Lecturer Notes for Abstract Algebra

Daniel K. Rui

September 6, 2025

Abstract

Abstraction requires a price. And most textbooks/courses on abstract algebra are in debt. Concrete examples/motivation, charting a Narrative beginning with counting lattice points on circles/spheres to develop basic ring and algebraic number theory, to Galois theory, to an example-dominated introduction to schemes, etc.

Contents

1 Appetizers **4**

2 The Circle **5**

2.1 Complex Numbers: Nouns and Verbs . . . 5
2.1.1 Subtle algebra-geometry interplay to develop algebra & geometry of $\mathbb{C}$ . . . 7
2.1.2 Simpler alternative: geometry of $\mathbb{C}$ assuming algebraic properties . . . 8
2.2 Pythagorean Triples . . . 8
2.3 Counting Lattice Points on Circles . . . 8
2.3.1 * Prologue/Foreshadowing (Mathologer & 3b1b video) . . . 9
2.3.2 STORY BEGINS: Base Representation, Division in $\mathbb{Z}, F[x]$ . . . 9
2.3.2.1 Euclidean algorithm, GCD, Bezout . . . 10
2.3.2.1.1 Leapfrogging (Explanation Variant) . . . 10
2.3.2.1.2 Continued fractions (Explanation Variant) . . . 11
2.3.2.1.3 PID proof of Bezout . . . 11
2.3.2.1.4 Saw division is heart of Bezout; polynomial division? . . . 12
2.3.2.1.5 Rings, Euclidean norms & division, Bezout, GCD . . . 13
2.3.3 Modular arithmetic (intro to quotients by ideals) . . . 14
2.3.3.1 Example: base-7, addition and multiplication algorithms . . . 14
2.3.3.1.1 "Blind" to multiples of 7 . . . 16
2.3.3.2 Definition of quotient ring (developing language) . . . 18
2.3.3.2.1 E.g.: indistinguishability class to identify formal syntax $\imath^2$ and -1 19
2.3.3.3 Quotient field iff maximal ideal . . . 20
2.3.3.4 Classify ideals of $\mathbb{Z}$ (abstract Bezout!) . . . 20
2.3.3.5 $\mathbb{Z}/p\mathbb{Z}$ is a field . . . 21
2.3.4 Factorization in $\mathbb{Z}, F[x]$, and $\mathbb{Z}[i]$ ("EuD $\implies$ UFD") . . . 21
2.3.4.1 Existence (irreducibility reigns supreme) . . . 22
2.3.4.2 Uniqueness (primality reigns supreme) . . . 22
2.3.4.2.1 Quotient integral domain iff prime ideal . . . 23
2.3.4.3 Irreducilibility vs. primality . . . 23
2.3.4.4 Redoing in $\mathbb{Z}[i]$ (or $F[x]$), or general PIR . . . 24
2.3.5 The Count . . . 24
2.4 Counting Lattice Points on Circles in Modulo-land . . . 25
2.4.1 Lattice points on circles mod n: geometry vs. algebra . . . 25
2.4.2 Factorization and Norm . . . 26
2.4.3 The Count (in dimension 2) . . . 28
2.4.3.1 Special (zero) fiber large $\iff \sqrt{-1} \in \mathbb{Z}/p\mathbb{Z}$! . . . 28

2.4.3.2 $p = 4k + 1$ characterization ($\Longleftarrow$): order, Lagrange's theorem 28
2.4.3.3 $p = 4k + 1$ characterization ($\Longrightarrow$): cyclicity of unit groups 29
2.4.3.3.1 Diffie-Hellman cryptographic key exchange, grand reveal 31
2.4.3.4 Concluding point count in $d = 2$ 31
2.5 Ditto on Spheres mod p, & Quadratic Reciprocity 32
2.5.1 Sample messy discovery process for $\#_p^d$ summations 32
2.5.2 * Alternative, elegant (convolution) proof of recursive formula 33
2.5.2.1 Convolution plays nice with averaging 35
2.5.3 Baby's first group action 35
2.5.3.1 * Group actions 36
2.5.4 Quadratic Reciprocity, Finale 37
2.5.5 * Unit group of $\mathbb{Z}/n\mathbb{Z}$ 38
TheCircle
2.6 §2 Algebra, Take 2: Abstraction Boogaloo 38

*** Oyster World ∼Layer 1∼** **40**

3 Galois **41**
3.1 Motivation & Set-up 41
3.1.1 Revisiting $\mathbb{C}$ 41
3.1.2 Universal property of quotient 41
3.1.2.1 * Universal property of quotients in general (sets) 42
3.1.3 Galois heart: permutations of alg. #s invisible to alg. eqs. 43
3.1.4 Formal polynomials, universal property of free ring extensions 43
3.1.4.1 Combine with previous univ. prop. for quotients 43
3.1.5 Example: $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ 44
3.1.6 Problem 1 44
3.1.7 Problem 2,3 44
3.1.8 Problem 4 44
3.1.9 Problem 5 44
3.1.10 Problem 6 44
3.1.11 Problem 7 45
3.1.12 Problem 8 45
3.1.13 Problem 9 45
3.2 Field automorphism extension/lifting 47
3.3 (Finite) cover concentration theorems 47

4 Dedekind **49**

5 Nullstellensatz (⇝Schemes...) **50**

6 Conics++ (Intro AG) **51**

7 Linear Algebra from Modules **52**

*** Oyster World ∼Layer 2∼** **53**

8 Group Theory **54**

9 Representations (Intro) **55**

10 Extras: Dirichlet **56**
10.1 Analytic (L-function, number theory) background 56

Alphabetical Index **57**

chap-1

1.0 Appetizers

Really, personally, I only care about direct mathematical motivations: how would one come up with various mathematical definitions, or various tricks/methods for proofs. But, I know many people care about things more along the lines of "applications".

Diffie-Hellman key exchange: I want to impress upon the reader just how impossible this task seems at first glance. Alice A and Bob B are separated, and EVERY communication they make is overheard by eavesdropper Eve E. How is it possible that they can agree upon a common string (key) k, that Eve doesn't know?! The only thing differentiating E from A, B is that E doesn't know the secrets of either A, B. Whereas A knows A's secrets (but not B's), and B knows B's secrets (but not A's).

The power of studying different "number systems" like $\mathbb{Z}/p\mathbb{Z}$ lies in the fact that in such number systems, there exist primitive pth roots of unity, i.e. $g \in \mathbb{Z}/p\mathbb{Z}$ that (multiplicatively) generate every non-zero number (i.e. just upon successive powers). This allows for an algebraically-very-nice [but very hard to invert!] scrambling $g^\bullet : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$.

I don't consider stuff like this to be really part of the mathematical story. I mean it's nice that abstract algebra tools can be applied in these situations, but these problems don't "*lead*" one to develop the theory, the way that the problems I'm about to present do. (The way a trail of breadcrumbs "*lead*" Hansel and Gretel to the witch's house.)

chap-2

2.0 The Circle

TheCircle

In the search for natural, universally-appreciable mathematical motivation, one can hardly do any better than arguably the first (and surprisingly ubiquitous) non-trivial pure mathematical theorem ever discovered by humankind: the Pythagorean theorem.

Already in $\approx$ 1800BCE, the Babylonians were already producing clay tablets like Plimpton 322 (the Wiki page is fascinating!), listing Pythagorean triples.

In the East, 周髀算經 (3rd century CE edition by Zhao Shuang) features a diagram of four 3-4-5 triangles in a square, essentially giving the modern visual proof.

Etc. etc. in a bunch of ancient civilizations.

Though, the Maya it seems did not develop the Pythagorean theorem, but they had other nice geometric constructions for right angles, and good calculational skills.

Pythagoras (quadratically!) relates **lengths** of any line segments to lengths of projections along any perpendicular basis of axes (our usually Cartesian coordinatization of Euclidean space), and so in particular give us a (quadratic!) algebraic equation for circles (constant **length** from one fixed point)

The problem of finding all Pythagorean triples (whose beauty was already apparent to some of the ancients) is exactly the problem of finding out which circles $x^2 + y^2 = r^2$ (which look like equally spaced concentric rings emanating from $(0, 0) \in \mathbb{R}^2$) have integer lattice points on them.

Of course, once we have an answer to whether or not there are integer lattice points on a given circle ("existence"), we can ask how many lattice points there are ("counting")

sec-2-1

2.1 Complex Numbers: Nouns and Verbs

sec-2-1
omplexNumbers

The/a "correct" viewpoint to tackle the above questions ends up coming from a breakthrough in pure algebra (see Veritasium mini-doc on history; the historical motivation is indeed for the purposes of pure algebra): the advent of complex numbers, or specifically the idea that there could be a "number" i satisfying $i^2 = 1$.

For a long time, (and even now, amongst the so-called "mathematical cranks"), people were very displeased with this notion i, which **syntatically** makes sense (as manipulations of pure algebra), but seemed to not have any **semantic** meaning. In the (aforementioned) short YouTube documentary on the subject, Veritasium pronounces these "new numbers" to "liberate algebra from geometry", and cites Decartes' terminology of "imaginary" from the 1600s (and Euler's choice of notation i in the 1700s) that persists to this day.

Centuries of this hand-wringing I think is very silly. Instead of sublimating algebra from its geometric foundation, complex numbers actually introduce the geometry of rotation to arithmetic,

which is the major power of complex numbers that we use to tackle the lattice point counting problems I introduced above.

Indeed, here is one way to introduce complex numbers (a) in a way that makes the geometry take center stage from the very beginning, and (b) concrete and semantically meaningful for even the most ardent of "cranks"/"haters" adverse to mathematical abstraction.

The key idea is to think of addition/multiplication not as something based on two "nouns" (e.g. $+$ acting on numbers 2 and 3), but rather as a "verb" (e.g. $2+$) acting on a "noun" (e.g. 3). Let's start the very basics: the integers $0, 1, 2, \ldots$, drawn say as dots on a line (evenly spaced, i.e. the label of the dot corresponds with the **length/distance** from 0). Those dots will be the "objects"/"nouns". (I guess it's not the complete basics, since I include 0 — very crucial for this whole discussion, since it serves as the reference point — and 0 took a while for humans to really grasp; but still, I think it's intuitive enough to teach to children: we can count[associate a pure mathematical concept to] things that get placed in front of us, but if there's nothing in front of us, we would still like to associate a pure mathematical concept, rather than just having no answer/blank stare.) Draw this (half) number line, and label say the dot/number 3.

Now for the "actions":

- The "verb" $2+$ (addition by 2) is just shift over by 2 units, i.e. **(rigid) translation** that moves (the dot) 0 to (the dot) 2. This whole operation moves (the dot) 3 to (the dot) 5, which is of course where we get $2 + 3 = 5$.

- The "verb" $2\ast$ (multiplication by 2) is just dilation by a factor of 2, i.e. **(rigid) dilation** that keeps (the dot) 0 fixed, and moves (the dot) 1 to (the dot) 2. This whole operation moves (the dot) 3 to (the dot) 6, which is of course where we get $2 \cdot 3 = 6$.

- What if we want to shift left instead of right? (Like in 5 subtract 1.) That would correspond to moving 0 to a dot 1 unit to its left, which we haven't drawn yet. So, we draw new dots to the left of 0 — the negative numbers!

- Then multiplication by e.g. -2 would fix 0 and move (the dot) 1 to (the new dot) -2, which we can think of dilation by factor of 2 and then reflect/flip of the number line around 0 (these 2 operations commute, i.e. can be done in either order).

  Illustrate this in front of the class by drawing the number line on the board, and holding one fist at 1 and one fist at -1, and then crossing your arms to show the transformation.

- The last step is to realize there's another way to think about negation (i.e. "verb" $(-1)\ast$). The reflect/flip/mirror image is very discrete. You can not do it halfway. But, on our numberline, reflection/flip is the same as rotation by 180°!

  Hype this moment up, keep interchanging fists at 1 and -1 by crossing arms (keeping everything along the numberline on the board). Then (I'm imagining accompanied by Stranger Things music as the camera rotates lol) dramatically start moving fists into the 2nd dimension in a rotation.

  Rotation is now something you *can* do halfway. So to achieve result of $(-1)\ast$, you can do 90° rotation twice. That is $i\ast$ (counterclockwise by convention; the choice does not matter).

  By analogy with previous cases, add new dot i above 0 so that 90° (counterclockwise) rotation (fixing 0) moves 1 to that new dot i. And by design of course, the "verb" $i\ast$ acting on the "noun" (the dot) i moves it to (the dot) -1; i.e. $i \cdot i = -1$!

I.e. we are thinking of "object"/"noun" i as vector $(0,1)$ in the $\mathbb{R}$-VS $\mathbb{R}^2$, and we are thinking of $i\star$ as the 90° counterclockwise rotation $\in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$ (the $\mathbb{R}$-linear transformations $\mathbb{R}^2 \to \mathbb{R}^2$).

2.1.1 Subtle algebra-geometry interplay to develop algebra & geometry of $\mathbb{C}$

If we define the syntax ("pure algebra"/"just symbols") $\{a+bi : a,b \in \mathbb{R}\}$ and obvious addition operation (i.e. standard $+$ in $\mathbb{C}$), then associating $a+bi \leftrightarrow$ vector $(a,b) \in \mathbb{R}^2$ gives a semantic "model" ("geometry") of this syntax.

One can then prove theorems, like $(a+bi)+ \ \in \mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^2)$ (defined as $[v \mapsto v+(a,b)]$) equals $(a+)+(bi+)$ as elements of the vector space $\mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^2)$.

[[[Draw a plane, so that there's a "dot"/"object" for every pair of x and y coordinates $(a,b) \in \mathbb{R}^2$, labelled as $a+bi$. The addition "verb" $(a+bi)+$ remains translation (rigid translate 0 to the dot $a+bi$).]]]

From what we said above about multiplication with real numbers and with i, we know we want to associate

- $a\star \leftrightarrow \left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right] \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$ for $a \in \mathbb{R}$ (dilation by a factor of $|a|$, and a rotation by 180° if $a<0$; again can be done in either order),
- $i\star \leftrightarrow \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$.

Then defining $(a+bi)\star := (a\star)+(bi\star)$ (by linearity), i.e. corresponding to $\left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0 & -b \\ b & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right] \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$, this becomes a semantic "model" for a

- syntactic multiplication operation on the above syntax/pure algebra $\{a+bi : a,b \in \mathbb{R}\}$ (previously equipped with addition):
- denoted $(a+bi)\cdot(c+di)$
- and defined by
  - distributativity
  - and $(i\star)$ commutes with multiplication by real numbers $\iff$ $(i\star)$ is $\mathbb{R}$-linear
  - and $i^2 = -1$

since linearity $\iff$ distributivity $(a+bi)\cdot z = a\cdot z + bi\cdot z$.

(In particular, this semantic model shows that the syntactic definition based on distributativity is consistent, i.e. does not result in logical contradictions.)

The beautiful thing about having both an algebraic and geometric understanding is that certain things are easier to do in one compared to the other.

- For instance, commutativity is easy on the algebraic side (just use distributivity and expand out the product written in both orders, use commutativity in $\mathbb{R}$ and that $(i\star)$ commutes with multiplication by real numbers $\iff$ $(i\star)$ is $\mathbb{R}$-linear, and match terms), and actually somewhat surprising/quite non-obvious from the geometric viewpoint: e.g. ["verb" $(a+bi)\star$ acting on "noun" i] is the same result as ["verb" $i\star$ acting on "noun" $a+bi$]!

  Here is a more detailed argument: $(a+bi)\cdot(c+di)\cdot 1 \leftrightarrow \left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]\left[\begin{smallmatrix} c & -d \\ d & c \end{smallmatrix}\right]\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, i.e. indeed $(a+bi)\star \circ (c+di)\star \ \leftrightarrow$ matrix multiplication! And the fact that in syntax-land, $(a+bi)\cdot(c+di) = (c+di)\cdot(a+bi)$, translates to the fact that the matrices must commute!

- On the other hand, associativity is easy on the geometric side, since matrix multiplication (i.e. function composition) is associative; and is quite tedious on the algebraic side! How beautiful, how balanced.

One more thing: inverse of multiplication in $\mathbb{C}$ (i.e. $\mathbb{C}$ is a field). We still do not need to do any explicit computation! We mix both algebra and the linear transformation perspectives ("geometric"): for $a+bi \neq 0$, because $\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$ is invertible (method (i): determinant $= a^2+b^2 \neq 0$; or method (ii): columns are orthogonal — geometry!), there is a solution to $\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, which shows that $\exists\, x+iy$ s.t. $(a+bi)\cdot(x+iy) = 1$.

2.1.2 Simpler alternative: geometry of $\mathbb{C}$ assuming algebraic properties

Above is a complete breakdown of the geometry and algebra of $\mathbb{C}$. If we're willing to take for granted the algebraic properties of $\mathbb{C}$ and accept that i is 90° counterclockwise rotation [e.g. in a class in which the above admittedly subtle discussion wouldn't go over well], here's another explanation for the geometry of the "verb" $(a+bi)*$.

- We know it fixes 0, and moves 1 to $a+bi$. By ($\mathbb{R}$-)linearity ("**geometry**") $\iff$ distributativity of multiplication ("**algebra**"), it suffices to know where it moves i.
- But the ["verb" $(a+bi)*$ acting on "noun" i] is the same result as ["verb" $i*$ acting on "noun" $a+bi$]! I think this is actually somewhat surprising/quite non-obvious from the geometric viewpoint; but of course it's literally "just" commutativity on the algebraic side.
- To conclude, "verb" $i*$ acting on "noun" $a+bi$ is by design is 90° (counterclockwise) rotation applied to $a+bi$, so in total, $(a+bi)*$ rotates and dilates the unit square (with the "0's counterclockwise side" being the line segment between 0 and 1) to a square with "0's counterclockwise side" being the line segment between 0 and $a+bi$.

sec-2-2

2.2 Pythagorean Triples

sec-2-2

Now that we understand the algebra and geometry of $\mathbb{C}$, can watch 3b1b video on Pythagorean triples. Geometry fact inscribe angle in circle motivates considering lines through $(-1,0)$. That I think is quite a clever leap.

Already in this ancient problem, we see a $\mathbb{P}^1(\mathbb{Q})$ parameterization of rational points (literal bijection with lines!), which foreshadows a lot of the theory we develop later for algebraic geometry.

sec-2-3

2.3 Counting Lattice Points on Circles

sec-2-3
intsOnCircles

Following the path I freestyled in 110A discussion 1. First, problem about (approximately) how many lattice points in a disk $D(0,R)$. [Try to get students to figure it out. Keyword: "area".] Mention Gauss circle problem, to emphasize that even simple questions are still open in math. And mention partial progress, to emphasize that in math and in life, just because we can't solve a problem completely doesn't mean there aren't still things to do — partial progress is still progress.

Then zoom in 1 dimension lower to counting lattice points on circle with equation $\texttt{C}_a \coloneqq [\texttt{x}^2+\texttt{y}^2=a]$. [Ask students to guess if they think problem is easier or harder. For 2D problem, we only have approximation. Do we have exact formula for 1D problem?]

Ask students to graph on say Desmos and start counting! Fill out table (up to 8, or 12, etc.):

a	1	2	3	4	5	6	7	8	9	10	11	12
$\#(\mathtt{C}_a)$	4	4	0	4	8	0	0	4	4	8	0	0

Everything is multiple of 4 because 4 symmetric quadrants of graph. **Maybe** with more data (emphasizing the sometimes experimental nature of mathematical discovery!), notice the pattern that 0 whenever $a \equiv 3 \pmod 4$ (3 more than a multiple of 4). Maybe students also notice $a = 5, 10$ both having 8 is a little suspicious...

2.3.1 * Prologue/Foreshadowing (Mathologer & 3b1b video)

In teaching, the circle lattice point problem is a "prologue", serving as an introduction, before the STORY begins in earnest. I think BEGIN with bases is a pretty good starting point, but I do think it's a good idea for an abstract algebra course to begin (prologue) as anchored as possible.

It is up to the discretion of the teacher to decide if it is better to show Mathologer Fermat Christmas theorem video and 3b1b's π hiding in primes video now (making sure to mention unique factorization blackbox!). I personally think it's a good idea to show these videos early, and then really dig in and revisit them later as well.

In particular, we will give independent proofs of the mod 4 dichotomy (factorization) theorem (using the language of primitive roots in $\mathbb{Z}/p\mathbb{Z}$), so if one was bothered by the "ad hoc" nature of the "ah, look mod 4 to rule out 3 (mod 4)" and Mathologer's windmill proof (though I think it's always good to keep in mind the OG geometric meaning of multiplication as areas of rectangles!!!), then one can rest assured a more Narratively satisfying proof is ahead.

2.3.2 STORY BEGINS: Base Representation, Division in $\mathbb{Z}, F[x]$

In order to make "3 more than a multiple of 4" immediately visually obvious, I want to write things in base-4. But the algorithm for writing things in base-4 is to do division with remainder algorithm, so let me review this (with an eye towards future generalizations) in the concrete settings of $\mathbb{Z}$ and $F[x]$, since we are familiar with those (the former from elementary school, and the latter from secondary school algebra — I mean I'm sure people have got to remember factoring at least quadratics... please...).

In $\mathbb{N}$ or $\mathbb{Z}$ or $F[x]$, given a and q, it may not be true that there is m s.t. $a = mq$, but it could be that there is m s.t. $a \approx mq$, i.e. $a = mq + r$ for some remainder r that we hope is "small". Like it's true that you can't split 184 cookies evenly amongst 13 people, but you can still give everyone 14 cookies and have 2 left over so things are as "fair as possible". [Prove division algo for $\mathbb{Z}$ and $F[x]$ is correct, in a way that is easy to generalize to $\mathbb{Z}[i]$.]

Base-10. Why do we have bases anyways? Can think about counting using tally marks. Lots of tally marks gets confusing, so we group tallies together, like how we learn in America for groups of 5, or the Chinese character 正. Then as there are many of those, then we want to group those together as well. In other words, "because unary is fundamentally extremely unwieldy", we naturally form these logarithmic/exponential hierarchies in representing numbers, which is what essentially leads to exactly the base system.

The choice of 10 is arbitrary, so let me go over this. How do we write a number in a certain base b? (Division algorithm)

th:divAlgoZ

Theorem 2.3.1: Division algorithm in $\mathbb{Z}$

For any $a, b \in \mathbb{Z}$, there is $q, r \in \mathbb{Z}$ s.t. $b = aq + r$ AND $0 \leqslant r < |a|$.
(In the $\mathbb{Z}$ case, we do have uniqueness.)

Using this base-b representation, let me remind people of the algorithms for addition and multiplication we know from childhood — or I guess that comes later in §2.3.3. Ugh, it seems like those algorithms would pack a better punch here earlier; but then that loses the train of thought on division ⇝ Euclidean algo. and Bezout below.

modularArithmetic

(One good thing about starting with this base system is it is essentially the chain $\ldots \twoheadrightarrow \mathbb{Z}/p^2\mathbb{Z} \twoheadrightarrow \mathbb{Z}/p\mathbb{Z}$, cf. 8/6/25 thoughts on zooming in/localization and p-adics)

2.3.2.1 Euclidean algorithm, GCD, Bezout

bezout

I present 2 different ways of teaching the Euclidean algorithm: leapfrogging division algorithm, and continued fractions.

2.3.2.1.1 Leapfrogging (Explanation Variant)

Can think of division as having say 2 points $0 < a < b$ on number line, and division is moving b "downward" ("small-wards") by increments of size a until right before it passes 0. So can picture division as "leapfrogging" b over a (by increments of size a).

So whenever see $0 < a < b$ (positive numbers, one on the left, one on the right), we should have an instinct to bounce our right hand over our left hand. But now again we have 2 positive numbers (unless previously one had cleanly divided the other), one on the right, one on the left, so our instinct should beat in our chest again.

This too (like below continued fractions) is very natural to iterate. It is clear that every point one can reach by this leapfrogging process must be a linear combination of the original a, b, i.e. inside $\langle a, b\rangle \subseteq \mathbb{Z}$ (which later when we develop the terminology, we will recognize as an "ideal of the ring $\mathbb{Z}$").

def:idealBracket

Definition 2.3.2: Ideal bracket / span

If $S \subseteq \mathbb{Z}$ (or more generally any ring R), then $\langle S\rangle := \mathrm{span}_R(S) :=$ {all finite R-linear combinations of elements in S}.

Process only ends when $0 < a' < b'$ and b' leapfrogs over a' to actually hit 0, i.e. $a' \mid b'$.
Do lots of examples, pictures.
Looking left to right, we see that everything is a multiple of a', i.e. everything in $\langle a'\rangle$.
Looking right to left, we see that everything is a linear combination of a, b, i.e. everything in $\langle a, b\rangle$.
So $a, b \in \langle a'\rangle$ and $a' \in \langle a, b\rangle$, so $\langle a, b\rangle = \langle a'\rangle$.
Can prove this last "so" statement directly easily, but let us distill out the following general lemmas:

lem:spanMonotone

Lemma 2.3.3: Span monotonicity

If $S, T \subseteq \mathbb{Z}$ (or more generally any ring R), then $\langle S\rangle \subseteq \langle T\rangle$.

n:spanIdempotent

Lemma 2.3.4: Span idempotence

If $S \subseteq \mathbb{Z}$ (or more generally any ring R), then $\langle\langle S\rangle\rangle = \langle S\rangle$. "Linear combinations of linear combinations are still just linear combinations."

2.3.2.1.2 Continued fractions (Explanation Variant)

~~Briefly mention continued fractions and connection with iterating Euclidean algorithms? Cool topic, but a little not relevant.~~ It's actually amazingly relevant.

There is a familiar notation that goes hand-in-hand with the division algorithm: improper fractions. Given a fraction $\frac{b}{a}$ (positive numbers, let's say), can always convert it to $q\frac{r}{a} = q + \frac{r}{a}$ where $(0 \leqslant)r < a$. So it's a pipeline that converts all top-heavy fractions (fractions abs. val. > 1) to bottom-heavy fractions (fractions abs. val. > 1), maybe with a $[\ldots]+$ in front.

This aligns very nicely with the **symmetry** of the fraction notation $\frac{\bullet}{\bullet}$, since top-heavy fractions are the same as bottom-heavy fractions, but reflected across the middle horizontal line of symmetry, which corresponds with the operation of "reciprocal"/"one-over"! (This highlights the importance of good notation; I don't think this connection would be so clear using say the decimal notation for fractions.)

So, by playing these 2 operations against each other (going from top-heavy to bottom-heavy by division algorithm, back to top-heavy by "one-over"), we can now **iteratively** apply the division algo.! In general, once one can do something non-trivial once, it's always a good idea to see (what happens) if you can do it again (cf. Goofy "I'll fuckin do it again" meme). (Most memorable example in recent memory is 223D apply analytic separation *twice*.) (Also, generally speaking interesting things happen upon iteration. See Mandelbrot set, or really rather the fantastic videos by 3b1b on Newton fractals.)

Continued fractions. When do we stop? When we get to a remainder of 0, i.e. get to a top-heavy fraction $\frac{b'}{a'}$ s.t. a' divides b', i.e. $b' \in a'\mathbb{Z} = \{a'm : m \in \mathbb{Z}\}$ (after seeing $\langle a, b\rangle$ notation below, we notice this follows that pattern, so write $=: \langle a'\rangle$).

Unwinding starting from the end/the very bottom, notice that everything stays in $a'\mathbb{Z}$ — in particular, a, b from very beginning $\in a'\mathbb{Z}$.

Unwinding starting from the top/the very beginning (I painstakingly illustrate on ME.SE here), notice everything stays in $\{am + bn : m, n \in \mathbb{Z}\} =: \langle a, b\rangle$ (all linear combinations of a, b, i.e. all sums of scalar multiples — remember this wording, it'll be important later for "ideals") — in particular, a' from the very end $\in \langle a, b\rangle$.

Conclude $\langle a'\rangle = \langle a, b\rangle$.

[Show some beautiful infinite continued fractions to end the discussion!]

2.3.2.1.3 PID proof of Bezout

Above, Bezout was proved by looking at one thing from 2 different directions. But quite fiddly, really dependent on understanding what happens at each step of the iterated algorithm.

Let us try to find a different proof, that doesn't need us to micromanage each step. Previously, we found a' by going step by step, getting smaller step by step until hit 0 (at which point we can't get smaller, i.e. we are at smallest possible). So instead of thinking of a' as the smallest positive element reached by our process (which depends on really understanding the fiddly process), let's just think of

a' as the smallest positive element reachable period (by linear combinations of a, b). We know this to indeed to be true (so we don't have to worry about barking up completely the wrong tree) because we already know $a'\mathbb{Z} = \langle a' \rangle = \langle a, b \rangle$).

prop:bezoutZ

Proposition 2.3.5: Abstract Bezout in $\mathbb{Z}$

For $a, b \in \mathbb{Z}$, there is $m \in \mathbb{Z}$ s.t. $\langle m \rangle = \langle a, b \rangle$
(We prove directly that $m := \min(\langle a, b \rangle \cap \mathbb{N}^+)$ works.)

Proof: one direction is trivial: $m \in \langle a, b \rangle \implies \langle m \rangle \subseteq \langle a, b \rangle$.

Now we want to show $a, b \in \langle m \rangle$. Division algorithm says $a = mq + r$ for $0 \leqslant r < m$. But $r = a - mq \in \langle a, b \rangle$, and $< m$ forces $= 0$ by minimality of m. Yay! ■

A quick reformulation (the more classical version of Bezout, taking advantage of people's possible familiarity with GCD in $\mathbb{Z}$):

prop:bezoutZgcd

Proposition 2.3.6: Bezout in $\mathbb{Z}$, gcd form

The minimal element m above (satisfying $\langle a, b \rangle = \langle m \rangle$) is actually $m = \gcd(a, b)$.
I.e. $\langle a, b \rangle = \langle \gcd(a, b) \rangle$.

Proof: prove the "cd" part, and then the "g" part. Common divisor c of a, b same as $a, b \in c\mathbb{Z} \iff \langle a, b \rangle \subseteq c\mathbb{Z}$. So in fact m is greatest common divisor (since $m \in \langle a, b \rangle \subseteq c\mathbb{Z}$ — m is a $\geqslant 1$ multiple of c — for all $c \mid a, b$).

2.3.2.1.4 Saw division is heart of Bezout; polynomial division?

The benefit of proving Bezout again but in the new way above, is that now is that we understand much clearer that the only tool we used was the division algorithm, no fiddly iterated process, no dependence on the specific geometry/picture of the number line $\mathbb{Z}$ and leapfrogging.

Perhaps the reader may remember (from middle school algebra, say) that polynomials also admit a division algorithm. It is very similar to division for integers we know from elementary school, and in fact in some sense easier:

- in $\mathbb{Z}$, dividing $52 = 5(10) + 2$ into $4793 = 4(10^3) + 7(10^2) + 9(10^1) + 3$ goes like: $52$00 doesn't go into 47−−, but $52$0 goes into 479− a maximum of 9 times with $479$3 − $468$0 = $011$3 leftover (in particular, the leftmost numbers, 4 in this case, cancel leaving 0 in leftmost place); iterating one step to the right, 52 goes into 113 a maximum of 2 times with $113 - 104 = 9$ leftover $\rightsquigarrow 4793 = 52 \cdot 92 + 9$.
- in $\mathbb{Q}[x]$, dividing $5x+2$ into $4x^3+7x^2+9x+3$ goes like: $(5x+2)x^2$ goes into $4x^3+7x^2+(-)x+(-)$ exactly $\frac{4}{5}$ times with $4x^3 + 7x^2 + 9x + 3 - (4x^3 + \frac{8}{5}x^2) = 0x^3 + \frac{27}{5}x^2 + 9x + 3$ leftover (in particular, the leftmost terms, $4x^3$ in this case, cancel leaving 0 in leftmost place); iterating one step to the right, ...

I.e. in polynomial division, all you have to do is match the highest order term, so that the leftover at that step has 0 in leftmost place. This is a necessary part of division in $\mathbb{Z}$ as well, but not sufficient (e.g. both $52$0 $\cdot 9$ and $52$0 $\cdot 8$ both yield leftovers $479$3 − $468$0 = $011$3, $479$3 − $416$0 = $063$3 with 0 in leftmost place; but 9 is correct and 8 is incorrect). Polynomials ("base x") also have no "carrying", unlike integers base b.

Anyways, having recalled polynomial division in $\mathbb{Q}[x]$ (or indeed any $F[x]$ for F a field), (Exc!) go back through the above Lemmas and Propositions and see that they continue to work in $F[x]$ in place of $\mathbb{Z}$.

We do have to make one change: before in the statement of the division algorithm for "b/a" Thm. th:divAlgoZ 2.3.1, we say remainder r has "strictly smaller size" than that of a. What is "size" for polynomials? Thinking through polynomial division, we see that the appropriate notion is "degree".

th:divAlgoFx

Theorem 2.3.7: Division algorithm in $F[x]$

For any $a, b \in F[x]$, there is $q, r \in F[x]$ s.t. $b = aq + r$ AND $\deg(r) < \deg(a)$.

exc:divFailZx

Exercise 2.3.8: Division algorithm fails in $\mathbb{Z}[x]$

I claim polynomial division fails in $\mathbb{Z}[x]$. Find explicit $a, b \in \mathbb{Z}[x]$ demonstrating this.

A careful analysis of the "minimal element" proof of Bezout reveals that we don't really have to worry about *positivity*; indeed, taking $m :=$ the minimal *non-zero* element of $\langle a, b \rangle$, the same proof goes through.

exc:bezoutFx

Exercise 2.3.9: Check Bezout, GCD for $F[x]$

Check that Prop. prop:bezoutZ 2.3.5 and Prop. prop:bezoutZgcd 2.3.6 continue to hold for $F[x]$ in place of $\mathbb{Z}$.

2.3.2.1.5 Rings, Euclidean norms & division, Bezout, GCD

Let us now give abstract framework encapsulating the above arguments/syntactic maneuvers. Think of this as proving the results for $\mathbb{Z}$ and $F[x]$ at the same time. (In general, it is helpful to think about abstraction/abstract math as a uniform method to deal with many concrete examples all at once. It is easy to get suckered down a rabbit hole, thinking "oh, should we get rid of that axiom too, to get the "best possible" definition?", but for us, let us not worry about optimization, but rather just trying to tweak known definitions (like fields) as little as possible to capture 2 specific concrete examples and the proofs above.)

Recall field: 2 columns $+, \cdot$; 5 rows (closure, commutativity, associativity, identity, inverse) + distributativity tying the 2 columns together.

def:cring

Definition 2.3.10: Cring

From the axioms of a field, removing multiplicative inverse gives us a list of axioms satisfied by $\mathbb{Z}, F[x], \mathbb{Z}[i]$.

Exercise 2.3.11: Check bracket notation works in any ring

Check that Lemma lem:spanMonotone 2.3.3 and Lemma lem:spanIdempotent 2.3.4 continue to hold in any ring R.

def:EuR

Definition 2.3.12: Euclidean ring

Given a ring R, suppose we have a notion of norm $|\bullet| : R \setminus \{0\} \to [0, \infty)$ s.t. $\forall a, b \in R$, $\exists q, r \in R$ s.t. $b = aq + r$ AND $|r| < |a|$. Then we say R is an Euclidean ring.

exc:bezoutEuR

Exercise 2.3.13: Check Bezout, GCD in any Euclidean ring

Check that Prop. prop:bezoutZ 2.3.5 and Prop. prop:bezoutZgcd 2.3.6 continue to hold for any Euclidean ring R in place of $\mathbb{Z}$.

2.3.3 Modular arithmetic (intro to quotients by ideals)

larArithmetic

Maybe can skip this part? If the base-b Story is compelling enough, then don't need the motivation from the Prologue for this modular arithmetic section. Else (if the Story is not compelling), then good idea to have this concrete "3 mod 4 impossibility result" application first.

So, we have a conjecture: no integer solutions to $x^2 + y^2 = a$ where a is any number that is 3 more than a multiple of 4. I think the best way to think about this at this early stage, is to represent all numbers in base-4 (so assumption on a is immediately visually apparent: the RHS is $[\cdots 3]$), and to realize that the addition and multiplication algorithms we know from childhood work equally well over base-10 as any other base.

[Maybe this would be a good time to review why those algorithms work in the first place... lol.]

In particular, to compute just the last (i.e. rightmost) digit of the answer, it suffices to know the rightmost digits of the inputs!

[Prove the conjecture, using the below visualization of addition and multiplication algorithms in base-4, boxing the rightmost digits.]

2.3.3.1 Example: base-7, addition and multiplication algorithms

Let us return to thinking about numbers in general bases. A key observation is that the addition and multiplication algorithms we learn in elementary school work equally well over base-10 as any other base! Here's a base-7 example $234234636 \cdot 340653562$:

$$
\begin{array}{rcccccccc}
 & 2 & 3 & 4 & 2 & 3 & 4 & 6 & 3 & \boxed{6} \\
\times & 3 & 4 & 0 & 6 & 5 & 3 & 5 & 6 & 2 \\
\hline
 & {\scriptstyle 1}5 & {\scriptstyle 1}0 & 1 & {\scriptstyle 1}5 & {\scriptstyle 1}0 & {\scriptstyle 1}2 & {\scriptstyle 1}6 & {\scriptstyle 1}0 & \boxed{5} \\
 & & & & & & \ldots & {\scriptstyle 5}0 & 1 & \\
 & & & & & & & \vdots & &
\end{array}
$$

I.e. on the set $[7] := \{0, 1, 2, 3, 4, 5, 6\}$ we can define an addition and multiplication operation by just taking regular addition and multiplication (in base-7) and taking the last digit, as in the box above.

exc:Z7Ztables

Exercise 2.3.14: "Last column" addition and multiplication table for [7)

Write out the "last column in base-7" addition and multiplication operation tables for [7). Notice any interesting patterns? Like every row/column has a permutation of the digits $0, 1, 2, 3, 4, 5, 6$? Does this hold for other b other than 7? Try to do the same things for $b = 2, 3, 4, 5, 6, 8, 9, 10$!

Essentially by construction, the 'last digit map' $\mathbb{N} \to [7)$ preserves addition and multiplication (like in a ring homomorphism way).

Actually, let's actually invent some new words for this phenomenon. We say $\mathbb{N}$ has am-structure (addition multiplication structure), or even aumu-structure (addition with additive unit 0, and multiplication with multiplicative unit 1).

And in the previous paragraph, we also gave [7) aumu-structure, and in fact, aimu-structure (addition with additive inverses — i.e. negation, where you add something to its negation to get additive unit 0, and multiplication with multiplicative unit 1).

And the 'last digit map' $\mathbb{N} \to [7)$ is an am-homomorphism (respects/preserves addition and multiplication), or even an aumu-homomorphism (respects/preserves addition with additive unit 0, and multiplication with multiplicative unit 1).

Going back to the above addition/multiplication algorithm with boxed rightmost digit, we have observed that: $[\cdots 6]$ (anything in base-7 ending with 3) times $[\cdots 2]$ equals $[\cdots 5]$ (something in base-7 ending with 5). Which in the language of **sumsets** and **sumset multiplication** says that

$$(6 + 7\mathbb{N}) \cdot (2 + 7\mathbb{N}) \subseteq (5 + 7\mathbb{N}).$$

def:setAddMult

Definition 2.3.15: Set addition and multiplication

For sets A, B (in a context where we can add elements — e.g. a (c)ring or VS), $A + B \coloneqq \{a + b : a \in A, b \in B\}$. When $A \coloneqq \{a\}$ is a singleton, often abbreviate $\{a\} + B$ as $a + B$.

For sets A, B (in a context where we can multiply elements — e.g. a (c)ring), $A \cdot B \coloneqq \{a \cdot b : a \in A, b \in B\}$.

mp:setAddMult7N

Computation 2.3.16: Set add. and mult. with $7\mathbb{N}$

Exc! Check the following set equalities and containments: $(6 + 7\mathbb{N}) \cdot (2 + 7\mathbb{N}) = 6 \cdot 2 + 7\mathbb{N} \cdot 2 + 6 \cdot 7\mathbb{N} + 7\mathbb{N} \cdot 7\mathbb{N} \subseteq 12 + 7\mathbb{N} + 7\mathbb{N} + 7\mathbb{N} \subseteq 5 + 7\mathbb{N}$.
Note: we used that $7\mathbb{N}$ multiplied by other natural numbers $\subseteq 7\mathbb{N}$ ("closed under scalar multiplication") and $7\mathbb{N} + 7\mathbb{N} \subseteq 7\mathbb{N}$ ("closed under addition").

Think of as saying "6 (up to ambiguity of multiple of 7) times 2 (up to ambiguity of multiple of 7) is equal to 5 (up to ambiguity of multiple of 7)".

It's a lot like continuity: addition in $\mathbb{R}$ being continuous means something like "adding 2 inputs who are determined up to ambiguity on the right of the 13th decimal place, will give a result determined up to ambiguity on the right of the 11th decimal place".

Whereas here, it's addition in $\mathbb{Z}$ of 2 inputs who are determined up to ambiguity on the left of the "10s" place base 7, will give a result determined up to ambiguity on the left of the "10s" place base 7.

2.3.3.1.1 "Blind" to multiples of 7

Another way of thinking about this is that we are working "blind" to multiples of 7; i.e. "looking at right-most digit/column" (i.e. looking at the world ($\mathbb{N}$ in this case, I guess) "through π goggles", i.e. the world goes through the function π before reaching our eyes) is the same thing as "not being able to see digits left of the rightmost digit".

So the numbers $12, 5$ (in base 10), "viewed through π", "look" exactly the same — namely "look" like 5, because in base 7, $12_{10} = 15_7$. This is of course because $12 \in 5 + 7\mathbb{N}$ (= all the (positive) integers that look like $[\dots 5]$ in base 7).

In general $a, b \in \mathbb{N}$ are indistinguishable "viewed through π" iff $a - b \in 7\mathbb{N}$ ($\iff a \in b + 7\mathbb{N}$ by (set-)adding b to both sides) or $b - a \in 7\mathbb{N}$ ($\iff b \in a + 7\mathbb{N}$). These conditions of course differ only by sign, and so a cleaner formulation which respects the negation symmetry between $a - b$ and $b - a$ is to phrase in terms of $\mathbb{Z}$ ("$\mathbb{N}$ made symmetric under negation"): $a, b \in \mathbb{N}$ are indistinguishable "viewed through π" iff $a - b \in 7\mathbb{Z}$ (which is $\iff b - a \in 7\mathbb{Z}$).

comp:7Nmanip

Computation 2.3.17: Element differences, set addition manipulations & equivalences in $7\mathbb{N}$

Extract out general equivalences from above paragraph: for all $a, b \in \mathbb{Z}$,

$$a \in b + 7\mathbb{Z} \iff a - b \in 7\mathbb{Z} \iff b - a \in 7\mathbb{Z} \iff b \in a + 7\mathbb{Z},$$

where the only thing we used about $7\mathbb{Z}$ was closure under negation.

Perhaps this feels a bit reminiscent of the "$m \in \langle a, b \rangle, a, b \in \langle m \rangle$" that led us to Lemmas 2.3.3 lem:spanMonotone and 2.3.4 lem:spanIdempotent; indeed something very analogous happens here (+7ℤ to both sides instead of ⟨•⟩):

$$\{a\} \subseteq b + 7\mathbb{Z} \implies \{a\} + 7\mathbb{Z} \subseteq b + 7\mathbb{Z} + 7\mathbb{Z} \subseteq b + 7\mathbb{Z},$$

(and of course the same swapping a, b; $\rightsquigarrow$ conclusion $a + 7\mathbb{Z} = b + 7\mathbb{Z}$); where the only thing we used about $7\mathbb{Z}$ was closure under addition.

Note that $a + 7\mathbb{Z} = b + 7\mathbb{Z}$ implies $a + 0 \in a + 7\mathbb{Z} = b + 7\mathbb{Z}$ (and with a, b swapped) — where we use that $7\mathbb{Z}$ contains 0) — so "$[a + 7\mathbb{Z} = b + 7\mathbb{Z}]$" is also equivalent to the 4 equivalent things in the 1st centered-display of this box.

Working with $\mathbb{Z}$ instead of $\mathbb{N}$ makes the relationship between 5 and 12 much more balanced, as it should be — "indistinguishable" should be a symmetric notion.

And these sumsets $5 + 7\mathbb{Z}$ and $12 + 7\mathbb{Z}$ being "all the numbers (integers, now) we can't distinguish from 5 and 12 resp. when wearing our π goggles" being exactly equal (unlike $5 + 7\mathbb{N} \supsetneq 12 + 7\mathbb{N}$) is harmonious with the fact that we can't distinguish 5 from 12 when wearing our π goggles.

I'm spending a lot of time on this transition from $\mathbb{N}$ to $\mathbb{Z}$. Perhaps this is silly, but my rational is that it is very natural to start this story from the perspective of the familiar addition and multiplication algorithms in $\mathbb{N}$ (but with arbitrary bases, which I find to be a good starting point philosophically too since we should always question arbitrary choices like our convention of base-10); however algorithmically we don't really work with negative numbers (e.g. when we multiply -317 with 42, we just take the negative sign out, do multiplication of positive integers, and put back the minus sign at the end), so I

have to pay the cost of a weird transition from $\mathbb{N}$ to $\mathbb{Z}$ (and then to abstract rings and ideals).

These more robust "indistinguishability classes" $0+7\mathbb{Z}, 1+7\mathbb{Z}, \ldots, 6+7\mathbb{Z}$ partition $\mathbb{Z}$ (distinct classes are disjoint exactly because of Comp. 2.3.17!), and the set addition and set multiplication operations of Comp. 2.3.16 still hold with $7\mathbb{Z}$ in place of $7\mathbb{N}$:

$$(6+7\mathbb{Z})\cdot(2+7\mathbb{Z}) \subseteq (5+7\mathbb{Z}) \quad \text{(in fact an equality — but this isn't important)};$$

"anything indistinguishable (a symmetric notion) from 6 times anything indistinguishable from 2 is indistinguishable from 5".

In other words, because of the special properties (isolated below, Def. 2.3.19) of $7\mathbb{Z}$ noted in Comp. 2.3.16, Comp. 2.3.17, set addition and set multiplication of 2 of these "indistinguishability classes" lands in another (unique!) such "indistinguishability class", thus forming in fact a ring structure (same as Exc. 2.3.14) on the "indistinguishability classes" (the ring elements are *entire sets*)! This is called the quotient of cring $\mathbb{Z}$ by ideal I. I.e. key idea is to see "$(6+7\mathbb{Z})\cdot(2+7\mathbb{Z}) \subseteq (5+7\mathbb{Z})$" and think of these $(a+7\mathbb{Z})$ as singular/individual objects (called "cosets"), instead of as a set/collection of many distinct individual elements — as "one-thing", not a collection of "many-things".

c:cosetZ7Ztables

Exercise 2.3.18: Addition and multiplication tables for $0,1,2,3,4,5,6+7\mathbb{Z}$

Write out the addition and multiplication operation tables for $0,1,2,3,4,5,6+7\mathbb{Z}$ (i.e., w.r.t. set addition and set multiplication operations described above). Check that it is the same as Exc. 2.3.14.

$$\begin{aligned}\mathbb{Z}/5\mathbb{Z} &= \{0+5\mathbb{Z}, 1+5\mathbb{Z}, 2+5\mathbb{Z}, 3+5\mathbb{Z}, 4+5\mathbb{Z}\} \\ &= \left\{ \{\ldots,-5,0,5,\ldots\}, \{\ldots,-4,1,6,\ldots\}, \{\ldots,-3,2,7,\ldots\}, \{\ldots,-2,3,8,\ldots\}, \{\ldots,-1,4,9,\ldots\} \right\}.\end{aligned}$$

def:ideal

Definition 2.3.19: Ideals

For cring R and subset $I \subset R$, we say I is an ideal when it has the special properties of $7\mathbb{Z}$ noted in Comp. 2.3.16, Comp. 2.3.17, namely:

- $0 \in I$,
- closed under negation $-I = I$,
- closed under addition $I + I \subseteq I$,
- closed under (R-)scalar multiplication $R \cdot I \subseteq I$.

— very vector-space like... in fact exactly same axioms of R-submodule $\subseteq R$)

def:cosets

Definition 2.3.20: Cosets

For cring R and a set $S \subseteq R$, the S-cosets are precisely the subsets of R of the form $r_0 + S$ for $r_0 \in R$. I.e. images of the set S under the bijective map $[r \mapsto r + r_0] : R \to R$. I.e. all cosets of S are just shifts of each other (or of S) by "translations" $r_0{+} : R \to R$ (bijection because inverse $-r_0{+} : R \to R$ exists).

Another common notation, when S is clear from context (to emphasize the idea that they are "one-thing" instead of a set) is $\overline{r_0} := r_0 + S$.

exc:cosetFacts

Exercise 2.3.21: Check cosets of ideal behave like 7ℤ: disjointness, arithmetic

Check that for cring R and ideal I, the manipulations from Comp. 2.3.17 still hold (I call this the coset equality lemma:

$$a \in b+I \iff a-b \in I \iff b-a \in I \iff b \in a+I \iff [a+I=b+I].$$

In particular, conclude that distinct I-cosets are disjoint.

Check that the manipulations from Comp. 2.3.16 still hold, i.e. set addition and set multiplication of two I-cosets lands inside (exactly) one I-coset.

Corollary of coset equality lemma in previous Exc:

llCosetEquality

Corollary 2.3.22: Null coset equality lemma

For cring R and ideal I, $[x+I=I \iff x \in I]$.

2.3.3.2 Definition of quotient ring (developing language)

I did prepare another example of quotient in §2.3.3.2.1 that I thought about doing before the general definitions, but honestly I think the $\mathbb{Z}/7\mathbb{Z}$ example above is complicated enough that it alone already justifies/motivates the search for a more "birds-eye view"/"higher(=more abstract) vantage point", to cut away the fluff and reveal the more streamlined core (like how the EuR definition captured exactly what was needed for the Bezout proof, leading to the cleanest possible arguments). It is up to the teacher to decide the order.

We have finally arrived at the modern understanding of modular arithmetic (or quotients in general), in which "elements" are entire equivalence classes/cosets.

Can now define quotient rings, ideals. Canonical "projection" aumu-homomorphism (official name is: ring homomorphism) to quotient, e.g. $\pi : \mathbb{Z} \to \mathbb{Z}/7\mathbb{Z}$.

def:quotientRing

Definition 2.3.23: Quotient ring R/I "R mod I" construction

For a cring R and ideal I, the set of "cosets" $R/I := \{r+I : r \in R\}$, equipped with set addition and set multiplication [w.r.t. the addition and multiplication on R] (or more precisely, the *unique [Exc. 2.3.21]* coset *containing* the set sum or set product), forms a ring (Exc!).
Note that in terms of elements, set multiplication for instance $(a+I)\cdot(b+I) \subseteq (c+I)$ means $\forall i_a, i_b \in I$, $\exists i_c \in I$ s.t. $(a+i_a)(b+i_b) = c+i_c$. Beware the "$+I$" is instantiated as $\forall$ on LHS and $\exists$ on RHS!

Exc: show that above definition of $+, \cdot$ in R/I (using set addition and set multiplication, $+$ a bit extra *subtlety*) actually can be written much more simply as: $(r_1+I)+(r_2+I) := (r_1+r_2)+I$ and $(r_1+I)(r_2+I) := (r_1r_2)+I$, or using the overline/bar notation for cosets, $\overline{r_1}+\overline{r_2} = \overline{r_1+r_2}$, $\overline{r_1}\,\overline{r_2} = \overline{r_1r_2}$. This is the standard definition; the benefit of my definition is that it is automatically well-defined (well after we prove the uniqueness of cosets), unlike this definition where we have to worry about choice of representative.
I.e. on a syntactic level, for anything we do in R (and with operations $R \times R \to R$), we can just apply bars on top of all the elements, and get the same things happening in R/I (and with

operations $R/I \times R/I \to R/I$). This is essentially the proof that R/I remains a cring.

I'm doing the Exercise here: using set plus, $(r_1+I)+(r_2+I) = (r_1+r_2)+I+I = (r_1+r_2)+I$ and $(r_1+I)\cdot(r_2+I) = r_1r_2 + Ir_2 + r_1I + II \subseteq (r_1 \cdot r_2) + I$.

:cosetsAsFibers

Proposition 2.3.24: Cosets are fibers

The natural projection $\pi : R \twoheadrightarrow R/I$ given by $r \mapsto \overline{r} := r + I \in R/I$ has fibers $r + I$ (almost sounds tautological, but think through it carefully). I.e. we WTS $\pi^{-1}(\overline{r}) = r + I$.

Proof: this really should be an Exercise, but I already typed up the following:

- To show $r + I \subseteq \pi^{-1}(\overline{r})$, consider arb. $r + i_0 \in$ LHS; then $\pi(r + i) := r + i_0 + I = r + I =: \overline{r} \in R/I$. (The last equality follows from null coset equality lemma Cor. cor:nullCosetEquality 2.3.22.)
- To show $\pi^{-1}(\overline{r}) \subseteq r + I$, suppose x in LHS, i.e. $x + I =: \pi(x) = \overline{r} \in R/I$; this means $x + I = r + I \implies x + 0 \in r + I$. ■ Alternative proofs below.

Fibers of any function f are always disjoint; can also think of fibers as the equivalence classes of pulling back literal equality on the codomain through the function f. Thus we have an alternative proof of coset disjointness (which we already saw from coset equality lemma Exc. exc:cosetFacts 2.3.21):

setDisjointness

Corollary 2.3.25: Coset disjointness, again (cf. Exc. exc:cosetFacts 2.3.21)

For cring R and ideal I, distinct cosets are disjoint.

Proof: distinct cosets $r_1 + I \neq r_2 + I$ are preimages of disjoint subsets $\{r_1 + I\}, \{r_2 + I\} \subseteq R/I$. I.e. we're using that $\{r_1 + I\}$ and $\{r_2 + I\}$ are disjoint subsets of R/I to show that $r_1 + I, r_2 + I$ are disjoint subsets of R. Troll! ■

Other ways of writing the proof (good to develop your own taste in proofs):

Alternative proof of disjoint coset lemma using coset equality lemma: recall the above characterization of coset equality $a = b \mod I :\iff a + I = b + I \iff a - b \in I \iff b - a \in I$. Thus $a + I \neq b + I \Leftrightarrow a - b \notin I \Leftrightarrow a \notin b + I \Leftrightarrow \forall i \in I, a + i \notin b + I + i = b + I \Leftrightarrow a + I \not\cap b + I$. If that was nonsense to you, read the more "algebra bash" proof below.

"Algebra bash" proof of disjoint coset lemma: contrapositive is $r_1 + I, r_2 + I$ overlap (at say point x), WTS they're the same, i.e. $r_1 + I = r_2 + I$. Let $y \in r_1 + I$, WTS $y \in r_2 + I$. Well, $x \in r_j + I$, the elements of set $r_j + I$ have guaranteed form, means we can write out explicit representation $x = r_j + i_j$. Well then $r_1 + i_1 = r_2 + i_2 \iff r_1 - r_2 = i_2 - i_1 \in I$. Again instantiating the form $y \in r_1 + I \implies y = r_1 + i$, get $y = r_2 + (r_1 - r_2) + i = r_2 + (i_2 - i_1) + i \in r_2 \in I$. ■

2.3.3.2.1 E.g.: indistinguishability class to identify formal syntax $\imath^2$ and -1

omplexNumbers

We saw above, we can do identify $12, 5$ ("blind to multiples of 7") and still do addition and multiplication ("modular arithmetic mod 7") by doing set addition and set multiplication with the cosets $0 + 7\mathbb{Z}, 1 + 7\mathbb{Z}, \ldots, 6 + 7\mathbb{Z}$.

This ability to identify certain classes of symbols/syntax and still do arithmetic is very useful; in fact one can view the syntax of complex numbers in this way: we have a formal symbol $\imath$, and we can

form formal polynomial expressions $\mathbb{R}[\imath] := \{a_n\imath^n + \ldots + a_1\imath + a_0 : n \in \mathbb{N}; a_n, \ldots, a_1, a_0 \in \mathbb{R}\}$ where we identify $\imath^2$ and -1.

Recall that in §2.1, we developed a "semantic model" of complex numbers (as actions=linear-transformations on $\mathbb{R}^2$), and then defined the syntax of complex numbers based off of that. It is of course possible to define the complex numbers purely syntactically, but the cost is then we have to verify all the field axioms, many of which (e.g. associativity) come for free from the semantics.

Our "cosets" construction is another way of defining the complex numbers (which inherits many properties from the formal polynomial ring $\mathbb{R}[x]$), and it has the advantage of being very general (so we can do the busywork once, and then get to use the machinery for free later).

Ok, enough philosophy. How do we actually identify $\imath^2$ and -1 in the formal polynomial ring $\mathbb{R}[\imath]$, and still have well-defined addition and multiplication? Well recall, to identify 5 and 12, we considered cosets $5 + 7\mathbb{Z} = 12 + 7\mathbb{Z}$ (where $7\mathbb{Z} = \langle 7\rangle_{\mathbb{Z}}$ came from $12 - 5 = 7$). I did enough work above laying out the general principles that we can do exactly the same thing:

Exercise 2.3.26: Quotient construction of $\mathbb{C}$

Figure out what ideal I we should quotient the formal polynomial ring $\mathbb{R}[\imath]$ by to get $\mathbb{C}$.

2.3.3.3 Quotient field iff maximal ideal

Massage the concepts until realize that R/I is a field $\iff$ I is a maximal ideal.

Example massage: every $r \in R$ not already in I has some x s.t. $rx = 1 \mod I$ so $rx + i = 1$ for some $i \in I \iff 1 \in \langle r, I\rangle := \{\text{all finite } R\text{-linear combinations of elements inside bracket}\} = \text{span}_R\{r, I\}$, i.e. I is a maximal ideal.

Try writing out your own proof before reading the polished one below!

Proposition 2.3.27: Quotient field iff maximal ideal

R/I is a field $\iff$ I is a maximal ideal.

Proof:

$$\begin{aligned} R/I \text{ field} &\iff \forall \overline{a} \neq 0, \exists \overline{b} : \overline{a}\overline{b} = \overline{1} \\ &\iff \forall a \notin I, \exists b : (a+I)(b+I) := \underbrace{ab + I = 1 + I}_{\substack{\iff ab+I \ni 1 \\ \iff \exists i \in I : ab + i = 1}} \\ &\iff \forall a \notin I, \quad \underbrace{\langle a, I\rangle \ni 1}_{\iff \langle a,I\rangle = \langle 1\rangle = R} \\ &\iff I \text{ maximal.} \blacksquare \end{aligned}$$

So arithmetic fact about quotient R/I equivalent to ideal theoretic fact in R!

2.3.3.4 Classify ideals of $\mathbb{Z}$ (abstract Bezout!)

Above saw that $n\mathbb{Z}$ is ideal (draw as evenly spaced grid in $\mathbb{Z}$), and understood quotients $\mathbb{Z}/n\mathbb{Z}$. Are there other ideals of $\mathbb{Z}$? Try to draw it out. Mysterious $I \subseteq \mathbb{Z}$ that I know is an idea; if $a \in I$, I get evenly spaced grid $a\mathbb{Z} \subseteq I$. If I also contains b not in the previous grid $a\mathbb{Z}$, then I still get evenly spaced grid $\gcd(a, b)\mathbb{Z} \subseteq I$. It seems like perhaps the only ideals are the evenly spaced grids:

prop:ZisPIR

Proposition 2.3.28: Generalized abstract Bezout lemma in $\mathbb{Z}$ (i.e. $\mathbb{Z}$ is PIR)

Let I be ideal in $\mathbb{Z}$. Then $\exists m \in \mathbb{Z}$ s.t. $I = m\mathbb{Z}$.

Proof: intuition: if I give you an evenly spaced grid, then how do you find m? It's the smallest non-zero element of the grid.

Case 1: suppose I contains non-zero element. Then let m be the minimal non-zero element of I (in this setting, we could write $m := \min(I \cap \mathbb{N}^+)$ but this doesn't generalize so I'm going to only keep this here as a gray remark).

def:PIR

Definition 2.3.29: PIR, principal ideals

A principal ideal I of a cring R just means an ideal of the form $I = rR$ for some $r \in R$.
A cring R is a PIR if every ideal I in R is principal.

exc:polysPIR

Exercise 2.3.30: $F[x]$ is PIR

or more generally,

c:EuRimpliesPIR

Exercise 2.3.31: EuR $\implies$ PIR

2.3.3.5 $\mathbb{Z}/p\mathbb{Z}$ is a field

WTS $p\mathbb{Z}$ is a maximal ideal.

The definition of prime people have in their head is actually "irreducible", which we can quickly massage into $\iff$ "maximal amongst principal ideals".

xPrincipalIdeal

Lemma 2.3.32: Irreducible element iff maximal amongst principal ideals

Let R be a cring. Then [$r \in R$ irreducible $\iff \langle r \rangle$ is maximal amongst the principal ideals of R (i.e. $\langle z \rangle \supsetneq \langle r \rangle \implies \langle z \rangle = R$)].

But we showed above $\mathbb{Z}$ is PIR (§2.3.2.1 bezout), so "maximal amongst principal ideals" is the same as "maximal".

Summary of what we've shown:

xPrincipalIdeal

Proposition 2.3.33: In PIR, irreducible element implies maximal amongst principal ideals

Let R be a cring that is PIR. Then $r \in R$ irreducible $\implies \langle r \rangle$ is a maximal ideal. (In particular, $R/\langle r \rangle$ is a field.)

2.3.4 Factorization in $\mathbb{Z}, F[x]$, and $\mathbb{Z}[i]$ ("EuD $\implies$ UFD")

Prologue circle lattice point counting question already convinced us that unique factorization is actually useful. Now we prove it, by thinking $\mathbb{Z}$ in our head but writing out something that works for $\mathbb{Z}, \mathbb{Z}[i], F[x]$.

2.3.4.1 Existence (irreducibility reigns supreme)

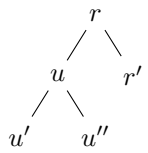
Key idea: irreducibility. Given an element r in $\mathbb{Z}$ (or you can imagine $F[x]$ or $\mathbb{Z}[i]$), it's either irreducible (which recall is defined up to units) or not. If not, by definition $r = ab$ for a, b both non-units.

$[a \text{ unit} \iff N(a) = 1]$

$[ab = r \implies N(a), N(b) \leqslant N(r), \text{ with strict inequality if } a, b \text{ are non-units}]$

In this case where we have Euclidean descent function, notion of "irreducible" is fairly natural (can't split into 2 strictly smaller pieces).

We can also give equivalent definition in the "language" of the binary tree and its only operation "$|$" ("divides"): from r we can grow a factorization tree, so all children divide/are factors of the parents; we declare a node "infertile" if its parent divides it; then an element is irreducible if every possible factorization tree grown out of it looks like (for units u, u', u''):



where the descent property $[ab = n \implies N(a), N(b) \leqslant N(n), \text{ with strict inequality if } a, b \text{ are non-units}]$ forces the branches of the binary tree to terminate, where an element is "terminal" if there are no smaller (w.r.t. N) a, b dividing n, in which case by contrapositive of strict part of descent property written above means at least one of a, b must be unit.

Stepping back a bit, what we really need in this Existence step is that no infinite chains $\ldots a'' \mid a' \mid a \mid r$, i.e. "minimal element" of chain $a^\star$ is s.t. $c \mid a^\star$ implies c equals 1 or $a^\star$ up to multiplication by units; so just a more general "no infinite descent" principle in place of the standard "no infinite descent" principle for $\mathbb{N}$.

Perhaps a helpful visualization: given $r = ab$ (root of binary tree) we can split into a, b (two branches: left and right), and repeat to get binary tree encoding factorization, where element-wise Noetherianity implies following each branch of the tree eventually terminates in a leaf node, i.e. some sort of irreducible element.

2.3.4.2 Uniqueness (primality reigns supreme)

So we saw above: starting at "root" n, if it is irreducible, then stop; else it is reducible, in which case we can split it into 2 children (that multiply to n). Iterating this process on each children, we build a tree (guaranteed finite because children decrease in Norm, and positive numbers can only decrease finitely many times and remain > 0) whose leaves are all irreducible.

Clearly, starting at n, we can build many trees. It seems almost unbelievable from this perspective that all the trees we build are guaranteed to have the same irreducible numbers in their leaves. We now show uniqueness of factorization by starting with

$$p_1 \cdots p_r = n = q_1 \cdots q_s$$

and showing that we can pair up the p_i and q_j together.

The rough plan is: start with p_1; want to show it shows up on RHS (that's **Step 1**). Then, we want to "cancel p_1 from both sides" (**Step 1**). This makes both sides smaller, so we can just say "induction",

or "repeat the argument" until both sides are empty.

For **Step 2**, let me distill out a definition: we want cancellation in $\mathbb{Z}$, i.e. we want $\mathbb{Z}$ to satisfy

Definition 2.3.34: Integral domain definitions (cancellation, zero divisors)

We say a (c)ring R is an integral domain exactly when: $\forall a \neq 0, b, c \in R$, $ab = ac \implies b = c$.
Equivalently (Exc!), exactly when: $xy = 0 \implies x = 0$ or $y = 0$.

Let me just assert $\mathbb{Z}$ is an integral domain; I don't want to get down in the Peano weeds.

Now, to **Step 1**:
Key idea: primality. We have p_1 on LHS, want p_1 to show up on RHS. In other words, looking mod p_1 (i.e. apply projection $R := \mathbb{Z} \twoheadrightarrow \mathbb{Z}/p_1\mathbb{Z}$), we have $0 = \overline{q_1} \cdots \overline{q_s}$, and want one of the $\overline{q_j}$ to be 0 (so that $p_1 \mid q_j$, forcing $p_1 = q_j$ up to ± 1 since by irreducibility, only divisors of q_j are 1 and itself — and $p_1 \neq \pm 1$).
In other words, we've massaged the problem to WTS the quotient ring $R/\langle p_1 \rangle$ is an integral domain! Good thing I just introduced that definition... hahaha
But we already showed $R/\langle p_1 \rangle$ is a field, so it's an integral domain! QED.

2.3.4.2.1 Quotient integral domain iff prime ideal

Recall we characterized quotient ring being field by a property of the ideal. Let us try to find such a characterization for quotient ring being an integral domain.

prop:primeDomain

Proposition 2.3.35: Integral domain quotient iff prime ideal

For cring R and ideal I,
$[R/I$ is integral domain $\iff$ I satisfies the property: $\forall x, y \in I$, $xy \in I \implies x \in I$ or $y \in I]$.
This property of ideals, is called "primality".

Proof:

$$\begin{aligned} R/I \text{ integral domain} &\iff \forall \overline{a}, \overline{b} \in R/I, [\overline{a}\overline{b} = \overline{0} \implies \overline{a} = \overline{0} \text{ or } \overline{b} = \overline{0}] \\ &\iff \forall a, b \in R, [\underbrace{\overline{a}\overline{b} = \overline{0}}_{\substack{\iff ab+I=I \\ \iff ab \in I}} \implies \underbrace{\overline{a} = \overline{0}}_{\substack{\iff a+I=I \\ \iff a \in I}} \text{ or } \underbrace{\overline{b} = \overline{0}}_{\substack{\iff b+I=I \\ \iff b \in I}}] \\ &\iff: I \text{ prime.} \end{aligned}$$

(The last " $\iff$ " in each underbrace is from Cor. cor:nullCosetEquality 2.3.22 $x + I = I \iff x \in I$ recall proof: ($\implies$) $x + 0 \in x + I = I$; ($\impliedby$) $x = i_0 \implies x + I = i_0 + I = I$.) ■

2.3.4.3 Irreducibility vs. primality

irredVsPrime

Let me distill out a proposition from what we discussed above:

Proposition 2.3.36: In PIR, irreducible element $\implies$ prime element

Best modern proof: irreducible $\iff$ [generated ideal is a maximal element amongst principal ideals, so PID $\rightsquigarrow$ maximal ideal] $\implies$ prime ideal.

bezout
§2.3.2.1 shows PID, so just need maximal ideal $\implies$ prime ideal.
quotFieldMaxIdeal
§2.3.3.3 already gave field characterization, and field $\implies$ integral domain.
Computational proof (maybe I'll even assign as homework for comparison sake): see MSE

2.3.4.4 Redoing in $\mathbb{Z}[i]$ (or $F[x]$), or general PIR

Let us think through above proof of existence and uniqueness of factorization in $\mathbb{Z}[i]$ (or $F[x]$; or I can bring that up later).

(a) Need a Norm that satisfies some properties (to show that the factor trees are always finite).

(b) Need integral domain (which recall we used to cancel p_1 from both sides).

(c) Need PIR (to apply Prop. prop:irredMaxPrincipalIdeal 2.3.33 to guarantee principal ideal of irreducible element is a maximal ideal).
- How did we show $\mathbb{Z}$ is PIR? For ideal I, we took m minimal non-zero element of I. Then how did we show $I = \langle m \rangle$? We used division algorithm!
- So need division algorithm in $\mathbb{Z}[i]$ (and $F[x]$).

th:divAlgoZi

Theorem 2.3.37: Division algorithm in $\mathbb{Z}[i]$

For $\beta, \alpha \in \mathbb{Z}[i]$, there is $q, r \in \mathbb{Z}[i]$ s.t. $\beta = \alpha q + r$ AND $0 \leqslant |r| < |a|$, where we take $|\bullet|$ to be ??? (later revealed to be magnitude (i.e. distance from 0) in $\mathbb{Z}[i] \subseteq \mathbb{C}$.)

What notion of length do we take? Taking magnitude, recall complexNumbers §2.1 that norm is multiplicative map $\mathbb{Z}[i] \to \mathbb{Z}$, so taking Norm to be magnitude, get item (a) above properties, so factor tree finite!

What about division algorithm? Can picture as follows: start at $\beta \in \mathbb{Z}[i] \subseteq \mathbb{C}$, can walk along $+\alpha, -\alpha, +i\alpha, -i\alpha$, and so forth. I.e. form a larger square grid $\alpha\mathbb{Z}[i]$, square sidelengths $|\alpha|$, and our goal is to find a point on this fatter lattice $\alpha\mathbb{Z}[i]$ close to the origin (within $< |\alpha|$ of the origin).

But we can! drawing picture, the farthest any point of $\mathbb{C}$ can be from a lattice point of the fatter grid $\alpha\mathbb{Z}[i]$ is $\leqslant \frac{\sqrt{2}}{2}|\alpha| < |\alpha|$! AMAZING! Somehow, an arithmetic fact comes from geometry (of lattices)!

2.3.5 The Count

Review/go over thoroughly 3b1b video "π hiding in prime regularities" to answer the question of $\#(\mathtt{C}_a)$. I've decided going all the way to χ_4 and Leibniz series is too far for now. I've pushed that to a later section, in which there's a section I can discuss more of the analytic number theory background, like divisor sums, Möbius inversion, etc.

First study $\#(\mathtt{C}_p)$ for p prime.

Fermat Christmas theorem (Mathologer windmill footprint 2-involutions proof), even though that proof is not too related to abstract algebra.

The hardest conceptual part is writing the pattern we find in terms of divisor sums of χ. 3b1b's example is $N = 2^2 \cdot 3^4 \cdot 5^3$. The idea is common theme, understand **flow/evolution**, as factors of N change one by one. Distinct prime factors are treated separately (multiply together at the end, i.e. $\#(N) = \#(2^2) \cdot \#(3^4) \cdot \#(5^3)$), so the only flow/evolution we need to understand is incrementing the power of a prime. I've decided to postpone this to a later section. Indeed, it is still quite an ingenious leap, as I've criticized before on Reddit, and MSE.

sec-2-4

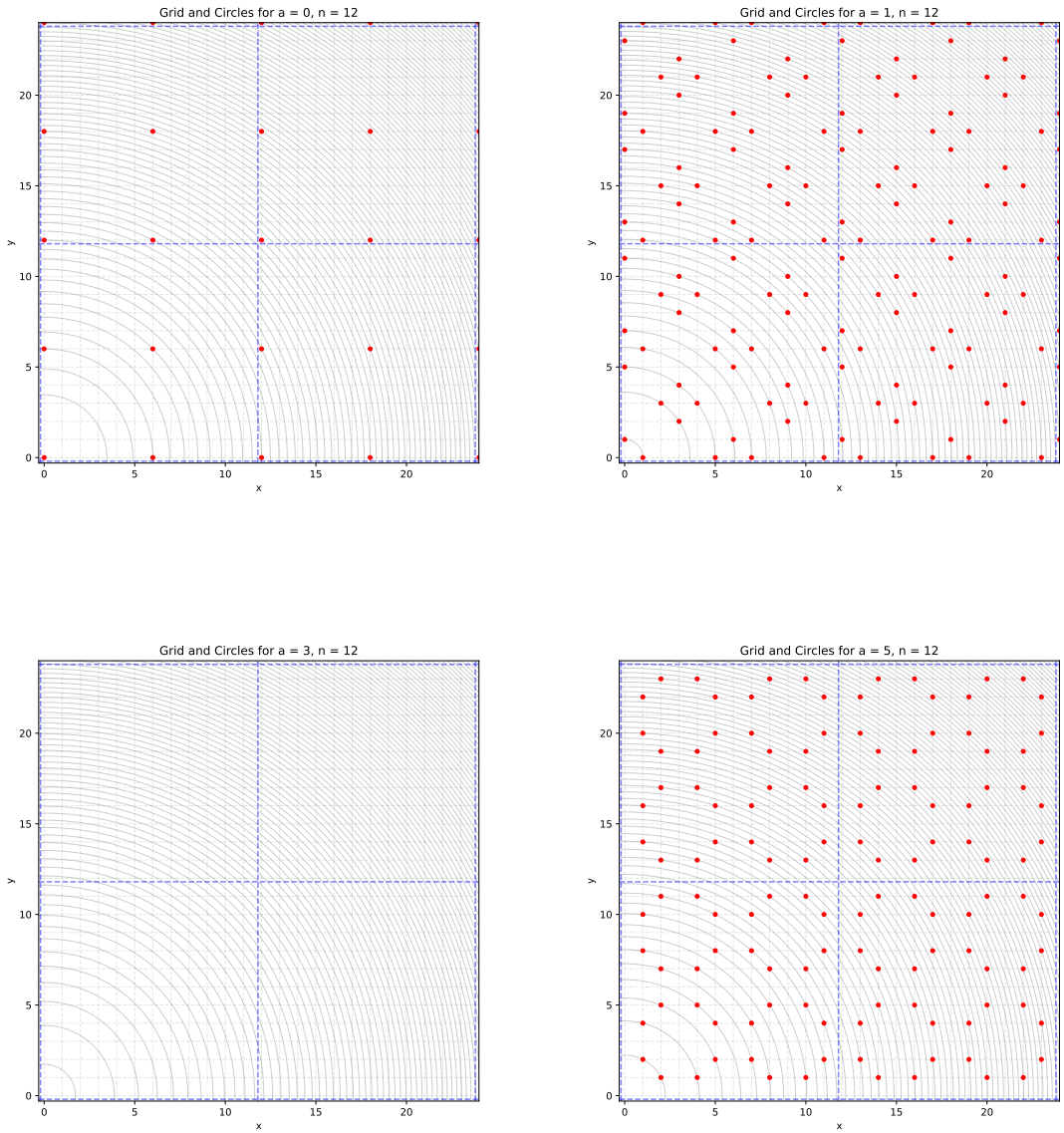
sec-2-4

2.4 Counting Lattice Points on Circles in Modulo-land

2.4.1 Lattice points on circles mod n: geometry vs. algebra

It is now natural to define $\#_n(\mathtt{C}_a) \coloneqq \#\{(x,y) \in (\mathbb{Z}/n\mathbb{Z})^2 : x^2+y^2 = a \pmod n\}$. Related to points in $\mathbb{Z}^2$ on circles $x^2+y^2 = a, a+n, a+2n, \ldots$.

Miraculous "conspiracy": **geometrically** (see concentric circles around 0), do not expect translation symmetry (radial symmetry, yes, but not translation). But that is trivial **arithmetically** (algebraically)!



This I think serves as a very good preliminary example for how arithmetic/algebra can tell us something very unexpected about geometry.

2.4.2 Factorization and Norm

The next instance of this geometry-algebra connection is the fact that we can rephrase lattice points on circles $\mathtt{C}_a$ as factorizations of a in $\mathbb{Z}[i] \subseteq \mathbb{C}$:

$$x^2 + y^2 = a \iff (x+iy)(x-iy) = a \iff Nm(x+iy) = a^2;$$

where we define $Nm(x+iy) = |x+iy|^2 = x^2+y^2$ (square of complex magnitude) — we choose to have an "extra" square because then $Nm : \mathbb{Z}[i] \to \mathbb{Z}$.

Recall (from my laborious discussion in §2.1) that the complex magnitude (length = "geometry") has a very nice algebraic/"homomorphism" property: it is multiplicative; and so the norm Nm is too. OR, $Nm(a+bi)$ is very tied to multiplication anyways because it's the same as "multiply by conjugate" operation; this interpretation makes multiplicativity trivial.

Recall our "fibers bijective" result for rings (see Prop. 2.3.24, Def. 2.3.20). Thinking about it again, all we really used was that $\pi : R \to S \coloneqq R/I$ preserved addition and so shifting a fiber F (say $\pi^{-1}(s)$) by $r \in R$ lands inside another fiber F': $r + \pi^{-1}(s) \subseteq \pi^{-1}(\pi(r)+s) =: F'$ (yes, the notation is disgusting but the idea is intuitive written in words); and shifting fiber F' by $-r$ lands back in fiber F; and $r+ : R \to R$ and $(-r)+ : R \to R$ are inverse bijections forcing F to be bijective to F'.

Here, we have Nm preserving multiplicative, but precisely the same idea applies. If we focus on multiplicatively "shifting" by elements which have (multiplicative) inverses, then we get fibers in bijection.

Let us focus on such invertible multiplicative shifts of the multiplicative identity: we get precisely the croup of units of $\mathbb{Z}[i]$.

def:croup

Definition 2.4.1: (Multiplicative) croup

One column (right for "multiplicative", left for "additive") of 2-column, 5-row table for definition of field.

def:unitCroup

Definition 2.4.2: Unit croup of cring

For R a cring, $R^\times \coloneqq \{r \in R : \exists r^{-1} \in R\}$.

Not many elements with multiplicative inverse in $\mathbb{Z}[i]$ or $\mathbb{Z}$, but remember that $\mathbb{Z}/n\mathbb{Z}$ (and espeically $\mathbb{Z}/p\mathbb{Z}$) has tons!

Note Nm factors through to multiplicative map

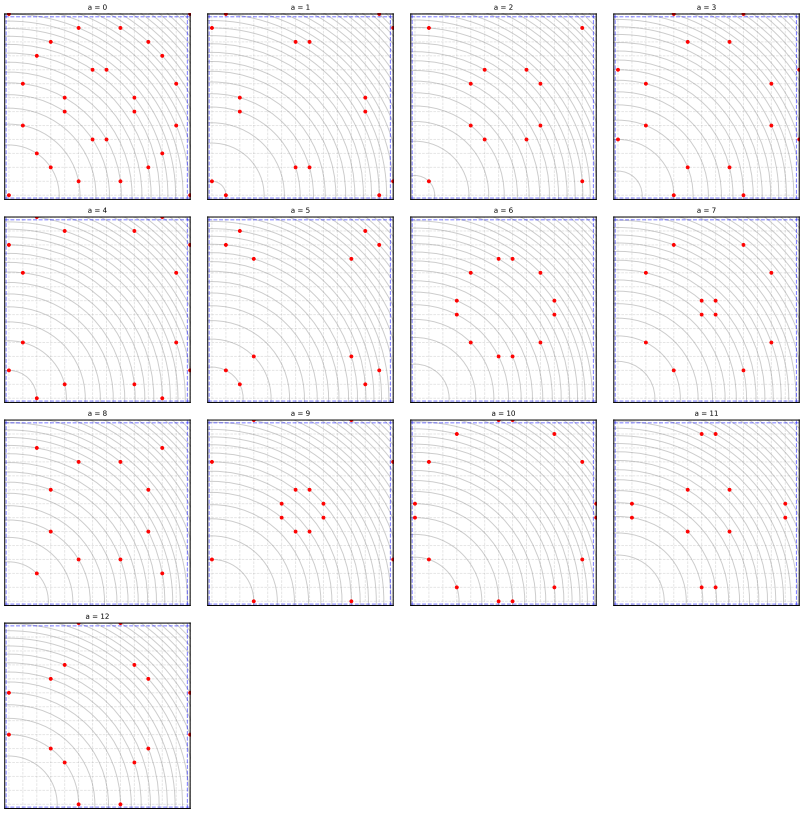
$$Nm : \mathbb{Z}/n\mathbb{Z}[i] \to \mathbb{Z}/n\mathbb{Z}. \qquad \text{eqn:NmZmodn}(2.1)$$

(Define operations on $\mathbb{Z}/n\mathbb{Z}[i]$ by copying definitions for $\mathbb{Z}[i]$ and putting bars over everything. Same with norm; proof of multiplicativity goes through just the same, but again with bars over every letter. The power of syntax!!!)

And the point counts $\#_n(\mathtt{C}_a)$ are precisely the size of the fibers of this multiplicative map!

Indeed one reason why the modular arithmetic version of the point-counting problem is easier, is because $Nm : (\mathbb{Z}/n\mathbb{Z}[i])^\times \to (\mathbb{Z}/n\mathbb{Z})^\times$ is non-trivial while $\mathbb{Z}[i]^\times \to \mathbb{Z}^\times$ is trivial!

Picture showing uniform size fibers:



oFibersBijective

Proposition 2.4.3: Fibers of group homomorphism are bijective

$\phi : G \to H$ (say surjective; otherwise non-empty fibers in bijection, and the rest are empty fibers) preserving multiplication (i.e. a group homomorphism) has fibers in bijection (via bijections $g* : G \to G$, $g^{-1}* : G \to G$).

The analogy with what we did with rings is very very strong; here is another repeat theorem (which we use later):

lagrangeTheorem

Theorem 2.4.4: Lagrange's theorem

For G a group, H subgroup, distinct cosets in $G/H := \{gH : g \in G\}$ partition G, and are the same size ($= |H|$) — in fact are in bijection to each other via left mutliplication bijections $g* : G \xrightarrow{\cong} G$. In particular, $|G| = |H| \cdot |G/H|$ so $|G|$ is divisible by $|H|$.
(Very similar to Prop. prop:groupFibersBijective 2.4.3, which (as mentioned in a previous paragraph) is in turn very similar to previous Prop. prop:cosetsAsFibers 2.3.24, Def. 2.3.20)

Proof:

exc:ZpZunitGroup

Exercise 2.4.5: Croup of units of $(\mathbb{Z}/p\mathbb{Z})[i]$

Check $\left((\mathbb{Z}/p\mathbb{Z})[i]\right)^\times = (\mathbb{Z}/p\mathbb{Z})[i] \setminus \{\overline{0} + \overline{0}i\}$. Hint: conjugate.

2.4.3 The Count (in dimension 2)

Following my GSO talk (cf. YouTube https://youtu.be/IXBp2UHbky0?si=VM0sBrrthBXU266-).

Once we have fibers of $Nm : (\mathbb{Z}/p\mathbb{Z}[i])^\times \to (\mathbb{Z}/p\mathbb{Z})^\times$ same size, arrive at

$$p^2 = Z + (p-1)A \implies Z \equiv A \mod p.$$

umRootsLeqDegree

Corollary 2.4.6: Number of roots bounded by degree for polys over field

For F a field, $p \in F[x]$ of degree d has $\leqslant d$ roots in F.

is a corollary of:

h:factorTheorem

Theorem 2.4.7: Factor theorem in $F[x]$

For F a field, $p \in F[x]$ has $\alpha \in F$ a root $\iff$ $(x-\alpha)$ is a factor of p, i.e. $p = (x-\alpha)q + 0$.

Proof: division algorithm! Exc! ■

Limits A, Z to 3 possibilities: $(p,p), (p-1, 2p-1), (p+1, 1)$.

2.4.3.1 Special (zero) fiber large $\iff$ $\sqrt{-1} \in \mathbb{Z}/p\mathbb{Z}$!

2.4.3.2 $p = 4k+1$ characterization ($\Longleftarrow$): order, Lagrange's theorem

We realize that 3 mod 4, 1 mod 4 behavior related to finding elements of certain order in $(\mathbb{Z}/p\mathbb{Z})^\times$. Investigate elements of multiplicative order d, relatedly solutions to $x^d = 1$ in $\mathbb{Z}/p\mathbb{Z}$. simple MSE proof (roots of unity, Euler totient function)

def:orderElement

Definition 2.4.8: Order of element

For G a (multiplicative) group, $x \in G$, the order of x is the smallest positive number d s.t. $x^d = \underbrace{x \cdot \ldots \cdot x}_{d \text{ times}} = 1$.

Usually (in examples) we calculate it by just calculating $x^1, x^2, \ldots$ and record the exponent that first leads to 1.

In fact, I claim that $\mathrm{ord}_G(x) = |\langle x \rangle|$, where $\langle x \rangle := \{x^0, x^1, x^2, \ldots\}$ — i.e. I'm claiming that $1 \leqslant m < n \leqslant d \implies x^m \neq x^n$; this is because $x^m = x^n \iff 1 = x^{n-m}$, but minimality of d forces $n - m \geqslant d$, ※.

Can also think about it like this: every x gives rise to a group homomorphism $\phi_x : \mathbb{Z} \to G$ mapping $n \mapsto x^n$. Then $\ker(\phi_x) \subseteq \mathbb{Z}$ is an ideal, so because $\mathbb{Z}$ is PIR (Prop. 2.3.28), $\ker(\phi_x) = \langle d \rangle_{\mathbb{Z}}$. 1st isomorphism theorem says $\mathbb{Z}/d\mathbb{Z} \cong \mathrm{im}(\phi_x) = \langle x \rangle \subseteq G$, so indeed $\langle x \rangle$ has d distinct elements.

prop:ZisPIR

More generally, we define:

subgroupBracket

Definition 2.4.9: "Subgroup generated by" brackets

For G a group, $S \subseteq G$, define $\langle S \rangle := \{\text{all finite products of elements in } S\}$.

Check (Exc!) this is actually a subgroup!

videsCardinality

Corollary 2.4.10: Order divides cardinality

For G a group, $x \in G$, I claim $\mathrm{ord}_G(x)$ divides $|G|$. Exc! (corollary of Lagrange's theorem Thm. th:lagrangeTheorem 2.4.4)

ex:Z13Zexample1

Example 2.4.11: Example/Exc.: $G := (\mathbb{Z}/13\mathbb{Z})^\times$

Let us go through elements of $G := (\mathbb{Z}/13\mathbb{Z})^\times$, calculate their orders, keeping in mind the question of order 4 elements.

1 has order 1. What about 2? We calculate the powers of 2 as in $\langle 2 \rangle$, and we observe that we actually generate all 12 elements of $(\mathbb{Z}/13\mathbb{Z})^\times$.
It is nice to visualize this by drawing $2^1, 2^2, \ldots$ on a circle (coincidentally, exactly a clock: 2^1 at hour 1, etc.)
Now that we have every element of $(\mathbb{Z}/13\mathbb{Z})^\times$ written as 2^e for some exponent e, it is easy to calculate orders. (I.e. we've found a group isomorphism $\exp_2 := [e \mapsto 2^e] : (\mathbb{Z}/12\mathbb{Z}, +) \to (\mathbb{Z}/13\mathbb{Z})^\times$; the order of 2^e is the smallest d s.t. $(2^e)^d = 1 \iff ed \in 12\mathbb{Z}$.)
It is also nice to visualize $\langle x \rangle$ for all elements x inside the "clock" — the form nice polygons/stars!

Exc.: illustrate Cor. cor:orderDividesCardinality 2.4.10/Lagrange's theorem by taking $H := \langle 5 \rangle$, and writing out G/H (a (finite) set whose elements are (finite) sets!).

What about for $p = 19$ instead of $p = 13$?

2.4.3.3 $p = 4k + 1$ characterization ($\Longrightarrow$): cyclicity of unit groups

For $U := (\mathbb{Z}/p\mathbb{Z})^\times$, define $O_d := \{x \in G : \mathrm{ord}(x) = d\} \subseteq \{x \in U : x^d = 1\} =: U_d \subseteq U$. Note that degree bound for number of roots (Cor. cor:numRootsLeqDegree 2.4.6) applies here, telling us $|U_d| \leqslant d$.

ex:Z13Zexample2

Example 2.4.12: Example: $U := (\mathbb{Z}/13\mathbb{Z})^\times$

Continuing Ex. ex:Z13Zexample1 2.4.11.
For $U^{p=13} := (\mathbb{Z}/13\mathbb{Z})^\times$, write out O_d, U_d for all $d|12$, and then read how I do it below.

Thinking again on what we said about group imor. $\exp_2 := [e \mapsto 2^e] : (\mathbb{Z}/12\mathbb{Z}, +) \to (\mathbb{Z}/13\mathbb{Z})^\times$ and "order of 2^e is the smallest d s.t. $(2^e)^d = 1 \iff ed \in 12\mathbb{Z}$", we see that

$$O_{12} \subseteq (\mathbb{Z}/13\mathbb{Z})^\times \overset{\exp_2}{\longleftrightarrow} \text{numbers in } (12] \text{ relatively prime to } 12, \text{ i.e. } (\mathbb{Z}/12\mathbb{Z})^\times.$$

So, we now understand all order 12 elements ("generating elements") of $U^{p=13}$.
Can easily find an order d element for any $d \mid n := |U^{p=13}| = 12$, namely $2^{n/d}$ (or $2^{-n/d}$).

Let me phrase what we discovered in the previous Example more generally/from a higher vantage point:

prop:ZnZgenerators

Proposition 2.4.13: Order n elements of $\mathbb{Z}/n\mathbb{Z}$ (called "generators")

Order n elements ("generators") of $\mathbb{Z}/n\mathbb{Z} = (\mathbb{Z}/n\mathbb{Z})^\times$. Euler totient function $\varphi(n) := |(\mathbb{Z}/n\mathbb{Z})^\times|$. Can use such order n elements to understand the elements $\in \mathbb{Z}/n\mathbb{Z}$ of all other orders; see Comp. comp:allOrdersCyclic 2.4.14.

allOrdersCyclic

Computation 2.4.14: Elements of all orders in cyclic groups

In general, if there is an order $n := |G|$ element $\mathfrak{g}$ of group G (written multiplicatively) — such groups are called cyclic (cf. "clock visualization" in Ex. ex:Z13Zexample1 2.4.11) — then we have converse to Lagrange theorem: for every d dividing $n := |G|$, the element $\mathfrak{g}^{|G|/d}$ will have order d.

We can actually figure out all the order d elements in G; in fact we have already done most of the work. I claim $\langle \mathfrak{g} \rangle^{n/d} = G_d := \{g \in G : g^d = 1\}$. The direction ($\subseteq$) is easy, and other direction ($\supseteq$) follows from $(\mathfrak{g}^e)^d = 1 \iff ed \in n \cdot \mathbb{Z}$ (cf. end of Def. def:orderElement 2.4.8).

Note we previously proved a similar result for U, but in that case the "other direction" followed from cardinality considerations coming from Cor. cor:numRootsLeqDegree 2.4.6 applied to $U \subseteq$ a field.

Now can apply Prop. prop:ZnZgenerators 2.4.13 to $\langle \mathfrak{g}^{n/d} \rangle \xleftrightarrow{\exp_{\mathfrak{g}}} \frac{n}{d}\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/d\mathbb{Z}$. This looks completely bonkers so see below Example.

ex:Z13Zexample3

Example 2.4.15: Example: $U := (\mathbb{Z}/13\mathbb{Z})^\times$

Continuing Ex. ex:Z13Zexample2 2.4.12.

$\mathfrak{g} = 2, d = 4, n = 12$: $\langle \mathfrak{g}^{n/d} \rangle = \langle 8 \rangle = \{8, 12, 5, 1\} = \{2^3, 2^6, 2^9, 2^{12}\} = U_4 \xleftrightarrow{\exp_2} \{3, 6, 9, 12\} \mod 12 =: 3\mathbb{Z}/12\mathbb{Z}$ (which is indeed a subgroup of $(\mathbb{Z}/12\mathbb{Z}, +)$!) $\cong \{1, 2, 3, 4\} \mod 4 =: \mathbb{Z}/4\mathbb{Z}$.

$O_3 = \{5, 8\} = \{2^9, 2^3\} \xleftrightarrow{\exp_2} \{3, 9\} \mod 12 \cong \{1, 3\} \mod 4$, where $1, 3$ are indeed the numbers in $(4]$ relatively prime to 4.

$O_4 = \{3, 9\} = \{2^4, 2^8\} \xleftrightarrow{\exp_2} \{4, 8\} \mod 12 \cong \{1, 2\} \mod 3$, where $1, 2$ are indeed the numbers in $(3]$ relatively prime to 3.

$O_6 = \{4, 10\} = \{2^2, 2^{10}\} \xleftrightarrow{\exp_2} \{2, 10\} \mod 12 \cong \{1, 5\} \mod 6$, where $1, 5$ are indeed the numbers in $(6]$ relatively prime to 6.

cor:totientSum

Corollary 2.4.16: Euler totient sum formula

Taking $G = (\mathbb{Z}/n\mathbb{Z}, +)$ for $n \geqslant 2$, the previous Computation comp:allOrdersCyclic 2.4.14 tells us

$$n = |\mathbb{Z}/n\mathbb{Z}| = \sum_{d|n} |O_d| = \sum_{d|n} \varphi(d).$$

This is a result about pure numbers! There is of course a purely number theoretic proof, but again I'm just using these more grounded results to teach you about abstract algebra. I also think anyways this abstract group theory proof is cleaner than the pure number theoretic proof.

Returning to $U \coloneqq (\mathbb{Z}/p\mathbb{Z})^\times = \mathbb{Z}/p\mathbb{Z} \setminus \{\overline{0}\}$, cardinality $n \coloneqq |U| = p-1$.

Because $O_d \subseteq U_d$, I claim that order d elements of U_d (which is a group) are exactly the order d elements of U.

We see that if $O_d \neq \varnothing$, then for any $x \in O_d$ we will have $\langle x \rangle \subseteq U_d$, which is in fact an equality by cardinality considerations:

$$d = |\langle x \rangle| \leqslant |U_d| \leqslant d.$$

But then we have subgroup $U_d \subseteq U$ (cardinality d) is generated by x, so Computation 2.4.14 comp:allOrdersCyclic tells us $|O_d| = \varphi(d)$.

Conclude O_d is either $\varnothing$ or size $\varphi(d)$ (Euler totient function), so

$$|U| = \sum_{d|n} |O_d| \leqslant \sum_{d|n} \varphi(d), \quad \text{with equality iff all } O_d \neq \varnothing. \qquad \text{eq:totientBound (✓)}$$

Applying Cor. 2.4.16 cor:totientSum on the RHS to $n \coloneqq |U| = p-1$, we get a sandwich inequality forcing equality in eq:totientBound, and hence all O_d for $d \mid n$ (in particular $d = 4$ when $p = 4k+1$) must be non-empty. ■

Taking $d \coloneqq n$, conclude

clicUnitGroupZpZ

Theorem 2.4.17: Unit group of $\mathbb{Z}/p\mathbb{Z}$ is cyclic

$(\mathbb{Z}/p\mathbb{Z})^\times$ is generated by some $\mathfrak{g}$, and $\exp_{\mathfrak{g}}$ gives a group isomorphism to $(\mathbb{Z}/\langle p-1\rangle, +)$ (yes, $\langle p-1\rangle = (p-1)\mathbb{Z}$; I just didn't use that notation in the quotient because it looks bad).

groupFieldCyclic

Exercise 2.4.18: Finite multiplicative subgroup of any field is cyclic

Go over above proof carefully; it in fact shows this stronger statement.

Very abstract and non-constructive proof. But unfortunately, humans literally do not have a formula/theorem for finding elements of desired orders in $(\mathbb{Z}/p\mathbb{Z})^\times$. It is even an open question whether or not $2 \in (\mathbb{Z}/p\mathbb{Z})^\times$ is a generator for infinitely many p.

2.4.3.3.1 Diffie-Hellman cryptographic key exchange, grand reveal

2.4.3.4 Concluding point count in $d = 2$

th:d2pointCount

Theorem 2.4.19: Point count mod p in $d = 2$

p	Z	A
2	2	2
$4k+1$	$2p-1$	$p-1$
$4k+3$	1	$p+1$

c:d2pointCountZ

Exercise 2.4.20: Point count in $\mathbb{Z}$ in $d = 2$

Use the above result to give an independent proof of Fermat's 2-square "Christmas" theorem (i.e. one without using Mathologer's windmill involutions). The only hard part is to show

$p = 4k+1 \implies$ solutions to $x^2+y^2=p$.
Hint (oops I guess I just did it for you): the mod p point count tells us we have $2p-1$ solutions to $x^2+y^2 = p\cdot c$ $(x,y>0)$ for $c = [0..2p)$ ($x=y=p-1$ is already top corner of 1st quadrant square that we focus on when studying mod p).
Thinking back to 3b1b's "split factors in $\mathbb{Z}[i]$ into 2 columns", ASSUMING p did NOT split in $\mathbb{Z}[i]$, then $x^2+y^2=pc$ would have to have p in one column, and therefore c would have to have a multiple of p to put in the other column. So $c = 0, p$ are the only possible circles for solutions. But $c=0$ has exactly 1 $\mathbb{N}^2$-solution $(0,0)$, and $c=p$ has exactly 2 $\mathbb{N}^2$-solutions $(p,0),(0,p)$ (there's no other way to split p^2 into 2 columns), which in total leads us to the 1 solution mod p scenario which we said above was characteristic of $p=4k+3$.

sec-2-5

2.5 Ditto on Spheres mod p, & Quadratic Reciprocity

sec-2-5

Following my GSO talk (cf. YouTube https://youtu.be/IXBp2UHbky0?si=VM0sBrrthBXU266-).

Know $\#_p^2(c)$ for all c, so actually know point count for 3D-sphere, since 3D-sphere is just many slices of 2D-circle! (Think 3D sphere passsing through 2D flatland; or 4D sphere passing through our 3D space.)

2.5.1 Sample messy discovery process for $\#_p^d$ summations

$$\#_p^3(c) = \sum_{z\in\mathbb{Z}/p\mathbb{Z}} \#_p^2(c-z^2)$$

Having z^2 inside of $\#_p^2$ is

With this form, can in fact find closed-form formula: if $c\in(\mathbb{Z}/p\mathbb{Z})^{\times 2}$, then we get $2Z+(p-2)A$. If $c=0$, then get $Z+(p-1)A$. And else get pA. I.e. $\#_p^1\cdot Z + (p-\#_p^1)A$, which $= pA + \#_p^1\cdot(Z-A)$.

Inspired by this, we see a more uniform description of $\#_p^2$, and the previous cases:

$$\#_p^2 = (Z-A)\cdot\delta_0 + A \quad \text{(as functions } \mathbb{Z}/p\mathbb{Z}\to\mathbb{Z})$$

$$Z - A = (-1)^{\frac{p-1}{2}}p, \qquad A = p - (-1)^{\frac{p-1}{2}}$$

$$\begin{aligned}\#_p^3(c) = \sum_{z\in\mathbb{Z}/p\mathbb{Z}} \#_p^2(c-z^2) = pA + (Z-A)\cdot\sum_{z\in\mathbb{Z}/p\mathbb{Z}}\delta_0(c-z^2) &= pA + \#_p^1(c)\cdot(Z-A)\\ &= p\cdot\left(p-(-1)^{\frac{p-1}{2}} + \#_p^1(c)\cdot(-1)^{\frac{p-1}{2}}\right)\\ &= p^2 + (-1)^{\frac{p-1}{2}}\cdot p\cdot(\#_p^1(c)-1)\end{aligned}$$

All of that was "following our nose", just plugging in our uniform formulas for $Z-A$ and A and doing obvious "reverse-distributativity" moves.

We see a "main term" p^2 and smaller order (order p^1) "error term"; notice that p^2 is the "average/expected" number of points in each 3D-sphere (p^3 many points in $(\mathbb{Z}/p\mathbb{Z})^3$, but p buckets=fibers) and similarly 1 is the "average/expected" value of $\#_p^1$, so we get a more uniform formula when we look

at the deviation (from average) quantities $\varepsilon_p^d := \#_p^d - \mathbb{E}\#_p^d = \#_p^d - p^{d-1}$:

$$\varepsilon_p^3 = (-1)^{\frac{p-1}{2}} \cdot p \cdot \varepsilon_p^1.$$

So, the deviation/error quantities have a very clean recursive formula. Our summation for $\#_p^d$ also leads to clean formula for ε_p^d:

$$\begin{aligned} \#_p^d(c) &= \sum_{z \in \mathbb{Z}/p\mathbb{Z}} \#_p^{d-1}(c - z^2) \\ \implies \varepsilon_p^d := \#_p^d(c) - p^{d-1} &= \sum_{z \in \mathbb{Z}/p\mathbb{Z}} \#_p^{d-1}(c - z^2) - p^{d-2} =: \sum_{z \in \mathbb{Z}/p\mathbb{Z}} \varepsilon_p^{d-1}(c - z^2). \end{aligned}$$

So

$$\varepsilon_p^4 = (-1)^{\frac{p-1}{2}} \cdot p \cdot \sum_{z \in \mathbb{Z}/p\mathbb{Z}} \varepsilon_p^1(c - z^2)$$

Looks hard to deal with, but use the same formula backwards to get

$$\varepsilon_p^4 = (-1)^{\frac{p-1}{2}} \cdot p \cdot \sum_{z \in \mathbb{Z}/p\mathbb{Z}} \varepsilon_p^1(c - z^2) = (-1)^{\frac{p-1}{2}} \cdot p \cdot \varepsilon_p^2(c)$$

Same argument (use $d \rightsquigarrow d-1$ summation formula, then inductive $d-1 \rightsquigarrow d-3$ recursion, then $(d-2) \rightsquigarrow (d-2)-1$ summation formula but backwards) to conclude

$$\varepsilon_p^d = (-1)^{\frac{p-1}{2}} \cdot p \cdot \varepsilon_p^{d-2},$$

and therefore by induction, for odd d,

$$\varepsilon_p^d = (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{d-1}{2}} \cdot p^{\frac{d-1}{2}} \cdot \varepsilon_p^1$$

pointCountOddDims

Theorem 2.5.1: Point count in all odd dimensions

For $d \in \mathbb{N}$ odd,

$$\#_p^d = p^{d-1} + \varepsilon_p^d = p^{d-1} + (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{d-1}{2}} \cdot p^{\frac{d-1}{2}} \cdot \varepsilon_p^1.$$

E.g. the point count of the "unit sphere in d-dimensions" $\#_p^d(1) = p^{d-1} + ((-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{d-1}{2}} \cdot p^{\frac{d-1}{2}}$ (because $\varepsilon_p^d(1) = 1$, since 1 is definitely a square mod p for any p lol)

2.5.2 * Alternative, elegant (convolution) proof of recursive formula

Additive combinatorics has very intuitive picture for convolution. Suppose we have $A, B \subseteq \mathbb{Z}/p\mathbb{Z}$ (or more generally, any additive croup G, say). Then we previously saw the sumset $A + B := \{a + b : a \in A, b \in B\}$. But what if we want to keep track of *how many* times a given final value appears, instead of just *what* final values (for the sum) appear? I.e. we want a function ("number of representations")

$$R_{A+B}(c) := |\{(a, b) \in A \times B : a + b = c\}|$$

More generally, let's say we have 2 functions $f_1, f_2 : G \to \mathbb{N}$ s.t. $f_i(x)$ is the number of ways some "process" P_i can produce the value x. (Just like $R_{A+B}(c)$ is the number of ways the "process" of "select any one $a \in A$, any one $b \in B$, and add" can produce the final value c; or $1_A(c)$ is the number of ways

the "process" of "select any one $a \in A$" can produce the final value c.)

Then the convolution of f_1, f_2, notated $[f_1 * f_2]$, evaluated at c, is precisely the number of ways one can get the final value c, by adding together values contributed by process P_1 and process P_2 (where again the value v can be contributed/output by process P_i possibly many times, namely $f_i(v)$ many times). Can think of this/call this the sum process of P_1 and P_2.

Example 2.5.2: Convolution

For example, $[1_A * 1_B](c)$ is exactly the number of ways you can cook up the final value c by adding the output of the "process" of "select any one $a \in A$", and the output of the "process" of "select any one $b \in B$"; i.e. $= R_{A+B}(c)$.

def:convolution

Definition 2.5.3: Convolution, additive combinatorics perspective

Because the "number of ways one can get the final value c, by adding together values contributed by process P_1 and process P_2 (where again the value v can be contributed/output by process P_i exactly $f_i(v)$ many times)" can be calculated by:

- for each possible sum representation $v = v_1 + v_2$ (i.e. for each pair $(v_1, v_2) \in G^2$ s.t. $v_1 + v_2 = v$),
- $f_1(v_1) \cdot f_2(v_2)$ (i.e. the number of ways P_1 can contribute/output v_1, times the number of ways P_2 can contribute/output v_2),
- is the total number of ways the "sum process of P_1 and P_2" can realize the sum representation $v = v_1 + v_2$;

we can write in mathematical notation (again recall $f_1, f_2 : G \to \mathbb{N}$, G is any additive croup),

$$[f_1 * f_2](x) \coloneqq \sum_{\substack{(x_1,x_2)\in G \\ x_1+x_2=x}} f_1(x_1) \cdot f_2(x_2) = \sum_{t\in G} f_1(x-t) \cdot f_2(t) = \sum_{s\in G} f_1(s) \cdot f_2(x-s)$$

(This definition gives commutativity of convolution for free.)

Exercise 2.5.4: Homomorphism property of convolution

(For a function $f : G \to \mathbb{N}$, I write $\sum f$ to mean $\sum_{x\in G} f(x)$.)

$$\sum [f_1 * f_2] = (\sum f_1) \cdot (\sum f_2)$$

(Give a combinatorial proof, i.e. don't work with the sum, but rather the counting contributions interpretation. Meditate on the definitions enough so that this fact feels obvious.)

convolutionRules

Exercise 2.5.5: More exercises to gain familiarity

- Linearity/distributativity $(\lambda_1 f + \lambda_2 g) * h = \lambda_1[f * h] + \lambda_2[g * h]$ (for any scalars λ_1, λ_2).
- $\delta_0 * f = f$.
- [const function] $* f =$ [that const] $\cdot (\sum f)$.

Why is this relevant to us? Well, because counting the solutions to $x^2+y^2+z^2 = c \pmod p$ is exactly the sum process of 3 copies of the "process" of "select $x \in \mathbb{Z}/p\mathbb{Z}$, and square it"; this process contributes the output value v, exactly $\#_p^1(v)$ many times (by defintion of $\#_p^1$)! So we have found that

$$\#_p^d = \underbrace{\#_p^1 * \ldots * \#_p^1}_{d \text{ times}}.$$

A more down-to-earth explanation is

$$\#_p^3(c) = \sum_{z\in\mathbb{Z}/p\mathbb{Z}} \#_p^2(c-z^2) = \sum_{s\in\mathbb{Z}/p\mathbb{Z}} \#p^2(c-s)\cdot\#_p^1(s)$$

(because different z will lead to the same $c-z^2$, or more specifically exactly $\#_p^1(s)$ many z will lead to $c-z^2 = c-s$).

2.5.2.1 Convolution plays nice with averaging

We can recover our recursive $d \rightsquigarrow d-2$ formula from before, using the following general fact about convolution:

lem:convMean0

Lemma 2.5.6: Convolution of mean-0 functions

For any $f, g : G \to \mathbb{R}$,

$$(f - \mathbb{E}f) * (g - [\text{any const}]) = (f * g) - (\mathbb{E}f)(\sum g).$$

Slightly more elegant in terms of $*_{\mathbb{E}}$ (take probabalistically normalized sum $\frac{1}{|G|}\sum$ instead of $\sum$):

$$(f - \mathbb{E}f) *_{\mathbb{E}} (g - [\text{any const}]) = (f *_{\mathbb{E}} g) - (\mathbb{E}f)(\mathbb{E}g).$$

And $\mathbb{E}[f *_{\mathbb{E}} g] = (\mathbb{E}f)\cdot(\mathbb{E}g)$.

Proof: Exc.! Hint: use the rules of Exc. exc:convolutionRules 2.5.5. ■

Exercise 2.5.7: Recursive formula for ε_p^d using convolution magic

Apply previous Lemma lem:convMean0 2.5.6 to $f := \#_p^{d-2}$ and $g := \#_p^2$ (and choose the value of "[any const]" to make $(g - [\text{any const}])$ as simple as possible!).

2.5.3 Baby's first group action

("baby's first" sentence/phrase pattern)

Let us look more closely at the exact point count for the unit sphere $\mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/\mathtt{p}\mathbb{Z}]$ in d-dimensions from Thm. th:pointCountOddDims 2.5.1:

$$|\mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]| =: \#_p^d(1) = p^{d-1} + (-1)^{\frac{p-1}{2}}\cdot(-1)^{\frac{d-1}{2}}\cdot p^{\frac{d-1}{2}}.$$

Lots and lots of factors of p, very very divisible by p. Lots of p vibes.

But recall we are counting d-tuples $(x_1,\ldots,x_d)$ obeying the equation $\mathtt{C}_1^{\mathtt{d}} =: [\mathtt{x}_1^2+\ldots+\mathtt{x}_{\mathtt{d}}^2 = 1]$. There is a natural "$d$-fold symmetry" to such tuples, namely we can cycle (with period d): $(x_1,x_2,\ldots,x_d) \rightsquigarrow (x_2,x_3,\ldots,x_d,x_1) \rightsquigarrow \ldots \rightsquigarrow (x_d,x_1,\ldots,x_{d-1}) \rightsquigarrow (x_1,x_2,\ldots,x_d)$.

A cycle-period of a tuple $\vec{x}$ is any $n \in \mathbb{Z}$ s.t. n cyclic shifts of $\vec{x}$ gets back to $\vec{x}$. (Negative n means cyclic shift in the opposite direction.)

ePeriodsAreIdeal

Exercise 2.5.8: Cycle-periods are an ideal of $\mathbb{Z}^d$

Check that the set of cycle-periods ($\subseteq \mathbb{Z}$) forms an ideal of $\mathbb{Z}$.

We thus conclude (recall Prop. prop:ZisPIR 2.3.28) the minimal positive cycle-period must divide d.

Thus if d is prime, the only possible minimal (positive) cycle-periods of these d-tuples $\vec{x}$ are 1 and d.

I.e. solutions to the equation $\mathtt{C}_1^{\mathtt{d}} := [\mathtt{x}_1^2 + \ldots + \mathtt{x}_{\mathtt{d}}^2 = 1]$ come in cycles of size d, or lie on the diagonal where $x_1 = \ldots = x_d$.

[Draw big circle/blob, tiled with smaller blobs but only of 2 sizes: singleton blob, or size-d blob. This picture shows the set of points $\mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]$ partitioned into its orbits, as defined below.]

We will show that the number of "diagonal solutions" is $\leqslant 2$. (You could try to do it now as an Exercise.)

So the set of points $\mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]$ has a lot of d vibes to it. The fact that it simultaneously has lots of d vibes, and lots of p vibes, is very unexpected, and is the key behind quadratic reciprocity which we finish proving in the next subsection.

2.5.3.1 * Group actions

In this subsubsection, I present a bigger-picture perspective on this "d-fold symmetry" idea we exploited above.

Example 2.5.9

The above "d-fold symmetry" of cyclic shifts is an example of a group action of the group $\mathbb{Z}/d\mathbb{Z}$ on the set $\mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]$.
The collection of all d-tuples (starting with say $\vec{x} \in \mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]$) you get in one such d-cycle is called the orbit of $x \in \mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]$ under the action of $\mathbb{Z}/d\mathbb{Z}$.

def:groupAction

Definition 2.5.10: Group action

A group action of a group G on a set X is a (binary=2-input) function $\curvearrowright : G \times X \to X$ written as "$g \curvearrowright x = y$", meaning "g acts on x by "turning" x into y". If you see $() \curvearrowright ()$, the thing on the left MUST be an element of the group G, and the thing on the right MUST be an element of the set X.
This function must obey the following intuitive rules: $g_1 \curvearrowright (g_2 \curvearrowright x) = (g_1 g_2) \curvearrowright x$ (for all $g_1, g_2 \in G$ and $x \in X$); and $e_G \curvearrowright x = x$ (for all $x \in X$).
Sometimes the notation $\curvearrowright$ is supressed.
The orbit of $x_0 \in X$ (under G), denoted Gx_0 or $\mathrm{Orb}_G(x_0)$, is the set of all elements of X one can reach by applying group actions to x_0; i.e. $Gx_0 := \{g \curvearrowright x_0 : g \in G\}$.

In a very similar spirit to Cor. cor:orderDividesCardinality 2.4.10 (which came from Lagrange Thm. th:lagrangeTheorem 2.4.4, which itself was already "repeat theorem" of older results Prop. prop:cosetsAsFibers 2.3.24), we have

videsCardinality

Corollary 2.5.11: Orbit sizes divide cardinality of group

Fix $x_0 \in X$. Then $|Gx_0|$ divides $|G|$.

which is a corollary of the famous

abilizerTheorem

Theorem 2.5.12: Orbit-stabilizer theorem

Fix $x_0 \in X$. Then $|G| = |Gx_0| \cdot |G^{\circlearrowleft x_0}|$.

Proof: after fixing $x_0 \in X$, we can now consider the function $[\bullet \curvearrowright x_0] : G \to X$. The image of this function is precisely the orbit Gx_0! Consider the fibers (subsets of G) of this surjective function $[\bullet \curvearrowright x_0] : G \twoheadrightarrow Gx_0$. The fiber "above" x_0 is called the stabilizer of x_0 in G; it is a subgroup of G, denoted $\mathrm{Stab}_G(x_0)$ or $G^{\circlearrowleft x_0}$.

I claim the fibers are precisely the (left-)cosets $G/G^{\circlearrowleft x_0}$ (Exc! — though not strictly necessary for next sentence). The point is then we know (Thm. th:lagrangeTheorem 2.4.4) that the left-multiplication bijections $g\ast : G \xrightarrow{\cong} G$ map each fiber/coset to another (possibly the same) fiber/coset; and thus all the fibers/cosets are in bijection hence the same size. ■

Me doing Exc: a left coset $gG^{\circlearrowleft x_0}$ clearly $\subseteq$ fiber f_{gx_0} above gx_0. Also by definition $G^{\circlearrowleft x_0} = \mathrm{f}_{x_0}$. Clearly map $g\ast : G \xrightarrow{\cong} G$ maps $\mathrm{f}_{x_0} \to \mathrm{f}_{gx_0}$, and $g^{-1}\ast : G \xrightarrow{\cong} G$ maps $\mathrm{f}_{gx_0} \to \mathrm{f}_{x_0}$, so we have bijection of the 2 fibers. To conclude fibers=cosets, observe each coset $\subseteq$ one fiber, and both cosets and fibers partition G so "$\subseteq$" must upgrade to equality.

2.5.4 Quadratic Reciprocity, Finale

Above, studying cyclic shifts of d-tuples (or appealing to the general theory of group actions) we see that for d prime (!!!), the unit sphere points $\mathtt{C}_1^{\mathtt{d}}[\mathbb{Z}/p\mathbb{Z}]$ decompose into some number of size d orbits, and some number of size 1 orbits = "diagonal" d-tuples $(\underbrace{z, \ldots, z}_{d \text{ times}})$ satisfying $d \cdot x^2 = 1 \iff (z^{-1})^2 = d \pmod p$.

Well, how many z's are there? It is precisely given by $\#_p^1(d) = 1 + (\frac{d}{p})$ (recall the Legendre symbol)!

So the unit sphere point count equation now reads

$$[1 + (\tfrac{d}{p})] + d \cdot (?) = p^{d-1} + (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{d-1}{2}} p^{\frac{d-1}{2}}.$$

We now reduce mod d, so we don't have to worry about the "(?)" term.

By Cor. cor:orderDividesCardinality 2.4.10,

$$p^{d-1} \equiv 1 \mod d.$$

Remains to figure out $p^{\frac{d-1}{2}} \mod d$. Must be ± 1, since it squares to 1 and the only solutions (in a field) to $\bullet^2 = 1$ are ± 1.

Ok, well look at equation $x^{\frac{d-1}{2}} - 1$. Must have $\leqslant \frac{d-1}{2}$ roots in $\mathbb{Z}/d\mathbb{Z}$ (Cor. cor:numRootsLeqDegree 2.4.6), but all non-zero squares $(\mathbb{Z}/d\mathbb{Z})^{\times 2} \subseteq (\mathbb{Z}/d\mathbb{Z})^\times$ (the squares of the lower half $1, \ldots, \frac{d-1}{2}$ are all the non-zero squares — since higher numbers are just negations of lower numbers; and they are unique because again by Cor. cor:numRootsLeqDegree 2.4.6 there are $\leqslant 2$ things that square to any given value) account for already $\frac{d-1}{2}$ roots! In conclusion,

ulerOrder4Square

Proposition 2.5.13: Euler theorem on order-4 elements and squares

For d prime (so that $\mathbb{Z}/d\mathbb{Z}$ is a field), Cor. cor:numRootsLeqDegree 2.4.6 (the main tool) gives us that $x^{\frac{d-1}{2}} \pmod d$ is precisely $= (\frac{x}{d})$ (the Legendre symbol).

Exercise 2.5.14

Alternative way of arguing previous paragraph. Consider group homomorphism $\bullet^2 : (\mathbb{Z}/d\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^{\times 2}$. Fibers are all the same size, easy to see each fiber size $\geqslant 2$, can't get larger by cardinality reasons, so must be exactly $= 2$.

So, the final conclusion is that

$$1 + \left(\frac{d}{p}\right) = 1 + (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{d-1}{2}} \cdot \left(\frac{p}{d}\right) \mod d,$$

where the possible values for each side are $-1, 0, 1$. So equality mod d (d odd prime!) in fact upgrades to equality in $\mathbb{Z}$:

raticReciprocity

Theorem 2.5.15: Gauss's law of quadratic reciprocity

For p, d odd primes,

$$\left(\frac{d}{p}\right) = (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{d-1}{2}} \cdot \left(\frac{p}{d}\right)$$

What a ride!!!

2.5.5 * Unit group of $\mathbb{Z}/n\mathbb{Z}$

We spent a while on proving that the unit group $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic; this was a major theorem on our quest to count $\#\mathtt{C}^{\mathtt{d}}_{\mathtt{a}}[\mathbb{Z}/\mathtt{p}\mathbb{Z}] =: \#^{\mathtt{d}}_{\mathtt{p}}(\mathtt{a})$, in fact shortly after we proved it we deduced Thm. 2.5.1. th:pointCountOddDims

One may then be curious, and ask what happens in general for $\mathbb{Z}/n\mathbb{Z}$.

Recall $(\mathbb{Z}/n\mathbb{Z})^\times$ has size $\varphi(n)$, or the number of elements of $\{1, \ldots, n\}$ relatively prime to n. For $n = p$, this is easy. For general n, this is quite hard, as it depends on what prime factors are in n. However, for $n = p^k$, this is still easy: $\varphi(p^k) = p^{k-1}(p-1)$ (Exc!), and the elements $\in \{1, \ldots, p^k\}$ coprime to p are exactly those of the form $** \text{❁}$ in base p (I take $k = 3$ when I write out examples), where $\text{❁} \in (\mathbb{Z}/p\mathbb{Z})^\times$ denotes a non-zero element. There is a simple product formula for $\varphi(n)$, but finding it by elementary methods is surprisingly tricky; I will prove it later when I develop product rings/groups and the Chinese remainder theorem, but for now, the fact that I am focusing on $\varphi(p^k)$ is justified "by necessity".

We build off of what we have with $\mathbb{Z}/p\mathbb{Z}$, which recall is the same as working only looking at the rightmost digit in base p.

$\mathbb{Z}/p^k\mathbb{Z}$ is working with the rightmost k digits in base p. We already know that in base p, we can multiply by $00a$ (I'll treat $k = 3$ in my written out examples) to get something that ends in 1.

Multiplying cba by $**1$ is same as $ba0 + cba$, so last digit is a and 2nd to last digit is $a + b$.

Another way of thinking about it is $(ap + 1)(bp + 1) = abp^2 + (a + b)p + 1$.

Ok, so we have $1 + p \in (1 + p\mathbb{Z})/p^k\mathbb{Z}$ have order p^k and so generate the whole (multiplicative) group. I.e. the map $[1 \mapsto 1 + p] : (\mathbb{Z}/p^k\mathbb{Z}, +) \xrightarrow{\cong} (1 + p\mathbb{Z})/p^{k+1}\mathbb{Z}$ is a group isomorphism (more generally, $a \mapsto (1 + p)^a$, i.e. the map is $\exp_{1+p}$).

Will expect logarithm to be isomorphism backward. Amazingly, $\ln = \log_e$ Taylor series is perfect

for the job, $\ln : (1+p\mathbb{Z})/p^{k+1}\mathbb{Z} \to p\mathbb{Z}/p^{k+1}\mathbb{Z} \cong \mathbb{Z}/p^k\mathbb{Z}$ makes perfect sense, i.e. the map is

$$\frac{1}{p}\ln(1+p\bullet) : (\mathbb{Z}/p^k\mathbb{Z}, \star) \xRightarrow{\cong} (\mathbb{Z}/p^k\mathbb{Z}, +).$$

Bijective because inverse

$$\frac{e^{p\bullet}-1}{p} = \frac{\exp(p\bullet)-1}{p} : (\mathbb{Z}/p^k\mathbb{Z}, +) \xRightarrow{\cong} (\mathbb{Z}/p^k\mathbb{Z}, \star).$$

Compare this to the other map we found:

$$\frac{(1+p)^\bullet - 1}{p} : (\mathbb{Z}/p^k\mathbb{Z}, +) \xRightarrow{\cong} (\mathbb{Z}/p^k\mathbb{Z}, \star).$$

sec-2-6

2.6 §2 Algebra, Take 2: Abstraction Boogaloo

sec-2-6

I'll take this space to give a more streamlined version of the abstract algebra theory development I scattered above. It's a difficult pedagogical situation. But maybe this way I can give you both worlds.

What I essentially showed above (in the "language" of specific special rings R like $\mathbb{Z}, F[x], \mathbb{Z}[i]$:

$$\text{EuD} \implies \left\{ \begin{array}{c} \text{finiteness of element divisibility chains} \\ \text{PID (Bezout)} \implies [\text{irreducible} \implies \text{prime}] \end{array} \right\} \implies \text{UFD}$$

I am reminded when I took 246C with Terry, covering the circle packing theorem. He told us that set of notes was the "excuse" he used to actually teach us about quasiconformal maps. Likewise, counting lattice points on circles is my "excuse" for teaching you about abstract algebra.

* Oyster World ~Layer 1~

Oyster World in reference to "the world's your oyster" After Chapter 1, we have many possible parallel routes. Galois theory is standard, but could equally well begin Dedekind theory, or Nullstellensatz, Linear Algebra from module-theoretic perspective, or so on. I wrote a flowchart in iPad for the possibilities.

Dirichlet Analytic Warning

However, I did put Dirichlet stuff after some analytic number theory. Kind of far off, not super belonging to this "pure algebra" textbook. But that is also a parallel track, just depends on whether or not the lecturer/students want care about number theory enough to first do some preliminary, non-algebraic work.

chap-3

3.0 Galois

sec-3-1

Greene 110A homework 3 was actually a pretty good practical introduction. See my solutions for that homework.

sec-3-1

3.1 Motivation & Set-up

3.1.1 Revisiting $\mathbb{C}$

exc:quotientComplexNumbers

(Similar to Exc. 2.3.26 in §2.3.3.2.1) Let us revisit definition of $\mathbb{C}$ again. We discussed semantics thoroughly in §2.1, but syntax was just manipulation rules on formal symbol i. In Exc. 2.3.26, we discussed how the general quotient ring construction provides a framework for "manipulation rules on formal symbols".

So, we now see that $\mathbb{C}$ can be formulated cleanly syntactically as a quotient ring $\mathbb{R}[x]/\langle x^2+1\rangle$, where x and $-x$ both square to -1. Yes, technically I should prove right now this is isomorphic to the semantic version of $\mathbb{C}$ (as a certain subring of linear transformations $\mathrm{Lin}_{\mathbb{R}}(\mathbb{R}^2)$), but it is easier to deal with simpler Ex. ?? (below) first.

ex:Qsqrt2

But now, which element should we call i? x or $-x$?

There is no way to distinguish algebraically, because $x \mapsto -x$ from $\mathbb{R}[x]/\langle x^2+1\rangle$ to itself is an isomorphism of fields/rings. (Geometrically, we may distinguish.)

For people familiar with analysis, perhaps this should be a bit of surprise, since all previous constructions, going from $\mathbb{N} \rightsquigarrow \mathbb{Z} \rightsquigarrow \mathbb{Q} \rightsquigarrow \mathbb{R}$ were all canonical ("unique up to unique isomorphism" in comparison to what's happening here). But for this last "expansion", we have 2 possible, distinct but isomorphic choices.

Some practical examples:

ex:Qsqrt2start

Example 3.1.1: Start of $\mathbb{Q}(\sqrt{2})$

We should expect $\mathbb{Q}[x]/\langle x^2-2\rangle$ to be (ring) isomorphic to $\mathbb{Q}(\sqrt{2}) \subset \mathbb{R}$, via $\overline{x} \mapsto \sqrt{2}$.
Let us discuss the technicalities of this next. Continued in Ex. 3.1.4

ex:Qsqrt2end

3.1.2 Universal property of quotient

In general, how can we/should we think about ring homomorphisms $R/I \to S$? Let us try to translate this back to a notion on the "level" of R (indeed, usually it is easier to work — at least syntactically, since elements of quotients are cosets which have multiple representatives — with R than R/I).

Any ring homomorphism $\Psi : R/I \to S$ can be precomposed with the canonical projection $\pi : R \twoheadrightarrow R/I$ to get a ring homomorphism $\psi : R \xrightarrow{\pi} R/I \xrightarrow{\Psi} S$. Indeed,

$$\Psi(r+I) = s \iff \psi(r+I) = \{s\}$$

(the former being function evaluated at a point of the domain; and the latter being the image of a function on a subset of the domain).

So, $\psi : R \to S$ "comes from" some $\Psi : R/I \to S$ (by precomposition with π) $\iff$ ψ applied to all cosets $r + I \subseteq R$ yields a singleton. Equivalently, because (by additivity) $\psi(r+I) = \psi(r) + \psi(I)$, the previous sentence is $\iff$ $\psi(I) = 0$.

univPropQuotRing

Theorem 3.1.2: Universal property of quotient in *Ring*

Or $\Psi : R/I \to S$ correspond to $\psi : R \to S$ s.t. $\psi(I) = 0$.
The above is a proof already, but Exc.: check that $\Xi := [\psi \mapsto [x + I \mapsto \psi(x)]]$ is well-defined, and also $\Upsilon := [\Psi \mapsto [x \mapsto \Psi(x + I)]]$ is well-defined, and they are inverses.

th:1stIsoRings

Theorem 3.1.3: 1st isomorphism theorem (rings)

ex:Qsqrt2end

Example 3.1.4: End of $\mathbb{Q}(\sqrt{2})$ (isomorphisms!)

Continuing from Ex. 3.1.1. ex:Qsqrt2start
Discuss isomorphism $\mathbb{Q}(\sqrt{2}) \to \mathbb{Q}(\sqrt{2})$ mapping $\sqrt{2} \mapsto -\sqrt{2}$.
Can define such a map as a $\mathbb{Q}$-VS linear map by specifying action on basis, and have additivity automatically, but checking multiplicativity is a pain. That's why the above ring-theoretic machinery is better suited.

3.1.2.1 * Universal property of quotients in general (sets)

$f : X \to Y$, partition $\mathcal{P}$ of X (equivalently an equivalence relation: $x \sim y :\iff x, y$ in same piece of partition; equivalence classes are pullback via $X \twoheadrightarrow \mathcal{P}$ of literal equality in $\mathcal{P}$), if f is "$\mathcal{P}$-measurable", then can factor though to $f : X \twoheadrightarrow \mathcal{P} \to Y$.

univPropQuotSet

Theorem 3.1.5: Universal property of quotient in *Set*

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow & & \downarrow \\ A' & \dashrightarrow & B' \end{array}$$

there is "push-forward" $f' : A' \to B'$ iff f maps fibers of π_A inside fibers of π_B.

(Throwback to norm map Nm mod n, Eq. 2.1) eqn:NmZmodn

3.1.3 Galois heart: permutations of alg. #s invisible to alg. eqs.

So, we have seen some examples where we can interchange elements in a way that is [invisible to]/[undetectable by] all (integer coefficient) polynomials — that is what ring homomorphism means.

Invisible to $\mathbb{Z}[x]$ polynomials $\iff$ invisible to $\mathbb{Q}[x]$ polynomials. Necessarily must fix $\mathbb{Q}$. Necessarily must map algebraic α to another root of its irreducible polynomial $\mathrm{Irr}(\alpha, \mathbb{Q})$.

Restricting attention to algebraic numbers $\overline{\mathbb{Q}}$ (can think of it as a subset of $\mathbb{C}$, supposing that the reader is familiar with say the (complex) analytic proofs of $\mathbb{C}$ being algebraically closed), we are interested in the phenomenon of maps $\overline{\mathbb{Q}} \to \overline{\mathbb{Q}}$ that preserve all $\mathbb{Z}$-polynomial (or equiv. $\mathbb{Q}$-polynomial) relations ($\iff$ preserve addition and multiplication and units, i.e. aumu-map).

def:absGaloisQ

Definition 3.1.6: Absolute Galois group

The set of such (aumu-)maps is called $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. This is perhaps the most important object in all of algebraic number theory.

We can/will give a sort of recipe to build up all such (aumu-)maps, by enumerating the countably many algebraic numbers $\alpha_1, \ldots$ in a "non-redundant" way, and then extending the trivial embedding $\mathbb{Q} \hookrightarrow \overline{\mathbb{Q}}$ to $\mathbb{Q}(\alpha_1) = \mathbb{Q}[\alpha_1] \hookrightarrow \overline{\mathbb{Q}}$, then $\mathbb{Q}(\alpha_1, \alpha_2) = \mathbb{Q}(\alpha_1)[\alpha_2] \hookrightarrow \overline{\mathbb{Q}}$, and so on.

Fields naturally come in because we'll see $\overline{\mathbb{Q}}$ is a field, and because preserving all $\mathbb{Z}$-polynomials in particular fixes sub*field* $\mathbb{Q}$, we will always here be working over base *field* $\mathbb{Q}$, where even adjoining algebraic α ring-theoretically: $\mathbb{Q}[\alpha]$, gives automatically a field.

Let me first lay out the machinery cleanly before fleshing the previous 2 paragraph out in detail (first in an example, $\alpha_1 = \sqrt{2}, \alpha_2 = \sqrt{3}$).

3.1.4 Formal polynomials, universal property of free ring extensions

Proper def. of formal polynomials. Before now, working over $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$, can think of polynomials as polynomial functions. We stop thinking about them primarily as functions now, and instead more as formal syntactic objects.

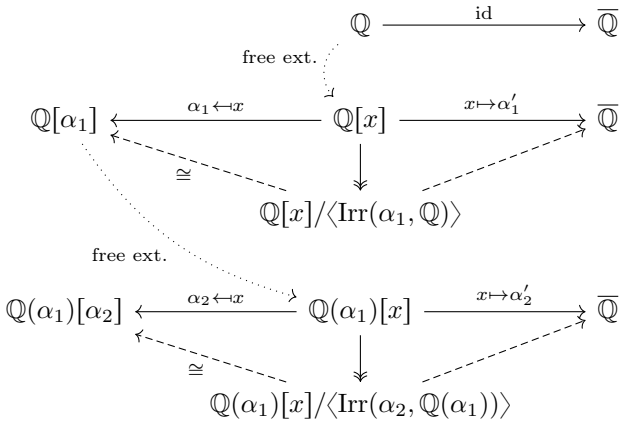
univPropFreeExt

Theorem 3.1.7: Universal property of free extension

For (c)ring R, the [ring homomorphisms $\mathrm{Hom}_{\underline{\underline{\mathbf{Ring}}}}(R[x], S)$] correspond to [ring homomorphisms $\mathrm{Hom}_{\underline{\underline{\mathbf{Ring}}}}(R, S)$ and an element $s \in S$ (to which to send x)].

3.1.4.1 Combine with previous univ. prop. for quotients

Brandenburg Lemma 2.1



"Non-redundant" enumeration of algebraic numbers $\alpha_1, \alpha_2, \ldots$ in the sense that α_{n+1} is not already in $\mathbb{Q}(\alpha_1, \ldots, \alpha_n)$.

3.1.5 Example: $\mathbb{Q}(\sqrt{2}, \sqrt{3})$

3.1.6 Problem 1

Prove: if $\frac{m}{n}$ is a reduced fraction (i.e. $m, n \in \mathbb{N}^+, \gcd(m, n) = 1$), then: $\frac{m}{n}$ is the square of a rational number if and only if m, n are both squares of integers, i.e. $\exists a, b \geqslant 1$ s.t. $m = a^2$ and $n = b^2$.

3.1.7 Problem 2,3

Use Problem 1 to show $\sqrt{3}$ can not be expressed in the form $a + b\sqrt{2}$ for $a, b \in \mathbb{Q}$.

I know this was on the previous problem set, but most students ended with "we know $\sqrt{\frac{3}{2}} \notin \mathbb{Q}$". So please use Problem 1 to give a **complete** proof of $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$, with NO missing steps.

3.1.8 Problem 4

Explain why $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$.

Suggestion: tower property $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] = [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}]$.

3.1.9 Problem 5

1. Explain why Problem 4 implies $\sqrt{2} + \sqrt{3}$ solves a polynomial equation with $\mathbb{Q}$-coefficients of degree $\leqslant 4$.
2. Show that $\sqrt{2} + \sqrt{3}$ does NOT solve an equation of degree 2 over $\mathbb{Q}$.
3. Find an **explicit** polynomial equation with $\mathbb{Q}$-coefficients of degree $\leqslant 4$ that $\sqrt{2} + \sqrt{3}$ solves

3.1.10 Problem 6

Assume (as we shall prove later) that $x^n - 2$ is irreducible over $\mathbb{Q}$ for all $n \in \mathbb{N}^+$. Then show that for all $n \in \mathbb{N}^+$, $[\mathbb{Q}(\sqrt[n]{2}) : \mathbb{Q}] = n$.

Interlude 3.1.8

Proof: $\geqslant n$ because $x^n - 2$ irreducible, so no smaller degree polynomials have $\sqrt[n]{2}$ as a root. I'll do this again after I discuss kernel $= \langle x^n - 2 \rangle$ below. Or maybe I'll discuss the Bezout stuff now: WTS $p(\alpha) = 0 \implies p \in \langle x^n - 2 \rangle$. Contrapositive: Using $\mathbb{Q}[x]$ UFD, knowing p is not multiple of $x^n - 2$ and $x^n - 2$ irreducible means p and $x^n - 2$ must be relatively prime. So Bezout gives $p(x) \cdot A(x) + (x^n - 2) \cdot B(x) = 1$, and applying $x = \alpha$ conclude $p(\alpha)$ can't possibly be 0. Fancy Proof later.

Show $\leqslant n$. We show $1, \sqrt[n]{2}, \ldots, (\sqrt[n]{2})^{n-1}$ are a spanning set of the $\mathbb{Q}$-VS $\mathbb{Q}(\sqrt[n]{2})$. To do this, I want to show $\mathbb{Q}(\sqrt[n]{2}) = \mathbb{Q}[\sqrt[n]{2}]$, i.e. show the RHS is a field. Let me denote $\alpha \coloneqq \sqrt[n]{2}$ for convenience.
Elements of $\mathbb{Q}[\alpha]$ are $\mathbb{Q}$-polynomials in α, so notate as $p(\alpha)$. Must assume it's non-zero. So $p(\alpha) \neq 0$. This already will force using $x^n - 2$ irreducible, i.e. $p(x)$ relatively prime to $x^n - 2$. So actually this proof is exactly the same proof as more abstract proof if you unwrap the fancy language.
Let me do Fancy Proof: we show the surjective ring homomorphism $\mathbb{Q}[x] \to \mathbb{Q}[\alpha]$ given by $p \mapsto p(\alpha)$ descends to ring isomorphism $\mathbb{Q}[x]/\langle x^n - 2 \rangle \to \mathbb{Q}[\alpha]$ (1st isomorphism theorem), and we know kernel is exactly $\langle x^n - 2 \rangle$ because PID [division algorithm forces everything in an ideal to be multiple of lowest norm element], and $x^n - 2$ irreducible is in kernel.
Irreducible so Bezout implies maximal ideal [indeed in general, Bezout tells us in PID, prime implies maximal], which implies LHS is a field.

3.1.11 Problem 7

Let $\overline{\mathbb{Q}}$ denote the set of all real numbers that are algebraic over $\mathbb{Q}$. Show that $\overline{\mathbb{Q}}$ is a field.

Interlude 3.1.9

Proof: comm, assoc, distr inherited from $\mathbb{C}$. Identities obviously contained. Additive inverse is easy, change explicit polynomial easy.
Multiplicative inverse easy too: α algebraic so $p(\alpha) = 0$ for $\deg(p) \geqslant 1$. Take p of lowest degree, so must have constant term nonzero. Rearrange.
Closure under addition, multiplication require a bit more cleverness, namely the result that $\alpha \in \mathbb{C}$ algebraic over $\mathbb{Q} \iff [\mathbb{Q}(\alpha) : \mathbb{Q}] < \infty$.
$\mathbb{Q}(a+b), \mathbb{Q}(ab) \subseteq \mathbb{Q}(a, b)$. Closure is now easy.

3.1.12 Problem 8

Use Problem 6 to show that $\overline{\mathbb{Q}}$ is NOT finite dimensional over $\mathbb{Q}$.

3.1.13 Problem 9

1. Find ALL the field automorphism of $M \coloneqq \mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$.
2. Show that $\prod_{A \in \mathrm{Aut}}(x - A\alpha)$ has coefficients in $\mathbb{Q}$, where $\alpha = \sqrt{2} + \sqrt{3}$. [The product notation means product of $(x - A\alpha)$ over all $A \in \mathrm{Aut}$ the finite set of automorphisms.]

3. Would this be true for all $\alpha \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$?

Interlude 3.1.10

Proof (a): $\{1, \sqrt{3}\}$ basis of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}(\sqrt{2})$, so $1, \sqrt{3}$ are $\mathbb{Q}(\sqrt{2})$-LI. $\{1, \sqrt{2}\}$ is similarly $\mathbb{Q}$-LI. So claim $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is $\mathbb{Q}$-LI. Proof: suppose $a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6} = 0$ for $a, b, c, d \in \mathbb{Q}$. By $\mathbb{Q}(\sqrt{2})$-LI, must be that $a+b\sqrt{2} = 0$ and $(c+d\sqrt{2}) = 0$. By $\mathbb{Q}$-LI, must be that $a, b, c, d = 0$. So $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is $\mathbb{Q}$-basis.
Field automorphism must send $\sqrt{2}$ to $\pm\sqrt{2}$, and $\sqrt{3}$ to $\pm\sqrt{3}$, because must preserve polynomial $x^2 - 2$, $x^2 - 3$.
So $\leqslant 4$ automorphisms already. Now, we want to show that all 4 are actually ring homomorphisms.
Of course, can do directly, but here's more abstract way. I already did this for $\mathbb{Q}(\sqrt{2})$ below. I'll bring it up.
As I described more in depth in Problem 6, using $x^2 - 2$ irreducible over $\mathbb{Q}$, get 1st iso theorem on $\mathbb{Q}[x] \twoheadrightarrow \mathbb{Q}[\sqrt{2}] \quad \rightsquigarrow \quad \mathbb{Q}[x]/\langle x^2 - 2\rangle \xrightarrow{\cong} \mathbb{Q}[\sqrt{2}]$, where again irreducible means LHS is a field.
Now prove essentially universal property: ring hom $\mathbb{Q}(\sqrt{2}) \cong \mathbb{Q}[x]/\langle x^2 - 2\rangle \to M$ biject with ring hom $\phi : \mathbb{Q}[x] \to M$ mapping $x \mapsto \phi(x) \in \{\pm\sqrt{2}\}$.
Or $\psi : R/I \to S$ correspond to $\phi : R \to S$ s.t. $\phi(I) = 0$. Check that $\Upsilon := [\phi \mapsto [x + I \mapsto \phi(x)]]$ is well-defined, and also $\Theta := [\psi \mapsto [x \mapsto \psi(x + I)]]$ is well-defined, and they are inverses.
This shows that $\mathbb{Q}(\sqrt{2}) \to M$ given by $\sqrt{2} \mapsto \pm\sqrt{2}$ are BOTH ring homs.
Next, we do same thing but now consider $\mathbb{Q}(\sqrt{2})[x]/\langle x^2 - 3\rangle$ (which of course $= M$ — again $x^2 - 3$ irreducible over $\mathbb{Q}(\sqrt{2})$).
Same Universal Property of Quotients says then that ring homs $\mathbb{Q}(\sqrt{2})[x]/\langle x^2 - 3\rangle \to M$ correspond to ring homs $\mathbb{Q}(\sqrt{2})[x] \to M$ mapping $x \mapsto \{\pm\sqrt{3}\}$. Putting together, conclude.

E.g. to construct the field hom. sending $\sqrt{2} \mapsto -\sqrt{2}$ and $\sqrt{3} \mapsto \sqrt{3}$, we start with the map $\mathbb{Q}[x] \to \mathbb{Q}[\sqrt{2}]$ mapping $x \mapsto -\sqrt{2}$ (which exists due to freeness of $\mathbb{Q}[x]$), which factors through (by univ. prop. of quotient) to $\mathbb{Q}(\sqrt{2}) \cong \mathbb{Q}[x]/\langle x^2 - 2\rangle \to \mathbb{Q}[\sqrt{2}]$, i.e. a field automorphism $\mathbb{Q}(\sqrt{2}) \xrightarrow{\cong} \mathbb{Q}(\sqrt{2})$ mapping $\sqrt{2} \mapsto -\sqrt{2}$.
Then, we can extend this (by freeness) to a map $\mathbb{Q}(\sqrt{2})[x] \to \mathbb{Q}(\sqrt{2}, \sqrt{3})$ mapping $x \mapsto \sqrt{3}$. This again factors through (by univ. prop. of quotient) to $\mathbb{Q}(\sqrt{2}, \sqrt{3}) \cong \mathbb{Q}(\sqrt{2})[x]/\langle x^2 - 3\rangle \to \mathbb{Q}(\sqrt{2}, \sqrt{3})$, i.e. a field automorphism $\mathbb{Q}(\sqrt{2}, \sqrt{3}) \xrightarrow{\cong} \mathbb{Q}(\sqrt{2}, \sqrt{3})$ mapping $\sqrt{2} \mapsto -\sqrt{2}, \sqrt{3} \mapsto \sqrt{3}$.

Interlude 3.1.11

Proof (b,c): for (b) can do direct calculation, that's fine; or refer to (c)
For (c), can see by hand that fixed field of $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \text{span}_{\mathbb{Q}}\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ (under total 4 automorphisms) is $\mathbb{Q}$.
Apply now to $p_\alpha := \prod_{A \in \text{Aut}}(x - A\alpha) \in \mathbb{Q}(\sqrt{2}, \sqrt{3})[x]$. Easy to see p_α fixed by all 4 automorphisms, so coefficients must be in $\mathbb{Q}$.

sec-3-2

sec-3-2

3.2 Field automorphism extension/lifting

~~So we have key idea that "averaging/balancing" over group $G := \mathrm{Aut}_{\mathbb{Q}}(\mathbb{Q}(\sqrt{2},\sqrt{3}))$ produces polynomials in $\mathbb{Q}$, the fixed field. Can play same game for subgroups $H \leqslant G$ to get polynomials in $L^{\circlearrowright H}[x]$.~~

See my [MSE] Lifting Lemma for field automorphisms, and simple proofs for the Galois correspondence.

MSE Pierre-Yves Gaillard has proof of FTGT for finite normal extensions generated by (finitely many) separable elements.

ghAutomorphisms

Theorem 3.2.1: Guaranteeing enough automorphisms

Suppose $L \supseteq K$ fields s.t. $L = K[\{\mathfrak{g}_\bullet\}]$ for a generating set $\{\mathfrak{g}_\bullet\}$ of elements algebraic over K. If we have that $\forall \mathfrak{g} \in \{\mathfrak{g}_\bullet\}$, the minimal polynomial $\mathrm{Irr}(\mathfrak{g}, K) \in K[t] \subseteq L[t]$ has $\geqslant 2$ distinct linear factors in $L[t]$ (one of which of course is $t - \alpha \in K[t] \subseteq L[t]$), then

$$L^{\circlearrowright \mathrm{Aut}_K(L)} = K$$

(i.e. for every $\alpha \in L \smallsetminus K$, can find some K-automorphism of L moving α off itself).

I.e. the map $K \mapsto \mathrm{Aut}_K(L)$ mapping subfield of L to subgroup of $\mathrm{Aut}_F(L)$ (for some "base field" $F \subseteq K \subseteq L$) can be undone, by "fixed points" operation.

cor:fmSubfield

Corollary 3.2.2: Finitely many intermediate subfields

sec-3-3

3.3 (Finite) cover concentration theorems

sec-3-3

The proof that is easier for me to remember uses a (finite) "big space big cover"/"small cover small space" theorem for fields (same genre as the Baire category theorem, which is a countable "big space big cover" theorem). EDIT 9/4/25: I came across Nakayama's lemma (cover M by scaled copies of M, namely $I \cdot M$; one copy must already cover M) which I thought fell in the same genre. I thought a better name was "cover singling theorem" (like singling-out a single element from a cover because it is somehow special, like covering everything, or in Baire case, covering an open set). I tried finding synonyms for "single out", found "winnow". Meh. I was then reminded of an exercise from M. Visan's 247B harmonic analysis class, a "(frequency) concentration compactness" result. Much better word. MO Martin Brandenburg

Other words for "big cover" may include "spacious cover", "coarse cover".

iteFieldCovering

Proposition 3.3.1: VS over big field not covered by n subspaces

(Instance of "cover concentration" result.)
Suppose F-VS V covered by n proper subspaces $V_1, \ldots, V_n$. Then $|F| \leqslant n - 1$.

Suffices to prove for *non-reducible covers*.

Proof: suppose we are given $V_1, \ldots, V_n$ a non-reducible cover of V. Picture V_1 as a plane in $\mathbb{R}^3$

(e.g. xy-plane).

Choose a point $n_1 \notin V_1$, so that the "coset" $n_1 + V_1$ is disjoint from V_1, hence covered by $V_2 \cup \ldots \cup V_n$. We aim to create a line passing through n_1 parallel to V_1 (i.e. pick some $v_1 \in V_1 \setminus \{0\}$, and define the line $\ell \coloneqq n_1 + F \cdot v_1 \subseteq V$) s.t. it intersects each of $V_2, \ldots, V_n$ at most once.

We can do this by choosing $v_1 \in V_1 \setminus (V_2 \cup \ldots \cup V_n)$ ($\exists$ by non-reducibility of the cover). This works because if ℓ hit subspace V_j twice (for $j \in [2..n]$), then $\ell \subseteq V_j \implies n_1, v_1 \in V_1$, ※.

In summary, the line ℓ definitely doesn't hit V_1, and intersects each V_j ($j \in [2..n]$]) at most once, so $|\ell| = |F| \leqslant n - 1$. ■

So we have proved the theorem if our non-reducible cover for (infinite) field K has intersection $K_1 \cap \ldots \cap K_n$ infinite.

We now show that this must always be the case. Perhaps one's first instinct is to try to work with vector spaces since then we might be able to utilize the power of linear algebra; but actually for this result we need only the (additive) group structure of K.

iteGroupCovering

Proposition 3.3.2: Non-reducible cover of infinite group has infinite intersection

(Instance of "cover concentration" result.) (I write as additive croup, but the result holds for any group G)
Suppose G is infinite additive croup, and $G_1, \ldots, G_n$ is a non-reducible cover of G (by proper subgroups). Then $G_1 \cap \ldots \cap G_n$ must also be infinite.
This is very good Exercise problem! Please try on your own before reading my thoughts below!

prop:VSinfiniteFieldCovering

Proof: the key idea (= heart of this proof by induction) is one we've already seen above in Prop. 3.3.1 (also in magenta): for subgroup H, take $g \notin H$ to get coset $g + H$ disjoint from H, hence covered by smaller number of G_j.

I guess we also use again the fact that a subgroup S containing 2 elements from a coset $g + H$ means you can subtract to get an element of H in S (in the above VS setting, i.e. S subspace and H a line, this phenomenon is extreme, in the sense that [S contains 2 points of $g+H$] $\implies g+H \subseteq S \implies g, H \subseteq S$). Said another way, "coset shifts" $H \rightsquigarrow g+H$ vanish upon taking difference sets: $(g+H)-(g+H) = H-H$.

Because G is infinite, w.l.o.g. say G_1 is infinite, i.e. $G_1 \subseteq_{\text{inf}} G$.

Choose $g \notin G_1$ ($\exists$ by non-reducibility of the cover). Then $g + G_1$ (an infinite coset) does not intersect G_1, hence must be covered by $G_2 \cup G_3 \cup \ldots \cup G_n$. Say $g + G_1 \subseteq_{\text{inf}} G_2$, i.e. $\exists G_{1,2} \subseteq_{\text{inf}} G_1$ s.t. $g + G_{1,2} \subseteq G_2$. Then $(g + G_{1,2}) - (g + G_{1,2}) = G_{1,2} - G_{1,2}$ is an infinite subset of *both* G_2 and G_1, so $G_1 \cap G_2$ is infinite.

Choose $g \notin G_1 \cup G_2$ ($\exists$ by non-reducibility of the cover). Then $g + (G_1 \cap G_2)$ (an infinite coset) does not intersect G_1 or G_2, hence must be covered by $G_3 \cup \ldots \cup G_n$. Say $g + (G_1 \cap G_2) \subseteq_{\text{inf}} G_3$. Similar argument as before shows $G_1 \cap G_2 \cap G_3$ is infinite (both contain some infinite difference-set $\subseteq G_1 \cap G_2$). Continue. Woohoo!

For (non-abelian) multiplicative group G, same as above: $gG_1 \subseteq_{\text{inf}} G_2$, i.e. $gG_{1,2} \subseteq G_2$, so $(gG_{1,2})^{-1}(gG_{1,2}) = G_{1,2}^{-1}G_{1,2}$ is an infinite subset of G_2 and also G_1, so $G_1 \cap G_2$ is infinite. Etc. ■

chap-4

4.0 Dedekind

chap-5

5.0 Nullstellensatz (⇝Schemes...)

See my sheafNSS.pdf

chap-6

6.0 Conics++ (Intro AG)

See Sandor https://mathoverflow.net/questions/59071/what-elementary-problems-can-you-solve-with-schemes/59136#59136
https://mathoverflow.net/questions/55244/why-must-nilpotent-elements-be-allowed-in-modern-algebraic-geometry/55258#55258

chap-7

7.0 Linear Algebra from Modules

Richard Elman's notes, ask others.

* Oyster World ~Layer 2~

chap-8

8.0 Group Theory

See my MSE post on Sylow theorems.

chap-9

9.0 Representations (Intro)

chap-10

10.0 Extras: Dirichlet

sec-10-1

10.1 Analytic (L-function, number theory) background

sec-10-1

See first third of http://danielrui.com/papers/PNTsurvey.pdf.

Reed notes for Dirichlet class number formula for quadratic number fields: https://people.reed.edu/~jerry/361/lectures/iqclassno.pdf. Covers a bit of Dedekind theory. There, ideal multiplication arises because of $\mathfrak{a}\overline{\mathfrak{a}} = Nm(\mathfrak{a})\mathcal{O}_K$. Easily shows multiplicativity of norm, which easily shows unique factorization into prime ideals.

So key is really multiplicativity of norm. I remembered from my MSE question and Terry Tao notes that norm is size of quotient field and also index of a subgroup. I wondered if there was a geometric interpretation (lattices, etc.) and ChatGPT gave me an interpretation in terms of matrices, in which multiplicativity of norm is exactly multiplicativity of determinant of matrices.

Alphabetical Index

absolute Galois group, 43

fibers
    cosets, 19

integers mod p: $\mathbb{Z}/p\mathbb{Z}$, 21

PIR, 21
    $F[x]$, 21
    testing, 21

universal property, 43
    free, 43
    quotient, 42