8.2. Commutative rings

We shall define FPS (= formal power series) and justify what we did to them in §8.1 (dividing, solving quadratic eqns, ∞ sums, ...).

First things first: FPS are not functions. You cannot substitute $x=2$ into $\frac{1}{1-x} = 1+x+x^2+\cdots$ and "obtain" $\frac{1}{-1} = 1+2+4+8+16+\cdots$.

Let us go back to abstract algebra to see what we can do.

Def. A commutative ring (CR) is, informally, a set K equipped with binary operations $\oplus$, $\ominus$, and $\odot$ and elements 0 and 1 that "behave" like addition, subtraction, multiplication (of numbers) and the numbers 0 and 1, respectively. For example, they should satisfy rules like

$$(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c).$$

Formally: A commutative ring is a set K equipped with

maps $\oplus : K \times K \to K$, $\ominus : K \times K \to K$, $\odot : K \times K \to K$ and elements $0 \in K$ and $1 \in K$ satisfying the following axioms:

(a) Commutativity of $\oplus$: $a \oplus b = b \oplus a$.

(b) Associativity of $\oplus$: $a \oplus (b \oplus c) = (a \oplus b) \oplus c$.

(c) Neutrality of 0: $a \oplus 0 = a = 0 \oplus a$.

(d) $\ominus$ undoes $\oplus$: $a \oplus b = c \iff a = c \ominus b$.

(e) Commutativity of $\odot$: $a \odot b = b \odot a$.

(f) Associativity of $\odot$: $a \odot (b \odot c) = (a \odot b) \odot c$.

(g) Distributivity: $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$;
$(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c)$.

(h) Neutrality of 1: $a \odot 1 = a = 1 \odot a$.

(i) Annihilation: $a \odot 0 = 0 = 0 \odot a$.

Note: Most authors do not include $\ominus$ in the definition of a CR, and require an "existence of additive inverses" axiom instead of (d).

Examples: • $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ (with usual operations).

• $\mathbb{N}$ is not a CR, as it has no $\ominus$.
This is called a semiring.

• $\mathbb{Q}^{m\times m}$ is not a CR, since it fails axiom (e) for $m>1$.
This is called a noncommutative ring.

• Polynomial rings are CRs (but we first need to define them).

• $\mathbb{Z}[\sqrt{5}] = \{a+b\sqrt{5} \mid a,b\in\mathbb{Z}\}$ is a CR with operations $+,-,\cdot$ inherited from $\mathbb{R}$. This is because

$$(a+b\sqrt{5})+(c+d\sqrt{5}) = (a+c)+(b+d)\sqrt{5};$$
$$(a+b\sqrt{5})-(c+d\sqrt{5}) = (a-c)+(b-d)\sqrt{5};$$
$$(a+b\sqrt{5})\cdot(c+d\sqrt{5}) = (ac+5bd)+(ad+bc)\sqrt{5}.$$

This is called a subring of R (i.e., a subset which is a CR with its operations $+, -, \cdot$ inherited from R).

- For each $m \in \mathbb{N}$, the set $\mathbb{Z}/m$ is a ring, with addition defined by $\bar{a} + \bar{b} = \overline{a+b}$, etc.
This ring is finite if $m > 0$.

- Fix a set S. Consider the power set $\mathcal{P}(S)$ of S.
Then, $\mathcal{P}(S)$ is a CR with addition Δ and multiplication $\cap$
(recall: $X \Delta Y = (X \setminus Y) \cup (Y \setminus X)$)
and subtraction Δ and $\textcircled{0} = \emptyset$ and $\textcircled{1} = S$.
Indeed, e.g. axiom (g) holds because

$$A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C);$$
$$(A \Delta B) \cap C = (A \cap C) \Delta (B \cap C).$$

This is called a Boolean ring.

- Here is another "not-quite-CR".

Let $\mathbb{T} = \mathbb{Z} \cup \{-\infty\}$, where $-\infty$ is just an extra symbol.

Define two operations $+_{\mathbb{T}}$ and $\cdot_{\mathbb{T}}$ on $\mathbb{T}$ by

$$a +_{\mathbb{T}} b = \max\{a, b\}$$

(where $\max\{n, -\infty\} := n$ $\forall n \in \mathbb{T}$)

and

$$a \cdot_{\mathbb{T}} b = a + b$$

(where $n + (-\infty) := (-\infty) + n = -\infty$ $\forall n \in \mathbb{T}$).

Then, $\mathbb{T}$ is almost a CR with these operations, except that it lacks a subtraction. So, again, $\mathbb{T}$ is a semiring (called the <u>tropical semiring of $\mathbb{Z}$</u>).

See [detnotes, Ch. 6] for more examples; see any abstract algebra book, too. We shall usually write $+, -$ and $\cdot$ for $\oplus$, $\ominus$ and $\odot$.

Good news: In any commutative ring, the standard rules of computation apply:

- You can compute finite sums without caring about the order of summation or parenthesis placement:
  $((a+(b+c))+d)+e = (a+b)+(c+(d+e))$
  ("general associativity"),
  so you can write $a+b+c+d+e$.
  Also, $a+b+c+d+e = d+b+a+e+c$
  ("general commutativity").

More formally: If $(a_s)_{s\in S}$ is any finite family of elements of a CR K, then $\sum_{s\in S} a_s$ is well-defined and satisfies the usual rules of sums (e.g., if $S = X \cup Y$ and $X \cap Y = \emptyset$, then $\sum_{s\in S} a_s = \sum_{s\in X} a_s + \sum_{s\in Y} a_s$).

For proofs, see [detnotes, Ch. 2] (but read "elements of K"

for "numbers"), or Spring 2018 Math 4707 Thm 2.8 (Feb. 7 notes).

- The same holds for products.
- If $n \in \mathbb{Z}$ and $a \in K$ (for $K \supseteq \mathbb{CR}$), then we can define $na \in K$ to be

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

(where $-b := 0 - b = \mathbb{0} \ominus b$).

- Standard rules hold:

$$-(a+b) = (-a) + (-b); \qquad -(-a) = a;$$

$$(nm)a = n(ma) \quad (\text{for } n, m \in \mathbb{Z});$$

. . . . . ;

$$(ab)^n = a^n b^n \quad (\text{for } n \in \mathbb{N});$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \qquad (\text{for } n \in \mathbb{N})$$

(the binomial theorem).

8.3. The definition of FPS

Fix a commutative ring K. (For example, $K = \mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{C}$.)

Def. A **FPS (formal power series)** (in 1 indeterminate over K) is a sequence $(a_0, a_1, a_2, \ldots) = (a_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$ of elements of K.

This answers "what is an FPS", but not "what we can do with them" or "why do the Examples in §8.1 work" or "what is x".

This will take us a while.

Def. (a) The **sum** of two FPSs $(a_0, a_1, a_2, \ldots)$ and $(b_0, b_1, b_2, \ldots)$ is

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots).$$

(b) The <u>difference</u> of two FPSs $(a_0, a_1, a_2, \ldots)$ and $(b_0, b_1, b_2, \ldots)$ is

$$(a_0, a_1, a_2, \ldots) - (b_0, b_1, b_2, \ldots) = (a_0 - b_0, a_1 - b_1, a_2 - b_2, \ldots).$$

(c) If $\lambda \in K$ and $(a_0, a_1, a_2, \ldots)$ is an FPS, then

$$\lambda(a_0, a_1, a_2, \ldots) := (\lambda a_0, \lambda a_1, \lambda a_2, \ldots).$$

(d) The <u>product</u> of two FPSs $(a_0, a_1, a_2, \ldots)$ and $(b_0, b_1, b_2, \ldots)$ is the FPS $(c_0, c_1, c_2, \ldots)$, where

$$c_n = \sum_{i=0}^{n} a_i b_{n-i} = \sum_{\substack{i,j \in \mathbb{N}; \\ i+j=n}} a_i b_j$$

$$= a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0.$$

(e) For each $a \in K$, the FPS $\underline{a}$ is defined to be $(a, 0, 0, 0, \ldots)$. This is called a <u>constant</u> FPS.

(f) The set of all FPSs is called $K[[x]]$.

Thm. 8.1. $K[[x]]$ is a CR (with $+, -, \cdot$ defined as above) with $\textcircled{0} = \underline{0} = (0,0,0,\ldots)$ and $\textcircled{1} = \underline{1} = (1,0,0,0,\ldots)$ and has K as a subring (if we identify each $a \in K$ with $\underline{a}$). This means:

(a) Addition in $K[[x]]$ is commutative & associative:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}; \qquad \vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}.$$

(b) $\underline{0} + \vec{a} = \vec{a} = \vec{a} + \underline{0}$.

(c) Multiplication in $K[[x]]$ is* commutative & associative:

$$\vec{a}\,\vec{b} = \vec{b}\,\vec{a}; \qquad \vec{a}(\vec{b}\,\vec{c}) = (\vec{a}\,\vec{b})\,\vec{c}.$$

(d) $\underline{1}\,\vec{a} = \vec{a} = \vec{a}\,\underline{1}$.

(e) $\underline{0}\,\vec{a} = \underline{0} = \vec{a}\,\underline{0}$.

(f) Distributivity holds:

$$\vec{a}(\vec{b} + \vec{c}) = \vec{a}\vec{b} + \vec{a}\vec{c}; \qquad (\vec{a} + \vec{b})\vec{c} = \vec{a}\vec{c} + \vec{b}\vec{c}.$$

(g) $\vec{a} + \vec{b} = \vec{c} \iff \vec{a} = \vec{c} - \vec{b}$.

---

* We write $\vec{a} \cdot \vec{b}$ or $\vec{a}\vec{b}$ for the product of two FPSs $\vec{a}$ and $\vec{b}$.

(h) $\forall a, b \in K$, we have $\underline{a} + \underline{b} = \underline{a+b}$ and $\underline{a} \cdot \underline{b} = \underline{ab}$.

Furthermore, $K[[x]]$ is a <u>K-module</u> (same as a K-vector space, except that K is not necessarily a field).

Concretely, this means:

(i) ~~$\lambda(\vec{a}+\vec{b}) = \lambda\vec{a} + \lambda\vec{b}$~~ $\lambda(\vec{a}+\vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ $\forall \lambda \in K$;

(j) $(\lambda+\mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ $\forall \lambda, \mu \in K$;

(k) $(\lambda\mu)\vec{a} = \lambda(\mu\vec{a})$ —"—;

(l) $1\vec{a} = \vec{a}$.

Finally,

(m) $\lambda\vec{a} = \underline{\lambda} \cdot \vec{a}$ $\forall \lambda \in K$ and $\vec{a} \in K[[x]]$.

The purpose of Thm. 8.1 is to justify computing with FPSs as with numbers, at least as far as $+, -$, and $\cdot$ are concerned.

Hence, e.g., we know that:

- sums & products in $K[[x]]$ need no parentheses and

don't depend on the order (so, e.g., we have

$$((\vec{a}\,\vec{b})\,\vec{c})\,\vec{d} = \vec{a}\,((\vec{b}\,\vec{c})\,\vec{d}) = \vec{a}\,(\vec{b}\,(\vec{c}\,\vec{d})),$$

so we can write $\vec{a}\,\vec{b}\,\vec{c}\,\vec{d}$ for all of these; furthermore,

$$\vec{a}\,\vec{b}\,\vec{c}\,\vec{d} = \vec{a}\,\vec{d}\,\vec{c}\,\vec{b} = \vec{c}\,\vec{a}\,\vec{d}\,\vec{b}).$$

- Finite sums & products ($\sum_{i=1}^{k} \vec{a}_i$, $\sum_{i\in I} \vec{a}_i$, $\prod_{i=1}^{k} \vec{a}_i$, $\prod_{i\in I} \vec{a}_i$) make sense & behave as usual.

- Powers exist: $\vec{a}^{\,n} = \underbrace{\vec{a}\,\vec{a}\,\cdots\,\vec{a}}_{n \text{ times}}$ $\forall n \in \mathbb{N}$.

  This includes $\vec{a}^{\,0} = \underline{1}$.

- standard rules hold: $\vec{a}^{\,n+m} = \vec{a}^{\,n}\vec{a}^{\,m}$, $(\vec{a}\,\vec{b})^n = \vec{a}^{\,n}\vec{b}^{\,n}$, etc.

- The binomial formula holds: $(\vec{a}+\vec{b})^n = \sum_{k=0}^{n} \binom{n}{k} \vec{a}^{\,k}\,\vec{b}^{\,n-k}$.

Def. If $n \in \mathbb{N}$ and $\vec{a} = (a_0, a_1, a_2, \ldots) \in K[[x]]$, then we set $[x^n]\vec{a} := a_n$. This is called the coefficient of x^n in $\vec{a}$, or the n-th coefficient of $\vec{a}$.

Thus, the definition of the sum of two FPSs rewrites as

$$[x^n](\vec{a} + \vec{b}) = [x^n]\vec{a} + [x^n]\vec{b}. \tag{81}$$

Also, the definition of the product of two FPSs rewrites as

$$[x^n](\vec{a}\vec{b}) = [x^0]\vec{a} \cdot [x^n]\vec{b} + [x^1]\vec{a} \cdot [x^{n-1}]\vec{b} + \ldots + [x^n]\vec{a} \cdot [x^0]\vec{b} \tag{82}$$

$$= \sum_{i=0}^{n} [x^i]\vec{a} \cdot [x^{n-i}]\vec{b} \tag{83}$$

$$= \sum_{j=0}^{n} [x^{n-j}]\vec{a} \cdot [x^j]\vec{b}. \tag{84}$$

Proof of Thm. 8.1. Most parts are strfwd.

(c) Associativity. Let $n \in \mathbb{N}$. Consider the two equalities

$$[x^n]((\vec{a}\,\vec{b})\,\vec{c}) \overset{(84)}{=} \sum_{j=0}^{n} \underbrace{[x^{n-j}](\vec{a}\,\vec{b})}_{\overset{(83)}{=}\sum_{i=0}^{n-j}[x^i]\vec{a}\cdot[x^{n-j-i}]\vec{b}} \cdot [x^j]\vec{c}$$

$$= \sum_{j=0}^{n}\sum_{i=0}^{n-j} [x^i]\vec{a}\cdot[x^{n-j-i}]\vec{b}\cdot[x^j]\vec{c}$$

and

$$[x^n](\vec{a}\,(\vec{b}\,\vec{c})) \overset{(83)}{=} \sum_{i=0}^{n} [x^i]\vec{a} \cdot \underbrace{[x^{n-i}](\vec{b}\,\vec{c})}_{\overset{(84)}{=}\sum_{j=0}^{n-i}[x^{n-i-j})\vec{b}\cdot[x^j]\vec{c}}$$

$$= \sum_{i=0}^{n}\sum_{j=0}^{n-i} [x^i]\vec{a}\cdot[x^{n-i-j}]\vec{b}\cdot[x^j]\vec{c}.$$

The RHSs are equal, since

$$\sum_{j=0}^{n}\sum_{i=0}^{n-j} = \sum_{\substack{i,j\in\mathbb{N};\\ i+j\leq n}} = \sum_{i=0}^{n}\sum_{j=0}^{n-i}$$

and $n-j-i = n-i-j$. Thus, $[x^n]((\vec{a}\,\vec{b})\vec{c}) = [x^n](\vec{a}(\vec{b}\,\vec{c}))$ $\forall n\in\mathbb{N}$. Hence, $(\vec{a}\,\vec{b})\vec{c} = \vec{a}(\vec{b}\,\vec{c})$, since an FPS is just the sequence of its coefficients.

□

Rest of Thm. 8.1 is LTTR.

Sometimes, infinite sums of FPSs make sense:

Example:

$$\begin{aligned}
&(1, 1, 1, 1, \ldots)\\
+\ &(0, 1, 1, 1, \ldots)\\
+\ &(0, 0, 1, 1, \ldots)\\
+\ &(0, 0, 0, 1, \ldots)\\
+\ &\ldots\\
&= (1, 2, 3, 4, \ldots).
\end{aligned}$$

Def. A (possibly infinite) family $(\vec{a}_i)_{i\in I}$ of FPSs is called **summable** if

(85) $\forall n \in \mathbb{N}$, only finitely many $i \in I$ satisfy $[x^n]\vec{a}_i \neq 0$.

In this case, the sum $\sum_{i\in I} \vec{a}_i$ is defined as the FPS with

$$[x^n]\left(\sum_{i\in I} \vec{a}_i\right) = \underbrace{\sum_{i\in I} [x^n]\vec{a}_i}_{\substack{\text{a sum with only} \\ \text{finitely many} \\ \text{nonzero addends,} \\ \text{hence well-defined in } \mathbb{K}}} \qquad \forall n \in \mathbb{N}.$$

Rmk. (85) $\Longleftrightarrow$ ($\forall n \in \mathbb{N}$, infinitely many $i \in I$ satisfy $[x^n]\vec{a}_i = 0$).

Prop. 8.2. Sums of summable families satisfy the usual rules

for summation (as long as all families involved are summable).

Caveat: $\sum_{i\in I}\sum_{j\in J}\overrightarrow{a_{i,j}} = \sum_{j\in J}\sum_{i\in I}\overrightarrow{a_{i,j}}$ requires the family $(\overrightarrow{a_{i,j}})_{(i,j)\in I\times J}$ to be summable.

(See the example after Thm. 2.17 for why this is needed, even for numbers!)

So the correct rule for interchanging summations is ("Discrete Fubini Theorem"):

If $(\overrightarrow{a_{i,j}})_{(i,j)\in I\times J}$ is a summable family of FPSs, then

$$\sum_{i\in I}\sum_{j\in J}\overrightarrow{a_{i,j}} = \sum_{(i,j)\in I\times J}\overrightarrow{a_{i,j}} = \sum_{j\in J}\sum_{i\in I}\overrightarrow{a_{i,j}}.$$

Def. x denotes the FPS $(0,1,0,0,0,\ldots)$.

Prop. 8.3. $x^k = (\underbrace{0, 0, \ldots, 0}_{k \text{ zeroes}}, 1, 0, 0, 0, \ldots) \quad \forall k \in \mathbb{N}$.

Prof. Induct on k. The ind. step relies on the observation that if $\vec{a} = (a_0, a_1, a_2, \ldots)$, then $x\vec{a} = (0, a_0, a_1, a_2, \ldots)$. □

Cor. 8.4. Any FPS $(a_0, a_1, a_2, \ldots) \in K[[x]]$ satisfies

$$(a_0, a_1, a_2, \ldots) = a_0 + a_1 x + a_2 x^2 + \ldots = \sum_{n \in \mathbb{N}} a_n x^n.$$

~~Proof~~ In particular, the RHS is well-defined, i.e., the family $(a_n x^n)_{n \in \mathbb{N}}$ is summable.

Proof.

$$\begin{aligned} a_0 + a_1 x + a_2 x^2 + \ldots &= (a_0, 0, 0, 0, \ldots) \\ &\quad + (0, a_1, 0, 0, \ldots) \\ &\quad + (0, 0, a_2, 0, \ldots) \\ &\quad + \ldots\ldots\ldots\ldots \\ &= (a_0, a_1, a_2, a_3, \ldots). \end{aligned}$$

□

So we have "found" our x & made sense of writing $(a_0, a_1, a_2, \ldots)$

as $a_0 + a_1 x + a_2 x^2 + \cdots$, without using analysis.

Thus, Example 3 is justified.

(recall: it was $(1+x)^a (1+x)^b = (1+x)^{a+b}$

$\Longrightarrow$ (comp. coeff.) $\sum_{i=0}^{n} \binom{a}{i}\binom{b}{n-i} = \binom{a+b}{n}$ $\forall n \in \mathbb{N}$.)

To justify Examples 1, 2, 4, we need to know:

- what we can substitute into an FPS;
- when & why can we divide FPSs by FPSs;
- when & why can we take $\sqrt{\text{FPS}}$ and solve quadr. eqns..

So we need to do more.

8.4. Dividing FPSs

From now on, we identify each $a \in K$ with $\underline{a} \in K[[x]]$.

This is harmless, by Thm. 8.1(h).

Also, let's no longer put $\rightarrow$'s on FPSs.

Def. Let L be a CR. Let $a \in L$. Then, an (multiplicative) inverse of a means a $b \in L$ such that $ab = ba = 1$. (Note: $ab = ba$ is always true.)

Thm. 8.5. Let L be a CR. Let $a \in L$. Then, there is at most one inverse of a.

Proof. Let b and c be two inverses of a. Then,

$ab = ba = 1$ & $ac = ca = 1$.

Now, $b(ac) = b1 = b$, but $(ba)c = 1c = c$.

Thus, $b = b(ac) \underset{\text{(assoc.)}}{=} (ba)c = c$. □

Def. Let L be a CR. Let $a \in L$. Then, the (mult.) inverse of a (if it $\exists$) is called a^{-1} or $1/a$.

Also, if $b \in L$, then $b/a := b \cdot a^{-1}$.

Thm. 8.6. Let $a \in K[[x]]$. Then, a has a (multipl.) inverse $\iff [x^0]a$ has an inverse in K.

Rmk. What elements have inverses in K?

- If $K = \mathbb{Z}$, then only 1 and -1.
- If $K = \mathbb{Q}$, then all nonzero numbers.
- If $K = \mathbb{R}$, ——— // ———.
- If $K = \mathbb{C}$, ——— // ———.

$\Rightarrow \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are fields

$\left(\frac{1}{a+bi} = \frac{a-bi}{a^2+b^2} \right)$

Proof of Thm. 8.6. $\Rightarrow$: ~~Write a ...~~ Let b be an inverse of a. Thus, $ab = 1$. $\Rightarrow [x^0](ab) = [x^0](1) = 1$.

Hence, $1 = [x^0](ab) \overset{(82)}{=} [x^0]a \cdot [x^0]b$

$\Rightarrow [x^0]a$ has an inverse in K (namely $[x^0]b$). $\square$