we have $n = 0$. Hence,

$$n\mathbb{Z} = 0\mathbb{Z} = \{\text{all multiples of } 0\} = \{0\} \qquad \text{(since the only multiple of 0 is 0 itself)}.$$

But the nonunital ring $\{0\}$ has a unity (namely, 0), since $\{0\}$ equipped with the unity 0 is a ring (viz., the zero ring). In other words, the nonunital ring $n\mathbb{Z}$ has a unity (since $n\mathbb{Z} = \{0\}$). Thus, we have proven in Case 2 that the nonunital ring $n\mathbb{Z}$ has a unity.

Let us finally consider Case 3. In this case, we have $n = -1$. Hence,

$$\begin{aligned} n\mathbb{Z} &= (-1)\,\mathbb{Z} = \{\text{all multiples of } -1\} = \{\text{all integers}\} \\ &\qquad \left(\begin{array}{c} \text{since all integers are multiples of } -1 \\ (\text{because each integer } a \text{ satisfies } a = (-1)\,(-a)\,) \end{array} \right) \\ &= \mathbb{Z}. \end{aligned}$$

But the nonunital ring $\mathbb{Z}$ has a unity (namely, 1), since $\mathbb{Z}$ equipped with the unity 1 is a ring. In other words, the nonunital ring $n\mathbb{Z}$ has a unity (since $n\mathbb{Z} = \mathbb{Z}$). Thus, we have proven in Case 3 that the nonunital ring $n\mathbb{Z}$ has a unity.

Now, in each of the three Cases 1, 2 and 3, we have shown that the nonunital ring $n\mathbb{Z}$ has a unity. Hence, this always holds. Thus, the "$\Longleftarrow$" direction of part **(a)** of this exercise is solved.

(b) Let $\mathbb{K}$ be a nonunital ring. We must prove that $\mathbb{K}$ has **at most one** unity. In other words, we must prove that any two unities of $\mathbb{K}$ are equal. In other words, we must prove that if u and v are two unities of $\mathbb{K}$, then $u = v$.

So let u and v be two unities of $\mathbb{K}$. We must prove that $u = v$.

We have assumed that u is a unity of $\mathbb{K}$. In other words, u is an element of $\mathbb{K}$ such that every $a \in \mathbb{K}$ satisfies $au = ua = a$ (by the definition of a unity).

We have assumed that v is a unity of $\mathbb{K}$. In other words, v is an element of $\mathbb{K}$ such that every $a \in \mathbb{K}$ satisfies $av = va = a$ (by the definition of a unity).

We know that every $a \in \mathbb{K}$ satisfies $au = ua = a$. Applying this to $a = v$, we obtain $vu = uv = v$.

We know that every $a \in \mathbb{K}$ satisfies $av = va = a$. Applying this to $a = u$, we obtain $uv = vu = u$.

Comparing $uv = u$ with $uv = v$, we obtain $u = v$. This completes our proof. Thus, part **(b)** of the exercise is solved.

(c) We have $0z = 0$ and $z0 = 0$, thus $0z = z0 = 0$.

The element 0 of $\mathbb{K}$ is an $r \in \mathbb{K}$ satisfying $rz = zr = r$ (since $0 \in \mathbb{K}$ and $0z = z0 = 0$). In other words, $0 \in \{r \in \mathbb{K} \mid rz = zr = r\}$. This rewrites as $0 \in U_z$ (since $U_z = \{r \in \mathbb{K} \mid rz = zr = r\}$).

Hence, it remains to prove that every $a, b \in U_z$ satisfy $a + b \in U_z$ and $ab \in U_z$.

So let $a, b \in U_z$. We must show that $a + b \in U_z$ and $ab \in U_z$.

We have $a \in U_z = \{r \in \mathbb{K} \mid rz = zr = r\}$. In other words, a is an $r \in \mathbb{K}$ satisfying $rz = zr = r$. In other words, a is an element of $\mathbb{K}$ and satisfies $az = za = a$. The same argument (applied to b instead of a) yields that b is an element of $\mathbb{K}$ and satisfies $bz = zb = b$. Now, using the distributivity axiom, we find

$$(a+b)\,z = \underbrace{az}_{=za} + \underbrace{bz}_{=zb} = za + zb = z\,(a+b) = \underbrace{za}_{=a} + \underbrace{zb}_{=b} = a + b.$$

Thus, $a + b$ is an element of $\mathbb{K}$ and satisfies $(a+b)\,z = z\,(a+b) = a + b$. In other words, $a + b$ is an $r \in \mathbb{K}$ satisfying $rz = zr = r$. In other words, $a + b \in \{r \in \mathbb{K} \mid rz = zr = r\}$. This rewrites as $a + b \in U_z$ (since $U_z = \{r \in \mathbb{K} \mid rz = zr = r\}$).

Also, using the associativity axiom, we find

$$(ab)\,z = a\,\underbrace{(bz)}_{=zb} = a\,(zb) = \underbrace{(az)}_{=za}\,b = (za)\,b = z\,(ab) = \underbrace{(za)}_{=a}\,b = ab.$$

Thus, ab is an element of $\mathbb{K}$ and satisfies $(ab)\,z = z\,(ab) = ab$. In other words, ab is an $r \in \mathbb{K}$ satisfying $rz = zr = r$. In other words, $ab \in \{r \in \mathbb{K} \mid rz = zr = r\}$. This rewrites as $ab \in U_z$ (since $U_z = \{r \in \mathbb{K} \mid rz = zr = r\}$).

So we have shown that $a + b \in U_z$ and $ab \in U_z$. This completes the solution to part **(c)** of the exercise.

(d) The element z of $\mathbb{K}$ is an $r \in \mathbb{K}$ satisfying $rz = zr = r$ (since $zz = z^2 = z$). In other words, $z \in \{r \in \mathbb{K} \mid rz = zr = r\}$. This rewrites as $z \in U_z$ (since $U_z = \{r \in \mathbb{K} \mid rz = zr = r\}$). Thus, z is an element of U_z. This element has the property that every $a \in U_z$ satisfies $az = za = a$ [1]. In other words, z is a unity of the nonunital ring U_z (by the definition of a unity). This solves part **(d)** of the exercise.

1.3 Remark

In the solution to part **(d)**, we have used the assumption $z^2 = z$ only in order to prove that z is an element of U_z. This is easy to overlook but nevertheless important: A unity of a nonunital ring must, first of all, be an element of this ring!

---

2 Exercise 2: Rings from nonunital rings

2.1 Problem

Let $\mathbb{K}$ be a nonunital ring. (See Exercise 1 for the definition of this notion.) Let $\mathbb{L}$ be the Cartesian product $\mathbb{Z} \times \mathbb{K}$ (so far, just a set). Define a binary operation $+$ on $\mathbb{L}$ by

$$(n, a) + (m, b) = (n + m, a + b) \qquad \text{for all } (n, a), (m, b) \in \mathbb{L}.$$

(This is an entrywise addition.) Define a binary operation $\cdot$ on $\mathbb{L}$ by

$$(n, a)\,(m, b) = (nm, nb + ma + ab) \qquad \text{for all } (n, a), (m, b) \in \mathbb{L}.$$

(Here, nb and ma are defined in the usual way: If $n \in \mathbb{Z}$ and $a \in \mathbb{K}$, then $na \in \mathbb{K}$ is defined by

$$na = \begin{cases} \underbrace{a + a + \cdots + a}_{n \text{ times}}, & \text{if } n \geq 0; \\ -\left(\underbrace{a + a + \cdots + a}_{-n \text{ times}}\right), & \text{if } n < 0 \end{cases}.$$

This does not require $\mathbb{K}$ to have a unity.)

Prove that $\mathbb{L}$, endowed with these two operations $+$ and $\cdot$ and the zero $(0, 0)$ and the unity $(1, 0)$, is a ring (in the usual sense of this word).

[**Hint:** You can use rules like $n\,(a + b) = na + nb$ and $(n + m)\,a = na + ma$ and $(nm)\,a = n\,(ma)$ (for $n, m \in \mathbb{Z}$ and $a, b \in \mathbb{K}$) without proof; they can be proven just as for usual rings. You can also use the fact that finite sums of elements of $\mathbb{K}$ are well-defined and

[1] *Proof.* Let $a \in U_z$. We must prove that $az = za = a$.

We have $a \in U_z = \{r \in \mathbb{K} \mid rz = zr = r\}$. In other words, a is an $r \in \mathbb{K}$ satisfying $rz = zr = r$. In other words, a is an element of $\mathbb{K}$ and satisfies $az = za = a$. Qed.

behave as we would expect them to (we already tacitly used that in writing "$\underbrace{a+a+\cdots+a}_{n \text{ times}}$" without parentheses).

You don't need to check the "additive" axioms (associativity of addition, commutativity of addition, neutrality of zero, and existence of additive inverses); as far as addition and zero are concerned, $\mathbb{L}$ is just a Cartesian product.]

2.2 Remark

This exercise gives a way to "embed" any nonunital ring $\mathbb{K}$ into a ring $\mathbb{L}$. This helps proving properties of nonunital rings, assuming that you can prove them for rings. The ring $\mathbb{L}$ is called the *Dorroh extension* of $\mathbb{K}$ (or, more precisely, the $\mathbb{Z}$*-Dorroh extension* of $\mathbb{K}$, since there are other possibilities as well).

There is also a much simpler notion of a Cartesian product of two nonunital rings (in which both addition and multiplication are defined entrywise). This lets us define a nonunital ring $\mathbb{Z} \times \mathbb{K}$. But this is **not** the ring $\mathbb{L}$; it does not generally have a unity.

2.3 Solution sketch

We first state a few basic rules for computing inside $\mathbb{K}$:

Proposition 2.1. *We have*

$$(n+m)\,a = na + ma \qquad \text{for all } a \in \mathbb{K} \text{ and } n, m \in \mathbb{Z}; \tag{1}$$

$$n\,(a+b) = na + nb \qquad \text{for all } a, b \in \mathbb{K} \text{ and } n \in \mathbb{Z}; \tag{2}$$

$$-\,(na) = (-n)\,a = n\,(-a) \qquad \text{for all } a \in \mathbb{K} \text{ and } n \in \mathbb{Z}; \tag{3}$$

$$(nm)\,a = n\,(ma) \qquad \text{for all } a \in \mathbb{K} \text{ and } n, m \in \mathbb{Z}; \tag{4}$$

$$n\,(ab) = (na)\,b = a\,(nb) \qquad \text{for all } a, b \in \mathbb{K} \text{ and } n \in \mathbb{Z}; \tag{5}$$

$$n0_{\mathbb{K}} = 0_{\mathbb{K}} \qquad \text{for all } n \in \mathbb{Z}; \tag{6}$$

$$1a = a \qquad \text{for all } a \in \mathbb{K} \tag{7}$$

(*here, the "1" means the integer* 1);

$$0a = 0_{\mathbb{K}} \qquad \text{for all } a \in \mathbb{K} \tag{8}$$

(*here, the "0" on the left hand side means the integer* 0);

$$(-1)\,a = -a \qquad \text{for all } a \in \mathbb{K}; \tag{9}$$

(*here, the "* -1*" means the integer* -1).

In particular:

- *The equality (3) shows that the expression "*$-na$*" (with* $a \in \mathbb{K}$ *and* $n \in \mathbb{Z}$*) is unambiguous (since its two possible interpretations, namely* $-(na)$ *and* $(-n)\,a$*, yield equal results).*
- *The equality (4) shows that the expression "*nma*" (with* $a \in \mathbb{K}$ *and* $n, m \in \mathbb{Z}$*) is unambiguous.*
- *The equality (5) shows that the expression "*nab*" (with* $a, b \in \mathbb{K}$ *and* $n \in \mathbb{Z}$*) is unambiguous.*

Proof of Proposition 2.1. If $\mathbb{K}$ is a ring (with a unity), then Proposition 2.1 is precisely Proposition 5.4.9 in the class notes. But the proof of Proposition 5.4.9 in the class notes

nowhere uses the unity of $\mathbb{K}$. Thus, this proof still applies to the case where $\mathbb{K}$ is a nonunital ring. This proves Proposition 2.1. $\square$

Now, let us return to the solution of the exercise. We have to prove that the set $\mathbb{L}$ (endowed with the two operations $+$ and $\cdot$ defined in the exercise, and with the zero $(0,0)$ and the unity $(1,0)$) is a ring. By our definition of a ring, this means that we have to verify the ring axioms. Here is a list of these axioms (specialized to the case of $\mathbb{L}$):

- **Commutativity of addition:** We have $a+b=b+a$ for all $a,b \in \mathbb{L}$.
- **Associativity of addition:** We have $a+(b+c)=(a+b)+c$ for all $a,b,c \in \mathbb{L}$.
- **Neutrality of zero:** We have $a+(0,0)=(0,0)+a=a$ for all $a \in \mathbb{L}$.
- **Existence of additive inverses:** For any $a \in \mathbb{L}$, there exists an element $a' \in \mathbb{L}$ such that $a+a'=a'+a=(0,0)$.
- **Associativity of multiplication:** We have $a\,(bc)=(ab)\,c$ for all $a,b,c \in \mathbb{L}$.
- **Neutrality of one:** We have $a\,(1,0)=(1,0)\,a=a$ for all $a \in \mathbb{L}$.
- **Annihilation:** We have $a\,(0,0)=(0,0)\,a=(0,0)$ for all $a \in \mathbb{L}$.
- **Distributivity:** We have

$$a\,(b+c)=ab+ac \qquad \text{and} \qquad (a+b)\,c=ac+bc$$

for all $a,b,c \in \mathbb{L}$.

Here, we are already using the standard notations (such as the abbreviation "ab" for $a \cdot b$, and the PEMDAS conventions), even before showing that $\mathbb{L}$ is indeed a ring.

So it remains to prove that the eight ring axioms listed above are indeed satisfied. Let us do this now:

[*Proof of the "Commutativity of addition" axiom:* Let $a,b \in \mathbb{L}$. We must prove that $a+b=b+a$.

We have $a \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write a in the form $a=(n,x)$ for some $n \in \mathbb{Z}$ and $x \in \mathbb{K}$. Consider these n and x.

We have $b \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write b in the form $b=(m,y)$ for some $m \in \mathbb{Z}$ and $y \in \mathbb{K}$. Consider these m and y.

We have the two equalities

$$\begin{aligned} \underbrace{a}_{=(n,x)} + \underbrace{b}_{=(m,y)} &= (n,x)+(m,y) \\ &= (n+m, x+y) \qquad \text{(by the definition of the addition on } \mathbb{L}) \end{aligned}$$

and

$$\begin{aligned} \underbrace{b}_{=(m,y)} + \underbrace{a}_{=(n,x)} &= (m,y)+(n,x) \\ &= (m+n, y+x) \qquad \text{(by the definition of the addition on } \mathbb{L})\,. \end{aligned}$$

Thus, in order to prove that $a+b=b+a$, it suffices to prove that $(n+m, x+y) = (m+n, y+x)$. But this is easy: The commutativity of addition of integers yields $n+m=$

$m+n$, whereas the commutativity of addition in $\mathbb{K}$ (which holds because $\mathbb{K}$ is a nonunital ring) yields $x+y=y+x$. Thus,

$$\left(\underbrace{n+m}_{=m+n}, \underbrace{x+y}_{=y+x}\right) = (m+n, y+x)\,.$$

In view of $a+b=(n+m, x+y)$ and $b+a=(m+n, y+x)$, this rewrites as $a+b=b+a$. Thus, the "Commutativity of addition" axiom is proven for $\mathbb{L}$.]

[*Proof of the "Associativity of addition" axiom:* This proof is analogous to the proof of the "Commutativity of addition" axiom, and thus is left to the reader.]

[*Proof of the "Neutrality of zero" axiom:* This proof is analogous to the proof of the "Commutativity of addition" axiom, and thus is left to the reader.]

[*Proof of the "Existence of additive inverses" axiom:* Let $a \in \mathbb{L}$. We must prove that there exists an $a' \in \mathbb{L}$ such that $a+a'=a'+a=(0,0)$.

We have $a \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write a in the form $a=(n,x)$ for some $n \in \mathbb{Z}$ and $x \in \mathbb{K}$. Consider these n and x.

Now, a straightforward computation (using $a=(n,x)$) shows that the element $(-n,-x)$ of $\mathbb{L}$ satisfies $a+(-n,-x)=(-n,-x)+a=(0,0)$. Hence, there exists an $a' \in \mathbb{L}$ such that $a+a'=a'+a=(0,0)$ (namely, $a'=(-n,-x)$). Hence, the "Existence of additive inverses" axiom is proven for $\mathbb{L}$.]

[*Proof of the "Associativity of multiplication" axiom:* Let $a, b, c \in \mathbb{L}$. We must prove that $a\,(bc) = (ab)\,c$.

We have $a \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write a in the form $a=(n,x)$ for some $n \in \mathbb{Z}$ and $x \in \mathbb{K}$. Consider these n and x.

We have $b \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write b in the form $b=(m,y)$ for some $m \in \mathbb{Z}$ and $y \in \mathbb{K}$. Consider these m and y.

We have $c \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write c in the form $c=(p,z)$ for some $p \in \mathbb{Z}$ and $z \in \mathbb{K}$. Consider these p and z.

Let us first notice that every $r \in \mathbb{Z}$ and $u, v, w \in \mathbb{K}$ satisfy

$$r\,(u+v+w) = ru+rv+rw. \tag{10}$$

(Indeed, this follows from

$$r\underbrace{(u+v+w)}_{=(u+v)+w} = r\,((u+v)+w) = \underbrace{r\,(u+v)}_{\substack{=ru+rv \\ \text{(by (2))}}} + rw \qquad \text{(by (2))}$$
$$= ru+rv+rw.$$

)

We have

$$\underbrace{a}_{=(n,x)}\,\underbrace{b}_{=(m,y)} = (n,x)\,(m,y) = (nm, ny+mx+xy)$$

(by the definition of the multiplication on $\mathbb{L}$) and thus

$$\underbrace{(ab)}_{=(nm,ny+mx+xy)}\,\underbrace{c}_{=(p,z)}$$
$$= (nm, ny+mx+xy)\,(p,z)$$
$$= ((nm)\,p, (nm)\,z + p\,(ny+mx+xy) + (ny+mx+xy)\,z) \tag{11}$$

(by the definition of the multiplication on $\mathbb{L}$). Also, we have

$$\underbrace{b}_{=(m,y)}\ \underbrace{c}_{=(p,z)} = (m,y)\,(p,z) = (mp, mz+py+yz)$$

(by the definition of the multiplication on $\mathbb{L}$) and thus

$$\begin{aligned}
&\underbrace{a}_{=(n,x)}\ \underbrace{(bc)}_{=(mp,mz+py+yz)} \\
&= (n,x)\,(mp, mz+py+yz) \\
&= (n\,(mp)\,, n\,(mz+py+yz) + (mp)\,x + x\,(mz+py+yz)) \qquad (12)
\end{aligned}$$

(by the definition of the multiplication on $\mathbb{L}$).

The equalities (11) and (12) are explicit formulas for $(ab)\,c$ and $a\,(bc)$ in terms of n, x, m, y, p, z. Thus, in order to prove that $a\,(bc) = (ab)\,c$, it suffices to prove that

$$\begin{aligned}
&((nm)\,p, (nm)\,z + p\,(ny+mx+xy) + (ny+mx+xy)\,z) \\
&= (n\,(mp)\,, n\,(mz+py+yz) + (mp)\,x + x\,(mz+py+yz))\,.
\end{aligned}$$

But this is easy: The associativity of multiplication of integers yields $(nm)\,p = n\,(mp)$. Meanwhile, comparing the equalities[2]

$$\begin{aligned}
&\underbrace{(nm)\,z}_{=nmz} + \underbrace{p\,(ny+mx+xy)}_{\substack{=pny+pmx+pxy \\ \text{(by (10))}}} + \underbrace{(ny+mx+xy)\,z}_{\substack{=nyz+mxz+xyz \\ \text{(by distributivity in } \mathbb{K})}} \\
&= nmz + pny + pmx + pxy + nyz + mxz + xyz
\end{aligned}$$

and

$$\begin{aligned}
&\underbrace{n\,(mz+py+yz)}_{\substack{=nmz+npy+nyz \\ \text{(by (10))}}} + \underbrace{(mp)\,x}_{=mpx} + \underbrace{x\,(mz+py+yz)}_{\substack{=x(mz)+x(py)+xyz \\ \text{(by distributivity in } \mathbb{K})}} \\
&= nmz + \underbrace{np}_{\substack{=pn \\ \text{(by the commutativity} \\ \text{of multiplication in } \mathbb{Z})}} y + nyz + \underbrace{mp}_{\substack{=pm \\ \text{(by the commutativity} \\ \text{of multiplication in } \mathbb{Z})}} x + \underbrace{x\,(mz)}_{\substack{=mxz \\ \text{(by (5))}}} + \underbrace{x\,(py)}_{\substack{=pxy \\ \text{(by (5))}}} + xyz \\
&= nmz + pny + nyz + pmx + mxz + pxy + xyz \\
&= nmz + pny + pmx + pxy + nyz + mxz + xyz,
\end{aligned}$$

we obtain

$$\begin{aligned}
&(nm)\,z + p\,(ny+mx+xy) + (ny+mx+xy)\,z \\
&= n\,(mz+py+yz) + (mp)\,x + x\,(mz+py+yz)\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\left(\underbrace{(nm)\,p}_{=n(mp)}, \underbrace{(nm)\,z + p\,(ny+mx+xy) + (ny+mx+xy)\,z}_{=n(mz+py+yz)+(mp)x+x(mz+py+yz)}\right) \\
&= (n\,(mp)\,, n\,(mz+py+yz) + (mp)\,x + x\,(mz+py+yz))\,.
\end{aligned}$$

[2]See Proposition 2.1 for the reason why we are able to write expressions like "nmz" and "nyz" without parentheses. The expression "xyz" can be written without parentheses because of the associativity of multiplication in $\mathbb{K}$.

In view of (11) and (12), this rewrites as $a\,(bc) = (ab)\,c$. Thus, the "Associativity of multiplication" axiom is proven for $\mathbb{L}$.]

[*Proof of the "Neutrality of one" axiom:* Let $a \in \mathbb{L}$. We must prove that $a\,(1,0) = (1,0)\,a = a$. Note that the symbol "0" will always stand for "$0_{\mathbb{K}}$" in this proof (rather than for the integer 0).

We have $a \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write a in the form $a = (n, x)$ for some $n \in \mathbb{Z}$ and $x \in \mathbb{K}$. Consider these n and x.

We have

$$\underbrace{a}_{=(n,x)}\,(1,0) = (n,x)\,(1,0) = (n1, n0 + 1x + x0)$$

(by the definition of the multiplication on $\mathbb{L}$). In view of $n1 = n$ and

$$\underbrace{n0}_{\substack{=0\\(\text{by (6)})}} + \underbrace{1x}_{\substack{=x\\(\text{by (7)})}} + \underbrace{x0}_{=0} = 0 + x + 0 = x,$$

this rewrites as

$$a\,(1,0) = (n,x)\,.$$

We also have

$$(1,0)\,\underbrace{a}_{=(n,x)} = (1,0)\,(n,x) = (1n, 1x + n0 + 0x)$$

(by the definition of the multiplication on $\mathbb{L}$). In view of $1n = n$ and

$$\underbrace{1x}_{\substack{=x\\(\text{by (7)})}} + \underbrace{n0}_{\substack{=0\\(\text{by (6)})}} + \underbrace{0x}_{=0} = x + 0 + 0 = x,$$

this rewrites as

$$(1,0)\,a = (n,x)\,.$$

Now, comparing the three equalities $a\,(1,0) = (n,x)$ and $(1,0)\,a = (n,x)$ and $a = (n,x)$, we obtain $a\,(1,0) = (1,0)\,a = a$. Thus, the "Neutrality of one" axiom is proven for $\mathbb{L}$.]

[*Proof of the "Annihilation" axiom:* Let $a \in \mathbb{L}$. We must prove that $a\,(0,0) = (0,0)\,a = (0,0)$. In order to avoid confusion, we shall write "$0_{\mathbb{Z}}$" for the integer 0 and write "$0_{\mathbb{K}}$" for the zero of the nonunital ring $\mathbb{K}$. Thus, the element $(0,0)$ of $\mathbb{L}$ rewrites as $(0_{\mathbb{Z}}, 0_{\mathbb{K}})$. Hence, we must prove $a\,(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (0_{\mathbb{Z}}, 0_{\mathbb{K}})\,a = (0_{\mathbb{Z}}, 0_{\mathbb{K}})$.

We have $a \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write a in the form $a = (n, x)$ for some $n \in \mathbb{Z}$ and $x \in \mathbb{K}$. Consider these n and x.

We have

$$\underbrace{a}_{=(n,x)}\,(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (n,x)\,(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (n0_{\mathbb{Z}}, n0_{\mathbb{K}} + 0_{\mathbb{Z}}x + x0_{\mathbb{K}})$$

(by the definition of the multiplication on $\mathbb{L}$). In view of $n0_{\mathbb{Z}} = 0_{\mathbb{Z}}$ and

$$\underbrace{n0_{\mathbb{K}}}_{\substack{=0_{\mathbb{K}}\\(\text{by (6)})}} + \underbrace{0_{\mathbb{Z}}x}_{\substack{=0_{\mathbb{K}}\\(\text{by (8)})}} + \underbrace{x0_{\mathbb{K}}}_{=0_{\mathbb{K}}} = 0_{\mathbb{K}} + 0_{\mathbb{K}} + 0_{\mathbb{K}} = 0_{\mathbb{K}},$$

this rewrites as

$$a\,(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (0_{\mathbb{Z}}, 0_{\mathbb{K}})\,.$$

We also have

$$(0_{\mathbb{Z}}, 0_{\mathbb{K}})\,\underbrace{a}_{=(n,x)} = (0_{\mathbb{Z}}, 0_{\mathbb{K}})\,(n,x) = (0_{\mathbb{Z}}n, 0_{\mathbb{Z}}x + n0_{\mathbb{K}} + 0_{\mathbb{K}}x)$$

(by the definition of the multiplication on $\mathbb{L}$). In view of $0_{\mathbb{Z}} n = 0_{\mathbb{Z}}$ and

$$\underbrace{0_{\mathbb{Z}} x}_{\substack{=0_{\mathbb{K}} \\ \text{(by (8))}}} + \underbrace{n 0_{\mathbb{K}}}_{\substack{=0_{\mathbb{K}} \\ \text{(by (6))}}} + \underbrace{0_{\mathbb{K}} x}_{=0_{\mathbb{K}}} = 0_{\mathbb{K}} + 0_{\mathbb{K}} + 0_{\mathbb{K}} = 0_{\mathbb{K}},$$

this rewrites as

$$(0_{\mathbb{Z}}, 0_{\mathbb{K}})\, a = (0_{\mathbb{Z}}, 0_{\mathbb{K}})\,.$$

Now, combining the two equalities $a\,(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (0_{\mathbb{Z}}, 0_{\mathbb{K}})$ and $(0_{\mathbb{Z}}, 0_{\mathbb{K}})\, a = (0_{\mathbb{Z}}, 0_{\mathbb{K}})$, we obtain $a\,(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (0_{\mathbb{Z}}, 0_{\mathbb{K}})\, a = (0_{\mathbb{Z}}, 0_{\mathbb{K}})$. In other words, $a\,(0,0) = (0,0)\, a = (0,0)$ (since $(0_{\mathbb{Z}}, 0_{\mathbb{K}}) = (0,0)$). Thus, the "Annihilation" axiom is proven for $\mathbb{L}$.]

[*Proof of the "Distributivity" axiom:* Let $a, b, c \in \mathbb{L}$. We must prove that $a\,(b+c) = ab + ac$ and $(a+b)\, c = ac + bc$.

We have $a \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write a in the form $a = (n, x)$ for some $n \in \mathbb{Z}$ and $x \in \mathbb{K}$. Consider these n and x.

We have $b \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write b in the form $b = (m, y)$ for some $m \in \mathbb{Z}$ and $y \in \mathbb{K}$. Consider these m and y.

We have $c \in \mathbb{L} = \mathbb{Z} \times \mathbb{K}$. Thus, we can write c in the form $c = (p, z)$ for some $p \in \mathbb{Z}$ and $z \in \mathbb{K}$. Consider these p and z.

Recall that $\mathbb{K}$ is a nonunital ring; thus, the distributivity axiom holds in $\mathbb{K}$. Hence, $x\,(y+z) = xy + xz$ and $(x+y)\, z = xz + yz$.

We have

$$\underbrace{a}_{=(n,x)}\ \underbrace{b}_{=(m,y)} = (n,x)\,(m,y) = (nm, ny + mx + xy) \tag{13}$$

(by the definition of the multiplication on $\mathbb{L}$). Similarly, using $c = (p, z)$, we obtain

$$ac = (np, nz + px + xz) \tag{14}$$

and

$$bc = (mp, mz + py + yz)\,. \tag{15}$$

Furthermore,

$$\underbrace{b}_{=(m,y)} + \underbrace{c}_{=(p,z)} = (m,y) + (p,z) = (m+p, y+z)$$

(by the definition of the addition on $\mathbb{L}$) and thus

$$\underbrace{a}_{=(n,x)}\ \underbrace{(b+c)}_{=(m+p,y+z)} = (n,x)\,(m+p, y+z) = (n\,(m+p)\,, n\,(y+z) + (m+p)\, x + x\,(y+z))$$

(by the definition of the multiplication on $\mathbb{L}$). In view of

$$n\,(m+p) = nm + np \qquad \text{(since distributivity holds for integers)}$$

and

$$\underbrace{n\,(y+z)}_{\substack{=ny+nz \\ \text{(by (2))}}} + \underbrace{(m+p)\, x}_{\substack{=mx+px \\ \text{(by (1))}}} + \underbrace{x\,(y+z)}_{=xy+xz} = ny + nz + mx + px + xy + xz,$$

this rewrites as

$$a\,(b+c) = (nm + np, ny + nz + mx + px + xy + xz)\,. \tag{16}$$

On the other hand, adding the equalities (13) and (14) together, we obtain

$$\begin{aligned} ab+ac &= (nm, ny+mx+xy) + (np, nz+px+xz) \\ &= \left(nm+np, \underbrace{ny+mx+xy+nz+px+xz}_{=ny+nz+mx+px+xy+xz} \right) \\ &\qquad \text{(by the definition of addition in } \mathbb{L}) \\ &= (nm+np, ny+nz+mx+px+xy+xz)\,. \end{aligned}$$

Comparing this with (16), we obtain $a(b+c) = ab+ac$.

Moreover,

$$\underbrace{a}_{=(n,x)} + \underbrace{b}_{=(m,y)} = (n,x)+(m,y) = (n+m, x+y)$$

(by the definition of the addition on $\mathbb{L}$) and thus

$$\underbrace{(a+b)}_{=(n+m,x+y)} \underbrace{c}_{=(p,z)} = (n+m, x+y)(p,z) = ((n+m)p, (n+m)z + p(x+y) + (x+y)z)$$

(by the definition of the multiplication on $\mathbb{L}$). In view of

$$(n+m)p = np+mp \qquad \text{(since distributivity holds for integers)}$$

and

$$\underbrace{(n+m)z}_{\substack{=nz+mz \\ \text{(by (1))}}} + \underbrace{p(x+y)}_{\substack{=px+py \\ \text{(by (2))}}} + \underbrace{(x+y)z}_{=xz+yz} = nz+mz+px+py+xz+yz,$$

this rewrites as

$$(a+b)c = (np+mp, nz+pz+px+py+xz+yz)\,. \tag{17}$$

On the other hand, adding the equalities (14) and (15) together, we obtain

$$\begin{aligned} ac+bc &= (np, nz+px+xz) + (mp, mz+py+yz) \\ &= \left(np+mp, \underbrace{nz+px+xz+mz+py+yz}_{=nz+mz+px+py+xz+yz} \right) \\ &\qquad \text{(by the definition of addition in } \mathbb{L}) \\ &= (np+mp, nz+pz+px+py+xz+yz)\,. \end{aligned}$$

Comparing this with (17), we obtain $(a+b)c = ac+bc$.

Now, we have proven that

$$a(b+c) = ab+ac \qquad \text{and} \qquad (a+b)c = ac+bc.$$

Thus, the "Distributivity" axiom is proven for $\mathbb{L}$.]

We have now proven all eight ring axioms for $\mathbb{L}$. Thus, $\mathbb{L}$ is a ring. The exercise is solved.

2.4 Remark

1. The construction of $\mathbb{L}$ can be viewed as a way of "adjoining" a unity to a nonunital ring $\mathbb{K}$. What happens if our original ring $\mathbb{K}$ already had a unity $1_{\mathbb{K}}$ to begin with? Is $\mathbb{L}$ then going to have two different unities?

No, because the original unity $1_{\mathbb{K}}$ of $\mathbb{K}$ will **not** in any way become a unity of $\mathbb{L}$. The unity $(1_{\mathbb{Z}}, 0_{\mathbb{K}})$ of $\mathbb{L}$ has nothing to do with $1_{\mathbb{K}}$. So we gain a new unity when we go from $\mathbb{K}$ to $\mathbb{L}$, but we "lose" our old unity in the process.

2. More can be said about the case when $\mathbb{K}$ has a unity $1_{\mathbb{K}}$. Indeed, in this case, the map

$$\begin{aligned} &\mathbb{Z} \times \mathbb{K} \to \mathbb{L}, \\ &(n, a) \mapsto (n, a - n1_{\mathbb{K}}) \end{aligned}$$

is a ring isomorphism (where the ring structure on $\mathbb{Z} \times \mathbb{K}$ is entrywise – i.e., we regard $\mathbb{Z} \times \mathbb{K}$ as the Cartesian product of the two rings $\mathbb{Z}$ and $\mathbb{K}$). This can be checked straightforwardly (we leave this to the reader).

3. The exercise can be generalized. Indeed, fix a further commutative ring $\mathbb{B}$, and assume that $\mathbb{K}$ is not just a nonunital ring but also a nonunital $\mathbb{B}$-algebra.[3] Then, we can replace $\mathbb{Z}$ by $\mathbb{B}$ in the above definition of $\mathbb{L}$, and the result will still be a well-defined ring, and furthermore (this is not hard to check) a well-defined $\mathbb{B}$-algebra (in the original sense, i.e., with a unity that satisfies the "Neutrality of one" axiom). Thus, any nonunital $\mathbb{B}$-algebra can be embedded into a $\mathbb{B}$-algebra.

---

3 Exercise 3: More sums from number theory

3.1 Problem

(a) Let n be a positive integer. Prove that

$$\sum_{j=1}^{n} \gcd(j, n) = \sum_{d \mid n} d\phi\left(\frac{n}{d}\right). \tag{18}$$

More generally, if $(a_1, a_2, a_3, \ldots)$ is a sequence of reals, then prove that

$$\sum_{j=1}^{n} a_{\gcd(j,n)} = \sum_{d \mid n} a_d \phi\left(\frac{n}{d}\right). \tag{19}$$

(b) Let $n \in \mathbb{N}$. Prove that

$$\begin{aligned} &\left(\text{the number of } (x, y) \in \mathbb{Z}^2 \text{ satisfying } x^2 + y^2 \leq n\right) \\ &= 1 + 4 \sum_{k \in \mathbb{N}} (-1)^k \left\lfloor \frac{n}{2k+1} \right\rfloor \\ &= 1 + 4\left(\left\lfloor \frac{n}{1} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor + \left\lfloor \frac{n}{5} \right\rfloor - \left\lfloor \frac{n}{7} \right\rfloor + \left\lfloor \frac{n}{9} \right\rfloor - \left\lfloor \frac{n}{11} \right\rfloor \pm \cdots\right). \end{aligned}$$

[3] A *nonunital $\mathbb{B}$-algebra* means exactly what you think: Take the definition of a $\mathbb{B}$-algebra that we gave in class, and remove the unity along with the "Neutrality of one" axiom.

(The infinite sums in this equality have only finitely many nonzero addends, and thus are well-defined.)

[**Hint:** Parts **(a)** and **(b)** have nothing to do with each other.

This is a good place for a reminder that results proven in the notes, as well as problems from previous homework sets and midterms, can be freely used. Both parts have rather short solutions if you remember the right results to use!]

3.2 Remark

Part **(b)** of this exercise is a "discrete" version of the famous Madhava–Gregory–Leibniz series

$$\frac{\pi}{4} = \frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \pm \cdots$$

(where π, at last, does denote the area of the unit circle). Indeed, if we divide the number of $(x, y) \in \mathbb{Z}^2$ satisfying $x^2 + y^2 \leq n$ by n, then we obtain an approximation to the area of the unit circle that gets better as n grows[4]. On the other hand, it appears reasonable that dividing

$$1 + 4\left(\left\lfloor \frac{n}{1} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor + \left\lfloor \frac{n}{5} \right\rfloor - \left\lfloor \frac{n}{7} \right\rfloor + \left\lfloor \frac{n}{9} \right\rfloor - \left\lfloor \frac{n}{11} \right\rfloor \pm \cdots\right)$$

by n, we obtain an approximation to $4\left(\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \pm \cdots\right)$. I am not sure whether this can be rigorously proven, however.[5]

3.3 Solution sketch

(a) Let $(a_1, a_2, a_3, \ldots)$ be a sequence of reals. We shall prove (19).

Indeed, we recall the following fact (which is Lemma 2.14.8 in the class notes):

Lemma 3.1. *Let n be a positive integer. Let d be a positive divisor of n. Then,*

$$(\textit{the number of } i \in \{1, 2, \ldots, n\} \textit{ such that } \gcd(i, n) = d) = \phi(n/d).$$

Now, we have the following equality of summation signs:

$$\sum_{i=1}^{n} = \sum_{i \in \{1,2,\ldots,n\}} = \sum_{d \mid n} \sum_{\substack{i \in \{1,2,\ldots,n\}; \\ \gcd(i,n)=d}}$$

[4] Just observe that the pairs $(x, y) \in \mathbb{Z}^2$ satisfying $x^2 + y^2 \leq n$, regarded as points in the Euclidean plane, are precisely the lattice points inside the circle with center 0 and radius $\sqrt{n}$. Thus, by counting these pairs, we are approximating the area of this circle. See [Clark18, Theorem 12.1] for a rigorous proof.

[5] Of course, for any given $k \in \mathbb{N}$, the number $\frac{1}{n}\left(\left\lfloor \frac{n}{2k+1} \right\rfloor - \frac{n}{2k+1}\right)$ does converge to 0 when $n \to \infty$. But here we are taking an alternating sum of infinitely many such numbers; we can ignore all but the first n, but even the first n may no longer converge to 0 when summed together.

(because if $i \in \{1,2,\ldots,n\}$, then $\gcd(i,n)$ is a positive divisor of n). Thus,

$$\underbrace{\sum_{i=1}^{n}}_{=\sum_{d\mid n}\ \sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(i,n)=d}}} a_{\gcd(i,n)} = \sum_{d\mid n}\ \sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(i,n)=d}} \underbrace{a_{\gcd(i,n)}}_{\substack{=a_d\\ \text{(since } \gcd(i,n)=d\text{)}}}$$

$$= \sum_{d\mid n} \underbrace{\sum_{\substack{i\in\{1,2,\ldots,n\};\\ \gcd(i,n)=d}} a_d}_{=(\text{the number of } i\in\{1,2,\ldots,n\} \text{ such that } \gcd(i,n)=d)\cdot a_d}$$

$$= \sum_{d\mid n} \underbrace{(\text{the number of } i \in \{1,2,\ldots,n\} \text{ such that } \gcd(i,n)=d)}_{\substack{=\phi(n/d)\\ \text{(by Lemma 3.1)}}} \cdot a_d$$

$$= \sum_{d\mid n} \phi\left(\underbrace{n/d}_{=\frac{n}{d}}\right) a_d = \sum_{d\mid n} \phi\left(\frac{n}{d}\right) a_d = \sum_{d\mid n} a_d \phi\left(\frac{n}{d}\right).$$

Hence,

$$\sum_{j=1}^{n} a_{\gcd(j,n)} = \sum_{i=1}^{n} a_{\gcd(i,n)} \qquad \text{(here, we have renamed the summation index } j \text{ as } i\text{)}$$

$$= \sum_{d\mid n} a_d \phi\left(\frac{n}{d}\right).$$

This proves (19).

Now, forget that we fixed $(a_1, a_2, a_3, \ldots)$. We have proven (19) for each sequence $(a_1, a_2, a_3, \ldots)$ of reals. Hence, we can apply (19) to $a_i = i$. We thus obtain

$$\sum_{j=1}^{n} \gcd(j,n) = \sum_{d\mid n} d\phi\left(\frac{n}{d}\right).$$

This proves (18). Hence, part **(a)** of the exercise is solved.

(b) Forget that we fixed n. We recall the following result (Proposition 2.8.3 in the class notes):

Proposition 3.2. *Let a and b be integers such that $b > 0$. Then, $\left\lfloor \frac{a}{b} \right\rfloor$ is well-defined and equals $a//b$.*

Let us use the Iverson bracket notation. We also recall the following result (Exercise 2.17.2 **(a)** in the class notes):

Proposition 3.3. *We have $n//k = \sum_{i=1}^{n} [k \mid i]$ for any $n \in \mathbb{N}$ and any positive integer k.*

Combining these two propositions, we easily obtain the following:

Corollary 3.4. *Let j be a positive integer. Let $n \in \mathbb{N}$. Then,*

$$\sum_{m=1}^{n} [j \mid m] = \left\lfloor \frac{n}{j} \right\rfloor .$$

Proof of Corollary 3.4. Proposition 3.2 (applied to $a = n$ and $b = j$) yields that $\left\lfloor \frac{n}{j} \right\rfloor$ is well-defined and equals $n//j$. Thus,

$$\begin{aligned} \left\lfloor \frac{n}{j} \right\rfloor = n//j &= \sum_{i=1}^{n} [j \mid i] \qquad \text{(by Proposition 3.3, applied to } k = j) \\ &= \sum_{m=1}^{n} [j \mid m] \qquad \text{(here, we have renamed the summation index } i \text{ as } m) . \end{aligned}$$

This proves Corollary 3.4. □

Finally, we recall the following result (Exercise 2 on homework set #5):

Proposition 3.5. *Let n be a positive integer. Then,*

$$\begin{aligned} &\left(\text{the number of pairs } (x, y) \in \mathbb{Z}^2 \text{ such that } n = x^2 + y^2\right) \\ &= 4 \left(\text{the number of positive divisors } d \text{ of } n \text{ such that } d \equiv 1 \bmod 4\right) \\ &\quad - 4 \left(\text{the number of positive divisors } d \text{ of } n \text{ such that } d \equiv 3 \bmod 4\right) . \end{aligned}$$

Let us restate this proposition in a more convenient (for our current goals) form:

Corollary 3.6. *Let n be a positive integer. Then,*

$$\left(\text{the number of pairs } (x, y) \in \mathbb{Z}^2 \text{ such that } n = x^2 + y^2\right) = 4 \sum_{i \in \mathbb{N}} (-1)^i [2i + 1 \mid n] .$$

(In particular, the sum $\sum_{i \in \mathbb{N}} (-1)^i [2i + 1 \mid n]$ is well-defined, since all but finitely many of its addends are 0.)

Proof of Corollary 3.6. Every $i \in \mathbb{N}$ satisfying $i \geq n$ must satisfy $(-1)^i [2i + 1 \mid n] = 0$ [6]. Hence, all but finitely many $i \in \mathbb{N}$ must satisfy $(-1)^i [2i + 1 \mid n] = 0$ (since all but finitely many $i \in \mathbb{N}$ satisfy $i \geq n$). Thus, all but finitely many addends of the sum $\sum_{i \in \mathbb{N}} (-1)^i [2i + 1 \mid n]$ are 0. Thus, this sum is well-defined.

[6] *Proof.* Let $i \in \mathbb{N}$ be such that $i \geq n$. We must prove that $(-1)^i [2i + 1 \mid n] = 0$.

The integer $2i + 1$ is positive (since $i \in \mathbb{N}$); thus, $|2i + 1| = 2i + 1$. The integer n is positive as well; thus, $|n| = n$. Thus,

$$|2i + 1| = 2i + 1 = i + \underbrace{(i + 1)}_{\substack{>0 \\ (\text{since } i \in \mathbb{N})}} > i \geq n = |n| .$$

But if we had $2i + 1 \mid n$, then Proposition 2.2.3 **(b)** in the class notes (applied to $a = 2i + 1$ and $b = n$) would yield $|2i + 1| \leq |n|$, which would contradict $|2i + 1| > |n|$. Hence, we cannot have $2i + 1 \mid n$. Thus, $[2i + 1 \mid n] = 0$ and therefore $(-1)^i \underbrace{[2i + 1 \mid n]}_{=0} = 0$. Qed.

Each $i \in \mathbb{N}$ is either even or odd (but never both simultaneously). Thus, we can split the sum $\sum_{i \in \mathbb{N}} (-1)^i [2i+1 \mid n]$ as follows:

$$\begin{aligned}
\sum_{i \in \mathbb{N}} (-1)^i [2i+1 \mid n] &= \sum_{\substack{i \in \mathbb{N}; \\ i \text{ is even}}} \underbrace{(-1)^i}_{\substack{=1 \\ (\text{since } i \text{ is even})}} [2i+1 \mid n] + \sum_{\substack{i \in \mathbb{N}; \\ i \text{ is odd}}} \underbrace{(-1)^i}_{\substack{=-1 \\ (\text{since } i \text{ is odd})}} [2i+1 \mid n] \\
&= \sum_{\substack{i \in \mathbb{N}; \\ i \text{ is even}}} [2i+1 \mid n] + \underbrace{\sum_{\substack{i \in \mathbb{N}; \\ i \text{ is odd}}} (-1) [2i+1 \mid n]}_{= - \sum\limits_{\substack{i \in \mathbb{N}; \\ i \text{ is odd}}} [2i+1 \mid n]} \\
&= \sum_{\substack{i \in \mathbb{N}; \\ i \text{ is even}}} [2i+1 \mid n] - \sum_{\substack{i \in \mathbb{N}; \\ i \text{ is odd}}} [2i+1 \mid n] . \qquad (20)
\end{aligned}$$

Exercise 2.7.3 **(b)** in the class notes shows that that the map

$$\begin{aligned}
\{i \in \mathbb{N} \mid i \text{ is odd}\} &\to \{d \in \mathbb{N} \mid d \equiv 3 \bmod 4\}, \\
i &\mapsto 2i+1
\end{aligned}$$

is well-defined and is a bijection.[7]

Hence, we can substitute d for $2i+1$ in the sum $\sum\limits_{\substack{i \in \mathbb{N}; \\ i \text{ is odd}}} [2i+1 \mid n]$, and thus obtain

$$\begin{aligned}
&\sum_{\substack{i \in \mathbb{N}; \\ i \text{ is odd}}} [2i+1 \mid n] \\
&= \sum_{\substack{d \in \mathbb{N}; \\ d \equiv 3 \bmod 4}} [d \mid n] = \sum_{\substack{d \in \mathbb{N}; \\ d \equiv 3 \bmod 4; \\ d \mid n}} \underbrace{[d \mid n]}_{\substack{=1 \\ (\text{since } d \mid n)}} + \sum_{\substack{d \in \mathbb{N}; \\ d \equiv 3 \bmod 4; \\ d \nmid n}} \underbrace{[d \mid n]}_{\substack{=0 \\ (\text{since } d \nmid n)}} \\
&\qquad \left(\begin{array}{c} \text{since each } d \in \mathbb{N} \text{ satisfies either } d \mid n \text{ or } d \nmid n \\ \text{(but never both at the same time)} \end{array} \right) \\
&= \sum_{\substack{d \in \mathbb{N}; \\ d \equiv 3 \bmod 4; \\ d \mid n}} 1 + \underbrace{\sum_{\substack{d \in \mathbb{N}; \\ d \equiv 3 \bmod 4; \\ d \nmid n}} 0}_{=0} = \sum_{\substack{d \in \mathbb{N}; \\ d \equiv 3 \bmod 4; \\ d \mid n}} 1 \\
&= (\text{the number of all } d \in \mathbb{N} \text{ satisfying } d \equiv 3 \bmod 4 \text{ and } d \mid n) \cdot 1 \\
&= (\text{the number of all } d \in \mathbb{N} \text{ satisfying } d \equiv 3 \bmod 4 \text{ and } d \mid n) \\
&= (\text{the number of all positive integers } d \text{ satisfying } d \equiv 3 \bmod 4 \text{ and } d \mid n) \\
&\qquad \left(\begin{array}{c} \text{since each } d \in \mathbb{N} \text{ satisfying } d \equiv 3 \bmod 4 \text{ must be a} \\ \text{positive integer (because 0 does not satisfy } 0 \equiv 3 \bmod 4) \end{array} \right) \\
&= (\text{the number of all positive integers } d \text{ satisfying } d \mid n \text{ and } d \equiv 3 \bmod 4) \\
&= (\text{the number of all positive divisors } d \text{ of } n \text{ such that } d \equiv 3 \bmod 4) \qquad (21)
\end{aligned}$$

(since the positive integers d satisfying $d \mid n$ are precisely the positive divisors d of n). A similar argument (but with the word "odd" replaced by "even", and with each appearance of

[7] This map sends $1, 3, 5, 7, 9, \ldots$ to $3, 7, 11, 15, 19, \ldots$, respectively.

"3" replaced by "1" [8]) shows that

$$\sum_{\substack{i\in\mathbb{N};\\ i \text{ is even}}} [2i+1 \mid n]$$
$$= (\text{the number of all positive divisors } d \text{ of } n \text{ such that } d \equiv 1 \bmod 4)\,. \tag{22}$$

Now, if we multiply both sides of the equality (20) by 4, then we obtain

$$\begin{aligned}
&4\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid n] \\
&= 4\left(\sum_{\substack{i\in\mathbb{N};\\ i \text{ is even}}} [2i+1 \mid n] - \sum_{\substack{i\in\mathbb{N};\\ i \text{ is odd}}} [2i+1 \mid n] \right) \\
&= 4 \underbrace{\sum_{\substack{i\in\mathbb{N};\\ i \text{ is even}}} [2i+1 \mid n]}_{\substack{=(\text{the number of all positive divisors } d \text{ of } n \text{ such that } d\equiv 1 \bmod 4)\\ (\text{by (22)})}} \\
&\quad - 4 \underbrace{\sum_{\substack{i\in\mathbb{N};\\ i \text{ is odd}}} [2i+1 \mid n]}_{\substack{=(\text{the number of all positive divisors } d \text{ of } n \text{ such that } d\equiv 3 \bmod 4)\\ (\text{by (21)})}} \\
&= 4\,(\text{the number of positive divisors } d \text{ of } n \text{ such that } d \equiv 1 \bmod 4) \\
&\quad - 4\,(\text{the number of positive divisors } d \text{ of } n \text{ such that } d \equiv 3 \bmod 4)\,.
\end{aligned}$$

Comparing this with

$$\begin{aligned}
&\left(\text{the number of pairs } (x,y) \in \mathbb{Z}^2 \text{ such that } n = x^2+y^2\right) \\
&= 4\,(\text{the number of positive divisors } d \text{ of } n \text{ such that } d \equiv 1 \bmod 4) \\
&\quad - 4\,(\text{the number of positive divisors } d \text{ of } n \text{ such that } d \equiv 3 \bmod 4) \\
&\qquad (\text{by Proposition 3.5})\,,
\end{aligned}$$

we obtain

$$\left(\text{the number of pairs } (x,y) \in \mathbb{Z}^2 \text{ such that } n = x^2+y^2\right) = 4\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid n]\,.$$

Thus, the proof of Corollary 3.6 is complete. □

Now, let $n \in \mathbb{N}$. Let $m \in \{1,2,\ldots,n\}$. Then, m is a positive integer. Hence, Corollary 3.6 (applied to m instead of n) shows that

$$\begin{aligned}
&\left(\text{the number of pairs } (x,y) \in \mathbb{Z}^2 \text{ such that } m = x^2+y^2\right) \\
&= 4\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]
\end{aligned} \tag{23}$$

(and that the sum $\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]$ is well-defined).

[8] and with "Exercise 2.7.3 **(b)** in the class notes" replaced by "Exercise 2.7.3 **(a)** in the class notes"

Forget that we fixed m. We thus have proven that for each $m \in \{1,2,\ldots,n\}$, the equality (23) holds (and the sum $\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]$ is well-defined).

Hence, we know that the sum $\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]$ is well-defined for each $m \in \{1,2,\ldots,n\}$. Thus, the sum of these n many sums can be rewritten as follows:

$$\sum_{m=1}^{n} \sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m] = \sum_{i\in\mathbb{N}} \underbrace{\sum_{m=1}^{n} (-1)^i [2i+1 \mid m]}_{=(-1)^i \sum_{m=1}^{n} [2i+1 \mid m]} = \sum_{i\in\mathbb{N}} (-1)^i \underbrace{\sum_{m=1}^{n} [2i+1 \mid m]}_{=\left\lfloor \frac{n}{2i+1} \right\rfloor \text{ (by Corollary 3.4, applied to } j=2i+1)}$$

$$= \sum_{i\in\mathbb{N}} (-1)^i \left\lfloor \frac{n}{2i+1} \right\rfloor. \tag{24}$$

But

$$\begin{aligned}
&\big(\text{the number of } (x,y) \in \mathbb{Z}^2 \text{ satisfying } x^2+y^2 \leq n\big) \\
&= \big(\text{the number of } (x,y) \in \mathbb{Z}^2 \text{ satisfying } x^2+y^2 \in \{0,1,\ldots,n\}\big) \\
&\qquad \left(\begin{array}{c} \text{because for any } (x,y) \in \mathbb{Z}^2, \text{ the statement "} x^2+y^2 \leq n \text{" is equivalent} \\ \text{to "} x^2+y^2 \in \{0,1,\ldots,n\} \text{" (since } x^2+y^2 \text{ is a nonnegative integer)} \end{array} \right) \\
&= \sum_{m=0}^{n} \big(\text{the number of } (x,y) \in \mathbb{Z}^2 \text{ satisfying } x^2+y^2 = m\big) \\
&= \underbrace{\big(\text{the number of } (x,y) \in \mathbb{Z}^2 \text{ satisfying } x^2+y^2 = 0\big)}_{\substack{=1 \\ \text{(since there exists only one } (x,y)\in\mathbb{Z}^2 \text{ satisfying } x^2+y^2=0 \\ \text{(namely, } (x,y)=(0,0)))}} \\
&\qquad + \sum_{m=1}^{n} \underbrace{\big(\text{the number of } (x,y) \in \mathbb{Z}^2 \text{ satisfying } x^2+y^2 = m\big)}_{\substack{=4\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m] \\ \text{(by (23))}}} \\
&= 1 + \underbrace{\sum_{m=1}^{n} 4 \sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]}_{=4\sum_{m=1}^{n}\sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]} = 1 + 4 \underbrace{\sum_{m=1}^{n} \sum_{i\in\mathbb{N}} (-1)^i [2i+1 \mid m]}_{\substack{=\sum_{i\in\mathbb{N}} (-1)^i \left\lfloor \frac{n}{2i+1} \right\rfloor \\ \text{(by (24))}}} \\
&= 1 + 4 \sum_{i\in\mathbb{N}} (-1)^i \left\lfloor \frac{n}{2i+1} \right\rfloor = 1 + 4 \sum_{k\in\mathbb{N}} (-1)^k \left\lfloor \frac{n}{2k+1} \right\rfloor \\
&\qquad \text{(here, we have renamed the summation index } i \text{ as } k) \\
&= 1 + 4 \left(\left\lfloor \frac{n}{1} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor + \left\lfloor \frac{n}{5} \right\rfloor - \left\lfloor \frac{n}{7} \right\rfloor + \left\lfloor \frac{n}{9} \right\rfloor - \left\lfloor \frac{n}{11} \right\rfloor \pm \cdots \right)
\end{aligned}$$

(since

$$\sum_{k\in\mathbb{N}} (-1)^k \left\lfloor \frac{n}{2k+1} \right\rfloor = \left\lfloor \frac{n}{1} \right\rfloor - \left\lfloor \frac{n}{3} \right\rfloor + \left\lfloor \frac{n}{5} \right\rfloor - \left\lfloor \frac{n}{7} \right\rfloor + \left\lfloor \frac{n}{9} \right\rfloor - \left\lfloor \frac{n}{11} \right\rfloor \pm \cdots$$

). This solves part **(b)** of the exercise.

---

4 Exercise 4: Squares in finite fields II

4.1 Problem

Let $\mathbb{F}$ be a finite field such that $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$. In Exercise 5 of homework set #6, we have seen that $|\mathbb{F}|$ is odd, and that the number of squares in $\mathbb{F}$ is $\frac{1}{2}\left(|\mathbb{F}|+1\right)$.

In the following, the word "square" shall always mean "square in $\mathbb{F}$".

A *nonsquare* shall mean an element of $\mathbb{F}$ that is not a square.

Prove the following:

(a) The product of two squares is always a square.

(b) The product of a nonzero square with a nonsquare is always a nonsquare.

(c) The product of two nonsquares is always a square.

[**Hint:** It is easiest to solve the three parts in this exact order. For **(c)**, recall that if a subset Y of a finite set X satisfies $|Y| \geq |X|$, then $Y = X$.]

4.2 Remark

If $\mathbb{F}$ is the field $\mathbb{Z}/p$ for some prime $p > 2$, then the nonzero squares in $\mathbb{F}$ are called *quadratic residues modulo p*, while the nonsquares in $\mathbb{F}$ are called the *quadratic nonresidues modulo p*. These two types of residue classes have a long history; in particular, one of the most famous results in mathematics – Gauss's law of quadratic reciprocity – is concerned with them (Franz Lemmermeyer's website lists 246 different proofs of this law). Quadratic residues also have applications; they are involved in the Solovay-Strassen primality test [GalQua17, Chapter 6], and have been used to build diffusors (architectural features of a room that cause sound to spread evenly through the room – used, e.g., in concert halls).

4.3 Solution sketch

Recall that a square in $\mathbb{F}$ is defined to be an element of the form a^2 for some $a \in \mathbb{F}$.

An element of $\mathbb{F}$ is a nonsquare if and only if it is not a square (because this is how nonsquares were defined). Thus, each element of $\mathbb{F}$ is either a square or a nonsquare (but never both at the same time).

The ring $\mathbb{F}$ is a field, and thus is a commutative skew field. Every nonzero element of $\mathbb{F}$ is invertible (since $\mathbb{F}$ is a skew field).

(a) We must prove that if $u \in \mathbb{F}$ and $v \in \mathbb{F}$ are two squares, then uv is a square.

Let $u \in \mathbb{F}$ and $v \in \mathbb{F}$ be two squares. We thus must prove that uv is a square.

The element $u \in \mathbb{F}$ is a square. In other words, $u = a^2$ for some $a \in \mathbb{F}$ (by the definition of a square). Similarly, $v = b^2$ for some $b \in \mathbb{F}$. Consider these a and b. Now, multiplying the equalities $u = a^2$ and $v = b^2$, we obtain

$$uv = a^2 b^2 = (ab)^2 \qquad (\text{since } \mathbb{F} \text{ is commutative}).$$

Hence, $uv = c^2$ for some $c \in \mathbb{F}$ (namely, for $c = ab$). In other words, uv is a square (by the definition of a square). This solves part **(a)** of the exercise.

[*Remark:* So far we have not used that $\mathbb{F}$ is a finite field satisfying $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$; we only used that $\mathbb{F}$ is a commutative ring.]

(b) We must prove that if $u \in \mathbb{F}$ is a nonzero square and $v \in \mathbb{F}$ is a nonsquare, then uv is a nonsquare.

Let $u \in \mathbb{F}$ be a nonzero square, and let $v \in \mathbb{F}$ be a nonsquare. We thus must prove that uv is a nonsquare.

Assume the contrary. Thus, uv is a square (since each element of $\mathbb{F}$ is either a square or a nonsquare). In other words, $uv = c^2$ for some $c \in \mathbb{F}$. Also, $u = a^2$ for some $a \in \mathbb{F}$ (since u is a square). Consider these c and a.

We have $aa = a^2 = u \neq 0$ (since u is nonzero) and thus $a \neq 0$ (since $a = 0$ would lead to $aa = 0 \cdot 0 = 0$, which would contradict $aa \neq 0$). Hence, the element a of $\mathbb{F}$ is nonzero, and therefore invertible (since every nonzero element of $\mathbb{F}$ is invertible). Thus, the quotient $\dfrac{c}{a} \in \mathbb{F}$ is well-defined (since $\mathbb{F}$ is commutative). Denote this quotient by q. Thus, $q = \dfrac{c}{a} \in \mathbb{F}$. From $q = \dfrac{c}{a}$, we obtain $qa = c$. Moreover,

$$q^2 \underbrace{u}_{=a^2} = q^2 a^2 = \left(\underbrace{qa}_{=c} \right)^2 \qquad \text{(since } \mathbb{F} \text{ is commutative)}$$
$$= c^2 = uv.$$

The element u of $\mathbb{F}$ is nonzero and thus invertible (since every nonzero element of $\mathbb{F}$ is invertible). Hence, we can divide the equality $q^2 u = uv$ by u. We thus obtain $q^2 = v$. Hence, $v = q^2$. Thus, v is a square (by the definition of a square). This contradicts the fact that v is a nonsquare. This contradiction shows that our assumption was false; hence, uv is a nonsquare. This solves part **(b)** of the exercise.

[*Remark:* So far we have not used that $\mathbb{F}$ is finite; nor have we used that $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$.]

(c) We must prove that if $u \in \mathbb{F}$ and $v \in \mathbb{F}$ are two nonsquares, then uv is a square.

Let $u \in \mathbb{F}$ and $v \in \mathbb{F}$ be two nonsquares. We thus must prove that uv is a square.

In the statement of the exercise, it was said that the number of squares in $\mathbb{F}$ is $\dfrac{1}{2}(|\mathbb{F}| + 1)$. In other words,

$$(\text{the number of all squares in } \mathbb{F}) = \frac{1}{2}(|\mathbb{F}| + 1).$$

Clearly, $0 \in \mathbb{F}$ is a square (since $0 = 0^2$). Thus, if we had $u = 0$, then u would be a square, which would contradict the fact that u is a nonsquare. Hence, we cannot have $u = 0$. Thus, $u \neq 0$. Hence, the element u of $\mathbb{F}$ is nonzero and thus invertible (since every nonzero element of $\mathbb{F}$ is invertible). Thus, it has an inverse u^{-1}.

Let $f : \mathbb{F} \to \mathbb{F}$ be the map sending each $a \in \mathbb{F}$ to ua. Then, it is easy to see that the map f is invertible (indeed, its inverse sends each $a \in \mathbb{F}$ to $u^{-1}a$). Hence, f is bijective and thus injective. Therefore, f satisfies

$$|f(S)| = |S| \qquad \text{for each subset } S \text{ of } \mathbb{F}. \tag{25}$$

Let N be the set of all nonsquares in $\mathbb{F}$. Thus,

$$\begin{aligned} |N| &= (\text{the number of all nonsquares in } \mathbb{F}) \\ &= |\mathbb{F}| - \underbrace{(\text{the number of all squares in } \mathbb{F})}_{= \frac{1}{2}(|\mathbb{F}|+1)} \\ &\qquad \text{(because the nonsquares in } \mathbb{F} \text{ are precisely the elements of } \mathbb{F} \text{ that are not squares)} \\ &= |\mathbb{F}| - \frac{1}{2}(|\mathbb{F}| + 1) = \frac{1}{2}(|\mathbb{F}| - 1). \end{aligned}$$

Recall that $0 \in \mathbb{F}$ is a square in $\mathbb{F}$. Thus, the nonzero squares in $\mathbb{F}$ differ from the squares in $\mathbb{F}$ only in that 0 belongs to the latter but not to the former. Hence,

$$\begin{aligned}&(\text{the number of all nonzero squares in } \mathbb{F}) \\ &= \underbrace{(\text{the number of all squares in } \mathbb{F})}_{= \frac{1}{2}(|\mathbb{F}|+1)} - 1 = \frac{1}{2}(|\mathbb{F}|+1) - 1 = \frac{1}{2}(|\mathbb{F}|-1).\end{aligned}$$

Let S be the set of all nonzero squares in $\mathbb{F}$. Thus,

$$|S| = (\text{the number of all nonzero squares in } \mathbb{F}) = \frac{1}{2}(|\mathbb{F}|-1) = |N|$$

(since $|N| = \frac{1}{2}(|\mathbb{F}|-1)$).

Let $w \in S$. Then, w is a nonzero square (by the definition of S). Hence, the product wu is a product of a nonzero square with a nonsquare (since u is a nonsquare), and thus itself is a nonsquare (by part **(b)** of this exercise). In other words, $wu \in N$ (by the definition of N). Now, the definition of f yields $f(w) = uw = wu \in N$.

Now, forget that we fixed w. We thus have shown that $f(w) \in N$ for each $w \in S$. In other words, $f(S) \subseteq N$. In other words, $f(S)$ is a subset of N.

But (25) yields $|f(S)| = |S| = |N|$. Hence, $|f(S)| \geq |N|$.

Now, recall the following fundamental fact: If a subset Y of a finite set X satisfies $|Y| \geq |X|$, then $Y = X$. Applying this to $Y = f(S)$ and $X = N$, we obtain $f(S) = N$ (since $f(S)$ is a subset of the finite set N and satisfies $|f(S)| \geq |N|$).

Recall that we must show that uv is a square. Assume the contrary. Thus, uv is a nonsquare (by the definition of "nonsquare"). In other words, $uv \in N$ (by the definition of N). Hence, $uv \in N = f(S)$ (since $f(S) = N$). In other words, $uv = f(q)$ for some $q \in S$. Consider this q. But the definition of f yields $f(v) = uv = f(q)$. Therefore, $v = q$ (since f is injective). Thus, $v = q \in S$. In other words, v is a nonzero square (by the definition of S). Thus, v is a square; but this contradicts the fact that v is a nonsquare. This contradiction shows that our assumption was false. Hence, we have shown that uv is a square. This solves part **(c)** of the exercise.

4.4 Remark

The condition "$2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$" in the above exercise is necessary in order to ensure that $|\mathbb{F}|$ is odd, and that the number of squares in $\mathbb{F}$ is $\frac{1}{2}(|\mathbb{F}|+1)$. This was used in our solution of part **(c)** of the exercise. However, it turns out that all three parts of the exercise (including part **(c)**) still hold if we drop the condition "$2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$". In other words, the following holds:

Theorem 4.1. *Let $\mathbb{F}$ be a finite field (which may and may not satisfy $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$). In the following, the word "square" shall always mean "square in $\mathbb{F}$". A nonsquare shall mean an element of $\mathbb{F}$ that is not a square. Then:*

(a) *The product of two squares is always a square.*

(b) *The product of a nonzero square with a nonsquare is always a nonsquare.*

(c) *The product of two nonsquares is always a square.*

We are going to prove this theorem – it turns out that the bulk of the work has already been done in our above solution. All we need is to prove Theorem 4.1 **(c)** in the case when $2 \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$. This will rely on the following fact:

Proposition 4.2. *Let $\mathbb{F}$ be a finite field such that $2 \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$. Then, every element of $\mathbb{F}$ is a square.*

Proof of Proposition 4.2. The ring $\mathbb{F}$ is a field, and thus is a commutative skew field. Every nonzero element of $\mathbb{F}$ is invertible (since $\mathbb{F}$ is a skew field).

Let $F : \mathbb{F} \to \mathbb{F}$ be the map defined by

$$\left(F(a) = a^2 \qquad \text{for all } a \in \mathbb{F}\right).$$

We shall now show that the map F is injective.

Indeed, let $a, b \in \mathbb{F}$ satisfy $F(a) = F(b)$. We want to show that $a = b$.

Assume the contrary. Thus, $a \neq b$. In other words, $a - b \neq 0_{\mathbb{F}}$. Thus, the element $a - b$ of $\mathbb{F}$ is nonzero, and therefore invertible (since every nonzero element of $\mathbb{F}$ is invertible). Hence, its inverse $(a-b)^{-1} \in \mathbb{F}$ is well-defined.

The definition of F yields $F(b) = b^2$ and $F(a) = a^2$. Thus, $a^2 = F(a) = F(b) = b^2$. In other words, $a^2 - b^2 = 0_{\mathbb{F}}$. Hence,

$$0_{\mathbb{F}} = a^2 - b^2 = (a+b)(a-b) \qquad (\text{since } \mathbb{F} \text{ is commutative}).$$

We can multiply both sides of this equality by $(a-b)^{-1}$ (since $(a-b)^{-1}$ is well-defined). We thus obtain

$$0_{\mathbb{F}} \cdot (a-b)^{-1} = (a+b)\underbrace{(a-b)\cdot(a-b)^{-1}}_{=1_{\mathbb{F}}} = (a+b)\,1_{\mathbb{F}} = a+b.$$

Hence, $a + b = 0_{\mathbb{F}} \cdot (a-b)^{-1} = 0_{\mathbb{F}}$, so that $a = -b$. But $b + b = 2\underbrace{b}_{=1_{\mathbb{F}}b} = \underbrace{2 \cdot 1_{\mathbb{F}}}_{=0_{\mathbb{F}}} b = 0_{\mathbb{F}} b = 0_{\mathbb{F}}$ and thus $b = -b$. Comparing this with $a = -b$, we obtain $a = b$; this contradicts $a \neq b$. Hence, we have found a contradiction. Thus, our assumption was wrong; hence, $a = b$.

Now, forget that we fixed a, b. We thus have proven that if $a, b \in \mathbb{F}$ satisfy $F(a) = F(b)$, then $a = b$. In other words, the map F is injective.

But F is a map from the finite set $\mathbb{F}$ to the finite set $\mathbb{F}$, and these two finite sets satisfy $|\mathbb{F}| \geq |\mathbb{F}|$. Hence, Theorem 2.15.5 in the class notes[9] (applied to $A = \mathbb{F}$, $B = \mathbb{F}$ and $f = F$) shows that the map F is bijective (since F is injective). Thus, F is surjective. In other words, $\mathbb{F} = F(\mathbb{F})$.

Now, let $u \in \mathbb{F}$. Then, $u \in \mathbb{F} = F(\mathbb{F})$. In other words, there exists some $a \in \mathbb{F}$ such that $u = F(a)$. Consider this a. Now, $u = F(a) = a^2$ (by the definition of F). Hence, u is a square (by the definition of a square).

Forget that we fixed u. We thus have shown that each $u \in \mathbb{F}$ is a square. In other words, every element of $\mathbb{F}$ is a square. This proves Proposition 4.2. □

Remark 4.3. The map F we constructed in the proof of Proposition 4.2 is actually a ring isomorphism. Indeed, it is easy to show that F is a ring homomorphism (because $2 \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$ yields $(a+b)^2 = a^2 + b^2$ for all $a, b \in \mathbb{F}$ – this is a particular case of Freshman's Dream), and then the bijectivity of F entails that F is a ring isomorphism.

Now, we are ready to prove Theorem 4.1:

Proof of Theorem 4.1. When solving parts **(a)** and **(b)** of our above exercise, we never used the assumption that $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$. Hence, our solutions to these two parts still apply if this assumption is omitted. Thus, these solutions qualify as proofs of parts **(a)** and **(b)** of Theorem 4.1. It therefore remains to prove Theorem 4.1 **(c)**.

[9] This is the Pigeonhole Principle for Injections.

(c) If $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$, then Theorem 4.1 **(c)** holds (by part **(c)** of the above exercise). Thus, for the rest of this proof of Theorem 4.1 **(c)**, we WLOG assume that we don't have $2 \cdot 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$. Hence, we have $2 \cdot 1_{\mathbb{F}} = 0_{\mathbb{F}}$. Thus, Proposition 4.2 shows that every element of $\mathbb{F}$ is a square. Hence, there exist no nonsquares in $\mathbb{F}$ (since a nonsquare is defined as an element of $\mathbb{F}$ that is not a square). Thus, the claim of Theorem 4.1 **(c)** is vacuously true (since this claim only concerns nonsquares). This proves Theorem 4.1 **(c)**. □

---

5 Exercise 5: Formal differential calculus

5.1 Problem

Let $\mathbb{K}$ be a commutative ring. For each FPS[10]

$$f = \sum_{k \in \mathbb{N}} a_k x^k = a_0 x^0 + a_1 x^1 + a_2 x^2 + \cdots \in \mathbb{K}\left[\left[x\right]\right] \qquad \text{(where } a_i \in \mathbb{K}\text{)},$$

we define the *derivative* f' of f to be the FPS

$$\sum_{k>0} k a_k x^{k-1} = 1a_1 x^0 + 2a_2 x^1 + 3a_3 x^2 + \cdots \in \mathbb{K}\left[\left[x\right]\right].$$

(This definition imitates the standard procedure for differentiating power series in analysis, but it does not require any analysis or topology itself. In particular, $\mathbb{K}$ may be any commutative ring – e.g., a finite field.)

Let $D : \mathbb{K}\left[\left[x\right]\right] \to \mathbb{K}\left[\left[x\right]\right]$ be the map sending each FPS f to its derivative f'. We refer to D as *(formal) differentiation*. As usual, for any $n \in \mathbb{N}$, we let D^n denote $\underbrace{D \circ D \circ \cdots \circ D}_{n \text{ times}}$ (which means id if $n = 0$).

Prove the following:

(a) If $f \in \mathbb{K}\left[x\right]$, then $f' \in \mathbb{K}\left[x\right]$ and $\deg\left(f'\right) \leq \deg f - 1$. (In other words, the derivative of a polynomial is again a polynomial of degree at least 1 less.)

(b) The map $D : \mathbb{K}\left[\left[x\right]\right] \to \mathbb{K}\left[\left[x\right]\right]$ is $\mathbb{K}$-linear (with respect to the $\mathbb{K}$-module structure on $\mathbb{K}\left[\left[x\right]\right]$ defined in class – i.e., both addition and scaling of FPSs are defined entrywise).

(c) We have $\left(fg\right)' = f'g + fg'$ for any two FPSs f and g. (This is called the *Leibniz rule*.)

(d) We have $D^n\left(x^k\right) = n!\dbinom{k}{n} x^{k-n}$ for all $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Here, the expression "$\dbinom{k}{n} x^{k-n}$" is to be understood as 0 when $k < n$.

[10] Just as in class, the abbreviation "FPS" stands for "formal power series". All FPSs and polynomials in this exercise are in 1 indeterminate over $\mathbb{K}$; the indeterminate is called x.

(e) If $\mathbb{Q}$ is a subring of $\mathbb{K}$, then every polynomial $f \in \mathbb{K}[x]$ satisfies[11]

$$f[x+a] = \sum_{n \in \mathbb{N}} \frac{1}{n!} \left(D^n(f)\right)[a] \cdot x^n \qquad \text{for all } a \in \mathbb{K}.$$

(The infinite sum on the right hand side has only finitely many nonzero addends.)

(f) If p is a prime such that $p \cdot 1_{\mathbb{K}} = 0$ (for example, this happens if $\mathbb{K} = \mathbb{Z}/p$), then $D^p(f) = 0$ for each $f \in \mathbb{K}[[x]]$.

Now, assume that $\mathbb{Q}$ is a subring of $\mathbb{K}$. For each FPS

$$f = \sum_{k \in \mathbb{N}} a_k x^k = a_0 x^0 + a_1 x^1 + a_2 x^2 + \cdots \in \mathbb{K}[[x]] \qquad (\text{where } a_i \in \mathbb{K}),$$

we define the *integral* $\int f$ of f to be the FPS

$$\sum_{k \geq 0} \frac{1}{k+1} a_k x^{k+1} = \frac{1}{1} a_0 x^1 + \frac{1}{2} a_1 x^2 + \frac{1}{3} a_2 x^3 + \cdots \in \mathbb{K}[[x]].$$

(This definition imitates the standard procedure for integrating power series in analysis, but again works for any commutative ring $\mathbb{K}$ that contains $\mathbb{Q}$ as subring.)

Let $J : \mathbb{K}[[x]] \to \mathbb{K}[[x]]$ be the map sending each FPS f to its integral $\int f$. Prove the following:

(g) The map $J : \mathbb{K}[[x]] \to \mathbb{K}[[x]]$ is $\mathbb{K}$-linear.

(h) We have $D \circ J = \mathrm{id}$.

(i) We have $J \circ D \neq \mathrm{id}$.

[**Hint:** Don't give too much detail; workable outlines are sufficient. Feel free to interchange summation signs without justification. For part **(c)**, it is easiest to first prove it in the particular case when $f = x^a$ and $g = x^b$ for some f and g, and then obtain the general case by interchanging summations.]

5.2 Remark

This exercise is just the beginning of "algebraic calculus". A lot more can be done: Differentiation can be extended to rational functions; partial derivatives can be defined for multivariate polynomials and FPSs; differential equations can be solved formally in FPSs (rather than functions); even a purely algebraic analogue of the classical $f'(x) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon}$ definition exists[12]. These algebraic derivatives play crucial roles in the study of fields (including finite fields!), in algebraic geometry (where they help define what a "singularity" of an algebraic variety is) and in enumerative combinatorics (where they aid in computing generating functions[13]).

[11] Just as in class, I am using the notation "$f[u]$" for the evaluation of f at u. The more common notation for this is $f(u)$, but is too easily mistaken for a product.
Note also that we need to require f to be a polynomial here, since $f[x+a]$ would not be defined if f was merely an FPS.

[12] See Theorem 5 in `https://math.stackexchange.com/a/2974977/` .

[13] see [Loehr11, §7.8 and further on], [Wilf94] and [GrKnPa94, §7.2 and further on] for examples

Part **(e)** is an algebraic counterpart of the well-known Taylor series from calculus (and it is much easier than the latter: no error terms, no smoothness requirements, no convergence issues).

The "integral" $\int f$ we defined above is, of course, only one possible choice of an FPS g satisfying $g' = f$. Just as in calculus, you can add any constant to it, and you get another. Part **(h)** is an algebraic analogue of the first Fundamental Theorem of Calculus. You can easily prove an analogue of the second Fundamental Theorem too: For each FPS f, the FPS $(J \circ D)(f)$ differs from f only in its constant term.

If $\mathbb{K}$ contains $\mathbb{Q}$ as a subring, then both J and D are elements of the $\mathbb{K}$-algebra $\operatorname{End}(\mathbb{K}[[x]])$ (by parts **(b)** and **(g)** of this exercise). Part **(h)** of this exercise shows that J is a right inverse of D; but part **(i)** shows that J is not a left inverse (and thus not an inverse) of D. This yields an example of a left inverse that is not a right inverse.

5.3 Solution sketch

We recall the following notation (which we introduced in the class notes): For each $n \in \mathbb{Z}$, we define a subset $\mathbb{K}[x]_{\leq n}$ of $\mathbb{K}[[x]]$ by

$$\begin{aligned}\mathbb{K}[x]_{\leq n} &= \{(a_0, a_1, a_2, \ldots) \in \mathbb{K}[[x]] \mid a_k = 0 \text{ for all } k > n\} \\ &= \left\{\mathbf{a} \in \mathbb{K}[[x]] \mid \left[x^k\right]\mathbf{a} = 0 \text{ for all } k > n\right\}.\end{aligned}$$

We know that a FPS $\mathbf{a} \in \mathbb{K}[[x]]$ belongs to $\mathbb{K}[x]_{\leq n}$ (for a given $n \in \mathbb{Z}$) if and only if $\mathbf{a}$ is a polynomial of degree $\leq n$. We also know that $\mathbb{K}[x]_{\leq n}$ is a $\mathbb{K}$-submodule of $\mathbb{K}[[x]]$ (for each $n \in \mathbb{N}$).

(a) Let $f \in \mathbb{K}[x]$. Write the FPS f in the form $f = \sum_{k\in\mathbb{N}} a_k x^k$ with $a_0, a_1, a_2, \ldots \in \mathbb{K}$. Thus,

$$\left[x^i\right] f = a_i \qquad \text{for each } i \in \mathbb{N} \tag{26}$$

(by the definition of $[x^i] f$). Furthermore, the definition of f' yields $f' = \sum_{k>0} k a_k x^{k-1}$.

We WLOG assume that $f \neq 0$ (because otherwise, we have $f = 0$ and thus $f' = 0' = 0 \in \mathbb{K}[x]$ and $\deg \underbrace{(f')}_{=0} = \deg 0 = -\infty \leq \deg f - 1$).

Define an $m \in \mathbb{N}$ by $m = \deg f$. (This is well-defined, since $f \neq 0$.) From $m = \deg f$, we conclude that $[x^i] f = 0$ for all integers $i > m$ (by the definition of degree). Hence, each integer $i > m$ satisfies

$$\begin{aligned}a_i &= \left[x^i\right] f \qquad \text{(by (26))} \\ &= 0.\end{aligned} \tag{27}$$

Renaming i as k in this statement, we obtain that

$$\text{each integer } k > m \text{ satisfies } a_k = 0. \tag{28}$$

Now,

$$f' = \sum_{k>0} k a_k x^{k-1} = \sum_{k=1}^{m} k a_k x^{k-1} + \sum_{k=m+1}^{\infty} k \underbrace{a_k}_{\substack{=0 \\ \text{(by (28))}}} x^{k-1} = \sum_{k=1}^{m} k a_k x^{k-1} + \underbrace{\sum_{k=m+1}^{\infty} k 0 x^{k-1}}_{=0}$$

$$= \sum_{k=1}^{m} k a_k x^{k-1} = 1 a_1 x^{1-1} + 2 a_2 x^{2-1} + \cdots + m a_m x^{m-1}.$$

Hence, f' is a $\mathbb{K}$-linear combination of the m polynomials $x^{1-1}, x^{2-1}, \ldots, x^{m-1}$ (since the coefficients ka_k belong to $\mathbb{K}$). But these m polynomials $x^{1-1}, x^{2-1}, \ldots, x^{m-1}$ all belong to $\mathbb{K}[x]_{\leq m-1}$. Thus, their $\mathbb{K}$-linear combination f' also belongs to $\mathbb{K}[x]_{\leq m-1}$ (since $\mathbb{K}[x]_{\leq m-1}$ is a $\mathbb{K}$-module). In other words, $f' \in \mathbb{K}[x]_{\leq m-1}$. In other words, f' is a polynomial of degree $\leq m-1$. Hence, f' is a polynomial and satisfies $\deg(f') \leq \underbrace{m}_{=\deg f} - 1 = \deg f - 1$. This solves part **(a)** of the exercise.

[*Remark:* It is not guaranteed that $\deg(f') = \deg f - 1$. First of all, this equality is false when $\deg f = 0$ (because in this case, $f' = 0$ and thus $\deg f' = -\infty \neq 0 - 1$). But even when $\deg f$ is positive, the equality $\deg(f') = \deg f - 1$ may fail. For example, if $\mathbb{K} = \mathbb{Z}/2$ and $f = x^2 + x$, then $f' = \underbrace{2}_{=0 \text{ in } \mathbb{K}} x + 1 = 1$ has degree $\deg(f') = 0 < 2 - 1$.

It is easy to see that $\deg(f') = \deg f - 1$ holds when $\mathbb{K}$ is a field satisfying $(\deg f) \cdot 1_{\mathbb{K}} \neq 0$.]

(b) According to the definition of a $\mathbb{K}$-linear map (also known as a $\mathbb{K}$-module homomorphism), we must prove the following three statements:

Statement 1: We have $D(a+b) = D(a) + D(b)$ for all $a, b \in \mathbb{K}[[x]]$.

Statement 2: We have $D(0) = 0$.

Statement 3: We have $D(\lambda a) = \lambda D(a)$ for all $\lambda \in \mathbb{K}$ and $a \in \mathbb{K}[[x]]$.

We shall only prove the first of these three statements; the other two are similar.

[*Proof of Statement 1:* Let $a, b \in \mathbb{K}[[x]]$. We must prove that $D(a+b) = D(a) + D(b)$.

Write the FPS a in the form $a = \sum_{k \in \mathbb{N}} a_k x^k$ with $a_0, a_1, a_2, \ldots \in \mathbb{K}$. Then, the definition of a' yields $a' = \sum_{k>0} k a_k x^{k-1}$.

Write the FPS b in the form $b = \sum_{k \in \mathbb{N}} b_k x^k$ with $b_0, b_1, b_2, \ldots \in \mathbb{K}$. Then, the definition of b' yields $b' = \sum_{k>0} k b_k x^{k-1}$.

Now, adding the equalities $a = \sum_{k \in \mathbb{N}} a_k x^k$ and $b = \sum_{k \in \mathbb{N}} b_k x^k$ together, we obtain

$$a + b = \sum_{k \in \mathbb{N}} a_k x^k + \sum_{k \in \mathbb{N}} b_k x^k = \sum_{k \in \mathbb{N}} \underbrace{\left(a_k x^k + b_k x^k\right)}_{=(a_k+b_k)x^k} = \sum_{k \in \mathbb{N}} (a_k + b_k) x^k.$$

(and the coefficients $a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots$ appearing on the right hand side of this equality all belong to $\mathbb{K}$). Thus, the definition of $(a+b)'$ yields

$$(a+b)' = \sum_{k>0} \underbrace{k(a_k + b_k) x^{k-1}}_{=k a_k x^{k-1} + k b_k x^{k-1}} = \sum_{k>0} \left(k a_k x^{k-1} + k b_k x^{k-1}\right) = \sum_{k>0} k a_k x^{k-1} + \sum_{k>0} k b_k x^{k-1}.$$

Comparing this with

$$\underbrace{a'}_{=\sum_{k>0} k a_k x^{k-1}} + \underbrace{b'}_{=\sum_{k>0} k b_k x^{k-1}} = \sum_{k>0} k a_k x^{k-1} + \sum_{k>0} k b_k x^{k-1},$$

we obtain $(a+b)' = a' + b'$. But the definition of D yields $D(a) = a'$ and $D(b) = b'$ and $D(a+b) = (a+b)'$. In light of these three equalities, we can rewrite our result $(a+b)' = a' + b'$ as $D(a+b) = D(a) + D(b)$. This proves Statement 1.]

Thus, we have proven Statement 1. The proofs of Statement 2 and Statement 3 are similar (but easier). This completes our solution of part **(b)**.

Before we move on to the next parts of the exercise, let us prove several auxiliary statements. The first is a mere restatement of the definition of the derivative f' of an FPS f:

Statement 4: Let $f \in \mathbb{K}[[x]]$. Then,

$$[x^n](f') = (n+1)\left[x^{n+1}\right] f \qquad \text{for each } n \in \mathbb{N}.$$

[*Proof of Statement 4:* We have $f = \sum_{k \in \mathbb{N}} \left(\left[x^k\right] f\right) x^k$ (since the $\left[x^k\right] f$ are the coefficients of f). Hence, the definition of f' yields

$$f' = \sum_{k>0} k\left(\left[x^k\right] f\right) x^{k-1} = \sum_{k \in \mathbb{N}} (k+1)\left(\left[x^{k+1}\right] f\right) x^k$$

(here, we have substituted $k+1$ for k in the sum). Hence, $[x^n](f') = (n+1)\left([x^{n+1}] f\right)$ for each $n \in \mathbb{N}$ (by the definition of $[x^n](f')$). This proves Statement 4.]

Our next auxiliary statement is an extension of Statement 1 (which we proved above, in our solution to part **(b)**) to arbitrary (possibly infinite) sums of summable families:

Statement 5: Let $(f_i)_{i \in I}$ be a summable family of FPSs in $\mathbb{K}[[x]]$. Then, we have $D\left(\sum_{i \in I} f_i\right) = \sum_{i \in I} D(f_i)$. (In particular, the family $(D(f_i))_{i \in I}$ is summable.)

[*Proof of Statement 5 (sketched):* For each $i \in I$ and $n \in \mathbb{N}$, we have

$$[x^n] \underbrace{(D(f_i))}_{\substack{=(f_i)' \\ \text{(by the definition of } D\text{)}}} = [x^n]\left((f_i)'\right) = (n+1)\left(\left[x^{n+1}\right](f_i)\right) \tag{29}$$

(by Statement 4, applied to $f = f_i$).

We shall now prove that the family $(D(f_i))_{i \in I}$ is summable.

Indeed, the family $(f_i)_{i \in I}$ is summable. In other words, for each $n \in \mathbb{N}$, the following requirement holds:

$$\text{only finitely many } i \in I \text{ satisfy } [x^n](f_i) \neq 0. \tag{30}$$

Now, let $n \in \mathbb{N}$. Then, only finitely many $i \in I$ satisfy $[x^{n+1}](f_i) \neq 0$ (by (30), applied to $n+1$ instead of n). Thus, only finitely many $i \in I$ satisfy $(n+1)\left([x^{n+1}](f_i)\right) \neq 0$ (because $(n+1)\left([x^{n+1}](f_i)\right) \neq 0$ can only hold if $[x^{n+1}](f_i) \neq 0$). In view of (29), this rewrites as follows: Only finitely many $i \in I$ satisfy $[x^n](D(f_i)) \neq 0$.

Now, forget that we fixed n. We thus have shown that for each $n \in \mathbb{N}$, only finitely many $i \in I$ satisfy $[x^n](D(f_i)) \neq 0$. In other words, the family $(D(f_i))_{i \in I}$ is summable.

The sum of a summable family is defined entrywise. Thus, each $n \in \mathbb{N}$ satisfies

$$[x^n]\left(\sum_{i \in I} D(f_i)\right) = \sum_{i \in I} \underbrace{[x^n](D(f_i))}_{\substack{=(n+1)\left(\left[x^{n+1}\right](f_i)\right) \\ \text{(by (29))}}} = \sum_{i \in I} (n+1)\left(\left[x^{n+1}\right](f_i)\right). \tag{31}$$

Also, Statement 4 (applied to $f = \sum_{i \in I} f_i$) shows that each $n \in \mathbb{N}$ satisfies

$$[x^n]\left(\left(\sum_{i\in I} f_i\right)'\right) = (n+1) \underbrace{\left[x^{n+1}\right]\left(\sum_{i\in I} f_i\right)}_{=\sum_{i\in I}\left[x^{n+1}\right](f_i)} = (n+1)\sum_{i\in I}\left[x^{n+1}\right](f_i)$$

(since the sum of a summable family is defined entrywise)

$$= \sum_{i\in I} (n+1)\left(\left[x^{n+1}\right](f_i)\right).$$

Comparing this with (31), we obtain that

$$[x^n]\left(\left(\sum_{i\in I} f_i\right)'\right) = [x^n]\left(\sum_{i\in I} D(f_i)\right) \qquad \text{for each } n \in \mathbb{N}.$$

In other words, $\left(\sum_{i\in I} f_i\right)' = \sum_{i\in I} D(f_i)$. But the definition of D yields $D\left(\sum_{i\in I} f_i\right) = \left(\sum_{i\in I} f_i\right)' = \sum_{i\in I} D(f_i)$. This completes the proof of Statement 5.]

Our next auxiliary statement extends Statement 5 from sums of FPSs to $\mathbb{K}$-linear combinations of FPSs:

Statement 6: Let $(f_i)_{i\in I}$ be a summable family of FPSs in $\mathbb{K}[[x]]$. Let $(\lambda_i)_{i\in I} \in \mathbb{K}^I$ be any family of scalars. Then, we have $D\left(\sum_{i\in I} \lambda_i f_i\right) = \sum_{i\in I} \lambda_i D(f_i)$. (In particular, the families $(\lambda_i f_i)_{i\in I}$ and $(\lambda_i D(f_i))_{i\in I}$ are summable.)

[*Proof of Statement 6 (sketched):* The map D is $\mathbb{K}$-linear (by part **(b)** of this exercise). The family $(f_i)_{i\in I}$ is summable. Thus, it is easy to see that the family $(\lambda_i f_i)_{i\in I}$ is summable[14]. Hence, Statement 5 (applied to $\lambda_i f_i$ instead of f_i) yields that we have $D\left(\sum_{i\in I} \lambda_i f_i\right) = \sum_{i\in I} D(\lambda_i f_i)$, and the family $(D(\lambda_i f_i))_{i\in I}$ is summable. Since each $i \in I$ satisfies $D(\lambda_i f_i) = \lambda_i D(f_i)$ (because D is a $\mathbb{K}$-linear map), we can rewrite this as follows: We have $D\left(\sum_{i\in I} \lambda_i f_i\right) = \sum_{i\in I} \lambda_i D(f_i)$, and that the family $(\lambda_i D(f_i))_{i\in I}$ is summable. This completes the proof of Statement 6.]

Furthermore, let us generalize Statement 6 from D to all powers D^k of D:

Statement 7: Let $(f_i)_{i\in I}$ be a summable family of FPSs in $\mathbb{K}[[x]]$. Let $(\lambda_i)_{i\in I} \in \mathbb{K}^I$ be any family of scalars. Let $k \in \mathbb{N}$. Then, we have $D^k\left(\sum_{i\in I} \lambda_i f_i\right) = \sum_{i\in I} \lambda_i D^k(f_i)$. (In particular, the families $(\lambda_i f_i)_{i\in I}$ and $\left(\lambda_i D^k(f_i)\right)_{i\in I}$ are summable.)

[14] because if some $i \in I$ and $n \in \mathbb{N}$ satisfy $[x^n](\lambda_i f_i) \neq 0$, then they must also satisfy $[x^n](f_i) \neq 0$ (because $\lambda_i \cdot [x^n](f_i) = [x^n](\lambda_i f_i) \neq 0$)

[*Proof of Statement 7 (sketched):* Statement 7 is easy to prove by induction on k. The induction base (i.e., the case $k = 0$) is obvious, and the induction step uses Statement 6.]

Finally, let us describe the derivative of a monomial:

Statement 8: Let $m \in \mathbb{N}$. Then, $D\left(x^m\right) = mx^{m-1}$. (Here, the expression "mx^{m-1}" is to be understood as 0 when $m = 0$.)

[*Proof of Statement 8 (sketched):* We have

$$\left[x^n\right]\left(x^m\right) = \begin{cases} 1, & \text{if } n = m; \\ 0, & \text{if } n \neq m \end{cases} \qquad \text{for each } n \in \mathbb{N}. \tag{32}$$

Also, we have

$$\left[x^n\right]\left(mx^{m-1}\right) = \begin{cases} m, & \text{if } n = m-1; \\ 0, & \text{if } n \neq m-1 \end{cases} \qquad \text{for each } n \in \mathbb{N}. \tag{33}$$

(This is clearly true when $m > 0$, but also holds for $m = 0$ by our definition of mx^{m-1} in that case.)

Now, for each $n \in \mathbb{N}$, we have

$$\begin{aligned}
&\left[x^n\right]\left(\left(x^m\right)'\right) \\
&= (n+1) \underbrace{\left[x^{n+1}\right]\left(x^m\right)}_{\substack{= \begin{cases} 1, & \text{if } n+1 = m; \\ 0, & \text{if } n+1 \neq m \end{cases} \\ \text{(by (32), applied to } n+1 \text{ instead of } n\text{)}}} \qquad \text{(by Statement 4, applied to } f = x^m\text{)} \\
&= (n+1) \cdot \begin{cases} 1, & \text{if } n+1 = m; \\ 0, & \text{if } n+1 \neq m \end{cases} \\
&= \begin{cases} (n+1)\cdot 1, & \text{if } n+1 = m; \\ (n+1)\cdot 0, & \text{if } n+1 \neq m \end{cases} = \begin{cases} n+1, & \text{if } n+1 = m; \\ 0, & \text{if } n+1 \neq m \end{cases} \\
&= \begin{cases} m, & \text{if } n+1 = m; \\ 0, & \text{if } n+1 \neq m \end{cases} \qquad \text{(since we have } n+1 = m \text{ in the case when } n+1 = m\text{)} \\
&= \begin{cases} m, & \text{if } n = m-1; \\ 0, & \text{if } n \neq m-1 \end{cases} \\
&\qquad \left(\begin{array}{c} \text{since the condition "}n+1 = m\text{" is equivalent to "}n = m-1\text{",} \\ \text{and the condition "}n+1 \neq m\text{" is equivalent to "}n \neq m-1\text{"} \end{array} \right) \\
&= \left[x^n\right]\left(mx^{m-1}\right) \qquad \text{(by (33))}.
\end{aligned}$$

In other words, we have $\left(x^m\right)' = mx^{m-1}$. But the definition of D yields $D\left(x^m\right) = \left(x^m\right)' = mx^{m-1}$. This proves Statement 8.]

Let us now resume solving the exercise.

(c) Here is one possible way to solve this. (Another can be found in [Grinbe18, proof of Proposition 0.2 **(c)**].)

We shall follow the same convention that we used in Statement 8: Namely, the expression "mx^{m-1}" is to be understood as 0 when $m = 0$.

Let $f, g \in \mathbb{K}[[x]]$ be two FPSs.

Write the FPS f in the form $f = \sum_{k \in \mathbb{N}} a_k x^k$ with $a_0, a_1, a_2, \ldots \in \mathbb{K}$.

Write the FPS g in the form $g = \sum_{k \in \mathbb{N}} b_k x^k$ with $b_0, b_1, b_2, \ldots \in \mathbb{K}$.

Multiplying the equalities $f = \sum_{k \in \mathbb{N}} a_k x^k = \sum_{i \in \mathbb{N}} a_i x^i$ and $g = \sum_{k \in \mathbb{N}} b_k x^k = \sum_{j \in \mathbb{N}} b_j x^j$, we obtain

$$fg = \left(\sum_{i \in \mathbb{N}} a_i x^i \right) \left(\sum_{j \in \mathbb{N}} b_j x^j \right) = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_i \underbrace{x^i b_j x^j}_{= b_j x^{i+j}} = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_i b_j x^{i+j}. \tag{34}$$

Applying the map D to both sides of this equality, we obtain

$$D(fg) = D\left(\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_i b_j x^{i+j} \right) = \sum_{i \in \mathbb{N}} D\left(\sum_{j \in \mathbb{N}} a_i b_j x^{i+j} \right) \tag{35}$$

(by Statement 5, applied to the set $I = \mathbb{N}$ and the summable family $(f_i)_{i \in I} = \left(\sum_{j \in \mathbb{N}} a_i b_j x^{i+j} \right)_{i \in \mathbb{N}}$ of FPSs). But each $i \in \mathbb{N}$ satisfies

$$D\left(\sum_{j \in \mathbb{N}} a_i b_j x^{i+j} \right) = \sum_{j \in \mathbb{N}} a_i b_j D\left(x^{i+j} \right) \tag{36}$$

(by Statement 6, applied to the set $I = \mathbb{N}$, the family[15] $(\lambda_j)_{j \in I} = (a_i b_j)_{j \in \mathbb{N}}$ of scalars and the summable family $(f_j)_{j \in I} = (x^{i+j})_{j \in \mathbb{N}}$ of FPSs). Thus, (35) becomes

$$D(fg) = \sum_{i \in \mathbb{N}} \underbrace{D\left(\sum_{j \in \mathbb{N}} a_i b_j x^{i+j} \right)}_{= \sum_{j \in \mathbb{N}} a_i b_j D\left(x^{i+j} \right) \text{ (by (36))}} = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_i b_j \underbrace{D\left(x^{i+j} \right)}_{= (i+j) x^{i+j-1} \text{ (by Statement 8, applied to } m = i+j)} = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_i b_j (i+j) x^{i+j-1}.$$

Comparing this with

$$D(fg) = (fg)' \qquad \text{(by the definition of } D\text{)},$$

we obtain

$$(fg)' = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_i b_j (i+j) x^{i+j-1}. \tag{37}$$

On the other hand, from $f = \sum_{k \in \mathbb{N}} a_k x^k$, we obtain

$$f' = \sum_{k > 0} k a_k x^{k-1} \qquad \text{(by the definition of } f'\text{)}.$$

[15] We are using "j" instead of "i" for the index of our families now, because the letter "i" already means something else.

Comparing this with

$$\sum_{k\in\mathbb{N}} a_k k x^{k-1} = a_0 \underbrace{0x^{0-1}}_{\substack{=0 \\ \text{(since we understand ``}mx^{m-1}\text{'' to mean 0 when } m=0)}} + \sum_{k>0} \underbrace{a_k k}_{=ka_k} x^{k-1}$$

(here, we have split off the addend for $k=0$ from the sum)

$$= \underbrace{a_0 0}_{=0} + \sum_{k>0} k a_k x^{k-1} = \sum_{k>0} k a_k x^{k-1},$$

we obtain

$$f' = \sum_{k\in\mathbb{N}} a_k k x^{k-1}. \tag{38}$$

The same argument (applied to g and b_k instead of f and a_k) yields

$$g' = \sum_{k\in\mathbb{N}} b_k k x^{k-1}. \tag{39}$$

Multiplying the equalities $f' = \sum\limits_{k\in\mathbb{N}} a_k k x^{k-1} = \sum\limits_{i\in\mathbb{N}} a_i i x^{i-1}$ and $g = \sum\limits_{k\in\mathbb{N}} b_k x^k = \sum\limits_{j\in\mathbb{N}} b_j x^j$, we obtain

$$f'g = \left(\sum_{i\in\mathbb{N}} a_i i x^{i-1}\right)\left(\sum_{j\in\mathbb{N}} b_j x^j\right) = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i \underbrace{i x^{i-1} b_j x^j}_{=b_j i x^{i-1} x^j} = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j i \underbrace{x^{i-1}x^j}_{=x^{(i-1)+j}=x^{i+j-1}}$$

$$= \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j i x^{i+j-1}. \tag{40}$$

Multiplying the equalities $f = \sum\limits_{k\in\mathbb{N}} a_k x^k = \sum\limits_{i\in\mathbb{N}} a_i x^i$ and $g' = \sum\limits_{k\in\mathbb{N}} b_k k x^{k-1} = \sum\limits_{j\in\mathbb{N}} b_j j x^{j-1}$, we obtain

$$fg' = \left(\sum_{i\in\mathbb{N}} a_i x^i\right)\left(\sum_{j\in\mathbb{N}} b_j j x^{j-1}\right) = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i \underbrace{x^i b_j j x^{j-1}}_{=b_j j x^i x^{j-1}} = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j j \underbrace{x^i x^{j-1}}_{=x^{i+(j-1)}=x^{i+j-1}}$$

$$= \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j j x^{i+j-1}.$$

Adding this equality to the equality (40), we obtain

$$f'g + fg' = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j i x^{i+j-1} + \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j j x^{i+j-1} = \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} \underbrace{\left(a_i b_j i x^{i+j-1} + a_i b_j j x^{i+j-1}\right)}_{=a_i b_j (i+j) x^{i+j-1}}$$

$$= \sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}} a_i b_j \left(i+j\right) x^{i+j-1}.$$

Comparing this with (37), we find $(fg)' = f'g + fg'$. This solves part **(c)** of the exercise.

(d) We shall solve part **(d)** of the exercise by induction on n:
Induction base: We have $D^0 = \mathrm{id}$. Thus,

$$\underbrace{D^0}_{=\mathrm{id}}\left(x^k\right) = \mathrm{id}\left(x^k\right) = x^k = 0!\binom{k}{0}x^k \qquad \left(\text{since } \underbrace{0!}_{=1}\underbrace{\binom{k}{0}}_{=1}x^k = x^k\right)$$

for each $k \in \mathbb{N}$. In other words, part **(d)** of the exercise holds for $n = 0$. This completes the induction base.

Induction step: Let $i \in \mathbb{N}$. Assume that part **(d)** of the exercise holds for $n = i$. We must prove that part **(d)** of the exercise holds for $n = i + 1$.

We have assumed that part **(d)** of the exercise holds for $n = i$. In other words, we have

$$D^i\left(x^k\right) = i!\binom{k}{i}x^{k-i} \qquad \text{for all } k \in \mathbb{N}. \tag{41}$$

Now, let $k \in \mathbb{N}$. Then, $D^{i+1} = D \circ D^i$. Applying both sides of this equality to x^k, we obtain

$$D^{i+1}\left(x^k\right) = \left(D \circ D^i\right)\left(x^k\right) = D\left(\underbrace{D^i\left(x^k\right)}_{\substack{=i!\binom{k}{i}x^{k-i} \\ \text{(by (41))}}}\right) = D\left(i!\binom{k}{i}x^{k-i}\right). \tag{42}$$

We shall now show that

$$D^{i+1}\left(x^k\right) = (i+1)!\binom{k}{i+1}x^{k-(i+1)}. \tag{43}$$

Indeed, three cases are possible:

Case 1: We have $k < i$.

Case 2: We have $k = i$.

Case 3: We have $k > i$.

Let us first consider Case 1. In this case, we have $k < i$. Thus, $\binom{k}{i}x^{k-i} = 0$ (according to our convention that the expression "$\binom{k}{n}x^{k-n}$" is to be understood as 0 when $k < n$). For the same reason, we have $\binom{k}{i+1}x^{k-(i+1)} = 0$ (since $k < i < i+1$). Now, (42) becomes

$$D^{i+1}\left(x^k\right) = D\left(i!\underbrace{\binom{k}{i}x^{k-i}}_{=0}\right) = D\left(\underbrace{i!0}_{=0}\right) = D(0) = 0 \qquad \text{(since } D \text{ is } \mathbb{K}\text{-linear)}$$

$$= (i+1)!\binom{k}{i+1}x^{k-(i+1)} \qquad \left(\text{since } (i+1)!\underbrace{\binom{k}{i+1}x^{k-(i+1)}}_{=0} = 0\right).$$

Thus, (43) is proven in Case 1.

Let us next consider Case 2. In this case, we have $k = i$. Thus, $k = i < i + 1$, so that $\binom{k}{i+1}x^{k-(i+1)} = 0$ (according to our convention that the expression "$\binom{k}{n}x^{k-n}$" is to be understood as 0 when $k < n$). Also, from $k = i$, we obtain $\binom{k}{i} = \binom{i}{i} = 1$ and

$x^{k-i} = x^{i-i} = x^0$. Now, (42) becomes

$$\begin{aligned}
&D^{i+1}\left(x^k\right) \\
&= D\left(i!\underbrace{\binom{k}{i}}_{=1}\underbrace{x^{k-i}}_{=x^0}\right) = D\left(i!x^0\right) = i!\underbrace{D\left(x^0\right)}_{\substack{=0x^{0-1} \\ \text{(by Statement 8,} \\ \text{applied to } m=0)}} \qquad \text{(since } D \text{ is } \mathbb{K}\text{-linear)} \\
&= \underbrace{i!0x^{0-1}}_{=0} = 0 = (i+1)!\binom{k}{i+1}x^{k-(i+1)} \qquad \left(\text{since } (i+1)!\underbrace{\binom{k}{i+1}x^{k-(i+1)}}_{=0} = 0\right).
\end{aligned}$$

Thus, (43) is proven in Case 2.

Finally, let us consider Case 3. In this case, we have $k > i$. It is easy to see that

$$(k-i)\binom{k}{i} = (i+1)\binom{k}{i+1}. \tag{44}$$

(Indeed, both sides of (44) can be simplified to $\dfrac{k(k-1)(k-2)\cdots(k-i)}{i!}$ by applying the definition of binomial coefficients and the fact that $(i+1)! = (i+1)\cdot i!$. An alternative way to prove (44) is by expanding both binomial coefficients $\binom{k}{i}$ and $\binom{k}{i+1}$ using Exercise 3 **(a)** on homework set #0.)

Now, $k > i$, so that $k - i \in \mathbb{N}$. The equality (42) becomes

$$\begin{aligned}
&D^{i+1}\left(x^k\right) \\
&= D\left(i!\binom{k}{i}x^{k-i}\right) = i!\binom{k}{i}\underbrace{D\left(x^{k-i}\right)}_{\substack{=(k-i)x^{k-i-1} \\ \text{(by Statement 8,} \\ \text{applied to } m=k-i)}} \qquad \text{(since } D \text{ is } \mathbb{K}\text{-linear)} \\
&= i!\underbrace{\underbrace{\binom{k}{i}}_{=(k-i)\binom{k}{i}}(k-i)}_{=(i+1)\binom{k}{i+1} \text{ (by (44))}}\underbrace{x^{k-i-1}}_{=x^{k-(i+1)}} = \underbrace{i!\cdot(i+1)}_{=(i+1)!}\binom{k}{i+1}x^{k-(i+1)} = (i+1)!\binom{k}{i+1}x^{k-(i+1)}.
\end{aligned}$$

Thus, (43) is proven in Case 3.

We have now proven (43) in each of the three Cases 1, 2 and 3. Thus, (43) always holds.

Now, forget that we fixed k. Thus, we have proven that (43) holds for each $k \in \mathbb{N}$. In other words, part **(d)** of the exercise holds for $n = i+1$. This completes the induction step. Thus, part **(d)** of the exercise is proven by induction.

(e) Let us first show two auxiliary statements:

Statement 9: Let $m \in \mathbb{N}$. Let $f \in \mathbb{K}[x]$ be a polynomial of degree $\leq m$. Then, for each $n \in \mathbb{N}$, we have $D^n(f) \in \mathbb{K}[x]$ and $\deg\left(D^n(f)\right) \leq m - n$.

[*Proof of Statement 9 (sketched):* This can be straightforwardly proven by induction on n. The induction base (i.e., the case $n = 0$) is obvious. The induction step proceeds by observing that if n is a positive integer, then

$$\underbrace{D^n}_{=D\circ D^{n-1}}(f) = \left(D \circ D^{n-1}(f)\right) = D\left(D^{n-1}(f)\right) = \left(D^{n-1}(f)\right)' \qquad \text{(by the definition of } D\text{)},$$

and applying part **(a)** of the exercise to $D^{n-1}(f)$ instead of f. Thus, Statement 9 is proven.]

Statement 10: Let $m \in \mathbb{N}$. Let $f \in \mathbb{K}[x]$ be a polynomial of degree $\leq m$. Then, $D^n(f) = 0$ for all integers $n > m$.

[*Proof of Statement 10 (sketched):* Let n be an integer such that $n > m$. Then, Statement 9 yields $D^n(f) \in \mathbb{K}[x]$ and $\deg\left(D^n(f)\right) \leq m - n$. Hence, $\deg\left(D^n(f)\right) \leq m - n < 0$ (since $n > m$), so that $D^n(f) = 0$. This proves Statement 10.]

Now, assume that $\mathbb{Q}$ is a subring of $\mathbb{K}$. Let $f \in \mathbb{K}[x]$ be a polynomial. Let $m = \deg f$. Then, f has degree $\leq m$. Hence, Statement 10 shows that $D^n(f) = 0$ for all integers $n > m$. Thus, $\frac{1}{n!}\underbrace{\left(D^n(f)\right)}_{=0}[a] \cdot x^n = \frac{1}{n!}\underbrace{0}_{=0}[a] \cdot x^n = 0$ for all integers $n > m$. Hence, all but finitely many addends of the sum $\sum_{n\in\mathbb{N}} \frac{1}{n!}\left(D^n(f)\right)[a] \cdot x^n$ are 0. Therefore, this sum is well-defined.

In the following, we shall use the same convention as we did in part **(d)** of this exercise: Namely, the expression "$\binom{k}{n}x^{k-n}$" is to be understood as 0 when $k < n$. Thus,

$$\binom{k}{n}x^{k-n} = 0 \qquad \text{for all } k \in \mathbb{N} \text{ and } n \in \mathbb{N} \text{ satisfying } k < n. \tag{45}$$

Write the polynomial f in the form $f = \sum_{k=0}^{m} b_k x^k$ with $b_0, b_1, \ldots, b_m \in \mathbb{K}$. (We can do this, since $\deg f = m$.) Then, for each $n \in \mathbb{N}$, we have

$$\begin{aligned}
D^n(f) &= D^n\left(\sum_{k=0}^{m} b_k x^k\right) \qquad \left(\text{since } f = \sum_{k=0}^{m} b_k x^k\right) \\
&= \underbrace{\sum_{k=0}^{m}}_{=\sum_{k\in\{0,1,\ldots,m\}}} b_k \underbrace{D^n\left(x^k\right)}_{\substack{=n!\binom{k}{n}x^{k-n} \\ \text{(by part (d)} \\ \text{of this exercise)}}} \\
&\qquad \text{(since the map } D^n \text{ is } \mathbb{K}\text{-linear (because the map } D \text{ is } \mathbb{K}\text{-linear))} \\
&= \sum_{k\in\{0,1,\ldots,m\}} b_k n! \binom{k}{n} x^{k-n} = \sum_{\substack{k\in\{0,1,\ldots,m\}; \\ k<n}} b_k n! \underbrace{\binom{k}{n}x^{k-n}}_{\substack{=0 \\ \text{(by (45))}}} + \underbrace{\sum_{\substack{k\in\{0,1,\ldots,m\}; \\ k\geq n}} b_k n! \binom{k}{n} x^{k-n}}_{=\sum_{k=n}^{m}} \\
&\qquad \left(\begin{array}{c}\text{since each } k \in \{0,1,\ldots,m\} \text{ satisfies either } k < n \text{ or } k \geq n \\ \text{(but never both at the same time)}\end{array}\right) \\
&= \underbrace{\sum_{\substack{k\in\{0,1,\ldots,m\}; \\ k<n}} b_k n! 0}_{=0} + \sum_{k=n}^{m} b_k n! \binom{k}{n} x^{k-n} = \sum_{k=n}^{m} b_k n! \binom{k}{n} x^{k-n}.
\end{aligned} \tag{46}$$

Let $a \in \mathbb{K}$. Then, for each $n \in \mathbb{N}$, we have

$$\begin{aligned}\left(D^n\left(f\right)\right)[a] &= \left(\sum_{k=n}^{m} b_k n! \binom{k}{n} x^{k-n}\right)[a] \\ &\qquad \text{(here, we have substituted } a \text{ for } x \text{ in the equality (46))} \\ &= \sum_{k=n}^{m} b_k n! \binom{k}{n} a^{k-n}\end{aligned}$$

and thus

$$\frac{1}{n!} \underbrace{\left(D^n\left(f\right)\right)[a]}_{=\sum\limits_{k=n}^{m} b_k n! \binom{k}{n} a^{k-n}} = \frac{1}{n!} \sum_{k=n}^{m} b_k n! \binom{k}{n} a^{k-n} = \sum_{k=n}^{m} b_k \binom{k}{n} a^{k-n}. \tag{47}$$

But substituting $x + a$ for x in the equality $f = \sum\limits_{k=0}^{m} b_k x^k$, we find

$$f[x+a] = \left(\sum_{k=0}^{m} b_k x^k\right)[x+a] = \sum_{k=0}^{m} b_k \underbrace{(x+a)^k}_{\substack{=\sum\limits_{n=0}^{k} \binom{k}{n} x^n a^{k-n} \\ \text{(by the binomial formula,} \\ \text{since } xa = ax)}} = \sum_{k=0}^{m} b_k \sum_{n=0}^{k} \binom{k}{n} x^n a^{k-n}$$

$$= \underbrace{\sum_{k=0}^{m} \sum_{n=0}^{k}}_{=\sum\limits_{n\in\mathbb{N}} \sum\limits_{k=n}^{m}} b_k \binom{k}{n} \underbrace{x^n a^{k-n}}_{=a^{k-n}\cdot x^n} = \sum_{n\in\mathbb{N}} \underbrace{\sum_{k=n}^{m} b_k \binom{k}{n} a^{k-n}}_{\substack{=\frac{1}{n!}\left(D^n(f)\right)[a] \\ \text{(by (47))}}} \cdot x^n = \sum_{n\in\mathbb{N}} \frac{1}{n!} \left(D^n\left(f\right)\right)[a] \cdot x^n.$$

This solves part **(e)** of the exercise.

(f) Let p be a prime such that $p \cdot 1_{\mathbb{K}} = 0$. Let $f \in \mathbb{K}[[x]]$.

Note that p is a positive integer (since p is a prime); thus, $p! = p \cdot (p-1)! = (p-1)! \cdot p$. Hence,

$$\underbrace{p!}_{=(p-1)!\cdot p} \cdot 1_{\mathbb{K}} = (p-1)! \cdot \underbrace{p \cdot 1_{\mathbb{K}}}_{=0} = (p-1)! \cdot 0 = 0.$$

Thus, each $u \in \mathbb{K}[[x]]$ satisfies

$$p! \underbrace{u}_{\substack{=1_{\mathbb{K}}\cdot u \\ \text{(since } \mathbb{K}[[x]] \text{ is a } \mathbb{K}\text{-module)}}} = \underbrace{p! \cdot 1_{\mathbb{K}}}_{=0} \cdot u = 0u = 0. \tag{48}$$

Write the FPS f in the form $f = \sum\limits_{i\in\mathbb{N}} a_i x^i$ with $a_0, a_1, a_2, \ldots \in \mathbb{K}$. Then, the family $\left(x^i\right)_{i\in\mathbb{N}}$ of FPSs is summable. Hence, Statement 7 (applied to $k = p$, $I = \mathbb{N}$, $\left(\lambda_i\right)_{i\in I} = \left(a_i\right)_{i\in\mathbb{N}}$ and $\left(f_i\right)_{i\in I} = \left(x^i\right)_{i\in\mathbb{N}}$) yields that $D^p\left(\sum\limits_{i\in\mathbb{N}} a_i x^i\right) = \sum\limits_{i\in\mathbb{N}} a_i D^p\left(x^i\right)$. In view of $f = \sum\limits_{i\in\mathbb{N}} a_i x^i$, this rewrites as

$$D^p\left(f\right) = \sum_{i\in\mathbb{N}} a_i D^p\left(x^i\right). \tag{49}$$

But let $i \in \mathbb{N}$. Then, part **(d)** of this exercise (applied to $k = i$ and $n = p$) yields

$$D^p\left(x^i\right) = p!\binom{i}{p}x^{i-p} = 0 \qquad \left(\text{by (48), applied to } u = \binom{i}{p}x^{i-p}\right). \tag{50}$$

Forget that we fixed i. We thus have proven the equality (50) for each $i \in \mathbb{N}$. Thus, (49) becomes

$$D^p(f) = \sum_{i\in\mathbb{N}} a_i \underbrace{D^p\left(x^i\right)}_{\substack{=0 \\ \text{(by (50))}}} = \sum_{i\in\mathbb{N}} a_i 0 = 0.$$

This solves part **(f)** of the exercise.

(g) This can be solved in the same way as we solved part **(b)** of the exercise (since the definition of J is similar to the definition of D).

(h) Let $f \in \mathbb{K}[[x]]$. We shall prove that $(D \circ J)(f) = f$.

Write the FPS f in the form $f = \sum_{k\in\mathbb{N}} a_k x^k$ with $a_0, a_1, a_2, \ldots \in \mathbb{K}$. Thus, the definition of $\int f$ yields

$$\begin{aligned}\int f &= \sum_{k\geq 0} \frac{1}{k+1} a_k x^{k+1} \\ &= \sum_{i\in\{1,2,3,\ldots\}} \frac{1}{i} a_{i-1} x^i \qquad \text{(here, we have substituted } i-1 \text{ for } k \text{ in the sum)}.\end{aligned}$$

Applying the map D to both sides of this equality, we find

$$\begin{aligned}D\left(\int f\right) &= D\left(\sum_{i\in\{1,2,3,\ldots\}} \frac{1}{i} a_{i-1} x^i\right) = \sum_{i\in\{1,2,3,\ldots\}} \frac{1}{i} a_{i-1} \underbrace{D\left(x^i\right)}_{\substack{=ix^{i-1} \\ \text{(by Statement 8, applied to } m=i\text{)}}} \\ &\qquad \left(\begin{array}{c}\text{by Statement 6, applied to the set } I = \{1,2,3,\ldots\}, \\ \text{the family } (\lambda_i)_{i\in I} = \left(\frac{1}{i}a_{i-1}\right)_{i\in\{1,2,3,\ldots\}} \text{ of scalars,} \\ \text{and the summable family } (f_i)_{i\in I} = \left(x^i\right)_{i\in\{1,2,3,\ldots\}} \text{ of FPSs}\end{array}\right) \\ &= \sum_{i\in\{1,2,3,\ldots\}} \underbrace{\frac{1}{i} a_{i-1} i x^{i-1}}_{=a_{i-1}x^{i-1}} = \sum_{i\in\{1,2,3,\ldots\}} a_{i-1} x^{i-1} = \sum_{k\in\mathbb{N}} a_k x^k\end{aligned}$$

(here, we have substituted k for $i-1$ in the sum). But the definition of J yields $J(f) = \int f$. Thus,

$$(D \circ J)(f) = D\left(\underbrace{J(f)}_{=\int f}\right) = D\left(\int f\right) = \sum_{k\in\mathbb{N}} a_k x^k = f = \mathrm{id}(f).$$

Now, forget that we fixed f. We thus have shown that $(D \circ J)(f) = \mathrm{id}(f)$ for each $f \in \mathbb{K}[[x]]$. In other words, $D \circ J = \mathrm{id}$. This solves part **(h)** of the exercise.

(i) It is easy to check that $(J \circ D)(1) = 0 \neq 1 = \mathrm{id}(1)$, and thus $J \circ D \neq \mathrm{id}$. This solves part **(i)** of the exercise.

5.4 Remark

As we have mentioned above, f needs to be a polynomial in part **(e)** of this exercise in order for $f[x+a]$ to be well-defined. But in the particular case when $a=0$, the evaluation $f[x+a]$ is well-defined for all $f \in \mathbb{K}[[x]]$ (indeed, in this case, we have $f[x+a] = f[x+0] = f[x] = f$). Thus, it is reasonable to ask whether the claim of part **(e)** holds for all FPSs $f \in \mathbb{K}[[x]]$ (rather than just for polynomials) when $a=0$. The answer is "yes": We have

$$f = \sum_{n \in \mathbb{N}} \frac{1}{n!} \left(D^n(f)\right)[0] \cdot x^n \qquad \text{for all } f \in \mathbb{K}[[x]]$$

(when $\mathbb{Q}$ is a subring of $\mathbb{K}$). This is an algebraic counterpart of the Maclaurin series; its proof is left to the reader (who can also look it up in [Loehr11, Theorem 7.55]).

The number p does not have to be a prime in part **(f)** of the exercise. It perfectly suffices that p is a positive integer. (The solution we gave above works perfectly in this generality.)

Part **(i)** of the exercise is a "near-miss": $J \circ D$ is not too far away from id. Indeed, every $f \in \mathbb{K}[[x]]$ satisfies

$$(J \circ D)(f) = f - a_0, \qquad \text{where } a_0 = \left[x^0\right] f.$$

This is the algebraic version of the second part of the Fundamental Theorem of Calculus. We leave the easy proof to the reader.

Most textbooks that are serious about introducing FPSs and their derivatives prove some parts of this exercise in some form (or leave the proofs to the reader). For example, [Loehr11, Theorem 7.54 **(a)** and **(b)**] is part **(b)** of the above exercise; [Loehr11, Theorem 7.54 **(d)**] is Statement 8 in our solution above; [Loehr11, Theorem 7.54 **(e)**] is part **(c)** of the above exercise; [Loehr11, Theorem 7.54 **(g)**] is a particular case of Statement 5 in our solution above.

---

6 Exercise 6: Formal difference calculus and integer-valued polynomials

6.1 Problem

Let $\mathbb{K}$ be a commutative ring.

For any polynomial $f \in \mathbb{K}[x]$, we define the *first finite difference* f^Δ of f to be the polynomial

$$f[x+1] - f[x] \in \mathbb{K}[x].$$

(This is a "discrete analogue" of the derivative, in case the analysis-free derivative from Exercise 5 was not discrete enough for you. It cannot be extended to FPSs, however, since you cannot substitute $x+1$ for x in an FPS.)

Let $\Delta : \mathbb{K}[x] \to \mathbb{K}[x]$ be the map sending each polynomial f to f^Δ. As usual, for any $n \in \mathbb{N}$, we let Δ^n denote $\underbrace{\Delta \circ \Delta \circ \cdots \circ \Delta}_{n \text{ times}}$ (which means id if $n=0$).

Prove the following:

(a) The map $\Delta : \mathbb{K}[x] \to \mathbb{K}[x]$ is $\mathbb{K}$-linear (with respect to the $\mathbb{K}$-module structure on $\mathbb{K}[x]$ defined in class – i.e., both addition and scaling of polynomials are defined entrywise).

(b) We have $(fg)^{\Delta} = f^{\Delta} g + f[x+1] g^{\Delta}$ for any two polynomials f and g.

Now, assume that $\mathbb{Q}$ is a subring of $\mathbb{K}$.
For any $n \in \mathbb{N}$, we define a polynomial[16]

$$\binom{x}{n} := \frac{x(x-1)(x-2)\cdots(x-n+1)}{n!} \in \mathbb{K}[x].$$

We also set $\binom{x}{n} := 0$ for every negative n.
Prove the following:

(c) We have $\Delta^n \left(\binom{x}{k} \right) = \binom{x}{k-n}$ for all $n \in \mathbb{N}$ and $k \in \mathbb{Z}$.

(d) If $m \in \mathbb{N}$, and if $f \in \mathbb{K}[x]$ is a polynomial of degree $\leq m$, then there exist elements $a_0, a_1, \ldots, a_m \in \mathbb{K}$ such that $f = \sum_{i=0}^{m} a_i \binom{x}{i}$.

(e) Every polynomial $f \in \mathbb{K}[x]$ satisfies

$$f[x+a] = \sum_{n \in \mathbb{N}} \left(\Delta^n (f) \right)[a] \cdot \binom{x}{n} \qquad \text{for all } a \in \mathbb{K}.$$

(The infinite sum on the right hand side has only finitely many nonzero addends.)

(f) Let $m \in \mathbb{N}$, and let $f \in \mathbb{K}[x]$ be a polynomial of degree $\leq m$. Assume that $f[k] \in \mathbb{Z}$ for each $k \in \{0, 1, \ldots, m\}$. Then, there exist integers $a_0, a_1, \ldots, a_m$ such that $f = \sum_{i=0}^{m} a_i \binom{x}{i}$.

[**Hint:** Part **(d)** is easiest to prove by induction on m. You can then prove part **(e)** for $f = \binom{x}{i}$ first (where $i \in \mathbb{N}$), and then extend it to arbitrary f by means of part **(d)**. Part **(f)**, in turn, can be derived from part **(e)** through a strategic choice of a.]

6.2 Remark

Just as Δ is an analogue of the differentiation operator D from Exercise 5, we can define an analogue of the integration operator J from Exercise 5. This will be a $\mathbb{K}$-linear map $\Sigma : \mathbb{K}[x] \to \mathbb{K}[x]$ that sends each polynomial $\sum_{i=0}^{m} a_i \binom{x}{i}$ to $\sum_{i=0}^{m} a_i \binom{x}{i+1}$ (this definition makes sense, because part **(d)** of this exercise shows that each polynomial can be written in the form $\sum_{i=0}^{m} a_i \binom{x}{i}$, uniquely except for "leading zeroes"). Again, we have $\Delta \circ \Sigma = \mathrm{id}$ but $\Sigma \circ \Delta \neq \mathrm{id}$.

[16] Note that we are within our rights to divide by $n!$ here, since $\mathbb{Q}$ is a subring of $\mathbb{K}$.

Moreover, if $f \in \mathbb{K}[x]$ is a polynomial, then $(\Sigma(f))[0] = 0$ and $(\Sigma(f))[n] = (\Sigma(f))[n-1] + f[n-1]$ for each $n \in \mathbb{Z}$ (indeed, the former equality follows easily from the definition of Σ, while the latter follows from $\Delta \circ \Sigma = \mathrm{id}$). Hence, by induction, we can see that

$$(\Sigma(f))[n] = f[0] + f[1] + \cdots + f[n-1] \qquad \text{for each polynomial } f \in \mathbb{K}[x] \text{ and each } n \in \mathbb{N}.$$

In other words, the value of $\Sigma(f)$ at an $n \in \mathbb{N}$ is the sum of the first n values of f on nonnegative integers! (Whence the notation Σ.) For example, if we set $f = x^2$, then it is easy to see that $\Sigma(f) = 2\binom{x}{3} + \binom{x}{2}$ (to see this, just expand f in the form $\sum_{i=0}^{m} a_i \binom{x}{i}$ – namely, $f = x^2 = 2\binom{x}{2} + \binom{x}{1}$ –, and then apply the definition of Σ); thus we obtain

$$2\binom{n}{3} + \binom{n}{2} = 0^2 + 1^2 + \cdots + (n-1)^2.$$

Similarly you can find a formula for $0^k + 1^k + \cdots + (n-1)^k$ whenever $k \in \mathbb{N}$.

Part **(e)** of this exercise is a result of Newton.

6.3 Solution sketch

We recall the following notation (which we introduced in the class notes): For each $n \in \mathbb{Z}$, we define a subset $\mathbb{K}[x]_{\leq n}$ of $\mathbb{K}[[x]]$ by

$$\begin{aligned}\mathbb{K}[x]_{\leq n} &= \{(a_0, a_1, a_2, \ldots) \in \mathbb{K}[[x]] \mid a_k = 0 \text{ for all } k > n\} \\ &= \left\{\mathbf{a} \in \mathbb{K}[[x]] \mid \left[x^k\right]\mathbf{a} = 0 \text{ for all } k > n\right\}.\end{aligned}$$

We know that a FPS $\mathbf{a} \in \mathbb{K}[[x]]$ belongs to $\mathbb{K}[x]_{\leq n}$ (for a given $n \in \mathbb{Z}$) if and only if $\mathbf{a}$ is a polynomial of degree $\leq n$. We also know that $\mathbb{K}[x]_{\leq n}$ is a $\mathbb{K}$-submodule of $\mathbb{K}[[x]]$ (for each $n \in \mathbb{N}$).

(a) According to the definition of a $\mathbb{K}$-linear map (also known as a $\mathbb{K}$-module homomorphism), we must prove the following three statements:

Statement 1: We have $\Delta(a+b) = \Delta(a) + \Delta(b)$ for all $a, b \in \mathbb{K}[x]$.

Statement 2: We have $\Delta(0) = 0$.

Statement 3: We have $\Delta(\lambda a) = \lambda \Delta(a)$ for all $\lambda \in \mathbb{K}$ and $a \in \mathbb{K}[x]$.

We shall only prove the first of these three statements; the other two are similar.

[*Proof of Statement 1:* Let $a, b \in \mathbb{K}[x]$. We must prove that $\Delta(a+b) = \Delta(a) + \Delta(b)$.

The definition of Δ yields

$$\Delta(a+b) = (a+b)^{\Delta} = \underbrace{(a+b)[x+1]}_{\substack{=a[x+1]+b[x+1] \\ \text{(by one of the basic} \\ \text{properties of evaluation)}}} - \underbrace{(a+b)[x]}_{\substack{=a[x]+b[x] \\ \text{(by one of the basic} \\ \text{properties of evaluation)}}}$$

$$\left(\text{by the definition of } (a+b)^{\Delta}\right)$$

$$= (a[x+1] + b[x+1]) - (a[x] + b[x]) = (a[x+1] - a[x]) + (b[x+1] - b[x]).$$

Comparing this with

$$\underbrace{\Delta(a)}_{\substack{=a^{\Delta} \\ \text{(by the definition of } \Delta\text{)}}} + \underbrace{\Delta(b)}_{\substack{=b^{\Delta} \\ \text{(by the definition of } \Delta\text{)}}} = \underbrace{a^{\Delta}}_{\substack{=a[x+1]-a[x] \\ \text{(by the definition of } a^{\Delta}\text{)}}} + \underbrace{b^{\Delta}}_{\substack{=b[x+1]-b[x] \\ \text{(by the definition of } b^{\Delta}\text{)}}}$$
$$= (a[x+1]-a[x]) + (b[x+1]-b[x]),$$

we obtain $\Delta(a+b) = \Delta(a) + \Delta(b)$. This proves Statement 1.]

Thus, we have proven Statement 1. The proofs of Statement 2 and Statement 3 are similar (but easier). This completes our solution of part **(a)** of the exercise.

(b) Let f and g be two polynomials. Then,

$$\underbrace{f^{\Delta}}_{\substack{=f[x+1]-f[x] \\ \text{(by the definition of } f^{\Delta}\text{)}}} \underbrace{g}_{\substack{=g[x] \\ \text{(since } g[x]=g\text{)}}} + f[x+1] \underbrace{g^{\Delta}}_{\substack{=g[x+1]-g[x] \\ \text{(by the definition of } g^{\Delta}\text{)}}}$$
$$= (f[x+1]-f[x])\, g[x] + f[x+1]\,(g[x+1]-g[x])$$
$$= f[x+1]\, g[x] - f[x]\, g[x] + f[x+1]\, g[x+1] - f[x+1]\, g[x]$$
$$= f[x+1]\, g[x+1] - f[x]\, g[x].$$

Comparing this with

$$(fg)^{\Delta} = \underbrace{(fg)[x+1]}_{\substack{=f[x+1]g[x+1] \\ \text{(by one of the basic} \\ \text{properties of evaluation)}}} - \underbrace{(fg)[x]}_{\substack{=f[x]g[x] \\ \text{(by one of the basic} \\ \text{properties of evaluation)}}} \qquad \left(\text{by the definition of } (fg)^{\Delta}\right)$$
$$= f[x+1]\, g[x+1] - f[x]\, g[x],$$

we obtain $(fg)^{\Delta} = f^{\Delta} g + f[x+1]\, g^{\Delta}$. Thus, part **(b)** of the exercise is solved.

Before we solve the rest of the exercise, let us lay some more groundwork. First, let us show a simple property of the Δ operator:

Statement 4: Let $f \in \mathbb{K}[x]$. Then, $\deg(\Delta(f)) \leq \deg f - 1$.

[*Proof of Statement 4:* This is obvious in the case when $f = 0$. Thus, we WLOG assume that $f \neq 0$. Hence, $\deg f \in \mathbb{N}$. Define $m \in \mathbb{N}$ by $m = \deg f$.

Write the polynomial f in the form $f = \sum_{k=0}^{m} b_k x^k$ with $b_0, b_1, \ldots, b_m \in \mathbb{K}$. (We can do this, since $\deg f = m$.) Thus,

$$f[x] = f = \sum_{k=0}^{m} b_k x^k = \sum_{i=0}^{m} b_i x^i. \tag{51}$$

Substituting $x+1$ for x in the equality $f = \sum_{k=0}^{m} b_k x^k$, we obtain

$$f[x+1] = \left(\sum_{k=0}^{m} b_k x^k\right)[x+1] = \sum_{k=0}^{m} b_k \underbrace{(x+1)^k}_{=\sum_{i=0}^{k}\binom{k}{i}x^i 1^{k-i}} = \sum_{k=0}^{m} b_k \sum_{i=0}^{k}\binom{k}{i}x^i \underbrace{1^{k-i}}_{=1}$$

(by the binomial formula)

$$= \sum_{k=0}^{m} b_k \sum_{i=0}^{k}\binom{k}{i}x^i = \underbrace{\sum_{k=0}^{m}\sum_{i=0}^{k}}_{=\sum_{i=0}^{m}\sum_{k=i}^{m}}\binom{k}{i}b_k x^i = \sum_{i=0}^{m} \underbrace{\sum_{k=i}^{m}\binom{k}{i}b_k x^i}_{=\binom{i}{i}b_i x^i + \sum_{k=i+1}^{m}\binom{k}{i}b_k x^i}$$

(here, we have split off the addend for $k=i$ from the sum)

$$= \sum_{i=0}^{m}\left(\binom{i}{i}b_i x^i + \sum_{k=i+1}^{m}\binom{k}{i}b_k x^i\right) = \sum_{i=0}^{m}\underbrace{\binom{i}{i}}_{=1}b_i x^i + \sum_{i=0}^{m}\sum_{k=i+1}^{m}\binom{k}{i}b_k x^i$$

$$= \underbrace{\sum_{i=0}^{m} b_i x^i}_{=f[x] \text{ (by (51))}} + \sum_{i=0}^{m}\sum_{k=i+1}^{m}\binom{k}{i}b_k x^i = f[x] + \sum_{i=0}^{m}\sum_{k=i+1}^{m}\binom{k}{i}b_k x^i.$$

Subtracting $f[x]$ from both sides of this equality, we obtain

$$f[x+1] - f[x] = \sum_{i=0}^{m}\sum_{k=i+1}^{m}\binom{k}{i}b_k x^i. \tag{52}$$

But the definition of Δ yields

$$\begin{aligned}\Delta(f) = f^{\Delta} &= f[x+1] - f[x] \qquad \left(\text{by the definition of } f^{\Delta}\right) \\ &= \sum_{i=0}^{m}\sum_{k=i+1}^{m}\binom{k}{i}b_k x^i \qquad (\text{by (52)}).\end{aligned} \tag{53}$$

But if $i \in \{0,1,\ldots,m\}$ and $k \in \{i+1, i+2, \ldots, m\}$, then $x^i \in \mathbb{K}[x]_{\leq m-1}$ [17]. Hence, the sum $\sum_{i=0}^{m}\sum_{k=i+1}^{m}\binom{k}{i}b_k x^i$ is a $\mathbb{K}$-linear combination of elements of $\mathbb{K}[x]_{\leq m-1}$ (since the coefficients $\binom{k}{i}b_k$ belong to $\mathbb{K}$), and thus itself lies in $\mathbb{K}[x]_{\leq m-1}$ (since $\mathbb{K}[x]_{\leq m-1}$ is a $\mathbb{K}$-module). In view of (53), this rewrites as follows: The polynomial $\Delta(f)$ lies in $\mathbb{K}[x]_{\leq m-1}$. In other words, the polynomial $\Delta(f)$ has degree $\leq m-1$. Hence, $\deg(\Delta(f)) \leq \underbrace{m}_{=\deg f} - 1 = \deg f - 1$. This proves Statement 4.]

From Statement 4, we can easily derive the following consequences:

Statement 5: Let $m \in \mathbb{N}$. Let $f \in \mathbb{K}[x]$ be a polynomial of degree $\leq m$. Then, for each $n \in \mathbb{N}$, we have $\deg(\Delta^n(f)) \leq m - n$.

[17] *Proof.* Let $i \in \{0,1,\ldots,m\}$ and $k \in \{i+1, i+2, \ldots, m\}$. Then, from $k \in \{i+1, i+2, \ldots, m\}$, we obtain $i+1 \leq k \leq m$, so that $i \leq m-1$ and therefore $x^i \in \mathbb{K}[x]_{\leq m-1}$. Qed.

[*Proof of Statement 5 (sketched):* This can be straightforwardly proven by induction on n. The induction base (i.e., the case $n = 0$) is obvious. The induction step proceeds by observing that if n is a positive integer, then

$$\underbrace{\Delta^n}_{=\Delta\circ\Delta^{n-1}}(f) = \left(\Delta \circ \Delta^{n-1}(f)\right) = \Delta\left(\Delta^{n-1}(f)\right),$$

and applying Statement 4 to $\Delta^{n-1}(f)$ instead of f. Thus, Statement 5 is proven.]

Statement 6: Let $m \in \mathbb{N}$. Let $f \in \mathbb{K}[x]$ be a polynomial of degree $\leq m$. Then, $\Delta^n(f) = 0$ for all integers $n > m$.

[*Proof of Statement 6 (sketched):* Let n be an integer such that $n > m$. Then, Statement 5 yields $\deg\left(\Delta^n(f)\right) \leq m - n$. Hence, $\deg\left(\Delta^n(f)\right) \leq m - n < 0$ (since $n > m$), so that $\Delta^n(f) = 0$. This proves Statement 6.]

From now on, we assume that $\mathbb{Q}$ is a subring of $\mathbb{K}$. Let us generalize our definition of the binomial coefficients $\binom{n}{k}$ to the case when n is not a rational number but a polynomial over $\mathbb{K}$ (that is, an element of $\mathbb{K}[x]$). This generalization goes as follows:

Definition 6.1. Let $n \in \mathbb{K}[x]$ and $k \in \mathbb{Q}$. Then, we define the *binomial coefficient* $\binom{n}{k}$ as follows:

(a) If $k \in \mathbb{N}$, then we set

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \frac{\prod_{i=0}^{k-1}(n-i)}{k!}.$$

(b) If $k \notin \mathbb{N}$, then we set $\binom{n}{k} = 0$.

This definition generalizes both

- the definition of $\binom{n}{k}$ in the case when $n \in \mathbb{Q}$ that we gave in class (Definition 2.17.1 in the class notes), and
- the definition of $\binom{x}{n}$ given in the above exercise (indeed, this is the particular case of Definition 6.1 when n and k are set to be x and n).

Now, we state a simple property of these generalized binomial coefficients:

Proposition 6.2. *Any $n \in \mathbb{K}[x]$ and $k \in \mathbb{Q}$ satisfy*

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Proof of Proposition 6.2. Proposition 6.2 can be proven using the same argument that was used to prove Theorem 2.17.8 in the class notes. (The only difference is that n is now an element of $\mathbb{K}[x]$ and not of $\mathbb{Q}$.) $\square$

Another property of these generalized binomial coefficients is a generalization of the Vandermonde convolution identity:

Proposition 6.3. *Let $a, b \in \mathbb{K}[x]$ and $n \in \mathbb{N}$. Then,*

$$\binom{a+b}{n} = \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k}.$$

Proposition 6.3 generalizes the well-known Vandermonde convolution identity (Theorem 2.17.14 in the class notes); in fact, the latter is obtained if we require a and b to be rational numbers (rather than polynomials in $\mathbb{K}[x]$).

Proof of Proposition 6.3. This is a bit tricky to derive from what we have done in class. The conceptually easiest proof is probably to say "read [Grinbe15, First proof of Theorem 3.29], and replace each appearance of "$\mathbb{Q}$" by "$\mathbb{K}[x]$", each appearance of "x" by "a", and each appearance of "y" by "b"". (The latter proof depends on a few results proven in [Grinbe15, §3.1], which too need to be generalized to elements of $\mathbb{K}[x]$ instead of $\mathbb{Q}$. But all this generalizing is straightforward.)

Unfortunately, the proof of the Vandermonde convolution identity that we gave in class cannot be generalized this easily. Indeed, $\mathbb{K}[x]$ is not necessarily a field, so we cannot directly apply the "polynomial identity trick". The easiest way out is by using polynomials in two indeterminates (even though I did not define them in class) and the corresponding analogue of the "polynomial identity trick". Here is this argument in a nutshell: Consider the two polynomials

$$P = \binom{u+v}{n} \qquad \text{and} \qquad Q = \sum_{k=0}^{n} \binom{u}{k}\binom{v}{n-k}$$

in two indeterminates u and v over the ring $\mathbb{Q}$ (not over the ring $\mathbb{K}$). Now, the Vandermonde convolution identity (Theorem 2.17.14 in the class notes) shows that $P[\alpha, \beta] = Q[\alpha, \beta]$ for all $\alpha, \beta \in \mathbb{N}$ (where "$P[\alpha, \beta]$" means the result of substituting $u = \alpha$ and $v = \beta$ in the polynomial P, and similarly for "$Q[\alpha, \beta]$"). Using a two-variable version of the "polynomial identity trick", we can thus conclude that $P = Q$ as polynomials. Substituting $u = a$ and $v = b$ in this equality, we obtain $P[a, b] = Q[a, b]$; but this rewrites as $\binom{a+b}{n} = \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k}$. Thus, Proposition 6.3 is proven (modulo the groundwork that we have skipped: defining two-variable polynomials and proving that they still satisfy a version of the "polynomial identity trick"). ☐

Let us furthermore state two utterly obvious properties of the polynomials $\binom{x}{n}$:

Statement 7: Let $n \in \mathbb{N}$. Then, the polynomial $\binom{x}{n} \in \mathbb{K}[x]$ satisfies

$$\deg \binom{x}{n} = n \qquad \text{and} \qquad [x^n]\left(\binom{x}{n}\right) = \frac{1}{n!}.$$

[*Proof of Statement 7 (sketched):* The definition of $\binom{x}{n}$ yields

$$\binom{x}{n} = \frac{x\left(x-1\right)\left(x-2\right)\cdots\left(x-n+1\right)}{n!}.$$

If we multiply out the numerator of this fraction, then we obtain $x^n + (\text{lower-degree terms})$. Thus,

$$\binom{x}{n} = \frac{x^n + (\text{lower-degree terms})}{n!} = \frac{1}{n!}x^n + (\text{lower-degree terms}).$$

Thus, the polynomial $\binom{x}{n}$ has degree n, and its x^n-coefficient is $\frac{1}{n!}$. In other words, $\deg\binom{x}{n} = n$ and $\left[x^n\right]\left(\binom{x}{n}\right) = \frac{1}{n!}$.

(An alternative proof of Statement 7 proceeds by induction on n. The induction step relies on observing that $\binom{x}{n} = \frac{x-n+1}{n}\binom{x}{n-1}$ when n is positive.)]

Statement 8: Let $u \in \mathbb{K}\left[x\right]$ and $i \in \mathbb{Q}$. Then,

$$\binom{x}{i}\left[u\right] = \binom{u}{i}.$$

[*Proof of Statement 8 (sketched):* If $i \notin \mathbb{N}$, then this equality boils down to $0\left[u\right] = 0$ (since both $\binom{x}{i}$ and $\binom{u}{i}$ are defined to be 0 in this case), which is clearly true. Hence, for the rest of this proof, we WLOG assume that $i \in \mathbb{N}$. Thus, the definition of $\binom{u}{i}$ yields $\binom{u}{i} = \frac{u\left(u-1\right)\left(u-2\right)\cdots\left(u-i+1\right)}{i!}$. But the definition of $\binom{x}{i}$ yields

$$\binom{x}{i} = \frac{x\left(x-1\right)\left(x-2\right)\cdots\left(x-i+1\right)}{i!}.$$

Substituting u for x in this equality, we find

$$\begin{aligned}
\binom{x}{i}\left[u\right] &= \frac{x\left(x-1\right)\left(x-2\right)\cdots\left(x-i+1\right)}{i!}\left[u\right] \\
&= \frac{u\left(u-1\right)\left(u-2\right)\cdots\left(u-i+1\right)}{i!} \qquad \left(\begin{array}{c}\text{since evaluation of polynomials at } u \\ \text{respects scaling and multiplication}\end{array}\right) \\
&= \binom{u}{i}.
\end{aligned}$$

This proves Statement 8.]

Finally, we shall use the following fact (proven as an exercise in the class notes):

Lemma 6.4. *Let $n \in \mathbb{N}$. Let $\mathbf{a} \in \mathbb{K}\left[x\right]_{\leq n}$. If $\left[x^n\right]\mathbf{a} = 0$, then $\mathbf{a} \in \mathbb{K}\left[x\right]_{\leq n-1}$.*

We now resume solving the exercise.

(c) We first prove that

$$\Delta\left(\binom{x}{k}\right) = \binom{x}{k-1} \qquad \text{for each } k \in \mathbb{Z}. \tag{54}$$

[*Proof of* (54)*:* Let $k \in \mathbb{Z}$. Then, Proposition 6.2 (applied to $n = x+1$) yields

$$\binom{x+1}{k} = \binom{(x+1)-1}{k} + \binom{(x+1)-1}{k-1} = \binom{x}{k} + \binom{x}{k-1}$$

(since $(x+1)-1 = x$). Thus,

$$\binom{x+1}{k} - \binom{x}{k} = \binom{x}{k-1}. \tag{55}$$

The definition of Δ shows that each $f \in \mathbb{K}[x]$ satisfies

$$\Delta(f) = f^{\Delta} = f[x+1] - f[x] \qquad \left(\text{by the definition of } f^{\Delta}\right).$$

Applying this to $f = \binom{x}{k}$, we obtain

$$\Delta\left(\binom{x}{k}\right) = \underbrace{\binom{x}{k}[x+1]}_{=\binom{x+1}{k}} - \underbrace{\binom{x}{k}[x]}_{=\binom{x}{k}} = \binom{x+1}{k} - \binom{x}{k} = \binom{x}{k-1}$$

(by Statement 8, applied to $i=k$ and $u=x+1$)

(by (55)). This proves (54).]

Now, we can solve part **(c)** of the exercise by induction on n:

Induction base: For each $k \in \mathbb{Z}$, we have

$$\underbrace{\Delta^0}_{=\mathrm{id}}\left(\binom{x}{k}\right) = \mathrm{id}\left(\binom{x}{k}\right) = \binom{x}{k} = \binom{x}{k-0} \qquad (\text{since } k = k-0).$$

In other words, part **(c)** of the exercise holds for $n = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that part **(c)** of the exercise holds for $n = m$. We must prove that part **(c)** of the exercise holds for $n = m+1$.

We have assumed that part **(c)** of the exercise holds for $n = m$. In other words, we have

$$\Delta^m\left(\binom{x}{k}\right) = \binom{x}{k-m} \qquad \text{for all } k \in \mathbb{Z}. \tag{56}$$

Now, for all $k \in \mathbb{Z}$, we have

$$\underbrace{\Delta^{m+1}}_{=\Delta\circ\Delta^m}\left(\binom{x}{k}\right) = (\Delta\circ\Delta^m)\left(\binom{x}{k}\right) = \Delta\left(\underbrace{\Delta^m\left(\binom{x}{k}\right)}_{=\binom{x}{k-m}\ (\text{by (56)})}\right) = \Delta\left(\binom{x}{k-m}\right)$$

$$= \binom{x}{k-m-1} \qquad (\text{by (54), applied to } k-m \text{ instead of } k)$$

$$= \binom{x}{k-(m+1)} \qquad (\text{since } k-m-1 = k-(m+1)).$$

In other words, part **(c)** of the exercise holds for $n = m + 1$. This completes the induction step. Thus, part **(c)** of the exercise is solved by induction.

(d) We shall prove the following statement:

Statement 9: Let $m \in \{-1, 0, 1, \ldots\}$. For each $f \in \mathbb{K}[x]_{\leq m}$, there exist elements $a_0, a_1, \ldots, a_m \in \mathbb{K}$ such that $f = \sum_{i=0}^{m} a_i \binom{x}{i}$.

[*Proof of Statement 9:* We shall prove Statement 9 by induction on m:

Induction base: For each $f \in \mathbb{K}[x]_{\leq -1}$, there exist elements $a_0, a_1, \ldots, a_{-1} \in \mathbb{K}$ such that $f = \sum_{i=0}^{-1} a_i \binom{x}{i}$ [18]. In other words, Statement 9 holds for $m = -1$. This completes the (extremely pedantic) induction base.

Induction step: Let $n \in \{-1, 0, 1, \ldots\}$ be such that $n > -1$. Assume that Statement 9 holds for $m = n - 1$. We must prove that Statement 9 holds for $m = n$.

We note that $n \geq 0$ (since n is an integer satisfying $n > -1$), thus $n \in \mathbb{N}$.

Now, let $f \in \mathbb{K}[x]_{\leq n}$. We shall construct elements $a_0, a_1, \ldots, a_n \in \mathbb{K}$ such that $f = \sum_{i=0}^{n} a_i \binom{x}{i}$.

Indeed, Statement 7 shows that the polynomial $\binom{x}{n} \in \mathbb{K}[x]$ satisfies $\deg \binom{x}{n} = n$ and $[x^n]\left(\binom{x}{n}\right) = \frac{1}{n!}$. From $\deg \binom{x}{n} = n \leq n$, we obtain $\binom{x}{n} \in \mathbb{K}[x]_{\leq n}$.

Now, define a scalar $\lambda \in \mathbb{K}$ by $\lambda = n! \cdot [x^n] f$. Then, $f - \lambda \binom{x}{n}$ is a $\mathbb{K}$-linear combination of the polynomials f and $\binom{x}{n}$. Since both of these polynomials f and $\binom{x}{n}$ belong to $\mathbb{K}[x]_{\leq n}$ (because $f \in \mathbb{K}[x]_{\leq n}$ and $\binom{x}{n} \in \mathbb{K}[x]_{\leq n}$), we thus conclude that their $\mathbb{K}$-linear combination $f - \lambda \binom{x}{n}$ also belongs to $\mathbb{K}[x]_{\leq n}$ (since $\mathbb{K}[x]_{\leq n}$ is a $\mathbb{K}$-module). Furthermore, since subtraction and scaling of polynomials are defined entrywise, we have

$$[x^n]\left(f - \lambda\binom{x}{n}\right) = [x^n] f - \underbrace{\lambda}_{= n! \cdot [x^n] f} \underbrace{[x^n]\left(\binom{x}{n}\right)}_{= \frac{1}{n!}}$$
$$= [x^n] f - n! \cdot ([x^n] f) \cdot \frac{1}{n!} = [x^n] f - [x^n] f = 0.$$

Hence, Lemma 6.4 (applied to $\mathbf{a} = f - \lambda \binom{x}{n}$) yields $f - \lambda \binom{x}{n} \in \mathbb{K}[x]_{\leq n-1}$. Thus, we can apply Statement 9 to $n - 1$ and $f - \lambda \binom{x}{n}$ instead of m and f (because we have

[18] *Proof.* Let $f \in \mathbb{K}[x]_{\leq -1}$. Then, $f \in \mathbb{K}[x]_{\leq -1} = \{0\} = 0$, so that $f = 0$. Thus, $f = \sum_{i=0}^{-1} 0 \binom{x}{i}$ (since $\sum_{i=0}^{-1} 0 \binom{x}{i} =$ (empty sum) $= 0 = f$). Hence, there exist elements $a_0, a_1, \ldots, a_{-1} \in \mathbb{K}$ such that $f = \sum_{i=0}^{-1} a_i \binom{x}{i}$ (namely, $a_i = 0$). Qed.

assumed that Statement 9 holds for $m = n-1$). We thus conclude that there exist elements $a_0, a_1, \ldots, a_{n-1} \in \mathbb{K}$ such that

$$f - \lambda \binom{x}{n} = \sum_{i=0}^{n-1} a_i \binom{x}{i}. \tag{57}$$

Consider these elements $a_0, a_1, \ldots, a_{n-1}$. We thus have obtained an n-tuple $(a_0, a_1, \ldots, a_{n-1}) \in \mathbb{K}^n$. We extend this n-tuple $(a_0, a_1, \ldots, a_{n-1})$ to an $(n+1)$-tuple $(a_0, a_1, \ldots, a_n) \in \mathbb{K}^{n+1}$ by setting $a_n = \lambda$. Then,

$$\begin{aligned}
\sum_{i=0}^{n} a_i \binom{x}{i} &= \sum_{i=0}^{n-1} a_i \binom{x}{i} + \underbrace{a_n}_{=\lambda} \binom{x}{n} \qquad \left(\begin{array}{c} \text{here, we have split off the} \\ \text{addend for } i = n \text{ from the sum} \end{array} \right) \\
&= \sum_{i=0}^{n-1} a_i \binom{x}{i} + \lambda \binom{x}{n} = f \qquad (\text{by (57)}) .
\end{aligned}$$

In other words, $f = \sum_{i=0}^{n} a_i \binom{x}{i}$.

Now, forget that we defined $a_0, a_1, \ldots, a_n$. We thus have found $n+1$ elements $a_0, a_1, \ldots, a_n \in \mathbb{K}$ such that $f = \sum_{i=0}^{n} a_i \binom{x}{i}$. Hence, we have shown that such elements exist. In other words, there exist elements $a_0, a_1, \ldots, a_n \in \mathbb{K}$ such that $f = \sum_{i=0}^{n} a_i \binom{x}{i}$.

Now, forget that we fixed f. We thus have shown that for each $f \in \mathbb{K}[x]_{\leq n}$, there exist elements $a_0, a_1, \ldots, a_n \in \mathbb{K}$ such that $f = \sum_{i=0}^{n} a_i \binom{x}{i}$. In other words, Statement 9 holds for $m = n$. This completes the induction step. Thus, Statement 9 is proven by induction.]

Now, let $m \in \mathbb{N}$, and let $f \in \mathbb{K}[x]$ be a polynomial of degree $\leq m$. Then, $m \in \mathbb{N} \subseteq \{-1, 0, 1, \ldots\}$. Furthermore, $f \in \mathbb{K}[x]_{\leq m}$ (since $f \in \mathbb{K}[x]$ is a polynomial of degree $\leq m$). Hence, Statement 9 shows that there exist elements $a_0, a_1, \ldots, a_m \in \mathbb{K}$ such that $f = \sum_{i=0}^{m} a_i \binom{x}{i}$. This solves part **(d)** of the exercise.

(e) Let $f \in \mathbb{K}[x]$ be a polynomial.

Let $m = \deg f$. Thus, f is a polynomial of degree $\leq m$ (since $\deg f = m \leq m$). Hence, part **(d)** of this exercise shows that there exist elements $a_0, a_1, \ldots, a_m \in \mathbb{K}$ such that $f = \sum_{i=0}^{m} a_i \binom{x}{i}$. Consider these $a_0, a_1, \ldots, a_m$. Substituting $x + a$ for x in the equality $f = \sum_{i=0}^{m} a_i \binom{x}{i}$, we obtain

$$\begin{aligned}
f[x+a] &= \left(\sum_{i=0}^{m} a_i \binom{x}{i} \right) [x+a] = \sum_{i=0}^{m} a_i \underbrace{\binom{x}{i}[x+a]}_{\substack{= \binom{x+a}{i} \\ \text{(by Statement 8,} \\ \text{applied to } u = x+a)}} \\
&= \sum_{i=0}^{m} a_i \binom{x+a}{i}.
\end{aligned} \tag{58}$$

Now, we claim that

$$\binom{x+a}{i} = \sum_{n=0}^{m} \binom{a}{i-n} \cdot \binom{x}{n} \tag{59}$$

for each $i \in \{0, 1, \ldots, m\}$.

[*Proof of* (59)*:* Let $i \in \{0, 1, \ldots, m\}$. Then, $0 \leq i \leq m$. We have

$$\begin{aligned}
&\sum_{n=0}^{m} \binom{a}{i-n} \cdot \binom{x}{n} \\
&= \sum_{n=0}^{i} \underbrace{\binom{a}{i-n} \cdot \binom{x}{n}}_{= \binom{x}{n}\binom{a}{i-n}} + \sum_{n=i+1}^{m} \underbrace{\binom{a}{i-n}}_{\substack{=0 \\ (\text{since } i-n<0 \\ (\text{because } n \geq i+1 > i))}} \cdot \binom{x}{n} \\
&\qquad \text{(here, we have split our sum at } n = i, \text{ since } 0 \leq i \leq m) \\
&= \sum_{n=0}^{i} \binom{x}{n}\binom{a}{i-n} + \underbrace{\sum_{n=i+1}^{m} 0 \cdot \binom{x}{n}}_{=0} = \sum_{n=0}^{i} \binom{x}{n}\binom{a}{i-n}.
\end{aligned} \tag{60}$$

But $x \in \mathbb{K}[x]$ and $a \in \mathbb{K} \subseteq \mathbb{K}[x]$. Hence, Proposition 6.3 (applied to x, a and i instead of a, b and n) yields

$$\binom{x+a}{i} = \sum_{k=0}^{i} \binom{x}{k}\binom{a}{i-k} = \sum_{n=0}^{i} \binom{x}{n}\binom{a}{i-n}$$

(here, we have renamed the summation index k as n). Comparing this with (60), we obtain $\binom{x+a}{i} = \sum_{n=0}^{m} \binom{a}{i-n} \cdot \binom{x}{n}$. This proves (59).]

But each $n \in \mathbb{N}$ satisfies

$$\begin{aligned}
\Delta^n(f) &= \Delta^n\left(\sum_{i=0}^{m} a_i \binom{x}{i}\right) \qquad \left(\text{since } f = \sum_{i=0}^{m} a_i \binom{x}{i}\right) \\
&= \sum_{i=0}^{m} a_i \underbrace{\left(\Delta^n\left(\binom{x}{i}\right)\right)}_{\substack{= \binom{x}{i-n} \\ (\text{by part } \textbf{(c)} \text{ of this exercise,} \\ \text{applied to } k=i)}} \qquad \left(\begin{array}{l}\text{since the map } \Delta^n \text{ is } \mathbb{K}\text{-linear} \\ (\text{since the map } \Delta \text{ is } \mathbb{K}\text{-linear})\end{array}\right) \\
&= \sum_{i=0}^{m} a_i \binom{x}{i-n}
\end{aligned}$$

and thus

$$\begin{aligned}
&\underbrace{\left(\Delta^n\left(f\right)\right)}_{=\sum_{i=0}^{m} a_i \binom{x}{i-n}}\ [a] \\
&= \left(\sum_{i=0}^{m} a_i \binom{x}{i-n}\right)[a] = \sum_{i=0}^{m} a_i \underbrace{\binom{x}{i-n}[a]}_{=\binom{a}{i-n}} \\
&\qquad \text{(by Statement 8, applied to } i-n \text{ and } a \text{ instead of } i \text{ and } u\text{)} \\
&\qquad \text{(since evaluation of polynomials at } a \text{ respects } \mathbb{K}\text{-linear combinations)} \\
&= \sum_{i=0}^{m} a_i \binom{a}{i-n}. \qquad (61)
\end{aligned}$$

On the other hand, Statement 6 shows that $\Delta^n\left(f\right) = 0$ for all integers $n > m$. Hence,

$$\underbrace{\left(\Delta^n\left(f\right)\right)}_{=0}[a] \cdot \binom{x}{n} = \underbrace{0\,[a]}_{=0} \cdot \binom{x}{n} = 0 \qquad (62)$$

for all integers $n > m$. Thus, all but finitely many addends of the sum $\sum_{n\in\mathbb{N}} \left(\Delta^n\left(f\right)\right)[a] \cdot \binom{x}{n}$ are 0. In other words, this sum $\sum_{n\in\mathbb{N}} \left(\Delta^n\left(f\right)\right)[a] \cdot \binom{x}{n}$ has only finitely many nonzero addends; thus, it is well-defined. Furthermore, we can split this sum at $n = m$; we thus find

$$\begin{aligned}
&\sum_{n\in\mathbb{N}} \left(\Delta^n\left(f\right)\right)[a] \cdot \binom{x}{n} \\
&= \sum_{n=0}^{m} \underbrace{\left(\Delta^n\left(f\right)\right)[a]}_{=\sum_{i=0}^{m} a_i \binom{a}{i-n} \text{ (by (61))}} \cdot \binom{x}{n} + \sum_{n=m+1}^{\infty} \underbrace{\left(\Delta^n\left(f\right)\right)[a] \cdot \binom{x}{n}}_{=0 \text{ (by (62), since } n \geq m+1 > m\text{)}} \\
&= \sum_{n=0}^{m} \sum_{i=0}^{m} a_i \binom{a}{i-n} \cdot \binom{x}{n} + \underbrace{\sum_{n=m+1}^{\infty} 0}_{=0} = \sum_{n=0}^{m} \sum_{i=0}^{m} a_i \binom{a}{i-n} \cdot \binom{x}{n} \\
&= \sum_{i=0}^{m} a_i \underbrace{\sum_{n=0}^{m} \binom{a}{i-n} \cdot \binom{x}{n}}_{=\binom{x+a}{i} \text{ (by (59))}} = \sum_{i=0}^{m} a_i \binom{x+a}{i}.
\end{aligned}$$

Comparing this with (58), we obtain

$$f\left[x+a\right] = \sum_{n\in\mathbb{N}} \left(\Delta^n\left(f\right)\right)[a] \cdot \binom{x}{n}.$$

This completes the solution of part **(e)** of this exercise.

(f) We have assumed that

$$f[k] \in \mathbb{Z} \qquad \text{for each } k \in \{0,1,\ldots,m\}. \tag{63}$$

Next, we claim that for each $n \in \{0,1,\ldots,m\}$, we have

$$\left(\Delta^n(f)\right)[k] \in \mathbb{Z} \qquad \text{for all } k \in \{0,1,\ldots,m-n\}. \tag{64}$$

[*Proof of* (64)*:* We shall prove (64) by induction on n:
Induction base: For each $k \in \{0,1,\ldots,m-0\}$, we have

$$\left(\underbrace{\Delta^0}_{=\mathrm{id}}(f)\right)[k] = \underbrace{\left(\mathrm{id}(f)\right)}_{=f}[k] = f[k] \in \mathbb{Z}$$

(by (63), since $k \in \{0,1,\ldots,m-0\} = \{0,1,\ldots,m\}$). Thus, we have shown that $\left(\Delta^0(f)\right)[k] \in \mathbb{Z}$ for all $k \in \{0,1,\ldots,m-0\}$. In other words, (64) holds for $n=0$. This completes the induction base.

Induction step: Let $N \in \{0,1,\ldots,m\}$ be positive. Assume that (64) holds for $n = N-1$. We must prove that (64) holds for $n = N$.

We have assumed that (64) holds for $n = N-1$. In other words, we have

$$\left(\Delta^{N-1}(f)\right)[k] \in \mathbb{Z} \qquad \text{for all } k \in \{0,1,\ldots,m-(N-1)\}. \tag{65}$$

Now, let $k \in \{0,1,\ldots,m-N\}$. Then, $k \in \{0,1,\ldots,m-N\} \subseteq \{0,1,\ldots,m-(N-1)\}$ (since $m-N \leq m-(N-1)$). Thus, (65) yields $\left(\Delta^{N-1}(f)\right)[k] \in \mathbb{Z}$. Furthermore, from $k \in \{0,1,\ldots,m-N\}$, we obtain

$$k+1 \in \{1,2,\ldots,m-N+1\} \subseteq \{0,1,\ldots,m-N+1\} = \{0,1,\ldots,m-(N-1)\}$$

(since $m-N+1 = m-(N-1)$). Thus, (65) (applied to $k+1$ instead of k) yields $\left(\Delta^{N-1}(f)\right)[k+1] \in \mathbb{Z}$. But

$$\underbrace{\Delta^N}_{=\Delta\circ\Delta^{N-1}}(f) = \left(\Delta\circ\Delta^{N-1}\right)(f) = \Delta\left(\Delta^{N-1}(f)\right) = \left(\Delta^{N-1}(f)\right)^{\Delta} \qquad \text{(by the definition of } \Delta\text{)}$$
$$= \left(\Delta^{N-1}(f)\right)[x+1] - \left(\Delta^{N-1}(f)\right)[x] \qquad \left(\text{by the definition of } \left(\Delta^{N-1}(f)\right)^{\Delta}\right).$$

Evaluating both sides of this equality at k, we find

$$\left(\Delta^N(f)\right)[k] = \left(\left(\Delta^{N-1}(f)\right)[x+1] - \left(\Delta^{N-1}(f)\right)[x]\right)[k]$$
$$= \underbrace{\left(\left(\Delta^{N-1}(f)\right)[x+1]\right)[k]}_{=(\Delta^{N-1}(f))[k+1]\in\mathbb{Z}} - \underbrace{\left(\left(\Delta^{N-1}(f)\right)[x]\right)[k]}_{=(\Delta^{N-1}(f))[k]\in\mathbb{Z}} \in \mathbb{Z}$$

(since the difference of two elements of $\mathbb{Z}$ must always belong to $\mathbb{Z}$).

Now, forget that we fixed k. We thus have shown that $\left(\Delta^N(f)\right)[k] \in \mathbb{Z}$ for all $k \in \{0,1,\ldots,m-N\}$. In other words, (64) holds for $n = N$. This completes the induction step. Thus, (64) is proven by induction.]

For each $n \in \{0,1,\ldots,m\}$, we have $0 \in \{0,1,\ldots,m-n\}$ (since $n \in \{0,1,\ldots,m\}$ yields $n \leq m$ and thus $m-n \geq 0$) and thus

$$\left(\Delta^n(f)\right)[0] \in \mathbb{Z}$$

(by (64), applied to $k = 0$). Renaming the index n as i in this statement, we obtain the following: For each $i \in \{0, 1, \ldots, m\}$, we have

$$\left(\Delta^i(f)\right)[0] \in \mathbb{Z}. \tag{66}$$

On the other hand, Statement 6 shows that $\Delta^n(f) = 0$ for all integers $n > m$. Hence,

$$\underbrace{\left(\Delta^n(f)\right)}_{=0}[0] \cdot \binom{x}{n} = \underbrace{0}_{=0}[0] \cdot \binom{x}{n} = 0 \tag{67}$$

for all integers $n > m$.

Now, part **(e)** of this exercise (applied to $a = 0$) yields

$$\begin{aligned} f[x+0] &= \sum_{n \in \mathbb{N}} \left(\Delta^n(f)\right)[0] \cdot \binom{x}{n} \\ &= \sum_{n=0}^{m} \left(\Delta^n(f)\right)[0] \cdot \binom{x}{n} + \sum_{n=m+1}^{\infty} \underbrace{\left(\Delta^n(f)\right)[0] \cdot \binom{x}{n}}_{\substack{=0 \\ \text{(by (67), since } n \geq m+1 > m\text{)}}} \\ &\qquad \text{(here, we have split the sum at } n = m\text{)} \\ &= \sum_{n=0}^{m} \left(\Delta^n(f)\right)[0] \cdot \binom{x}{n} + \underbrace{\sum_{n=m+1}^{\infty} 0}_{=0} = \sum_{n=0}^{m} \left(\Delta^n(f)\right)[0] \cdot \binom{x}{n} = \sum_{i=0}^{m} \left(\Delta^i(f)\right)[0] \cdot \binom{x}{i} \end{aligned}$$

(here, we have renamed the summation index n as i). Comparing this with $f[x+0] = f[x] = f$, we obtain

$$f = \sum_{i=0}^{m} \left(\Delta^i(f)\right)[0] \cdot \binom{x}{i}.$$

The coefficients $\left(\Delta^i(f)\right)[0]$ appearing in this sum are integers (because of (66)). Hence, there exist integers $a_0, a_1, \ldots, a_m$ such that $f = \sum_{i=0}^{m} a_i \binom{x}{i}$ (namely, $a_i = \left(\Delta^i(f)\right)[0]$). This solves part **(f)** of the exercise.

References

[Clark18] Pete L. Clark, *Number Theory: A Contemporary Introduction*, 8 January 2018. `http://math.uga.edu/~pete/4400FULL.pdf`

[GalQua17] Jean Gallier, Jocelyn Quaintance, *Notes on Primality Testing And Public Key Cryptography, Part 1*, 27 February 2019. `https://www.cis.upenn.edu/~jean/RSA-primality-testing.pdf`

[GrKnPa94] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics, Second Edition*, Addison-Wesley 1994. See `https://www-cs-faculty.stanford.edu/~knuth/gkp.html` for errata.

[Grinbe15] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 10 January 2019. `http://www.cip.ifi.lmu.de/~grinberg/primes2015/sols.pdf`

The numbering of theorems and formulas in this link might shift when the project gets updated; for a "frozen" version whose numbering is guaranteed to match that in the citations above, see `https://github.com/darijgr/detnotes/releases/tag/2019-01-10` .

[Grinbe18] Darij Grinberg, *Why the log and exp series are mutually inverse*, May 11, 2018. `http://www.cip.ifi.lmu.de/~grinberg/t/17f/logexp.pdf`

[Loehr11] Nicholas A. Loehr, *Bijective Combinatorics*, Chapman & Hall/CRC 2011.

[Wilf94] Herbert S. Wilf, *generatingfunctionology*, 1999. `https://www.math.upenn.edu/~wilf/DownldGF.html`