Lecture Notes for Abstract Algebra I

James S. Cook
Liberty University
Department of Mathematics

Fall 2016

preface

Abstract algebra is a relatively modern topic in mathematics. In fact, when I took this course it was called *Modern Algebra*. I used the fourth ed. of *Contemporary Abstract Algebra* by Joseph Gallian. It happened that my double major in Physics kept me away from the lecture time for the course. I learned this subject first from reading Gallian's text. In my experience, it was an excellent and efficient method to initiate the study of abstract algebra. Now, the point of this story is not that I want you to skip class and just read Gallian. I will emphasize things in a rather different way, but, certainly reading Gallian gives you a second and lucid narrative to gather your thoughts on this fascinating topic. I provide these notes to gather ideas from Gallian and to add my own.

sources

I should confess, I have borrowed many ideas from:

1. *Contemporary Abstract Algebra* by Joseph Gallian
2. the excellent lectures given by Professor Gross of Harvard based loosely on Artin's *Algebra*
3. Dummit and Foote's *Abstract Algebra*
4. Fraleigh
5. Rotman

style guide

I use a few standard conventions throughout these notes. They were prepared with LaTeX which automatically numbers sections and the hyperref package provides links within the pdf copy from the Table of Contents as well as other references made within the body of the text.

I use color and some boxes to set apart some points for convenient reference. In particular,

1. definitions are in green.
2. remarks are in red.
3. theorems, propositions, lemmas and corollaries are in blue.
4. proofs start with a **Proof:** and are concluded with a □.

However, I do make some definitions within the body of the text. As a rule, I try to put what I am defining in **bold**. Doubtless, I have failed to live up to my legalism somewhere. If you keep a list of these transgressions to give me at the end of the course it would be worthwhile for all involved.

The symbol □ indicates that a proof is complete. The symbol ▽ indicates part of a proof is done, but it continues.

As I add sections, the Table of Contents will get longer and eventually change the page numbering of the later content in terms of the pdf. When I refer to page number, it will be the document numbering, not the pdf numbering.

Contents

Chapter 1

Group Theory

The organization of these notes loosely follows Gallian. For the most part I include every theorem which Gallian includes. However, I include some extra examples and background. I lack interesting quotes[1][2]

1.1 Lecture 1: an origin story: groups, rings and fields

In a different notation, but with the same essential idea, the fact that solutions to $ax^2+bx+c=0$ are given by $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$ has been known for millenia. In contrast, the formula for solutions of the cubic equation ax^3+bx^2+cx+d is only about a half-millenia old. Del Ferro solve the cubic[3] circa 1500, Tartaglia solved it around 1530 then it was published by Cardano in his *Ars Magna* in 1545. Cardano's student Ferrari solved quartic[4] and that can also be found in the *Ars Magna*. Nearly the same tricks give closed form equations for the cubic and quartic. Euler, Lagrange and other 18th century mathematicians knew volumes about how to factor and solve polynomial equations. It seemed it was just a matter of time to find a formula for the solution of

$$ax^5+bx^4+cx^3+dx^2+ex+f=0.$$

But, after a great effort by Lagrange there was no formula forthcoming. Moreover, it began to be clear that such a formula would be impossible due to the structure of Lagrange's study. At the dawn of the nineteenth century Ruffini gave the first (incomplete in 1799 and again in 1813) proofs that there could not exist a general quintic formula. Abel, at the age of 19, gave a complete proof of the non-existence of the quintic formula in 1821. In 1831 a young Frenchman named Evariste Galois found a way to explain when it was possible to find the solutions to a 5-th order polynomial equation (for example, $x^5-1=0$ is easy to solve). Galois insight was to identify the patterns in Lagrange's work which involved permutations of the roots of the equation. In retrospect, this was the birth of Group Theory. In short, Galois said there was a nice solution to a quintic if the Galois group is *solvable*. If a group is *simple*[5] then it cannot be broken down further, they're sort of atomic[6]. So, in particular, if you show the Galois group of a polynomial is simple then, game-over,

[1] I make up for these with odd footnotes.
[2] for example, this or this. No Rickroll, I promise.
[3] forgive me if I don't reproduce the formula here. See this for example
[4] this is quite a formula, it takes about a page, for example see this
[5] we later define simple and solvable groups, the details are not too important for our current discussion.
[6] more a bit later on how the term atom breaks down: Neutrons, Protons, electrons then on to quarks and such...

no solution[7]. This helps you understand why mathematicians were so happy we finally[8] classified all finite simple groups in 2004[9]. To give a specific example of Galois' Theory's power,

$$3x^5 - 15x + 5 = 0$$

is not solvable by radicals. Gallian gives the group theoretic argument on why that is on page 559 of our text. Interestingly, Galois' contribution was not recognized until several decades after his death. In 1846 Lioville understood the importance of Galois' work and began to promote Galois' group concept. By 1870, Jordan[10] understood Galois' well-enough to write a text on it. That said, I don't have much more to say about Galois theor in this course. It is interesting, powerful, and motivational to the study of group theory. But, our focus is on more elementary material.

Initially, groups were all about permutations, but, as the story continues mathematicians discovered the structure of a group was not unique to permutations. For example, the symmetry groups promoted by Klein and Lie in the late nineteenth century. Thinking of groups abstractly came a bit later. Gallian credits this to Dyck and Weber circa 1883. Dyck, a student of Klein, emphasized the importance of invertibility in a paper about Tesselations[11]. Let pause our historical tour to examine the definition of a group and a few elementary examples.

Definition 1.1.1. *A set G with an operation[12] $\star : G \times G \to G$ forms a* **group** *if*

(i.) Associativity: $(a \star b) \star c = a \star (b \star c)$ *for all* $a, b, c \in G$,

(ii.) Identity: *there exists* $e \in G$ *such that* $a \star e = e \star a = a$ *for each* $a \in G$,

(iii.) Invertibility: *for each* $g \in G$ *there exists* $h \in G$ *such that* $h \star g = g \star h = e$.

If $a \star b = b \star a$ *for all* $a, b \in G$ *then we say G is an* **abelian** *or* **commutative** *group. If there exist* $a, b \in G$ *for which* $a \star b \neq b \star a$ *then G is a* **non-abelian** *group.*

The notation $\star$ is not typically used as we study specific examples. In fact, to denote $a \star b$ we typically use either **juxtaposition** (ab) or in the case of an abelian group we use **additive notation** $(a+b)$. It is customary to only use $+$ for a commutative operation.

Example 1.1.2. *Let $G = \mathbb{Z}$ is a group under addition with identity 0: In particular, we know for $a, b, c \in \mathbb{Z}$ there exists $-a \in \mathbb{Z}$ and $0 \in \mathbb{Z}$ for which:*

$$\textbf{(i.)}\ (a+b)+c = a+(b+c), \quad \textbf{(ii.)}\ a+0 = a = 0+a, \quad \textbf{(iii.)}\ a+(-a) = 0 = (-a)+a.$$

Moreover, we know whenever $a, b \in \mathbb{Z}$ the sum $a + b \in \mathbb{Z}$.

---

[7] ok, to be precise, no closed-form solution in terms of radicals and such, a fifth order polynomial with real coefficients has a zero by the intermediate value theorem. But, the existence of such a zero is not the same as the existence of a nice formula for the zero

[8] in 2004, Aschbacher and Smith published a 1221-page proof for the missing quasithin case

[9] we wont get to that in this course, its about 10,000 pages, including for example the paper of Feit-Thompson which alone is 250 pages, but, I will loosely cover the appropriate section later in Gallian in due time

[10] of the Jordan form, yes, sorry bad memories for my Math 321 class

[11] see Dr. Runion's office for an example

[12] this notation indicates that $\star$ is a **function** from $G \times G$ to G. In other words, $\star$ is a **binary operation**. This is sometimes identified as an axiom of a group known as **closure**.

You might wonder *how* we know such properties hold for $\mathbb{Z}$. To be precise, we could build the integers from scratch using set-theory, but, to properly understand that construction it more or less begs an understanding of this course. Consequently, we will be content[13] to use $\mathbb{Z}, \mathbb{C}, \mathbb{R}$ and $\mathbb{Q}$ as known objects complete with their standard properties. That said, as our understanding of abstract algebra increases we will begin to demonstrate how these standard number systems can be *constructed*.

Example 1.1.3. *Actually, this is a pair of non-examples. First, $\mathbb{Z}$ with subtraction is not a group. Second, $\mathbb{Z}$ with multiplication is not a group.* WHY ?

The next example is a bit *meta*.

Example 1.1.4. *Let V be a vector space then $V, +$ where $+$ denoted vector addition forms a group where the identity element is the zero vector 0. The definition of a vector space includes the assumption $(x+y)+z = x+(y+z)$ for all $x, y, z \in V$ hence Axiom (i.) holds true. Axiom (ii.) is satisfied since $x + 0 = 0 + x = 0$ for each $x \in V$. Finally, Axiom (iii.) for each $x \in V$ there exists $-x \in V$ such that $x + (-x) = 0$. In summary, any vector space is also an abelian group where the operation is understood to be vector addition*[14]

I should pause to note, the examples considered thus far are not the sort of interesting examples which motivated and caused mathematicians to coin the term group. These examples are just easy and make for short discussion. Let me add a few more to our list:

Example 1.1.5. *Let $\mathbb{Q}^\times = \mathbb{Q} - \{0\}$ denote the set of nonzero rational numbers. $\mathbb{Q}^\times$ forms a group with respect to multiplication. The identity element is 1.*

Example 1.1.6. *Let $\mathbb{R}^\times = \mathbb{R} - \{0\}$ denote the set of nonzero real numbers. $\mathbb{R}^\times$ forms a group with respect to multiplication. The identity element is 1.*

Example 1.1.7. *Let $\mathbb{C}^\times = \mathbb{C} - \{0\}$ denote the set of nonzero complex numbers. $\mathbb{C}^\times$ forms a group with respect to multiplication. The identity element is 1.*

Example 1.1.8. *Let $\mathbb{Z}^\times = \mathbb{Z} - \{0\}$ denote the set of nonzero integers. $\mathbb{Z}^\times$ does not form a group since $2x = 1$ has solution $x = 1/2 \notin \mathbb{Z}$.*

Let me give at least one interesting explicit example in this section. This group is closely tied to invertible linear transformations on $\mathbb{R}^n$:

Example 1.1.9. *Let $GL(n, \mathbb{R}) = \{A \in \mathbb{R}^{n \times n} \mid det(A) \neq 0\}$. We call $GL(n, \mathbb{R})$ the* **general linear group** *of $n \times n$ matrices over $\mathbb{R}$. We can verify $GL(n, \mathbb{R})$ paired with matrix multiplication forms a nonabelian group. Notice, matrix multiplication is associative; $(AB)C = A(BC)$ for all $A, B, C \in GL(n, \mathbb{R})$. Also, the identity matrix I defined*[15] *by $I_{ij} = \delta_{ij}$ has $AI = A = IA$ for each $A \in GL(n, \mathbb{R})$. It remains to check closure of multiplication and inversion. Both of these questions are nicely resolved by the theory of determinants: if $A, B \in GL(n, \mathbb{R})$ then*

$$det(AB) = det(A)det(B) \neq 0$$

[13] in an intuitive sense, numbers exist independent of their particular construction, so, not much is lost here. For example, I can construct $\mathbb{C}$ using vectors in the plane, particular 2×2 matrices, or via equivalence classes of polynomials. Any of these three could reasonably be called $\mathbb{C}$

[14] Of course, there is more structure to a vector space, but, I leave that for another time and place.

[15] δ_{ij} is one of my favorite things. This Kronecker delta is zero when $i \neq j$ and is one when $i = j$.

thus $AB \in GL(n,\mathbb{R})$ hence we find matrix multiplication forms a binary operation on $GL(n,\mathbb{R})$. Finally, we know $det(A) \neq 0$ implies there exists A^{-1} for which $AA^{-1} = I = A^{-1}A$ and $det(AA^{-1}) = det(A)det(A^{-1}) = det(I) = 1$ thus $det(A^{-1}) = 1/det(A) \neq 0$. Therefore, we find $A \in GL(n,\mathbb{R})$ implies $A^{-1} \in GL(n,\mathbb{R})$

The previous example is more in line with Klein and Lie's investigations of *transformation groups*. Many of those groups will appear as subgroups[16] of the example above. At this point I owe you a few basic theorems about groups.

Theorem 1.1.10. *In a group G there can be only one identity element.*

Proof: let G be a group with operation $\star$. Suppose e and e' are identity elements in G. We have (i.) $e \star a = a = a \star e$ and (ii.) $e' \star a = a = a \star e'$ for each $a \in G$. Thus, by (i.) with $a = e'$ and (ii.) with $a = e$,

$$e \star e' = e' = e' \star e \quad \& \quad e' \star e = e.$$

We observe $e' \star e = e' = e$. In summary, the identity in a group is unique. □

An examination of the proof above reveals that the axiom of associativity was not required for the uniqueness of the identity. As a point of trivia, a group without the associativity axiom is a *loop*. Here is a table[17] with other popular terms for various weakenings of the group axioms:

Group-like structures

	Totality[α]	Associativity	Identity	Divisibility	Commutativity
Semicategory	Unneeded	Required	Unneeded	Unneeded	Unneeded
Category	Unneeded	Required	Required	Unneeded	Unneeded
Groupoid	Unneeded	Required	Required	Required	Unneeded
Magma	Required	Unneeded	Unneeded	Unneeded	Unneeded
Quasigroup	Required	Unneeded	Unneeded	Required	Unneeded
Loop	Required	Unneeded	Required	Required	Unneeded
Semigroup	Required	Required	Unneeded	Unneeded	Unneeded
Monoid	Required	Required	Required	Unneeded	Unneeded
Group	Required	Required	Required	Required	Unneeded
Abelian Group	Required	Required	Required	Required	Required

^α Closure, which is used in many sources, is an equivalent axiom to totality, though defined differently.

Relax, I only expect you to know the definition of group for the time being[18].

Theorem 1.1.11. Cancellation Laws: *In a group G right and left cancellation laws hold. In particular, $ba = ca$ implies $b = c$ and $ab = ac$ implies $b = c$.*

Proof: let G be a group with operation denoted by juxtaposition. Suppose $a, b, c \in G$ and $ba = ca$. Since G is a group, there exists $a^{-1} \in G$ for which $aa^{-1} = e$ where e is the identity. Multiply $ba = ca$ by a^{-1} to obtain $baa^{-1} = caa^{-1}$ hence $be = ce$ and we conclude $b = c$. Likewise, if $ab = ac$

[16] ask yourself about this next lecture

[17] I borrowed this from the fun article on groups at Wikipedia

[18] As my adviser would say, I include the table above for the most elusive creature, the interested reader

then $a^{-1}ab = a^{-1}ac$ hence $eb = ec$ and we find $b = c$. □

Cancellation is nice. Perhaps this is also a nice way to see certain operations cannot be group multiplications. For example, the cross product in $\mathbb{R}^3$ does not support the cancellation property. For those who have taken multivariate calculus, quick question, which group axioms fail for the cross product?

Theorem 1.1.12. *Let G be a group with identity e. For each $g \in G$ there exists a unique element h for which $gh = e = hg$.*

Proof: let G be a group with identity e. Suppose $g \in G$ and $h, h' \in G$ such that

$$gh = e = hg \qquad \& \qquad gh' = e = h'g$$

In particular, we have $gh = gh'$ thus $h = h'$ by the cancellation law. □

At this point, I return to our historical overview of abstract algebra[19] Returning to Lagrange and Euler once more, they also played some with algebraic integers which were things like $a + b\sqrt{n}$ in order to attack certain questions in number theory. Gauss instead used modular arithmetic in his master work *Disquisitiones Arithmeticae* (1801) to attack many of the same questions. Gauss also used numbers[20] of the form $a + b\sqrt{-1}$ to study the structure of primes. Gauss' mistrust of Lagrange's algebraic numbers was not without merit, it was known that unique factorization broke down in some cases, and this gave cause for concern since many arguments are based on factorizations into primes. For example, in $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$ we have:

$$(2)(3) = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

It follows the usual arguments based on comparing prime factors break down. Thus, much as with Abel and Ruffini and the quintic, we knew something was up. Kummer repaired the troubling ambiguity above by introducing so-called *ideal numbers*. These ideal numbers were properly constructed by Dedekind who among other things was one of the first mathematicians to explicitly use congruence classes. For example, it was Dedekind who constructed the real numbers using so-called Dedekind-cuts in 1858. In any event, the ideals of Kummer and Dedekind and the modular arithmetic of Gauss all falls under the general concept of a ring. What is a ring?

Definition 1.1.13. *A set R with addition $+ : R \times R \to R$ and multiplication $\star : R \times R \to R$ is called a* **ring** *if*

(i.) *$(R, +)$ forms an abelian group*

(ii.) *$(a + b) \star c = a \star c + b \star c$ and $a \star (b + c) = a \star b + a \star c$ for all $a, b, c \in R$.*

If there exists $1 \in R$ such that $a \star 1 = a = 1 \star a$ for each $a \in R$ then R is called a **ring with unity**. *If $a \star b = b \star a$ for all $a, b \in R$ then R is a* **commutative ring**.

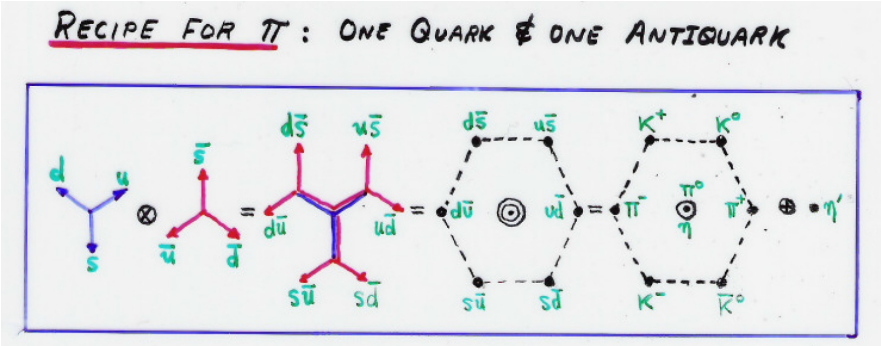
Rings are everywhere, so many mathematical objects have both some concept of addition and multiplication which gives a ring structure. Rings were studied from an abstract vantage point by Emmy Noether in the 1920's. Jacobson, Artin, McCoy, many others, all added depth and application of ring theory in the early twentieth century. If $ab = 0$ and neither a nor b is zero then a

[19] I have betrayed Cayley in this story, but, have no fear well get back to him and many others soon enough

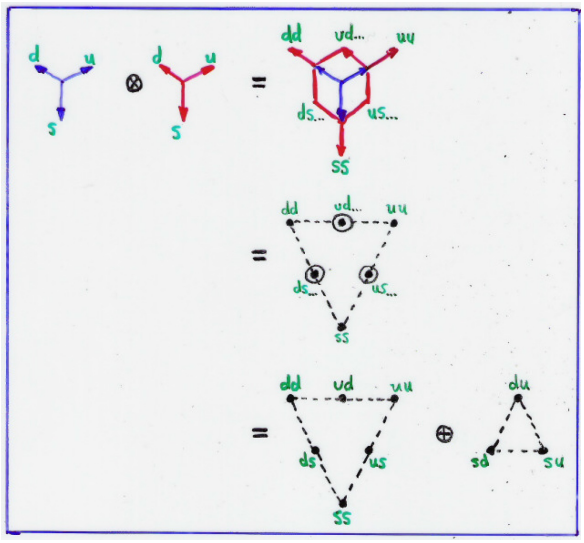
[20] if $a, b \in \mathbb{Z}$ then $a + bi$ is known as a Gaussian integer

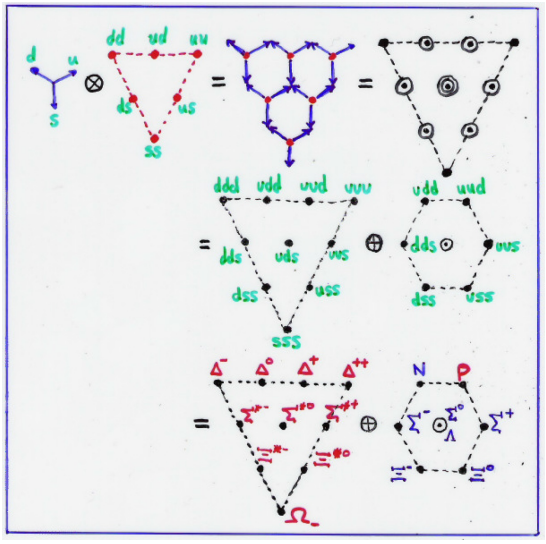
and b are nontrivial zero-divisors. If $ab = c$ then we say that b divides c. Notice, zero always is a divisor of zero. Anyway, trivial comments aside, if a ring has no zero divisors then we say the ring is an **integral domain**. Ok at this point, it becomes fashionable (unless youre McCoy) to assume R is commutative. A good example of an integral domain is the integers. Next, if a has b for which $ab = 1$ then we say a is a **unit**. If every nonzero element of a ring is a unit then we call such a ring a **field**. Our goal this semester is to understand the rudiments of groups, rings and fields. Well focus on group structure for a while, but, truth be told, some of our examples have more structure. We return to the formal study of rings after Test 2. Finally, if you stick with me until the end, Ill explain what an *algebra* is at the end of this course.

Since we have a minute, let me show you a recent application of group *representation* theory to elementary particle physics. First, the picture below illustrates how a quark and an antiquark combine to make Pions, and Kaons:



These were all the rage in early-to-mid-twentieth century nuclear physics. But, perhaps the next pair of examples will bring us to something you have heard of previously. Let's look at how quarks can build Protons and Neutrons:





Let me briefly explain the patterns. These are drawn in the isospin-hypercharge plane. They show how the isospin and hypercharge of individual up, down or strange quarks combine together to make a variety of hadronic particles. The **N** and **P** stand for **Neutron** and **Proton**. These patterns were discovered before quarks. Then, the mathematics of *group representations* suggested

the existence of quarks. The $\otimes$ is a *tensor product*. These pictures are taken from a talk I gave in graduate school in Dr. Misra's *Representation Theory* course. Incidentally, on the suspicion the pattern continued, Gell-Mann predicted the Ω^- particle existed in 1962. It was experimentally verified in 1964. Murray Gell-Mann won the Nobel Prize in Physics for this work on what he called the *eight-fold way*. Gell-Mann and Zweig (independently) proposed the quark model in 1964. It took about three decades for physicsists to experimentally confirm the existence of the quarks[21].

I recommend working on Chapter 2 Gallian problems such as:

$$\#7, 11, 14, 15, 16, 17, 18, 19, 20, 26, 28, 30, 31, 34,$$

these should help bring the Group concept to life. I do not collect these, but, I will keep them in mind as I construct tests.

Problems for Lecture 1: (these are collected at the beginning of Lecture 3)

Problem 1: Determine which of the following sets with operations are groups. If it is a group, state its identity, what the inverse of a typical element looks like, and determine if it is Abelian. If it is not a group, state which axioms hold and give counter-examples for those which fail (don't forget closure).

(a) $(\mathbb{Z}_{\geq 0}, +)$ non-negative integers with addition

(b) $(3\mathbb{Z}, +)$ multiples of 3 (i.e. $0, \pm 3, \pm 6, \dots$) with addition

(c) $(\mathbb{R}_{<0}, \cdot)$ negative reals with multiplication

(d) $(\mathbb{R}_{\neq 0}, \div)$ non-zero reals with division

(e) $(\mathbb{Q}_{>0}, \cdot)$ positive rationals with multiplication

Problem 2: explain why $\mathbb{Z}$ does not form a group with respect to subtraction. Also, explain why $\mathbb{Z}$ does not form a group with respect to multiplication.

Problem 3: I claimed $\mathrm{GL}(n, \mathbb{R})$ was nonabelian. Prove my claim in the case $n = 2$.

Problem 4: solve number 37 from page 56 of Gallian.

[21] I'll let our physics department explain the details of those experiments for you...

1.2 Lecture 2: on orders and subgroups

In this section we introduce more groups and we initiate our study of subgroups and orders. Next, for our final new example before I get to the meat and potatoes of Gallian Chapter 3, I take a logical loan from our future: in particular, the assumption that $\mathbb{Z}_n$ forms a commutative ring with respect to the operations of addition and multiplication given below:

Definition 1.2.1. modular arithmetic *Suppose $n \in \mathbb{N}$ where $n \geq 2$. Let $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ and define* **addition** *and* **multiplication** *on $\mathbb{Z}_n$ as follows: if $a, b \in \mathbb{Z}_n$ then $a + b = c$ and $a \times b = ab = d$ where $c, d \in \mathbb{Z}_n$ and $a + b - c, ab - d \in n\mathbb{Z}$ where $n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$.*

In Lectures 4 and 5 we'll study this structure carefully. Essentially, the idea is just to add or multiply in $\mathbb{Z}$ then remove multiples of n until we have some integer in $\{0, 1, \ldots, n-1\}$. This means multiples of n serve as zero in $\mathbb{Z}_n$.

Example 1.2.2. *The addition and multiplication tables for $\mathbb{Z}_3$:*

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

&

$\times$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Notice, $(\mathbb{Z}_3, +)$ forms an abelian group whereas $(\mathbb{Z}_3, \times)$ is not a group because 0 is not invertible. However, $G = \{1, 2\}$ with multiplication modulo 3 does form a group with two elements. Notice, $2 \times 2 = 1$ hence $2^{-1} = 2$.

The multiplicative group we exhibited in the example above is part of a much larger story. In particular, it can be shown that if the $\gcd(n, k) = 1$ then there exists $j \in \mathbb{Z}_n$ for which $jk = 1$ which is to say $k^{-1} = j \in \mathbb{Z}_n$. Again, we'll prove such assertions in Lectures 4 and 5 for the sake of the students who do not already know these fun facts from previous course work. That said, for the sake of conversation, let us define:

Definition 1.2.3. *The* **group of units** *in $\mathbb{Z}_n$ is the set of all $x \in \mathbb{Z}_n$ such that $gcd(x, n) = 1$. We denote this group by $U(n)$.*

For example, the G of Example 1.2.2 is $U(3)$.

Example 1.2.4. *In $\mathbb{Z}_{10}$ we have $1, 3, 7, 9$ relatively prime to 10. Hence, $U(10) = \{1, 3, 7, 9\}$. The multiplication (aka Cayley) table for $U(10)$ is:*

$\times$	1	3	7	9
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

Example 1.2.5. *In $\mathbb{Z}_{11}$ since 11 is prime, all smaller integers are relatively prime to 11:*

$$U(11) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Notice, as a sample calculation, $9 \times 5 = 1$ modulo 11 hence $9^{-1} = 5$ in $U(11)$.

In fact, whenever p is prime, we should notice $\mathbb{Z}_p$ with addition and multiplication modulo p provides the structure of a **field** because addition and multiplication behave as they ought and every nonzero element has a multiplicative inverse[22]. We say $\mathbb{Z}_p$ is a *finite field* in constrast to the infinite fields $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ etc. I should mention, an example we studied last lecture generalizes to any field:

Definition 1.2.6. *The* **general linear group** *of $n \times n$ matrices over a field $\mathbb{F}$ is the set of all invertible matrices in $\mathbb{F}^{n\times n}$ which is denoted $GL(n, \mathbb{F})$.*

The group operation is given by the natural matrix multiplication of $n \times n$ matrices with entries in $\mathbb{F}$. With appropriate modifications the usual linear algebra for invertible matrices equally well applies in this context. In particular, in linear algebra we found a formula for the inverse in terms of the classical adjoint. This formula continues to make sense over $\mathbb{F}$ with the understanding that division is replaced by multiplication by multiplicative inverse. Perhaps you will have a homework on this after we discuss modular arithmetic a bit more.

Definition 1.2.7. *The number of elements of a group (finite or infinite) is called the* **order**. *The order of G is denoted $|G|$.*

Notice this allows $|G| \in \mathbb{N}$ or $|G| = \infty$. Notice $|G| \geq 1$ since every group has at least the identity element.

Example 1.2.8. *If $G = \{e\}$ where $e \star e = e$ then G forms a group where $e^{-1} = e$ and $|G| = 1$.*

Example 1.2.9. *The order of $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ is simply n.*

To discuss the order of an element we should define some notation[23]. The additive notation ng is a notation for successive group additions:

$$\underbrace{g + g + \cdots + g}_{n-summands} = ng$$

to be precise, $g + g = 2g$ and for $n \in \mathbb{N}$ we *recursively* define $(n+1)g = ng + g$. Likewise, the multiplicative notation

$$\underbrace{gg \cdots g}_{n-factors} = g^n$$

is defined recursively via $g^{n+1} = g^n g$ for each $n \in \mathbb{N}$.

Definition 1.2.10. *The* **order** *of an element g in a group G is the smallest $n \in \mathbb{N}$ such that $g^n = e$. If no such $n \in \mathbb{N}$ exists then g has* **infinite order**. *We denote the order of an element g by $|g|$. In additive notation, with additive identity $0 \in G$, if $|g| = n$ then $ng = 0$.*

Example 1.2.11. *In the context of $\mathbb{Z}_4$ as an additive group we have*

$$1 + 1 + 1 + 1 = 0, \quad 2 + 2 = 0, \quad 3 + 3 + 3 + 3 = 0$$

thus $|1| = |3| = 4$ whereas $|2| = 2$.

[22] that is, every nonzero element is a unit

[23] my apologies if I used this previously without explaination, we should be careful to not assume usual math automatically applies to the abstract group context. Since group operations are abstract, we must define things like powers in terms of the group operation alone.

Example 1.2.12. *Consider $G = \mathbb{R}^\times$. We note $a \in G$ with $a \neq 1$ has $a^n \neq 1$ for all $n \in \mathbb{N}$. Thus $|a| = \infty$. Every element of G, except 1, has infinite order. Moreover, $|G| = \infty$. We can make the same observation about $\mathbb{C}^\times$ or $\mathbb{Q}^\times$.*

Definition 1.2.13. Subgroup *If H is a subset of a group G and H is itself a group with respect to the operations of G then we say H is a subgroup of G. We denote[24] $H \leq G$.*

A silly example, $G \subseteq G$ hence $G \leq G$. If we wish to indicate H is a subgroup of G and $H \neq G$ then we write $H < G$. If $H < G$ then we say H is a proper subgroup of G. A second silly example, if e is the identity in G, then $\{e\} = H$ forms the **trivial subgroup**. If $H \leq G$ and $H \neq \{e\}$ then H is a **nontrivial subgroup** of G. Notice, I wrote $\{e\}$ not e because one is a set and the other is not. I expect you to do likewise.

Example 1.2.14. *Is $\mathbb{Z}_n$ a subgroup of $\mathbb{Z}$ under addition ?* WHY NOT ?

Example 1.2.15. *In $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ we have $H = \{0, 2\}$ with $2 + 2 = 0$ hence $H \leq \mathbb{Z}_4$ as it clearly forms a group.*

Clearly is always suspicious, but, I really say it because of the wonderful theorems which follow next:

Theorem 1.2.16. one-step subgroup test[25]: *Let G be a group. If $H \subseteq G$ and $H \neq \emptyset$ then $H \leq G$ if $ab^{-1} \in H$ whenever $a, b \in H$. Equivalently, in additive notation, $H \leq G$ if $a - b \in H$ whenever $a, b \in H$.*

Proof: I will give the proof in multiplicative notation. Suppose H a nonempty subset of G with the property that $ab^{-1} \in H$ whenever $a, b \in H$. To show $H \leq G$ we must prove H satisfies the axioms of a group where the operation is the multiplication of G suitably restricted to H.

Identity: Notice, since $H \neq \emptyset$ there exists $a \in H$ hence $aa^{-1} = e \in H$. Observe $he = h = eh$ for each $h \in H$ by the given group structure of G and the fact $h \in H$ implies $h \in G$ since $H \subseteq G$. Thus e is the identity in H.

Invertibility: let $a \in H$ and note $aa^{-1} = e \in H$ thus $a^{-1} \in H$. It follows every element of H has an inverse in H.

Closure: suppose $a, b \in H$ and note by invertiblility $b^{-1} \in H$. Moreover, we can prove, $(b^{-1})^{-1} = b$. Thus $ab = a(b^{-1})^{-1} \in H$ and we have shown the operation on G restricts to a binary operation on H as desired.

Associativity: of multiplication in H is easy, if $a, b, c \in H$ then $a, b, c \in G$ thus $a(bc) = (ab)c$. □

Theorem 1.2.17. two-step subgroup test: *Let G be a group. If $H \subseteq G$ and $H \neq \emptyset$ then $H \leq G$ if H is closed under multiplication and inversion. That is, $H \leq G$ if (1.) $ab \in H$ whenever $a, b \in H$ and (2.) $a^{-1} \in H$ whenever $a \in H$.*

Proof: suppose H is a nonempty subset of a group G with properties (1.) and (2.) as described in the theorem. Suppose $a, b \in H$ then by (2.) we have $b^{-1} \in H$. Thus, as $a, b^{-1} \in H$ we have $ab^{-1} \in H$ using (1.). Therefore, Theorem 1.2.16 applies and we conclude $H \leq G$. □

It is important to prove $H \neq \emptyset$ as we use the subgroup theorems to analyze potential subgroups.

[24] this is read " H is a subgroup of G "

[25] depending on how you parse things, you might see more steps here, see page 60 of Gallian for commentary

Example 1.2.18. *In problem 37 of page 56, you are asked to show* $G = \{\begin{bmatrix} a & a \\ a & a \end{bmatrix} | a \in \mathbb{R}, a \neq 0\}$ *forms a group with respect to matrix multiplication. Is this a subgroup of* $GL(2,\mathbb{R})$ *?*

Example 1.2.19. *Let* $G = GL(n,\mathbb{R})$ *and define* $H = \{A \in \mathbb{R}^{n\times n} \mid det(A) = 1\}$*. Notice,* $det(I) = 1$ *thus* $I \in H \neq \emptyset$*. If* $A, B \in H$ *then notice* $det(AB) = det(A)det(B) = 1(1) = 1$ *thus* $AB \in H$*. Also if* $A \in H$ *then* $det(A) = 1$ *thus* $A^{-1} \in G$ *exists with* $AA^{-1} = I$*. Note, as* $det(I) = 1$*,*

$$1 = det(AA^{-1}) = det(A)det(A^{-1})$$

We find $det(A^{-1}) = 1$ *and conclude* $A^{-1} \in H$*. We conclude by the two-step subgroup test* $H \leq G$*.*

The example above shows the following name-calling is warranted. Also, while I stated the example for $\mathbb{R}$ we can just as well use any field $\mathbb{F}$.

Definition 1.2.20. Special Linear Group: *of* $n \times n$ *matrices over the field* $\mathbb{F}$ *is given by*

$$SL(n,\mathbb{F}) = \{A \in \mathbb{F}^{n\times n} \mid det(A) = 1\}.$$

It is interesting to note that in the case $|G| < \infty$ it suffices to check that a nonempty subset H is closed under the operation of G. See Theorem 3.3 in Gallian. The remainder of this section is devoted to special subgroups which we can construct for any given group. We begin with the subgroup generated by a particular element. First some notation[26]

Definition 1.2.21. cyclic subgroup generated by an element: *Let* G *be a multiplicative group and* $g \in G$ *then* $\langle g\rangle = \{g^n \mid n \in \mathbb{Z}\}$*. If* G *is an additive group and* $g \in G$ *then* $\langle g\rangle = \{ng \mid n \in \mathbb{Z}\}$

Of course, the definition above would be very out of place if we didn't have the following theorem:

Theorem 1.2.22. *Let* $a \in G$ *where* G *is a group. Then* $\langle a\rangle \leq G$*. In other words, the cyclic subgroup generated by* a *is indeed a subgroup.*

Proof: let $a \in G$ with multiplicative notation. We define $a^{-n} = (a^{-1})^n$ for $n \in \mathbb{N}$ and $a^0 = e$. Observe $a \in \langle a\rangle \neq \emptyset$. Suppose $a^j, a^k \in \langle a\rangle$ where $j, k \in \mathbb{Z}$. Notice[27] $a^j a^k = a^{j+k}$ and as $j + k \in \mathbb{Z}$ we find $a^j a^k \in \langle a\rangle$. Moreover, $a^k \in \langle a\rangle$ with $k \in \mathbb{Z}$ has $-k \in \mathbb{Z}$ and thus $a^{-k} \in \langle a\rangle$ where $a^k a^{-k} = a^0 = e$. Thus $\langle a\rangle$ is closed under multiplication and inversion in G and we conclude by the two-step subgroup test that $\langle a\rangle \leq G$. □

Of course, we could prove the theorem above with the one-step test if we prefer that course of action. See Gallians slick argument at top of page 63.

Definition 1.2.23. Center of Group: *let* G *be a group then we define the* **center** *of* G *by* $Z(G) = \{a \in G \mid ax = xa \text{ for each } x \in G\}$*.*

In an additive group notation, $Z(G) = \{a \in G \mid a + x = x + a \text{ for each } x \in G\}$. For example, $Z(\mathbb{Z}_n) = Z_n$ since addition in $\mathbb{Z}_n$ commutes. The center is more interesting for nonabelian groups. This much we can say with relatively little effort:

Theorem 1.2.24. *If* G *is a group then* $Z(G) \leq G$*.*

[26] this is important!

[27] This is a pretty big leap! As it happens, this takes some work to prove... see solution to Lecture 2 problems.

Proof: see Gallian page 64. □

We can also study a similar object for a particular element in the group:

Definition 1.2.25. Centralizer of a group element: *let g be a fixed element in a group G. We define $C(x) = \{g \in G \mid gx = xg\}$ to be the* **centralizer** *of x in G.*

A silly example, in an abelian group G we have $C(x) = G$ for any $x \in G$ since x commutes with all elements of G.

Theorem 1.2.26. *For each $x \in G$ the centralizer of x is a subgroup of G; $C(x) \leq G$.*

Proof: Suppose $x \in G$. Note $xe = ex$ thus $e \in C(x) \neq \emptyset$. Let $a, b \in C(x)$. We are given $ax = xa$ and $bx = xb$. Notice, $bx = xb$ implies $b^{-1}bxb^{-1} = b^{-1}xbb^{-1}$ thus $xb^{-1} = b^{-1}x$. Therefore,

$$(ab^{-1})x = a(b^{-1}x) = a(xb^{-1}) = (ax)b^{-1} = (xa)b^{-1} = x(ab^{-1})$$

hence $ab^{-1} \in C(x)$. We conclude $C(x) \leq G$ by the one-step subgroup test. □

I recommend working on Chapter 2 Gallian problems such as:

$$\#12, 13, 21, 23, 29$$

also, Chapter 3 Gallian problems such as:

$$\#1, 6, 8, 9, 10, 14, 15, 16, 20, 21, 24, 28, 34, 35, 38, 39, 40, 41, 42, 44, 45, 46, 54$$

these should help bring the Group concept to life. I do not collect these, but, I will keep them in mind as I construct tests.

Problems for Lecture 2: (these are collected at the beginning of Lecture 4)

Problem 5: solve number 24 from page 55 of Gallian (construct Cayley table for $U(12)$.)

Problem 6: solve number 4 from page 67 of Gallian. ($|g| = |g^{-1}|$)

Problem 7: solve number 45 from page 70 of Gallian. (subgroup problem)

Problem 8: solve number 51 from page 71 of Gallian. (centralizer and center $\mathrm{GL}(2, \mathbb{R})$)

1.3 Lecture 3: on the dihedral group and symmetries

In this section we discuss groups which are tied to the concept of distance in euclidean space. I'll focus on the context of $\mathbb{R}^n$, and I will state without proof some of the deeper theorems in this section to euclidean n-space $\mathbb{R}^n$. Before I get into that, I should make a general comment. There is a group we can construct for any set S.

Definition 1.3.1. *Let S be a nonempty set a bijections on S is called a* **permutation** *of S.*

If G is the set of permutations on a nonempty set S then it is not hard to show that G forms a group with respect to function composition. For example, the mapping $Id_S : S \to S$ defined by $Id_S(x) = x$ for each $x \in S$ serves as the identity of G. I'll let you complete the proof in your homework. Ultimately in this section we explain that the set of distance preserving functions on $\mathbb{R}^n$ form a subgroup of the permutations on $\mathbb{R}^n$.

We denote **euclidean norm** or **vector length** by $||(x_1, \ldots, x_n)|| = \sqrt{x_1^2 + \cdots + x_n^2}$. This norm satisfies the needed axioms for a norm:

$$\underbrace{||x+y|| \leq ||x|| + ||y||}_{\text{triangle inequality}}, \quad \underbrace{||cx|| = |c|\,||x||}_{\text{absolute homogeneity}} \ , \quad \underbrace{||x|| = 0 \ \text{ only if } x = 0, \quad ||x|| \geq 0}_{\text{positive definite}}.$$

The **distance** between $P, Q \in \mathbb{R}^n$ is naturally given by the length of the displacement vector from P to Q; $d(P,Q) = ||Q - P|| = ||P - Q||$. Let us define:

Definition 1.3.2. *If $P, Q \in \mathbb{R}^n$ then we define* **distance** *between P, Q as $d(P,Q) = ||P - Q||$.*

Of course, we can also express $||v|| = \sqrt{v \bullet v}$ thus $d(P,Q) = \sqrt{(P-Q) \bullet (P-Q)}$. This is nice since we already know many nice properties of the dot-product from our work in linear algebra. The term *isometry* means *same measure*. In particular, an isometry of a space is a mapping on the space which preserves the distance between points.

Definition 1.3.3. Isometry: *if $\phi : \mathbb{R}^n \to \mathbb{R}^n$ has*

$$||\phi(P) - \phi(Q)|| = ||P - Q||$$

for each $P, Q \in \mathbb{R}^n$ then ϕ is an **isometry**.

Notice the definition above allows ϕ to be any function in principle. However, after some study, we'll find isometries are quite rigid in their construction. I follow §2.3 of Rotman's *A First Course in Abstract Algebra*, although, he focuses entirely on $n = 2$.

Theorem 1.3.4. *Let ϕ be an isometry of $\mathbb{R}^n$. Then ϕ preserves dot-products*[28] *iff $\phi(0) = 0$.*

Proof: Suppose $\phi(P) \bullet \phi(Q) = P \bullet Q$ for all $P, Q \in \mathbb{R}^n$. Thus $\phi(P) \bullet \phi(P) = P \bullet P = ||P||^2$. Hence, $\phi(0) \bullet \phi(0) = ||0||^2 = 0$ and it follows $\phi(0) = 0$.

To prove the converse direction we begin by assuming ϕ is an isometry for which $\phi(0) = 0$. As ϕ is an isometry we have $d(P,0) = d(\phi(P), \phi(0))$ hence $||P|| = ||\phi(P) - \phi(0)|| = ||\phi(P)||$. Consider

[28] this means $\phi(P) \bullet \phi(Q) = P \bullet Q$ for all $P, Q \in \mathbb{R}^n$

then, by the usual algebra of dot-products,

$$\begin{aligned} ||\phi(P)||^2 + ||\phi(Q)||^2 - 2\phi(P)\bullet\phi(Q) &= [\phi(P)-\phi(Q)]\bullet[\phi(P)-\phi(Q)] \\ &= ||\phi(P)-\phi(Q)||^2 \\ &= ||P-Q||^2 \\ &= (P-Q)\bullet(P-Q) \\ &= ||P||^2 + ||Q||^2 - 2P\bullet Q. \end{aligned}$$

Thus $\phi(P)\bullet\phi(Q) = P\bullet Q$ for all $P, Q \in \mathbb{R}^n$ and the theorem follows. □

Since the Cauchy Schwarz inequality says $|v\bullet w| \leq ||v||\,||w||$ it is reasonable to define the angle between nonzero vectors. Notice from the Cauchy Schwarz inequality we have, for $v, w \neq 0$,

$$\left| \frac{v\bullet w}{||v||\,||w||} \right| < 1$$

hence **define** $\theta \in [0, \pi]$ to be the value for which

$$\frac{v\bullet w}{||v||\,||w||} = \cos\theta.$$

In this way we provide a definition for angle between vectors in n-dimensions.

Theorem 1.3.5. *If ϕ is an isometry of $\mathbb{R}^n$ and θ is the angle between $v, w \in \mathbb{R}^n$ then θ is also the angle between $\phi(v), \phi(w)$. In other words, ϕ preserves angles.*

Proof: simply apply Theorem 1.3.4. In particular, if θ' is the angle between $\phi(v), \phi(w)$ for $v, w \neq 0$ then, by definition,

$$\cos\theta' = \frac{\phi(v)\bullet\phi(w)}{||\phi(v)||\,||\phi(w)||} = \frac{v\bullet w}{||v||\,||w||} = \cos\theta$$

thus $\theta' = \theta$ as both $\theta, \theta' \in [0, \pi]$ by definition of angle between vectors. □

In summary, isometries of euclidean space preserve both the euclidean distance between points and the usual angle between line-segments. It is good to have a notation for the set of all isometries, and, also those special isometries which fix the origin:

Definition 1.3.6. *The set of all isometries is denoted* $\mathbf{Isom}(\mathbb{R}^n) = \{\phi \mid \phi \text{ an isometry of } \mathbb{R}^n\}$. *We also denote the origin-fixing isometries of $\mathbb{R}^n$ by* $\mathbf{Orth}(n, \mathbb{R}) = \{\phi \in \mathbf{Isom}(\mathbb{R}^n) \mid \phi(0) = 0\}$.

Distinct points $P, Q \in \mathbb{R}^n$ determine a **line** $L[P, Q]$ which we define via

$$L[P, Q] = \{tP + (1-t)(Q-P) \mid t \in \mathbb{R}\}$$

in contrast, the **line-segment** PQ is given by simply restricting t to the unit-interval $[0, 1]$,

$$PQ = \{tP + (1-t)(Q-P) \mid t \in [0, 1]\}.$$

We need these terminologies to discuss some of the isometries below. There are three types we wish to discuss: (feel free to visualize in $\mathbb{R}^2$ for geometric clarity, but, these are also defined in $\mathbb{R}^n$ with a bit more imagination)

1. **Rotations:** about the origin, $R_\theta(0) = 0$ whereas $R_\theta(P) = P'$ where P' is rotated angle θ in CCW direction from P. Observe this is distance preserving hence rotations are isometries.

2. **Reflections:** about a line L is denoted ρ_L. If $x \in L$ then $\rho_L(x) = x$. Otherwise, if $y \neq L$ then $\rho_L(y) = y'$ where y' is on the perpendicular bisector of L through y and y' is the same distance from L as y. Once again, geometrically, it is clear these are distance preserving hence reflections are isometries.

3. **Translations:** given a point Q, a **translation** by Q is the function $\tau_Q : \mathbb{R}^n \to \mathbb{R}^n$ by $\tau_Q(x) = x + Q$ for each $x \in \mathbb{R}^n$. I leave it to the reader to prove translations are isometries.

There are also *glide reflections* and you can read more about those in Chapter 28 of Gallian, however, don't do that right now, it's not the right time[29]. Independent of whatever ambiguities exist in my brief descriptions of rotations, reflections and translations above, it can be shown from properties of euclidean geometry (dot-products, linear algebra, collinearity arguments,... see page 140-141 of Rotman's *A First Course in Abstract Algebra* or see my video where I argue that every isometry of euclidean n-space is the composition of a linear map and a translation in this video which is based on Barret Oneill's Elementary Differential Geometry[30] Ok, fond memories aside:

Theorem 1.3.7. *Every isometry of $\mathbb{R}^n$ is a bijection. Moreover, every isometry fixing 0 is a nonsingular linear transformation.*

With the result above given, it's not too hard to prove the following:

Theorem 1.3.8. $\mathbf{Isom}(\mathbb{R}^n) = \{\phi \mid \phi \text{ an isometry of } \mathbb{R}^n\}$ *forms a group with respect to function composition. Moreover,* $\mathbf{Orth}(n, \mathbb{R}) \leq \mathbf{Isom}(\mathbb{R}^n)$.

Proof: left to reader. □

Since orthogonal transformations fix the origin we know they are nonsingular linear transformations by Theorem 1.3.7. It follows that we can write $T \in \mathbf{Orth}(n, \mathbb{R})$ via multiplication by its **standard matrix**; that is $T(x) = Rx$ for some $R \in \mathbb{R}^{n \times n}$. Such matrices are naturally called **orthogonal matrices**.

Definition 1.3.9. *The set of all standard matrices of orthogonal transformations on $\mathbb{R}^n$ is denoted $O(n, \mathbb{R})$. That is, $O(n, \mathbb{R}) = \{[T] \mid T \in \mathbf{Orth}(n, \mathbb{R})\}$.*

You should show that $O(n, \mathbb{R}) = \{R \in \mathbb{R}^{n \times n} \mid R^T R = I\}$ where R^T denotes the **transpose** of the matrix R. Just as the set of orthogonal transformations forms a subgroup of the set of all bijections on $\mathbb{R}^n$ we will see that $O(n, \mathbb{R}) \leq GL(n, \mathbb{R})$. We pick up this discussion again in the problems at the conclusion of this section. For now, we turn to the discussion of the dihedral group and its origin.

We turn our focus to $n = 2$. The isometries of the plane are particularly nice to visualize and study. In particular, you can envision what happens to shapes as they are transported by an isometry. A circle maps to a circle. A line maps to a line. Things made by gluing lines together at regular angles are sent to likewise constructed objects. In short, isometries preserve the shape of objects in the plane. With this in mind, it is interesting to study those isometries which leave a particular shape invariant. These are examples of **symmetry**. To be precise,

[29] also, for future edification past my course, you ought to watch the four lectures given by Professor Gross of Harvard on the structure of isometries and discrete symmetries of the plane. I have not included all of his wonderful arguments here, he uses group actions which we have yet to discuss. See this lecture approximately

[30] The proof I skip here is not abstract algebra, it is geometry, beautiful analytic geometry mixed with linear algebra.

Definition 1.3.10. *Let $\Omega \subseteq \mathbb{R}^2$. The symmetry group of Ω is defined via:*

$$\Sigma(\Omega) = \{\phi \in \mathbf{Isom}(\mathbb{R}^2) \mid \phi(\Omega) = \Omega\}.$$

To show the symmetry group of Ω is indeed a group we can easily verify the identity map is a symmetry of Ω. Can we prove the rest? Is it clear to you that the inverse of a symmetry is a symmetry and is the product of any two symmetries once more a symmetry of Ω. What else should we check to be careful?

In any event, it is true the symmetry group of a figure in the plane is a subgroup of the isometries of the plane. In particular, the symmetry groups of regular polygons are known as the **dihedral groups** [31] I hope this larger discussion has put Chapter 1 of Gallian in a bit more context.

Definition 1.3.11. *Let D_n denote the symmetry group of a regular, unit-side-length, n-polygon. We call this the* **dihedral group of order 2n**.

It can be shown that all elements in D_n are produced by a rotation and a reflection. In particular, Problem 32 on page 56 of Gallian is very helpful towards calculation in dihedral groups. The key is:

$$fRf = R^{-1}.$$

I intend to show how we argue that in this lecture, but, I'll abstain here as I have no pretty pictures to include. In short, for D_n the rotation which is fundamental is the CCW rotation R by $2\pi/n$ radians; it is geometrically clear that $R^n = e$. Furthermore, if f is a reflection of the n-gon then we can list the elements of D_n as:

$$\{e, R, R^2, \ldots, R^{n-1}, f, fR, R^2f, \ldots, R^{n-1}f\}$$

you can easily count there are $n + n = 2n$ elements above and it follows that $|D_n| = 2n$. This is not a proof! Of course, we can exhibit this in $n = 3$ or $n = 4$ without much trouble.

Example 1.3.12. *Let R be a rotation about the origin by 120^o or $2\pi/3$ radians if you must. Let f be the reflection about the vertical axis of the equilateral triangle where it has one side horizontal. We can verify the symmetry group of the triangle is precisely:*

$$\{e, R, R^2, f, Rf, R^2f\}$$

Furthermore, we can either fill out a Cayley Table for D_3 via geometry or we can use $fRf = R^{-1}$ to algebraically subdue the task. For example,

$$(R^2f)(Rf) = R^2(fRf) = R^2R^{-1} = R.$$

whereas,

$$(Rf)(R^2f) = (RfR)(Rf) = R^{-1}(Rf) = f.$$

There you have it, D_3 is nonabelian. We should try to fill out the Cayley table for D_3 some time.

It is more subtle to actually prove that every element of D_n has the form I exhibit above. You by now should not be surprised that I tell you to see page 144-145 of Rotman's *A First Course in*

[31] The term Dihedral is due to Klein who initiated a larger study of symmetries known as the **Erlangen Program**. See this Wikipedia article for a sense of the scope and importance of Klein's program as it continues to this day

Abstract Algebra for the gory and somewhat subtle details.

I recommend working on Chapter 1 Gallian problems such as:

$$\#1, 2, 3, 12$$

I recommend working on Chapter 2 Gallian problems such as:

$$\#32$$

also, Chapter 3 Gallian problems such as:

$$\#7.$$

these should help bring the Group concept to life. I do not collect these, but, I will keep them in mind as I construct tests.

Problems for Lecture 3: (these are collected at the beginning of Lecture 5)

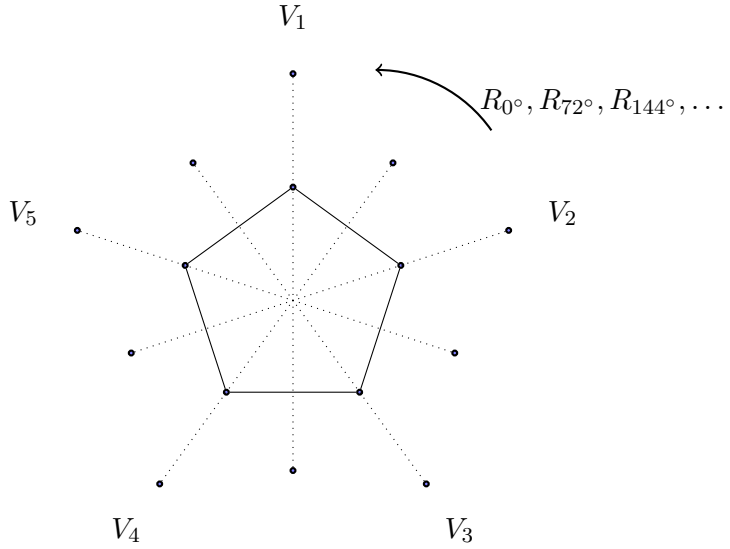
Problem 9: Let S be a nonempty set. Let G be the set of permutations on S. Prove G forms a group under function composition.

Problem 10: Prove Theorem 1.3.8 of this section. Keep in mind that Theorem 1.3.7 is known as you construct the proof.

Problem 11: Using Definition 1.3.9 as the definition of $O(n, \mathbb{R})$, show that

$$O(n, \mathbb{R}) = \{R \in \mathbb{R}^{n \times n} \mid R^T R = I\}.$$

Problem 12: Consider the dihedral group $D_5 = \{R_{0^\circ}, R_{72^\circ}, R_{144^\circ}, R_{216^\circ}, R_{288^\circ}, V_1, V_2, V_3, V_4, V_5\}$ (symmetries of a regular pentagon). [Rotations are done counter-clockwise and reflections are labeled in the picture below.]



(a) Compute $V_1 R_{72^\circ}$, $R_{144^\circ} V_3$, and $V_2 V_5$.

(b) Is D_5 Abelian? Why or why not?

(c) Find the inverse of each element ($R_{0^\circ}^{-1} =???$, $R_{72^\circ}^{-1} =???$, etc.).

(d) Find the order of each element.

1.4 Lecture 4: back to $\mathbb{Z}$-number theory

In the interest of getting to nice examples of groups I have delayed the proper discussion of the topics in this section. For most students of this course the content of this section is a review. I will not cover the entirety during the classtime, but, I include all the gory details here since we need these details to give proper, careful, arguments in the upcoming lectures about cyclic groups. This material overlaps Chapter 0 of Gallian.

1.4.1 $\mathbb{Z}$-Basics

Let's start at the very beginning, it is a good place to start.

Definition 1.4.1. *The integers $\mathbb{Z}$ are the set of natural numbers $\mathbb{N}$ together with 0 and the negatives of $\mathbb{N}$. It is possible to concretely construct (we will not) these from sets and set-operations.*

From the construction of $\mathbb{Z}$ it is clear (we assume these to be true)

1. the sum of integers is an integer
2. the product of integers is an integer
3. the usual rules of arithmetic hold for $\mathbb{Z}$

Much is hidden in (3.): let me elaborate, we assume for all $a, b, c \in \mathbb{Z}$,

$$\begin{aligned} a + b &= b + a \\ ab &= ba \\ a(b + c) &= ab + ac \\ (a + b)c &= ac + bc \\ (a + b) + c &= a + (b + c) \\ (ab)c &= a(bc) \\ a + 0 &= 0 + a = a \\ 1a &= a1. \end{aligned}$$

Where we assume the **order of operations** is done multiplication then addition; so, for example, $ab + ac$ means to first multiply a with b and a with c then you add the result.

Let me comment briefly about our standard conventions for the presentation of numbers. If I write 123 then we understand this is the **base-ten** representation. In particular,

$$123 = 1 \times 10^2 + 2 \times 10 + 3.$$

On the other hand, $1 \cdot 2 \cdot 3$ denotes the product of $1, 2$ and 3 and $1 \cdot 2 \cdot 3 = 6$. By default, algebraic variables juxtaposed denote multiplication; xy denotes x multiplied by y. If we wish for symbolic variables to denote digits in a number then we must explain this explicitly. For example, to study all numbers between 990 and 999 I could analyze $99x$ where $x \in \{0, 1, \ldots, 9\}$. But, to be clear I ought to preface such analysis by a statement like: let $99x$ be the base-ten representation of a number where x represents the 1's digit.

1.4.2 division algorithm

Division is repeated subtraction. For example, consider 11/3. Notice repeated subtraction of the dividing number[32] 3 gives:

$$11-3=8 \qquad 8-3=5 \qquad 5-3=2$$

then we cannot subtract anymore. We were able to subtract 3 copies of 3 from 11. Then we stopped at 2 since $2<3$. To summarize,

$$\boxed{11=3(3)+2}$$

We say 2 is the **remainder**; the remainder is the part which is too small to subtract for the given *dividing number*. Divide the boxed equation by the divisor to see:

$$\frac{11}{3}=3+\frac{2}{3}.$$

The generalization of the boxed equation for an arbitrary pair of natural numbers is known as the **division algorithm**.

Theorem 1.4.2. positive division algorithm: *If $a,b\in\mathbb{Z}$ with $b>0$ then there is a unique quotient $q\in\mathbb{Z}$ and remainder $r\in\mathbb{Z}$ for which $a=qb+r$ and $0\leq r<b$.*

Proof (existence): suppose $a,b\in\mathbb{Z}$ and $b>0$. Construct $R=\{a-nb \mid q\in\mathbb{Z},\quad a-nb\geq 0\}$. The set R comprises all non-negative integers which are reached from a by integer multiples of b. Explicitly,

$$R=\{a,a\pm b,a\pm 2b,\dots\}\cap\{0,1,2,\dots\}.$$

To prove R is non-empty we consider $n=-|a|\in\mathbb{Z}$ yields $a-nb=a+|a|b$. If $a\geq 0$ then clearly $a+|a|b\geq 0$. If $a<0$ then $|a|=-a$ hence $a+|a|b=-|a|+|a|b=|a|(b-1)$ but $b\in\mathbb{N}$ by assumption hence $b\geq 1$ and we find $a+|a|b\geq 0$. Therefore, as R is a non-empty subset of the non-negative integers. We apply the **Well-Ordering-Principle** to deduce there exists a smallest element $r\in R$.

Suppose r is the smallest element in R and $r\geq b$. In particular, $r=a-nb$ for some $n\in\mathbb{Z}$. Thus $a-nb\geq b$ hence $r'=a-(n+1)b\geq 0$ hence $r'\in R$ and $r'<r$. But $r'<r$ contradicts r being the smallest element. Thus, using proof by contradiction, we find $r<b$.

Proof (uniqueness): assume $q,q'\in\mathbb{Z}$ and $r,r'\in\mathbb{Z}$ such that $a=qb+r$ and $a=q'b+r'$ where $0\leq r,r'<b$. We have $qb+r=q'b+r'$ hence $(q-q')b=r-r'$. Suppose towards a contradiction $q\neq q'$. Since $q,q'\in\mathbb{Z}$ the inequality of q and q' implies $|q-q'|\geq 1$ and thus $|r-r'|=|(q-q')b|\geq|b|=b$. However, $r,r'\in[0,b)$ thus the distance[33] between r and r' cannot be larger than or equal to b. This is a contradiction, therefore, $q=q'$. Finally, $qb+r=q'b+r'$ yields $r=r'$. □

We can say more about q and r in the case $b>0$. We have

$$\frac{a}{b}=q+\frac{r}{b} \qquad \& \qquad q=\lfloor a/b\rfloor$$

[32] my resident Chinese scholar tells me in Chinese a/b has the "dividing" number b and the "divided" number a. I am tempted to call b the divisor, but the term "divisor" has a precise meaning, if b is a divisor of a then $a=mb$ for some $n\in\mathbb{Z}$. In our current discussion, to say b is a divisor assumes the remainder is zero.

[33] for a non-geometric argument here: note $0\leq r<b$ and $0\leq r'<b$ imply $-r'<r-r'<b-r'\leq b$. But, $r'<b$ gives $-b<-r'$ hence $-b<r-r'<b$. Thus $|r-r'|<b$. Indeed, the distance between r and r' is less than b.

That is q is the greatest integer which is below a/b. The function $x \mapsto \lfloor x \rfloor$ is the **floor function**. For example,

$$\lfloor -0.4 \rfloor = -1, \qquad \lfloor \pi \rfloor = 3, \qquad \lfloor n + \varepsilon \rfloor = n$$

for all $n \in \mathbb{Z}$ provided $0 \leq \varepsilon < 1$. It is easy to calculate the floor function of x when x is presented in decimal form. For example,

$$\frac{324}{11} = 29.4545... \quad \Rightarrow \quad \frac{324}{11} = 29 + 0.4545... \quad \Rightarrow \quad 324 = 29(11) + (0.4545...)(11)$$

We can calculate, $0.4545 \cdot 11 = 4.9995$. From this we find

$$324 = 29(11) + 5$$

In other words, $\frac{324}{11} = 29 + \frac{5}{11}$. The decimal form of numbers and the floor function provides a simple way to find quotients and remainders.

Consider $456/(-10) = -45.6 = -45 - 0.6$ suggests $456 = (-10)(-45) + 6$. In the case of a negative divisor ($b < 0$) the division algorithm needs a bit of modification:

Theorem 1.4.3. nonzero division algorithm: *If $a, b \in \mathbb{Z}$ with $b \neq 0$ then there is a unique quotient $q \in \mathbb{Z}$ and remainder $r \in \mathbb{Z}$ for which*

$$a = qb + r \qquad \& \qquad 0 \leq r < |b|.$$

Proof: Theorem 1.4.2 covers case $b > 0$. Thus, assume $b < 0$ hence $b' = -b > 0$. Apply Theorem 1.4.2 to $a, b' \in \mathbb{Z}$ to find q', r' such that $a = q'b' + r'$ with $0 \leq r' < b'$. However, $b' = -b = |b|$ as $b < 0$. Thus,

$$a = -q'b + r'$$

with $0 \leq r' < |b|$. Identify $q = -q'$ and $r = r'$ in the case $b < 0$. Uniqueness is clear from the equations which define q and r from the uniquely given q' and r'. This concludes the proof as $b \neq 0$ means either $b < 0$ or $b > 0$. □

The selection of the quotient in the negative divisor case is given by the **ceiling** function $x \mapsto \lceil x \rceil$. The notation $\lceil x \rceil$ indicates the next integer which is greater than or equal to x. For example,

$$\lceil 456/(-10) \rceil = -45, \quad \lceil 3.7 \rceil = 4, \quad \lceil n - \varepsilon \rceil = n$$

for all $n \in \mathbb{Z}$ given $0 \leq \varepsilon < 1$.

Remark 1.4.4. The division algorithm proves an assertion of elementary school arithmetic. For example, consider the **improper fraction** $10/3$ we can write it as the sum of 3 and $1/3$. When you write $3\frac{1}{3}$ what is truly meant is $3 + \frac{1}{3}$. In fact, the truth will set you free of a myriad of errors which arise from the poor notation $3\frac{1}{3}$. With this example in mind, let $a, b \in \mathbb{N}$. The division algorithm simply says for a/b there exists $q, r \in \mathbb{N} \cup \{0\}$ such that $a = qb + r$ hence $a/b = q + r/b$ where $0 \leq r < b$. This is merely the statement that any improper fraction can be reduced to the sum of a whole number and a proper fraction. In other words, you already knew the division algorithm. However, thinking of it without writing fractions is a bit of an adjustment for some of us.

1.4.3 divisibility in $\mathbb{Z}$

Consider $105 = 3 \cdot 5 \cdot 7$. We say 3 is a *factor* or *divisor* of 105. Also, we say 35 *divides* 105. Furthermore, 105 is a *multiple* of 3. Indeed, 105 is also a multiple of 5, 7 and even 21 or 35. Examples are nice, but, definitions are crucial:

Definition 1.4.5. *Let $a, b \in \mathbb{Z}$ then we say b* **divides** *a if there exists $c \in \mathbb{Z}$ such that $a = bc$. If b divides a then we also say b is a* **factor** *of a and a is a* **multiple** *of b.*

The notation $b \mid a$ means b divides a. If b is does not divide a then we write $b \nmid a$. The divisors of a given number are not unique. For example, $105 = 7(15) = (3)(35) = (-1)(-105)$. However, the prime divisors are unique up to reordering: $105 = (3)(5)(7)$. Much of number theory is centered around the study of primes. We ought to give a proper definition:

Definition 1.4.6. *If $p \in \mathbb{N}$ such that $n \mid p$ implies $n = p$ or $n = 1$ then we say p is* **prime**.

In words: a prime is a positive integer whose only divisors are 1 and itself.

There are many interesting features of divisibility. Notice, every number $b \in \mathbb{Z}$ divides 0 as $0 = b \cdot 0$. Furthermore, $b \mid b$ for all $b \in \mathbb{Z}$ as $b = b \cdot 1$. In related news, 1 is a factor of every integer and every integer is a multiple of 1[34]

Proposition 1.4.7. *Let $a, b, c, d, m \in \mathbb{Z}$. Then,*

(i.) *if $a \mid b$ and $b \mid c$ then $a \mid c$,*

(ii.) *if $a \mid b$ and $c \mid d$ then $ac \mid bd$,*

(iii.) *if $m \neq 0$, then $ma \mid mb$ if and only if $a \mid b$*

(iv.) *if $d \mid a$ and $a \neq 0$ then $|d| \leq |a|$.*

Proof (i.) : suppose $a \mid b$ and $b \mid c$. By the definition of divisibility there exist $m, n \in \mathbb{Z}$ such that $b = ma$ and $c = nb$. Hence $c = n(ma) = (nm)a$. Therefore, $a \mid c$ as $nm \in \mathbb{Z}$.

Proof (ii.) : suppose $a \mid b$ and $c \mid d$. By the definition of divisibility there exist $m, n \in \mathbb{Z}$ such that $b = ma$ and $d = nc$. Subsitution yields $bd = (ma)(nc) = mn(ac)$. But, $mn \in \mathbb{Z}$ hence we have shown $ac \mid bd$.

Proof (iii.) : left to the reader.

Proof (iv.) : if $d \mid a$ and $a \neq 0$ then $a = md$ for some $m \in \mathbb{Z}$. Suppose $m = 0$ then $a = (0)d = 0$ which contradicts $a \neq 0$. Therefore, $m \neq 0$. Recall that the absolute value function is multiplicative; $|md| = |m||d|$. As $m \neq 0$ we have $|m| \geq 1$ thus $|a| = |m||d| \geq |d|$. □

I hope you see these proofs are not too hard. You ought to be able to reproduce them without much effort.

Theorem 1.4.8. *Let $a_1, \ldots, a_k, c \in \mathbb{Z}$. Then,*

(i.) *if $c \mid a_i$ for $i = 1, \ldots, k$ then $c \mid (u_1 a_1 + \cdots + u_k a_k)$ for all $u_1, \ldots, u_k \in \mathbb{Z}$,*

[34] I should mention, I am partly following the excellent presentation of Jones and Jones *Elementary Number Theory* which I almost used as the text for Math 307 in Spring 2015. We're on page 4.

(ii.) *$a \mid b$ and $b \mid a$ if and only if $a = \pm b$.*

Proof (i.): suppose $c \mid a_1, c \mid a_2, \ldots, c \mid a_k$. It follows there exist $m_1, m_2, \ldots, m_k \in \mathbb{Z}$ such that $a_1 = cm_1$, $a_2 = cm_2$ and $a_k = cm_k$. Let $u_1, u_2, \ldots, u_k \in \mathbb{Z}$ and consider,

$$u_1a_1 + \cdots + u_ka_k = u_1(cm_1) + \cdots + u_k(cm_k) = c(u_1m_1 + \cdots + u_km_k).$$

Notice $u_1m_1 + \cdots + u_km_k \in \mathbb{Z}$ thus the equation above shows $c \mid (u_1a_1 + \cdots + u_ka_k)$.

Proof (ii.): suppose $a \mid b$ and $b \mid a$. If $a = 0$ then $a \mid b$ implies there exists $m \in \mathbb{Z}$ such that $b = m(0) = 0$ hence $b = 0$. Observe $a = \pm b = 0$. Continuing, we suppose $a \neq 0$ which implies $b \neq 0$ by the argument above. Notice $a \mid b$ and $b \mid a$ imply there exist $m, n \in \mathbb{Z} - \{0\}$ such that $a = mb$ and $b = na$. Multiply $a = mb$ by $n \neq 0$ to find $na = mnb$. But, $b = na$ hence $na = mn(na)$ which implies $1 = mn$. Thus, $m = n = 1$ or $m = n = -1$. These cases yield $a = b$ and $a = -b$ respective hence $a = \pm b$. □

The proof above is really not much more difficult than those we gave for Proposition 1.4.7. The most important case of the Theorem above is when $k = 2$ in part (i.).

Corollary 1.4.9. *If $c \mid x$ and $c \mid y$ then $c \mid (ax + by)$ for all $a, b \in \mathbb{Z}$.*

The result above is used repeatedly as we study the structure of common divisors.

Definition 1.4.10. *If $d \mid a$ and $d \mid b$ then d is a* **common divisor** *of a and b.*

Proposition 1.4.7 part (iv.) shows that a divisor cannot have a larger magnitude than its multiple. It follows that the largest a common divisor could be is $\max\{|a|, |b|\}$. Furthermore, 1 is a divisor of all nonzero integers. If both a and b are not zero then $\max\{|a|, |b|\} \geq 1$. Therefore, if both a and b are not zero then there must be a largest number between 1 and $\max\{|a|, |b|\}$ which divides both a and b. Thus, the definition to follow is reasonable:

Definition 1.4.11. *If $a, b \in \mathbb{Z}$, not both zero, then the* **greatest common divisor** *of a and b is denoted $gcd(a, b)$.*

The method to find the greatest common divisor which served me well as a child was simply to a and b in their prime factorization. Then to find the gcd I just selected all the primes which I could pair in both numbers.

Example 1.4.12.

$$gcd(105, 90) = gcd(\underline{3 \cdot 5} \cdot 7,\ 2 \cdot 3 \cdot \underline{3 \cdot 5}) = 3 \cdot 5 = 15.$$

The method above faces several difficulties as we attempt to solve non-elementary problems.

1. it is not an easy problem to find the prime factorization of a given integer. Indeed, this difficulty is one of the major motivations RSA cryptography.

2. it is not so easy to compare lists and select all the common pairs. Admittedly, this is not as serious a problem, but even with the simple example above I had to double-check.

Thankfully, there is a better method to find the gcd. It's old, but, popular. Euclid (yes, the same one with the parallel lines and all that) gave us the **Euclidean Algorithm**. We prove a Lemma towards developing Euclid's Algorithm.

Lemma 1.4.13. *Let $a, b, q, r \in \mathbb{Z}$. If $a = qb + r$ then $gcd(a, b) = gcd(b, r)$.*

Proof: by Corollary 1.4.9 we see a divisor of both b and r is also a divisor of a. Likewise, as $r = a - qb$ we see any common divisor of a and b is also a divisor of r. It follows that a, b and b, r share the same divisors. Hence, $gcd(a, b) = gcd(b, r)$. $\square$

We now work towards Euclid's Algorithm. Let $a, b \in \mathbb{Z}$, not both zero. Our goal is to calculate $gcd(a, b)$. If $a = 0$ and $b \neq 0$ then $gcd(a, b) = |b|$. Likewise, if $a \neq 0$ and $b = 0$ then $gcd(a, b) = |a|$. Note $gcd(a, a) = |a|$ hence we may asssume $a \neq b$ in what follows. Furthermore,

$$gcd(a, b) = gcd(-a, b) = gcd(a, -b) = gcd(-a, -b).$$

Therefore, suppose $a, b \in \mathbb{N}$ with $a > b$[35]. Apply the division algorithm (Theorem 1.4.2) to select q_1, r_1 such that

$$a = q_1 b + r_1 \qquad \text{such that} \qquad 0 \leq r_1 < b.$$

If $r_1 = 0$ then $a = q_1 b$ hence $b \mid a$ and as b is the largest divisor of b we find $gcd(a, b) = b$. If $r_1 \neq 0$ then we continue to apply the division algorithm once again to select q_2, r_2 such that

$$b = q_2 r_1 + r_2 \qquad \text{such that} \qquad 0 \leq r_2 < r_1.$$

If $r_2 = 0$ then $r_1 \mid b$ and clearly $gcd(b, r_1) = r_1$. However, as $a = q_1 b + r_1$ allows us to apply Lemma 1.4.13 to obtain $gcd(a, b) = gcd(b, r_1) = r_1$. Continuing, we suppose $r_2 \neq 0$ with $r_1 > r_2$ hence we may select q_3, r_3 for which:

$$r_1 = q_3 r_2 + r_3 \qquad \text{such that} \qquad 0 \leq r_3 < r_2.$$

Once again, if $r_3 = 0$ then $r_2 \mid r_1$ hence it is clear $gcd(r_1, r_2) = r_2$. However, as $b = q_2 r_1 + r_2$ gives $gcd(b, r_1) = gcd(r_1, r_2)$ and $a = q_1 b + r_1$ gives $gcd(a, b) = gcd(b, r_1)$ we find that $gcd(a, b) = r_2$. This process continues. It cannot go on forever as we have the conditions:

$$0 < \cdots < r_3 < r_2 < r_1 < b.$$

There must exist some $n \in \mathbb{N}$ for which $r_{n+1} = 0$ yet $r_n \neq 0$. All together we have:

$$\begin{aligned} a &= q_1 b + r_1, \\ b &= q_2 r_1 + r_2, \\ r_1 &= q_3 r_2 + r_3, \ldots, \\ r_{n-2} &= q_n r_{n-1} + r_n, \\ r_{n-1} &= q_{n+1} r_n. \end{aligned}$$

The last condition yields $r_n \mid r_{n-1}$ hence $gcd(r_{n-1}, r_n) = r_n$. Furthermore, we find, by repeated application of Lemma 1.4.13 the following string of equalities

$$gcd(a, b) = gcd(b, r_1) = gcd(r_1, r_2) = gcd(r_2, r_3) = \cdots = gcd(r_{n-1}, r_n) = r_{n-1}.$$

In summary, we have shown that repeated division of remainders into remainder gives a strictly decreasing sequence of positive integers whose last member is precisely $gcd(a, b)$.

[35] the equation above shows we can cover all other cases once we solve the problem for positive integers.

Theorem 1.4.14. Euclidean Algorithm: *suppose $a, b \in \mathbb{N}$ with $a > b$ and form the finite sequence $\{b, r_1, r_2, \ldots, r_n\}$ for which $r_{n+1} = 0$ and $b, r_1, \ldots, r_n$ are defined as discussed above. Then $gcd(a, b) = r_n$.*

Example 1.4.15. *Let me show you how the euclidean algorithm works for a simple example. Consider $a = 100$ and $b = 44$. Euclid's algorithm will allow us to find $gcd(100, 44)$.*

1. $100 = 44(2) + 12$ *divided 100 by 44 got remainder of 12*
2. $44 = 12(3) + 8$ *divided 44 by 12 got remainder of 8*
3. $12 = 8(1) + \boxed{4}$ *divided 12 by 8 got remainder of 4*
4. $8 = 4(2) + 0$ *divided 4 by 1 got remainder of zero*

The last nonzero remainder will always be the gcd when you play the game we just played. Here we find $\boxed{gcd(100, 44) = 4}$. *Moreover, we can write 4 as a $\mathbb{Z}$-linear combination of 100 and 44. This can be gleaned from the calculations already presented by working backwards from the gcd:*

3. $4 = 12 - 8$
2. $8 = 44 - 12(3)$ *implies* $4 = 12 - (44 - 12(3)) = 4(12) - 44$
1. $12 = 100 - 44(2)$ *implies* $4 = 4(100 - 44(2)) - 44 = 4(100) - 9(44)$

I call this a "$\mathbb{Z}$-linear combination of 100 and 44 since $4, -9 \in \mathbb{Z}$. We find $\boxed{4(100) - 9(44) = 4}$.

The fact that we can always work euclid's algorithm backwards to find how the $gcd(a, b)$ is written as $ax + by = gcd(a, b)$ for some $x, y \in \mathbb{Z}$ is remarkable. I continue to showcase this side-benefit of the Euclidean Algorithm as we continue. We will give a general argument after the examples. I now shift to a less verbose presentation:

Example 1.4.16. *Find $gcd(62, 626)$*

$$626 = 10(62) + 6$$
$$62 = 10(6) + 2$$
$$6 = 3(2) + 0$$

From the E.A. I deduce $gcd(62, 626) = 2$. Moreover,

$$2 = 62 - 10(6) = 62 - 10[626 - 10(62)] = 101(62) - 10(626)$$

Example 1.4.17. *Find $gcd(240, 11)$.*

$$240 = 11(21) + 9$$
$$11 = 9(1) + 2$$
$$9 = 2(4) + 1$$
$$2 = 1(2)$$

Thus, by E.A., $gcd(240, 11) = 1$. Moreover,

$$1 = 9 - 2(4) = 9 - 4(11 - 9) = -4(11) + 5(9) = -4(11) + 5(240 - 11(21))$$

That is,

$$\boxed{1 = -109(11) + 5(240)}$$

Example 1.4.18. *Find $gcd(4, 20)$. This example is a bit silly, but I include it since it is an exceptional case in the algorithm. The algorithm works, you just need to interpret the instructions correctly.*

$$20 = 4(5) + 0$$

Since there is only one row to go from we identify 4 as playing the same role as the last non-zero remainder in most examples. Clearly, $gcd(4, 20) = 4$. Now, what about working backwards? Since we do not have the gcd appearing by itself in the next to last equation (as we did in the last example) we are forced to solve the given equation for the gcd,

$$20 = 4(4 + 1) = 4(4) + 4 \implies \boxed{20 - 4(4) = 4}$$

The following result also follows from the discussion before Theorem 1.4.14. I continue to use the notational set-up given there.

Theorem 1.4.19. Bezout's Identity: *if $a, b \in \mathbb{Z}$, not both zero, then there exist $x, y \in \mathbb{Z}$ such that $ax + by = gcd(a, b)$.*

Proof: we have illustrated the proof in the examples. Basically we just back-substitute the division algorithms. For brevity of exposition, I assume $r_3 = gcd(a, b)$. It follows that:

$$\begin{aligned} a = q_1 b + r_1 &\Rightarrow r_1 = a - q_1 b \\ b = q_2 r_1 + r_2 &\Rightarrow r_2 = b - q_2 r_1 \\ r_1 = q_3 r_2 + r_3 &\Rightarrow r_3 = r_1 - q_3 r_2 \end{aligned}$$

where $gcd(a, b) = r_3$. Moreover, $r_2 = b - q_2(a - q_1 b)$ implies $r_3 = r_1 - q_3[b - q_2(a - q_1 b)]$. Therefore,

$$gcd(a, b) = a - q_1 b - q_3[b - q_2(a - q_1 b)] = a - (q_1 - q_3[1 - q_2(a - q_1)]b.$$

Identify $x = 1$ and $y = q_1 - q_3[1 - q_2(a - q_1)]$. □

We should appreciate that x, y in the above result are far from unique. However, as we have shown, the method at least suffices to find a solution of the equation $ax + by = gcd(a, b)$. One nice application of Bezout's identity is seen in *Euclid's Lemma*: intuitively, Euclid's Lemma testifies to the indestructibly of primes.

Lemma 1.4.20. (Euclid): *Let $a, b \in \mathbb{Z}$. If $p \in \mathbb{Z}$ is prime and $p \mid ab$ then $p \mid a$ or $p \mid b$.*

Proof: Suppose $a, b, p \in \mathbb{Z}$ and p is prime. Further, suppose $p \mid ab$ but $p \nmid a$. Since p does not divide a we have $\gcd(a, p) = 1$ and by Bezout's identity there exist $x, y \in \mathbb{Z}$ for which $ax + py = 1$. Multiply by b to obtain $bax + bpy = b$ (label this by $\star$. Since $p \mid ab$ we know there exists $c \in \mathbb{Z}$ for which $ab = cp$. Hence, returning to $\star$,

$$b = cpx + bpy = p(cx + by)$$

since $cx + by \in \mathbb{Z}$ the result above clearly shows $p \mid b$ and Euclid's Lemma follows. □

I recommend working on Chapter 0 Gallian problems such as:

$$\#4, 7, 8, 16, 17$$

these should help bring the Group concept to life. I do not collect these, but, I will keep them in mind as I construct tests.

Problems for Lecture 4: (these are collected at the beginning of Lecture 6)

Problem 13: The Euclidean Algorithm

(a) Use the Euclidean Algorithm to find the greatest common divisor (gcd) of 1234 and 542.

(b) Use the (extended) Euclidean Algorithm to find the greatest common divisor of $a = 1001$ and $b = 53$, say $d = \gcd(a, b)$. Then determine integers x and y such that $ax + by = d$.

(c) Use the (extended) Euclidean Algorithm to find 9^{-1} in $U(1000)$.

Problem 14: Let $d = \gcd(a, b)$. If $a = da'$ and $b = db'$, show that $\gcd(a', b') = 1$. [Of course, $a, a', b, b', d \in \mathbb{Z}$.]

Problem 15: Prove the rule of 9's: that is, show that $n \in \mathbb{N}$ is divisible by 9 iff the sum of the digits in its base-ten representation is divisible by 9. *e.g.* $136,098$ *is divisible by* 9 *since* $1 + 3 + 6 + 0 + 9 + 8 = 27$. *You may use results from the next Lecture. In fact, I would encourage you to offer a solution in terms of congruence.*

Problem 16: Dihedral groups: generators and relations style. Recall that ...

$$D_4 = \langle x, y \mid x^4 = 1,\ y^2 = 1,\ \text{and } (xy)^2 = 1 \rangle = \{1, x, x^2, x^3, y, xy, x^2y, x^3y\}$$

(a) Write down the Cayley table for D_4.

(b) Find the inverse of each element (i.e. $1^{-1} =???$, $x^{-1} =???$, etc.).

(c) Find the order of each element.

(d) Find all of the **distinct** cyclic subgroups of D_4.

(e) What is in $Z(D_4)$ (recall that $Z(G)$ is the *center* of G)?

(f) Simplify $x^6y^{-3}x^3y^8x^{-5}yxy$.

1.5 Lecture 5: modular arithmetic and groups

This material overlaps Chapter 0 of Gallian, although, I have nothing to say about ISBN error digits. It is an interesting topic and I assigned a token problem so you can appreciate that application of modular arithmetic. In this section we assume $n \in \mathbb{N}$ throughout. In summary, we develop a careful model for $\mathbb{Z}_n$ in this section.

Remark 1.5.1. I use some notation in this section which we can omit elsewhere for the sake of brevity. In particular, in the middle of this section I might use the notation $[2]$ or $\bar{2}$ for $2 \in \mathbb{Z}_n$ whereas in later work we simply use 2 with the understanding that we are working in the context of modular arithmetic. I have a bit more to say about this notational issue and the deeper group theory it involves at the conclusion of this section.

Definition 1.5.2. *$a \equiv b \ mod(n)$ if and only if $n \mid (b-a)$.*

The definition above is made convenient by the simple equivalent criteria below:

Theorem 1.5.3. *Let $a, b \in \mathbb{Z}$ then we say a is* **congruent** *to $b \ mod(n)$ and write $a \equiv b \ mod(n)$ if a and b have the same remainder when divided by n.*

Proof: Suppose $a \equiv b \ mod(n)$ then a and b share the same remainder after division by n. By the Division Algorithm, there exist $q_1, q_2 \in \mathbb{Z}$ for which $a = q_1 n + r$ and $b = q_2 n + r$. Observe, $b - a = (q_2 n + r) - (q_1 n + r) = (q_2 - q_1)n$. Therefore, $n \mid (b-a)$.

Conversely, suppose $n \mid (b-a)$ then there exists $q \in \mathbb{Z}$ for which $b - a = qn$. Apply the Division Algorithm to find q_1, q_2 and r_1, r_2 such that: $a = q_1 n + r_1$ and $b = q_2 n + r_2$ with $0 \leq r_1 < n$ and $0 \leq r_2 < n$. We should pause to note $|r_2 - r_1| < n$. Observe,

$$b - a = qn = (q_2 n + r_2) - (q_1 n + r_1) = (q_2 - q_1)n + r_2 - r_1.$$

Therefore, solving for the difference of the remainders and taking the absolute value,

$$|q - q_2 + q_1|n = |r_2 - r_1|$$

Notice $|q - q_2 + q_1| \in \mathbb{N} \cup \{0\}$ and $|r_2 - r_1| < n$. It follows $|q - q_2 + q_1| = 0$ hence $|r_2 - r_1| = 0$ and we conclude $r_1 = r_2$. □

Congruence has properties you might have failed to notice as a child.

Proposition 1.5.4. *Let n be a positive integer, for all $x, y, z \in \mathbb{Z}$,*

(i.) *$x \equiv x \ mod(n)$,*

(ii.) *$x \equiv y \ mod(n)$ implies $y \equiv x \ mod(n)$,*

(iii.) *if $x \equiv y \ mod(n)$ and $y \equiv z \ mod(n)$ then $x \equiv z \ mod(n)$.*

Proof: we use Definition 1.5.2 throughout what follows.
(i.) Let $x \in \mathbb{Z}$ then $x - x = 0 = 0 \cdot n$ hence $n \mid (x - x)$ and we find $x \equiv x \ mod(n)$.
(ii.) Suppose $x \equiv y \ mod(n)$. Observe $n \mid (x - y)$ indicates $x - y = nk$ for some $k \in \mathbb{Z}$. Hence $y - x = n(-k)$ where $-k \in \mathbb{Z}$. Therefore, $n \mid (y - x)$ and we find $y \equiv x \ mod(n)$.

(iii.) Suppose $x \equiv y \ mod(n)$ and $y \equiv z \ mod(n)$. Thus $n \mid (y-x)$ and $n \mid z-y$. Corollary 1.4.9 indicates n also divides the sum of two integers which are each divisible by n. Thus, $n \mid [(y-x)+(z-y)]$ hence $n \mid (z-x)$ which shows $x \equiv z \ mod(n)$. □

I referenced the Corollary to prove part (iii.) to remind you how our current discussion fits naturally with our previous discussion.

Corollary 1.5.5. *Let $n \in \mathbb{N}$. Congruence modulo n forms an equivalence relation on $\mathbb{Z}$.*

This immediately informs us of an interesting **partition** of the integers. Recall, a **partition** of a set S is a family of subsets $U_\alpha \subseteq S$ where $\alpha \in \Lambda$ is some index set such that $U_\alpha \cap U_\beta = \emptyset$ for $\alpha \neq \beta$ and $\cup_{\alpha\in\Lambda} U_\alpha = S$. A partition takes a set and parses it into disjoint pieces which cover the whole set. The partition induced from an equivalence relation is simply formed by the **equivalence classes** of the relation. Let me focus on $\mathbb{Z}$ with the equivalence relation of congruence modulo a positive integer n. We define:[36]:

Definition 1.5.6. equivalence classes of $\mathbb{Z}$ modulo $n \in \mathbb{N}$:

$$[x] = \{y \in \mathbb{Z} \mid y \equiv x \ mod(n)\}$$

Observe, there are several ways to characterize such sets:

$$[x] = \{y \in \mathbb{Z} \mid y \equiv x \ mod(n)\} = \{y \in \mathbb{Z} \mid y - x = nk \text{ for some } k \in \mathbb{Z}\} = \{x + nk \mid k \in \mathbb{Z}\}.$$

I find the last presentation of $[x]$ to be useful in practical computations.

Example 1.5.7. *Congruence mod(2) partititions $\mathbb{Z}$ into even and odd integers:*

$$[0] = \{2k \mid k \in \mathbb{Z}\} \qquad \& \qquad [1] = \{2k+1 \mid k \in \mathbb{Z}\}$$

Example 1.5.8. *Congruence mod(4) partititions $\mathbb{Z}$ into four classes of numbers:*

$$[0] = \{4k \mid k \in \mathbb{Z}\} = \{\ldots, -8, -4, 0, 4, 8, \ldots\}$$

$$[1] = \{4k+1 \mid k \in \mathbb{Z}\} = \{\ldots, -7, -3, 1, 5, 9, \ldots\}$$

$$[2] = \{4k+2 \mid k \in \mathbb{Z}\} = \{\ldots, -6, -2, 2, 6, 10, \ldots\}$$

$$[3] = \{4k+3 \mid k \in \mathbb{Z}\} = \{\ldots, -5, -1, 3, 7, 11, \ldots\}$$

The patterns above are interesting, there is something special about [0] and [2] in comparison to [1] and [3]. Patterns aside, the notation of the previous two example can be improved. Let me share a natural notation which helps us understand the structure of congruence classes.

Definition 1.5.9. Coset Notation: *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$ we define:*

$$n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\} \qquad a + n\mathbb{Z} = \{a + nk \mid k \in \mathbb{Z}\}.$$

Observe, in the notation just introduced, we have

$$\boxed{[a] = a + n\mathbb{Z}}$$

[36] there are other notations, the concept here is far more important than the notation we currently employ

Example 1.5.10. *Congruence mod*(2) *partitions* $\mathbb{Z}$ *into even and odd integers:*

$$[0] = 2\mathbb{Z} \qquad \& \qquad [1] = 1 + 2\mathbb{Z}.$$

Example 1.5.11. *Congruence mod*(4) *partitions* $\mathbb{Z}$ *into four classes of numbers:*

$$[0] = 4\mathbb{Z}, \quad [1] = 1 + 4\mathbb{Z}, \quad [2] = 2 + 4\mathbb{Z}, \quad [3] = 3 + 4\mathbb{Z}.$$

We should pause to appreciate a subtle aspect of the notation. It is crucial to note $[x] = [y]$ does **not** imply $x = y$. For example, modulo 2:

$$[1] = [3] = [7] = [1000037550385987987987971] \qquad \& \qquad [2] = [-2] = [-42].$$

Or, modulo 9:

$$[1] = [10] = [-8], \qquad \& \qquad [3] = [12] = [-6], \qquad \& \qquad [0] = [90] = [-9].$$

Yet, modulo 9, $[1] \neq [3]$. Of course, I just said $[1] = [3]$. How can this be? Well, context matters. In some sense, the notation $[x]$ is dangerous and $[x]_n$ would be better. We could clarify that $[1]_2 = [3]_2$ whereas $[1]_9 \neq [3]_9$. I don't recall such notation used in any text. What is more common is to use the *coset notation* to clarify:

$$1 + 2\mathbb{Z} = 3 + 2\mathbb{Z} \qquad \text{whereas} \qquad 1 + 9\mathbb{Z} \neq 3 + 9\mathbb{Z}.$$

I'm not entirely sure the Proposition below is necessary.

Proposition 1.5.12. *Let* $n \in \mathbb{N}$. *We have* $[x] = [y]$ *if and only if* $x \equiv y \ mod(n)$. *Or, in the coset notation* $x + n\mathbb{Z} = y + n\mathbb{Z}$ *if and only if* $y - x \in n\mathbb{Z}$.

Proof: Observe $x \in [x]$. If $[x] = [y]$ then $x \in [y]$ hence there exists $k \in \mathbb{Z}$ for which $x = y + nk$ hence $x - y = nk$ and we find $x \equiv y \ mod(n)$. Conversely, if $x \equiv y \ mod(n)$ then there exists $k \in \mathbb{Z}$ such that $y - x = nk$ thus $x = y - nk$ and $y = x + nk$. Suppose $a \in [x]$ then there exists $j \in \mathbb{Z}$ for which $a = nj + x$ hence $a = nj + y - nk = n(j - k) + y \in [y]$. We have shown $[x] \subseteq [y]$. Likewise, if $b \in [y]$ then there exists $j \in \mathbb{Z}$ for which $b = nj + y$ hence $b = nj + x + nk = n(j + k) + x \in [x]$. Thus $[y] \subseteq [x]$ and we conclude $[x] = [y]$. □

Notice the proposition above allows us to calculate as follows: for $n \in \mathbb{N}$

$$na + b + n\mathbb{Z} = b + n\mathbb{Z} \qquad \text{or} \qquad [na + b] = [b]$$

for $a, b \in \mathbb{Z}$. There is more.

Proposition 1.5.13. *Let* $n \in \mathbb{N}$. *If* $[x] = [x']$ *and* $[y] = [y']$ *then*

(i.) $[x + y] = [x' + y']$,

(ii.) $[xy] = [x'y']$

(iii.) $[x - y] = [x' - y']$

Proof: Suppose $[x] = [x']$ and $[y] = [y']$. It follows there exists $j, k \in \mathbb{Z}$ such that $x' = nj + x$ and $y' = nk + y$. Notice $x' \pm y' = nj + x \pm (nk + y) = n(j \pm k) + x \pm y$. Therefore, $x \pm y \equiv x' \pm y'$ $mod(n)$ and by Proposition 1.5.12 we find $[x \pm y] = [x' \pm y']$. This proves $(i.)$ and $(iii.)$. Next, consider:

$$x'y' = (nj + x)(nk + y) = n(jkn + jy + xk) + xy$$

thus $x'y' \equiv xy$ $mod(n)$ we apply Proposition 1.5.12 once more to find $[xy] = [x'y']$. □

We ought to appreciate the content of the proposition above as it applies to congruence modulo n. In fact, the assertions below all apear in the proof above.

Corollary 1.5.14. *Let $n \in \mathbb{N}$. If $x \equiv x'$ and $y \equiv y'$ modulo n then*

(i.) $x + y \equiv x' + y'$ $mod(n)$,

(ii.) $xy \equiv x'y'$ $mod(n)$,

(iii.) $x - y \equiv x' - y'$ $mod(n)$,

Example 1.5.15. *Suppose $x + y \equiv 3$ and $x - y \equiv 1$ modulo 4. Then, by Corollary 1.5.14 we add and substract the given congruences to obtain:*

$$2x \equiv 4 \qquad 2y \equiv 2$$

There are 4 cases to consider. Either $x \in [0]$, $x \in [1]$, $x \in [2]$ or $x \in [3]$. Observe,

$$\begin{array}{ll} 2(0) \equiv 0 \equiv 4, & 2(0) \not\equiv 2 \\ 2(1) \equiv 2 \not\equiv 4, & 2(1) \equiv 2 \\ 2(2) \equiv 4, & 2(2) \equiv 4 \not\equiv 2 \\ 2(3) \equiv 2 \not\equiv 4, & 2(3) \equiv 2. \end{array}$$

It follows that $x \in [0] \cup [2]$ and $y \in [1] \cup [3]$ forms the solution set of this system of congruences.

The method I used to solve the above example was not too hard since there were just 4 cases to consider. I suppose, if we wished to solve the same problem modulo 42 we probably would like to learn a better method.

Proposition 1.5.13 justifies that the definition below does give a **binary operation** on the set of equivalence classes modulo n. Recall, a *binary operation* on a set S is simply a *function* from $S \times S$ to S. It is a single-valued assignment of pairs of S-elements to S-elements.

Definition 1.5.16. modular arithmetic: *let $n \in \mathbb{N}$, define*

$$[x] + [y] = [x + y] \qquad \& \qquad [x][y] = [xy]$$

for all $x, y \in \mathbb{Z}$. Or, if we denote the set of all equivalence classes modulo n by $\mathbb{Z}/n\mathbb{Z}$ then write: for each $x + n\mathbb{Z}, y + n\mathbb{Z} \in \mathbb{Z}/n\mathbb{Z}$

$$(x + n\mathbb{Z}) + (y + n\mathbb{Z}) = x + y + n\mathbb{Z} \qquad \& \qquad (x + n\mathbb{Z})(y + n\mathbb{Z}) = xy + n\mathbb{Z}.$$

Finally, we often use the notation $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.

Notice the operation defined above is a binary operation on $\mathbb{Z}/n\mathbb{Z}$ (not $\mathbb{Z}$). Many properties of integer arithmetic transfer to $\mathbb{Z}/n\mathbb{Z}$:

$$\begin{aligned}
[a]+[b] &= [b]+[a] \\
[a][b] &= [b][a] \\
[a]([b]+[c]) &= [a][b]+[a][c] \\
([a]+[b])[c] &= [a][c]+[b][c] \\
([a]+[b])+[c] &= [a]+([b]+[c]) \\
([a][b])[c] &= [a]([b][c]) \\
[a]+[0] &= [0]+[a]=[a] \\
[1][a] &= [a][1].
\end{aligned}$$

Furthermore, for $k \in \mathbb{N}$,

$$\begin{aligned}
[a_1]+[a_2]+\cdots+[a_k] &= [a_1+a_2+\cdots+a_k] \\
[a_1][a_2]\cdots[a_k] &= [a_1a_2\cdots a_k] \\
[a]^k &= [a^k].
\end{aligned}$$

Example 1.5.17. *Simplify* $[1234]$ *modulo* 5. *Notice,*

$$1234 = 1\times 10^3 + 2\times 10^2 + 3\times 10 + 4.$$

However, $10 = 2(5)$ *thus,*

$$1234 = 1\times 2^3 5^3 + 2\times 2^2 5^2 + 3\times 2\cdot 5 + 4.$$

Note, $[5]=[0]$ *hence* $[5^k]=[0]$ *for* $k\in\mathbb{N}$. *By the properties of modular arithmetic it is clear that the* $10's$, $100's$ *and* $1000's$ *digits are irrelevant to the result. Only the first digit matters,* $[1234]=[4]$.

It is not hard to see the result of the example above equally well applies to larger numbers; if $a_k,\ a_{k-1},\ldots,a_2,a_1$ are the digits in a decimal representation of an integer then $[a_ka_{k-1}\cdots a_2a_1] = [a_1]$ *mod*(5).

Example 1.5.18. *Calculate the cube of* 51 *modulo* 7.

$$[51^3] = [51][51][51] = [51]^3 = [49+2]^3 = [2]^3 = [8].$$

Of course, you can also denote the same calculation via congruence:

$$51^3 = 51\cdot 51\cdot 51 \equiv 2\cdot 2\cdot 2 = 8 \quad\Rightarrow\quad [51^3]=[8].$$

The next example is a cautionary tale:

Example 1.5.19. *Simplify* 7^{100} *modulo* 6. *Consider,*

$$[7^{100}] = [7]^{100} = [1]^{100} = [1^{100}] = [1].$$

or, (incorrectly !)

$$[7^{100}] = [7^{[100]}] = [7^{[6(16)+4]}] = [7^4] = [28] = [4].$$

The point is this: it is **not** *true that* $[a^k] = [a^{[k]}]$.

Naturally, as we discuss $\mathbb{Z}_n$ it is convenient to have a particular choice of representative for this set of residues. Two main choices: the *set of least non-negative residues*

$$\mathbb{Z}_n = \{[0], [1], [2], \ldots, [n-1]\}$$

alternatively, *set of least absolute value residues* or simply *least absolute residues*

$$\mathbb{Z}_n = \{[0], [\pm 1], [\pm 2], \ldots\}$$

where the details depend on if n is even or odd. For example,

$$\mathbb{Z}_5 = \{[0], [1], [2], [3], [4]\} = \{[-2], [-1], [0], [1], [2]\}$$

or,

$$\mathbb{Z}_4 = \{[0], [1], [2], [3]\} = \{[-2], [-1], [0], [1]\}$$

Honestly, if we work in the particular context of $\mathbb{Z}_n$ then there is not much harm in dropping the $[\cdot]$-notation. Sometimes, I use $[x] = \bar{x}$. Whichever notation we choose, we must be careful to not fall into the trap of assuming the usual properties of $\mathbb{Z}$ when calculating in the specific context of modular arithmetic. The example that follows would be very clumsy to write in the $[\cdot]$-notation.

Example 1.5.20. *Consider $f(x) = x^2 + 2x + 3$ for $x \in \mathbb{Z}_5$. We can determine if f has a zero by explicit calculation modulo 5:*

$$f(-2) = (-2)^2 + 2(-2) + 3 = 3$$

$$f(-1) = (-1)^2 + 2(-1) + 3 = 2$$

$$f(0) = (0)^2 + 2(0) + 1 = 3$$

$$f(1) = 1 + 2 + 3 \equiv 1$$

$$f(2) = 4 + 4 + 3 \equiv 1$$

Therefore, $f(x)$ has no zero for $x \in \mathbb{Z}_5$.

The examples below are from Jones and Jones' *Elementary Number Theory* pages 42-43.

Example 1.5.21. *Calculate the least positive residue of 28×33 modulo 35. Note that $28 \equiv 28 - 35 = -7$ and $33 \equiv 33 - 35 = -2$ hence $28 \times 33 \equiv (-7) \times (-2) = 14$. Or, $[28][33] = [14]$.*

Example 1.5.22. *Calculate the least absolute residue of 15×59 mod(75). Observe $59 \equiv 59 - 75 = -16$ thus*

$$59 \times 15 \equiv -16 \times 15 = (-1 - 15) \times 15 = -15 - 3(75) \equiv -15.$$

Since $|-15| = 15 \leq 75/2$ it is clear -15 is the least absolute residue modulo 75.

Example 1.5.23. *To calculate 3^8 modulo 13 we break the problem into several doublings; $3^8 = ((3^2)^2)^2$. At each stage we take care to use modular arithmetic to simplify:*

$$3^2 = 9 \equiv -4$$

modulo 13. Next,

$$3^4 = (3^2)^2 \equiv (-4)^2 = 16 \equiv 3$$

thus

$$3^8 = (3^4)^2 \equiv 3^2 = 9.$$

Example 1.5.24. *Prove that $a(a+1)(a+2)$ is divisible by 6 for each integer a. In other words, we wish to show $a(a+1)(a+2) \equiv 0$ mod(6). Note $\mathbb{Z}_6 = \{[0], [\pm 1], [\pm 2], [3]\}$ so consider:*

$$\begin{aligned} a = 0: &\quad a(a+1)(a+2) = 0, \\ a = \pm 1: &\quad a(a+1)(a+2) = (\pm 1)(1 \pm 1)(2 \pm 1) = \{6, 0\} \equiv 0, \\ a = \pm 2: &\quad a(a+1)(a+2) = (\pm 2)(1 \pm 2)(2 \pm 2) = \{12, 0\} \equiv 0, \\ a = 3: &\quad a(a+1)(a+2) = (3)(3+1)(3+2) = 60 \equiv 0. \end{aligned}$$

Therefore, $a(a+1)(a+2) \equiv 0$ modulo 6 for all $a \in \mathbb{Z}$ hence $6 \mid a(a+1)(a+2)$ for all $a \in \mathbb{Z}$.

The claim in the example above is very obviously true if we just think about some cases $1 \cdot 2 \cdot 3$,$2 \cdot 3 \cdot 4$,...$10 \cdot 11 \cdot 12$, $11 \cdot 12 \cdot 13$ etc. You can see the reason a 6 appears is that in any triple of successive integers you have at least one number divisible by 3 and at least one number divisible by 2. This suggests a different method of proof.

Example 1.5.25. *Prove that $a(a+1)(a+2)$ is divisible by 6 for each integer a. Once again, we wish to show $a(a+1)(a+2) \equiv 0$ mod(6). Observe, if $2 \mid x$ and $3 \mid x$ then $x = 2j$ and $x = 3k$ for some $j, k \in \mathbb{Z}$. It follows from the prime factorization of integers that $3 \mid j$ and $2 \mid k$ hence[37] there exists $m \in \mathbb{Z}$ for which $j = 3m$ and we find $x = 2j = 2(3m) = 6m$ which proves $6 \mid x$. Therefore, if we are able to show $a(a+1)(a+2)$ is divisible by 2 and 3 it follows $a(a+1)(a+2)$ is divisible by 6. Consider congruence modulo 2:*

$$\begin{aligned} a = 0: &\quad a(a+1)(a+2) = 0, \\ a = 1: &\quad a(a+1)(a+2) = (1)(2)(3) \equiv 0. \end{aligned}$$

Next, the modulo 3 case:

$$\begin{aligned} a = 0: &\quad a(a+1)(a+2) = 0, \\ a = 1: &\quad a(a+1)(a+2) = (1)(2)(3) \equiv 0, \\ a = 2: &\quad a(a+1)(a+2) = (2)(3)(4) \equiv 0. \end{aligned}$$

Thus $a(a+1)(a+2) \equiv 0$ modulo 6 and we conclude $6 \mid a(a+1)(a+2)$ for each $a \in \mathbb{Z}$.

Notice I had to invoke the Fundalmental Theorem of Arithmetic in the example above. Let me state it without proof here:

Theorem 1.5.26. *Let $n \in \mathbb{N}$ then there exist a unique set of distinct primes $p_1, p_2, \ldots, p_k$ and multiplicities $r_1, r_2, \ldots, r_k$ for which $n = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$.*

Proof: to be found in Math 307 (the number theory course). □

We already saw a specific case of the theorem below in action to solve Example 1.5.25.

Theorem 1.5.27. *Let $n \in \mathbb{N}$ such that there exist a unique set of distinct primes $p_1, p_2, \ldots, p_k$ and multiplicities $r_1, r_2, \ldots, r_k$ for which $n = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$. Then $a \equiv b$ mod(n) if and only if $a \equiv b$ mod$(p_i^{r_i})$ for each $i = 1, 2, \ldots k$.*

[37] yes, I could just as well have messed with k

Proof: to be found in Math 307(the number theory course). □

The theme of this section is illustrate the structure and utility of modular arithmetic. The Theorem below is certainly a showcase of the technique. The problem of determining if $f(x) = 0$ for some $x \in \mathbb{Z}$ is somewhat daunting as there are infinitely many integers. However, for polynomial $f(x)$ we are able to answer this question by analyzing the corresponding polynomial over $\mathbb{Z}_n$. Let's study an example before I state the general theorem.

Example 1.5.28. *Show $f(x) = x^5 - x^2 + x - 3$ has no integer roots. Consider, modulo 4,*

$$f(0) = -3, \qquad f(1) = 1 - 1 + 1 - 3 = -2,$$

$$f(-1) = -1 - 1 - 1 - 3 = -6 \equiv 2, \qquad f(2) = 32 - 4 + 2 - 3 \equiv -1.$$

This means there is no integer for which $f(x) = 0$. Why? Because $\mathbb{Z} = 4\mathbb{Z} \cup (4\mathbb{Z}+1) \cup (4\mathbb{Z}+2) \cup (4\mathbb{Z}+3)$ and we have shown each partition gives no value in $4\mathbb{Z}$ hence no integer input into $f(x)$ returns a value of 0.

Theorem 1.5.29. *Let $f(x) \in \mathbb{Z}[x]$, that is let $f(x)$ be a polynomial with integer coefficients, and suppose $n \in \mathbb{N}$. If $a \equiv b \ mod(n)$ then $f(a) \equiv f(b) \ mod(n)$.*

Proof: Suppose $a \equiv b \ mod(n)$ and $f(x) = c_m x^m + \cdots + c_1 x + c_0$ where $c_m, \ldots, c_1, c_0 \in \mathbb{Z}$. Consider then, by repeated application of Corollary 1.5.14 we have:

$$f(a) = c_m a^m + \cdots + c_1 a + c_0 \equiv c_m b^m + \cdots + c_1 b + c_0 = f(b). \quad \square$$

To solve Example 1.5.28 we used the **contrapositive**. Let me remind you: the contrapositive allows us to know that when $P \Rightarrow Q$ is true then $\widetilde{Q} \Rightarrow \widetilde{P}$ is true. Here I use P, Q to denote statements and $\widetilde{P}, \widetilde{Q}$ to denote the negation of those statements. Suppose $f(a) = 0$ for some some $a \in \mathbb{Z}$. Then a clear implication is that $f(a) \equiv 0 \ mod(n)$ for all $n \in \mathbb{N}$. In this case P is the statement about integer zeros whereas Q is the statement about the congruence of $f(a)$ modulo n for all $n \in \mathbb{N}$. The contrapositive negates Q to the statement *there exists* $n \in \mathbb{N}$ for which $f(a) \not\equiv 0 \ mod(n)$. On the other hand, the negation of P is simply $f(a) \neq 0$. To finish the thought, the contrapositive of the theorem suggests that if we can find an n such that $f(a) \not\equiv 0$ for all $a \in \mathbb{Z}$ then it follows $f(a) \neq 0$ for all $a \in \mathbb{Z}$.

This method is not generally successful in proving the non-existence of integer zeros for polynomials over the integers. See page 45 of Jones and Jones' *Elementary Number Theory* for comments[38].

There is a large difference between ordinary arithmetic in $\mathbb{Z}$ and that of $\mathbb{Z}_n$. We already saw in Example 1.5.15 the solution set of a system of equations in $\mathbb{Z}_4$ had four distinct solutions. In the context of systems of equations over $\mathbb{Z}$ we either obtain no solutions, one solution, or infinitely many. This distinction is largely tied to the fact that some numbers in $\mathbb{Z}_n$ do not have multiplicative inverses. For example, in $\mathbb{Z}_4$ the fact that $[2][2] = [0]$ implies there cannot be $[x]$ such that $[2][x] = [1]$ since that would give us $[2][2][x] = [0][x]$ implying $[2][1] = [2] = [0]$ which is absurd. Apparently, only certain numbers in $\mathbb{Z}_n$ have multiplicative inverses. Let us characterize which numbers have inverses modulo n. Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$ we seek to solve:

$$[a][x] = [1] \quad \Rightarrow \quad ax - 1 = nk$$

[38] give me a warning via email if you want to look at this book, I might need to grab it from home

for some $k \in \mathbb{Z}$. This gives,

$$ax + nk = 1$$

If a and n have a common factor larger than 1 then we obtain a contradiction since 1 has no divisors. Thus, in the case there is a solution, we must have $gcd(a, n) = 1$. This is fortunate news since we have a nice method to calculate $gcd(a, n)$ and the criteria that a^{-1} exist in $\mathbb{Z}_n$ is simply that a is **relatively prime** or, if you prefer, **coprime**.

Example 1.5.30. *In Example 1.4.16 we found* $gcd(62, 626) = 2$. *This shows* 62 *does not have a multiplicative inverse modulo* 626. *Also, it shows* 626 *does not have a multiplicative inverse modulo* 62.

Example 1.5.31. *In Example 1.4.17 we found* $gcd(11, 240) = 1$ *and* $1 = -109(11) + 5(240)$. *From this we may read several things:*

$$[-109]^{-1} = [11] \ mod(240) \qquad \& \qquad [-109]^{-1} = [11] \ mod(5)$$

and,

$$[5]^{-1} = [240] \ mod(11) \qquad \& \qquad [5]^{-1} = [240] \ mod(109).$$

In terms of least positive residues the last statement reduces to $[5]^{-1} = [22]$. *Of course, we can check this;* $[5][22] = [110] = [1]$.

Remark 1.5.32. At this point our work on the model $\mathbb{Z}/n\mathbb{Z}$ for $\mathbb{Z}_n$ comes to an end. From this point forward, we return to the less burdensome notation

$$\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$$

as a default. However, we are open-minded, if you wish, you can define

$$\mathbb{Z}_n = \{1, 2, \ldots, n\}$$

where n serves as the additive identity of the group. Furthermore, we wish to allow calculations such as: working modulo 5 we have:

$$3(17) = 3(15 + 2) = 3(2) = 6 = 1$$

thus $17^{-1} = 2$ or $3^{-1} = 2$ etc. *Technically, if we define* $G_1, G_2, G_3 \subseteq \mathbb{Z}$ *where*

$$G_1 = \{0, 1, 2\}, \quad G_2 = \{1, 2, 3\}, \quad G_3 = \{10, 11, 12\}$$

and addition is defined modulo 3 *then* G_1, G_2 *and* G_3 *are distinct* **point sets** *and hence are different groups. However, these are all models of* $\mathbb{Z}_3$. *In fact,* G_1, G_2, G_3 *are all* **isomorphic**[39]. *Algebraists will say things like, all of the sets* G_1, G_2, G_3 *are* $\mathbb{Z}_3$. *What they mean by that is that these set are all to the inutition of a group theorist the same thing. What we define* $\mathbb{Z}_n$ *to be is largely a matter of convenience. Again, the two main models:*

(1.) $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} = \{k + n\mathbb{Z} \mid k \in \mathbb{Z}\}$ *makes* $\mathbb{Z}_n$ *a set of sets of integers, or, a set of cosets of* $\mathbb{Z}$.

(2.) $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ *where* $\mathbb{Z}_n \subset \mathbb{Z}$.

[39] we will define this carefully in due time, for Fall 2016, this is after Test 1

In either case, $\mathbb{Z}_n$ is not a subgroup of $\mathbb{Z}$, but, for slightly different reasons. In case (1.) the set of cosets is not a subset of $\mathbb{Z}$ so it fails the subset criterion for subgroup. In case (2.) while it is a subset it fails to have the same additive operation as $\mathbb{Z}$.

We now have all the tools we need to prove $U(n)$ is a group.

Theorem 1.5.33. *The set $U(n) = \{x \mid gcd(x, n) = 1\}$ with multiplication modulo n forms a group.*

Proof: We suppose $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ and $U(n) \subset \mathbb{Z}_n$ for the sake of specificity. Observe $\gcd(1, n) = 1$ thus $1 \in U(n)$. Moreover, $1x = x1 = x$ modulo n for each $x \in U(n)$ so 1 is the multiplicative identity for $U(n)$. Suppose $x, y \in U(n)$ hence by Bezout's identity there exist $a, b, c, d \in \mathbb{Z}$ for which

$$ax + bn = 1 \qquad \& \qquad cy + dn = 1$$

thus $ax = 1 - bn$ and $cy = 1 - dn$ so

$$(ax)(cy) = (1 - bn)(1 - dn) = 1 + (n - b - d)n$$

Thus $(xy)(ac) + n(b + d - n) = 1$ and as $ac, b + d - n \in \mathbb{Z}$ we see that $\gcd(xy, n) = 1$. Therefore, $xy \in U(n)$. Moreover, we also see that

$$xa + bn = 1 \quad \Rightarrow \quad \gcd(x, n) = 1$$

thus $x \in U(n)$. Furthermore, $xa + bn = 1$ modulo n yields $xa \equiv 1$ which means $x = a^{-1}$. In summary, we have shown $U(n)$ has a binary operation with identity and inverses. I suppose to complete the argument we ought to have shown that multiplication in $\mathbb{Z}_n$ is associative. The proof is similar to that which was given for Proposition 1.5.13 and we leave it to the reader. □

Matrices with entries in $\mathbb{Z}_n$ are multiplied and added in the usual fashion. In particular,

$$(A + B)_{ij} = A_{ij} + B_{ij}, \quad (cA)_{ij} = cA_{ij}, \quad (XY)_{ij} = \sum_{k=1}^{r} X_{ir}Y_{rj}$$

where $A, B \in \mathbb{Z}_n^{p\times q}$, $X \in \mathbb{Z}_n^{p\times r}$ and $Y \in \mathbb{Z}_n^{r\times q}$. We can show $I_{ij} = \delta_{ij}$ has $XI = IX = X$ for any matrix. Naturally, the addition and multiplications above are all done modulo n. This has some curious side-effects:

$$\underbrace{A + A + \cdots + A}_{n-summands} = nA = 0$$

A square $M \in \mathbb{Z}_n^{p\times p}$ is invertible only if $\det(M) \in U(n)$. To prove this we could go through all the usual linear algebra simply replacing regular addition and multiplication with the modular equivalent. In particular, we can show the classical adjoint of M satisfies

$$M\text{adj}(M) = \det(M)I$$

If $\det(M) \in U(n)$ then there exists $\det(M)^{-1} \in U(n)$ for which $\det(M)^{-1} \cdot \det(M) = 1$. Multiplying the classical adjoint equation we derive

$$M(\det(M)^{-1} \cdot \text{adj}(M)) = \det(M)^{-1} \cdot \det(M)I = I.$$

Thus,

$$\boxed{M^{-1} = \det(M)^{-1} \cdot \text{adj}(M).}$$

Calculation of the inverse in 3×3 or larger cases requires some calculation, but, for the 2×2 case we have a simple formula: if $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{Z}_n^{2\times 2}$ and $ad - bc \in U(n)$ then

$$\boxed{\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = (ad-bc)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}$$

In the case $n = p$ a prime, $U(p) = \mathbb{Z}_p^\times$ and the inverse of a matrix M over $\mathbb{Z}_p$ exists whenever $\det(M) \neq 0$. In fact, we can define the general linear group over matrices even when the entries are not taken from a field.

Definition 1.5.34. *The general linear group of $p \times p$ matrices over $\mathbb{Z}_n$ is defined by:*

$$GL(p, \mathbb{Z}_n) = \{A \in \mathbb{Z}_n^{p\times p} \mid det(A) \in U(n)\}.$$

Moreover, $GL(p, \mathbb{Z}) = \{A \in \mathbb{Z}^{p\times p} \mid det(A) = \pm 1\}$.

I will forego proof that the general linear groups are indeed groups at the moment. In fact, we can define the general linear group for matrices built over any ring in a similar fashion. These suffice for our current purposes.

Example 1.5.35. *Let $M = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$ modulo 10. Note $det(M) = 1(1) - 2(3) = -5 = 5 \notin U(10)$. I claim we can find $X \neq 0$ for which $MX = 0$. Let's calculate:*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

we want to solve,

$$a + 3b = 0, \quad 2a + b = 0, \quad c + 3d = 0, \quad 2c + d = 0$$

thus $d = -2c$ and $c + 3(-2c) = -5c = 0$ we choose $c = 2$ hence $d = -2c = -4 = 6$. Continuing, $b = -2a$ this $a + 3(-2a) = -5a = 0$ so for fun we choose $a = 4$ and find $b = -2a = -8 = 2$. Finally, check our work, does $MX = 0$ as we wished?

$$\begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 10 \\ 20 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Do you understand why the existence of X for which $MX = 0$ forbids the possibility that M^{-1} exists? Suppose M^{-1} existed for the sake of discussion, notice:

$$M^{-1}MX = M^{-1}0 \quad \Rightarrow \quad X = 0$$

but, $X \neq 0$ hence no multiplicative inverse for M can exist.

Example 1.5.36. *Let $M = \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}$ in $\mathbb{Z}_6^{2\times 2}$. Let us find the order of M in the additive sense.*

$$2M = \begin{bmatrix} 0 & 4 \\ 0 & 2 \end{bmatrix}, \quad 3M = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}, \quad 4M = \begin{bmatrix} 0 & 2 \\ 0 & 4 \end{bmatrix}, \quad 5M = \begin{bmatrix} 0 & 4 \\ 3 & 2 \end{bmatrix}, \quad 6M = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Example 1.5.37. *Let* $X = \begin{bmatrix} 2 & 4 \\ 6 & 0 \end{bmatrix}$ *in* $\mathbb{Z}_8^{2\times 2}$. *Let us find the order of* X *in the additive sense.*

$$2X = \begin{bmatrix} 4 & 0 \\ 4 & 0 \end{bmatrix}, \quad 3X = \begin{bmatrix} 6 & 4 \\ 2 & 0 \end{bmatrix}, \quad 4X = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus the order of X *is* 4.

There are additive groups of matrices and multiplicative groups of matrices. Let us consider a pair of examples in the multiplicative realm.

Example 1.5.38. *Let* $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ *in* $\mathbb{Z}_4^{2\times 2}$. *Notice,* $det(A) = 4 - 1 = -1 = 3$ *and* $3^{-1} = 3$ *as* $3(3) = 9 = 1$ *modulo* 4. *Thus,*

$$A^{-1} = 3\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ -3 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Indeed, you can check, $AA = I$ *thus* $A = A^{-1}$. *Moreover, this shows* A *has order* 2.

Example 1.5.39. *Let* $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ *in* $\mathbb{Z}_5^{2\times 2}$. *Notice,* $det(A) = 4 - 1 = 3$ *and* $3^{-1} = 2$ *as* $2(3) = 6 = 1$ *modulo* 5. *Thus,*

$$A^{-1} = 2\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$$

Next, we determine the order of A *by direct calculation:*

$$A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$$

as a quick check on my calculation, note $det(A^2) = -16 = 4$ *and* $[det(A)]^2 = 3^2 = 9 = 4$. *Continuing,*

$$A^3 = AA^2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}.$$

Observe $A^3 = A^{-1}$ *hence* $A^4 = I$ *and find* $|A| = 4$.

I recommend working on Chapter 0 Gallian problems such as:

$$\#3, 9, 14, 15, 28, 46, 47, 48,$$

I recommend working on Chapter 2 Gallian problems such as:

$$\#5, 8,$$

also, Chapter 3 Gallian problems such as:

$$\#13, 37.$$

these should help bring the Group concept to life. I do not collect these, but, I will keep them in mind as I construct tests.

Problems for Lecture 5: (these are collected at the beginning of Lecture 7)

Problem 17: Gallian Chapter 0 number 40 (ISBN code, based on 39 in part)

Problem 18: Workin' mod 14.

(a) Find the additive inverse and order of each element in $\mathbb{Z}_{14}$.

(b) Find the multiplicative inverse or indicate "DNE" (does not exist) for each element in $\mathbb{Z}_{14}$. If the multiplicative inverse exists, that element belongs to $U(14)$. In this case, find the order of that element (in $U(14)$).

(c) Compute $5^{-2} \cdot (4-10) \cdot 13^{999} + 11 \pmod{14}$.

(d) Compute A^{-1} given $A = \begin{bmatrix} 1 & 5 \\ 4 & 9 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Z}_{14})$.

Problem 19: The Matrix problem

(a) Compute $A^{-1}B^2$ where $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Z}_9)$

(b) Find the cyclic subgroup generated by A. What is the order of A?

Problem 20: A function problem

(a) Let $f : \mathbb{Z} \to \mathbb{Z}$ be defined by $f(x) = 2x^2 - 3$.

i. Show f is not 1-1.

ii. Show f is not onto.

iii. Let $A = \{-1, 0, 1, 2, 3\}$. Find $f(A) = \{f(x) \mid x \in A\}$ (the image of the set A under the map f).

iv. Let $A = \{-1, 0, 1, 2, 3\}$. Find $f^{-1}(A) = \{x \in \mathbb{Z} \mid f(x) \in A\}$ (the inverse image of A).

(b) Let $g : X \to Y$. Prove that g is onto if and only if $g^{-1}(B) \neq \phi$ (the inverse image of B is non-empty) for all non-empty subsets of Y: $\phi \neq B \subset Y$.

Recall that for $A \subseteq X$ and $B \subseteq Y$ ...

$$f(A) = \{f(x) \mid x \in A\} \subseteq Y \qquad \text{and} \qquad f^{-1}(B) = \{x \in X \mid f(x) \in B\} \subseteq X$$

1.6 Lecture 6: cyclic groups

We studied the subgroup $\langle a \rangle$ in Lecture 3. We now continue that discussion. This and the next lecture are set aside to unravel the basic theory of cyclic groups. They are largely based on Chapter 4 of Gallian. What is a cyclic group? It[40]

Definition 1.6.1. *If a group $G = \langle a \rangle$ for some $a \in G$ then we say G is a* **cyclic group**. *Moreover, any element b for which $\langle b \rangle = G$ is called a* **generator** *of G.*

To prove G is not cyclic we can demonstrate that no element of G generates all of G.

Example 1.6.2. *Consider $D_3 = \{1, r, r^2, f, rf, r^2f\}$ where $frf = r^{-1}$ and $r^3 = 1$ and $f^2 = 1$. We also may derive that $fr^{-1}f = r$ and $fr = r^{-1}f$ and $fr^{-1} = rf$. With these relations in mind it is not much trouble to calculate the subgroup of D_3 generated by various elements.*

$$\langle r \rangle = \{1, r, r^2\}, \qquad \& \qquad \langle f \rangle = \{1, f\}, \qquad \& \qquad \langle rf \rangle = \{1, rf\}$$

as $(rf)(rf) = r(frf) = rr^{-1} = 1$. Consider r^2f, note $r^2r = 1$ hence $r^2 = r^{-1}$,

$$(r^2f)^2 = (r^{-1}f)(r^{-1}f) = r^{-1}(fr^{-1}f) = r^{-1}r = 1 \quad \Rightarrow \quad \langle r^2f \rangle = \{1, r^2f\}$$

Notice, every element of D_3 is covered by the 5 cyclic subgroups we have explicitly given. Do you understand why this implies D_3 is not a cyclic group? If not, then what is the remaining possibility? Notice, $\langle r^2 \rangle = \{1, r^2, r\}$ as $(r^2)^2 = r^4 = r$. Visibly, D_3 is not cyclic.

The method used in the example above is what I would call **brute force**. We shall learn a few labor saving devices as we continue our study. That said, let me just give a few examples to get used to the idea of a generator.

Example 1.6.3. *A nice infinite group example is found in $(\mathbb{Z}, +)$. Observe,*

$$\langle 1 \rangle = \{n(1) \mid n \in \mathbb{Z}\} = \mathbb{Z}.$$

likewise, $\langle -1 \rangle = \{n(-1) \mid n \in \mathbb{Z}\} = \mathbb{Z}$. Thus $\mathbb{Z}$ is generated by both 1 and -1.

Example 1.6.4. *$\mathbb{Z}_4 = \{0, 1, 2, 3\}$ has $\langle 1 \rangle = \{0, 1, 2, 3\}$ and $\langle 2 \rangle = \{0, 2\}$ and $\langle 3 \rangle = \{0, 3, 2, 1\}$ hence 1 and $3 = -1$ serve as generators for $\mathbb{Z}_4$.*

Example 1.6.5. *$\mathbb{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$ has $6 = -1$ and for much the same reasons as the past two examples, $\langle 1 \rangle = \langle -1 \rangle = \mathbb{Z}_7$. However, we also have,*

$$\langle 2 \rangle = \{0, 2, 4, 6, 1, 3, 5\} = \mathbb{Z}_7 \quad \& \quad \langle 3 \rangle = \{0, 3, 6, 2, 5, 1, 4\} = \mathbb{Z}_7$$

in fact, $\langle 4 \rangle = \langle 5 \rangle = \langle 6 \rangle = \mathbb{Z}_7$. Every nonzero element in $\mathbb{Z}_7$ serves as a generator of the group.

Example 1.6.6. *$\mathbb{Z}_{10}$ has $\langle 1 \rangle = \langle -1 \rangle = \mathbb{Z}_{10}$. However,*

$$\langle 2 \rangle = \{0, 2, 4, 6, 8\} \quad \& \quad \langle 4 \rangle = \{0, 4, 8, 2, 6\} \quad \& \quad \langle 6 \rangle = \{0, 6, 2, 8, 4\} \quad \& \quad \langle 8 \rangle = \{0, 8, 6, 4, 2\}$$

and $\langle 5 \rangle = \{0, 5\rangle$. In contrast, $1, 3, 7, 9$ all serve as generators of $\mathbb{Z}_{10}$.

A clear pattern begins to immerge. The generators of $\mathbb{Z}_n$ are found in $U(n)$. This is not a proof, this is merely a conjecture at this point! (see Corollary 3 on page 77 of Gallian)

[40] is time for...

Example 1.6.7. *$U(10) = \{1, 3, 7, 9\}$ has $\langle a \rangle = \{1, a, a^2, \dots\}$*

$$\langle 1 \rangle = \{1\}, \quad \& \quad \langle 3 \rangle = \{1, 3, 9, 7\}, \quad \& \quad \langle 7 \rangle = \{1, 7, 9, 3\}, \quad \& \quad \langle 9 \rangle = \{1, 9\}$$

Apparently, 3 and 7 serve as generators for $U(10)$ however, 9 does not. How are these numbers different as they relate to 10?

I ask the question in the example above not hoping you find an answer. It is not that obvious which integers serve as generators for the group of units. In fact, not all $U(n)$ are cyclic.

Example 1.6.8. *$U(8) = \{1, 3, 5, 7\}$ we have $\langle 3 \rangle = \{1, 3\}$ as $3^2 = 9 = 1$. Likewise, $5^2 = 25 = 1$ and $7^2 = 49 = 1$ hence $\langle 5 \rangle = \{1, 5\}$ and $\langle 7 \rangle = \{1, 7\}$. By brute force we have shown $U(8)$ is not cyclic.*

Example 1.6.9. *$U(20) = \{1, 3, 7, 9, 11, 13, 17, 19\}$ is not cyclic. I invite the reader to verify this through explicit computation. Or, you can watch me do it here see minute 35 onward. I show every element of $U(20)$ has order $1, 2$ or 4.*

I think we've seen enough examples for now. Let us begin our work on the theory of cyclic groups. Before that, I must state a theorem:

Theorem 1.6.10. *Let G be a group and $g \in G$ if $m, n \in \mathbb{Z}$ then*

(1.) *$g^m g^n = g^{m+n}$,*

(2.) *$(g^m)^n = g^{mn}$.*

Proof: I gave a proof of (1.) in the solution to Lecture 3 homework. I offer a partial proof of (2.) here, fix $m \in \mathbb{Z}$ and note,

$$(g^m)^0 = e = g^{m(0)}.$$

Inductively suppose $(g^m)^n = g^{mn}$ for some $n \in \mathbb{N}$. Consider,

$$(g^m)^{n+1} = (g^m)^n g^m = g^{mn} g^m$$

by the induction hypothesis. Next, use (1.) to add exponents,

$$(g^m)^{n+1} = g^{mn+m} = g^{m(n+1)}$$

thus $(g^m)^n = g^{mn}$ for all $n \in \mathbb{N}$. Next, let $k \in \mathbb{N}$ and set $n = -k$,

$$(g^m)^n = (g^m)^{-k} = ((g^m)^{-1})^k$$

where we used the definition of negative power in the last step. Note, by applying (1.) once more we obtain $g^m g^{-m} = g^{m-m} = g^0 = e$ thus $(g^m)^{-1} = g^{-m}$. Hence,

$$(g^m)^n = ((g^m)^{-1})^k = (g^{-m})^k = g^{(-m)k} = g^{m(-k)} = g^{mn}. \qquad \square$$

This could have been done much earlier. My apologies. In any event, we are free to use the laws of exponents above going forward. Naturally, I might ask you to prove a law of exponents on an exam, but, otherwise, they are for us all to enjoy. This matters a fair amount in this section since we face many calculations involving products of powers.

Theorem 1.6.11. *Let G be a group and $a \in G$. If $|a| = \infty$ then $i \neq j$ implies $a^i \neq a^j$. If $|a| = n \in \mathbb{N}$ then $\langle a \rangle = \{1, a, \ldots, a^{n-1}\}$ where $a^i = a^j$ if and only if $n \mid (i-j)$.*

Proof: begin with the case $|a| = \infty$. Suppose $i \neq j$ for some $i, j \in \mathbb{Z}$ and suppose $a^i = a^j$. Without loss of generality we may suppose $i > j$. Multiply by a^{-j} and obtain, $a^i a^{-j} = a^j a^{-j}$ hence $a^{i-j} = e$ thus $|a| \leq i - j$ which contradicts the infinite order of a. Therefore, we find $a^i \neq a^j$.

Suppose $|a| = n$ for some $n \in \mathbb{N}$. Suppose $a^i = a^j$ for $1 \leq j < i < n$. Multiply by a^{-j} as to obtain $a^{i-j} = e$. Notice, $j < i < n$ implies $i - j < n - j < n$ hence $a^{i-j} = e$ contradicts $|a| = n$. Therefore, we find $a^i \neq a^j$ for $1 \leq j < i < n$.

Suppose $k \in \mathbb{Z}$ then by the division algorithm there exists $q, r \in \mathbb{Z}$ with $k = qn + r$ where $0 \leq r < n$. Observe,

$$a^k = a^{qn+r} = a^{qn} a^r = (a^n)^q a^r = e^q a^r = a^r.$$

Notice, $0 \leq r < n$ allows $r = 0, 1, \ldots, n-1$ and thus the cyclic subgroup generated by a is simply:

$$\langle a \rangle = \{1, a, a^2, \ldots, a^{n-1}\}$$

Suppose $a^i = a^j$ for some $i, j \in \mathbb{Z}$. Multiply by a^{-j} as to obtain $a^{i-j} = e$. Apply the division algorithm to obtain $q, r \in \mathbb{Z}$ with $0 \leq r < n$ and $i - j = qn + r$. Hence,

$$e = a^{i-j} = a^{qn+r} = (a^n)^q a^r = e^q a^r = a^r \quad \Rightarrow \quad a^r = e.$$

since $r < n$ we must conclude $r = 0$ as say otherwise contradicts $|a| = n$. Therefore, $i - j = qn$ and we conclude $n \mid i - j$. Conversely, if $n \mid i - j$ then $i - j = qn$ for some $q \in \mathbb{Z}$ hence $a^{i-j} = a^{qn} = (a^n)^q = e^q = 1$ thus, multiplying by a^j on the equation we derive $a^i = a^j$. □

Remark 1.6.12. The proof I give here is fairly similar to that given for Theorem 4.1 in Gallian. The key is the division algorithm appropriately applied.

Corollary 1.6.13. *If G be a group and $a \in G$ then $|a| = |\langle a \rangle|$.*

Proof: if G is a group and $a \in G$ and $|a| = \infty$ then Theorem 1.6.11 shows $a^i \neq a^j$ for all $i, j \in \mathbb{Z}$ hence $\langle a \rangle = \{1, a, a^{-1}, a^2, a^{-2}, \ldots\}$ is not a finite set. If $|a| = n$ then Theorem 1.6.11 shows $\langle a \rangle = \{1, a, \ldots, a^{n-1}\}$ with listed elements distinct. Thus, by counting, $|\langle a \rangle| = n = |a|$. □.

Notice, all the hard work for the Corollary above really is done by Theorem 1.6.11. As a point of etymology the terms THEOREM and COROLLARY are quite old. According to Rotman page xii. in his *A First Course in Abstract Algebra with Applications*, 3rd edition,

1. the term **theorem** is from a Greek word meaning *to watch* or *to contemplate*. In other words, the term theorem indicates something worthy of contemplation.

2. the term **corollary** is from a Latin word meaning *to flower*, possibly because flowers were a common gift in ancient Rome, so, the corollary is a gift from the theorem.

Rotman also explains *mathematics* is actually a classical Greek term which means *to learn*.

Corollary 1.6.14. *If G be a group and $a \in G$ with $|a| = n \in \mathbb{N}$. If $a^k = e$ then $n \mid k$.*

Proof: if $a \in G$ with $|a| = n$ and $a^k = e$. Then $a^k = a^0$ hence $n \mid (k - 0)$ by Theorem 1.6.11. □

Discussion on factoring: How does this Corollary apply? Suppose $|a| = 3$ then the only way $a^k = e$ is if k is a multiple of 3. Recall, $n \mid k$ implies $k = nm$ for some $m \in \mathbb{Z}$. Again, $a \mid b$ only if b is a **multiple** of a. Working backwards, if $a^{50} = e$ then the order of a must be some factor of 50. It could be, $a^2 = e$ or $a^5 = e$ or $a^{10} = e$ or $a^{25} = e$. What can you say if we compute $a^{26} \neq e$? What must the order of a be if $a^{26} \neq e$ and $a^{50} = e$?

Discussion on similarity to $\mathbb{Z}_n$: If $a^n = e$ and $i = j + n$ then $a^i = a^{j+n} = a^j a^n = a^j$. Notice, if $|a| = n$ then Theorem 1.6.11 provides that $a^i = a^j$ if and only if i is congruent to j modulo n. Of course, in the case $G = \mathbb{Z}_n$ and $\langle a \rangle = \{0, a, 2a, \ldots, na\}$ we have $|a| = n$ and $ia = ja$ only if $i \equiv j \bmod n$. The addition in $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ and exponent laws for multiplication in $\langle a \rangle = \{1, a, a^2, \ldots, a^{n-1}\}$ are nicely connected. We state the theorems and corollaries in this section in multiplicative notation, however, there are additive restatements of all our results. In fact, anything fact we prove for a generic cyclic group we also know for $\mathbb{Z}_n$. Why? Because $\mathbb{Z}_n$ is the quintessential finite cyclic group. Figure 4.1 on page 76 of Gallian is helpful.

Theorem 1.6.15. *If G is a group and $a \in G$ with $|a| = n \in \mathbb{N}$ and $k \in \mathbb{N}$ then*

$$\langle a^k \rangle = \langle a^{gcd(n,k)} \rangle \qquad \text{and} \qquad |a^k| = \frac{n}{gcd(n,k)}.$$

Proof: assume G is a group and $a \in G$ with $|a| = n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let $d = \gcd(n, k)$. Observe, as d is a divisor of k, there exists $z \in \mathbb{Z}$ for which $k = zd$.

Suppose $x \in \langle a^k \rangle$ then there exists $y \in \mathbb{Z}$ for which $x = (a^k)^y$. Observe $x = a^{ky} = a^{zdy} = (a^d)^{zy} \in \langle a^{\gcd(n,k)} \rangle$. Thus $\langle a^k \rangle \subseteq \langle a^{\gcd(n,k)} \rangle$.

Let $w \in \langle a^{\gcd(n,k)} \rangle$ hence $w = (a^d)^u$ for some $u \in \mathbb{Z}$. By Bezout's Identity[41] there exist $p, q \in \mathbb{Z}$ for which $pn + qk = d$. Thus, noting $a^n = e$ is given for the fifth equality,

$$w = (a^d)^u = (a^{pn+qk})^u = (a^{pn} a^{qk})^u = ((a^n)^p a^{qk})^u = (e a^{qk})^u = a^{qku} = (a^k)^{qu} \in \langle a^k \rangle.$$

Thus $\langle a^{\gcd(n,k)} \rangle \subseteq \langle a^k \rangle$ and we conclude $\langle a^{\gcd(n,k)} \rangle = \langle a^k \rangle$

It remains to show $|a^k| = \frac{n}{d}$. Observe $(a^d)^{n/d} = a^n = e$ hence $|a^d| \leq n/d$. If $0 < i < n/d$ then $di < n$ and hence $(a^d)^i = a^{di} \neq e$ since $|a| = n$. Therefore, $|a^d| = n/d$. Recall Corollary 1.6.13 assures us the order of an element is the same as the order of the cyclic subgroup it generates; $|a^d| = |\langle a^{\gcd(n,k)} \rangle|$ and $|a^k| = |\langle a^k \rangle|$. We already proved $\langle a^{\gcd(n,k)} \rangle = \langle a^k \rangle$ thus $|a^k| = n/d$. □

Corollary 1.6.16. *If $|a| = n$. Then $\langle a^i \rangle = \langle a^j \rangle$ if and only if $gcd(n, i) = gcd(n, j)$.*

Proof: let G be a group and $a \in G$ with $|a| = n \in \mathbb{N}$. By Theorem 1.6.15 we have

$$\langle a^i \rangle = \langle a^{\gcd(n,i)} \rangle \qquad \& \qquad \langle a^j \rangle = \langle a^{\gcd(n,j)} \rangle$$

hence $\gcd(n, i) = \gcd(n, j)$ yields immediately the equality $\langle a^i \rangle = \langle a^j \rangle$. Conversely, suppose $\langle a^i \rangle = \langle a^j \rangle$. Hence $|a^i| = |a^j|$ and by Theorem 1.6.15 we find $|a^i| = \frac{n}{\gcd(n,i)}$ and $|a^j| = \frac{n}{\gcd(n,j)}$ thus

$$\frac{n}{\gcd(n,i)} = \frac{n}{\gcd(n,j)}$$

[41] Gallian refers to this as the GCD theorem

and we derive $\gcd(n,i) = \gcd(n,j)$. □

And now some corollaries to the Corollary above.

Corollary 1.6.17. *If $G = \langle a \rangle$ is a cyclic group of order n then $G = \langle a^k \rangle$ if and only if $gcd(n,k) = 1$.*

Proof: assume $G = \langle a \rangle$. Observe, by Corollary 1.6.16, $\langle a^k \rangle = \langle a^1 \rangle$ iff $\gcd(n,k) = \gcd(n,1)$. Clearly $\gcd(n,1) = 1$ thus $\gcd(n,k) = 1$. □

Corollary 1.6.18. *If $k \in \mathbb{Z}_n$ then k is a generator of $\mathbb{Z}_n$ iff $gcd(n,k) = 1$.*

Proof: almost the same as Corollary 1.6.17. Observe $\mathbb{Z}_n = \langle 1 \rangle$. By Corollary 1.6.16, $\langle 1 \rangle = \langle k \rangle$ iff $\gcd(n,k) = \gcd(n,1)$ hence $\gcd(n,k) = 1$. □

I you return to all the examples we've done in the past few lectures you'll see this Corollary at work.

Example 1.6.19. *In the dihedral group D_n we may study the subgroup of rotations $H = \{1, r, r^2, \ldots, r^{n-1}\}$ where $r^n = 1$ and $|r| = n$ (this indicates the list defining H is comprised of distinct elements). Notice,*

$$H = \langle r \rangle$$

thus $H \leq G$ by Theorem 1.2.22. Furthermore, the possible generators for H are simply r^k where $gcd(k,n) = 1$. For example, in D_{16} denoting $H = \langle r \rangle$ where $r^{16} = 1$ then $r^3, r^5, r^7, r^9, r^{11}, r^{13}, r^{15}$ are other generators for H.

A more powerful example is given in Gallian on page 78. I give a different example here (less impressive, I do $U(14)$ as compared to his $U(20)$ example)

Example 1.6.20. *Consider $U(14) = \{1, 3, 5, 9, 11, 13\}$ we can show $\langle 3 \rangle = U(14)$ since, modulo 14 we calculate (I maintain order across the equality)*

$$1,\ 3,\ 3^2,\ 3^3,\ 3^4,\ 3^5 = 1,\ 3,\ 9,\ 13,\ 11,\ 5.$$

The order of $U(14)$ is $n = 6$. Notice, only 1 and 5 are relatively prime to 6 hence the only other generator of $U(14)$ is given by $3^5 = 5$. I am using Corollary 1.6.17.

I recommend working on Chapter 4 Gallian problems such as:

$$\#1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15$$

I do not collect these, but, I will keep them in mind as I construct tests. There might be further problems which are appropriate given this lecture. That said, by the end of Lecture 7 we should be ready to try all the problems in Chapter 4 of Gallian.

Problems for Lecture 6: (these are collected at the beginning of Lecture 8)

Problem 21: Exercise 13 of Chapter 4 in Gallian.

Problem 22: Orders of elements and number of such elements.

(a) Make a table which lists the possible orders of elements of $\mathbb{Z}_{294}$. List of the number such elements in the second row. [I'll get you started: There is 1 element of order 1] How many generators does $\mathbb{Z}_{294}$ have?

(b) Repeat part (a) for D_{294}.

(c) How many elements of order 8 are there in $\mathbb{Z}_{1440000}$? What are they?

(d) How many elements of order 7 are there in $\mathbb{Z}_{1440000}$?

Problem 23: Let $g \in G$ (for some group G). Suppose $|g| = 120$. List the *distinct* elements of $\langle g^{100} \rangle$. Is $g^{30} \in \langle g^{100} \rangle$?

Problem 24: Let $g, x \in G$ (for some group G).

i. Show that $|x| = |gxg^{-1}|$ (i.e. conjugates have the same order).

ii. Prove or give a counterexample: $\langle x \rangle = \langle gxg^{-1} \rangle$.

1.7 Lecture 7: classification of cyclic subgroups

We begin with several examples.

Example 1.7.1. *If $G = \langle a \rangle$ where $a^5 = e$ then the trivial subgroup $\langle e \rangle = \{e\}$. However, there are no nontrivial proper subgroups. Notice:*

$$\begin{aligned} \langle a \rangle &= \{e, a, a^2, a^3, a^4\} \\ \langle a^2 \rangle &= \{e, a^2, a^4, a\} \\ \langle a^3 \rangle &= \{e, a^3, a, a^4, a^2\} \\ \langle a^4 \rangle &= \{e, a^4, a^3, a^2, a\} \end{aligned}$$

Thus every non-identity element serves as a generator for G.

The behaviour above is typical of groups of prime order. For a cyclic group, if the order is the square of a prime then we get just one proper subgroup.

Example 1.7.2. *Consider $G = \mathbb{Z}_{25}$. As always, the identity generates a subgroup containing itself alone; $\langle 0 \rangle = \{0\}$. Also, every element of G except for $0, 5, 10, 15, 20$ generates $\mathbb{Z}_{25}$ since all other numbers in $\mathbb{Z}_{25}$ are relatively prime to 25. We find the subgroup*

$$\langle 5 \rangle = \{0, 5, 10, 15, 20\}$$

note $2, 3$ and 4 are relatively prime to the order of $|\langle 5 \rangle| = 5$ and so $10 = 2(5)$, $15 = 3(5)$ and $20 = 4(5)$ are also generators for $\langle 5 \rangle$. Here we see the additive version of Corollary 1.6.17 at play.

Cyclic groups of composite order can have many subgroups.

Example 1.7.3. *Consider $G = \langle a \rangle$ where $a^8 = e$. Observe:*

$$\begin{aligned} \langle a^2 \rangle &= \{e, a^2, a^4, a^6\} = \langle a^6 \rangle \\ \langle a^4 \rangle &= \{e, a^4\} \\ \langle e \rangle &= \{e\} \end{aligned}$$

whereas a, a^3, a^5, a^7 serve as generators for G itself $\langle a \rangle = \langle a^3 \rangle = \langle a^5 \rangle = \langle a^7 \rangle$. For example,

$$\langle a^3 \rangle = \{e, a^3, (a^3)^2, (a^3)^3, (a^3)^4, (a^3)^5, (a^3)^6, (a^3)^7\} = \{e, a^3, a^6, a, a^4, a^7, a^2, a^5\}$$

Example 1.7.4. *Consider $G = \mathbb{Z}_{20}$. In this case we have a few more subgroups to consider. Corollary 1.6.18 provides that $U(20) = \{1, 3, 7, 9, 11, 13, 17, 19\}$ contains generators for $\mathbb{Z}_{20}$. There are several proper subgroups,*

$$\begin{aligned} \langle 2 \rangle &= \{0, 2, 4, 6, 8, 10, 12, 14, 16, 18\} \\ \langle 4 \rangle &= \{0, 4, 8, 12, 16\} \\ \langle 5 \rangle &= \{0, 5, 10, 15\} \\ \langle 10 \rangle &= \{0, 10\} \\ \langle 0 \rangle &= \{0\}. \end{aligned}$$

Incidentally, using the additive version of Corollary 1.6.17 we find from $|\langle 2 \rangle| = 10$ and the fact that $3, 7$ and 9 are relatively prime to 10 that $3(2) = 6$, $7(2) = 14$ and $9(2) = 18$ are generators of $\langle 2 \rangle$.

The patterns in the examples above turn out to be general for cyclic groups.

(1.) Every subgroup of a cyclic group is cyclic.

(2.) The order of a subgroup must divide the order of the group.

(3.) For each divisor of the order of the group there just one subgroup with that order

We later learn that (2.) holds for finite groups in general whereas (3.) is not generally true. Gallian calls the following the *Fundamental Theorem of Cyclic Groups*.

Theorem 1.7.5. *Let $G = \langle a \rangle$ be a cyclic group.*

(1.) *If $H \leq G$ then H is cyclic.*

(2.) *If $|G| = n$ and $H \leq G$ with $|H| = k$ then $k \mid n$.*

(3.) *If $|G| = n$ and $k \in \mathbb{N}$ with $k \mid n$ then $H = \langle a^{n/k} \rangle$ is the unique subgroup of order n/k in G.*

Proof: **(1.)** suppose $G = \langle a \rangle = \{a^k \mid k \in \mathbb{Z}\}$ and $H \leq G$. Notice $H = \{e\}$ is cyclic thus assume $H \neq \{e\}$ in what follows. Consider, if $a^k \in H$ then $a^{-k} \in H$ as $a^k a^{-k} = a^0 = e$. Let $\Lambda = \{n \in \mathbb{N} \mid a^n \in H\}$ and note that Λ has a smallest element by the well-ordering-principle. Let $t \in \Lambda$ be the smallest element of Λ. We suspect $H = \langle a^t \rangle$.

Since $a^t \in H$ it follows $(a^t)^s \in H$ for $s \in \mathbb{Z}$ by closure of the group operations of H. Thus, $\langle a^t \rangle \subseteq H$. Suppose $a^k \in H$. By the division algorithm, there exist q, r for which $k = qt + r$ where $0 \leq r < t$. Observe:

$$a^k = a^{qt+r} = a^{qt}a^r \quad \Rightarrow \quad a^r = a^{-qt}a^k.$$

Note $a^t \in H$ implies $a^{-qt} = ((a^t)^{-1})^q \in H$. Thus $a^{-qt}, a^k \in H$ and $a^r = a^{-qt}a^k \in H$. Thus, $r = 0$ as t is the smallest element of Λ. Therefore, $k = qt$ and we find $a^k = (a^t)^q \in \langle a^t \rangle$ thus $H \subseteq \langle a^t \rangle$. We conclude $H = \langle a^t \rangle$ which shows H is cyclic with generator a^t.

(2.) Suppose $G = \langle a \rangle$ and $|G| = n$. Let $H \leq G$ and $|H| = k$. Following the proof of (1.) we know there exists $a^t \in G$ for which $H = \langle a^t \rangle$. Notice, $(a^t)^n = (a^n)^t = e^t = e$. Observe a^t has $(a^t)^n = e$ in the group H thus Corollary 1.6.14 provides $|a^t| \mid n$ which is to say $k \mid n$.

(3.) Suppose $|G| = n$ and k is a positive divisor of n. Theorem 1.6.15 provides that $|\langle a^{n/k} \rangle| = \frac{n}{\gcd(n, n/k)} = \frac{n}{n/k} = k$ hence $\langle a^{n/k} \rangle$ is a subgroup of order k. Suppose $H \leq G$ is another subgroup of order k. By the proof of (1.) we know there exists $a^t \in H$ with $\langle a^t \rangle = H$ and t is a divisor of n. Consider, $t = \gcd(t, n)$ and by Theorem 1.6.15

$$k = |a^t| = |a^{\gcd(t,n)}| = \frac{n}{\gcd(t, n)} = \frac{n}{t}.$$

Therefore, $t = \frac{n}{k}$ and we conclude the unique subgroup of order k is precisely $H = \langle a^{n/k} \rangle$. □

The order of the group of units in $\mathbb{Z}_n$ is given by the *Euler ϕ function*:

Definition 1.7.6. *The* **Euler ϕ function** *is defined by $\phi(n) = |U(n)|$ for each $n \in \mathbb{N}$.*

In other words, $\phi(n)$ is the number of relative prime positive integers to n. For example,

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\phi(n)$	1	1	2	2	4	2	6	4	6	4	10	4	12	6	8	8	16	6	18	8

Notice Gallian's Exercises 38,39 and 40 of Chapter 3 indicate certain formulas may hold for the Euler-phi function. For example, $\phi(20) = \phi(4 \cdot 5) = \phi(4) \cdot \phi(5)$ whereas $\phi(8) = \phi(2 \cdot 4) \neq \phi(2) \cdot \phi(4)$.

Perhaps you can determine some relations for the Euler ϕ function from the table on the past page. Finding efficient methods to calculate $\phi(n)$ for larger n is important as it allows us to determine the number of elements of a given order in a cyclic group.

Theorem 1.7.7. *Let $G = \langle a \rangle$ be a cyclic group of order n. If d is a positive divisor of n then $\phi(d)$ is the number of elements of order d in G*

Proof: by Theorem 1.7.5 if G is cyclic with order n and d is a positive divisor of n then there is a unique subgroup H of order d in G. Moreover, $H = \langle a \rangle$ for some $a \in G$ and $|a| = d$. Recall, Corollary 1.6.17 told us $\langle a \rangle = \langle a^k \rangle$ only if $\gcd(k, d) = 1$. The number of distinct choices for k is precisely $\phi(d)$. Thus the number of elements of order d is precisely $\phi(d)$. □
Let's apply this Theorem.

Example 1.7.8. *If G is a cyclic group of order $n = 19k$ for some $k \in \mathbb{N}$ then $19 \mid n$ and we find there are $\phi(19) = 18$ elements of order 19 in G. This is true for $\mathbb{Z}_{19}$ or $\mathbb{Z}_{19,000,000}$.*

It is important to notice the qualifier **cyclic** as it appears in most of the results in this section. Up to isomorphism[42] all cyclic groups are just $\mathbb{Z}_n$ so the structure is fairly simple to decipher[43].

Corollary 1.7.9. *In a finite group the number of elements of order d is divisible by $\phi(d)$.*

Proof: If G is a finite group and G has no elements of order d then $\phi(d) \mid 0$. Otherwise, suppose d has an element of order d, say $a \in G$. Observe $\langle a \rangle$ has $\phi(d)$ elements of order d since $|\langle a \rangle| = d$ and Theorem 1.7.7 applies. Next, suppose $b \in G$ also has order d but $b \notin \langle a \rangle$. Once again, we argue $\langle b \rangle$ has $\phi(d)$ elements of order d. Suppose $x \in \langle a \rangle$ and $x \in \langle b \rangle$ where $|x| = d$ then $\langle a \rangle = \langle x \rangle = \langle b \rangle$ which implies $b \in \langle a \rangle$ a contradiction. Thus the elements of order d in $\langle a \rangle$ and $\langle b \rangle$ are distinct. Hence we count $2\phi(d)$ elements of order d thus far in G. Continuing this process[44] yields the number of elements of order d is a multiple of $\phi(d)$. □

We have considered groups which are not cyclic. For example, the Dihedral groups, $GL(n, \mathbb{R})$ or for certain n even $U(n)$ is not cyclic.

Example 1.7.10. *Consider $U(8) = \{1, 3, 5, 7\}$ we have*

$$3^2 = 5^2 = 7^2 = 1$$

In this group of order 4 we find 3 elements of order 2. Of course, $3 = 3\phi(2)$ since $\phi(2) = 1$.

Admittedly, the example above is not terribly exciting.

Example 1.7.11. *Note $U(20) = \{1, 3, 7, 9, 11, 13, 17, 19\}$ has elements $9, 11$ and 19 with order 2. This gives us three distinct subgroups of order 2 in $U(20)$. That alone shows $U(20)$ is not cyclic as it violates the Fundamental Theorem of Cyclic groups. In contrast, my brute-force argument given in the help-session (see 39:30 or so) required much more work. In fact, if we can show there are two elements of order 2 that suffices to disprove G is cyclic! A bit of arithmetic shows that $|3| = |7| = |13| = |17| = 4$ thus there are 4 elements of order 4 in $U(20)$. Notice, $\phi(4) = 2$ and $4 = 2(2)$. Observe this demonstrates the result put forth in Corollary 1.7.9.*

[42] you hopefully discussed this concept in linear algebra, and, intuitively is has the same meaning here, more later...
[43] well, fortunately, when n gets big enough encryption works, but, the math of encryption is fairly simple.
[44] G is finite, we eventually must run out of new elements of order d

At this point we reach the end of Chapter 4. I will probably draw a subgroup lattice or two for you in class. There is a subgroup lattice diagram for $\mathbb{Z}_{30}$ given on page 81 of Gallian. In the last two sections we have learned much about the structure of subgroups and the order of elements in finite cyclic groups. There are further theorems about subgroups of a finite group and more can be said about the Euler-phi function. In any event, I hope you realize as you attack the homework from these sections that you should use a mixture of explicit calculation **and** the theorems we discuss here. In particular, since we know many things about the structure of a cyclic group, it begins to be easy to see when a group is **not cyclic**. For example, G is not cyclic when:

(1.) G has two elements of order 2.

(2.) G has two elements $a \neq b$ of order 3 with $b \notin \langle a \rangle$.

(3.) G has more than one element of order d and they generate different subgroups.

(4.) G has more than one subgroup of a given order

(5.) G has d a positive divisor of $|G|$ and yet there is no subgroup of order d in G.

Basically, all I'm getting at in the above, is, if we see the results of the theorems for cyclic groups fail in a given example, then the given example is not cyclic[45]

I recommend working on Chapter 4 Gallian problems such as:

$$\#1 - 61, 63 - 65.$$

I do not collect these, but, I will keep them in mind as I construct tests. I recommended some of these in the last lecture and I don't expect you actually do all of them. But, the more you do, the more you know.

Problems for Lecture 7: (these are collected at the beginning of Lecture 9)

Problem 25: Gallian exercise #24 of page 83 (subgroup and centeralizer)

Problem 26: Gallian exercises #32, 33 and 34 of page 84 (subgroup lattice diagrams)

Problem 27: Gallian exercise #41 of Chapter 4 (intersection of cyclic subgroups)

Problem 28: Gallian exercise #49 of Chapter 4 (extrapolation on cyclic data)

[45] or we miscalculated, but, that won't happen...

1.8 Lecture 8: permutations and cycle notation

I follow §2.2 of Rotman's *A First Course in Abstract Algebra*, which, is like Chapter 5 of Gallian. We defined a permutation in Definition 1.3.1. A permutation on a set S is simply a bijection which takes S as its domain and range. This definition is quite broad. It implies all invertible linear transformations on a vector space are permutations. In fact, any nonempty set of bijections on a space forms a subgroup of the group of permutations on the space provided the given set of bijections is closed under composition and inversion. All of this said, our primary focus in this section and the next will be on permutations of $S = \{1, 2, \ldots, n\}$. These have tremendous application in problem solving and they provide beautiful formulas for very complicated ideas.

Remark 1.8.1. A good amount of this Section is notation. Your main goal here is to understand the special cycle notation we develop for permutations.

Definition 1.8.2. *The* **symmetric group** *on n-symbols is the set of bijections on $\{1, 2, \ldots, n\}$ with the operation of function composition. We denote the symmetric group by S_n.*

The proof that S_n forms a group stems from three easily verified facts: let $\mathbb{N}_n = \{1, \ldots, n\}$

- **(i.)** the identity map $Id(x) = x$ for $x \in \mathbb{N}_n$ is a bijection,
- **(ii.)** the composite of bijections on $\mathbb{N}_n$ is once more a bijection on $\mathbb{N}_n$
- **(iii.)** the inverse of a bijection on $\mathbb{N}_n$ is a bijection on $\mathbb{N}_n$.

Example 1.8.3. *Consider $n = 2$. $S_n = \{Id, \alpha\}$ where $\alpha(x) = \begin{cases} 2 & \text{if } x = 1 \\ 1 & \text{if } x = 2 \end{cases}$. We calculate, $\alpha^2 = \alpha \circ \alpha = Id$ since $\alpha(\alpha(1)) = \alpha(2) = 1$ and $\alpha(\alpha(2)) = \alpha(1) = 2$.*

Customarily, instead of writing $\alpha \circ \beta$ we simply write $\alpha\beta$ when dealing with permutations. We adopt this convention (as does Gallian) for permutations.

Definition 1.8.4. *Let $\alpha \in S_n$ then α* **fixes** *i if $\alpha(i) = i$. In contrast, α* **moves** *i if $\alpha(i) \neq i$.*

Since $\alpha \in S_n$ has domain $\mathbb{N}_n = \{1, \ldots, n\}$ we either have that α moves or fixes each $i \in \mathbb{N}_n$. Next, we study S_3 and use it to introduce both **array** and **cycle notation**.

Example 1.8.5. *For $n = 3$ it is convenient to introduce a notation. Suppose $\alpha \in S_3$ is defined by:*

$$\alpha(1) = 2, \quad \alpha(2) = 3, \quad \alpha(3) = 1.$$

Denote the same permutation by $\alpha = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$. In **array** *notation,*

$$S_3 = \left\{ \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \right\}$$

Honestly, I can't bear this notation any longer. Let me show you the better way: the right-hand-side of the equations below is what is known as **cycle notation**

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = (123), \qquad \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix} = (132).$$

Generally, $\alpha = (abc)$ means $\alpha(a) = b$, $\alpha(b) = c$ and $\alpha(c) = a$. In other words,

The cycles are read from left to right and they loop back.

This means there is some ambiguity to the cycle notation:

$$(123) = (231) = (312) \qquad \& \qquad (132) = (213) = (321).$$

Furthermore, if one of the numbers is unmoved by the permutation then we may omit it as follows:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} = (13), \qquad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} = (12), \qquad \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = (23).$$

Again, there is some ambiguity, we have $(13) = (31)$ *and* $(12) = (21)$ *and* $(23) = (32)$*. In the cycle notation, we write* $Id = (1)$*. Let's see how group multiplication works in cycle notation:*

$$(21)(23) = (123) \qquad \textit{or} \qquad (31)(32) = (132).$$

To recap, in cycle notation the symmetric group in 3 *symbols is*

$$S_3 = \{(1), (123), (132), (13), (12), (23)\}$$

Let's calculate some products. Note, $(12)(12) = (1)$*,* $(13)(13) = (1)$ *and* $(23)(23) = (1)$*.*

$$(123)(123) = (132)$$

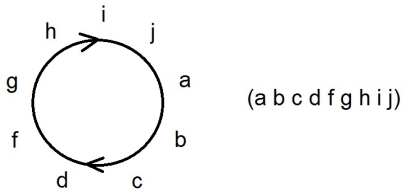
$$\begin{aligned} (123)(132) &= (1) \qquad \& \qquad (132)(123) = (1) \\ (123)(13) &= (32) \qquad \& \qquad (13)(123) = (12) \\ (123)(12) &= (13) \qquad \& \qquad (12)(123) = (23) \\ (123)(23) &= (12) \qquad \& \qquad (23)(123) = (13) \end{aligned}$$

I'll start a Cayley table for S_3*:*

$\circ$	(1)	(123)	(132)	(13)	(12)	(23)
(1)	(1)	(123)	(132)	(13)	(12)	(23)
(123)	(123)	(132)	(1)	(23)	(13)	(12)
(132)	(132)	(1)				
(13)	(13)	(12)		(1)		
(12)	(12)	(23)			(1)	
(23)	(23)	(13)				(1)

Definition 1.8.6. *Let* $\alpha \in S_n$ *and suppose* $\alpha(a_j) = a_{j+1}$ *for* $j = 1, \ldots, r$ *and* α *fixes all* $x \in \mathbb{N}_n$ *for which* $x \neq a_j$ *for* $j \in \mathbb{N}_r$*. We say* α *is an* r**-cycle** *and we denote* $\alpha = (a_1 \ldots a_r)$*. In the case* $r = 2$ *we call* α *either a* 2*-cycle or a* **transposition**

We can make pictures as on page 96 of Gallian. The diagram below equally well describes $(bcdfghija)$ or $(cdfghijab)$ and so forth. You can start wherever you like and travel around the circle (cycle)



If we consider permutations in cycle notation it is relatively simple to prove certain claims. As you read the example below, consider how it would have looked in the array notation!

Example 1.8.7. *Let $\alpha = (123)$ and $\beta = (456)$ in S_6. Consider, for $i = 1, 2, 3$*

$$\alpha(\beta(i)) = \alpha(i) \qquad \& \qquad \beta(\alpha(i)) = \alpha(i)$$

since $\beta(x) = x$ for $x = 1, 2, 3$. On the other hand, for $i = 4, 5, 6$ we have $\alpha(i) = i$ and

$$\alpha(\beta(i)) = \beta(i) \qquad \& \qquad \beta(\alpha(i)) = \beta(i)$$

Thus, as we have considered all inputs from $\mathbb{N}_6$, we find:

$$\alpha\beta = (123)(456) = (456)(123) = \beta\alpha.$$

The example above generalizes: if $\alpha = (a_1 a_2 \dots a_j)$ and $\beta = (b_1 b_2 \dots b_k)$ have no common symbol then $\alpha\beta = \beta\alpha$. The argument is the same as the example above. For $\alpha, \beta \in S_n$ all the symbols in neither α nor β are fixed by both. Then, the symbols in α are fixed by β and conversely the symbols in β are fixed by α. That is, $\alpha(b_i) = b_i$ and $\beta(a_i) = a_i$. Observe $\alpha(a_i) \in \{a_1, \dots, a_j\}$ thus $\beta(\alpha(a_i)) = \alpha(a_i)$ for each a_i. Likewise, as $\beta(b_i) \in \{b_1, \dots, b_k\}$, we calculate $\alpha(\beta(b_i)) = \beta(b_i)$. Therefore, for a_i, b_i as above and for $x \in \mathbb{N}_n - \{a_1, \dots, a_j, b_1, \dots, b_k\}$

$$\alpha(\beta(b_i)) = \beta(b_i) = \beta(\alpha(b_i)), \qquad \alpha(\beta(a_i)) = \alpha(a_i) = \beta(\alpha(a_i)), \qquad \alpha(\beta(x) = x = \beta(\alpha(x)).$$

Thus $\alpha\beta = \beta\alpha$ and we have proved the following:

Theorem 1.8.8. *If $\alpha, \beta \in S_n$ and $\alpha = (a_1 a_2 \dots a_j)$ and $\beta = (b_1 b_2 \dots b_k)$ have no common symbol then $\alpha\beta = \beta\alpha$. That is,* **disjoint cycles commute**.

Not every permutation is a cycle. For example,

$$\alpha = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 7 & 5 & 4 & 6 & 1 \end{bmatrix} = (137)(2)(45)(6) = (137)(45).$$

Let me walk through how I calculated the assertion above:

1. begin with 1 being sent to 3, write (13..,
2. next follow 3 to 7, write (137..,
3. next follow 7 to 1 this closes the first cycle (137),
4. pick a number not in (137), seems like 2 is good, note 2 goes to 2 hence write (137)(2),
5. pick a number not in (137)(2) say 4 and note it goes to 5, so we add (45.. to obtain (137)(2)(45..,
6. follow 5 back to 4, the circle is complete (137)(2)(45)
7. only (6) remains, hence $\alpha = (137)(2)(45)(6)$. But, writing (2) and (6) is superfluous thus cut it back to our answer $\alpha = (137)(45)$.

I found a website that seems fairly reliable, but, backwards!. You can try it to check your answers, but, keep in mind it swaps $\alpha\beta$ for $\beta\alpha$: Andrew G. Crowell's permutation calculator. Or see the

clunky but fun Cycle Notation Generator by James Hamblin. Finally, I would point out the legendary Arturo Magidin agrees with my general approach.

It is probably obvious from the examples thus far that any permutation can be written as the product of cycles. We study the existence and structure of cycle-decompositions of permutations in Section 1.9.

I recommend working on Chapter 5 Gallian problems such as:

$$\#17, 18, 23, 29, 30.$$

I do not collect these. But, I will keep them in mind as I construct tests.

Problems for Lecture 8: (these are collected in the Question Day before Test 1)

Problem 29: Complete the Cayley table for S_3 in Example 1.8.5.

Problem 30: For each of the following permutations:

i. Write the permutation as a product of disjoint cycles.
ii. Find its inverse.
iii. Find its order.
iv. Write it as a product of transpositions and state whether it is even or odd.
v. Conjugate it by $\sigma = (123)(45)$ (i.e. compute $\sigma\tau\sigma^{-1}$).
vi. Compute τ^{99}.

(a) $\tau = (124)(35)(24)(132)$
(b) $\tau = (1253)(354)(135)$
(c) $\tau = (12435)(134)(45)$

Problem 31: Gallian exercise 10 on page 112.

Problem 32: Gallian exercise 35 on page 113.

1.9 Lecture 9: theory of permutations

Example 1.9.1. *Consider,*

$$\alpha = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 5 & 7 & 9 & 1 & 6 & 3 & 8 & 2 \end{bmatrix}$$

To write this in cycle-notation we identify 6 *and* 8 *are fixed whereas* $Y = \{1, 2, 3, 4, 5, 7, 9\}$ *are moved. Indeed,*

$$\alpha = (14925)(37)$$

We could look at this in terms of sets which are fixed by α*. To say* U *is fixed by* α *is to say* $\alpha(U) \subseteq U$*, but, as* α *is a bijection we have* $\alpha(U) = U$*. There are two nonempty sets which are fixed by* α *here;* $U_1 = \{3, 7\}$ *and* $U_2 = \{1, 2, 4, 5, 9\}$*.*

The Example above helps inspire the proof of the Lemma below:

Theorem 1.9.2. *Each permutation in* S_n *can be expressed as a product of disjoint cycles.*

Proof: let $\alpha \in S_n$.
Step One: Suppose $a_1 \in \mathbb{N}_n$ is the smallest element in $\mathbb{N}_n$ moved by α. If no such a_1 exists then α fixes every element in $\mathbb{N}_n$ and we find $\alpha = (1) = Id$. Otherwise, let $\alpha(a_1) = a_2$ where $a_2 \neq a_1$. If $\alpha(a_2) = a_1$ then note $j = 2$ the go to **Step 2**. However, if $\alpha(a_2) \neq a_1$ then define $\alpha(a_2) = a_3$ and continue to define $\alpha(a_i) = a_{i+1}$ until either we exhaust the set $\mathbb{N}_n$ or we find $\alpha(a_j) = a_1$ for some $j > 1$. Note, we cannot have $\alpha(a_j) = a_i$ for some $i > 1$ as $\alpha(a_{i-1}) = a_i = \alpha(a_j)$ contradicts the injectivity of α. Continue to **Step Two**,

Step Two: let b_1 denote the smallest element moved by α in $\mathbb{N}_n - \{a_1, a_2, \ldots, a_j\}$. If b_1 does not exist then $\alpha = (a_1 a_2 \ldots a_j)$ as all elements except $a_1, a_2, \ldots, a_j$ are fixed. Otherwise, define recursively $b_{i+1} = \alpha(b_i)$ and note there must exist k for which $\alpha(b_k) = \alpha(b_1)$ just as in the argument for the last case. If each element in $\mathbb{N}_n - \{a_1, \ldots, a_j, b_1, \ldots, b_k\}$ is fixed then we find

$$\alpha = (a_1 \ldots a_j)(b_1 \ldots b_k)$$

Continuing: otherwise, we continue to select subsets of α-moved elements in $\mathbb{N}_n$ which are set-level fixed by α. Each such set forms a cycle, thus as $\mathbb{N}_n$ is finite, we eventually cover all the α-moved elements in $\mathbb{N}_n$ by a product of cycles. In our current notation:

$$\alpha = (a_1 \ldots a_j)(b_1 \ldots b_k) \cdots (c_1 \ldots c_l)$$

Furthermore, by construction, the cycles above are disjoint. □

If this is not convincing. Feel free to read Rotman's page 112-113 where he proves the Theorem above with an explicit induction argument.

Theorem 1.9.3. *Let* $n \geq 2$*. Each permutation in* S_n *can be expressed as a product of* 2*-cycles. That is, any element of* S_n *can be expressed as the product of transpositions.*

Proof: by Theorem 1.9.2 we know $\alpha \in S_n$ has the form $\alpha = \gamma_1 \ldots \gamma_s$ where each γ_i is a cycle. Thus, it suffices to prove a k-cycle can be written as a product of 2-cycles for each $k \in \mathbb{N}$. We use induction on cycle-length, note $(1) = (12)(12)$ for S_n where $n \geq 2$ hence a 1-cycle is the product of two transpositions. Then for a k-cycle with $k \geq 2$ we have the identity:

$$(a_1 a_2 \ldots a_k) = (a_1 a_k)(a_1 a_{k-1}) \ldots (a_1 a_2)$$

which is easily verified by case-by-case analysis. □

Example 1.9.4. *Observe,* $(12345) = (15)(14)(13)(12)$. *However,* $(ab)(ab) = (1)$ *thus the decomposition is far from unique,* $(12345) = (15)(14)(23)(23)(13)(12)$ *etc.*

You might notice, inserting or deleting $(ab)(ab) = Id$ changes the number of transpositions by 2. Our next example points out several general tricks for transpositions

Example 1.9.5. *Let* a, b, c, d *be distinct elements in* $\mathbb{N}_n$ *then since disjoint cycles commute,*

$$(ab)(cd) = (cd)(ab)$$

of course you could prove this directly without using the general disjoint cycles commute result we proved earlier. Next, the multiply-out then cyclically permute tricks:

$$(bc)(ab) = (acb) = (cba) = (ca)(cb)$$

and

$$(ac)(ab) = (abc) = (bca) = (ba)(bc)$$

and finally, as we noted in the previous example, $(ab)(ab) = 1$.. *Notice, we also find*

$$(bc)(ab) = (ca)(cb) \quad \Rightarrow \quad (ca)(bc)(ab) = (cb)$$

and

$$(ac)(ab) = (ba)(bc) \quad \Rightarrow \quad (ba)(ac)(ab) = (bc)$$

Some of the identities above are important for our proof that the identity permutation is a product of an even number of transpositions. Before I get to that result, let me exhibit a few more examples of cycle calculation.

Example 1.9.6. *for* $a \in \mathbb{N}$,

$$(12) = (1a)(21)(a2) \qquad \& \qquad (12) = (2a)(a2)(a1)$$

or, if numbers help you check out,

$$(12) = (13)(21)(32) \qquad \& \qquad (12) = (24)(42)(41).$$

Theorem 1.9.3 showed that every permutation can be expressed as a product of 2-cycles and the examples above illustrate how the decomposition into 2-cycles is far from unique. Despite the non-uniqueness there is still a fairly clear pattern to discern:

> **Conjecture:** If a 2-cycle decomposition of a permutation has an even number of transpositions then any other decomposition also has an even number of transpositions. Likewise, if the permutation permits a decomposition into an odd number of 2-cycles then any other decomposition will also have an odd number of 2-cycles.

The next page or so of notes is devoted to proving the conjecture above is true.

Lemma 1.9.7. *Suppose $n \geq 2$. If the identity permutation is written as a product of transpositions then the number of transpositions in the decomposition is even.*

Proof: First, we know that $(1) = (12)(12)$, so the identity is even. Now suppose that $(1) = (a_1a_2)\cdots(a_{\ell-1}a_\ell)$. We want to show that there **must** be an even number of these transpositions. First, let's see how to push transpositions past each other. There are 4 cases of interest: Let a, b, c, d be distinct elements of the set $\{1, 2, \ldots, n\}$.

- $(cd)(ab) = (ab)(cd)$ — disjoint cycles commute.
- $(bc)(ab) = (acb) = (cba) = (ca)(cb)$ — multiply out, cyclicly permute, transposition trick.
- $(ac)(ab) = (abc) = (bca) = (ba)(bc)$ — same as before.
- $(ab)(ab) = (1)$

Notice that in the first 3 cases, we can move a to the left. In the last case, we cancel a out completely.

Now suppose a is the largest number appearing among all the transpositions in $(a_1a_2)\cdots(a_{\ell-1}a_\ell)$. We can take the right-most occurrence of a and move it to the left. As we move all of the a's to the left, at some point, the a's must cancel out (we have to end up with the "$(ab)(ab)$" case). If not, we would have $(1) = (ab)\tau$ with no a's appearing in τ. But this is impossible since τ maps a to a (no occurrences of a in τ) and (ab) maps a to b so that $(ab)\tau$ is not the identity! Therefore, we can get rid of all of the occurrences of a by canceling out transpositions in **pairs**. Continuing in this fashion (after a is gone pick the next smallest remaining number), we will eventually cancel out all of the transpositions. Since cancelations always occur in pairs, it must be that (1) was written as an even number of transpositions. Therefore, (1) cannot be odd. □

Definition 1.9.8. *Let $\alpha \in S_n$ then α is an* **even** *permutation if it can be written as the product of an even number of transpositions. Likewise, α is an* **odd** *permutation if it can be written as the product of an odd number of transpositions.*

The theorem below asserts that the categories of even and odd are mutually exclusive and cover all possible permutations in S_n.

Theorem 1.9.9. *Every permutation in S_n is either even or odd.*

Proof: Let $\sigma \in S_n$. We know by the transposition trick (Theorem 1.9.3) above that σ can be written as a product of transpositions. Suppose $\sigma = (a_1a_2)\cdots(a_{2\ell-1}a_{2\ell}) = (b_1b_2)\cdots(b_{2k-1}b_{2k})$. Then

$$\begin{aligned}(1) = \sigma\sigma^{-1} &= (a_1a_2)\cdots(a_{2\ell-1}a_{2\ell})[(b_1b_2)\cdots(b_{2k-1}b_{2k})]^{-1} \\ &= (a_1a_2)\cdots(a_{2\ell-1}a_{2\ell})(b_{2k-1}b_{2k})^{-1}\cdots(b_1b_2)^{-1} \\ &= (a_1a_2)\cdots(a_{2\ell-1}a_{2\ell})(b_{2k-1}b_{2k})\cdots(b_1b_2)\end{aligned}$$

So we have written (1) as the product of $\ell + k$ transpositions. Our lemma says that $\ell + k$ must be even. Therefore, either both k and ℓ are even or both are odd. □

Thanks to my brother Bill for the proofs above. They are taken from his handout on permutations.

Corollary 1.9.10. *Consider permutations in S_n. The product of two even or two odd permutations is an even permutation. The product of an even with an odd permutation is an odd permutation.*

Proof: let $\sigma, \beta \in S_n$ and note by Theorem 1.9.9 σ is formed from k-transpositions and β is formed from j-transpositions. We find $\sigma\beta$ is formed from $j + k$ transpositions. If j and k are even or odd then $j + k$ is even. If just one of j and k is odd then $j + k$ is odd. Since j and k are either even or odd by Theorem 1.9.9 the Corollary follows. □

Example 1.9.11. *The identity* $(1) = (12)(12)$ *is an even permutation. The permutation* $(123) = (13)(12)$ *is even. The permutation* $(12345) = (15)(14)(13)(12)$ *is even. You see the pattern. An* r*-cycle with* r **odd** *is in fact an* **even** *permutation. Remember, the terms even and odd refer to the* 2*-cycle decomposition of a given permutation.*

Example 1.9.12. *Note,* (12) *is odd. Also* $(1234) = (14)(13)(12)$ *is odd. Indeed,* $(123456) = (16)(15)(14)(13)(12)$ *is odd. In summary, if we consider an* r*-cycle with* r **even** *then the permutation is* **odd**.

It is convenient to define a function which captures the parity of a permutation. Theorem 1.9.9 indicates this function is well-defined for any choice[46] of n.

Definition 1.9.13. *The* **sign** *of a permutation is denoted* $sgn(\sigma)$ *or* $(-1)^\sigma$ *for each* $\sigma \in S_n$*. In particular, we define*

$$sgn(\sigma) = (-1)^\sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

Notice Corollary 1.9.10 implies that $(-1)^{\sigma\beta} = (-1)^\sigma(-1)^\beta$. In fact, we can use the sign of a permutation to define the determinant we studied in linear algebra: Let $A = (A_{ij})$ be an $n \times n$ matrix with entries A_{ij}. Then

$$\det(A) = \sum_{\sigma \in S_n} (-1)^\sigma A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}$$

In particular, consider a 2×2 matrix. $S_2 = \{(1), (12)\}$. Let $\sigma = (1)$. σ is even so $(-1)^\sigma = +1$. Also, let $\tau = (12)$. τ is odd so $(-1)^\tau = -1$. Thus $\det(A) = (-1)^\sigma A_{1\sigma(1)} A_{2\sigma(2)} + (-1)^\tau A_{1\tau(1)} A_{2\tau(2)} = A_{11}A_{22} - A_{12}A_{21}$ (the regular determinant formula).

Our next result is computationally important.

Lemma 1.9.14. *If* $\alpha = (a_1 a_2 \ldots a_k)$ *is a* k*-cycle then* $\alpha^{-1} = (a_k \cdots a_2 a_1)$ *and the order of* α *is* k*.*

Proof: exercise for the reader. □

The following result is due to Ruffini in 1799 according to Gallian. This is interesting given that the cycle notation is due to Cayley in an 1815 paper.

Theorem 1.9.15. *If a permutation is formed from the product of disjoint cycles then the order of a permutation is the least common multiple of the lengths of the disjoint cycle.*

Proof: Suppose $\gamma_1, \ldots, \gamma_k$ are disjoint cycles of lengths $m_1, \ldots, m_k$ respective. Let $\sigma = \gamma_1\gamma_2 \cdots \gamma_k$. We proved disjoint cycles commute in pairs, I invite the reader to prove that inductively extends to higher products. It follows that:

$$\sigma^n = \underbrace{(\gamma_1\gamma_2 \cdots \gamma_k)(\gamma_1\gamma_2 \cdots \gamma_k) \cdots (\gamma_1\gamma_2 \cdots \gamma_k)}_{n-copies} = \gamma_1^n \gamma_2^n \cdots \gamma_k^n$$

If $m = \text{lcm}(m_1, \ldots, m_k)$ then m is a multiple of each of the lengths $m_1, \ldots, m_k$. But, by Lemma 1.9.14 we know the order of a m_i-cycle is simply m_i hence $|\gamma_i| = m_i$ and $m_i \mid n$ for each $i = 1, \ldots, k$.

[46] S_1 is rather silly though

Therefore, $\gamma_i^m = (1)$ for each $i = 1, \ldots, k$ and we find $\sigma^m = (1)$. We have shown $|\sigma| \leq m$. It remains to show no smaller power than m produces the identity permutation. To see why it is not possible, suppose there was a smaller power $j < m$ for which $\sigma^j = (1)$. By calculation above that implies $\gamma_i^j = (1)$ for $j = 1, \ldots, k$. Hence $m_1 \mid j$ for each $i = 1, \ldots, k$. Hence j is a common multiple of $m_1, \ldots, m_k$ and $j < \text{lcm}(m_1, \ldots, m_k)$ which is **impossible**. Thus, by contradiction, $|\sigma| = \text{lcm}(m_1, \ldots, m_k)$. □

If you don't care for my proof, feel free to read page 100 of Gallian.

Theorem 1.9.16. *The set of even permutations forms a subgroup of S_n.*

Proof: homework. □

Definition 1.9.17. *We call $A_n = \{\sigma \in S_n \mid sgn(\sigma) = 1\}$ the* **alternating group** *of degree n*

It is important to use the word **degree** in the above since the order of A_n is not n.

Theorem 1.9.18. *For $n > 1$, the order of the alternating group is $n!/2$.*

Proof: see Gallian page 103, I will give a different proof a bit later in the course so I merely mention this as a point of trivia for the moment. □

Example 1.9.19. *Let's exhibit $A_3 \leq S_3$. In cycle notation,*

$$S_3 = \{(1), (123), (132), (13), (12), (23)\}$$

or, noting $(123) = (13)(12)$ and $(132) = (12)(13)$ we find the even cycles in S_3 are just:

$$A_3 = \{(1), (13)(12), (12)(13)\}.$$

If $\alpha = (13)(12)$ and $\beta = (12)(13)$ then $\alpha^3 = \beta^3 = (1)$ and $\alpha\beta = (1) = \beta\alpha$. What is A_3 like ?

Example 1.9.20. *We expect $A_4 \leq S_4$ will have $4!/2 = 12$ elements. The basic two-cycles we have to build with in S_4 are*

$$(12), (13), (14), (23), (24), (34)$$

Disjoint cycles commute hence $(12)(34) = (34)(12)$ and $(13)(24) = (24)(13)$ and $(14)(23) = (23)(14)$ are all in A_4 since they are even. Since $(132) = (321) = (213)$ the products below are equal

$$(12)(13) = (132), \quad (321) = (31)(32) = (13)(23), \quad (213) = (23)(21) = (12)(23).$$

Likewise, $(234) = (342) = (423)$ hence

$$(24)(23) = (32)(34) = (43)(42) \quad \Rightarrow \quad (24)(23) = (23)(34) = (34)(24)$$

All the three-cycles can be formed in such a way from products of appropriately chosen 2-cycles. Let me conclude by listing the elements in A_4 in a somewhat natural order:

$$A_4 = \{(1), (12)(34), (13)(24), (14)(23), (123), (132), (234), (243), (314), (341), (412), (421)\}$$

In summary, we have the identity permutation, the elements of order two which are formed by products of disjoint transpositions and finally the eight 3-cycles each of which fix one number while moving the remaining 3. You can count, $|A_4| = 12$. Incidentally, you can show A_4 is formed by products of $(12)(34)$ and (123).

We have more to say about alternating groups as the semester continues. They provide important examples for a number of interesting questions. Incidentally, this document contains Cayley Graphs of S_3 and A_4. I haven't talked about Cayley Graphs yet, but, they're another fun way to understand the structure of a group. Gallian has a section on them much later in the text. There is more to say about permutations, but, I defer further study for a time when we have more technology.

Example 1.9.21. *Let $\alpha = (1234)$ and $\beta = (174)$. Notice, following Lemma 1.9.14 we have:*

$$\alpha^{-1} = (4321) \qquad \& \qquad |\alpha| = 4$$

and

$$\beta^{-1} = (471) \qquad \& \qquad |\beta| = 3$$

We cannot apply Ruffini's Theorem to $\sigma = \alpha\beta = (1234)(174)$ directly since α and β are not disjoint. But, we can use the socks-shoes inverse identity to derive:

$$\sigma^{-1} = (\alpha\beta)^{-1} = \beta^{-1}\alpha^{-1} = (471)(4321).$$

Of course, we could multiply these to disjoint cycle notation:

$$\sigma^{-1} = (17)(243)$$

and Ruffini's Theorem applies to the above as to show $|\sigma^{-1}| = lcm(2,3) = 6$. But, in retrospect, as $|\sigma| = |\sigma^{-1}|$ we find the order of σ is also 6. Moreover, while I'm at it,

$$(\sigma^{-1})^{-1} = ((17)(243))^{-1} = (243)^{-1}(17)^{-1} = (342)(71) = (17)(234).$$

where in the last step I used that disjoint cycles commute as well as the loop-back feature of cycle notation. Naturally, the result above agrees with direct calculation of $(1234)(174) = (17)(234)$.

I recommend working on Chapter 5 Gallian problems such as:

$$\#2, 3, 6, 8, 9, 12, 19, 25, 27, 35, 45, 51$$

I do not collect these. But, I will keep them in mind as I construct tests.

Problems for Lecture 9: (these are collected in the Question Day before Test 1)

Problem 33: Gallian exercise 1 on page 111.

Problem 34: Orders in S_n.

(a) What are the orders of the elements in S_5? Give an example of an element with each order.

(b) Does S_{11} have an element of order 24? If so, find one. If not, explain why not.

(c) Does S_{11} have an element of order 16? If so, find one. If not, explain why not.

Problem 35: Given $n > 1$, prove A_n is a subgroup of S_n.

Problem 36: Gallian exercise 40 on page 113.

Chapter 2

On the Structure of Groups

In the next series of lectures we study how to identify when two seemingly different groups are the same. In particular, we study homomorphisms and isomorphisms. In order to create new examples we also introduce several methods to create new groups from old. The construction of the factor group is given and we also study internal and external direct products. Finally we discuss some deeper results due to Sylow as well as the Fundamental Theorem of Finite Abelian Groups. Many new constructions are given in these Chapters, but, largely these build off our work in the previous chapter. The examples of $\mathbb{Z}_n, U(n), S_n, A_n$ and the matrix groups provide players for the game we now play. However, we introduce a few new groups in this chapter, for example the quaternion 8 group. I will likely intersperse parts of Chapter 10 along side the study of Chapter 6. This Chapter in my notes contains concepts from Gallian Chapters 6-11 and 24-25. I also take some inspiration from Rotman and other soruces as we study group actions and counting problems.

2.1 Lecture 10: homomorphism and isomorphism

Isomorphisms are everywhere if we look for them. I'll begin by defining the analog of a linear transformation in our context. Recall, a linear transformation is a mapping of vector spaces which preserved linear combinations. In other words, a linear transformation preserved the structure of a vector space. In the same way, a **homomorphism** preserves the structure of a group.

Definition 2.1.1. *Let $(G_1, \star)$ and $(G_2, \bullet)$ be groups. We say $\phi : G_1 \to G_2$ is a* **homomorphism** *if $\phi(x \star y) = \phi(x) \bullet \phi(y)$ for all $x, y \in G_1$.*

Example 2.1.2. *Let $\phi(x) = e^x$ define a map from $\mathbb{R}$ with addition to $(0, \infty)$ with multiplication. Notice, $\phi(x + y) = e^{x+y} = e^x e^y = \phi(x)\phi(y)$ thus ϕ is a homomorphism.*

Example 2.1.3. *Let $\phi : GL(n, \mathbb{F}) \to \mathbb{F}^\times$ where $\mathbb{F}$ is a field and $\phi(A) = det(A)$. Notice, $det(A) \neq 0$ given A^{-1} exists hence ϕ is* **into** *$\mathbb{F}^\times$. Furthermore, by the theory of determinants, $\phi(AB) = det(AB) = det(A)det(B) = \phi(A)\phi(B)$ for all $A, B \in textGL(n, \mathbb{F})$. Thus ϕ is a homomorphism from the multiplicative group of invertible matrices over $\mathbb{F}$ to the multiplicative group $\mathbb{F}^\times = \mathbb{F} - \{0\}$.*

Example 2.1.4. *Let $\phi : S_n \to \mathbb{Z}_2$ be defined by $\phi(\sigma) = 0$ iff $\sigma \in A_n$ and $\phi(\sigma) = 1$ if $\sigma \notin A_n$. There are four cases to consider.*

- **(i.)** *If $\sigma, \beta \in A_n$ then $\sigma\beta \in A_n$ and then $\phi(\sigma) = 0$ and $\phi(\beta) = 0$ thus $\phi(\sigma\beta) = 0 = 0 + 0 = \phi(\sigma)\phi(\beta)$.*
- **(ii.)** *If $\sigma, \beta \notin A_n$ then $\sigma\beta \in A_n$ and then $\sigma\beta \in A_n$ and $\phi(\sigma) = 1$ and $\phi(\beta) = 1$ thus $\phi(\sigma\beta) = 0 = 1 + 1 = \phi(\sigma)\phi(\beta)$.*
- **(iii.)** *If $\sigma \in A_n$ and $\beta \notin A_n$ then $\sigma\beta \notin A_n$ and $\phi(\sigma) = 0$ and $\phi(\beta) = 1$ thus $\phi(\sigma\beta) = 1 = 0 + 1 = \phi(\sigma)\phi(\beta)$.*
- **(iv.)** *If $\sigma \notin A_n$ and $\beta \in A_n$ then $\sigma\beta \notin A_n$ and $\phi(\sigma) = 1$ and $\phi(\beta) = 0$ thus $\phi(\sigma\beta) = 1 = 1 + 0 = \phi(\sigma)\phi(\beta)$.*

Sometimes we have no alternative but to break into cases. It is one of the things that working problems helps you gain a better sense of. What is the proper notation to attack a given problem. Incidentally, I'm not certain there is not a more clever way to do the previous example. Perhaps the next example is it?

Example 2.1.5. *Define $\phi : S_n \to U(3) = \{-1, 1\}$ by $\phi(\sigma) = sgn(\sigma)$ for each $\sigma \in S_n$. We understand that $-1 = 2$ in $\mathbb{Z}_3$. We should note Corollary 1.9.10 we have $sgn(\sigma\beta) = sgn(\sigma)sgn(\beta)$. Hence $\phi(\sigma\beta) = sgn(\sigma\beta) = sgn(\sigma)sgn(\beta) = \phi(\sigma)\phi(\beta)$.*

Example 2.1.6. *Let V and W be vector spaces over $\mathbb{R}$ then V and W are additive groups with respect to vector addition. If $T : V \to W$ is a linear transformation then $T(x + y) = T(x) + T(y)$ for all $x, y \in V$ thus T is a homomorphism. Indeed, T has additional structure, but, I merely add this example to make an explicit connection with your previous thinking on linear algebra.*

Naturally, homomorphisms have nice properties:

Proposition 2.1.7. *If G_1 has identity e_1 and G_2 has identity e_2 and $\phi : G_1 \to G_2$ is a homomorphism then $\phi(e_1) = e_2$. In addition, $\phi(g^{-1}) = (\phi(g))^{-1}$.*

Proof: Let $(G_1, \star)$ and $(G_2, \bullet)$ be groups with identities e_1, e_2 respective. Observe,

$$e_1 = e_1 \star e_1 \quad \Rightarrow \quad \phi(e_1) = \phi(e_1 \star e_1) = \phi(e_1) \bullet \phi(e_1).$$

But, $e_2 \bullet \phi(e_1) = \phi(e_1)$ thus $e_2 \bullet \phi(e_1) = \phi(e_1) \bullet \phi(e_1)$ and by cancellation we deduce $\phi(e_1) = e_2$. Next, consider $g \star g^{-1} = e_1$ hence $\phi(g \star g^{-1}) = \phi(g) \bullet \phi(g^{-1}) = \phi(e_1) = e_2 \;\Rightarrow\; \phi(g^{-1}) = (\phi(g))^{-1}$. □

The inverse of a homomorphism need not exist, however, we can calculate the set-theoretic inverse image of any function. For example, if $\phi(x) = x^2$ for $x \in \mathbb{R}^\times$ then $\phi^{-1}\{4\} = \{-2, 2\}$ which goes to show you ϕ^{-1} is not a function. Is $\phi : \mathbb{R}^\times \to (0, \infty)$ even a homomorphism?

Proposition 2.1.8. *If $H_1 \leq G_1$ and $K_1 \leq G_2$ and $\phi : G_1 \to G_2$ is a homomorphism then:*

$$(\mathbf{1}.)\ \phi(H_1) \leq G_2, \qquad (\mathbf{2}.)\ \phi^{-1}(K_1) \leq G_1.$$

Proof: to prove (1.). Notice $e_1 \in H_1$ and $\phi(e_1) = e_2$ thus $e_2 \in \phi(H_1)$. Suppose $x, y \in \phi(H_1)$ then there exist $h_x, h_y \in H_1$ for which $x = \phi(h_x)$ and $y = \phi(h_y)$. Notice, $h_x h_y^{-1} \in H_1$ since $H_1 \leq G_1$ thus $\phi(h_x h_y^{-1}) = \phi(h_x)\phi(h_y^{-1}) = \phi(h_x)(\phi(h_y))^{-1} = xy^{-1}$ where we used Proposition 2.1.7 to pull out the inverse. Thus, $x, y \in \phi(H_1)$ implies $xy^{-1} \in \phi(H_1)$ and by the one-step subgroup test we have shown $\phi(H_1) \leq G_2$.▽

To prove (2.), notice $e_1 \in \phi^{-1}(K_1)$ since $\phi(e_1) = e_2 \in K_1$. Suppose $a, b \in \phi^{-1}(K_1)$ hence there exist $x, y \in K_1$ for which $\phi(a) = x$ and $\phi(b) = y$. Hence

$$\phi(ab^{-1}) = \phi(a)\phi(b^{-1}) = \phi(a)(\phi(b))^{-1} = xy^{-1} \in K_1$$

where we again use Proposition 2.1.7 to pull out the inverse and also $K_1 \leq G_2$. Therefore, $\phi(ab^{-1}) \in K_1$ which means $ab^{-1} \in \phi^{-1}(K_1)$ and we conclude $\phi^{-1}(K_1) \leq G_1$ by the one-step subgroup test. □

I ask you to prove a special case of the above Proposition in your homework. Please don't just quote the notes. I want you to work through it for yourself. The special cases are $K_1 = \{e_2\}$ and $H_1 = G_1$. These have names:

Definition 2.1.9. *Let $(G_1, \star)$ and $(G_2, \bullet)$ be groups and $\phi : G_1 \to G_2$ a function then the* **kernel** *of ϕ is given by: $Ker(\phi) = \phi^{-1}\{e_2\} = \{x \in G_1 \mid \phi(x) = e_2\}$. The* **image** *of ϕ is given by: $Im(\phi) = \phi(G_1) = \{\phi(x) \mid x \in G_1\}$.*

The fact that the kernel and image are subgroups provide us with powerful, efficient, methods to prove various subsets are subgroups.

Example 2.1.10. *To see $A_n \leq S_n$ simply observe that $Ker(\phi) = A_n$ for the homomorphism $\phi : S_n \to U(3)$ with $\phi(\sigma) = sgn(\sigma)$.*

Example 2.1.11. *Continuing Example 2.1.3 where we argued $\phi : GL(n, \mathbb{F}) \to \mathbb{F}^\times$ where $\phi(A) = det(A)$ is homomorphism, we find the special linear group is a subgroup since:*

$$Ker(\phi) = \{A \in GL(n, \mathbb{F}) \mid det(A) = 1\} = SL(n, \mathbb{F}).$$

Hence $SL(n, \mathbb{F}) \leq GL(n, \mathbb{F})$.

In case you forgot, or I forgot to tell you, the **special linear group** is $\mathrm{SL}(n, \mathbb{F})$ whereas $\mathrm{GL}(n, \mathbb{F})$ is the **general linear group**.

Theorem 2.1.12. *If $\phi : G_1 \to G_2$ is a homomorphism of groups ϕ is injective if and only if $Ker(\phi) = \{e_1\}$.*

Proof: let G_1 and G_2 be groups and $\phi : G_1 \to G_2$ a homomorphism.

Suppose ϕ is injective. Recall $\phi(e_1) = e_2$ hence $e_1 \in \text{Ker}(\phi)$. Suppose $x \in \text{Ker}(\phi)$ then $\phi(x) = e_2$ thus $\phi(x) = \phi(e_1)$ and by injectivity we find $x = e_1$. Thus $\text{Ker}(\phi) = \{e_1\}$.

Conversely, suppose $\text{Ker}(\phi) = \{e_1\}$. Let $x, y \in G_1$ and $\phi(x) = \phi(y)$. Multiply by $\phi(x^{-1})$ on both sides and use the homomorphism property:

$$\phi(x^{-1})\phi(x) = \phi(x^{-1})\phi(y) \quad \Rightarrow \quad e_2 = \phi(x^{-1}y).$$

Thus $x^{-1}y \in \text{Ker}(\phi)$ and we find $x^{-1}y = e_1$. Multiply by x to obtain $xx^{-1}y = xe_1$ and hence $y = x$. We conclude that ϕ is injective. □

We saw the kernel of the sign-homormorphism and the determinant homomorphism provided non-trivial subgroups. In contrast:

Example 2.1.13. *Consider $\phi(x) = e^x$ note $Ker(\phi) = \{x \in \mathbb{R} \mid e^x = 1\} = \{0\}$. Thus ϕ is an injective homomorphism.*

Isomorphisms are homomorphisms which are both injective and surjective.

Definition 2.1.14. *If $\phi : G \to \bar{G}$ is a homomorphism of groups and ϕ is a bijection then ϕ is an* **isomorphism**. *Moreover, we say G is* **isomorphic** *to $\bar{G}$ under ϕ and we write $G \approx \bar{G}$.*

In the case G and $\bar{G}$ are finite and $|G| = |\bar{G}|$ then we know the map $\phi : G \to \bar{G}$ is injective if and only if ϕ is surjective. However, in most cases, we need to explicitly verify injectivtity (perhaps by a kernel calculation) and surjectivity.

Example 2.1.15. *For $\phi(x) = e^x$ for each $x \in \mathbb{R}$ if $y \in (0, \infty)$ then note $\phi(\ln(y)) = e^{\ln(y)} = y$ thus ϕ is onto $(0, \infty)$ and as $Ker(\phi) = \{0\}$ we have ϕ is a bijection. Indeed, ϕ is an isomorphism from $(\mathbb{R}, +)$ to $((0, \infty), \cdot)$. In other words, $\mathbb{R} \approx (0, \infty)$ under the isomorphism ϕ.*

Proposition 2.1.16. *Suppose $\phi : G_1 \to G_2$ and $\gamma : G_2 \to G_3$ are isomorphisms. Then*

- **(i.)** *$Id_{G_1} : G_1 \to G_1$ and $Id_{G_1}(x) = x$ for each $x \in G_1$ is an isomorphism on G_1,*
- **(ii.)** *ϕ^{-1} is an isomorphism,*
- **(iii.)** *$\gamma \circ \phi$ is an isomorphism.*

Proof: it is simply to verify $Id_{G_1}(xy) = Id_{G_1}(x)Id_{G_1}(y)$ for all $x, y \in G_1$ and $Id_{G_1}^{-1} = Id_{G_1}$ hence Id_{G_1} is a bijection. Next, to prove ϕ^{-1} is an isomorphism first notice the inverse of a bijection is a bijection[1]. It remains to show ϕ^{-1} is a homomorphism. Let $a, b \in G_2$ and consider $a = \phi(\phi^{-1}(a))$ and $b = \phi(\phi^{-1}(b))$. Furthermore, by the homomorphism property,

$$\phi(\phi^{-1}(a)\phi^{-1}(b)) = \phi(\phi^{-1}(a))\phi(\phi^{-1}(b)) = ab$$

[1] if you're not sure about how to prove this then you should prove it!

Therefore, $\phi^{-1}(a)\phi^{-1}(b) = \phi^{-1}(ab)$ for all $a, b \in G_2$ and we have thus shown ϕ^{-1} is a homomorphism[2][3]. Finally, to prove (iii.) we recall the composite of bijections is a bijection. It remains to show $\gamma \circ \phi$ is operation-preserving: let $x, y \in G_1$,

$$(\gamma \circ \phi)(xy) = \gamma(\phi(xy)) = \gamma(\phi(x)\phi(y)) = \gamma(\phi(x))\gamma(\phi(y)) = (\gamma \circ \phi)(x)(\gamma \circ \phi)(y).$$

where we have used the homomorphism property first of ϕ on $x, y \in G_1$ and then of γ on $\phi(x), \phi(y) \in G_2$. □

Notice that the Proposition above allows us to see that $G \approx G$, $G \approx H$ implies $H \approx G$ and finally, $G \approx H$ and $H \approx K$ then $G \approx K$. In short, isomorphism forms an **equivalence relation** on groups. Each group fits into a particular equivalence class of isomorphic groups. Intuitively, two isomorphic groups are the *same*. Essentially, isomorphic groups have the same group structure. We put further detail on this thought in the next Lecture. For now, we conclude with an interesting theorem due to Cayley[4]

Theorem 2.1.17. *Let G be a group then G is isomorphic to a subgroup of permutations on G*

Proof: we noted $Id : G \to G$ is a bijection. Furthermore, define **left multiplication** by g via

$$L_g(x) = gx$$

for each $x \in G$. Suppose $g, h \in G$ and calculate for $x \in G$, by associativtity of the multiplication in G:

$$(L_g \circ L_h)(x) = g(hx) = (gh)x = L_{gh}(x)$$

for each $x \in G$. Thus $L_g \circ L_h = L_{gh}$. Moreover, $L_g \circ L_{g^{-1}} = L_{gg^{-1}} = L_e = Id$. Observe,

$$\overline{G} = \{L_g \mid g \in G\}$$

forms a subgroup of $\text{Perm}(G) = \{f : G \to G \mid f \text{ a bijection}\}$. Let $\phi : G \to \overline{G}$ be defined by $\phi(g) = L_g$. By construction, L_g is into. Note, $\phi(f \circ g) = L_{f \circ g} = L_f \circ L_g = \phi(f) \circ \phi(g)$ thus ϕ is a homomorphism. Moreover, ϕ is a surjection since each $L_f \in \overline{G}$ has $\phi(f) = L_f$. Finally,

$$\phi(g) = L_g = Id \quad \Rightarrow \quad L_g(x) = Id(x) \quad \Rightarrow \quad gx = x \quad \Rightarrow \quad g = e.$$

Hence $\text{Ker}(\phi) = \{e\}$ and we find ϕ is injective. Therefore, ϕ is an isomorphism and we conclude $G \approx \overline{G}$. □

I should mention, the construction of $\overline{G}$ is called the **regular representation** of G on permutations.

[2]Perhaps you recall this theorem from linear algebra as well; if a bijection is linear then its inverse is automatically linear as well. That is a particular instance of this group theoretic theorem

[3]doubtless this theorem transcends mere group theory to some categorical uber theorem, but, I leave that to Nathan BeDell for the moment

[4]this is our first time meeting Cayley's Theorem, we return to this once or twice more and add further baggage

I recommend working on Chapter 6 Gallian problems such as:

$$\#1 - 43.$$

I do not collect all of these. But, I will keep them in mind as I construct tests.

Problems for Lecture 10: (these are collected at the start of Lecture 12)

Problem 37: Let $\phi : G_1 \to G_2$ be a homomorphism of groups. Prove $\mathrm{Ker}(\phi) \leq G_1$ and $\mathrm{Im}(\phi) \leq G_2$,

Problem 38: Gallian number 5 on page 130

Problem 39: Gallian number 7 on page 130

Problem 40: Gallian number 24 on page 130-131.

2.2 Lecture 11: isomorphism preserves structure

When groups are isomorphic they share the same group structure. There are many facets to this statement; order of the group, number of elements of each order, abelian, size of center and types of subgroups. I'm being a bit vague when I say *types of subgroups*, later we use collections of subsets to characterize groups.

Example 2.2.1. *Any pair of cyclic groups of order n are isomorphic. In particular, suppose $G = \langle a \rangle$ and $H = \langle x \rangle$ where $|a| = |x| = n$. Define $\phi(a^j) = x^j$. Notice,*

$$\phi(a^j a^k) = \phi(a^{j+k}) = b^{j+k} = b^j b^k = \phi(a^j)\phi(a^k)$$

Notice $\phi(a^j) = b^j = 1_H$ implies $j = 0$ hence $Ker(\phi) = \{1_G\}$. I'm using Theorem 1.6.11 in the preceding sentence. In any event, as $|G| = |H| = n$ and we've shown ϕ is injective thus ϕ is surjective. Thus ϕ establishes the isomorphism $G \approx H$.

A similar argument can be used to prove any infinite cyclic group is isomorphic to $\mathbb{Z}$. There are several additional examples on page 121 in Gallian I hope you study.

Example 2.2.2. *Let $M = \{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{R} \}$. The mapping $\psi(a + ib) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ defines two isomorphisms when suitably interpreted. Considering M as an additive group,*

$$\psi((a+ib)+(c+id)) = \begin{bmatrix} a+c & -(b+d) \\ b+d & a+c \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \psi(a+ib) + \psi(c+id).$$

hence M and $\mathbb{C}$ are isomorphic as additive groups (I leave the easy, but tedious, proof that ψ is a bijection to the reader). On the other hand, for $a + ib \neq 0$ the corresponding matrix is also nonzero and we calculate

$$\psi((a+ib)(c+id)) = \psi((ac-bd)+i(ad+bc)) = \begin{bmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{bmatrix}$$

on the other hand,

$$\psi(a+ib)\psi(c+id) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{bmatrix}$$

Thus $\psi((a+ib)(c+id)) = \psi(a+ib)\psi(c+id)$ for all $a+ib, c+id \in \mathbb{C}^\times$. Again, I leave showing ψ is a bijection to the reader, but conclude this map demonstrates that $\mathbb{C}^\times$ is isomorphic to the nonzero group of matrices in M with respect to matrix multiplication.

The Example above illustrates a **representation** of the complex numbers as 2×2 real matrices. Generally, a **representation** of a group is a homomorphic group of matrices which serves to give a concrete computationally useful model of the abstract group. In physics, the concept of a group representation gains extra importance as these matrices are allowed to act on physical states and in this way symmetry groups interact with quantum states. Much of the energy and progress in group theory is tied to the connections which are known between the representation theory of groups and its role in modern physics. We don't get far enough in this course to do this topic justice, but, I mention it for your future studies as appropriate.

Theorem 2.2.3. *Let G and H be groups and let $\phi : G \to H$ be an isomorphism.*

(i.) *$Z(H) = \phi(Z(G))$*

(ii.) *G is abelian if and only if H is abelian.*

Proof: let $\phi : G \to H$ be an isomorphism. Suppose $a \in Z(G)$ then $ag = ga$ for all $g \in G$. Suppose $h \in H$ and consider

$$h\phi(a) = \phi(\phi^{-1}(h))\phi(a) = \phi(\phi^{-1}(h)a)$$

Now, $\phi^{-1}(h) \in G$ hence it commutes with a and we find

$$h\phi(a) = \phi(a\phi^{-1}(h)) = \phi(a)\phi(a\phi^{-1}(h)) = \phi(a)h.$$

Since h is arbitrary we've shown $\phi(a) \in Z(H)$ thus $\phi(Z(G)) \subseteq Z(H)$. Conversely, suppose $b \in Z(H)$ then as ϕ is surjective there exists $x \in G$ for which $\phi(x) = b$ (of course, $x = \phi^{-1}(b)$). Consider, for $g \in G$,

$$gx = g\phi^{-1}(b) = \phi^{-1}(\phi(g))\phi^{-1}(b) = \underbrace{\phi^{-1}(\phi(g)b) = \phi^{-1}(b\phi(g))}_{b\in Z(H)\ \&\ \phi(g)\in H} = \phi^{-1}(b)\phi^{-1}(\phi(g)) = xg.$$

Thus, as g was arbitrary, $x \in Z(G)$ and hence $b = \phi(x) \in \phi(Z(G))$ and we find $Z(H) \subseteq \phi(Z(G))$. Therefore, $Z(H) = \phi(Z(G))$ and this completes the proof of $(i.)$. The proof of (ii.) follows easily since G is abelian iff $G = Z(G)$. Note, by (i.), $Z(H) = \phi(Z(G)) = \phi(\{e_G\}) = \{e_H\}$. □

I called part (i.) of this the lemmito in class.

Lemma 2.2.4. *Let G and H be groups with $\phi : G \to H$ an isomorphism. For each $g \in G$ we have:*

(i.) *$\phi(g^{-1}) = (\phi(g))^{-1}$*

(ii.) *$\phi(g^n) = (\phi(g))^n$ for all $n \in \mathbb{Z}$*

Proof: to prove (i.) note $gg^{-1} = e_G$ thus

$$\phi(e_G) = \phi(gg^{-1}) \quad \Rightarrow \quad e_H = \phi(g)\phi(g^{-1}) \quad \Rightarrow \quad (\phi(g))^{-1} = \phi(g^{-1}).$$

To prove (ii.) we begin by proving the claim for $n \in \mathbb{N}$. Notice $n = 1$ is notation. Suppose $\phi(g^n) = (\phi(g))^n$ for some $n \in \mathbb{N}$ and $g \in G$. Consider,

$$\phi(g^{n+1}) = \phi(g^n g) = \underbrace{\phi(g^n)\phi(g) = (\phi(g))^n}_{using\ induction\ hypothesis}\ \phi(g) = (\phi(g))^{n+1}$$

Thus the induction hypothesis is true for $n+1$ and we conclude by induction $\phi(g^n) = (\phi(g))^n$ for all $n \in \mathbb{N}$. It remains to prove (ii.) for non-positive powers. Since $g^0 = e_G$ and $\phi(e_G) = e_H = (\phi(g))^0$ we are left with negative powers. Suppose $n = -m$ for $m \in \mathbb{N}$. Calculate,

$$\phi(g^n) = \phi(g^{-m}) = \phi((g^{-1})^m) = \underbrace{(\phi(g^{-1})^m = ((\phi(g))^{-1})^m}_{by(i.)} = (\phi(g))^{-m} = (\phi(g))^n$$

where we have used the definition $g^{-m} = (g^{-1})^m$ throughout. This completes the proof of (ii.). □

I hope you understand my convention to denote the identity in G by e_G where there might be danger of confusion. Sometimes we will be less careful and use e for the identity in both e_G and e_H, but, I thought distinguishing them would be wise in the proofs above. And now, the main event:

Theorem 2.2.5. *Let G and H be groups and let $\phi : G \to H$ be an isomorphism.*

(i.) *$\phi(\langle a\rangle) = \langle\phi(a)\rangle$ for each $a \in G$.*

(ii.) *G is cyclic if and only if H is cyclic.*

Proof: suppose $\phi : G \to H$ is an isomorphism and $a \in G$. Recall $\langle a\rangle = \{a^n \mid n \in \mathbb{Z}\}$. Let $x \in \phi(\langle a\rangle)$ then there exists $n \in \mathbb{Z}$ for which $x = \phi(a^n)$. Thus, by Lemma 2.2.4 part (ii.),

$$x = \phi(a^n) = (\phi(a))^n \in \langle\phi(a)\rangle$$

thus $\phi(\langle a\rangle) \subseteq \langle\phi(a)\rangle$. Conversely, if $y \in \langle\phi(a)\rangle$ then there exists $n \in \mathbb{N}$ for which $y = (\phi(a))^n$ and again by the Lemma 2.2.4 part (ii.) we find $y = \phi(a^n)$. Noting $a^n \in \langle a\rangle$ we have $y \in \phi(\langle a\rangle)$ and thus $\langle\phi(a)\rangle \subseteq \phi(\langle a\rangle)$. Therefore, $\langle\phi(a)\rangle = \phi(\langle a\rangle)$ and we have proved (i.). To prove (ii.) simply note G cyclic implies $G = \langle a\rangle$ for some generator $a \in G$. Then $\phi(G) = H = \langle\phi(a)\rangle$ by (i.) hence H is cyclic. The converse follows immediately as $\phi^{-1} : H \to G$ is an isomorphism so H cyclic implies G cyclic by the argument just given. □

The task of proving two groups are **not** isomorphic is often quickly accomplished via:

Theorem 2.2.6. *If G and H are groups and $\phi : G \to H$ is an isomorphism then $|a| = |\phi(a)|$*

Proof: let $\phi : G \to H$ be an isomorphism of groups and $a \in G$. Recall $|\langle a\rangle| = |a|$ and, by the same theorem, $|\langle\phi(a)\rangle| = |\phi(a)|$. Consider, if $|\langle a\rangle| = n < \infty$ then $\phi(\langle a\rangle)$ also is a set with n-elements[5]. Likewise, if $|\langle a\rangle| = \infty$ then $\langle a\rangle$ is a set with infinitely many elements and as ϕ is a bijection the image $\phi(\langle a\rangle)$ is also an infinite set; $|\langle a\rangle| = |\phi(\langle a\rangle)|$. In conclusion, $|\langle a\rangle| = |\phi(\langle a\rangle)|$ in all cases and:

$$|\langle a\rangle| = |\underbrace{\phi(\langle a\rangle)| = |\langle\phi(a)\rangle}_{\text{Theorem 2.2.6 part(i.)}}| = |\phi(a)|. \quad \square$$

The argument I sketched in class was at the level of elements and that is also a reasonable way to prove the preservation of order. That said, I enjoy the argument above.

Theorem 2.2.7. *If G and H are groups and $\phi : G \to H$ is an isomorphism. If $b \in G$ and $k \in \mathbb{Z}$ then the equation $x^k = b$ has the same number of solutions in G as does the equation $y^k = \phi(b)$ in H (we suppose $x \in G$ whereas $y \in H$).*

Proof: suppose G, H, ϕ are as in the statement above. Let $k \in \mathbb{Z}$ and $b \in G$. Suppose $x^k = b$. Observe,

$$\phi(x^k) = \phi(b) \quad \Rightarrow \quad (\phi(x))^k = \phi(b)$$

Thus $y = \phi(x)$ is a solution of $y^k = \phi(b)$ whenever x is a solution of $x^k = b$. Conversely, suppose $y^k = \phi(b)$ for some $y \in H$. Notice,

$$\phi^{-1}(y^k) = \phi^{-1}(\phi(b)) \quad \Rightarrow \quad (\phi^{-1}(y))^k = b$$

thus $\phi^{-1}(y)$ is a solution of $x^k = b$ whenever y solves $y^k = \phi(b)$. In conclusion, there is a one-to-one correspondance between the solution sets and the Theorem follows. □

[5] to say otherwise violates both injecivity and surjectivty of ϕ

Example 2.2.8. *Observe $G = \mathbb{R}^\times = \mathbb{R} - \{0\}$ is a multiplicative group and likewise $H = (0, \infty)$ is also a multiplicative group. In fact, $H \leq G$. Notice the equation $x^2 = 1$ has solution $x = \pm 1$ in G. If there was an isomorphism $\phi : G \to H$ then $\phi(1) = 1$ necessarily and the equation $y^2 = \phi(1) = 1$ would need to have two solutions. But, $y^2 = 1$ has only the $y = 1$ solution for $y \in (0, \infty)$. Therefore, no isomorphism exists between $\mathbb{R}^\times$ and $(0, \infty)$.*

Gallian gives the following example on page 125.

Example 2.2.9. *The equation $x^4 = 1$ has solutions $x = \pm 1$ for $x \in \mathbb{R}^\times$. However, the equation $y^4 = 1$ has solutions $y = \pm 1, \pm i$ for $y \in \mathbb{C}^\times$. Thus $\mathbb{R}^\times \not\approx \mathbb{C}^\times$ as the* **same** *equation has different sized solution sets in $\mathbb{R}^\times$ verses $\mathbb{C}^\times$.*

What is the mathematics of thinking about the **same** equation with different choices of variable data? Can we define an equation independent of the location of the variables? I think so. But, I'll leave the formalities of that for another time.

Changing gears considerably, we next study a special kind of isomorphism.

Definition 2.2.10. *An isomorphism from a group G to G is called an* **automorphism**. *The set of all automorpisms of G is denoted $Aut(G)$.*

Denote the set of permutations on G by $\text{Perm}(G)$. We can argue $\text{Aut}(G)$ is a subgroup of $\text{Perm}(G)$. Notice the identity $Id_G : G \to G$ is an automorphism and if $\phi, \psi \in \text{Aut}(G)$ then

$$\phi \circ \psi^{-1} : G \to G$$

is again an automorphism (by Proposition 2.1.16) thus $\phi \circ \psi^{-1} \in \text{Aut}(G)$ and we conclude by the one-step-subgroup test that $\text{Aut}(G) \leq \text{Perm}(G)$.

Example 2.2.11. *Let G be a group and define the* **conjugation by g map** *by $\phi_g(x) = gxg^{-1}$. We can show (exercise for the reader[6]) that*

$$\phi_e = Id_G \qquad \& \qquad \phi_{gh} = \phi_g \circ \phi_h \qquad \& \qquad \phi_{g^{-1}} = (\phi_g)^{-1}$$

It follows that $H = \{\phi_g \mid g \in G\} \leq Aut(G)$.

We covered the example on page 121-122 in class. Of course, the Example above warrants a definition:

Definition 2.2.12. *An automorphism $\phi : G \to G$ for which $\phi(x) = gxg^{-1}$ is called an* **inner automorphism**. *Moreover, the set of all inner automorphisms is denoted $Inn(G)$.*

We have $\text{Inn}(G) \leq \text{Aut}(G) \leq \text{Perm}(G)$. In general, there are many bijections which do not preserve the structure of the group. The fact that automorphisms are isomorphisms strictly limits their contruction.

Example 2.2.13. *Consider automorphisms of $\mathbb{Z}_3 = \{0, 1, 2\}$. We have two generators; $\langle 1 \rangle = \langle 2 \rangle = \mathbb{Z}_3$. Since generators much map to generators under an isomorphism we have two choices:*

$$\alpha_1(1) = 1 \qquad \& \qquad \alpha_2(1) = 2$$

[6] I'll be nice, $\phi_{gh}(x) = (gh)x(gh)^{-1} = g(hxh^{-1})g^{-1} = g\phi_h(x)g^{-1} = \phi_g(\phi_h(x))$ for each $x \in G$...

Notice, (although, I feel a bit silly for the abstraction below, note either $k = 0, 1, 2$)

$$\alpha_1(k) = \underbrace{\alpha_1(1) + \cdots + \alpha_1(1)}_{k-summands} = k\alpha_1(1) = k$$

likewise, $\alpha_2(k) = k\alpha_2(1) = 2k$. Of course $\alpha_1 = Id$ whereas we calculate:

$$(\alpha_2 \circ \alpha_2)(k) = \alpha_2(2k) = 2(2k) = (3+1)k = k.$$

Thus $\alpha_2^2 = Id$. Observe,

$Aut(\mathbb{Z}_3)$	α_1	α_2
α_1	α_1	α_2
α_2	α_2	α_1

&

$U(3)$	1	2
1	1	2
2	2	1

It is not hard to see $\phi(\alpha_j) = j$ for $j = 1, 2$ defines an isomorphism of $Aut(\mathbb{Z}_3)$ and $U(3)$.

Gallian gives a more exciting example which exhibits the isomorphism of the order-four groups $U(8)$ and $\mathrm{Aut}(\mathbb{Z}_8)$. In addition, Gallian provides the following generalization of these examples:

Theorem 2.2.14. *Automorphisms and the group of units for $\mathbb{Z}_n$ are isomorphic; $Aut(\mathbb{Z}_n) \approx U(n)$.*

Proof: an automorphism of $\mathbb{Z}_n$ must send generators of $\mathbb{Z}_n$ to generators[7] of $\mathbb{Z}_n$. Recall j is a generator of $\mathbb{Z}_n$ iff $j \in U(n)$. We prove that the assignment of 1 to $j \in U(n)$ naturally induces an isomorphism for each $j \in U(n)$. Let α_j be the isomorphism with $\alpha_j(1) = j$. Calculate:

$$\alpha_j(k) = \alpha_k(\underbrace{1 + \cdots + 1}_{k-summands}) = \underbrace{\alpha_j(1) + \cdots + \alpha_j(1)}_{k-summands} = k\alpha_j(1) = jk.$$

Observe, for $l, j \in U(n)$ since $U(n)$ is a group $lj \in U(n)$ and

$$(\alpha_l \circ \alpha_j)(x) = \alpha_l(\alpha_j(x)) = ljx = \alpha_{lj}(x) \quad \star$$

Thus $\alpha_l \circ \alpha_j = \alpha_{lj}$ for each $l, j \in U(n)$. Define $\Psi : \mathrm{Aut}(\mathbb{Z}_n) \to U(n)$ by $\Psi(\alpha) = \alpha(1)$ for each $\alpha \in \mathrm{Aut}(\mathbb{Z}_n)$. Observe $\alpha(1) \in U(n)$ since isomorphisms preserve order and each element of $U(n)$ has order n. Thus $\Psi : \mathrm{Aut}(\mathbb{Z}_n) \to U(n)$ is into. Moreover, Ψ is surjective since for each $j \in U(n)$ we have $\Psi(\alpha_j) = \alpha_j(1) = j$. Suppose $\Psi(\alpha) = \Psi(\beta)$ then $\alpha(1) = \beta(1)$ thus $\alpha(k) = \alpha(1 + \cdots + 1) = \alpha(1) + \cdots + \alpha(1) = \beta(1) + \cdots + \beta(1) = \beta(k)$ for each $k \in \mathbb{Z}_n$. Therefore, $\alpha = \beta$ and we've shown Ψ is injective. Finally, to see Ψ is a homomorphism, if $\alpha_l, \alpha_j \in \mathrm{Aut}(\mathbb{Z}_n)$ then (following $\star$)

$$\Psi(\alpha_l \circ \alpha_j) = (\alpha_l \circ \alpha_j)(1) = lj = \Psi(\alpha_l)\Psi(\alpha_j).$$

Thus $\mathrm{Aut}(\mathbb{Z}_n) \approx U(n)$. □

Are all the automorphisms inner automorphisms for $\mathbb{Z}_n$? When a group is not cyclic, but, we have a generators and relations presentation of the group it may still be relatively easy to calculate inner automorphisms.

[7] recall isomorphisms preserve the order of elements so an element of order n must be sent to an element of order n. That is, a generator of $\mathbb{Z}_n$ must be sent to another generator of $\mathbb{Z}_n$

Example 2.2.15. *Let* $D_3 = \{1, x, x^2, y, xy, x^2y \mid x^3 = 1, y^2 = 1, (xy)^2 = 1\}$. *Notice,* $z \in D_3$ *has the form* $z = x^k$ *or* $z = x^k y$ *where* $k = 0, 1, 2$. *Therefore, to study an automorphism* ϕ *on* D_3 *it suffices to check these two generic cases. Consider, if* $\phi(x) = x$ *and* $\phi(y) = x^2y$ *then*

$$\phi(x^m) = (\phi(x))^m = x^m = xx^m x^{-1} = \phi_x(x^m)$$

so ϕ *agrees with the inner automorphism* ϕ_x *on rotations in* D_3. *What about reflections?*

$$\phi(x^m y) = \phi(x^m)\phi(y) = (\phi(x))^m x^2 y = x^m xyx^{-1} = x(x^m y)x^{-1} = \phi_x(x^m y).$$

Therefore $\phi = \phi_x$ *and we have shown* $\phi \in Inn(D_3)$.

I expect you will have a good computational command of this idea after completing Problem 42.

Problems for Lecture 11: (these are collected at the start of Lecture 13)

Problem 41: Show $Inn(G)$ is a subgroup of $Aut(G)$.

Problem 42: Calculate $Inn(D_4)$ using the notation

$$D_4 = \{1, x, x^2, x^3, y, xy, x^2y, x^3y \mid x^4 = 1, y^2 = 1, (xy)^2 = 1\}$$

The answer is given in Gallian in the geometric notation.

Problem 43: Gallian number 30 on page 131

Problem 44: Gallian number 33 on page 131.

2.3 Lecture 12: cosets and Lagrange's Theorem

The concept of a coset is not new. We already considered this idea in the particular context of building $\mathbb{Z}_n$. Recall,

$$\mathbb{Z}_n = \{[0], [1], \ldots, [n-1]\}$$

where $[k] = k + n\mathbb{Z} = \{k + nj \mid j \in \mathbb{Z}\}$. Two numbers in $\mathbb{Z}_n$ are equal only if their difference is a multiple of n. If we pay attention to the substructure of $\mathbb{Z}_n$ then we will notice that the addition in $\mathbb{Z}_n$ is a method of adding sets of integers. Naturally, we would like to understand this concept in more generality. The abstract coset construction is a relatively new idea in mathematics. The concept of elevating sets in an initial object to points in a new object is rather imaginative. I think Dedekind was one of the first mathematicians to really think at this level[8] There are two main directions we utilize these: (1.) towards counting problems in groups, (2.) in the construction of factor groups. In this Lecture we only see (1.) in the proof of Lagrange's Theorem. In our next Lecture we see how it is sometimes possible to take the set of cosets and give it a natural group structure.

Definition 2.3.1. *Let G be a group and H a nonempty subset of G then we define*

$$aH = \{ah \mid h \in H\}$$

as the **left-H-coset** *with representative a. Also define:*

$$Ha = \{ha \mid h \in H\}$$

as the **right-H-coset** *with representative a. We denote the number of elements in aH or Ha by $|aH|$ and $|Ha|$ respectively*[9].

In the case we work with an additive group then left cosets have the form $a + H$ whereas right cosets have the form $H + a$. I should emphasize H does not have to be a subgroup, but, most of the fun results we soon cover do assume $H \leq G$. I'll follow the organization of Gallian, this is his Lemma on page 135 essentially.

Proposition 2.3.2. *Let G be a group with $H \leq G$ and suppose $a, b \in G$. We have:*

- **(i.)** $a \in aH$,
- **(ii.)** $aH = H$ *if and only if* $a \in H$,
- **(iii.)** *either* $aH = bH$ *or* $aH \cap bH = \emptyset$,
- **(iv.)** $aH = bH$ *if and only if* $a^{-1}b \in H$,
- **(v.)** $|aH| = |bH|$
- **(vi.)** $aH = Ha$ *if and only if* $H = aHa^{-1} = \{aha^{-1} \mid h \in H\}$
- **(vii.)** $aH \leq G$ *if and only if* $a \in H$

[8] In 1858 Dedekind gave a construction of the real numbers which used sets to describe numbers. We call them Dedekind-cuts in his honor. Before that point, certain questions could not be rigorously asked or answered in real analysis. Personally, I prefer the construction of $\mathbb{R}$ as the completion of the rational numbers. We will develop the rational numbers with a fair amount of rigor by the completion of Math 422 if all goes as I hope

[9] we later define the number of distinct H-cosets to be the **index** of H in G which is denoted $[G : H]$

Proof: **(i.)** to see (i.) simply notice $a = ae$ thus $a \in aH$.
(ii.) I'll prove (ii.) in two steps. First, assume $aH = H$ thus, for any $h \in H$ we have $ah \in aH$. But, $aH = H$ so $ah \in H$ and $ah = h_2$. Hence, $a = h_2h^{-1} \in H$. Second, we suppose $a \in H$. If $x \in aH$ then $x = ah_2$ for some $h_2 \in H$ thus $x \in H$ as $a, h_2 \in H \leq G$. Consequently, $aH \subseteq H$. If $x \in H$ then note $x = aa^{-1}x \in aH$ as $a^{-1}x \in H$ hence $H \subseteq aH$ and we have shown $aH = H$ which completes the proof of (ii.)

(iii.) suppose $aH \cap bH \neq \emptyset$. In particular, there exists $x \in aH \cap bH$ thus $x \in aH$ and $x \in bH$. Thus, there exist $h, k \in H$ for which $x = ah = bk$. Note, $ah = bk$ gives $a = bkh^{-1}$. Suppose $z \in aH$ then $z = ah_2$ for some $h_2 \in H$. However,

$$z = ah_2 = bkh^{-1}h_2 \in bH$$

as $kh^{-1}h_2 \in H$. We have shown $aH \subseteq bH$. By symmetry of argument, $bH \subseteq aH$ and we deduce $aH = bH$ which completes the proof of (iii.).

(iv.) suppose $aH = bH$. From (i.) $b \in aH$ hence we find there exists $h \in H$ for which $b = ah$. Thus, $a^{-1}b = h \in H$. Conversely, suppose $a^{-1}b \in H$ thus $a^{-1}b = h$ for some $h \in H$ and $b = ah$. Note,

$$bH = ahH = aH.$$

where I used (ii.) which tells us $hH = H$ for any $h \in H$.

(v.) Consider the function $f(x) = ba^{-1}x$ for each $x \in aH$. Notice, if $bh \in bH$ then $f(ah) = ba^{-1}(ah) = bh$ and we find f is surjective. If $f(x) = f(y)$ then $ba^{-1}x = ba^{-1}y$ hence $x = y$ and we find f is injective. Since $f : aH \to bH$ and f is a bijection we find the cardnality $|aH| = |bH|$.

(vi.) suppose $aH = Ha$. Let $x \in aHa^{-1}$ hence $x = aha^{-1}$ for some $h \in H$. But, $ah \in aH = Ha$ thus $ah = ka$ for some $k \in H$ hence $x = kaa^{-1} = k \in H$ thus $aHa^{-1} \subseteq H$. Conversely, suppose $x \in H$ and consider $x = a^{-1}ax = a^{-1}ha$ since $ax \in aH = Ha$ implies there exists $h \in H$ for which $ax = ha$. Thus, $x = a^{-1}ha$ for some $h \in H$ which means $x \in a^{-1}Ha$. Hence, $H \subseteq a^{-1}Ha$ and we conclude $a^{-1}Ha = H$.

Conversely, assume $H = a^{-1}Ha$. Notice,

$$aH = a(a^{-1}Ha) = Ha.$$

To be more explicit, what is meant by the statement above is:

$$aH = a(a^{-1}Ha) = a\{a^{-1}ha \mid h \in H\} = \{aa^{-1}ha \mid h \in H\} = \{ha \mid h \in H\} = Ha.$$

(vii.) if $aH \leq G$ then $e \in aH$ since aH is a subgroup. Therefore, there exists $h \in H$ such that $ah = e$ and we learn $h = a^{-1} \in aH$ hence $h^{-1} = a \in aH$ as aH is closed under inversion. Conversely, if $a \in aH$ then by (ii.) $aH = H \leq G$. $\square$

Admittedly, my proofs are not as efficient as some of Gallian's. Once you're comfortable with the coset notation and understand what it means then perhaps his proofs are superior to the rather explicit proofs I offer above.

Example 2.3.3. *Consider* $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ *notice* $H = \langle 3 \rangle = \{0, 3\}$. *Note the distinct cosets:*

$$0 + H = \{0, 3\}, \quad 1 + H = \{1, 4\}, \quad 2 + H = \{2, 5\}$$

we also may note $0 + H = 3 + H$ *and* $1 + H = 4 + H$ *and* $2 + H = 5 + H$. *Furthermore, there is no distinction* $a + H = H + a$ *for each* $a \in \mathbb{Z}_6$.

Definition 2.3.4. *If* G *is a group and* $H \leq G$ *then we define the number of distinct* H*-cosets to be the* **index** *of* H *in* G. *We denote the index by* $[G : H]$.

Example 2.3.5. *In* $\mathbb{Z}_{100}$ *we have subgroup* $H = \{0, 20, 40, 60, 80\}$ *of order* 5 *and we obtain* 20 *distinct cosets:*

$$H, 1 + H, 2 + H, \ldots, 19 + H = \{19, 39, 59, 79, 99\}.$$

Hence the index of H *is* 20*;* $[\mathbb{Z}_{100} : H] = 20$.

The next two examples are very greedy. I attempt to outline natural cosets which appear in the study of linear algebra.

Example 2.3.6. *Let* $A \in \mathbb{R}^{m \times n}$ *and recall the soution set to* $Ax = b$ *has the form* $x = x_p + x_h$ *where* $Ax_p = b$ *(the particular solution) and* $Ax_h = 0$*(the homogenous solution). In our current notation, the solution set has the form* $x_p + Null(A)$ *where the* **null space** *is defined as* $Null(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$. *For a nonhomogeneous linear system the solution set is a coset.*

I reference some material from the study of orthogonal complements with respect to an inner product. If you are unfamilar with these concepts feel free to ask me for more details.

Example 2.3.7. *Let* W *be a subspace of a vector space* V. *Then the coset* $p + W$ *is a coset of* W. *Geometrically, this is a parallel object to* W *where we have shifted the origin to* p.

If W *is a line through the origin and* $V = \mathbb{R}^2$ *then* $p + W$ *is precisely a parallel line to* W. *It is simple to see* $\mathbb{R}^2$ *is* **foliated** *by such cosets*[10]. *Notice* $W^\perp$ *is the normal line through the origin and we learned* $\mathbb{R}^2 = W \oplus W^\perp$. *There is always some* $p_2 \in W^\perp$ *for which* $p + W = p_2 + W$ *since* $p = p_1 + p_2$ *for* $p_1 \in W$ *and* $p_2 \in W^\perp$ *for each* $p \in V$. *In our current context,* $W^\perp$ *is just a line so each coset is uniquely given by the intersection point of* $W^\perp$ *and* $p + W$.

If W *is a line through the origin and* $V = \mathbb{R}^3$ *then once again* $p + W$ *is a parallel line. Now* $W^\perp$ *is the plane through the origin with normal line* W. *It follows that the coset* $p + W$ *is uniquely labeled by the intersection point of the line* $p + W$ *with the plane* $W^\perp$.

This story continues for finite dimensional vector spaces. We can always[11] *define an inner product on* V *and so define a perpendicular space to a given subspace* W. *In particular,*

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}$$

The theory of orthogonal complements told us $dim(W) + dim(W^\perp) = dim(V)$. *The perpendicular space serves to label the* W*-cosets* $p + W$. *Notice, there are as many cosets of* W *as there are points in* $W^\perp$.

[10] a foliation is essentially a partition of the space into equidimensional submanifolds which fit together nicely, see this wikipedia article for some less trivial foliations

[11] if $\beta = \{v_1, \ldots, v_n\}$ is a basis for V then $\langle v_i, v_j \rangle = \delta_{ij}$ extended bilinearly defines an inner product for the real vector space V

Example 2.3.8. *Consider $G = GL(n, \mathbb{R})$ the general linear group of $n \times n$ matrices over $\mathbb{R}$. The special linear group is defined by $SL(n, \mathbb{R}) = \{A \in \mathbb{R}^{n \times n} \mid det(A) = 1\}$. Let $H = SL(n, \mathbb{R})$ and consider for $g \in G$,*

$$gSL(n, \mathbb{R}) = \{gA \mid det(A) = 1\} \qquad \& \qquad SL(n, \mathbb{R})g = \{Ag \mid det(A) = 1\}$$

But, $det(gA) = det(Ag) = det(A)det(g) = det(g)$ and we conclude:

$$gSL(n, \mathbb{R}) = SL(n, \mathbb{R})g = \{B \in \mathbb{R}^{n \times n} \mid det(B) = det(g)\}.$$

The cosets of the special linear group are sets of equal determinant matrices.

18th-century work by Lagrange and Euler set the stage for the 19-th century work of Galois, Abel, Cayley and others. This theorem was found by Lagrange before group theory had been completely formed. I suppose this is much like the result of Ruffini we saw earlier. In the study of math, we often find parts of a larger story before the whole story in known. There is something similar happening right now with what is known as *Langlands Program*.

Theorem 2.3.9. Lagrange's Theorem: *If G is a finite group and $H \leq G$ then $|H| \mid |G|$. Moreover, the number of distinct left (or right) cosets in G is $|G|/|H|$; that is, $[G : H] = |G|/|H|$.*

Proof: from (iii.) of Proposition 2.3.2 we know that $H \leq G$ gives a partition of G into distinct H-cosets . Moreover, by (v.) we know $|H| = |aH|$ for each $a \in G$ which means the distinct cosets each have $|H|$ elements. It follows there can only be finitely many cosets as G is a finite group. Suppose $H, a_2H, \ldots, a_kH$ are the distinct cosets which partition G. Thus,

$$G = H \cup a_2H \cup \cdots \cup a_kH$$

counting,

$$|G| = |H| + |a_2H| + \cdots + |a_kH| = \underbrace{|H| + \cdots + |H|}_{k-summands} = k|H|$$

Thus $|H| \mid |G|$ and $[G : H] = k = |G|/|H|$. □

You should notice, in our previous work we only knew this result for *cyclic* groups. You might recall my reaction to your work on Problem 28 where I was not happy you assumed G was cyclic when it was only given that G was abelian. In retrospect, you should understand, your (then) wrong solution may well have been correct if you already knew Lagrange's Theorem. Part of the challenge of this course is keeping track of what is in our toolkit. The addition of Lagrange's Theorem is a game-changer.

We have a few flowers to pick.

Corollary 2.3.10. *In a finite group G, $|a| \mid |G|$ for each $a \in G$.*

Proof: Let $|G| < \infty$ and take $a \in G$. Observe $|\langle a \rangle| = |a|$ and $\langle a \rangle \leq G$ thus we find $|a| \mid |G|$ by Lagrange's Theorem. □

Corollary 2.3.11. *A group of prime order is cyclic.*

Proof: suppose $|G| = p$ where p is prime. Let $a \in G$ and $a \neq e$. Observe $|a| = |\langle a \rangle| \mid p$ by Lagrange's Theorem. Hence $|a| = 1$ or $|a| = p$ since p is prime. But, $a \neq e$ hence $|a| \neq 1$ thus $|a| = p$ and we conclude $\langle a \rangle = G$. □

You might also notice the theorem above allows us to prove every non-identity element of G serves as a generator in the case $|G|$ is prime.

Corollary 2.3.12. *Let G be a finite group and $a \in G$ then $a^{|G|} = e$.*

Proof: by Corollary 2.3.10 we know there exists $k \in \mathbb{N}$ for which $|G| = k|a|$. Thus, $a^{|G|} = a^{k|a|} = (a^{|a|})^k = e^k = e$. □

There are a few results in Chapter 7 of Gallian we have yet to cover. Fermat's Little Theorem, the classification of groups of order $2p$ and most notably the Orbit Stabelizer Theorem. We will return to those in a later lecture if all goes as planned.

I recommend working on Chapter 7 Gallian problems such as:

$$\#1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 24.$$

I do not collect all of these. But, I will keep them in mind as I construct tests.

Problems for Lecture 12: (these are collected at the start of Lecture 14)

Problem 45: Gallian number 7 on page 145

Problem 46: The following pairs of groups are isomorphic. Prove it.

(a) $U(7)$ and $\mathbb{Z}_6$

(b) $H = \left\{ \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \middle| \, n \in \mathbb{Q} \right\}$ and $\mathbb{Q}$

Problem 47: The following pairs of groups are **not** isomorphic. Prove it.

(a) $(\mathbb{Z}_5)^{2\times 2}$ and $\mathrm{GL}_2(\mathbb{R})$

(b) $\mathbb{Z}_{222}$ and D_{111}

(c) A_4 and D_6

(d) $\mathbb{R}_{\neq 0}$ and $\mathbb{C}_{\neq 0}$ (non-zero reals and complex numbers both under multiplication)

Problem 48: Gallian number 42 on page 132.

2.4 Lecture 13: on dividing and multiplying groups

I'll begin with how to divide groups since it fits into the storyline we began last lecture. In particular, we should take another look at cosets. If you study the examples in the last lecture carefully, you'll notice that there is no difference between the left and right cosets. This is a quirk of the examples I chose. It is not uncommon for there to be a difference.

Example 2.4.1. *Consider* $S_3 = \{(1), (12), (13), (23), (123), (132)\}$ *and the subgroup* $H = \{(1), (13)\}$. *Note:* $(12)(13) = (132)$ *and* $(123)(13) = (23)$ *and* $(13)(12) = (123)$ *and* $(13)(132) = (23)$ *thus:*

Left H cosets	*Right H cosets*
$(1)H = \{(1), (13)\} = (13)H$	$H(1) = \{(1), (13)\} = H(13)$
$(12)H = \{(12), (132)\} = (132)H$	$H(12) = \{(12), (123)\} = H(123)$
$(123)H = \{(123), (23)\} = (23)H$	$H(23) = \{(23), (132)\} = H(132)$

You can see the left and right cosets which don't contain (1) *are not the same.*

Suppose we tried to define an operation on cosets by multiplying representatives; that is, suppose $(aH)(bH) = abH$. Would this make sense for the cosets of $H = \{(1), (13)\}$ in S_3? Notice,

$$(12)H(123)H = (12)(123)H = (23)H$$

yet, $(12)H = (132)H$ and so,

$$(12)H(123)H = (132)H(123)H = (132)(123)H = (1)H$$

Apparently, the multiplication of the cosets $(12)H$ and $(123)H$ by the proposed rule does not yield a single result. In short, the proposed operation is **not a binary operation**. It turns out the missing ingredient is that the left and right cosets don't match.

Definition 2.4.2. *If* G *is a group and* $H \leq G$ *then we say* H *is a* **normal subgroup** *iff* $aH = Ha$ *for each* $a \in G$. *We indicate a subgroup is normal by writing* $H \trianglelefteq G$.

Example 2.4.3. *If* G *is an abelian group and* $H \leq G$ *then*

$$a + H = \{a + h \mid h \in H\} = \{h + a \mid h \in H\} = H + a$$

thus every subgroup of an abelian group is normal.

As an application of the above example, note every subspace of a vector space is a normal subgroup of the additive group formed by the vector space with respect to vector addition.

Example 2.4.4. *In retrospect, we showed* $SL(n, \mathbb{R})$ *was a normal subgroup of* $GL(n, \mathbb{R})$ *in Example 2.3.8.*

Example 2.4.5. *The dihedral group* $D_n = \{1, x, \ldots, x^{n-1}, y, xy, \ldots, x^{n-1}y \mid x^n = 1, y^2 = 1, (xy)^2 = 1\}$ *has a subgroup of rotations* $\langle x \rangle = \{1, x, \ldots, x^{n-1}\}$. *Notice,*

$$y\langle x \rangle = \{yx^k \mid k = 0, 1, \ldots, n-1\} = \{x^{-k}y \mid k = 0, \ldots, n-1\} = \langle x \rangle y$$

since $x^{-k} = x^{n-k}$. *Here* $[D_n : \langle x \rangle] = 2n/n = 2$ *and we have just two cosets. In fact,* $\langle x \rangle \trianglelefteq D_n$

If $H \leq G$ then $eH = He$ so we only have to check $aH = Ha$ for $aH \neq H$. In the example above there was just one such coset to check.

Example 2.4.6. *It can be shown that* $A_n \trianglelefteq S_n$.

I'll illustrate how this happens for $n = 3$.

Example 2.4.7. *Consider* $S_3 = \{(1), (12), (13), (23), (123), (132)\}$ *and the subgroup* $H = A_3 = \{(1), (123), (132)\}$.

Left H cosets	*Right H cosets*
$(1)H = \{(1), (123), (132)\} = (123)H = (132)H$	$H(1) = \{(1), (123), (132)\} = H(123) = H(132)$
$(12)H = \{(12), (13), (23)\} = (13)H = (23)H$	$H(12) = \{(12), (13), (23)\} = H(13) = H(23)$

Note $H \trianglelefteq S_3$. *Multiply* $(12)H$ *by itself using different representatives:*

$$(12)H(12)H = (1)H,$$
$$(13)H(12)H = (13)(12)H = (123)H$$
$$(23)H(12)H = (23)(12)H = (132)H$$

However, there is no problem this time since $(1)H = (123)H = (132)H = H$. *I won't show how all calculations are unambiguous since we're about to prove it follows directly from normality of the subgroup in general. For now, let me just record the Cayley table of the group of cosets*[12] *of* $H = A_3$

S_3/H	H	$(12)H$
H	H	$(12)H$
$(12)H$	$(12)H$	H

Since $|A_n| = n!/2$ and $|S_n| = n!$ the index of A_n is always 2; $[S_n : A_n] = n!/(n!/2) = 2$. It follows the example above generalizes fairly easily. In S_n/A_n we'd have the coset A_n and the coset $(12)A_n$.

The following was found by O. Hölder in 1889 according to Gallian. Also, apparently the notation G/H for the *factor group of* G *by* H is due to Jordan.

Theorem 2.4.8. *Factor Group: Let* $H \trianglelefteq G$ *and denote the set of* H*-cosets by* G/H. *The operation* $(aH, bH) \mapsto abH$ *defines a binary operation which makes* G/H *a group with identity* H.

Proof: we begin by showing the operation is well-defined. Note that the rule $(aH)(bH) = abH$ does assign at least one element $abH \in G/H$ for each pair (aH, bH) in $G/H \times G/H$. It remains to show the assignment is single-valued. Suppose H is normal and $aH = a'H$ and $bH = b'H$,

$$\begin{aligned}
(a'H)(b'H) &= a'b'H && \text{definition of operation}\\
&= a'bH && \text{we assumed } b'H = bH\\
&= a'Hb && \text{we have } bH = Hb\\
&= aHb && \text{we assumed } a'H = aH\\
&= abH && \text{once again } Hb = bH\\
&= (aH)(bH) && \text{definition of operation}
\end{aligned}$$

thus $(aH)(bH) = abH$ defines a binary operation on G/H. Next, we claim H serves as the identity. Notice $e \in H$ thus $eH = H$. Let $aH \in G/H$ and consider[13]

$$(aH)(H) = (aH)(eH) = aeH = aH \quad \& \quad (H)(aH) = (eH)(aH) = eaH = aH$$

[12] You can easily verify $\Psi(0) = H$ and $\Psi(1) = (12)H$ defines an isomorphism of $\mathbb{Z}_2$ and S_3/A_3.

[13] I'm using Proposition 2.3.2 part (i.) and (ii.) to absorb e into H

thus H serves an identity for G/H. Next, observe $(aH)^{-1} = a^{-1}H$ for each $aH \in G/H$ as

$$(aH)(a^{-1}H) = aa^{-1}H = eH = H \quad \& \quad (a^{-1}H)(aH) = a^{-1}aH = eH = H$$

where we knew a^{-1} existed for $a \in G$ and $aa^{-1} = a^{-1}a = e$ as G is a group. It remains to verify the associativity of the product on G/H. Suppose $aH, bH, cH \in G/H$ and observe:

$$(aH)\big((bH)(cH)\big) = (aH)(bcH) = a(bc)H = (ab)cH = (abH)(cH) = \big((aH)(bH)\big)(cH).$$

Thus $G/H = \{aH \mid a \in G\}$ forms a group with identity H. □

Remark 2.4.9. The idea of the quotient group or factor group by H is to glue all the points in H together into a single element. Because of the structure of group multiplication we are forced to glue all points in each distinct coset of H together in the same fashion. The result is a smaller group. We found in the previous lecture the number of cosets was $[G : H] = |G|/|H|$ which we know realize means $|G/H| = |G|/|H|$. A victory of notation I suppose.

The examples on pages 175-179 of Gallian help to bring my remark above to life. I hope you'll study those.

Example 2.4.10. *Consider* $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ *we have subgroup* $H = \langle 3 \rangle = \{0, 3\}$ *with cosets* $1 + H = \{1, 4\}$ *and* $2 + H = \{2, 5\}$.

$\mathbb{Z}_6/H$	H	$1+H$	$2+H$
H	H	$1+H$	$2+H$
$1+H$	$1+H$	$2+H$	H
$2+H$	$2+H$	H	$1+H$

compare with

$\mathbb{Z}_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

You can see $\mathbb{Z}_6/\langle 3 \rangle \approx \mathbb{Z}_3$.

Example 2.4.11. *Consider* $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ *we have subgroup* $K = \langle 2 \rangle = \{0, 2, 4\}$ *with coset* $1 + K = \{1, 3, 5\}$.

$\mathbb{Z}_6/K$	K	$1+K$
K	K	$1+K$
$1+H$	$1+K$	K

compare with

$\mathbb{Z}_2$	0	1
0	0	1
1	1	0

You can see $\mathbb{Z}_6/\langle 2 \rangle \approx \mathbb{Z}_2$.

Naturally, we have much more to say about quotient groups in future lectures. I merely intiate the discussion here. Likewise, let me introduce how we may form the product of groups. There are two kinds of products to consider at this point:

1. **Internal:** recognize two or more subgroups of a given group may be multiplied to generate the entire group.
2. **External:** take two distinct groups and multiply them to form a new group.

Definition 2.4.12. *If* G *is a group and* $H, K \leq G$ *with* $H \cap K = \{e\}$ *and*

$$H \oplus K = \{hk \mid h \in H, k \in K\} = G$$

then we say G *is the* **internal direct product of** H **and** K *and we write* $G = H \oplus K$.

Remark 2.4.13. In the case that G is an additive group we say $G = H \oplus K$ is a **direct sum** decomposition. In addition, we write $H + K = \{h + k \mid h \in H, k \in K\}$ and say G is formed by the sum of H and K if $G = H + K$. The added condition $H \cap K = \{0\}$ makes the sum a direct sum. Likewise, for a multiplicative group we have the concept of a non-direct product. We write $HK = \{hk \mid h \in H, k \in K\}$ and if $G = HK$ then we say G is formed by the product of H and K. If in addition we have $H \cap K = \{e\}$ for $H, K \trianglelefteq G$ then we say G is the direct product of H and K and write $G = H \oplus K$. You might want me to write something like $H \otimes K$ here, but, that notation usually means something else we'll discuss in Math 422.

Example 2.4.14. *Note $G = \mathbb{Z}_6$ is formed by the internal direct product of $H = \langle 2 \rangle = \{0, 2, 4\}$ and $K = \langle 3 \rangle = \{0, 3\}$. Clearly H and K are normal since G is abelian and $H \cap K = \{0\}$. It is routine arithmetic to verify $H \oplus K = G$. Note the elements of $H \cup K$ are clearly in $H + K$ and mod 6:*

$$1 = 3 + 4, \quad 5 = 2 + 3$$

thus $H + K = \mathbb{Z}_6$ and we conclude $G = H \oplus K$.

Gallian's Chapter 8 also tells us how to recognize a groups is formed the internal direct product of three or more subgroups, but, I wish to return to that more complicated discussion in a future lecture. However, I can tolerate the general external product here:

Definition 2.4.15. *Let $G_1, G_2, \ldots, G_n$ be groups then we define the* **external direct product of** *$G_1, G_2, \ldots, G_n$ to be $G_1 \times G_2 \times \cdots \times G_n = \{(g_1, g_2, \ldots, g_n) \mid g_i \in G_i \text{ for } i = 1, \ldots, n\}$ with:*

$$(x_1, x_2, \ldots, x_n)(y_1, y_2, \ldots, y_n) = (x_1 y_1, x_2 y_2, \ldots, x_n y_n).$$

If $e_i \in G_i$ is the identity of G_i for each $i = 1, 2, \ldots, n$ then calculate:

$$(e_1, e_2, \ldots, e_n)(x_1, x_2, \ldots, x_n) = (e_1 x_1, e_2 x_2, \ldots, e_n x_n) = (x_1, x_2, \ldots, x_n)$$

for each $(x_1, x_2, \ldots, x_n) \in G$ where I let $G = G_1 \times G_2 \times \cdots \times G_n$. Therefore, $e_G = (e_1, e_2, \ldots, e_n)$. Furthermore,

$$(a_1, a_2, \ldots, a_n)^{-1} = (a_1^{-1}, a_2^{-1}, \ldots, a_n^{-1})$$

is easily verifed:

$$(a_1^{-1}, a_2^{-1}, \ldots, a_n^{-1})(a_1, a_2, \ldots, a_n) = (a_1^{-1} a_1, a_2^{-1} a_2, \ldots, a_n^{-1} a_n) = (e_1, e_2, \ldots, e_n).$$

I'll leave associativity to the reader.

Remark 2.4.16. The reason I choose to ignore Gallian's non-standard notation (he trades $\times$ for $\oplus$) is simply that I want our notation for products of groups to naturally fit with our already decided notation for Cartesian products. The Definition and discussion above simply say that when we take the Cartesian product of groups there is a natural group structure which is inherited from the factors in the product. Futhermore, my notation here is now consistent with that of Math 321.

Example 2.4.17. *Consider $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ has $(a, b) + (c, d) = (a + c, b + d)$ this is the direct product of the additive group of $\mathbb{R}$ and itself. We can also write $\mathbb{R}^2 = (1, 0)\mathbb{R} \oplus (0, 1)\mathbb{R}$ since $(1, 0)\mathbb{R} = \{(x, 0) \mid x \in \mathbb{R}\}$ and $(0, 1)\mathbb{R} = \{(0, y) \mid y \in \mathbb{R}\}$ have $(1, 0)\mathbb{R} \cap (0, 1)\mathbb{R} = \{(0, 0)\}$ and $(1, 0)\mathbb{R} + (0, 1)\mathbb{R} = \mathbb{R}^2$.*

Example 2.4.18. *Let $G = \mathbb{Z}_2 \times \mathbb{Z}_3$. Explicitly,*

$$G = \{(0,0),\ (0,1),\ (0,2),\ (1,0),\ (1,1),\ (1,2)\}$$

For the sake of curiousity, let's build the Cayley table:

$+$	$(0,0)$	$(0,1)$	$(0,2)$	$(1,0)$	$(1,1)$	$(1,2)$
$(0,0)$	$(0,0)$	$(0,1)$	$(0,2)$	$(1,0)$	$(1,1)$	$(1,2)$
$(0,1)$	$(0,1)$	$(0,2)$	$(0,0)$	$(1,1)$	$(1,2)$	$(1,0)$
$(0,2)$	$(0,2)$	$(0,0)$	$(0,1)$	$(1,2)$	$(1,0)$	$(1,1)$
$(1,0)$	$(1,0)$	$(1,1)$	$(1,2)$	$(0,0)$	$(0,1)$	$(0,2)$
$(1,1)$	$(1,1)$	$(1,2)$	$(1,0)$	$(0,1)$	$(0,2)$	$(0,0)$
$(1,2)$	$(1,2)$	$(1,0)$	$(1,1)$	$(0,2)$	$(0,0)$	$(0,1)$

You can check $|(1,1)| = 6$ and $|(1,2)| = 6$ whereas $|(1,0)| = 2$ and $|(0,1)| = |(0,2)| = 3$. Compare this to $\mathbb{Z}_6$ which also has $|1| = |5| = 6$, $|3| = 2$ and $|2| = |4| = 3$. We could find an explicit isomorphism of $\mathbb{Z}_2 \times \mathbb{Z}_3 \approx \mathbb{Z}_6$.

There is much theory to discover, prove and use here. I relegate all of that to a future Lecture. I recommend working on Chapter 8 Gallian problems such as:

$$\#1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.$$

also Chapter 9 Gallian problems such as:

$$\#1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 44.$$

I do not collect all of these. But, I will keep them in mind as I construct tests.

Problems for Lecture 13: (these are collected at the start of Lecture 15)

Problem 49: Cayley's theorem tells us that $Q = \{\pm 1, \pm i, \pm j, \pm k\}$ is isomorphic to a subgroup of S_8. Find such a subgroup using the ordering of elements: $1, -1, i, -i, j, -j, k, -k$ (so, for example, i is element #3). To help you get started, left multiplication by -1 sends 1 to -1, -1 to 1, i to $-i$, etc. so it sends 1 to 2, 2 to 1, 3 to 4, etc. Thus -1 corresponds with $(12)(34)(56)(78)$.

Problem 50: Let H and K be subgroups of G.

(a) Suppose that H and K are normal subgroups of G. Show that $H \cap K$ is a normal subgroup of G as well.

(b) Let $|G| = 36$, $|H| = 12$, and $|K| = 18$. Using Lagrange's Theorem, what are the possible orders of $H \cap K$?

Problem 51: Let $H = \{1, x^3, x^6\} \subseteq D_9 = \{1, x, \ldots, x^8, y, xy, \ldots, x^8y\} = \langle x, y \mid x^9 = 1, y^2 = 1, xyxy = 1\rangle$. Notice $H = \langle x^3 \rangle$, so H is a subgroup of D_9. Quickly compute $[D_9 : H]$ (i.e. the index of H in D_9). Then find all of the left and right cosets of H in D_9. Is H a normal subgroup of D_9?

Problem 52: Let G and H be groups.

(a) Show $\{e\} \times H = \{(e, h) \mid h \in H\}$ is a normal subgroup of $G \times H$ (where e is the identity of G).
Note: You need to show that $\{e\} \times H$ is a subgroup AND that it's normal.

(b) Show $G \times H \cong H \times G$.

2.5 Lecture 14: on the first isomorphism theorem

In fact, to be honest, this Lecture is on an assortment of things I probably should have included here and there earlier. But, better late than never. I begin with a discussion of the classification of groups up to order 7. Then I detail a few important observations about normal subgroups, including the definition of *simple* groups. We examine a few factor groups to showcase the utility of the classification result. Then we present the main theorems of this section which naturally lead to the first isomorphism theorem.

2.5.1 classification of groups up to order 7

First, recall Corollary 2.3.11 tells us that groups of prime order are cyclic. Therefore, up to isomorphism, there is just one group of order $1, 2, 3, 5$ and 7. For orders 4 and 6 we need further analysis.

Groups of Order 4: suppose $|G| = 4$. By Corollary 2.3.10 the order of each element of G must divide the order of G. Hence, if $a \in G$ then $|a| = 1, 2$ or 4. If there is an element of order 4 then

$$G = \{e, a, a^2, a^3\}$$

and $G = \langle a \rangle$ is cyclic. However, if G does not have an element of order 4 then G must have an element of order 2 since it cannot have more than one identity element ($|e| = 1$). Thus, in the case G has no element of order 4, it must be that G has 3 elements, say a, b, c of order 2. That is, $a^2 = b^2 = c^2 = e$. If we write a multiplication table (aka Cayley table) for this potential group we are forced to write: (everything except for the **red terms** is already forced by the assumption a, b, c have order 2, then, you can see we have to write the red terms where they are otherwise we can't have a Cayley Table)

G	e	a	b	c
e	e	a	b	c
a	a	e	**c**	**b**
b	b	**c**	e	**a**
c	c	**b**	**a**	e

compare with

$\mathbb{Z}_2 \times \mathbb{Z}_2$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(0,0)$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(1,0)$	$(1,0)$	$(0,0)$	$(1,1)$	$(0,1)$
$(0,1)$	$(0,1)$	$(1,1)$	$(0,0)$	$(1,0)$
$(1,1)$	$(1,1)$	$(0,1)$	$(1,0)$	$(0,0)$

It is not immediately obvious from the table for G that the proposed group is associative. However, as we compare with $\mathbb{Z}_2 \times \mathbb{Z}_2$ the correspondence of $a \mapsto (1,0)$, $b \mapsto (0,1)$ and $c \mapsto (1,1)$ and of course $e \mapsto (0,0)$ defines a bijection of G and $\mathbb{Z}_2 \times \mathbb{Z}_2$. We know $\mathbb{Z}_2 \times \mathbb{Z}_2$ (the Klein 4-group) is associative hence it follows the table for G also represents a associative group structure. The argument is that G has the same pattern as $\mathbb{Z}_2 \times \mathbb{Z}_2$. Incidentally, this is one of the primary utilities of the Cayley table. It allows us to prove isomorphisms for groups of small order with nice organization and relatively little writing. In summary, any group of order 4 is either isomorphic to $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Groups of Order 6: suppose $|G| = 6$ then G has elements of order $1, 2, 3$ or 6 by the corollary to Lagrange's Theorem. If G has an element of order 6 then $G \approx \mathbb{Z}_6$. Otherwise, G must have elements of order 2 and/or 3. I leave the details as a homework, but, we can argue that either G is isomorphic to $\mathbb{Z}_6$ or S_3.

Let us summarize the results we've found or you will soon work out. These will be labor saving as we decide on whether a particular group of small order is isomorphic to another.

Order of G	Representative Example of G
1	$\{e\}$
2	$\mathbb{Z}_2$
3	$\mathbb{Z}_3$
4	$\mathbb{Z}_4$ $\mathbb{Z}_2 \times \mathbb{Z}_2$
5	$\mathbb{Z}_5$
6	$\mathbb{Z}_6$ S_3
7	$\mathbb{Z}_7$

As an example of the utility of this discussion, consider $A_3 \leq S_3$. We saw $|A_3| = 6/2 = 3$ thus $A_3 \approx \mathbb{Z}_3$. No need for an explicit isomorphism now that we know all groups of order three are isomorphic. We choose $\mathbb{Z}_3$ as the quintessiential example, but, this is merely a choice. We could just as well use the group of the third roots of unity in the complex numbers:

$$S = \{\cos(2\pi j/3) + i\sin(2\pi j/3) \mid j = 0, 1, 2\} \leq \mathbb{C}^\times \approx \mathbb{Z}_3$$

2.5.2 a discussion of normal subgroups

I should point out, for any group G we have several standard normal subgroups: it is simple to verify $gHg^{-1} \subseteq H$ for any $g \in G$ in the case that $H = \{e\}, \mathrm{Z}(G)$ and G. These groups are related:

$$\{e\} \leq \mathrm{Z}(G) \leq G$$

However, these are not always distinct. For example, when G is abelian $Z(G) = G$. If there is a $H \triangleleft G$ where $H \neq G$ and $H \neq \{e\}$ then the factor group G/H is interesting. Why is it not that interesting in the case $H = \{e\}$? Well, the cosets are merely points so the factor group is just the group again up to isomorphism; $\phi(x) = x\{e\}$ for each $x \in G$ gives[14] $G/\{e\} \approx G$. In contrast, if we quotient by G then $G/G \approx \{e\}$ since G is the only G-coset in G and a group of order 1 is the identity group. In summary, for a quotient to form an interesting factor group there must be some subgroup which is nontrivial and proper. If this is not possible then in some sense the group cannot be factored so it is as basic as it can be. To give an analogy, 25/100 is not simple because we can reduce it to 5/20 which is also not simple because we can break it further to 1/4. The fraction 1/4 is in least terms, it's as simple a representative of the fraction 25/100 we can obtain in the world of fractions. The idea of a simple group is somewhat the same. You probably have some informal sense of when a fraction is reduced to lowest terms. The problem of enumerating all possible simple groups is a harder problem which we will discuss further[15] in a later lecture.

Definition 2.5.1. *A group G is called* **simple** *if the only normal subgroups of G are $\{e\}$ and G.*

Example 2.5.2. *Consider $\mathbb{Z}_p$ where p is prime. Since the only subgroups of $\mathbb{Z}_p$ are $\{0\}$ and $\mathbb{Z}_p$ we observe $\mathbb{Z}_p$ is a simple group.*

Example 2.5.3. *It can be shown (with considerable effort) that A_n is a simple group for $n \leq 5$. In contrast, A_4 is not simple, but, this would take some effort to illustrate directly.*

[14] this is a very silly quotient, note $x\{e\} = y\{e\}$ only if $x = y$. In fact, $x\{e\} = \{x\}$ so in total $G/\{e\} = \{\{x\} \mid x \in G\}$

[15] don't worry, I'm not going to reproduce the 5-10 thousand page proof of the enumeration. Famously, the Feit-Thompson result alone takes 250 pages to prove.

We'll talk about the simplicity of A_n for $n \geq 5$ some other time, but, for now let us study a relatively *simple*[16] example of a group which is not simple. First, let me make an observation:

> **Observation** in a factor group G/H the order of aH (as an element of G/H) is the smallest positive power k for which $a^k \in H$. This is due to the identity $(aH)^k = a^k H$.

There is a distinction between the order of the coset aH and the order of the element $aH \in G/H$. For example, the order of $(12) + A_3$ is 2 in S_3/A_3 whereas the number of elements in the coset $(12) + A_3$ is simply 3. Unfortunately, we use $|aH|$ to denote both concepts of order. If there is danger of ambiguity we could adopt the alternate notation $\#(aH)$ for the cardnality of the coset aH. I'm not sure if this will actually be an issue as we continue.

Example 2.5.4. *Let $G = \mathbb{Z}_4 \times \mathbb{Z}_4$. Observe,*

$$H = \{(0,0), (2,0), (0,2), (2,2)\} = 2\mathbb{Z}_4 \times 2\mathbb{Z}_4$$

is a normal subgroup[17] *with factor group:*

$$G/H = \{H,\ (1,0) + H,\ (0,1) + H,\ (1,1) + H\}.$$

We find every non-identity element in G/H has order 2:

$$2((1,0) + H) = (2,0) + H = H,$$

$$2((0,1) + H) = (0,2) + H = H,$$

$$2((1,1) + H) = (2,2) + H = H.$$

Thus $G/H \approx \mathbb{Z}_2 \times \mathbb{Z}_2$. The factor group of G by H is the Klein 4-group up to isomorphism. In contrast, we can study the subgroup

$$K = \{(1,0), (2,0), (3,0), (4,0)\} = \mathbb{Z}_4 \times \{0\}$$

which gives factor group

$$G/K = \{K, (0,1) + K, (0,2) + K, (0,3) + K\}$$

Notice the order of $(0,1) + K$ is 4 hence G/K is isomorphic to $\mathbb{Z}_4$. Our results thus far:

$$\frac{\mathbb{Z}_4 \times \mathbb{Z}_4}{2\mathbb{Z}_4 \times 2\mathbb{Z}_4} \approx \mathbb{Z}_2 \times \mathbb{Z}_2 \qquad \& \qquad \frac{\mathbb{Z}_4 \times \mathbb{Z}_4}{\mathbb{Z}_4 \times \{0\}} \approx \mathbb{Z}_4$$

I'll prove in the next lecture that $\mathbb{Z}_n/m\mathbb{Z}_n \approx \mathbb{Z}_m$ provided m is a divisor of n so with that future knowledge in hand we note $\mathbb{Z}_4/2\mathbb{Z}_4 \approx \mathbb{Z}_2$ and our results are quite nice:

$$\frac{\mathbb{Z}_4 \times \mathbb{Z}_4}{2\mathbb{Z}_4 \times 2\mathbb{Z}_4} \approx \frac{\mathbb{Z}_4}{2\mathbb{Z}_4} \times \frac{\mathbb{Z}_4}{2\mathbb{Z}_4} \qquad \& \qquad \frac{\mathbb{Z}_4 \times \mathbb{Z}_4}{\mathbb{Z}_4 \times \{0\}} \approx \frac{\mathbb{Z}_4}{\mathbb{Z}_4} \times \frac{\mathbb{Z}_4}{\{0\}}.$$

where I used that $\{0\} \times \mathbb{Z}_4 \approx \mathbb{Z}_4$ as well as $\{0\} = \mathbb{Z}_4/\mathbb{Z}_4$ to rewrite the $G/K = \mathbb{Z}_4$ result as a product of factor groups. In fact, in the next lecture we'll show if $N_1 \trianglelefteq G_1$ and $N_2 \trianglelefteq G_2$ then

[16] in the untechnical sense of the term

[17] G is abelian so every subgroup of G is normal

$\frac{G_1 \times G_2}{N_1 \times N_2} \approx \frac{G_1}{N_1} \times \frac{G_2}{N_2}$. *That said, not every factor group must be obtained in this way for a product group such as* $G = \mathbb{Z}_4 \times \mathbb{Z}_4$. *Consider,*

$$N = \{(0,0), (1,1), (2,2), (3,3)\}$$

we have factor group,

$$G/N = \{N, (1,0) + N, (2,0) + N, (3,0) + N\}$$

where $(1,0) + N$ *generates* G/N *hence* $G/N \approx \mathbb{Z}_4$.

We came across a subgroup of index 2 in our last lecture when we studied how A_3 is a normal subgroup of S_3. The reason that A_3 has matching left and right cosets generalizes to other subgroups with index 2.

Example 2.5.5. *Suppose* G *is a group and* $H \leq G$ *with* $|H| = |G|/2$ *then* $[G : H] = 2$. *Note,* $eH = He = H$ *so* H *is both a left and right coset. However, we also know either left or right cosets partition* G *into equal-sized parts. It follows that* $G - H$ *is both a left and right coset and so* $H \trianglelefteq G$.

Thus, a group which has a subgroup which is half as large as the group cannot be simple. We just saw such a subgroup is necessarily normal hence G/H is interesting. Well, not that interesting, $G/H \approx \mathbb{Z}_2$ in such a case.

2.5.3 first isomorphism theorem

Let us begin by discussing the natural homomorphism that comes with any factor group.

Definition 2.5.6. *If* $H \trianglelefteq G$ *then define* $\pi : G \to G/H$ *by* $\pi(x) = xH$ *for each* $x \in G$.

We say π is the **coset map** or **fundamental homomorphism** of the factor group G/H.

Theorem 2.5.7. *If* $H \trianglelefteq G$ *and* $\pi : G \to G/H$ *is defined by* $\pi(x) = xH$ *for each* $x \in G$ *then* π *is a homomorphsm from* G *to* G/H. *Moreover,* $Ker(\pi) = H$.

Proof: observe $\pi(x) = xH$ is a mapping into the group G/H provided H is a normal subgroup of G. Moreover,

$$\pi(xy) = xyH = (xH)(yH) = \pi(x)\pi(y)$$

by the definition of multiplication in the factor group G/H. Thus $\pi : G \to G/H$ is a homomorphism. Furthermore, $x \in \text{Ker}(\pi)$ implies $\pi(x) = H$. Hence, $xH = H$ which means $x \in H$ and we conclude $\text{Ker}(\pi) = H$. □

Example 2.5.8. *Consider the subgroup* $H = n\mathbb{Z}$ *of* $\mathbb{Z}$ *then the factor group* $\mathbb{Z}/n\mathbb{Z} = \{[k] \mid k \in \mathbb{Z}\} = \{[0], [1], \ldots, [n-1]\}$ *has coset map* $\pi(x) = [x] = x + n\mathbb{Z}$. *The kernel of* π *is* $n\mathbb{Z}$.

Normal subgroups play nicely with homomorphisms. The essential concept of the theorem below is that normal subgroups are a part of the group structure which is preserved by homomorphism.

Theorem 2.5.9. *Suppose* $\phi : G_1 \to G_2$ *is a homomorphism of groups. Then*

- **(i.)** *if* ϕ *is a surjection and* $H_1 \trianglelefteq G_1$ *then* $\phi(H_1) \trianglelefteq G_2$,
- **(ii.)** *if* $H_2 \trianglelefteq G_2$ *then* $\phi^{-1}(H_2) \trianglelefteq G_1$

Proof:(i.) suppose $\phi : G_1 \to G_2$ is a surjective homomorphism and $H_1 \trianglelefteq G_1$. Suppose $y \in G_1$ and suppose $z \in y\phi(H_1)y^{-1}$ hence there exists $h \in H_1$ for which $z = y\phi(h)y^{-1}$. Note, by surjectivity of ϕ there exists $x \in G_1$ for which $y = \phi(x)$. Thus, as ϕ is a homomorphism:

$$z = \phi(x)\phi(h)\phi(x)^{-1} = \phi(x)\phi(h)\phi(x^{-1}) = \phi(xhx^{-1})$$

By normality of H_1 we know $xhx^{-1} \in H_1$ thus $z = \phi(xhx^{-1}) \in \phi(H_1)$. We find $y\phi(H_1)y^{-1} \subseteq \phi(H_1)$ for each $y \in G_1$ and we conclude $\phi(H_1) \trianglelefteq G_2$.

(ii.) suppose $H_2 \trianglelefteq G_2$ and $\phi : G_1 \to G_2$ is a homomorphism. Suppose $g \in G_1$ and consider $x \in g\phi^{-1}(H_2)g^{-1}$. Hence, suppose there exists $h \in G_1$ for which $\phi(h) \in H_2$ and $x = ghg^{-1}$. Calculate, using the homomorphism property of ϕ,

$$\phi(x) = \phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g)^{-1}$$

by normality of H_2 we find $\phi(g)\phi(h)\phi(g)^{-1} \in H_2$. Thus $\phi(x) \in H_2$ which means $x \in \phi^{-1}(H_2)$. Hence $g\phi^{-1}(H_2)g^{-1} \subseteq \phi^{-1}(H_2)$ and we find $\phi^{-1}(H_2) \trianglelefteq G_1$. □

Corollary 2.5.10. *If $\phi : G_1 \to G_2$ is a homomorphism then $Ker(\phi) \trianglelefteq G_1$.*

Proof: observe $\mathrm{Ker}(\phi) = \phi^{-1}\{e_2\}$ where $\{e_2\} \trianglelefteq G_2$ hence by Theorem 2.5.9 part (ii.) we find the kernel is a normal subgroup of G_1. □

Surely this is prefered over checking whether left and right cosets match! Consider,

Example 2.5.11. *Consider $sgn : S_n \to \{1, -1\}$ where $sgn(\sigma) = 1$ if $\sigma \in A_n$. Hence $Ker(sgn) = A_n$ and we conclude $A_n \trianglelefteq S_n$ as we argued before that sgn is a homomorphism.*

Note also, for $n \geq 2$, S_n has two kinds of permutations. The even permutations in A_n and the odd permutations in $(12)A_n$. However, these are cosets of A_n hence $|A_n| = |(12)A_n|$ which shows $2|A_n| = |S_n| = n!$ hence the order of A_n is $n!/2$. This counting is made easy by the uniform size of cosets. We could have pointed this out in the coset lecture.

Theorem 2.5.12. *Suppose $\phi : G_1 \to G_2$ is a homomorphism of groups. Then $\bar{\phi} : G_1/Ker(\phi) \to \phi(G_1)$ defined by $\bar{\phi}(xKer(\phi)) = \phi(x)$ is an isomorphism which shows $G_1/Ker(\phi) \approx \phi(G_1)$.*

Proof: suppose $\phi : G_1 \to G_2$ is a homomorphism. By Corollary 2.5.10 we have that $K = \mathrm{Ker}(\phi)$ is a normal subgroup of G_1 hence by Theorem 2.4.8 the set of cosets G_1/K forms a group with operation $(aK)(bK) = abK$ for all $aK, bK \in G_1/K$. Suppose $\bar{\phi}(xK) = \phi(x)$ for each $xK \in G_1/K$. Clearly $\phi(x) \in \phi(G_1)$ since $x \in G_1$ whenever $xK \in G_1/K$. To see $\bar{\phi}$ is single-valued, consider $xK = yK$ thus $xy^{-1} \in K = \mathrm{Ker}(\phi)$ hence $\phi(xy^{-1}) = e_2$ and $\phi(x)\phi(y^{-1}) = e_2$ as ϕ is a homomorphism. Furthermore, homomorphisms have $\phi(y^{-1}) = \phi(y)^{-1}$ thus $\phi(x)\phi(y)^{-1} = e_2$ and, multiplying $\phi(y)$ on the right, we find $\phi(x) = \phi(y)$. Therefore, $\bar{\phi}(xK) = \phi(x) = \phi(y) = \bar{\phi}(yK)$ and we have shown $\bar{\phi} : G_1/K \to \phi(G_1)$ is a function. To see $\bar{\phi}$ is a homomorphism, let $aK, bK \in G_1/K$ and calculate:

$$\bar{\phi}((aK)(bK)) = \bar{\phi}(abK) = \phi(ab) = \phi(a)\phi(b) = \bar{\phi}(aK)\bar{\phi}(bK).$$

It remains to prove $\bar{\phi}$ is a bijection. If $z \in \phi(G_1)$ then by definition of image there exists $x \in G_1$ for which $\phi(x) = z$ and we note $\bar{\phi}(xK) = \phi(x) = z$ hence $\bar{\phi}$ is onto $\phi(G_1)$. To show $\bar{\phi}$ is injective we study its kernel: (remember we set $K = \mathrm{Ker}(\phi)$)

$$\mathrm{Ker}(\bar{\phi}) = \{xK \mid \bar{\phi}(xK) = \phi(x) = e_2\} = \{xK \mid x \in K\} = K.$$

Thus $\bar{\phi}$ is injective. Thus, $\bar{\phi} : G_1/\text{Ker}(\phi) \to \phi(G_1)$ is an isomorphism and $G_1/\text{Ker}(\phi) \approx \phi(G_1)$. $\square$

There are many applications of the first isomorphism theorem. Many of those we explore in the next lecture. Let me conclude with a non-example:

Example 2.5.13. *Let $L_{x^2}(g) = x^2 g$ for all $g \in D_4$ where the dihedral group of order 8 is given by*

$$D_4 = \{1, x, x^2, x^3, y, xy, x^2y, x^3y \mid x^4 = 1, y^2 = 1, (xy)^2 = 1\}.$$

You can calculate $Ker(\phi) = \{1, x^2\}$. Thus $\{1, x^2\} \trianglelefteq D_4$. You may recall $\{1, x^2\} = Z(D_4)$ so we are not surprised this is a normal subgroup, but, the method of demonstration is new to us. Furthermore, we note L_{x^2} is a surjective map. That is, $L_{x^2}(D_4) = D_4$. Hence, by the first isomorphism theorem

$$D_4/\{1, x^2\} \approx D_4 \qquad \leftarrow \textbf{INCORRECT}$$

What is wrong with this ? We should have $[K : D_4] = 8/2 = 4$ so this suggestion that the factor group of D_4 by $\{1, x^2\}$ is absurd. Let $\{1, x^2\} = N$. Explicitly, we calculate:

$$D_4/N = \{N, xN, yN, xyN\}$$

where $xN = \{x, x^3\}$, $yN = \{y, x^2y\}$ and $xyN = \{xy, x^3y\}$. In fact, while we're at it, note $|xN| = |yN| = |xyN| = 2$ since $(xN)^2 = x^2N = N$ and $(yN)^2 = y^2N = N$ and $(xyN)^2 = (xy)^2N = N$. It follows $D_4/N \approx \mathbb{Z}_2 \times \mathbb{Z}_2$ since D_4 is a group of order 4 with 3 elements of order 2.

Again, I ask the reader, what is wrong with the Example above? How have I abused the first isomorphism theorem?

I recommend working on Chapter 10 Gallian problems such as:

$$\#1, 2, 3, 4, 5, 6, 7, 8, 9, 10.$$

I do not collect all of these. But, I will keep them in mind as I construct tests.

Problems for Lecture 14: (these are collected at the start of Lecture 16)

Problem 53: Let G be a group of order 6. Prove G must have an element of order 2. However, also prove not every element beside the identity may have order 2.

Problem 54: Let G be a group of order 6 and $a, b \in G$ with $|a| = 3$ and $|b| = 2$. Show either G is cyclic or $ab \neq ba$.

Problem 55: Let G be a group of order 6. Show that if G is not cyclic then its multiplication table matches that of S_3. *Hint: use the previous two exercises to argue for G not cyclic, there are $a, b \in G$ with $a^3 = e$ and $b^2 = e$ and $ba = a^b$*

Problem 56: Suppose N_1 is a normal subgroup of G_1 and N_2 is a normal subgroup of G_2. Prove that $N_1 \times N_2$ is a normal subgroup of $G_1 \times G_2$. Also, prove $\frac{G_1 \times G_2}{N_1 \times N_2} \approx \frac{G_1}{N_1} \times \frac{G_2}{N_2}$.

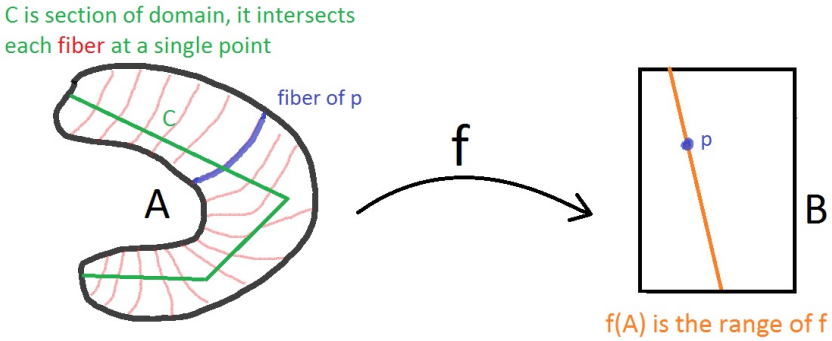
2.6 Lecture 15: first isomorphism theorem, again!

I'll begin with a discussion to attempt to bring some intuition and context to the first isomorphism theorem. I'll begin by discussing how we can create a bijection which naturally stems from any given function.

First, note to make f a surjection we simply swap B for $f(B)$. In words, trade the codomain for the range if need be. The problem of obtaining injectivity is a bit more involved. Consider, for $f : A \to B$ we can partition the domain A into **fibers**. The inverse image of a singleton is a fiber. In particular, for $b \in f(A)$,

$$f^{-1}\{b\} = \{a \in A \mid f(a) = b\}$$

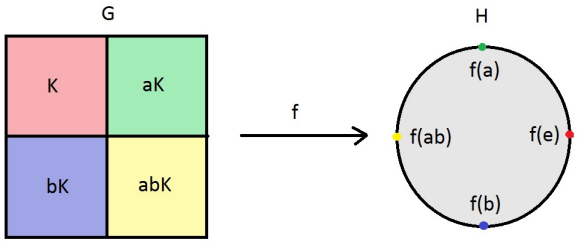
if f is not a surjection then $b \notin f(A)$ has $f^{-1}(b) = \emptyset$. I'm primarily interested in the nontrivial fibers. For an injective map the fibers are singletons as well; $a_1, a_2 \in f^{-1}(b)$ implies $f(a_1) = b = f(a_2)$ hence $a_1 = a_2$ if f is injective. However, when f is not injective the fibers can be larger, sometimes very large. For example, $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(x, y) = y$ has fibers of the form $\mathbb{R} \times \{y\}$. In any event, you can prove the fibers partition the domain: for each $a \in A$ we have $a \in f^{-1}(f(a))$ and if $a \in f^{-1}(b_1)$ and $a \in f^{-1}(b_2)$ then $f(a) = b_1$ and $f(a) = b_2$, but, f is a function and hence $b_1 = b_2$. We've shown fibers cover the domain and are disjoint. You could look at this as an equivalence relation, two points in the domain are related if they map to the same point in the range. Finally, to obtain injectivity, we simply select just one point from each fiber and restrict the function to this **section** of the domain. Let C denote the section then the map $\bar{f} : C \to f(A)$ defines a **bijection** where $\bar{f}(x) = f(x)$ for each $x \in C$.



For a homomorphism $\phi : G_1 \to G_2$ the fibers are precisely the cosets of the kernel. Suppose $x \in \phi^{-1}(p)$ where $p \in \phi(G_1)$ then we may argue $\phi^{-1}(p) = x\text{Ker}(\phi)$. Here's how:

If $z \in x\text{Ker}(\phi)$ then $z = xy$ where $y \in \text{Ker}(\phi)$ and $\phi(z) = \phi(xy) = \phi(x)\phi(y) = pe_2 = p$ hence $z \in \phi^{-1}(p)$ and we find $x\text{Ker}(\phi) \subseteq \phi^{-1}(p)$. Conversely, if $z \in \phi^{-1}(p)$ then $\phi(z) = p = \phi(x)$ thus $\phi(x)^{-1}\phi(z) = e_2$ and we find $\phi(x^{-1}z) = e_2$ which gives $x^{-1}z \in \text{Ker}(\phi)$. Hence $x^{-1}z\text{Ker}(\phi) = \text{Ker}(\phi)$ which gives $z\text{Ker}(\phi) = x\text{Ker}(\phi)$. But, $z \in z\text{Ker}(\phi)$ hence $z \in x\text{Ker}(\phi)$ and we conclude $\phi^{-1}(p) \subseteq x\text{Ker}(\phi)$ and thus $\phi^{-1}(p) = x\text{Ker}(\phi)$.

For general functions there is no reason the fibers should be so nicely related. The fact that the fibers of a homomorphism are cosets of the kernel of the map is very special. Consider $f : G \to H$ which is a homomorphism with $\text{Ker}(f) = K$. Suppose $[G : K] = 4$ and the distinct cosets are K, aK, bK, abK. There are just 4 distinct points in the codomain H which are attained by f: $f(a), f(b), f(ab)$ and $f(e) = e_H$. We can picture this data as follows:



The pictured map envisions $f : G \to H$ as being far from surjective. We propose $\bar{f}(gK) = f(g)$ for $gK = K, aK, bK, abK$. It's easy to see this makes $\bar{f} : G/K \to f(G)$ an homomorphism: for example:

$$\bar{f}(aK)\bar{f}(bK) = f(a)f(b) = f(ab) = \bar{f}(abK).$$

It is clear $\bar{f}$ is onto $f(A) = \{e_H, f(a), f(b), f(ab)\}$ and $\text{Ker}(\bar{f}) = K$ hence $\bar{f} : G/K \to f(A)$ is an isomorphism. Of course, we proved the First Isomorphism Theorem in detail in the last Lecture, so I'll refer you to Theorem 2.5.12 for a complete proof. This discussion is merely to attempt some intuition for why it works.

Example 2.6.1. *Consider $\phi : GL(n, \mathbb{R}) \to \mathbb{R}^\times$ defined by $\phi(A) = det(A)$ for each $A \in GL(n, \mathbb{R})$. Note, $det(AB) = det(A)det(B)$ hence ϕ is a homomorphism. Moreover, $Ker(\phi) = \{A \in GL(n, \mathbb{R}) \mid \phi(A) = 1\} = SL(n, \mathbb{R})$. Furthermore, $A = E_{11}(k-1) + I$ clearly has $det(A) = k$ for any $k \in \mathbb{R}^\times$ thus ϕ is a surjection. Hence, by the first isomorphism theorem,*

$$GL(n, \mathbb{R})/SL(n, \mathbb{R}).$$

Of course, there is little reason to keep it real in the example above. We could just as well replace $\mathbb{R}$ with another field $\mathbb{F}$.

Example 2.6.2. *Consider $Id : G \to G$ where $Id(x) = x$ for each $x \in G$. Clearly Id is a surjection with $Ker(Id) = \{e\}$ and we find $G/\{e\} \approx G$.*

The set $G/\{e\} = \{g\{e\} \mid g \in G\} = \{\{g\} \mid g \in G$ is rather silly. We mentioned $\phi(\{g\}) = g$ gives the isomorphism of $G/\{e\}$ and G directly. I added the Example above to show you another way to think about this result. Likewise,

Example 2.6.3. *Let G be a group and define $\phi : G \to G$ by $\phi(g) = e$ for each $g \in G$. Clearly $\phi(G) = \{e\}$ and $Ker(\phi) = G$ hence $G/G \approx \{e\}$ by the first isomorphism theorem.*

Again, the isomorphism $G/G \approx \{e\}$ can be established by any number of easy arguments beside the one I offer above. For example, G/G has one coset thus is a group with one element hence is isomorphic to the trivial group $\{e\}$.

Example 2.6.4. *Let $G = \langle a \rangle$ where $|a| = \infty$ then define $\phi : \mathbb{Z} \to G$ by $\phi(m) = a^m$. Note,*

$$\phi(x + y) = a^{x+y} = a^x a^y = \phi(x)\phi(y)$$

and $\phi(x) = a^x = e$ iff $x = 0$ hence $Ker(\phi) = \{0\}$. As ϕ is clearly a surjection we have $\mathbb{Z}/\{0\} \approx G$. But, we know $\mathbb{Z}/\{0\} \approx \mathbb{Z}$ hence $G \approx \mathbb{Z}$.

Next, consider cyclic groups of finite order.

Example 2.6.5. *Let $G = \langle a \rangle$ where $|a| = n \in \mathbb{N}$. Define $\phi : \mathbb{Z} \to G$ by $\phi(k) = a^k$. Note*

$$\phi(j+k) = a^{j+k} = a^j a^k = \phi(j)\phi(k)$$

hence ϕ is homomorphism. Also,

$$\phi(k) = a^k = e \quad \Rightarrow \quad n \mid k \quad \Rightarrow \quad Ker(\phi) = n\mathbb{Z}$$

If $a^k \in G$ then $\phi(k) = a^k$ thus $\phi(\mathbb{Z}) = G$. Hence the first isomorphism theorem provides $\mathbb{Z}/n\mathbb{Z} \approx G$.

It might be helpful to revisit the example above in the case that $G = \mathbb{Z}_n$ explicitly:

Example 2.6.6. *Define $\phi : \mathbb{Z} \to \mathbb{Z}_n$ by $\phi(x) = [x]_n$ where $[x]_n = x + n\mathbb{Z}$. Observe,*

$$\phi(x+y) = [x+y]_n = [x]_n + [y]_n$$

hence ϕ is a homomorphism. Furthermore, it is clear ϕ is a surjection since $[k]_n = \phi(k)$ for $k = 0, 1, \ldots, n-1$. Also, $\phi(x) = [0]_n$ iff $x \in n\mathbb{Z}$ which is to say $Ker(\phi) = n\mathbb{Z}$. We conclude, by the first isomorphism theorem,

$$\mathbb{Z}/n\mathbb{Z} \approx \mathbb{Z}_n$$

Of course, this should not be surprising as our very construction of $\mathbb{Z}_n$ was to parse $\mathbb{Z}$ into cosets of $n\mathbb{Z}$. In fact, we can replace $\approx$ with $=$ and write $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n$ given our construction of $\mathbb{Z}_n$.

It is not usually the case we can replace $\approx$ with $=$. In this course I use equality to indicate the objects are the same as point-sets. Isomorphism is a looser requirement. For example, $U(8) \approx \mathbb{Z}_2 \times \mathbb{Z}_2$ but $U(8) \neq \mathbb{Z}_2 \times \mathbb{Z}_2$.

Example 2.6.7. *Let $\phi : \mathbb{R} \to \mathbb{C}^\times$ be defined by*

$$\phi(\theta) = \cos\theta + i\sin\theta$$

for $\theta \in \mathbb{R}$. It can be shown with some trigonometry that $\phi(\theta + \beta) = \phi(\theta)\phi(\beta)$ hence ϕ is a homomorphism. Moreover, $\phi(\theta) = 1$ only if both $\cos\theta = 1$ and $\sin\theta = 0$ hence $\theta = 2\pi k$ for some $k \in \mathbb{Z}$. That is,

$$Ker(\phi) = \{2\pi k \mid k \in \mathbb{Z}\} = 2\pi\mathbb{Z}\}$$

Notice[18]*, $|\phi(\theta)|^2 = |\cos\theta + i\sin\theta|^2 = \cos^2\theta + \sin^2\theta = 1$. We find the image of ϕ falls on the collection of points in $\mathbb{C}^\times$ which are distance one from the origin. In other words, $\phi(\mathbb{R}) = S^1$ the* **unit-circle** *in the complex plane. To see ϕ is onto S^1 simply pick any point on S^1, calculate its standard angle θ and notice that $\phi(\theta)$ is precisely the point in question. Thus $\mathbb{R}/2\pi\mathbb{Z} \approx S^1$ by the first isomorphism theorem. In this case, we can visualize the homomorphism by imagining wrapping $\mathbb{R}$ around S^1 over and over again. As we wind 2π-length of the number line we arrive back at the same point once more. It follows that all the points which differ by an integer-multiple of 2π end up over the same point on S^1. The process of geometrically identifying standard angles which differ by a multiple of 2π radians is precisely the concept of working with the quotient $\mathbb{R}/2\pi\mathbb{Z}$.*

In other words, the reason angles are a little tricky is that the proper understanding of them necessitates the concept of a coset. A geometric angle is not really a single number, it's a whole collection of numbers each of which gives the same geometric direction... such numbers must differ by $2\pi k$ for some $k \in \mathbb{Z}$ in the case we use radians. For degree-based geometry we identify angles which differ by $360^o k$ for some $k \in \mathbb{Z}$.

[18] in case you never had the talk, $|x+iy| = \sqrt{x^2+y^2}$ is called the **modulus** of $x+iy$. The modulus is simply the distance from the origin to $x+iy$ in the complex plane

Example 2.6.8. *Suppose G, H are groups and form the direct product*

$$G \times H = \{(g, h) \mid g \in G, h \in H\}$$

Define **projections** $\pi_1 : G \times H \to G$ *and* $\pi_2 : G \times H \to H$ *by*

$$\pi_1(x, y) = x \qquad \& \qquad \pi_2(x, y) = y$$

for each $(x, y) \in G \times H$*. We calculate,*

$$Ker(\pi_1) = \{e_G\} \times H \qquad \& \qquad Ker(\pi_2) = G \times \{e_H\}$$

since the projections are clearly surjective the first isomorphism theorem provides:

$$\frac{G \times H}{\{e_G\} \times H} \approx G \qquad \& \qquad \frac{G \times H}{G \times \{e_H\}} \approx H$$

I asked you to prove a general version of this result in Problem 56.

I recommend working on Chapter 10 Gallian problems such as:

$$\#11 - 36.$$

You may need to consult Chapter 8 for some of these, and keep in mind the notation $\oplus_{Gallian} = \times_{me}$.

Problems for Lecture 15: (these are collected at the start of Lecture 17)

Problem 57: Prove that a nontrivial surjective homomorphism from $\mathbb{Z}_n$ to $\mathbb{Z}_k$ requires that $k \mid n$. *this problem modified from its original form, I restate the problem as Problem 72*

Problem 58: Let $m\mathbb{Z}_n = \{mx \mid x \in \mathbb{Z}_n\}$. Give a condition under which $\mathbb{Z}_n/m\mathbb{Z}_n \approx \mathbb{Z}_m$ (you supply the condition and the proof these are isomorphic). *Hint: $\psi : \mathbb{Z}_n \to \mathbb{Z}_m$ defined by $\psi([x]_n) = [x]_m$ has kernel $m\mathbb{Z}_n$*

Problem 59: Show that $\phi(\theta) = \cos\theta + i\sin\theta$ defines a homomorphism from $\mathbb{R}$ under addition to $\mathbb{C}^\times$. Also, prove that $S^1 = \{\cos\theta + i\sin\theta \mid \theta \in \mathbb{R}\} \leq \mathbb{C}^\times$.

Problem 60: Gallian problem 7 on page 169. *yes, I know I sent you a link to a page with multiple solutions to this, the point here is not that you can find an answer, the point is for you to make one of those proofs your own... own it.*

2.7 Lecture 16: direct products inside and outside

My goal in this section is to show the basic interaction between external and internal direct products and to prove the major results from Chapter 8 of Gallian (although, I will be translating his $\oplus$ to our $\times$). Let me remind the reader, the external direct product of H and K is given by:

$$G' = H \times K = \{(h,k) \mid h \in H, k \in K\}$$

whereas the internal direct product assumes $H, K \trianglelefteq G$ with $H \cap K = \{e\}$ and

$$G = H \oplus K = \{hk \mid h \in H, k \in K\}$$

Clearly $G \neq G'$ since $H \times K \subset G \times G$ whereas $H \oplus K \subset G$.

Why do we require the normality of the subgroups forming the internal direct product? Consider the following example:

Example 2.7.1. *Consider* $D_3 = \{1, x, x^2, y, xy, x^2y \mid x^3 = 1, y^2 = 1, (xy)^2 = 1\}$. *Let* $H = \langle x \rangle = \{1, x, x^2\}$ *and* $K = \{1, y\}$ *then* $HK = D_3$ *and* $H \cap K = \{1\}$. *We have* $H \trianglelefteq D_3$ *however* $xK = \{x, xy\}$ *and* $Kx = \{x, yx\} = \{x, x^2y\}$ *thus* $xK \neq Kx$ *which shows* K *is not a normal subgroup of* D_3. *Let's study the external direct product of* H *and* K*:*

$$H \times K = \{(1,1), (1,y), (x,1), (x,y), (x^2,1), (x^2,y)\} \leq D_3 \times D_3$$

It happens that $|HK| = |H \times K| = 6$. *However, these are not isomorphic. Notice, since* $(x,y)^n = (x^n, y^n)$ *and* $x^3 = 1$ *and* $y^2 = 1$ *we have:*

$$(x,y)^2 = (x^2,1), \quad (x,y)^3 = (1,y), \quad (x,y)^4 = (x,1), \quad (x,y)^5 = (x^2,y), \quad (x,y)^6 = (1,1)$$

thus (x,y) *is an element or order* 6 *in* $H \times K$ *which indicates* $H \times K \approx \mathbb{Z}_6$. *We know* D_3 *is not cyclic thus* $D_3 = HK \not\approx H \times K$.

We wish for $H \times K$ to be isomorphic to $H \oplus K$. We show now this is indeed the case. But, first we need a lemma:

Lemma 2.7.2. *Suppose* $H, K \trianglelefteq G$ *and* $H \cap K = \{e\}$.

- **(i.)** *if* $ab = a'b'$ *where* $a, a' \in H$ *and* $b, b' \in K$ *then* $a = a'$ *and* $b = b'$,
- **(ii.)** *if* $a \in H$ *and* $b \in K$ *then* $ab = ba$.

Proof: let H, K be normal subgroups with $H \cap K = \{e\}$. Suppose $a, a' \in H$ and $b, b' \in K$ with $ab = a'b'$. Notice, $bb'^{-1} \in K$ and $a^{-1}a' \in H$. Multiplying on the left by a^{-1} and on the right by b'^{-1} we derive $bb'^{-1} = a^{-1}a'$ which shows $bb'^{-1}, a^{-1}a' \in H \cap K$. However, $H \cap K = \{e\}$ so we find $bb'^{-1} = e$ and $a^{-1}a' = e$ which yield $a = a'$ and $b = b'$ which completes the proof of (i.)

Suppose $a \in H$ and $b \in K$. To show $ab = ba$ we must show $aba^{-1}b^{-1} = e$. Notice, $K \trianglelefteq G$ provides that $aba^{-1} \in K$ once more as $gKg^{-1} \subseteq K$. Hence, $(aba^{-1})b^{-1} \in K$ as it is the product of $aba^{-1}, b^{-1} \in K$. Likewise, by the normality of H we find $ba^{-1}b^{-1} \in H$ once more and hence $a, ba^{-1}b^{-1} \in H$ and thus the product $a(ba^{-1}b^{-1}) \in H$. Hence $aba^{-1}b^{-1} \in H \cap K$ which shows $aba^{-1}b^{-1} = e$ thus $ab = ba$ and this completes the proof of (ii.). $\square$

Theorem 2.7.3. *If* $G = H \oplus K$ *then* $G \approx H \times K$. *Here the notation* $H \oplus K = HK = \{hk \mid h \in H, k \in K\}$ *where* $H, K \trianglelefteq G$ *and* $H \cap K = \{e\}$.

Proof: assume $H, K \trianglelefteq G$ and $H \cap K = \{e\}$ and $HK = G$. Define $\phi(x, y) = xy$ for each $(x, y) \in H \times K$. Consider,

$$\phi((x, y)(a, b)) = \phi((xa, yb)) = (xa)(yb) = (xy)(ab) = \phi(x, y)\phi(a, b)$$

where we used part (ii.) of Lemma 2.7.2 to commute $a \in H$ with $y \in K$. Hence ϕ is a homomorphism from $H \times K$ to G. If $g \in G = HK$ then $g = hk$ for some $h \in H$ and $k \in K$ thus $\phi(h, k) = hk = g$ and we find $\phi(H \times K) = G$. Next we study the kernel of ϕ. Suppose $\phi(x, y) = e$ thus $xy = e = ee$ and by part (i.) of Lemma 2.7.2 we find $x = e$ and $y = e$ hence $\mathrm{Ker}(\phi) = \{(e, e)\}$ which completes the proof that $\phi : H \times K \to H \oplus K$ is an isomorphism; $H \times K \approx H \oplus K$. □

Theorem 2.7.4. *If $|x|, |y|$ are finite then $(x, y) \in G \times H$ has order $|(x, y)| = lcm(|x|, |y|)$.*

Proof: Suppose $|x| = m$ and $|y| = n$ hence $x^m = e$ and $y^n = e$. Let $\mathrm{lcm}(|x|, |y|) = s$ hence $s = km$ and $s = ln$ for some $l, k \in \mathbb{N}$. Calculate,

$$(x, y)^s = (x^s, y^s) = (x^{km}, y^{ln}) = ((x^m)^k, (y^n)^l) = (e^k, e^l) = (e, e).$$

Thus $|(x, y)| \leq s$. Suppose $(x, y)^j = (e, e)$ for some $j < s$ then,

$$(x, y)^j = (x^j, y^j) = (e, e)$$

hence $x^j = e$ and $y^j = e$ thus $m \mid j$ and $n \mid j$ and if $j < s$ we have a common multiple of m and n which is smaller than the least common multiple. Of course, that is absurd, hence no such j exists and we conclude $|(x, y)| = s = \mathrm{lcm}(|x|, |y|)$. □

Example 2.7.5. *We noticed that $(x, y) \in D_3$ has order 6. This makes sense as $|x| = 3$ and $|y| = 2$ hence $|(x, y)| = lcm(2, 3) = 6$.*

Since $|a| = |\langle a \rangle|$ we can say much about subgroups. If $\langle a \rangle$ has order m and $\langle b \rangle$ has order n then $\langle (a, b) \rangle$ has order $\mathrm{lcm}(m, n)$.

Example 2.7.6. *Observe $\langle 10 \rangle$ is a group of order 4 in $\mathbb{Z}_{40}$ and $\langle 3 \rangle$ is a group of order 10 in $\mathbb{Z}_{30}$ hence $\langle (4, 3) \rangle$ is a subgroup of order 20 in $\mathbb{Z}_{40} \times \mathbb{Z}_{30}$.*

Example 2.7.7. *How many elements of order 6 are there in $\mathbb{Z}_6 \times \mathbb{Z}_2$? We obtain 6 as the least common multiple of 6 and 1 or as 3 and 2. In $\mathbb{Z}_6$ we have 1 and 5 with order 6 and the identity 0 is the element of order 1 in $\mathbb{Z}_2$. Thus, $(1, 0), (5, 0)$ have order 6. Next, the elements of order 3 in $\mathbb{Z}_6$ are precisely 2 and 4. Of course 1 is the only element of order 2 in $\mathbb{Z}_2$, thus $(2, 1), (4, 1)$ have order 6. In total,*

$$(1, 0), (5, 0), (2, 1), (4, 1)$$

are the elements of order 6 in $\mathbb{Z}_6 \times \mathbb{Z}_2$. You can see $\langle (1, 0) \rangle = \langle (5, 0) \rangle$ and $\langle (2, 1) \rangle = \langle (4, 1) \rangle$. In particular,

$$\langle (2, 1) \rangle = \{(2, 1), (4, 0), (0, 1), (2, 0), (4, 1), (0, 0)\}$$

Cyclic subgroups of order 6 come with $\phi(6) = 2$ generators each. So, to count the number of subgroups of order 6 we have to divide 4 by 2.

I'll work on a slight twist of Example 3 on page 152.

Example 2.7.8. *Find the number of elements of order 7 in $\mathbb{Z}_{49} \times \mathbb{Z}_7$. For $|(a, b)| = lcm(|a|, |b|) = 7$ we have 3 distinct cases:*

(i.) $|a| = 7$, $|b| = 1$ *note* $\langle a \rangle$ *is cyclic group of order* 7 *hence there are* 6 *generators. In total we have 6 elements of order* 7 *in this case.*

(ii.) $|a| = 1$, $|b| = 7$. *Again,* $\langle b \rangle$ *contains* 6 *generators and this provides us* 6 *elements of order* 7.

(iii.) $|a| = 7$, $|b| = 7$. *We have 6 choices for* a *and* b *hence* 36 *total choices for elements of order* 7 *formed in this fashion*

In summary, $6 + 6 + 36 = 48$ *elements of order* 7.

Example 2.7.9. *Gallian explains how* $\mathbb{Z}_{100} \times \mathbb{Z}_{25}$ *has* 24 *distinct elements of order* 10. *Then, since every cyclic subgroup of order* 10 *has* $\phi(10) = 4$ *generators. It follows there are* $24/4 = 6$ *distinct subgroups of order* 10.

I enjoy how Gallian explains this counting in terms of counting the number of legs of sheep then dividing by 4 to count sheep. So, I guess if we're counting subgroups of order 20 then as $\phi(20) = \phi(4)\phi(5) = 2(4) = 8$ we should think of the elements of order 20 like legs of a spider. I'm not sure I know enough zoology[19] to generalize this method.

Theorem 2.7.10. *If* G *and* H *are finite cyclic groups then* $G \times H$ *is cyclic if and only if* $|G|$ *and* $|H|$ *are relatively prime.*

Proof: suppose G and H have order m, n respective with $\gcd(m, n) = 1$. Note, $\text{lcm}(m, n) = mn$ in this case. Futher, if $G = \langle a \rangle$ and $H = \langle b \rangle$ then $|a| = m$ and $|b| = n$. Hence, $|\langle (a, b) \rangle| = mn$ by Theorem 2.7.15. But, $|G \times H| = mn$ by counting hence $G \times H = \langle (a, b) \rangle$.

Conversely, suppose $G \times H$ is cyclic and $|G| = m$ and $|H| = n$. Let $d = \gcd(m, n)$. Suppose (x, y) is a generator of $G \times H$. Notice, $x^m = e$ and $y^n = e$, hence:

$$(x, y)^{mn/d} = ((x^m)^{n/d}, (y^n)^{m/d}) = (e^{n/d}, e^{m/d}) = (e, e).$$

Thus, $mn = |(x, y)| \leq mn/d$ from which we find $d = 1$ hence m, n are relatively prime. □

Corollary 2.7.11. $\mathbb{Z}_{mn} \approx \mathbb{Z}_m \times \mathbb{Z}_n$ *iff* m, n *are relatively prime.*

Proof: apply Theorem 2.7.10. □

We should generalize to products of more than two groups.

Definition 2.7.12. *If* $H_1, H_2, \ldots, H_n \trianglelefteq G$ *and*

(i.) $G = H_1 H_2 \cdots H_n = \{x_1 x_2 \cdots x_n \mid x_i \in H_i, i = 1, 2, \ldots, n\}$

(ii.) $(H_1 \cdots H_i) \cap H_{i+1} = \{e\}$ *for* $i = 1, 2, \ldots, n-1$.

then we say $G = H_1 \oplus H_2 \oplus \cdots \oplus H_n$.

Condition (i.) provides that the product of the subgroups generate the entire group. Condition (ii.) provides the subgroups are independent. Once again, this definition is made so that the internal direct product be isomorphic to the external direct product which is defined in the natural fashion (see Definition 2.4.15). Lemma 2.7.2 generalizes nicely:

[19] and you thought I'd never work this in

Lemma 2.7.13. *$G = H_1 \oplus H_2 \oplus \cdots \oplus H_n$*

(i.) *if $a_1a_2\cdots a_n = x_1x_2\cdots x_n$ where $a_i, x_i \in H_i$ for each $i = 1, 2, \ldots, n$ then $a_i = x_i$ for each $i = 1, 2, \ldots, n$*

(ii.) *if $a_i \in H_i$ for $i = 1, 2, \ldots, n$ then $a_1a_2 \ldots a_n = a_{\sigma(1)}a_{\sigma(2)} \ldots, a_{\sigma(n)}$ for any $\sigma \in S_n$.*

Proof: notice the $n = 2$ case was shown Lemma 2.7.2 hence $a_1a_2 = x_1x_2$ for $a_1, x_1 \in H_1$ and $a_2, x_2 \in H_2$ where $H_1 \cap H_2 = \{e\}$ provides $a_1 = x_1$ and $a_2 = x_2$ and $a_1a_2 = a_2a_1$. For $n = 3$, note $H_1H_2 \trianglelefteq H_1H_2H_3$ and $H_1H_2 \cap H_3 = \{e\}$ hence we apply Lemma 2.7.2 using $H = H_1H_2$ and $K = H_3$; if $(a_1a_2)a_3 = (x_1x_2)x_3$ where $a_i, x_i \in H_i$ then we deduce $a_1a_2 = x_1x_2$ and $a_3 = x_3$. Moreover,

$$(a_1a_2)a_3 = a_3(a_1a_2).$$

But, recalling our work from $n = 2$ we have:

$$a_1 = x_1, \;\; a_2 = x_2, \;\; a_3 = x_3,$$

$$(a_1a_2)a_3 = a_3(a_1a_2) = (a_2a_1)a_3 = a_3(a_2a_1), \quad a_1a_2 = a_2a_1, \;\; a_2a_3 = a_3a_2, \quad a_1a_3 = a_3a_1.$$

In other words, $a_1a_2a_3 = a_{\sigma(1)}a_{\sigma(2)}a_{\sigma(3)}$ for any σ. Suppose inductively (i.) and (ii.) hold for all $n = 1, 2, \ldots, m$. If $a_i, x_i \in H_i$ for $i = 1, 2, \ldots, m+1$ and

$$a_1a_2\cdots a_ma_{m+1} = x_1x_2\cdots x_mx_{m+1}$$

then applying Lemma 2.7.2 to the normal subgroup $H = H_1H_2\cdots H_m$ and $K = H_{m+1}$ we find

$$a_1a_2\cdots a_m = x_1x_2\cdots x_m \quad \& \quad a_{m+1} = x_{m+1}$$

and

$$(a_1a_2\cdots a_m)a_{m+1} = a_{m+1}(a_1a_2\cdots a_m)$$

Then, by the induction hypothesis,

$$a_1 = x_1, a_2 = x_2, \ldots, a_m = x_m, a_{m+1} = x_{m+1} \quad \& \quad a_1a_2\cdots a_m = a_{\sigma(1)}a_{\sigma(2)}\cdots a_{\sigma(m)}$$

for all $\sigma \in S_m$. I leave it to the reader to complete the proof that (ii.) holds for the $n = m+1$ case hence by induction the Lemma follows. □

The Lemma above makes easy work of the Theorem to follow:

Theorem 2.7.14. *$G = H_1 \oplus H_2 \oplus \cdots \oplus H_n$ then $G \approx H_1 \times H_2 \times \cdots \times H_n$.*

Proof: let $\phi(x_1, \ldots, x_n) = x_1\cdots x_n$ define a map from $H_1 \times \cdots \times H_n$ to $G = H_1 \oplus \cdots \oplus H_n$. Consider,

$$\begin{aligned}
\phi((a_1, a_2, \ldots, a_n)(x_1, x_2, \ldots, x_n)) &= \phi((a_1x_1, a_2x_2, \ldots, a_nx_n)) \\
&= (a_1x_1)(a_2x_2)\cdots(a_n)(x_n) \\
&= (a_1a_2\cdots a_n)(x_1x_2\cdots x_n) \\
&= \phi((a_1, a_2, \ldots, a_n))\phi((x_1, x_2, \ldots, x_n))
\end{aligned}$$

Thus ϕ is a homomorphism. Moreover, ϕ is surjective since each $g \in G = H_1 \oplus \cdots \oplus H_n$ can be written as $g = a_1\cdots a_n$ and $\phi(a_1, \ldots, a_n) = g$. Part (ii.) of Lemma 2.7.13 provides injectivity since:

$$\begin{aligned}
\phi(a_1, a_2, \ldots, a_n) = \phi(x_1, x_2, \ldots, x_n) \quad &\Rightarrow \quad a_1a_2\cdots a_n = x_1x_2\cdots x_n \\
&\Rightarrow \quad a_1 = x_1, a_2 = x_2, \ldots, a_n = x_n.
\end{aligned}$$

Thus ϕ is a bijective homomorphism and we conclude $H_1 \oplus H_2 \oplus \cdots \oplus H_n \approx H_1 \times H_2 \times \cdots \times H_n$. □

There is also a natural generalization of the order of an element theorem we had for 2-tuples. I'll skip the proof of this Theorem since it is nearly identical to the $n = 2$ case.

Theorem 2.7.15. *If $|x_1|, |x_2|, \ldots, |x_n|$ are finite then $(x_1, x_2, \ldots, x_n) \in H_1 \times H_2 \times \cdots \times H_n$ has order $|(x_1, x_2, \ldots, x_n)| = lcm(|x_1|, |x_2|, \ldots, |x_n|)$.*

Example 2.7.16. *Consider $G = \mathbb{Z}_6 \times \mathbb{Z}_3 \times \mathbb{Z}_4$. How elements of order 6 and how many cyclic subgroups of order 6 ? Considering $(a, b, c) \in G$ has $|(a, b, c)| = lcm(|a|, |b|, |c|)$ we have the following cases:*

(1.) $|a| = 6, |b| = 1, 3, |c| = 1, 2$**:** *Let $\#(a)$ denote the number of choices for a with $|a| = 6$ in $\mathbb{Z}_6$ we have $\#(a) = 2$. For $|b| = 1, 3$ in $|ZN_3$ we obtain all of $\mathbb{Z}_3$ hence $\#(b) = 3$. On the other hand $\#(c) = 2$ since $c = 0, 2$ have $|c| = 1, 2$ in $\mathbb{Z}_4$. In total, the number of elements of order 6 in this case are $\#(a)\#(b)\#(c) = (2)(3)(2) = 12$.*

(2.) $|a| = 3$**,** $|b| = 1$**,** $|c| = 2$**:** *we determine $\#(a)\#(b)\#(c) = (2)(1)(1) = 2$*

(3.) $|a| = 1$**,** $|b| = 3$**,** $|c| = 2$**:** *we determine $\#(a)\#(b)\#(c) = (1)(2)(1) = 2$*

In total, there are $12 + 2 + 2 = 16$ elements of order 6 in $\mathbb{Z}_6 \times \mathbb{Z}_3 \times \mathbb{Z}_4$. It follows there are 8 cyclic subgroups of order 6 since each cyclic subgroup of order 6 has 2 generators.

There is also a nice generalization of Theorem 2.7.10 to three or more factors.

Theorem 2.7.17. *If $H_1, H_2, \ldots, H_n$ are cyclic groups of finite order then $H_1 \times H_2 \times \cdots \times H_n$ is cyclic if and only if $|H_i|, |H_j|$ is relatively prime whenever $i \neq j$.*

Proof: exercise for reader. □

Corollary 2.7.18. *$\mathbb{Z}_{n_1 n_2 \ldots n_k} \approx \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$ if and only if $gcd(n_i, n_j) = 1$ whenever $i \neq j$*

Proof: since $|\mathbb{Z}_n| = n$ and $\mathbb{Z}_n$ is cyclic we find $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$ is cyclic if and only if $\gcd(n_i, n_j) = 1$ whenever $i \neq j$ by Theorem 2.7.17. Moreover, as the order of $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$ is given by $|\mathbb{Z}_{n_1}||\mathbb{Z}_{n_2}| \cdots |\mathbb{Z}_{n_k}| = n_1 n_2 \cdots n_k$ we determine $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_k}$ is isomorphic to a cyclic group of order $n_1 n_2 \cdots n_k$ and the Corollary follows. □

Example 2.7.19. *Since $105 = 3 \cdot 5 \cdot 7$ (relatively prime factors) we have $\mathbb{Z}_{105} \approx \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_7$.*

Example 2.7.20. *Consider $\mathbb{Z}_{20}$ since 2 and 10 are not relatively prime it is not the case that $\mathbb{Z}_{20} \approx \mathbb{Z}_2 \times \mathbb{Z}_{10}$. On the other hand, $20 = 4 \cdot 5$ and $gcd(4, 5) = 1$ hence $\mathbb{Z}_{20} \approx \mathbb{Z}_4 \times \mathbb{Z}_5$.*

2.7.1 classification of finite abelian groups

I choose to not prove this result in the first semester. Honestly, it's part of a much larger story which includes the rational cannonical form and modules. The proper proof is significantly harder than most of what we've been up to in this course thus far and I'd rather invest our energy in other pursuits (like group actions and counting). That said, I should communicate the result we will likely prove next semester. There is a proof in Chapter 11 of Gallian if you cannot wait until Math 422.

Theorem 2.7.21. *Every finite abelian group is the direct product of cyclic groups of prime-power order. Moreover, the number of terms in the product and the orders of the cyclic groups are uniquely determined by the group.*

It is helpful to discuss a generic prime p to appreciate the content of the Theorem.

order	isomorphism classes
p	$\mathbb{Z}_p$
p^2	$\mathbb{Z}_{p^2}$ $\mathbb{Z}_p \times \mathbb{Z}_p$
p^3	$\mathbb{Z}_{p^3}$ $\mathbb{Z}_{p^2} \times \mathbb{Z}_p$ $\mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$

The listed cases are clearly not isomorphic. For example, $\mathbb{Z}_{p^2}$ is cyclic whereas $\mathbb{Z}_p \times \mathbb{Z}_p$ is not cyclic. Being formed from the product of cyclic groups is not the same as being cyclic!

Example 2.7.22. *What are the possible types of abelian groups of order* 100 *up to isomorphism? Well, notice* $100 = 2^2 \cdot 5^2$ *hence we have the following choices:*

$$\mathbb{Z}_4 \times \mathbb{Z}_{25} \approx \mathbb{Z}_{100},$$
$$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{25} \approx \mathbb{Z}_2 \times \mathbb{Z}_{50},$$
$$\mathbb{Z}_4 \times \mathbb{Z}_5 \times \mathbb{Z}_5 \approx \mathbb{Z}_{20} \times \mathbb{Z}_5,$$
$$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_5 \times \mathbb{Z}_5 \approx \mathbb{Z}_{10} \times \mathbb{Z}_{10}$$

where I have made ample use of Corollary 2.7.11.

I recommend working on Chapter 11 Gallian problems such as:

$$\#11 - 36.$$

also Chapter 8 Gallian problems such as:

$$\#4 - 63$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know.

Problems for Lecture 16: (these are collected at the start of Lecture 18)

Problem 61: Prove the product $H_1 H_2 \cdots H_k$ of normal subgroups $H_1, \ldots, H_k$ of a group G is once more a normal subgroup.

Problem 62: Gallian Chapter 8 #32 from page 163

Problem 63: Gallian Chapter 8 #41 from page 163.

Problem 64: Gallian Chapter 11 #8 from page 219.

2.8 Lecture 17: a little number theory and encryption

I claimed without proof a bit earlier in this course that[20] for a prime p and relatively prime s, t,

$$\phi(p) = p - 1 \qquad \phi(p^k) = p^k - p^{k-1} \qquad \& \qquad \phi(st) = \phi(s)\phi(t).$$

In this section I intend to investigate these identities via studying the $U(n)$. Note, $\phi(n) = |U(n)|$ since the group of units is the set of integers relatively prime to n under the operation of multiplication modulo n and the **Euler-phi-function** $\phi(n)$ counts the number of integers which are relatively prime to n and not larger than n.

Proposition 2.8.1. *If p is prime then $|U(p)| = p - 1$ hence $\phi(p) = p - 1$.*

Proof: if p is prime then $\gcd(p, j) = 1$ for $j = 1, 2, \ldots, p - 1$ hence $|U(n)| = p - 1$. □

Proposition 2.8.2. *If p is prime then $|U(p^k)| = p^k - p^{k-1}$ hence $\phi(p^k) = p^k - p^{k-1}$.*

Proof: Let $n = p^k$ where p is prime. Observe the divisors of n include 1 and multiples of p from p to $p^k = p^{k-1}p$. In particular the list has p^{k-1} members:

$$1, p, 2p, \ldots, pp, (p + 1)p, \ldots, p^{k-1}p$$

Thus there are p^{k-1} integers in $\mathbb{Z}_n$ which are not relatively prime to $n = p^k$. It follows that the remaining numbers in $\mathbb{Z}_n$ are relatively prime to n. Hence, $|U(n)| = p^k - p^{k-1} = \phi(p^k)$. □

The results above are mostly just counting and the definition of prime in $\mathbb{Z}$. Our next result is a bit more group-theoretic.

Theorem 2.8.3. *Suppose s, t are relatively prime then $U(st) \approx U(s) \times U(t)$*

Proof: consider the mapping $\psi([x]_{st}) = ([x]_s, [x]_t)$ for each $[x]_{st} \in U(st)$. We seek to show this gives an isomorphism from $U(st)$ to $U(s) \times U(t)$. Let us begin by showing ψ is single-valued. If $[x]_{st} = [y]_{st}$ then $y = x + j(st)$ hence

$$\psi([y]_{st}) = ([y]_s, [y]_t) = ([x + jst]_s, [x + jst]_t) = ([x]_s, [x]_t) = \psi([x]_{st}).$$

To show ψ is into $U(s) \times U(t)$ we need to demonstrate the inverse of $[x]_s, [x]_t$ exist whenever $[x]_{st}$ has a multiplicative inverse. Consider, $[x]_{st} \in U(st)$ implies there exists $[y]_{st}$ such that

$$xy - 1 = n(st)$$

for some $n \in \mathbb{Z}$ hence $xy - 1 = (nt)s$ and $xy - 1 = (ns)t$ hence $[x]_s[y]_s = [1]_s$ and $[x]_t[y]_t = [1]_t$ which shows ψ is into. The homomorphism property of ψ follows naturally from modular arithmetic:

$$\begin{aligned} \psi([x]_{st}[y]_{st}) &= \psi([xy]_{st}) = ([xy]_s, [xy]_t) \\ &= ([x]_s[y]_s, [x]_t[y]_t) \\ &= ([x]_s, [x]_t)([y]_s, [y]_t) \\ &= \psi([x]_{st})\psi([y]_{st}). \end{aligned}$$

Finally, since $U(st)$ and $U(s) \times U(t)$ are finite we show ψ is a bijection if we either show injectivity or surjectivity. Consider the kernel: $\mathrm{Ker}(\psi) = \{[x]_{st} \mid [x]_s = [1]_s \ \& \ [x]_t = [1]_t\}$. Thus $[x]_{st} \in \mathrm{Ker}(\psi)$ iff $x \equiv 1 (\mathrm{mod}\, s)$ and $x \equiv 1 (\mathrm{mod}\, t)$ for s, t relatively prime. By the **Chinese Remainder Theorem** there is a simultaneous solution to these congruences and $x \equiv 1 (\mathrm{mod}\, st)$. In other words, by the Chinese Remainder Theorem, we obtain $\mathrm{Ker}(\psi) = \{[1]_{st}\}$ and we find ψ is an isomorphism of $U(st)$ and $U(s) \times U(t)$. □

[20] I will take care to not use ϕ as an isomorphism in this section as it always means the euler-phi-function in this Lecture

Multiplicativity of the euler-phi-function follows easily:

Corollary 2.8.4. *If $s, t \in \mathbb{N}$ are relatively prime then $\phi(st) = \phi(s)\phi(t)$.*

Proof: by Theorem 2.8.7, if s, t are relatively prime then $U(st) \approx U(s) \times U(t)$. Hence $|U(st)| = |U(s)| \cdot |U(t)|$ but, $\phi(n) = |U(n)|$ thus $\phi(st) = \phi(s)\phi(t)$. □

I would wager not all of you are familar with the Chinese Remainder Theorem. In short, it gives us a result which links calculations with respect to several moduli. The simplest form of the Theorem is for just two moduli which are relatively prime. We used that result to prove Theorem 2.8.7. I'll state a bit more general version which we need to prove the extended version of the Theorem.

Theorem 2.8.5. *Suppose $m_1, m_2, \ldots, m_r$ have $gcd(n_i, n_j) = 1$ for $i \neq j$ then the system of congruences*

$$x \equiv a_1 \ (mod m_1), \ x \equiv a_2 \ (mod m_2), \ldots, x \equiv a_r \ (mod m_r)$$

has a unique solution modulo $M = m_1 m_2 \ldots m_r$

Proof: First we construct a solution of the system. Define $M_k = M/m_k$ for each $k = 1, 2, \ldots, r$. Observe $\gcd(M_k, m_k) = 1$ since by construction all the factors composing M_k are relatively prime to m_k. For each k, By Bezout's Theorem, the observation $\gcd(M_k, m_k) = 1$ earns the existence of $y_k, b_k \in \mathbb{Z}$ for which $y_k M_k + b_k m_k = 1$ thus $y_k M_k \equiv 1 \bmod m_k$. By math magic, consider:

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_r M_r y_r$$

clearly $M_j \equiv 0 \bmod m_i$ for each $i \neq j$ hence

$$x \equiv a_i M_i y_i \equiv a_i$$

modulo m_i. But, i was arbitrary hence x solves all r of the congruences. Suppose y is another solution of all the congruences then $x \equiv y \equiv a_i \bmod i$ for $i = 1, 2, \ldots, r$. Hence $m_i \mid (y - x)$ for each i and hence $M = m_1 m_2 \cdots m_r \mid (y - x)$ and we conclude $y \equiv x \bmod M$ as the Theorem claims. □

The proof above is more than a proof. It's a template for how to solve these multiple congruence problems. I use the notation of the proof to guide my calculation in the example below:

Example 2.8.6. *Solve $x \equiv 2$ mod 3 and $x \equiv 7$ mod 11. Observe $M = 3(11) = 33$ and $m_1 = 3$ whereas $M_1 = 11$ and conversely $m_2 = 11$ and $M_2 = 3$. We calculate,*

$$(mod\ 3): \quad 11^{-1} = 2 = y_1$$

$$(mod\ 11): \quad 3^{-1} = 4 = y_2$$

Hence, noting $a_1 = 2$ and $a_2 = 7$ we construct:

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 = 2(11)(2) + (7)(3)(4) = 44 + 84 = 128 \equiv 29 \ \ mod\ 33$$

as $128 - 3(33) = 29$. We can check that $x = 29$ is indeed congruent to 2 mod 3 and 7 mod 11.

The other way to solve these sort of problems is by substitution, but, I'm not trying to be general here at the moment. In number theory perhaps you saw other methods to solve such a problem as well as how we deal with the case the moduli are not relatively prime. These questions are interesting, but, I leave them for another time.

Theorem 2.8.7. *Suppose $m_1, m_2, \ldots, m_r$ have $gcd(m_i, m_j) = 1$ whenever $i \neq j$ then*

$$U(m_1 m_2 \cdots m_r) \approx U(m_1) \times U(m_2) \times \cdots \times U(m_r).$$

Proof: let $M = m_1 m_2 \cdots m_r$ and define

$$\psi([x]_M) = ([x]_{m_1}, [x]_{m_2}, \ldots, [x]_{m_r})$$

the proof that ψ is single-valued, into and a homomorphism are very similar to that given in Theorem 2.8.7. We'll examine the kernel calculation in detail. If $[x]_M \in \text{Ker}(\psi)$ then

$$([x]_{m_1}, [x]_{m_2}, \ldots, [x]_{m_r}) = ([1]_{m_1}, [1]_{m_2}, \ldots, [1]_{m_r})$$

which is to say

$$x \equiv 1 \bmod m_1, \quad x \equiv 1 \bmod m_2, \quad \ldots \quad , x \equiv 1 \bmod m_r$$

By the Theorem 2.8.5 the simultaneous solution $x = 1$ is unique modulo $m_1 m_2 \cdots m_r$ which is to say $\text{Ker}(\psi) = \{[1]_M\}$ and we conclude ψ is an isomorphism of $U(m_1 m_2 \cdots m_r)$ and $U(m_1) \times U(m_2) \times \cdots \times U(m_r)$. □

Corollary 2.8.8. *If $m_1, m_2, \ldots, m_r$ have $gcd(m_i, m_j) = 1$ for all $i \neq j$ then*

$$\phi(m_1 m_2 \ldots m_r) = \phi(m_1)\phi(m_2) \cdots \phi(m_r).$$

Proof: by Theorem 2.8.7 we observe for $m_1, m_2, \ldots, m_r$ with $\gcd(m_i, m_j) = 1$ we have

$$|U(m_1 m_2 \cdots m_r)| = |U(m_1) \times U(m_2) \times \cdots \times U(m_r)| = |U(m_1)||U(m_2)| \cdots |U(m_r|$$

thus $\phi(m_1 m_2 \cdots m_r) = \phi(m_1)\phi(m_2) \cdots \phi(m_r)$. □

Discussion: the classification of $U(n)$: the fundamental theorem of arithmetic states n can be expressed uniquely, up to reordering, as the product of prime powers. Denote,

$$n = p^{k_1} p^{k_2} \cdots p^{k_s}$$

By Theorem 2.8.7 we can decompose $U(n)$ into the product below:

$$U(n) \approx U(p^{k_1}) \times U(p^{k_2}) \times \cdots U(p^{k_s})$$

It can be shown, Gallian credits Gauss circa 1801 as one source, that $U(2) \approx \{0\}, U(4) \approx \mathbb{Z}_2$ and $U(2^n) \approx \mathbb{Z}_2 \times \mathbb{Z}_{2^{n-2}}$ for $n \geq 3$. Moreover, $U(p^n) \approx \mathbb{Z}_{p^n - p^{n-1}}$ for any odd prime power. With these results and the decomposition scheme above this means we can express any $U(n)$ as the direct product of copies of $\mathbb{Z}_n$.

Example 2.8.9.

$$U(200) = U(5^2 \cdot 2^3) \approx U(5^2) \times U(2^3) \approx \mathbb{Z}_{25-5} \times \mathbb{Z}_2 \times \mathbb{Z}_2 = \mathbb{Z}_{20} \times \mathbb{Z}_2 \times \mathbb{Z}_2$$

Example 2.8.10.

$$U(405) = U(5 \cdot 3^4) \approx U(5) \times U(3^4) \approx \mathbb{Z}_4 \times \mathbb{Z}_{81-27} = \mathbb{Z}_4 \times \mathbb{Z}_{54}$$

Example 2.8.11.

$$U(195) = U(5 \cdot 3 \cdot 13) \approx U(5) \times U(3) \times U(13) \approx \mathbb{Z}_4 \times \mathbb{Z}_2 \times \mathbb{Z}_{12}.$$

We should recall, Theorem 2.2.14 showed us the automorphisms of $\mathbb{Z}_n$ were isomorphic to $U(n)$. Now that we have a method of picking apart $U(n)$ into a product of cyclic groups this old result gains new utility.

Example 2.8.12. *Find how many automorphisms of $\mathbb{Z}_{100}$ have order k. Since $Aut(\mathbb{Z}_{100}) \approx U(100)$ we can trade the given question for a related question of how many elements of order k are there in $U(100)$? Note,*

$$U(100) \approx U(4) \times U(25) \approx \mathbb{Z}_2 \times \mathbb{Z}_{20}$$

Now, let's be specific, suppose we look at $k = 4$ then to get $(a, b) \in \mathbb{Z}_2 \times \mathbb{Z}_{20}$ we need either $|a| = 1$ and $|b| = 4$ or $|a| = 2$ and $|b| = 4$. There are 2 elements of order 4 in $\mathbb{Z}_{20}$ and there is just one element of order 1 (or 2) in $\mathbb{Z}_2$ hence there are 4 elements of order 4 in $\mathbb{Z}_2 \times \mathbb{Z}_{20}$ (not that it matters, but they are $(0, 5), (0, 15), (1, 5), (1, 15)$) thus there exist 4 automorphisms of $\mathbb{Z}_{100}$ with order 4.

Remark 2.8.13. Gallian's example on page 156 is more impressive. But, I go on since I believe the example above suffices to illustrate how we can combine the various tools we've developed in this part of the course.

We now return to a little Theorem from Chapter 7 which is terribly useful for calculations. Although, I prove a slightly different version,

Theorem 2.8.14. *Fermat's Little Theorem: if p is prime and $a \neq 0$ mod p then $a^{p-1} = 1$ mod p.*

Proof: Consider $\mathbb{Z}_p$ has p-elements and $U(p)$ has $p - 1$ elements. Observe $\langle a \rangle \leq U(p)$ hence by Lagrange's Theorem we require $|a| \mid (p - 1)$. Thus, there exists n for which $p - 1 = n|a|$ and mod p we calculate that $a^{p-1} = a^{n|a|} = (a^{|a|})^n = 1^n = 1$. □

Theorem 2.8.15. Euler's Theorem: *if $x \in U(k)$ then $x^{\phi(k)} = 1$ mod k.*

Proof: as $|U(k)| = \phi(k)$ this result follows immediately from Corollary 2.3.12 with $G = U(k)$. □

2.8.1 encryption

Both of the results above are useful for modular calculation. A good example of nontrivial modular calculation is given by the RSA encryption scheme. A bit of history, the basic idea of a **trapdoor** function goes to Diffie and Hillman around 1976, then the method I describe here was proposed by Rivest, Shamir and Adleman (RSA) in 1978. Apparently, a British Intelligence's Clifford Cocks also invented the same trick in the 1970's. I learned this material from *Elements of Number Theory* by Stillwell. Honestly, if you think about mathematical magic tricks where the magician does some complicated sequence of operations to the number and then ciphers the original guess... such tricks are based on a reversable algorithm much like the RSA algorithm I describe in this subsection. Only a computer can feasibly implement the encryption of RSA, even my silly toy example takes considerable effort to work through with a traditional handheld calculator. You can get a lot further with Wolframalpha etc. Anyway, let's get to it: the RSA algorithm is roughly as follows:

(1.) Choose a pair of large prime numbers p_1 and p_2 and define $n = p_1 p_2$ (this is huge in real applications)

(2.) Calculate $n = p_1 p_2$ and choose an **encryption exponent** e with $\gcd(e, \phi(n)) = 1$

(3.) Publish e and n so anyone can send you an encrypted message subject to commonly held conventions.

(4.) Your friend takes a message and uses simple transcription to translate it into a coded number. For example, the message "cast" might be traded for 03011920 using $a = 01, b = 02, \ldots z = 26$ and taking 4-letters at a time. Obviously, there is more involved here in real applications, but, I think you can use your imagination to see we can trade words and symbols for a string of numbers with the appropriate scheme. This sort of raw coding is not hard to break and it has been used for centuries. In any event, let us agree that m denotes the message and $m \in \mathbb{N}$ where $m < n$.

(5.) Your friend takes the n and e you published and runs their message through the map $f(m) = m^e$ which scrambles the message m to the encrypted message $f(m)$. They communicate $f(m)$ to you without hiding it, anyone could intercept $f(m)$ mid-transit and it matters not. **Even if the interceptor knows n and e !**

(6.) Since you know p_1, p_2 you can calculate $\phi(n) = (p_1 - 1)(p_2 - 1)$. Suppose $ed = 1 \mod \phi(n)$ then note that $(m^e)^d = m^{1+k\phi(n)} = m^1(m^{\phi(n)})^k = m \mod n$ by Euler's Theorem. Thus, raising $f(m)$ to the d power reveals the message $m \mod n$.

We can try this out using some artificially small primes just to appreciate the algorithm better.

Example 2.8.16. *I tried this example when I last taught number theory. I hope it goes better now:*

(1.) *Consider* $p_1 = 13$ *and* $p_2 = 17$ *hence* $n = 221$.

(2.) $\phi(221) = \phi(13)\phi(17) = 12(16) = 192$ *choose* $e = 37$ *as* $gcd(37, 192) = 1$.

(3.) *my public key is* $n = 221$ *and* $e = 37$

(4.) *my friend chooses a single-letter message of "J" which translates by the alphabet code to* $m = 10$

(5.) *my friend encrypts the message* $m = 10$ *by calculating* $m^{37} = 10^{37}$ *mod 221. There are various tricks to actually perform this calculation. Notice* $37 = 32 + 4 + 1$ *so* $(37)_2 = (100101)$ *which inspires us to look at* m^{37} *as*

$$((((m^2)^2)^2)^2)^2((m^2))^2m$$

Note $10^2 = 100$ *then* $10^4 = (10^2)^2 = 10000 = 221(45) + 55 = 55$ *mod* 221. *Hence,*

$$10^8 = ((10^2)^2)^2 = 55^2 = 3025 = 13(221) + 152 = 152 = -69$$

mod 221. Continuing,

$$10^{16} = (-69)^2 = 4761 = 21(221) + 120 = 120$$

$$10^{32} = (120)^2 = 14400 = 65(221) + 35 = 35$$

$$10^{37} = 10^{32}10^410^1 = (35)(55)(10) = 19250 = 87(221) + 23 = 23 \quad \Rightarrow \quad \boxed{f(m) = 23}$$

(6.) *my private decryption key requires me to calculate the multiplicative inverse of* $e = 37$ *modulo* $\phi(221) = 192$. *We can use the extended euclidean algorithm to accomplish this generally.*

$$192 = 37(5) + 7, \quad 37 = 5(7) + 2, \quad 7 = 3(2) + 1$$

hence working backwards,

$$1 = 7 - 3(2) = 7 - 3(37 - 5(7)) = 16(7) - 3(37) = 16(192 - 5(37)) - 3(37)$$

thus $1 = 16(192) - 83(37)$ *hence* $37^{-1} = -83 = 109$. *Hence, to decrypt the message* $f(m) = 23$ *I simply raise the message to the* 109*-th power mod* 221. *Notice,* $109 = 64 + 32 + 8 + 4 + 1$ *so we can calculate* 23^{109} *mod* 221 *systematically as follows:*

$$23^4 = 279841 = 1266(221) + 55 = 55$$

$$23^8 = 55^2 = 3025 = 13(221) + 152 = 152$$

$$23^{16} = (23^8)^2 = 152^2 = 23104 = 104(221) + 120 = 120$$

$$23^{32} = 120^2 = 14400 = 65(221) + 35 = 35$$

$$23^{64} = 35^2 = 1225 = 5(221) + 120 = 120$$

Hence, mod 221 *we have:*

$$\begin{aligned} 23^{109} &= 23^{64}23^{32}23^{8}23^{4}23 \\ &= (120)(35)(152)(55)(23) \\ &= (4200)(192280) \\ &= (1)(10) \\ &= 10. \end{aligned}$$

As you can see, my friend sent the coded message of 23 *and I was able to cipher it was the hidden message of* 10. *In order to calculate the decryption* $d = 109$ *it was necessary to calculate* $\phi(n)$ *which is simple when the factorization of* n *is known. Of course, for* $n = 221$ *you can easily find* 13 *and* 17 *as factors and hence* $\phi(n) = 12(16) = 192$ *was easy to find hence the inverse of the public* $e = 37$ *is also very much possible to calculate in my toy example. The difference with real encryption is the* p_1, p_2 *are typically hundreds of digits long so the modulus* p_1p_2 *is pragmatically impossible to factor and the problem of calculating* $\phi(n)$ *directly is also rather daunting. It is neat that the difficulty of finding large primes allows us to communicate securely. It's not without holes, and from what I read, the algorithm I describe here has further refinements in practice. I just thought it would be fun to run through this as a real world example of sorts.*

Remark 2.8.17. When I did this toy example the first time I used $e = 49$. This was a most unfortunate choice since $U(221) \approx U(13) \times U(17) \approx \mathbb{Z}_{12} \times \mathbb{Z}_{16} \approx \langle a \rangle \times \langle b \rangle$ where $|a| = 12$ and $|b| = 16$ and you can easily calculate $(x, y)^{48} = ((x^{12})^4, (y^{16})^3) = (e, e)$ hence $(x, y)^{49} = (x, y)$. In other words, my encryption exponent was the worst possible choice. I gave the students some time to encrypt a message then after leaving the room and returning the gave me $f(m)$ and told me m was the same. If I had picked most any other number it would have been ok. I thought $\gcd(49, 192) = 1$ sufficed to make the method go, but, as my story shows, there are hidden dangers and the algorithm I sketch in this section is just the most rudimentary version.

I recommend working on Supplemental Exercises for Gallian Chapters 5-8 such as: (page 169-171)

$$\#1 - 47.$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know.

Problems for Lecture 17: (these are collected at the start of Lecture 19)

Problem 65: Gallian page 169 #3

Problem 66: Gallian #6 from page 169 .

Problem 67: Gallian #54 from page 164 .

Problem 68: Prove the following:

(a) Suppose G, H, K are finite. Prove: if $G \oplus H = G \oplus K$ then $H \approx K$.

(b) Show that removing the finiteness condition in the previous part makes the claim false. Give an explicit counterexample to be precise.

(c) Show that $G \oplus H = G \oplus K$ does not imply $H = K$ even for abelian groups. Give an infinite and finite counterexample (of course, logically just one is needed, but, I want you to think about both cases)

2.9 Lecture 18: group actions

The concept of a group action has thus far been underplayed in our course. If you watch the videos by Professor Gross of Harvard based on Artin's *Algebra* then you might notice he introduces the concept of a group action much earlier and he centers much of the course around the concept. Group actions do put groups into action and they help bring the application of group theory into the foreground. I'm following a combination of §7.3 in Beachy and Blair's *Abstract Algebra* as well as §2.7 in Rotman's *A First Course in Abstract Algebra* both in their 3^{rd} edition.

Definition 2.9.1. *Let G be a group and S a set. A multiplication of G on S defined by $\star : G \times S \to S$ is called a* **group action** *of G on S if for each $x \in S$,*

(i.) *$e \star x = x$ for e the identity of G,*

(ii.) *$a \star (b \star x) = (ab) \star x$ for all $a, b \in G$*

Let's look at a few examples to get a sense of the breadth of this concept.

Example 2.9.2. *Consider G a subgroup of the group of permutations on some set S; $G \leq Perm(S)$. We define $\star : G \times S$ as follows for $\sigma \in G$,*

$$\sigma \star x = \sigma(x)$$

for each $x \in S$. Recall, $\sigma : S \to S$ is a bijection by the definition of permutations. Moreover, as G is a group the identity map $\sigma = Id$ is in G and

$$Id \star x = Id(x) = x$$

for each $x \in S$. Furthermore, if $\sigma, \beta \in G$ then

$$(\sigma \circ \beta) \star x = (\sigma \circ \beta)(x) = \sigma(\beta(x)) = \sigma(\beta \star x) = \sigma \star (\beta \star x)$$

for each $x \in S$. Thus $\star$ defines a group action of G on S.

Notice, we could take $S = G$ and use the permutations induced from left-multiplications on G as a particular instance of the above Example. In other words, we can view the proof of Cayley's Theorem 2.1.17 as an example of using a group action to study the group. Many interesting applications appear when S and G are not the same set.

Example 2.9.3. *Suppose $H \leq G$ then H acts on G according to the following action:*

$$h \star x = hx$$

for each $h \in H$ and $x \in G$. Naturally,

$$e \star x = ex = x, \qquad \& \qquad (ab) \star x = (ab)x = a(bx) = a \star (b \star x)$$

for all $a, b \in H$ and $x \in G$ using the associativity of group multiplication and the existence of the identity $e \in H \leq G$.

With a bit more imagination, we can define a second action of H on G:

Example 2.9.4. *Suppose $H \leq G$ then H acts on G according to the following action:*

$$h \bullet x = xh^{-1}$$

for each $h \in H$ and $x \in G$. Naturally,

$$e \bullet x = xe^{-1} = x, \qquad \& \qquad (ab) \bullet x = x(ab)^{-1} = xb^{-1}a^{-1} = (b \bullet x)a^{-1} = a \bullet (b \bullet x).$$

for all $a, b \in H$ and $x \in G$. This action is brought to you courtesy of the socks-shoes formula for the inverse.

Example 2.9.5. *Scalar multiplication by $\mathbb{F}^\times$ gives an action on a vector space over $\mathbb{F}$. Recall $\mathbb{F}^\times = \mathbb{F} - \{0\}$ forms a group and note*

$$1 \cdot x = x, \qquad \& \qquad (ab) \cdot x = a \cdot (b \cdot x)$$

for each $x \in V$ and $a, b \in \mathbb{F}^\times$. These identities given in the definition of a vector space over $\mathbb{F}$ serve to show scalar multiplication forms a group action by the nonzero-scalars.

Example 2.9.6. *Let $S = \mathbb{F}^n$ and consider $G = GL(n, \mathbb{F})$ the general linear group of invertible $n \times n$ matrices over the field $\mathbb{F}$. A natural group action of $GL(n, \mathbb{F})$ on $\mathbb{F}^n$ is given by* **matrix multiplication***:*

$$A \star x = Ax$$

for each $A \in GL(n, \mathbb{F})$ and $x \in \mathbb{F}^n$. Observe,

$$I \star x = Ix = x, \qquad \& \qquad (AB) \star x = (AB)x = A(Bx) = A(B \star x) = A \star (B \star x)$$

for all $A, B \in GL(n, \mathbb{F})$ and $x \in \mathbb{F}^n$ where I denotes the $n \times n$ identity matrix.

Of course, you can replace $\mathrm{GL}(n, \mathbb{F})$ with a suitable subgroup and still obtain a group action. Indeed, if you think about any of our group actions if we have an action by G on S then you can easily see how to create a corresponding action by $H \leq G$ on S simply by restricting the given action. In particular, Problem 69 investigates the action of a particular subgroup of $\mathrm{GL}(2, \mathbb{R})$.

Theorem 2.9.7. *Let G be a group and suppose S is a set. Any group homomorphism from G into $Perm(S)$ corresponds to an action of G on S. Conversely, every group action of G on S arises in this manner.*

Proof: was given in class on 10-14-16. I hope to type it up sometime soon... □

Definition 2.9.8. *Let G be a group and S a set and $\star : G \times S \to S$ a group action. We define:*

(i.) *for each $x \in S$ the* **orbit of x under G** *is $\mathcal{O}(x) = \{g \star x \mid g \in G\}$*

(ii.) *for each $x \in S$ the* **stabilizer of x in G** *is $G_x = \{g \in G \mid g \star x = x\}$*

(iii.) *the* **subset of S fixed by G** *is denoted $S^G = \{x \in S \mid g \star x = x, \text{ for all } g \in G\}$*

The stabilizer G_x is also known as the **isotropy subgroup** of x. Notice, $e \in G_x$ is immediate from the definition of a group action.

Example 2.9.9. *Continuing Example 2.9.3, we note for $H \leq G$*

$$\mathcal{O}(x) = \{hx \mid h \in H\} = Hx.$$

the orbits are right-H-cosets. Let $x \in G$ then the stabilizer of x in H is given by:

$$H_x = \{h \in H \mid hx = x\}$$

notice $hx = x$ implies $h = e$ thus $H_x = \{e\}$ for each $x \in G$. There is no subset of G fixed by H unless $H = \{e\}$ in which case all of G is fixed by H.

Example 2.9.4 gives orbits which are left-H-cosets.

Example 2.9.10. *The scalar multiplication action of Example 2.9.5 gives interesting orbits. In particular, if $v \in \mathbb{F}^n$ then $c \cdot v$ gives the line with direction-vector v with origin removed since $c \in \mathbb{F}^\times$ forbids $c = 0$. The orbit of $v = 0$ is just the origin. In contrast, if $v \neq 0$ the stabilizer is $(\mathbb{F}^\times)_v = \{1\}$ and the stabilizer of the origin is the $\mathbb{F}^\times$ ($(\mathbb{F}^\times)_0 = \mathbb{F}^\times$).*

Conjugation provides an important group action of a group on itself.

Example 2.9.11. *Consider $S = G$ a group and define an action of G on itself as follows:*

$$g \star x = gxg^{-1}$$

for all $g, x \in G$. Clearly $e \star x = exe = x$ for each $x \in G$. Moreover, if $a, b \in G$ then

$$(ab) \star x = (ab)x(ab)^{-1} = a(bxb^{-1})a^{-1} = a(b \star x)a^{-1} = a \star (b \star x)$$

for each $x \in G$. Furthermore, the orbit of x is its **conjugacy class**[21]*:*

$$\mathcal{O}(x) = \{gxg^{-1} \mid g \in G\}$$

and the stabilizer of x in G is the **centralizer** *of x*

$$G_x = \{g \in G \mid gxg^{-1} = x\} = \{g \in G \mid gx = xg\}.$$

The centralizer of x is the set of all group elements which commute with x. Finally, the fixed subset of this group action is the **center of** *G:*

$$S^G = \{g \in G \mid xgx^{-1} = g \text{ for all } x \in G\} = \{g \in G \mid xg = gx \text{ for all } x \in G\} = Z(G).$$

The action of G on itself by conjugation is at the heart of many of the major theorems which are derived from the theory of group actions. I won't prove all of these theorems this semester, but, I hope I show you enough you can see the spirit of the arguments.

[21] we soon prove that the orbits define a partition of S and they are the equivalence classes for a natural equivalence relation given by the group action

I recommend working on Supplemental Exercises for Gallian Chapters 5-8 such as: (page 169-171)

$$\#1 - 47.$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know.

Problems for Lecture 18: (these are collected at the start of Lecture 20)

Problem 69: Let $G = \mathrm{O}(2, \mathbb{R}) = \{R \in \mathbb{R}^{2\times 2} \mid R^T R = I\}$. Consider the action of G on $\mathbb{R}^2$ by matrix multiplication. Find the orbits and stabilizers for the given action.

Problem 70: Gallian Chapter 7, #27 from page 146. *(this is not particularly attached to the topic of group actions, but this is a useful identity to know and use)*

Problem 71: Find all homomorphisms from $\mathbb{Z}_5$ to $\mathbb{Z}_7$.

Problem 72: Show any homomorphism from $\mathbb{Z}_n$ to $\mathbb{Z}_k$ must have the form $\phi([x]_n) = [mx]_k$ for all $[x]_n \in \mathbb{Z}_n$ for some m where $k \mid mn$.

2.10 Lecture 19: orbit stabilizer theorem and conjugacy

In this Lecture we develop the main tools we need to solve counting problems. The orbit stabilizer theorem gives a simple connection between the number of cosets of the stabilizer and the orbit of a particular point: they match. When we apply this theorem to the conjugation action of G on itself we obtain interesting new insight into conjugacy classes. Ultimately, this allows us to prove Cauchy's Theorem in the next Lecture.

Theorem 2.10.1. *If G is a group which acts on a set S and $x \in S$ then $G_x \leq G$.*

Proof: let $x \in S$ and consider $G_x = \{g \in G \mid g \star x = x\}$ where $\star$ is a group action on S. Note $e \star y = y$ for all $y \in S$ hence $e \star x = x$ and we find $e \in G_x \neq \emptyset$. Suppose $a, b \in G_x$ and calculate:

$$(ab) \star x = a \star (b \star x) = a \star x = x$$

thus $ab \in G_x$. Consider $a \in G_x$ note that as G is a group there exists a^{-1} for which $aa^{-1} = e$. Remember $a \star x = x$ as $a \in G_x$ hence:

$$x = e \star x = (a^{-1}a) \star x = a^{-1} \star (a \star x) = a^{-1} \star x$$

which shows a^{-1} fixes x thus $a^{-1} \in G_x$ and we conclude $G_x \leq G$ by the two-step subgroup test. □

Theorem 2.10.2. *If G is a group which acts on a set S then the orbits $\mathcal{O}(x)$ partition S. Moreover, for a finite set,*

$$|S| = \sum_i |\mathcal{O}(x_i)|,$$

where one x_i is selected for each orbit.

Proof: define $x \sim y$ if $x = g \star y$ for some $g \in G$. We claim $\sim$ is an equivalence relation on S. Notice, $x = e \star x$ hence $x \sim x$ for each $x \in S$. If $x \sim y$ then $x = g \star y$ for some $g \in G$ hence $g^{-1} \star x = g^{-1} \star (g \star y) = y$ and we find $y \sim x$. If $x \sim y$ and $y \sim z$ then there exist $g, h \in G$ for which $x = g \star y$ and $y = h \star z$ thus $x = g \star (h \star z) = (gh) \star z$ whence $x \sim z$. In summary, $\sim$ is reflexive, symmetric and transitive. It follows the equivalence classes of $\sim$ partition S. Notice, the equivalence class containing x is given by:

$$\{y \in S \mid y \sim x\} = \{y \in S \mid y = g \star x \text{ for some } g \in G\} = \{g \star x \mid g \in G\} = \mathcal{O}(x).$$

Hence the orbits of the group action partition S. Counting gives us the formula for $|S|$. □

Theorem 2.10.3. *If G is a group which acts on a set S and $x \in S$ then the elements of the orbit $\mathcal{O}(x)$ are in one-to-one correspondence with the left-cosets of G_x in G. Moreover, $|\mathcal{O}(x)| = [G : G_x]$ which is to say the size of the orbit $|\mathcal{O}(x)|$ is the index of the stabilizer G_x in G.*

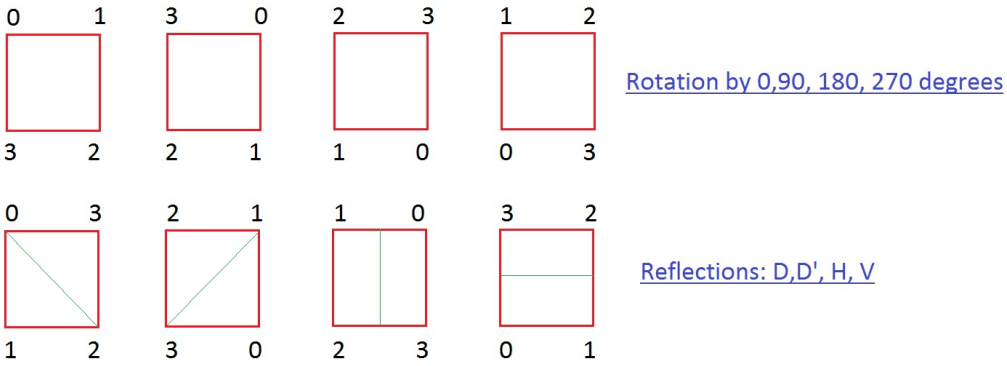
Proof: Let $x \in S$ and suppose G acts on S. Let G/G_x denote the family of G_x-cosets in G. Define $f : \mathcal{O}(x) \to G/G_x$ as follows: for $y \in \mathcal{O}(x)$ there exists $g \in G$ with $y = g \star x$ we define $f(y) = gG_x$. To see f is well-defined note it is clearly into G/G_x. Suppose $y = h \star x$ for some $h \in G$. Then $h \star x = g \star x$ and hence $(h^{-1}g) \star x = x$ which provides $h^{-1}g \in G_x$ and hence $hG_x = gG_x$ thus f is single-valued. To see that f is injective suppose $f(y) = f(z)$ hence there exist $h, g \in G$ for which $y = h \star x$ and $z = g \star x$ where $hG_x = gG_x$. Thus $h^{-1}g \in G_x$ and $(h^{-1}g) \star x = x$ or $h \star x = g \star x$ hence $y = z$ and we conclude f is injective. If $gG_x \in G/G_x$ then observe $y = g \star x \in \mathcal{O}(x)$ and $f(y) = gG_x$. Thus f is a bijection and we can use it to **count**: $|\mathcal{O}(x)| = |G/G_x| = [G : G_x]$. □

Corollary 2.10.4. Orbit Stabilizer Theorem: *If a finite group G acts on a set S then the number of elements in any orbit $\mathcal{O}(x)$ must divide the order of the group. Moreover, $|G| = |\mathcal{O}(x)||G_x|$ for each $x \in S$.*

Proof: by Theorem 2.10.3 we know $|\mathcal{O}(x)| = |G/G_x| = [G : G_x]$. But, for a finite group the number of G_x-cosets is the order of the group G divided by the order of G_x: $[G : G_x] = |G|/|G_x|$. Hence $|G| = [G : G_x]|G_x| = |\mathcal{O}(x)||G_x|$. □

The orbit stabilizer theorem gives us a nice tool for a variety of counting problems.

Example 2.10.5. *Consider $S = \{v_o, v_1, v_2, v_3\}$ the vertices of a square. We have a natural action of D_4 acting on S. The following picture makes the discussion easiest to follow:*



In cycle notation the rotations are $(1), (v_o v_1 v_2 v_3), (v_0 v_2)(v_1 v_3), (v_0 v_3 v_2 v_1)$ whereas the reflections are $(v_1 v_3), (v_0 v_2), (v_0 v_1)(v_2 v_3), (v_0 v_3)(v_1 v_2)$. It is not hard to see which vertex is fixed or moved by each element of D_4. Observe,

$$\mathcal{O}(v_0) = \{g \star v_0 \mid g \in D_4\} = \{v_0, v_1, v_2, v_3\}$$

indeed, you could start with any vertex and find the same orbit. This makes the given action a **transitive action**. *Furthermore, observe:*

$$G_{v_0} = \{g \in D_4 \mid g \star v_0 = v_0\} = \{(1), (v_1 v_3)\}$$

Note, $|D_4| = 8 = |\mathcal{O}(v_0)||G_{v_o}|$. In fact, we can do the same for any vertex. For example,

$$G_{v_1} = \{(1), (v_0 v_2)\}, \quad \& \quad \mathcal{O}(v_1) = \{v_0, v_1, v_2, v_3\}$$

Example 2.10.6. *Think about a cube. There are six faces to a cube. If we think about the symmetries of the cube, notice that the stabilizer of the face is given by four the rotations about the center of the face. So, thinking of our set S as the set of faces we find the stabilizer of a particular face x has $|G_x| = 4$. On the other hand, there is a symmetry of the square which moves any face to another face hence the group action is transitive; $|\mathcal{O}(x)| = 6$. It follows the group of symmetries on the cube has order 24.*

Following Example 2.9.2 we consider some rather special subgroups of the permutation group on $\mathbb{N}_n$. Begin with $n = 6$ to get warmed-up:

Example 2.10.7. *Consider* $S = \{1, 2, 3, 4, 5, 6\}$ *and* $\sigma = (123)(56)$. *The cyclic group* $\langle\sigma\rangle = \{\sigma^k \mid k \in \mathbb{Z}\} \leq S_6$ *acts on* S *in the natural way. Observe,*

$$\mathcal{O}(1) = \{1, 2, 3\}, \quad \mathcal{O}(5) = \{5, 6\}, \quad \mathcal{O}(4) = \{4\}$$

As you can see the size of the orbits divide the order of G *which is of course the order of* $\sigma = lcm(2, 3) = 6$. *Furthemore, we see Theorem 2.10.2 in action:*

$$|S| = 6 = |\mathcal{O}(1)| + |\mathcal{O}(5)| + |\mathcal{O}(4)|$$

The example above naturally generalizes.

Example 2.10.8. *Consider* $S = \{1, 2, \ldots, n\}$ *and* $\sigma \in S_n$. *Let* $G = \langle\sigma\rangle$ *act on* S *and note:*

$$\mathcal{O}(i) = \{\sigma^k(i) \mid i \in \mathbb{Z}\}$$

If we know the disjoint cycle factorization of $\sigma = \beta_1\beta_2\cdots\beta_t$ *where* $\beta_j = (i_{j1}i_{j2}\cdots i_{jr_j})$ *for* $j = 1, \ldots, t$ *then* $i \in S$ *is only moved by the particular* β_j *which contains it. Moreover, the orbit is simply the entries in* β_j *listed: in the notation I chose,*

$$\mathcal{O}(i_{j1}) = \{i_{j1}, i_{j2}, \ldots, i_{jr_j}\}$$

for $j = 1, 2, \ldots, t$. *Here we know* $|G| = lcm(r_1, r_2, \ldots, r_t)$. *The size of the orbits are just* $r_1, r_2, \ldots, r_t$ *and* 1 *for the numbers left out. Again, we see that the orbit sizes divide the order of the group in action.*

I recommend working on Supplemental Exercises for Gallian Chapters 5-8 such as: (page 169-171)

$$\#1 - 47.$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know.

Problems for Lecture 19: (these are collected at the start of Lecture 21)

Problem 73: Let G be a group and let S be the set of all subgroups of G. If $a \in G$ and $H \leq G$ then define $a \star H = aHa^{-1}$. Show that $\star$ forms a group action and describe the orbits and stabilizers.

Problem 74: Let H be a proper subgroup of a group G and let $S = G/H$ denote the set of left cosets of H. Define $g \star (xH) = (gx)H$ for each $g, x \in G$. Show $\star$ forms a group action and describe the orbits, stabilizers, and fixed subset of this action.

Problem 75: Let H, K be subgroups of G and let H act on the left cosets of K in the natural way; $h \star (gK) = (hg)K$ for each $h \in H$ and $g \in G$. Show that $|HK| = \frac{|H||K|}{|H \cap K|}$. *Hint: study the orbit of* K *under the given action*

Problem 76: How many symmetries does a tetrahedron have? Prove your result in three ways by considering the set of faces, edges or vertices of the regular tetrahedron. There are 4 faces, 6 edges and 4 vertices for this platonic solid.

2.11 Lecture 20: Cauchy and Sylow theorems

In Example 2.10.8 we saw that the action of $\alpha \in \langle \sigma \rangle \leq S_n$ on $\mathbb{N}_n = \{1, \ldots, n\}$ defined by $\alpha \star x = \alpha(x)$ for $x \in \mathbb{N}_n$ gave orbits which directly related to the disjoint cycle factorization of σ. For example, if $\sigma = (1234)(56)(78)$ acts on $\mathbb{N}_{10}$ then orbits of the action of $\langle \sigma \rangle$ are just:

$$\mathcal{O}(1) = \{1, 2, 3, 4\}, \quad \mathcal{O}(5) = \{5, 6\}, \quad \mathcal{O}(7) = \{7, 8\}, \quad \mathcal{O}(9) = \{9\}, \quad \mathcal{O}(10) = \{10\}.$$

The conjugation action of S_n on itself is also interesting. Let me point out a feature of cycle calculation we have not yet appreciated in our study:

Example 2.11.1. *Let us study conjugation of $\alpha = (12)(34)$ in S_4. Conjugate by (123),*

$$(123)\alpha(123)^{-1} = (123)(12)(34)(321) = (14)(23)$$

$$(124)\alpha(124)^{-1} = (124)(12)(34)(421) = (13)(24)$$

$$(23)\alpha(23)^{-1} = (23)(12)(34)(23) = (13)(24)$$

$$(1234)\alpha(1234)^{-1} = (1234)(12)(34)(4321) = (14)(23)$$

$$(12)\alpha(12)^{-1} = (12)(12)(34)(12) = (12)(34)$$

In short, α is fixed under conjugation by any cycle which is in its own cycle decomposition whereas conjugation by other permutations moves α to some other permutation in its conjugacy class. Notice the **cycle type** *is the same; we still have a product of two transpositions. Let $\beta = (1234)$ then*

$$(12)\beta(12)^{-1} = (12)(1234)(12) = (1342)$$

or $\gamma = (123)$,

$$(124)\gamma(124)^{-1} = (124)(123)(421) = (1)(243) = (243)$$

Conjugation of β produces another 4-cycle and conjugation of γ produces another 3-cycle. You can try other possible conjugations, the result will be the same. Conjugation preserves **cycle-type**.

We can generalize the observation of the example above. You could view the next example as an informal definition of cycle-type.

Example 2.11.2. *Let $\alpha = \beta_1\beta_2 \cdots \beta_t$ be a disjoint cycle factorization of $\alpha \in S_n$. If we conjugate α by any $\sigma \in S_n$ then the result is a permutation $\sigma\alpha\sigma^{-1}$ which has a disjoint cycle factorization $\sigma\alpha\sigma^{-1} = \gamma_1\gamma_2 \cdots \gamma_t$ where $|\beta_i| = |\gamma_i|$ for $i = 1, 2, \ldots, t$. In other words, the conjugation of α by σ produces another permutation with the same cycle-type. Furthermore, if σ is formed by the product of some subset of the cycles $\beta_1, \ldots, \beta_t$ then $\sigma\alpha\sigma^{-1} = \alpha$.*

Corollary 2.11.3. *(to the orbit stabilizer theorem) Let G be a finite group and $x \in G$. The number of conjugates to x is the index of the centralizer of x; $|\{gxg^{-1} \mid g \in G\}| = [G : C(x)]$.*

Proof: consider the action of G on itself by conjugation. Notice,

$$\mathcal{O}(x) = \{gxg^{-1} \mid g \in G\} \quad \& \quad G_x = \{g \in G \mid gxg^{-1} = x\} = \{g \in G \mid gx = xg\} = C(x).$$

The orbit stabilizer theorem provides $|\mathcal{O}(x)| = [G : G_x]$ hence $|\{gxg^{-1} \mid g \in G\}| = [G : C(x)]$. □
The notation $\{gxg^{-1} \mid g \in G\}$ is a bit cumbersome. The following is a common notation to avoid all that writing:

Definition 2.11.4. *The* **conjugacy class** *of x in a group G is denoted $x^G = \{gxg^{-1} \mid g \in G\}$.*

In the language just introduced, we find $|x^G| = [G : C(x)]$. Since the index of a subgroup necessarily divdes the order of a group we find an application of Corollary 2.11.3:

Corollary 2.11.5. *The number of permutations in S_n with a particular cycle type must divide $n!$.*

Proof: consider the group action of S_n on itself by conjugation. Observe the conjugacy classes are formed by permutations of the same cycle type. Thus $|\sigma^{S_n}|$ is the number of permutations with the same cycle-type as σ. Apply Corollary 2.11.3 to find $|\sigma^{S_n}| = [S_n : C(\sigma)]$ but we know $|C(\sigma)|[S_n : C(\sigma)] = |S_n| = n!$ and the Corollary follows. □

Example 2.11.6. *Consider S_4, we find 5 distinct cycle-types:*

representative cycle σ	$\lvert\sigma^{S_4}\rvert$ = *# of cycles with same type*
(1)	1
(12)	6
(123)	8
(1234)	6
(12)(34)	3

Since the $\mathcal{O}(\sigma) = \sigma^{S_4}$ under the conjugation action and we know the orbits partition the set on which the group acts (Theorem 2.10.2) we are not surprised to notice:

$$4! = 24 = 1 + 6 + 8 + 6 + 3.$$

I invite the reader to think through the analog of the Example above for S_5. Here's a hint:

$$120 = 5! = 1 + 10 + 20 + 30 + 24 + 20 + 15.$$

Theorem 2.11.7. *If G is a finite group with order divisible by a prime p then G contains an element of order p.*

Proof: assume G is a finite group with order divisible by a prime p. Our proof will proceed by induction on $|G|$. Note, for $|G| = 1$ the theorem is trivially true as 1 has no prime divisors. Suppose the theorem is true for groups upto order $n - 1$ and consider $|G| = n$. If $x \in G$ then $|x^G| = [G : C(x)]$ where $C(x)$ is the centralizer of x in G. If $x \notin Z(G)$ then x^G has more than one element[22] hence

$$|C(x)| < |G|$$

If $p \mid |C(x)|$ for some[23] $x \notin Z(G)$ then by the inductive hypothesis $C(x) \leq G$ has an element of order p and thus G has an element of order p. It remains to study the case $p \nmid C(x)$ for all noncentral x. Recall we assume $p \mid |G|$ and note

$$|G| = [G : C(x)]|C(x)|$$

thus by Euclid's Lemma (as $p \nmid C(x)$) it must be that $p \mid [G : C(x)]$. Since the conjugacy classes partition G we have: (applying Theorem 2.10.2 to the action by conjugation on G)

$$|G| = |Z(G)| + \sum_i [G : C(x_i)]$$

where one x_i is selected from each conjugacy class containing more than one element. To be clear, the $|Z(G)|$ counts the elements which fit into conjugacy classes which are mere singletons; $x \in Z(G)$ implies $x^G = \{x\}$. Now, $p \mid G$ and $p \mid [G : C(x_i)]$ for each i thus $p \mid |Z(G)|$. Claim[24], an abelian

[22] think about why $x \in Z(G)$ makes $x^G = \{x\}$.

[23] we say such an x is **non-central**

[24] seems like a good bonus homework problem

group whose order is divisible by p contains an element of order p. Since $Z(G)$ is abelian and $p \mid Z(G)$ we conclude that $Z(G)$ contains an element of order p hence G contains an element of order p and we are done. □.

Definition 2.11.8. *The* **class equation** *of a finite group G is:*

$$|G| = |Z(G)| + \sum_i [G : C(x_i)]$$

where one x_i is selected from each conjugacy class containing more than one element.

I recommend working on Supplemental Exercises for Gallian Chapters 5-8 such as: (page 169-171)

$$\#1 - 47.$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know.

Remark 2.11.9. Notes currently missing statement of Sylow Theorems and some discussion here, we did a bit more in Lecture, I'll have to leave the notes a bit unfinished here for the time being. 10-30-16.

Problems for Lecture 20: (these are collected at the start of Test 2 Question Day)

Problem 77: Find the distinct cycle types for S_5 and determine the number of each type. Make a table to organize your results in similar fashion to the table given for S_4. Explain how your data is consistent with the class equation.

Problem 78: **Prove:** a finite abelian group whose order is divisible by a prime p has an element of order p. *Hint: the proof is by induction on the order of the group.*

Problem 79: Solve $[x]_2 = [3]_2$ and $[x]_7 = [4]_7$ and $[x]_{37} = [20]_{37}$ simultaneously. Here $[x]_n = x + n\mathbb{Z}$ denotes $x \in \mathbb{Z}_n$ in our usual less cumbersome notation. This problem involves several moduli so I invoke the clumsy notation as to reduce ambiguity.

Problem 80: Find the conjugacy classes for A_4 and verify the class equation.

2.12 Lecture 21: lattice theorem, finite simple groups

Remark 2.12.1. Notes currently missing statement of the second, third and lattice isomorphism theorems and the overview of the classification of simple groups. I gave some handouts in class to fill the gap here, unfortunately the notes will have to remain incomplete here for the time begin 10-30-16.

I recommend working on Supplemental Exercises for Gallian Chapters 5-8 such as: (page 169-171)

$$\#1 - 47.$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know.

Problems for Lecture 21: (these are collected at the start of Test 2 Question Day)

Problem 81: **Prove:** if A, B are subgroups of G and $A \leq N_G(B)$ then $AB \leq G$.

Problem 82: Prove the Third Isomorphism Theorem. In particular, suppose G is a group with normal subgroups H and K where $H \leq K$.

(a.) Prove $K/H \trianglelefteq G/H$

(b.) Define $\phi : G/H \to G/K$ by $\phi(gH) = gK$ for each $gH \in G/H$. Show ϕ is well-defined and ϕ is a homomorphism with $\text{Ker}(\phi) = K/H$

(c.) apply the first isomorphism theorem to the data you found in the last part to argue that K/H is a normal subgroup of G/H and $(G/H)/(K/H) \approx G/K$

Problem 83: Find the subgroup lattice for $G = \mathbb{Z}_{12}$. Apply the lattice isomorphism theorem to find lattice diagrams for all possible factor groups of G.

Problem 84: Gallian number 51 from page 208.

2.13 Lecture 22: Boolean group, rank nullity theorem

Remark 2.13.1. There was no homework assigned in this Lecture which was given on 10-26-16.

Chapter 3

Introduction to Rings and Fields

The groups we have thus far studied are sometimes just half of a larger system. We don't just have addition or multiplication, no, we have both. The addition is commutative and the multiplication is sometimes commutative. However, with matrices we know $AB \neq BA$. Furthermore, sometimes there is some object which serves as the multplicative identity and sometimes there is no such object. For example, even integers are closed under multiplication and addition in the natural sense, yet, there is no number which serves as a multiplicative identity. The concept of a ring is given here to collect nearly all the examples you have ever encountered which involve both an addition and multiplication. By abstracting the concept of a ring we are able to prove common truths about a myriad of examples by one sweeping proof.

However, the ring concept is more than just an efficiency of argument. We learn there are different kinds of rings which merit our study; integral domains, euclidean domains, principal ideal domains, unique factorization domains and fields. These are interconnected and there theory generalizes the story we already know for prime factorization of integers. Formal power series ith coefficients in a ring and the simple polynomials contained within them provide a canvas on which we can paint much of the mathematics we hold dear. Highschool algebra is studied from the vantage of modern algebra. We even learn how to invent new number systems where solutions exist to polynomial equations which were before unsolvable.

Finally, this portion of the course gives the student another chance to assimilate concepts such as a **subring**, **coset**, or the **first isomorphism theorem**. Nearly everything we did for groups has an analog in the study of rings. Ideally this next exposure will help the main concepts of the course finally ring true to every student.

Roughly, we cover Chapters 12-21 of Gallian's 5-th edition in this Chapter.

3.1 Lecture 23: rings and integral domains

As usual, we follow Gallian, this definition is taken from the beginning of Chapter 12. Gallian mentions the term **ring** is due to Hilbert who introduced the term in 1897.

Definition 3.1.1. *A* **ring** R *is a nonempty set with two binary operations known as* **addition** *and* **multiplication** *deoted by* $+$ *and juxtaposition respectively; addition has* $(a,b) \mapsto a+b$ *whereas multiplication has* $(a,b) \mapsto ab$*. These operations satisfy:*

(1.) $a+b = b+a$ *for all* $a,b \in R$

(2.) $(a+b)+c = a+(b+c)$ *for all* $a,b,c \in R$

(3.) *there exists* $0 \in R$ *(known as* **zero***) for which* $a+0 = a$ *for each* $a \in R$*.*

(4.) *for each* $a \in R$ *there exists an additive inverse denoted* $-a$ *for which* $a+(-a) = 0$*.*

(5.) $(ab)c = a(bc)$ *for all* $a,b,c \in R$

(6.) $a(b+c) = ab+ac$ *and* $(b+c)a = ba+ca$ *for all* $a,b,c \in R$

In words, (1.) says addition commutes, (2.) says addition is associative, (3.) say an additive identity exists, (4.) says R is closed under additive inversion. Collectively, (1.)-(4.) and the fact that $+ : R \times R \to R$ is a binary operation make $(R,+)$ an **Abelian group**. Condition (5.) says multiplication of R is associative. The conditions comprising (6.) are known as the **right-distributive property** $a(b+c) = ab+ac$ and the **left-distributive property** $(b+c)a = ba+ca$.

We recognize (6.) as the main facts which power factoring in precollegiate algebra. Read one direction (6.) gives you the pattern needed to factor out a, read the other way, (6.) tells us how to multiply a across a sum $b+c$. We need to set a few more terms:

Definition 3.1.2. *Let* R *be a ring.*

(1.) *If* $ab = ba$ *for all* $a,b \in R$ *then* R *is a* **commutative ring***. Otherwise,* R *is known as a* **noncommutative ring***.*

(2.) *If there exists* $1 \in R$ *for which* $a\,1 = 1\,a = a$ *for each* $a \in R$ *then we say* R *is a* **ring with unity** *or we say* R *is* **unital** *with unity* 1*.*

Most of the rings we study are commutative rings with unity. Some authors insist that the definition of ring includes the existence of a multiplicative identity. For example, Rotman's elementary abstract text uses such a definition. Next, we generalize divisibility terminology: I'll assume the ring is commutative as I do not desire to invest energy studying left verses right factors. If you wish to read about noncommutative ring theory then you should consult the classic text *Rings and Ideals* by McCoy which develops ring theory in surprising generality.

Definition 3.1.3. *Let* R *be a commutative ring.*

(1.) *Let* $a,b \in R$*. If there exists* $k \in R$ *for which* $a = kb$ *then we say* a *is a* **multiple** *of* b *or equivalently that* b **divides** a *and we write* $b \mid a$*. When* $b \mid a$ *we also may state that* b *is a* **factor** *of* a*. If no* $k \in R$ *exists for which* $a = kb$ *then we write* $b \nmid a$ *to express that* b *does not divide* a*.*

(2.) *If* R *is a unital ring and* $a \in R$ *has* $b \in R$ *for which* $ab = 1$ *then we write* $b = a^{-1}$ *and say that* a *is a* **unit***.*

(3.) *Suppose* $a,b \in R$ *and* $a,b \neq 0$*. If* $ab = 0$ *then we say* a *and* b *are* **zero-divisors***.*

You might worry the notation a^{-1} is ambiguous, I mean, what if there are two multiplicative inverses to a? Fortunately, for the same reasons as we covered in Lecture 1:

Theorem 3.1.4. *If R is a unital ring then the multiplicative identity is unique and each unit in R has a unique multiplicative inverse.*

Proof: see Theorems 1.1.12 and 1.1.10. □

It is convenient to use our usual notation for repeated addition or subtraction: for $n \in \mathbb{N}$,

$$n \cdot a = \underbrace{a + a + \cdots + a}_{n-summands} \quad \& \quad -n \cdot a = n \cdot (-a) = \underbrace{(-a) + (-a) + \cdots + (-a)}_{n-summands}.$$

We could define the operation above recursively, $(n+1)\cdot a = n\cdot a + a$ for each $n \in \mathbb{N}$ and $0 \cdot a = 0$. Of course, in the case $R = \mathbb{Z}$ we have $n \in \mathbb{Z}$ and what is described above is merely the multiplication of the ring.

Example 3.1.5. *$\mathbb{Z}$ forms a ring with respect to the usual addition and multiplication.*

Example 3.1.6. *$\mathbb{Z}_n$ forms a ring with respect to the usual modular addition and multiplication.*

Example 3.1.7. *Polynomials in x with integer coefficients are denoted $\mathbb{Z}[x]$. Since the sum and product of polynomials with integer coeffients is once more a polynomial with integer coefficients the usual addition and multiplication of polynomials provide binary operations on $\mathbb{Z}[x]$.*

Example 3.1.8. *Let S be a set and let $\mathcal{F}(S, R)$ denote functions with domain S and range a ring R. Add and multiply functions by the usual point-wise rules; $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for each $x \in S$. Because $f(x), g(x) \in R$ we can show $\mathcal{F}(S, R)$ forms a ring where the zero is the constant zero function. Moreover, if R is commutative then $\mathcal{F}(S, R)$ is commutative. Likewise, if R is unital with $1 \in R$ then $\mathbb{I}(x) = 1$ for each $x \in S$ defines $1 \in \mathcal{F}(S, R)$ as $(\mathbb{I}f)(x) = \mathbb{I}(x)f(x) = 1f(x) = f(x)$ for each $x \in S$ hence $\mathbb{I}f = f$ for each $f \in \mathcal{F}(S, R)$. In invite the reader to verify the rest of the ring properties for $\mathcal{F}(S, R)$ naturally follow from those given for R.*

In an initial discussion of rings the example below would be out of place, but, given our discussion thus far we should mention:

Example 3.1.9. *Any field $\mathbb{F}$ is a ring with respect to the usual multiplication and addition in the field. For example, $\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ or $\mathbb{Z}_p$ where p is prime.*

The following pair of examples give us a two interesting ways of building new rings from old.

Example 3.1.10. *Let R be a ring then define $R^{n\times n}$ to be square matrices with entries from R. The sum and product of such matrices are naturally defined by the usual formulas from linear algebra:*

Example 3.1.11. *Let $R_1, R_2, \ldots, R_k$ be a rings then define*

$$R = R_1 \times R_2 \times \cdots \times R_k$$

with the i-th component's addition and multiplication given by R_i for each $i = 1, \ldots, k$. That is:

$$(x_1, x_2, \ldots, x_k) + (y_1, y_2, \ldots, y_k) = (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k)$$

and

$$(x_1, x_2, \ldots, x_k)(y_1, y_2, \ldots, y_k) = (x_1y_1, x_2y_2, \ldots, x_ky_k)$$

for each $(x_1, \ldots, x_k), (y_1, \ldots, y_k) \in R_1 \times \cdots \times R_k$. We say R is the direct product ring of $R_1, R_2, \ldots, R_k$.

Th wealth of examples above just scratches the surface of rings. In any event, the axioms of the ring once verified for any of the examples above then immediately reward us with the properties given below. The reasoning is simple; the properties below are those of rings, so, once you have a ring you get the power contained within it. Of course this is Theorem 12.1 on page 231 of Gallian. I often use this Theorem without explicit reference in the remainder of this work.

Theorem 3.1.12. *Suppose $a, b, c \in R$ where R is a ring. Then,*

(1.) $a0 = 0a = 0$

(2.) $a(-b) = (-a)b = -(ab)$

(3.) $(-a)(-b) = ab$

(4.) $a(b-c) = ab - ac$ *and* $(b-c)a = ba - ca$

(5.) *if R is unital with unity 1 then* $(-1)a = -a$ *and* $(-1)(-1) = 1$.

Proof: to prove (1.) notice that by the definition of a ring,

$$0 + a0 = a0 = a(0+0) = a0 + a0 \quad \Rightarrow \quad a0 = 0.$$

where I used the cancellation property for the additive group $(R, +)$ in the last implication. On the other hand,

$$0 + 0a = 0a = (0+0)a = 0a + 0a \quad \Rightarrow \quad 0a = 0.$$

To prove (2.), use (.1) with $0 = a + (-a)$ hence

$$0 = 0b = [a + (-a)]b = ab + (-a)b$$

by distributive law and hence $-(ab) = (-a)b$. The proof that $a(-b) = -(ab)$ is similar. I leave (3.) and (4.) to the reader. Consider R with unity 1 and $a \in R$, note $-1 \in R$ and $1 + (-1) = 0$ by the fact $(R, +)$ is an additive group. Notice, by (1.),

$$0a = 0 \quad \Rightarrow \quad [1 + (-1)]a = 0 \quad \Rightarrow \quad (1)a + (-1)a = 0 \quad \Rightarrow \quad a + (-1)a = 0$$

hence $-a + a + (-1)a = -a + 0$ or $0 + (-1)a = -a$ which gives $(-1)a = -a$. The proof that $(-1)(-1) = 1$ is left to the reader. □

I will not say it's an *easy* exercise. Rather, it is an exercise which *could* be easy. As a general rule on these sort of proofs, if you spend longer than 10 minutes, then you should stop and do something different for a while before trying again. They're not that hard, but, you have to choose the right path.

You might be wondering, if we study subgroups, is there also such a thing as a *subring*. Indeed!

Definition 3.1.13. *Let R be a ring. A subset $S \subseteq R$ is a* **subring** *of R if it is a ring with respect to the addition and multiplication of R.*

Naturally, there is a generalization of the one-step-subgroup test Theorem 1.2.16:

Theorem 3.1.14. *A nonempty subset S of a ring R is a subring if S is closed under subtraction and multiplication. That is, if for each $a, b \in S$ we have $a - b, ab \in S$ then S is a subring of R.*

Proof: suppose $S \subseteq R$ and $S \neq \emptyset$. Also, assume if $a, b \in S$ then $a - b, ab \in S$. Observe the one-step-subgroup test provides $(S, +)$ is a subgroup of $(R, +)$. Moreover, multiplication restricted to S is a binary operation since know $(a, b) \mapsto ab$ is a function on $S \times S \subseteq R \times R$ and $ab \in S$ is given hence multiplication on S is a binary operation. Furthermore, the multiplication on S satisfies the ring axioms since it the multiplication on R satisfies the ring axioms. For example, for $a, b, c \in S$ we note $a, b, c \in R$ as well hence

$$a(b + c) = ab + ac \qquad (b + c)a = ba + ca.$$

I leave the remaining details to the reader, they should very unsurprising. □

Notice, there is nothing about unity in the subring test. Only zero is certain to be in a subring.

Example 3.1.15. *Consider $R = \mathbb{Z}$ and $S = 2\mathbb{Z}$. Notice if $2x, 2y \in S$ then $2x - 2y = 2(x - y) \in S$ and $(2x)(2y) = 2(2xy) \in S$. Since $2 \in 2\mathbb{Z} = S$ we note $S \neq \emptyset$ hence by subring test S is a subring of R. Notice, R is a unital ring whereas S is not unital. It is also nice to notice $2\mathbb{Z} + 1$ does not form a subring of the integers as $0 \notin 2\mathbb{Z} + 1$.*

I think the next example is helpful to remove a likely misconception.

Example 3.1.16. *Consider $R = \mathbb{Z}_6$ and $S = 2\mathbb{Z}_6 = \{0, 2, 4\}$. For essentially the same reasons as the last example, S is a subring of R. Of course, in this case you could make a pair of Cayley tables to check out the way addition and multiplication work for S. Let's look at the tables for fun:*

$+$	0	2	4
0	0	2	4
2	2	4	0
4	4	0	2

&

	0	2	4
0	0	0	0
2	0	4	2
4	0	2	4

Notice, $4s = s4 = s$ for each $s \in S$. In other words, 4 is the multiplicative identity in the subring S. In contrast, $1r = r1 = r$ for each $r \in \mathbb{Z}_6$.

Remark 3.1.17. The unity of R need not be the unity of a subring S of R. In constrast, R and all its subrings share the same zero. Multiplication is a more subtle in rings than addition.

I leave the proof of these claims to the reader:

Example 3.1.18. *If $R = \mathbb{Z}_n$ then $S = m\mathbb{Z}_n$ forms a subring.*

Example 3.1.19. *If $R = \mathbb{C}$ then $S = \mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ forms the ring of* **Gaussian Integers**. *We can show S is a subring of the complex numbers.*

What follows is an abstraction of Example 12 in Gallian on page 233-234.

Example 3.1.20. *Following Example 3.1.8 let S be a nonempty set and R a ring. Define $R' = \mathcal{F}(S, R)$ to be the set of R-valued functions of S. Pick $s_o \in S$ and define:*

$$S = \{f \in R' \mid f(s_o) = 0\}$$

you can show S is a subring of R'.

Example 3.1.21. *Following Example 3.1.10 we study $R' = R^{n \times n}$ where R is a given ring. Define S to be the set of diagonal matrices in R'. You can show the difference and product of diagonal matrices in R' is once more in R' thus S is a subring of R'.*

Of course, we could give many more examples, but this will suffice for our current needs. I should mention the subring lattice diagram on page 234 is pretty. I will try to draw that in class. This brings us to Chapter 13.

Definition 3.1.22. *Let R be a commutative ring with unity. We say R is an* **integral domain** *if R has no zero-divisors.*

Recall, zero-divisors are nonzero elements in R which multiply to produce 0. The term **integral** domain is largely due to the fact that $\mathbb{Z}$ is an integral domain. It is helpful to contrast zero divisors and units. In fact, I should make a definition before we go further, I leave the proof that $\mathrm{U}(R)$ forms a group to your homework.

Definition 3.1.23. *Let R be a commutative ring with unity. The set of all units in R is denoted $U(R)$. In particular, $U(R) = \{r \in R \mid \text{there exists } s \in R \text{ with } sr = 1\}$.*

Example 3.1.24. *$\mathbb{Z}$ forms a ring with respect to the usual addition and multiplication. $\mathbb{Z}$ has no zero-divisors as $ab = 0$ implies either a or b is zero. Furthermore, 1 is the multiplicative identity and the equation $ab = 1$ has only the solutions $a = b = 1$ or $a = b = -1$. You could say $U(\mathbb{Z}) = \{-1, 1\}$. In other words, the group of units in $\mathbb{Z}$ is just the two-element multiplicative group $\{-1, 1\}$.*

Example 3.1.25. *$\mathbb{Z}_n$ forms a ring with respect to the usual modular addition and multiplication. There may be zero divisors in the case that n is not prime. For example, in $\mathbb{Z}_6$ we have $3(2) = 0$ with $2, 3 \neq 0$. On the other hand, $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ has no zero divisors. In $\mathbb{Z}_n$ you can prove that each element is either 0, a unit, or a zero-divisor. We have already spent considerable effort to study $U(n)$. Now we recognize $U(n)$ is the set of units in $\mathbb{Z}_n$.*

Example 3.1.26. *Recall, polynomials in x with integer coefficients are denoted $\mathbb{Z}[x]$. If $f(x)g(x) = 0$ the either $f(x) = 0$ or $g(x) = 0$ hence $\mathbb{Z}[x]$ has no zero-divisors and we observe $\mathbb{Z}[x]$ is an integral domain. The polynomial $\mathbb{I}(x) = 1$ serves as the multiplicative identity and you can see $f(x)g(x) = \mathbb{I}(x) = 1$ forces us to set $f(x) = g(x) = \pm 1$. In short $U(\mathbb{Z}[x]) = \{-1, 1\}$.*

Example 3.1.27. *The ring $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$ forms an integral domain.*

Example 3.1.28. *The ring $\mathbb{Z}_p$ where p is prime forms an integral domain. But, $\mathbb{Z}_n$ where $n = mk$ for $m, k \neq 1$ does not form an integral domain since $mk = 0$ in $\mathbb{Z}_n$ yet $m, k \neq 0$.*

Example 3.1.29. *If R is a ring then it is not generally the case that $R^{n\times n}$ forms an integral domain. We can have the product of nonzero matrices form a zero matrix. The group of units in $R^{n\times n}$ was defined in an earlier lecture,*

$$U(R^{n\times n}) = GL(n, R) = \{A \in R^{n\times n} \mid det(A) \in U(R)\}$$

Example 3.1.30. *$\mathbb{Z} \times \mathbb{Z}$ is not an integral domain since $(a, 0)(0, b) = (0, 0)$ for any $a, b \in \mathbb{Z}$. The ring $\mathbb{Z} \times \mathbb{Z}$ has many zero-divisors. This is not special to $\mathbb{Z}$, generally direct product rings are not integral domains.*

An important property of integral domains is **cancellation**.

Theorem 3.1.31. *Let a, b and c belong to an integral domain. If $a \neq 0$ and $ab = ac$ then $b = c$.*

Proof: suppose a, b, c are elements of an integral domain and $a \neq 0$. If $ab = ac$ then $ab - ac = 0$ hence $a(b-c) = 0$ hence $b - c = 0$ as $a \neq 0$ and there are no zero-divisors in an integral domain. □

Notice a need not be a unit for cancellation to hold. We may not be able to multiply $ab = ac$ by a^{-1}, yet, the cancellation still is valid for $a \neq 0$ in an integral domain. For example, if $(x+3)(x^2+1) = (x+3)f(x)$ then the fact that $\mathbb{Z}[x]$ is an integral domain allows us to immediately conclude $f(x) = x^2 + 1$. Of course, if you pay close attention to what I have carefully shown about polynomials thus far, you should realize I haven't earned this claim in all honesty. We will later spend some time developing polynomials with some care. Until then, I will continue to make vague claims about what you know from past classes.

Definition 3.1.32. *A commutative ring with unity for which every nonzero element is a unit is called a* **field**.

It is easy to see a field is an integral domain. Suppose $ab = 0$ for $a \neq 0$. In a field, a^{-1} exists hence $a^{-1}ab = a^{-1}0$ which yields $b = 0$.

Theorem 3.1.33. *Each finite integral domain is a field.*

Proof: suppose D is a finite integral domain with unity $1 \in D$. Let $a \in D$ with $a \neq 0$. If $a = 1$ then $a^{-1} = 1$ since $1(1) = 1$. Otherwise, $a \neq 1$ and we notice the list $a, a^2, a^3, \ldots$ must eventually loop back to itself since D is finite. In other words, finiteness of D suggests the existence of i, j for which $a^i = a^j$. Thus, $a^{i-j} = 1$ which provides $a^{i-j-1}a = 1$ hence $a^{-1} = a^{i-j-1}$. Thus, every nonzero element of D is a unit and we conclude D is a field. □

The proof above is charmingly simple. I suppose we already know the following from our work on $\mathrm{U}(n)$, but, it's nice to see it as part of the natural flow of ideas here:

Corollary 3.1.34. *If p is prime then $\mathbb{Z}_p$ is a field.*

Proof: in $\mathbb{Z}_p$ if $ab = 0$ then $ab = pk$ for some $k \in \mathbb{Z}$ hence by Euclid's Lemma either p divides a or b which implies either $a = 0$ or $b = 0$ in $\mathbb{Z}_p$. Thus $\mathbb{Z}_p$ is a finite integral domain and is hence a field by Theorem 3.1.33. □

I'll follow Gallian page 243 where he introduces $\mathbb{Z}_3[i]$ the Gaussian integers modulo 3 and $\mathbb{Q}[\sqrt{2}]$ the rational numbers with the square root of two adjoined.

Example 3.1.35. *Define $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$ hence*

$$\mathbb{Z}_3[i] = \{0, 1, 2, i, 1+i, 2+i, 2i, 1+2i, 2+2i\}$$

To see this is an integral domain, suppose

$$(x + yi)(a + bi) = xa - yb + i(xb + ya) = 0$$

hence $xa - yb = 0$ and $xb + ya = 0$. Let's look at these linear equations as:

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \frac{1}{x^2+y^2} \begin{bmatrix} x & y \\ -y & x \end{bmatrix} \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} = 0.$$

where $x + yi \neq 0$ implies $x^2 + y^2 \neq 0$. In short, if $x + yi \neq 0$ and $(x + yi)(a + bi) = 0$ then we are forced to conclude $a + bi = 0$, hence no zero-divisors in $\mathbb{Z}_3[i]$ exist and thus $\mathbb{Z}_3[i]$ is a field with 9 elements. The multiplication table for the 8 nonzero elements is given on page 243 of Gallian.

Example 3.1.36. *Define* $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. *This is not a finite integral domain! Yet,*

$$(a + b\sqrt{2})(x + y\sqrt{2}) = ax + 2by + (ay + bx)\sqrt{2}$$

and of course $(a + b\sqrt{2}) + (x + y\sqrt{2}) = (a + x) + (b + y)\sqrt{2}$ *hence* $\mathbb{Q}[\sqrt{2}]$ *is closed under addition and multiplication. Furthermore, if* $a + b\sqrt{2} \neq 0$ *then we can solve* $(a + b\sqrt{2})(x + y\sqrt{2}) = 1$ *in* $\mathbb{R}$ *and derive*

$$x + y\sqrt{2} = \frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{(a + b\sqrt{2})(a - b\sqrt{2})} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$$

hence $(a+b\sqrt{2})^{-1} = \frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2}$ *and as* $a^2-2b^2 \neq 0$ *for* $a, b \in \mathbb{Q}$ *we note* $\frac{a}{a^2-2b^2}, -\frac{b}{a^2-2b^2} \in \mathbb{Q}$. *Therefore, we've shown every nonzero element in* $\mathbb{Q}[\sqrt{2}]$ *is a unit. The field* $\mathbb{Q}[\sqrt{2}]$ *is larger than* $\mathbb{Q}$ *but, still much smaller than* $\mathbb{R}$ *which contains many more irrational numbers.*

Definition 3.1.37. *The* **characteristic** *of a ring* R *is the smallest positive integer for which* $nx = 0$ *for all* $x \in R$. *We denote the character of* R *by* $char(R) = n$. *If no such integer exists then we say* $char(R) = 0$.

In practice, we usually can judge the character of a ring by how its identity behaves.

Theorem 3.1.38. *If* R *is a ring with unity* 1 *then* R *has characteristic zero if* 1 *has infinite order. If* 1 *has additive order* n *then* $char(R) = n$.

Proof: If 1 has infinite additive order then there is no positive n for which $n \cdot 1 = 0$ and hence R has characteristic zero. Otherwise, suppose $|1| = n$ in the additive sense. That is $n \cdot 1 = 0$ and n is the least positive integer for which we obtain 0. Calculate,

$$n \cdot x = \underbrace{x + x + \cdots + x}_{n-summands} = 1x + 1x + \cdots + 1x = (1 + 1 + \cdots + 1)x = (n \cdot 1)x = 0x = 0.$$

therefore $\text{char}(R) = n$. □

Theorem 3.1.39. *The characteristic of an integral domain is either* 0 *or a prime.*

Proof: notice if 1 has infinite order than $\text{char}(R) = 0$ and we're done. So, suppose $n \cdot 1 = 0$ where $|1| = n$ in the additive sense. Let us suppose $n = st$ for some $1 \leq s, t \leq n$. Calculate,

$$0 = n \cdot 1 = (st) \cdot 1 = (s \cdot 1)(t \cdot 1)$$

hence either $s \cdot 1 = 0$ or $t \cdot 1 = 0$ thus either $s = n$ and $t = 1$ or $s = 1$ and $t = n$ since $|1| = n$. We've determined factors of n are 1 and n hence n is prime. □

Please study the table on page 246. It is a great summary of examples with attention to unity, commutativity, integral domain, field and characteristic.

I recommend working on Exercises for Gallian Chapter 12 such as: (page 234-237)

$$\#1 - 49 \ (odds).$$

and Exercises for Gallian Chapter 13 such as: (page 246-249)

$$\#1 - 59 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 23: (these are collected at the start of Lecture 25)

Problem 85: Gallian number 19 from page 235.

Problem 86: Gallian number 22 from page 236.

Problem 87: Gallian number 23 from page 236.

Problem 88: Gallian number 40 from page 237.

Problem 89: Gallian number 43 from page 237.

Problem 90: Gallian number 48 from page 237.

3.2 Lecture 24: ideals and factor rings

As usual, we follow Gallian, we're in Chapter 14 now. In this lecture we study the concept of quotients of rings. What follows you might expect us to call a *normal subring*, but, that is not a commonly used term. It happens that group theory is the oddball here. As you study other abstract algebraic systems, it is typically the case that subobjects which allow natural quotients are called **ideals**. This terminology goes back to Kummer in the late 1800's who introduced **ideal numbers** to repair the failure of certain algebraic numbers. It was left to Dedekind to clarify and make rigorous the somewhat fuzzy work of Kummer[1]

Definition 3.2.1. *A subring A of a ring R is called an* **ideal** *if for every $r \in R$ and every $a \in A$ both $ar \in A$ and $ra \in A$. We say A is a proper ideal if $A \neq R$.*

Later in this Lecture we will see that this definition allows us well-defined operations on R/A. Notice, in the case R is commutative $ar = ra$ so we only have to check one thing. This closure of A by elements inside and outside A has been called *uber-closure* by some.

Theorem 3.2.2. *A nonempty subset A of a ring R is an ideal of R if*

- **(i.)** *if $a, b \in A$ then $a - b \in A$*
- **(ii.)** *if $a \in A$ and $r \in R$ then $ar, ra \in A$*

Proof: observe $A \subseteq R$ hence $(i.)$ and $(ii.)$ provide $a, b \in A$ implies $a - b, ab \in A$ hence A is a subring by Theorem 3.1.14. Furthermore, (ii.) provides that A is an ideal. □

Example 3.2.3. *Observe for R a ring we always have R and $\{0\}$ as ideals since conditions (i.) and (ii.) of Theorem 3.2.2 are easily verified.*

Example 3.2.4. *Consider $R = \mathbb{Z}$ then $n\mathbb{Z}$ forms an ideal for any $n \in \mathbb{N}$. Suppose $a, b \in n\mathbb{Z}$ then $a = na'$ and $b = nb'$ for some $a', b' \in \mathbb{Z}$ and*

$$a - b = na' - nb' = n(a' - b') \in n\mathbb{Z}$$

and for $x \in \mathbb{Z}$,

$$ax = xa = na'x \in n\mathbb{Z}$$

Thus, noting $0 = n(0) \in n\mathbb{Z} \neq \emptyset$, $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$ by Theorem 3.2.2.

Gallian hid this definition within Example 3 on page 254, but, I know this needs further emphasis. The concept of a set generated by an element in a ring is a central idea for many future lectures. It is important to carefully understand this concept as soon as possible.

Definition 3.2.5. *Let R be a commutive ring with unity and let $a \in R$ then we denote the* **principal ideal generated by** *a by $\langle a \rangle = \{ar \mid r \in R\} = aR = Ra$*

Example 3.2.6. *Let $n \in \mathbb{N}$ then $\langle n \rangle = \{nz \mid z \in \mathbb{Z}\} = n\mathbb{Z}$. We have shown this is an ideal of $\mathbb{Z}$.*

Theorem 3.2.7. *Let R be a commutative ring with unity and $a \in R$ then $\langle a \rangle$ is an ideal of R.*

[1] this is what I gather from reading Stillwell's *Elements of Number Theory* where there is much discussion of the ideal numbers and algebraic integers.

Proof: if $x, y \in \langle a \rangle$ then $x = ax'$ and $y = ay'$ for some $x', y' \in R$. Thus, $x - y = ax' - ay' = a(x' - y') \in \langle a \rangle$ as $x' - y' \in R$. Also, for $r \in R$,

$$xr = ax'r \in \langle a \rangle$$

since $x', r \in R$ implies $x'r \in R$. Finally, $a(0) = 0 \in \langle a \rangle \neq \emptyset$ thus $\langle a \rangle$ forms an ideal of R by Theorem 3.2.2. □

You should appreciate this gives us a very nice way to prove certain ideals are ideal. This is very much the analog of the span is a subspace theorem in linear algebra.

Example 3.2.8. *Consider the ring of polynomials with real coefficients: $\mathbb{R}[x]$. The ideal $\langle x \rangle$ is the set of polynomials with zero constant term.*

Example 3.2.9. *If we consider $\mathbb{R}[x, y]$ to be bivariate polynomials then $\langle x \rangle$ would be the ideal of polynomials which have zero constant term and no terms with just y. For example, $f(x, y) = y + y^3 + x^2y^2 \in \mathbb{R}[x, y]$, but, $f(x, y) \notin \langle xy \rangle$. Anything in $\langle xy \rangle$ has the form $xyg(x, y)$ for some $g(x, y) \in \mathbb{R}[x, y]$.*

This definition is given inside Example 5 on page 254 of Gallian.

Definition 3.2.10. *Let R be a commutive ring with unity and let $a_1, a_2, \ldots, a_n \in R$ then we denote the* **ideal generated by** *$a_1, a_2, \ldots, a_n$ by*

$$\langle a_1, a_2, \ldots, a_n \rangle = \{a_1r_1 + a_2r_2 + \cdots + a_nr_n \mid r_1, r_2, \ldots, r_n \in R\}.$$

It is not wrong to call $\langle a_1, a_2, \ldots, a_n \rangle$ an ideal:

Theorem 3.2.11. *If R is a commutive unital ring and $a_1, \ldots, a_n \in R$ then $\langle a_1, \ldots, a_n \rangle$ is an ideal.*

Proof: I leave this to the reader as an exercise. □

The notation $\langle a_1, \ldots, a_n \rangle$ allows concise description of many interesting ideals.

Example 3.2.12. *Consider $\mathbb{Z}[x]$ which is a commutative, unital ring. The ideal $\langle x, 2 \rangle$ is the set of polynomials with even constant term. To see this, note: $f(x) \in \langle x, 2 \rangle$ means there exist $h(x), j(x) \in \mathbb{Z}[x]$ for which*

$$f(x) = xh(x) + 2j(x)$$

If $h(x) = h_o + h_1x + \cdots + h_kx^k$ and $j(x) = j_o + j_1x + \cdots + j_lx^l$ then

$$f(x) = x(h_o + h_1x + \cdots + h_kx^k) + 2(j_o + j_1x + \cdots + j_lx^l) = 2j_o + (h_o + 2j_1)x + \cdots$$

Example 3.2.13. *Observe $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ a function}\}$ forms a ring by the pointwise addition and multiplication of functions. Since the product and difference of differentiable functions on $\mathbb{R}$ is once more differentiable it follows the set of differentiable functions forms a subring. However, it is not the case that the product of* **any** *function in $\mathcal{F}(\mathbb{R}, \mathbb{R})$ with a differentiable function is differentiable. For example, $f(x) = 1$ has $f'(x) = 0$ and $g(x) = |x|$ defines a function yet $f(x)g(x) = |x|$ is not everywhere differentiable. In short, the set of differentiable functions is a subring, but, it is not an ideal of $\mathcal{F}(\mathbb{R}, \mathbb{R})$.*

Theorem 3.2.14. *Let R be a ring and A a subring of R then define for set of cosets $R/A = \{r + A \mid r \in R\}$ the following possible operations*

$$(s + A) + (t + A) = (s + t) + A \qquad \& \qquad (s + A)(t + A) = st + A$$

for each $s, t \in R$. Then R/A forms a ring with respect to the above operations if and only if A is an ideal.

Proof: suppose A is an ideal of R then A is a normal subgroup of $(R.+)$ thus $(R/A, +)$ forms a factor group with respect to addition. It remains to show R/A has a multiplication which is well-defined. Suppose there exist $s, s', t, t' \in R$ for which $s + A = s' + A$ and $t + A = t' + A$. Hence, $t - t' = a \in A$ and $s - s' = b \in A$. Consider,

$$st = (b + s')(a + t') = ba + s'a + bt' + s't'$$

now, $a, b \in A$ gives $ba \in A$ as A is a subring. Moreover, using the closure of A under multiplication on the left or right by ring elements we find $s'a \in A$ and $bt' \in A$. Since A is a subring it follows that $ba + s'a + bt' \in A$ thus $st - s't' \in A$ and we find $st + A = s't' + A$ which shows the proposed multiplication on R/A is well-defined. Verification of the associative and distributive properties are straightforward and rest squarely on the respective properties of R: for $r, s, t \in R$,

$$(r + A)((s + A)(t + A)) = (r + A)(st + A) = r(st) + A = (rs)t + A = ((r + A)(s + A))(t + A).$$

and distributivity is similarly shown: the key step is where I use corresponding distributivity of R

$$\begin{aligned}(r + A)((s + A) + (t + A)) &= (r + A)((s + t) + A)\\ &= r(s + t) + A\\ &= rs + rt + A\\ &= (rs + A) + (rt + A)\\ &= (r + A)(s + A) + (r + A)(t + A)\end{aligned}$$

The proof of left distributivity is similar and I omit it. Thus R/A forms a ring known as the **Factor Ring of R by A**.

Conversely, suppose A is not an ideal. If A is not a subring then R/A is not an additive group with respect to addition of cosets. So, suppose A is subring, but, not an ideal. We have the additive structure on R/A, but there exist $r \in A$ and $a \in A$ for which ra or ar are not in A. Take the case $ar \notin A$. Observe, $a + A = 0 + A$ and $r + A$ are elements of R/A and yet $(a + A)(r + A) = ar + A$ and $(0 + A)(r + A) = 0r + A = A$ are at odds since $ar + A \neq A$ since we have assumed $ar \notin A$. Similar argument can be offered for the case $ra \notin A$. In any event, the multiplication on R/A is spoiled when A is not an ideal. □

Example 3.2.15. *Consider $R = \mathbb{Z}$ and $A = 3\mathbb{Z}$ then $R/A = \{3\mathbb{Z}, 1 + 3\mathbb{Z}, 2 + 3\mathbb{Z}\}$. We have:*

$+$	$3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$
$3\mathbb{Z}$	$3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$
$1+3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$	$3\mathbb{Z}$
$2+3\mathbb{Z}$	$2+3\mathbb{Z}$	$3\mathbb{Z}$	$1+3\mathbb{Z}$

&

$\cdot$	$3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$
$3\mathbb{Z}$	$3\mathbb{Z}$	$3\mathbb{Z}$	$3\mathbb{Z}$
$1+3\mathbb{Z}$	$3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$
$2+3\mathbb{Z}$	$3\mathbb{Z}$	$2+3\mathbb{Z}$	$1+3\mathbb{Z}$

You can compare $\mathbb{Z}/3\mathbb{Z}$ to the Cayley tables for $\mathbb{Z}_3$. In fact, these are the same object as we have defined them. Our usual notation is $3\mathbb{Z} = [0]_3$ and $1 + 3\mathbb{Z} = [1]_3$ or simply $0, 1 \in \mathbb{Z}_3$ as is oft convenient exposition. Notation aside, $\mathbb{Z}/n\mathbb{Z}$ is a quotient ring of $\mathbb{Z}$ by the principal ideal $n\mathbb{Z}$.

Example 3.2.16. *Let $R = \mathbb{Z}^{2\times 2}$ then consider $A = 2\mathbb{Z}^{2\times 2}$ which is the set of 2×2 matrices with even entries. If $X, Y \in A$ then $X = 2X'$ and $Y = 2Y'$ where $X', Y' \in R$ thus,*

$$X - Y = 2X' - 2Y' = 2(X' - Y') \in A$$

and for $Z \in R$,

$$XZ = 2X'Z \in A$$

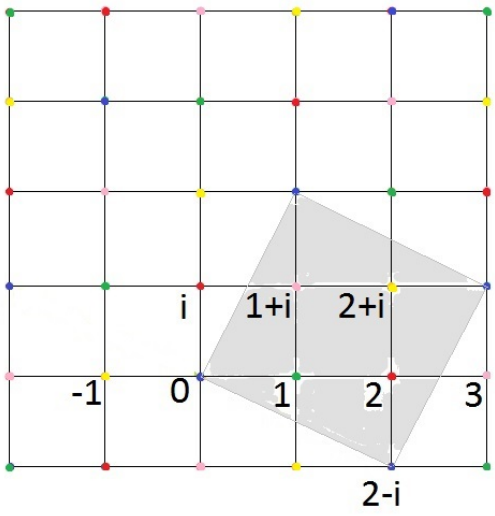
as $0 \in A \neq \emptyset$ we find A is an ideal of R. Gallian claims that R/A is a ring with 16 elements. Think about the uniqueness of representatives, the following are distinct since they differ by a matrix which is not in A:

$$R/A = \{A, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + A, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + A, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + A, \cdots, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + A, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + A\}$$

We have 4 entries with 2 choices each so $2^4 = 16$ distinct cosets.

Example 3.2.17. *Consider the Gaussian integers $\mathbb{Z}[i]$ and the principal ideal $\langle 2 - i \rangle$. Let us try to understand the structure of cosets $a + bi + \langle 2 - i \rangle$. The key here is that $2 + \langle 2 - i \rangle = i + \langle 2 - i \rangle$ since $2 - i \in \langle 2 - i \rangle$. So, for simplifying representatives we have the very simple rule $i = 2$. Thus, $a + ib$ and $a + 2b$ are representatives of the same coset in $\mathbb{Z}[i]/\langle 2 - i \rangle$. That said, I think a geometric approach is most clear for this example. Consider $\mathbb{Z}[i]$ as the lattice $\mathbb{Z} \oplus i\mathbb{Z}$ in the complex plane. Observe $\langle 2 - i \rangle$ has elements $0, 2 - i, -1(2 - i), i(2 - i), -i(2 - i)$ which simplify to $0, 2 - i, i - 2, 1 + 2i, -1 - 2i$. Any representative $a + ib$ can be shifted by some sum of the basic vectors $2 - i, i - 2, 1 + 2i, -1 - 2i$ as to obtain a different representative of the same coset. It turns out there are five such cosets. I used color coding to indicate these cosets*

- **(a.)** *blue is $\langle 2 - i \rangle$,*
- **(b.)** *green is $1 + \langle 2 - i \rangle$,*
- **(c.)** *red is $2 + \langle 2 - i \rangle$,*
- **(d.)** *yellow is $2 + i + \langle 2 - i \rangle = -1 + \langle 2 - i \rangle$*
- **(e.)** *pink is $1 + i + \langle 2 - i \rangle = 3 + \langle 2 - i \rangle$*



In summary, $\mathbb{Z}[i]/\langle 2 - i \rangle = \{[0], [1], [2], [3], [4]\}$ where we introduce the notation $[x] = x + \langle 2 - i \rangle$.

My way of understanding the quotient of $\mathbb{Z}[i]$ is rather different than that given by Gallian. Of course, I can't draw such pictures for most quotient rings. Often we have to calculate directly to understand the structure of the cosets. But, when our ring is a subring of $\mathbb{C}$ it is at least possible to do some direct visualization.

Example 3.2.18. *Let $\mathbb{R}[x]$ denote polynomials with real coefficients and consider the principal ideal $\langle x^2+1\rangle$:*

$$\langle x^2+1\rangle = \{(x^2+1)f(x) \mid f(x) \in \mathbb{R}[x]\}$$

Since $x^2 + \langle x^2+1\rangle = x^2+1-1+\langle x^2+1\rangle = -1+\langle x^2+1\rangle$ we find x^2 is congruent to -1 modulo x^2+1. It follows we can reduce all the higher powers of a given representative and just keep the linear terms. For example, using the notation $f(x)+\langle x^2+1\rangle = [f(x)]$,

$$[1+2x+x^4+x^2] = [1+2x+(x^2)^2-1] = [2x+(-1)^2] = [1+2x].$$

More generally, if $f(x) \in \mathbb{R}[x]$ then we may use the division algorithm for polynomials to find $q(x)$ and $r(x)$ such that

$$f(x) = q(x)(x^2+1)+r(x)$$

and $r(x) = ax+b$ for some $a, b \in \mathbb{R}$. Thus,

$$f(x)+\langle x^2+1\rangle = ax+b+\langle x^2+1\rangle$$

as $q(x)(x^2+1) \in \langle x^2+1\rangle$ hence we absorb it into the ideal. In summary,

$$\mathbb{R}[x]/\langle x^2+1\rangle = \{a+bx+\langle x^2+1\rangle \mid a, b \in \mathbb{R}\}$$

Observe,

$$[a+bx][c+dx] = [ac+adx+bcx+bdx^2] = [ac-bd+(bd+bc)x].$$

Compare this to the multiplication of $a+bi$ with $c+di$. Basically, x functions as i in this construction. This is one of the many ways to construct the complex number system, it was given by Cauchy in 1847.

The definition below is very important. We need to remember and absorb these terms for the remainder of our study of rings.

Definition 3.2.19. *Let R be a commutative ring and A a proper ideal of R,*

(i.) *A is a* **prime ideal** *of R if $a, b \in R$ and $ab \in A$ implies $a \in A$ or $b \in A$.*

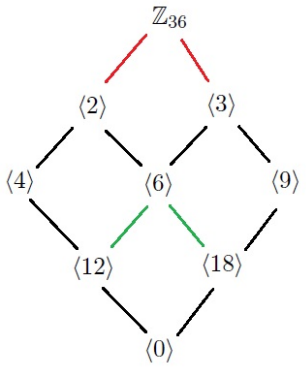
(ii.) *A is a* **maximal ideal** *of R if any ideal B of R with $A \subseteq B \subseteq R$ has $B = A$ or $B = R$.*

The terminology of prime naturally ties into the concept of prime we know from our work in $\mathbb{Z}$. Recall that Euclid's Lemma states that if a prime $p \mid ab$ then $p \mid a$ or $p \mid b$.

Example 3.2.20. *Let p be a prime and consider the ideal $p\mathbb{Z}$. If $a, b \in \mathbb{Z}$ and $ab \in p\mathbb{Z}$ then $ab = pk$ for some $k \in \mathbb{Z}$ hence $p \mid ab$ and thus $p \mid a$ or $p \mid b$ by Euclid's Lemma. If $p \mid a$ then $a = pn$ for some $n \in \mathbb{Z}$ and hence $a \in p\mathbb{Z}$. Likewise, $p \mid b$ then $b \in p\mathbb{Z}$. In summary, if p is prime then $p\mathbb{Z}$ is a prime ideal.*

I suppose I should mention, there is another way of defininig a prime ideal which helps make the correspondence between containment of ideals and divisibility of integers a bit more clear. See Lecture 22 of my Math 307 notes. You don't have to study that just yet, I mostly mention this to remind my Math 307 students how we treated prime ideals in terms of containment. I like Gallian's definition, pragmatically it is simple enough to use.

Example 3.2.21. *Consider $\mathbb{Z}_{36}$ the ideals $\langle 2\rangle$ and $\langle 3\rangle$ are maximal ideals in $\mathbb{Z}_{36}$. On the other hand, we also note $\langle 12\rangle$ and $\langle 18\rangle$ are maximal ideals in $\langle 6\rangle$. You can see the maximality in the lattice diagram below:*



You might notice $\mathbb{Z}_{36}/2\mathbb{Z}_{36} \approx \mathbb{Z}_2$ and $\mathbb{Z}_{36}/3\mathbb{Z}_{36} \approx \mathbb{Z}_3$ are both fields. What about $\langle 6\rangle/\langle 12\rangle$? I'll be explicit,

$$\langle 6\rangle = \{0, 6, 12, 18, 24, 30\} \qquad \& \qquad \langle 12\rangle = \{0, 12, 24\}$$

So, you can see,

$$\langle 6\rangle/\langle 12\rangle = \{\langle 12\rangle, 6 + \langle 12\rangle\} \approx \mathbb{Z}_2$$

Showing $\langle x^2 + 1\rangle$ is maximal in $\mathbb{R}[x]$ requires some careful calculation:

Example 3.2.22. *Let A be an ideal of $\mathbb{R}[x]$ for which $\langle x^2 + 1\rangle \subseteq A \subseteq \mathbb{R}[x]$ and $A \neq \langle x^2 + 1\rangle$. In other words, suppose $\langle x^2+1\rangle$ is properly contained in A. There exists $f(x) \in A$ and $f(x) \notin \langle x^2+1\rangle$. By the division of polynomials, there exists $q(x), r(x) \in \mathbb{R}[x]$ for which*

$$f(x) = q(x)(x^2 + 1) + r(x)$$

and $r(x) \neq 0$ and $r(x) = ax+b$. Note $r(x) \neq 0$ indicates at least one of a, b is nonzero. Furthermore,

$$ax + b = f(x) - q(x)(x^2 + 1) \in A$$

since $f(x) \in A$ and $q(x)(x^2 + 1) \in \langle x^2 + 1\rangle \subseteq A$ and A is an ideal. Moreover,

$$a^2x^2 - b^2 = (ax + b)(ax - b) \in A$$

since the produce of $ax + b \in A$ and $ax - b \in \mathbb{R}[x]$ must be in A again as A is an ideal. As $\langle x^2 + 1\rangle$ is contained in A we also may note $a^2(x^2 + 1) \in A$. Therefore,

$$0 \neq a^2 + b^2 = (a^2x^2 + a^2) - (a^2x^2 - b^2) \in A$$

But, $1 = \frac{1}{a^2+b^2}(a^2 + b^2) \in A$ hence $\langle 1\rangle \subset A$ and $\langle 1\rangle = \{(1)f(x) \mid f(x) \in \mathbb{R}[x]\} = \mathbb{R}[x]$. Therefore, $\langle x^2 + 1\rangle$ is a maximal ideal.

I followed Gallian on page 258-259 for the most part in the example above. Likewise, the next example is Gallian's Example 16 on page 259.

Example 3.2.23. *In $\mathbb{Z}_2[x]$ the ideal $\langle x^2+1\rangle$ is not a prime ideal as $(x+1)^2 = x^2+2x+1 = x^2+1 \in \langle x^2 + 1\rangle$ yet $x + 1 \notin \langle x^2 + 1\rangle$. To elaborate on the noncontainment claim, suppose $x + 1 \in \langle x^2 + 1\rangle$ for some $f(x) \in \mathbb{Z}_2[x]$ we need*

$$x + 1 = f(x)(x^2 + 1)$$

why can we not solve the above for appropriate $f(x) \in \mathbb{Z}_2[x]$?

Theorem 3.2.24. *Let R be a commutative ring with unity and let A be an ideal of R. The quotient ring R/A is an integral domain if and only if A is prime.*

Proof: suppose R is a unital commutative ring with ideal A in R. Suppose R/A is an integral domain. Let $a, b \in R$ and $ab \in A$. Note,

$$A = ab + A = (a + A)(b + A)$$

thus $a + A = A$ or $b + A = A$ as R/A has no zero divisors (here A serves as zero in R/A). Hence $a \in A$ or $b \in A$.

Conversely, suppose A is a prime ideal. We need to show R/A has no zero divisors. Suppose $(a + A)(b + A) = A$ then $ab + A = A$ hence $ab \in A$. But, A is prime hence $a \in A$ or $b \in A$ thus $a + A = A$ or $b + A = A$. Furthermore, denoting the unity of R as 1 we note that $(1 + A)(r + A) = 1r + A = r + A$ for each $r + A \in R/A$. Also, calculate $(r+A)(s+A) = rs+A = sr+A = (s+A)(r+A)$ hence R/A is a commutative ring. Therefore, R/A is an integral domain. □

Theorem 3.2.25. *Let R be a commutative ring with unity and let A be an ideal of R. The quotient ring R/A is a field if and only if A is maximal.*

Proof: suppose R is a commutative ring with unity $1 \in R$ and suppose A is an ideal of R. Assume R/A is a field. Consider an ideal B of R for which $A \subseteq B \subseteq R$ with $A \neq B$. It follows there exists $x \in B$ for which $x \notin A$ hence $x + A \neq A$ which means $x + A$ is a nonzero element in R/A. Since R/A is a field and $1 + A$ serves as the unity we have the existence of $y + A$ for which $(x + A)(y + A) = 1 + A$. Thus, $xy + A = 1 + A$ and we find $1 - xy \in A$. However, $x \in B$ implies $xy \in B$ as B is an ideal. Since $A \subseteq B$ we find $1 - xy \in B$. Thus,

$$xy + (1 - xy) = 1 \in B$$

But, $x = 1(x) \in B$ for each $x \in R$ hence $B = R$ and we find A is a maximal ideal.

Conversely, suppose A is a maximal ideal. Suppose $x \in R$ yet $x \notin A$. In other words, we consider a nonzero element $x + A$ in R/A. Construct,

$$B = \{xr + a \mid r \in R, a \in A\}$$

I'll leave it to the reader to verify that B is indeed an ideal of R. Moreover, if $a \in A$ then note $a = x(0) + a \in B$ thus $A \subseteq B$. By maximality of A we have $B = R$. Therefore, $1 \in B$ and we find there exists $r \in R, a \in A$ for which $xr + a = 1$ or $1 - xr = a \in A$. Observe, $(x+A)(r+A) = xr+A = 1+A$. Thus $x+A$ has multiplicative inverse $r+A$ in R/A. Furthermore, we note that $(1+A)(r+A) = 1r+A = r+A$ for each $r+A \in R/A$. Also, calculate $(r+A)(s+A) = rs+A = sr+A = (s+A)(r+A)$ hence R/A is a commutative ring with unity where every nonzero element has a multiplicative inverse. That is, R/A forms a field. □

Example 3.2.26. *Since a field is an integral domain it follows that a maximal ideal must be a prime ideal in view of Theorems 3.2.24 and 3.2.25. On the other hand, we can exhibit an ideal which is prime, but, not maximal. Consider $\langle x \rangle$ in $\mathbb{Z}[x]$ if $f(x), g(x) \in \mathbb{Z}[x]$ and $f(x)g(x) \in \langle x \rangle$ then $f(x)g(x) = xh(x)$ for some $h(x) \in \mathbb{Z}[x]$. It follows that x must be a factor in $f(x)$ or $g(x)$ thus $f(x) \in \langle x \rangle$ or $g(x) \in \langle x \rangle$ and we find $\langle x \rangle$ is a prime ideal of $\mathbb{Z}[x]$. Consider, $\langle x, 2 \rangle$ contains $\langle x \rangle$ since $\langle x, 2 \rangle = \{xf(x) + 2g(x) \mid f(x), g(x) \in \mathbb{Z}[x]\}$ so to obtain $\langle x \rangle$ simply select elements with $g(x) = 0$. On the other hand, $2 \in \langle x, 2 \rangle$ and $2 \notin \langle x \rangle$. Also, $1 \in \mathbb{Z}[x]$ and $1 \notin \langle 2, x \rangle$ hence $\langle x \rangle \subset \langle 2, x \rangle \subset \mathbb{Z}[x]$. This proves $\langle x \rangle$ is not maximal.*

I recommend working on Exercises for Gallian Chapter 14 such as: (page 260-263)

$$\#1 - 57 \ (odds).$$

and supplementary exercises for Chapters 12-14 such as: (page 267-269)

$$\#1 - 47 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 24: (these are collected at the start of Lecture 26)

Problem 91: Gallian number 16 from page 247.

Problem 92: Gallian number 24 from page 247.

Problem 93: Gallian number 35 from page 247.

Problem 94: Gallian number 34 from page 262.

Problem 95: Gallian number 45 from page 262.

Problem 96: Prove Theorem 3.2.11.

3.3 Lecture 25: ring homomorphism

We saw the concept of homomorphism allowed us connect groups which seemed the same in terms of their group structure. In the same way, the concept of ring homomorphism gives us a precise method to describe when two rings share similar structure. Or, in the case of isomorphism, the rings in question are, from the viewpoint of algebraic structure, the same. Much of this section directly echoes our previous work on groups, as such I will omit some proofs. In contrast, the field of quotients construction at the end of this Lecture is fascinating and new.

Definition 3.3.1. *A* **ring homomorphism** ϕ *from a ring* R *to a ring* S *is a function* $\phi : R \to S$ *which preserves the ring operations:*

(i.) $\phi(a+b) = \phi(a) + \phi(b)$ *for all* $a, b \in R$,

(ii.) $\phi(ab) = \phi(a)\phi(b)$ *for all* $a, b \in R$.

If ϕ *is a bijective ring homomorphism then* ϕ *is a ring isomorphism and we write* $R \approx S$

The meaning of $R \approx S$ should be clear from the context. We use $\approx$ to indicate an isomorphism of groups or rings as appropriate[2].

Example 3.3.2. *Consider* $\phi : \mathbb{Z} \to \mathbb{Z}_n$ *defined by* $\phi(x) = [x]_n$. *Observe,* ϕ *is a function since the domain is* $\mathbb{Z}$ *so there is no ambiguity in* $x \in \mathbb{Z}$[3].

$$\phi(x+y) = [x+y]_n = [x]_n + [y]_n = \phi(x) + \phi(y) \qquad \& \qquad \phi(xy) = [xy]_n = [x]_n[y]_n = \phi(x)\phi(y)$$

for all $x, y \in \mathbb{Z}$. *Thus* $\mathbb{Z}$ *and* $\mathbb{Z}_n$ *are homomorphic rings under the ring homomorphism* ϕ. *Incidentally, this is the* **natural homomorphism** *which also call the* **coset map** *since* $\mathbb{Z}_n$ *is the factor ring of* $\mathbb{Z}$ *by* $n\mathbb{Z}$ *and* $[x]_n = x + n\mathbb{Z}$, *so we could write* $\phi(x) = x + n\mathbb{Z}$.

Example 3.3.3. *The map* $\phi(z) = z^*$ *is a ring isomorphism from* $\mathbb{C}$ *to* $\mathbb{C}$ *with respect to the usual complex arithmetic where I intend the complex conjugate given by* $(x+iy)^* = x - iy$ *for* $x, y \in \mathbb{R}$. *You can check:*

$$(zw)^* = z^* w^* \qquad \& \qquad (z+w)^* = z^* + w^*$$

thus ϕ *is a ring homomorphism. In fact,* $\phi : \mathbb{C} \to \mathbb{C}$ *is an* **automorphism of** $\mathbb{C}$ *since* $\phi^{-1} = \phi$ *as* $(z^*)^* = z$ *for each* $z \in \mathbb{C}$. *You can verify,* $\phi^2 = Id$ *thus* ϕ *is an automorphism of order* 2.

My next example is an deeper version of Gallian's Example 3 on page 271.

Example 3.3.4. *The* **evaluation map** *is an important homomorphism which connects a ring* R *with polynomials* $R[x]$. *Pick* $a \in R$ *and define* $\phi_a(f(x)) = f(a)$ *for each* $f(x) \in R[x]$. *Observe,*

$$\phi_a((f+g)(x)) = (f+g)(a) = f(a) + g(a) = \phi_a(f(x)) + \phi_a(g(x))$$

and

$$\phi_a((fg)(x)) = (fg)(a) = f(a)g(a) = \phi_a(f(x))\phi_a(g(x))$$

thus $\phi_a : R[x] \to R$ *is a ring homomorphism.*

[2] currently we have no notation for homomorphism of groups or rings, I was toying with idea of changing $\approx$ to $\cong$ but, it was denied by the students... maybe next year

[3] in contrast, $g([x]_n) = x$ is rather disfunctional

Gallian's Examples 4,5,6,7,8 and 9 on page 271-272 are interesting. I will cover some of those in class.

Theorem 3.3.5. *Let $\phi : R \to S$ be a ring homomorphism from a ring R to a ring S. Let A be a subring of R and B an ideal of S*

- **(i.)** *for any $r \in R$ and $n \in \mathbb{N}$, $\phi(nr) = n\phi(r)$ and $\phi(r^n) = (\phi(r))^n$,*
- **(ii.)** *$\phi(A)$ is a subring of S*
- **(iii.)** *if A is an ideal and $\phi(R) = S$ then $\phi(A)$ is an ideal of S*
- **(iv.)** *$\phi^{-1}(B)$ is an ideal of R*
- **(v.)** *if R is commutative then $\phi(R)$ is commutative*
- **(vi.)** *if R has unity 1 and $S \neq \{0\}$ and ϕ is a surjection then $\phi(1)$ is the unity of S*
- **(vii.)** *ϕ is an isomorphism iff ϕ is surjective and $Ker(\phi) = \{r \in R \mid \phi(r) = 0\} = \{0\}$.*
- **(viii.)** *If $\phi : R \to S$ is a ring isomorphism of then $\phi^{-1} : S \to R$ is a ring isomorphism.*

Proof: similar to those given for groups. Main difference, for the multiplicative properties we cannot use the existence of inverses. However, if you study our proofs for the corresponding group claims then you'll see we can adopt those proofs with little modification. I will leave these proofs as exercises for the reader. □

Notice the **additive** kernel determines injectivity of the ring homomorphism. This is not surprising as $(R, +)$ enjoys the structure of an abelian group so the injectivity from trivial kernel is precisely our group theoretic theorem.

Theorem 3.3.6. *Let $\phi : R \to S$ be a ring homomorphism from a ring R to a ring S. Then $Ker(\phi) = \{r \in R \mid \phi(r) = 0\}$ is an ideal of R.*

Proof: suppose $\phi : R \to S$ is a ring homomorphism. Suppose $a, b \in \text{Ker}(\phi)$ then $\phi(a) = 0$ and $\phi(b) = 0$ consequently,

$$\phi(a - b) = \phi(a) - \phi(b) = 0 - 0 = 0,$$

and for $r \in R$,

$$\phi(ra) = \phi(r)\phi(a) = \phi(r)0 = 0 \qquad \& \qquad \phi(ar) = \phi(a)\phi(r) = 0\phi(r) = 0.$$

Thus $a - b \in \text{Ker}(\phi)$ and $ar, ra \in \text{Ker}(\phi)$ for all $a, b \in \text{Ker}(\phi)$ and $r \in R$. We find $\text{Ker}(\phi)$ is an ideal via Theorem 3.2.2. □

The first isomorphism theorem is also available for rings:

Theorem 3.3.7. *Let $\phi : R \to S$ be a ring homomorphism. Then the mapping from $R/Ker(\phi)$ to $\phi(R)$ given by $r + Ker(\phi) \mapsto \phi(r)$ is a ring isomorphism; $R/Ker(\phi) \approx \phi(R)$.*

Proof: exercise for the reader. □

The next theorem is also available for groups, but, I don't think I emphasized it much if at all this semester: I follow Gallian as usual, this is his Theorem 15.4 on page 274.

Theorem 3.3.8. *Every ideal of a ring R is the kernel of a ring homomorphism of R. In particular, an ideal A is the kernel of the mapping $r \mapsto r + A$ from R to R/A.*

Proof: if A is an ideal of R then the quotient ring R/A is well-defined and we construct $\pi : R \to R/A$ by $\pi(r) = r + A$. Observe,

$$\pi(r+s) = r+s+A = (r+A)+(s+A) = \pi(r)+\pi(s)$$

and

$$\pi(rs) = rs+A = (r+A)(s+A) = \pi(r)\pi(s)$$

for each $r, s \in R$. Morover, $\text{Ker}(\pi) = A$ hence A is the kernel of a ring homomorphism. □

Example 3.3.9. *Consider $\phi : \mathbb{Z}[x] \to \mathbb{Z}$ defined by $\phi(f(x)) = f(0)$. Since ϕ is a surjective ring homomorphism with $Ker(\phi) = \langle x \rangle$ we have by the first isomorphism theorem $\mathbb{Z}[x]/\langle x \rangle \approx \mathbb{Z}$. However, we know $\mathbb{Z}$ is an integral domain hence by Theorem 3.2.24 we find $\langle x \rangle$ is a prime ideal of $\mathbb{Z}[x]$. Indeed, by Theorem 3.2.25 we also see $\langle x \rangle$ is* **not** *maximal as $\mathbb{Z}$ is not a field.*

Theorem 3.3.10. *If R is a ring with unity 1 then the mapping $\phi : \mathbb{Z} \to R$ defined by $\phi(n) = n \cdot 1$ is a ring homomorphism.*

Proof: recall $n \cdot 1$ is a notation for n-fold additions of 1 for $n \in \mathbb{N}$ or k-fold additions of -1 if $k = -n \in \mathbb{N}$. The proof is given on page 274-275 of Gallian. Essentially, this affirms that:

$$(m+n)\cdot 1 = m\cdot 1 + n \cdot 1 \qquad \& \qquad (m\cdot 1)(n\cdot 1) = (mn)\cdot 1 \qquad \square$$

Corollary 3.3.11. *If R is a ring with unity 1 and $Char(R) = n > 0$ then R contains a subring which is isomorphic to $\mathbb{Z}_n$. If $Char(R) = 0$ then R contains a subring which is isomorphic to $\mathbb{Z}$.*

Proof: suppose R is unital. Construct

$$S = \{k \cdot 1 \mid k \in \mathbb{Z}\}$$

in view of from Theorem 3.3.10 we note $\phi(k) = k \cdot 1$ is a homomorphism of $\mathbb{Z}$ and R and by construction $\phi(R) = S$. Suppose $\text{Char}(R) = n$, then $\text{Ker}(\phi) = \{k \in \mathbb{Z} \mid k \cdot 1 = 0\} = n\mathbb{Z}$. Hence, by the first isomorphism theorem, $\mathbb{Z}/\text{Ker}(\phi) \approx \phi(R)$ which gives $\mathbb{Z}/n\mathbb{Z} \approx S$. If R has characteristic zero then $S \approx \mathbb{Z}/\langle 0 \rangle \approx \mathbb{Z}$. □

Corollary 3.3.12. *For any positive integer m, the mapping $\phi : \mathbb{Z} \to \mathbb{Z}_m$ defined by $\phi(x) = [x]_m$ is a ring homomorphism.*

Proof: note $[x]_m = [1 + 1 + \cdots + 1]_m = x \cdot [1]_m$ hence $\phi(x) = [x]_m$ is a mapping with the same form as that given in Theorem 3.3.10. □

The calculation in the Corollary above, the main point is that $[x]_m = x \cdot [1]_m$. We needed to make this same calculational observation in several past problems. For example, it is the heart of why homomorphisms from $\mathbb{Z}_n$ to $\mathbb{Z}_k$ have the form $[x]_n \mapsto [mx]_k$ where $k \mid mn$ (Problem 72).

Corollary 3.3.13. *(Steinitz, 1910): If $\mathbb{F}$ is a field of characteristic p then $\mathbb{F}$ contains a subfield which is isomorphic to $\mathbb{Z}_p$. If $\mathbb{F}$ is a field of characteristic 0, then $\mathbb{F}$ contains a subfield isomorphic to the rational numbers.*

Proof: if $\mathbb{F}$ is a field of characteristic p then as a field is also a ring by Corollary 3.3.11. Thus $\mathbb{F}$ contains a subring isomorphic to $\mathbb{Z}_p$. If $\mathbb{F}$ has characteristic 0 then $\mathbb{F}$ has a subring S isomorphic to $\mathbb{Z}$ and we can construct a copy of $\mathbb{Q}$ from S as follows:

$$S_{\mathbb{Q}} = \{ab^{-1} \mid a, b \in S \text{ with } b \neq 0\} \quad \square$$

Definition 3.3.14. *Given a field $\mathbb{F}$ the subfield of $\mathbb{F}$ which is contained in all other subfields of $\mathbb{F}$ is called the* **prime** *subfield of $\mathbb{F}$.*

We can argue from Steinitz Theorem that the prime subfield of $\mathbb{F}$ is either $\mathbb{Q}$ or $\mathbb{Z}_p$. Any field of characteristic zero has $\mathbb{Q}$ as its *smallest* subfield. Any field of prime p characteristic has $\mathbb{Z}_p$ as its smallest subfield.

Theorem 3.3.15. *Let D be an integral domain. Then, there exists a field F that contains a subring isomorphic to D.*

Proof: an explicit and beautiful construction, see page 277-278 of Gallian. I may change the notation a bit. The notation which Gallian uses is the notation we wish to use in eventuality, but, to begin we should divorce our thinking from the familar so we don't assume more than we ought from the notation.

Let D be an integral domain with 1 the unity in D. Let $S = \{(a,b) \mid a,b \in D,\ b \neq 0\}$. Define $(a,b) \sim (c,d)$ if[4] $ad = bc$. We prove $\sim$ forms an equivalence relation on S:

(i.) let $(a,b) \in S$ then $(a,b) \sim (a,b)$ since $ab = ba$ (D is a commutative ring)

(ii.) if $(a,b) \sim (c,d)$ then $ad = bc$ hence $cb = da$ thus $(c,d) \sim (a,b)$.

(iii.) if $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$ then $ad = bc$ and $cf = de$. Consider, by associativity of multiplication and the known data on a,b,c,d,e,f,

$$(ad)f = (bc)f = b(cf) = b(de)$$

Thus $(af)d = (be)d$ where $(c,d) \in S$ hence $d \neq 0$ and by the cancellation property of integral domains we find $af = be$ hence $(a,b) \sim (e,f)$

Therefore, $\sim$ is a reflexive, symmetric and transitive relation on S. Denote the equivalence class containing (a,b) by $[a,b] = \{(c,d) \mid (c,d) \sim (a,b)\}$. We claim that $S/\sim$ the set of equivalence classes of S under $\sim$ forms a field with respect to the following operations of addition and multiplication:

$$[a,b] + [c,d] = [ad + bc, bd] \qquad \& \qquad [a,b][c,d] = [ac, bd].$$

We must show these operations are well-defined since we used a representative to define the rule for an equivalence class. Suppose $(a,b) \sim (a',b')$ and $(c,d) \sim (c',d')$ hence $ab' = ba'$ and $cd' = dc'$. Observe that

$$[ad + bc, bd] = [a'd' + b'c', b'd'] \qquad \text{if and only if} \qquad (ad + bc)b'd' = bd(a'd' + b'c').$$

Thus consider:

$$(ad + bc)b'd' = (ab')(dd') + (cd')(bb') = (ba')(dd') + (dc')(bb') = bd(a'd' + b'c').$$

Therefore addition on $S/\sim$ is well-defined. Next, observe that

$$[ac, bd] = [a'c', b'd'] \qquad \text{if and only if} \qquad (ac)(b'd') = (bd)(a'c')$$

Consider then,

$$(ac)(b'd') = (ab')(cd') = (ba')(dc') = (bd)(a'c')$$

[4] yes, intuitively, we want (a,b) to model the fraction a/b whatever that means... surely $a/b = c/d$ gives $ad = bc$ hence this definition

Therefore, multiplication on $S/\sim$ is well-defined. It remains to verify addition and multiplication satisfy the field axioms. I'll begin by noting the operations are commutative since D is a commutative ring:

$$[a,b]+[c,d]=[ad+bc,bd]=[cb+da,db]=[c,d]+[a,b]$$

likewise,

$$[a,b][c,d]=[ac,bd]=[ca,db]=[c,d][a,b].$$

Let $x\in D$ be nonzero, and $[a,b]\in S/\sim$. Note:

$$[a,b]+[0,x]=[ax+b(0),bx]=[ax,bx]=[a,b]$$

as $(ax,bx)\sim(a,b)$ is easy to verify (remember $x\neq 0$). We find $[0,x]$ serves as the additive identity of $S/\sim$. Next, consider $[1,1]$ and $[a,b]\in S/\sim$,

$$[a,b][1,1]=[a(1),b(1)]=[a,b]$$

hence $[1,1]$ is the unity of $S/\sim$. Multiplicative inverse is easy $[a,b]\neq 0$ has $a,b\neq 0$ hence $[b,a]$ is in $S/\sim$ and

$$[a,b][b,a]=[ab,ba]=[1,1]$$

as $(ab,ba)\sim(1,1)$ is easy to verify. Associativity,

$$[a,b]+\big([c,d]+[e,f]\big)=[a,b]+[cf+de,df]=[a(df)+(cf+de)b,bdf]$$

and

$$\big([a,b]+[c,d]\big)+[e,f]=[ad+bc,bd]+[e,f]=[(ad+bc)f+e(bd),bdf]$$

Thus addition is associative. I leave it to the reader to prove associativity of multiplication as well as the needed distributive properties linking addition and mulitplication. In summary, we have shown $S/\sim$ is a field. It remains to explain how it contains a subring which is isomorphic to D. You should not be surprised when I tell you that $\phi:D\to S/\sim$ defines an injective ring homomorphism if we set $\phi(x)=[x,1]$. Notice, $\phi(x)=[x,1]=0$ implies $x=0$ hence $\mathrm{Ker}(\phi)=\{0\}$. Moreover,

$$\phi(x+y)=[x+y,1]=[x(1)+1(y),1(1)]=[x,1]+[y,1]=\phi(x)+\phi(y)$$

and

$$\phi(xy)=[xy,1]=[xy,1(1)]=[x,1][y,1]=\phi(x)\phi(y)$$

for all $x,y\in D$. Thus $D/\{0\}\approx\!\approx\phi(D)$ by the first isomorphism theorem of rings and hence $D\approx\phi(D)$. □

Definition 3.3.16. *The field F constructed from an integral domain D as in the proof above is called the* **field of quotients of** *D. We use the notation a/b or $\frac{a}{b}$ for the equivalence class $[a,b]$. We have shown,*

$$\mathbb{F}=\left\{\frac{a}{b}\mid a,b\in D,b\neq 0\right\}$$

is a field where we **define**

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}\qquad\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}.$$

You can trace back through the proof of the field of quotients construction to see we have proved all the usual properties of rational numbers:

$$\frac{0}{a} = 0, \quad \frac{a}{b} \cdot \frac{b}{a} = 1, \quad \frac{ax}{bx} = \frac{a}{b}.$$

So, on the one hand, this proof we went over just now proves that $\mathbb{Q}$ exists if we are given $\mathbb{Z}$. On the other hand, it allows us to construct abstract fields which play the same role for a given integral domain as does $\mathbb{Q}$ for $\mathbb{Z}$. Personally, I view this construction and the clarity it can bring to **what** rational numbers **are** as a high point of abstract algebra. Is $1/2$ and $3/6$ the same number? I say emphatically **yes**. We have shown $1/2 = 3/6$ because the rigorous definition of $\mathbb{Q}$ says $a/b = c/d$ only if $ad = bc$ and surely we can agree $1(6) = 2(3)$. Now, does a given rational number have many different **fractions** which represent the same number? Yes. We also can agree about that. The pair $(1, 2) \neq (3, 6)$. In any event, we should keep in mind, equivalence classes are always with us whether we understand them or not. You might read this post by Paul Garrett.

Example 3.3.17. *If $D = \mathbb{Z}[x]$ then the field of quotients for D is the set $\{f(x)/g(x) \mid f(x), g(x) \in \mathbb{Z}[x], g(x) \neq 0\}$*

Example 3.3.18. *If $D = \mathbb{F}[x]$ then the field of quotients for D is the set $\{f(x)/g(x) \mid f(x), g(x) \in \mathbb{F}[x], g(x) \neq 0\} = \mathbb{F}(x)$ the rational functions over $\mathbb{F}$. For $\mathbb{F} = \mathbb{R}$ this is just the usual rational functions.*

Example 3.3.19. *The notation $\mathbb{Z}_p[x]$ is polynomials with $\mathbb{Z}_p$-coefficients. In contrast, $\mathbb{Z}_p(x) = \{f(x)/g(x) \mid f(x), g(x) \in \mathbb{Z}_p[x], g(x) \neq 0\}$. This gives an example of an infinite field with characteristic p.*

Outside this conversation, I might be tempted to agree that fields with finite characteristic are finite fields. This is clearly false by our last example !

I recommend working on Exercises for Gallian Chapter 15 such as: (page 260-263)

$$\#1 - 57 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 25: (these are collected at the start of Lecture 27)

Problem 97: Gallian number 11 from page 261.

Problem 98: Gallian number 28 from page 261.

Problem 99: Gallian number 29 from page 261.

Problem 100: Gallian number 34 from page 262.

Problem 101: Gallian number 43 from page 262.

Problem 102: Gallian number 47 from page 262.

3.4 Lecture 26: polynomials in an indeterminant

It seems to me something is missing here in Gallian and I need to add a bit of material from Rotman (and many other texts) to build up the foundations of polynomials over a ring.

We use the phrase **indeterminant form** in early calculus to capture the idea of a limit whose form does not indicate its eventual convergence or divergence. The term **indeterminant** here is given mainly to divorce the concept of a **polynomial function** from a **polynomial expression**. This much I should say, when x is an indeterminant this means x is not a variable. We do not have in mind some bucket of things which we can pour into x as our imagination warrants. We wish instead to think of x as a sort of place-holder. Of course, x and x^2 are different. Moreover, $1, x, x^2, x^3, \ldots$ are distinct. I could go on about the idea here, but, the best way to be clear is to give the actual definition. Before we define polynomials we first define **formal power series**[5].

Definition 3.4.1. *Suppose R is a commutative ring, then a* **formal power series over** *R is a function $\sigma : \mathbb{N} \cup \{0\} \to R$. Write $\sigma(j) = s_j$ for $j \in \mathbb{N} \cup \{0\}$ and we use the sequential notation:*

$$\sigma = (s_0, s_1, \ldots, s_j, \ldots)$$

where we call $s_j \in R$ the **coefficients**[6] *of the formal power series.*

So, what is a polynomial?

Definition 3.4.2. *A formal power series $\sigma = (s_o, s_1, \ldots, s_j, \ldots)$ over a commutative ring R is called a* **polynomial over** *R if there is some integer $m \geq 0$ with $s_j = 0$ for all $j > m$; that is $\sigma = (s_0, s_1, \ldots, s_m, 0, 0, \ldots)$. Furthermore, the* **zero polynomial** *is $\sigma = (0, 0, \ldots)$. If $\sigma = (s_0, s_1, \ldots)$ is a nonzero polynomial and $n \in \mathbb{N}$ is the smallest integer for which $s_j = 0$ for all $j > n$ then we say $deg(\sigma) = n$ and s_n is the* **leading coefficient**.

We are using sequences to build polynomial expressions. Our next step is to define addition and multiplication of such sequences:

Definition 3.4.3. *Denote the set of polynomials with coefficients in R by $R[x]$. If $\sigma, \tau \in R[x]$ then*

$$\sigma + \tau = (s_0 + t_0, s_1 + t_1, \ldots, s_j + t_j, \ldots)$$

where $\sigma = (s_j)$ and $\tau = (t_j)$. Moreover,

$$\sigma\tau = (s_0t_0, s_ot_1 + s_1t_0, s_0t_2 + s_1t_1 + s_2t_0, \ldots),$$

where to be precise $\sigma\tau = (a_0, a_1, \ldots, a_k, \ldots)$ and $a_k = \sum_{i+j=k} s_it_j = \sum_{i=0}^{k} s_it_{k-i}$.

To be careful, we should explain why this definition is reasonable. Let me outline the argument:

(1.) $\deg(\sigma + \tau) \leq \max(\deg(\sigma), \deg(\tau))$. It follows that the number of nonzero entries in $\sigma + \tau$ is finite. Hence $\sigma + \tau$ is a polynomial.

(2.) either $\sigma\tau = 0$ or $\deg(\sigma\tau) \leq \deg(\sigma)\deg(\tau)$. Therefore the product of two polynomials is once more a polynomial.

[5] These are known as **formal** power series because there is no expectation of convergence. For example, $\sum_{j=0}^{\infty} s_jx^j = s_0 + s_1x + s_2x^2 + \cdots$ is a formal power series. But, I'm getting a bit ahead of the story here.

[6] Rotman, page 236 of *First Course in Abstract Algebra* shares that the term **coefficient** means **acting together to some single end**, here the coefficients together form the formal power series.

Next, we should show $R[x]$ forms a commutative ring with respect to the addition and multiplication just defined. Consider,

$$(s_0, s_1, \dots, s_n, 0, \dots) + (0, 0, \dots) = (s_0 + 0, s_1 + 0, \dots, s_n + 0, 0 + 0, \dots) = (s_0, s_1, \dots, s_n, 0, \dots)$$

hence $0 = (0, 0, \dots)$. Moreover, setting $n = max(\deg(\sigma), \deg(\sigma))$,

$$\sigma + \tau = (s_0 + t_0, s_1 + t_1, \dots, s_n + t_n, 0, \dots) = (t_0 + s_0, t_1 + s_1, \dots, t_n + s_n, 0, \dots) = \tau + \sigma$$

hence addition is commutative. Clearly, $\sigma = (s_j)$ has additive inverse $-\sigma = (-s_j)$. Addition of sequences is addition of functions from $\mathbb{N} \cup \{0\}$ and we know that is associative. It remains to prove multiplication is associative and distributive. I leave those to the reader. Let me explain how x comes into the picture. We need to assume R is unital for our convenience at this point.

Definition 3.4.4. *Let R be a commutative ring with unity 1 then in the polynomials $R[x]$ we define $x = (0, 1, 0, \dots)$.*

We finally learn why the notation $R[x]$ is warranted. Also, it should be fairly clear we cannot make x a variable in this context. Is $(0, 1, 0, \dots)$ a variable ?

Theorem 3.4.5. *Let R be a commutative unital ring and $\sigma \in R[x]$ with $\sigma = (s_j)$ then $\sigma = \sum_{j=0}^{\infty} \vec{s}_j x^j$ where we define $\vec{r} = (r, 0, \dots)$ for each $r \in R$.*

Proof: first, we note a property of the multiplication, if $\vec{c} = (c, 0, 0, \dots)$ and $\tau = (t_0, t_1, \dots, t_n, 0 \dots)$ then $\vec{c}\tau = (ct_0, ct_1, \dots, ct_n, 0, \dots)$. Second, notice $x^2 = xx$ is calculated by:

$$x^2 = (0, 1, 0, \dots)(0, 1, 0, \dots) = (0, 0, 1, 0, \dots)$$

since $\alpha = (0, 1, 0, \dots) = x$ and $\beta = (0, 1, 0, \dots) = x$ has $\alpha = (a_i)$ and $\beta = (b_j)$ with $a_i = b_i = 0$ for $i \neq 1$ hence:

$$\alpha\beta = (a_0b_0, a_0b_1 + a_1b_0, a_0b_2 + a_1b_1 + a_2b_0, \dots) = (0, 0, 1, 0, \dots).$$

Furthermore, if we suppose inductively for some $n \in \mathbb{N}$, $x^n = e_{n+1}$ where $(e_i)_j = \delta_{ij}$ defines the sequence which is everywhere zero except in the i-th entry where we find 1. Then, $xx^n = e_{n+2}$ by the definition of the multiplication, only the $(n + 2) - th$ entry is nontrivial since x has $x_1 = 1$ whereas $(x^n)_{n+1} = 1$ and all other entries are zero. Hence inductively $x^n = e_{n+1}$ for all $n \in \mathbb{N}$. We also define $x^0 = \vec{1}$ and $x^1 = x$ where we may note $x^0x = \vec{1}x = x$ as we should expect. Now that we have the structure of x and powers of x sorted out we can produce the main result. Observe, we can write a polynomial as a sum of mostly zero sequences: σ with $\deg(\sigma) = n$,

$$\begin{aligned}
\sigma &= (s_0, s_1, \dots, s_n, \dots, 0) \\
&= (s_0, 0, \dots) + (0, s_1, 0, \dots) + \cdots + (0, \dots, 0, s_n, 0, \dots) \\
&= \vec{s}_0(1, 0, \dots) + \vec{s}_1(0, 1, 0, \dots) + \cdots + \vec{s}_n e_{n+1} \\
&= \vec{s}_0 x^0 + \vec{s}_1 x + \cdots + \vec{s}_n x^n \\
&= \sum_{j=0}^{\infty} \vec{s}_j x^j
\end{aligned}$$

where we threw in a few zeros in the last step. □

At this point, we tire of the notation $\vec{s}_j$. It is customary to simply write s_j in place of $\vec{s}_j$. With this notation, a typical polynomial in $R[x]$ can be expressed as:

$$\sigma = s_0 + s_1 x + s_2 x^2 + \cdots + s_n x^n$$

where $s_0, s_1, \ldots, s_n \in R$ and $\deg(\sigma) = n$. I hope you appreciate how removed this is from our standard viewpoint in previous math courses. Notice this is merely notation to overlay sequences with finitely many nonzero entries. In any event, what we should take with us going forward is that $R[x]$ behaves precisely as we have assumed thus far in this course. The construction I've outlined merely shows you how we can construct indeterminants and expressions without use of functions on R. At this point we return to Gallian and follow his presentation going forward from Theorem 16.1 on page 286. Gallian has a concrete example worth including from page 284:

Example 3.4.6. *The polynomials $f(x) = x^3+2x$ and $g(x) = x^5+2x$ are distinct in $\mathbb{Z}_3[x]$. However, if we consider f, g as functions on $\mathbb{Z}_3$ notice*

$$\begin{aligned} f(1) = 1^3 + 2(1) = 1 + 2 = 0, \quad & \& \quad g(1) = 1^5 + 2(1) = 1 + 2 = 0 \\ f(2) = 2^3 + 2(2) = 8 + 4 = 0, \quad & \& \quad g(2) = 2^5 + 2(2) = 32 + 4 = 0 \\ f(3) = 3^3 + 2(3) = 0, \quad & \& \quad g(1) = 3^5 + 2(3) = 0 \end{aligned}$$

Thus, as polynomial functions on $\mathbb{Z}_3$, $f = g$.

I should also mention, Example 3.3.4 is a bit more interesting with our new view of $R[x]$. In fact, when I write $\phi_a(f(x)) = f(a)$ we mean to define the value $f(a)$ as if f was a function of R. Very sneaky.

Definition 3.4.7. *Let R be a commutative unital ring. Define the* **evaluation map** *for $a \in R$ by:*

$$\phi_a(s_0 + s_1 x + \cdots + s_n x^n) = s_0 + s_1 a + \cdots + s_n a^n.$$

for each $s_0 + s_1 x + \cdots s_n x^n \in R[x]$.

Pragmatically, it doesn't matter for many applications if we think of $R[x]$ as polynomial functions, but, algebraically, we take the viewpoint $R[x]$ is the set of polynomials in **indeterminant** x. If we wish to obtain the corresponding function then we simply make use of the evaluation map (in fact, $\phi_a : R[x] \to R$ is a ring homomorphism).

Theorem 3.4.8. *If D is an integral domain then $D[x]$ is an integral domain.*

Proof: suppose $f(x), g(x) \in D[x]$ are nonzero polynomials $f(x) = a_n x^n + \cdots + a_0$ and $g(x) = b_m x^m + \cdots + b_0$ where a_n, b_m are the leading coefficients of $f(x), g(x)$ respective. Observe,

$$f(x)g(x) = a_n b_m x^{m+n} + \cdots + a_0 b_0.$$

Note $a_n, b_m \neq 0$ in integral domain D hence $a_n b_m \neq 0$ and we find $f(x)g(x) \neq 0$. Therefore, there are no zero divisors in $D[x]$. Furthermore, $D[x]$ is a commutative ring with unity $f(x) = 1$ hence $D[x]$ is an integral domain. □

The proof of the following is really not much removed from standard highschool algebra.

Theorem 3.4.9. *Let F be a field and $f(x), g(x) \in F[x]$ with $g(x) \neq 0$. Then there exist unique $q(x), r(x) \in F[x]$ such that $f(x) = g(x)q(x) + r(x)$ and either $r(x) = 0$ or $deg(r(x)) < deg(g(x))$. We call $q(x)$ the* **quotient** *and $r(x)$ the* **remainder** *in the division of $f(x)$ by $g(x)$.*

Proof: see page 286-287 of Gallian. If you don't understand it when you read it, try getting out a piece of paper and writing it out. It's not too hard to follow. □

Corollary 3.4.10. *Let F be a field and $a \in F$ and $f(x) \in F[x]$. Then $f(a)$ is the remainder in the division of $f(x)$ by $x - a$.*

Proof: by the division algorithm, there exists $g(x), r(x)$ for which $f(x) = (x-a)g(x)+r(x)$ where either $r(x) = 0$ or $\deg(r(x)) < \deg(x-a) = 1$. It follows $r(x) = r \in R$. Moreover, by the evaluation homomorphism at a we find,

$$\phi_a(f(x)) = f(a) = (a-a)g(a) + r = r \quad \Rightarrow \quad r = f(a). \qquad \square$$

Definition 3.4.11. *Let F be a field. Let $f(x) \in F[x]$, we say $c \in F$ is a* **zero** *of $f(x)$ if $\phi_c(f(x)) = f(c) = 0$. If $(x-c)^k$ is a factor of $f(x)$ and $(x-c)^{k+1}$ is not a factor of $f(x)$ then we say c* **is a zero with multiplicity** *k.*

There are pretty connections between the algebra of calculus and the existence of repeated zeros. But, we save that for another time.

Corollary 3.4.12. *Let F be a field and $a \in F$ and $f(x) \in F[x]$. Then a is a zero of $f(x)$ if and only if $x - a$ is a factor of $f(x)$.*

Proof: left to reader. □

Example 3.4.13. *An interesting counterpoint to the Corollary below is found in the polynomials with coefficients in $\mathbb{Z}_6$. The polynomial $f(x) = x^2 + 3x + 2$ has four zeros. Gallian mentions Lagrange proved the Corollary below for $\mathbb{Z}_p$ where p is prime. Another interesting point, $\mathbb{Z}_6[x]$ is also not an integral domain; $(2x+2)(3x^2+3) = 0$ yet $2x+2, 3x^2+3 \neq 0$. The study of zero divisors in $D[x]$ for D which is not integral is a nice topic to investigate. Perhaps we'll look at that further in a future lecture.*

Corollary 3.4.14. *A polynomial of degree n over a field F has at most n zeros counting multiplicity.*

Proof: the proof is by induction on degree. If $f(x) \in F[x]$ has $\deg(f(x)) = 0$ then $f(x) = c \neq 0$ hence there are zero zeroes for $f(x)$. Suppose inductively that each polynomial up to degree $n-1$ has at most $n-1$ zeros. Consider $f(x)$ with degree n. Suppose a is a zero with multiplicity k then $f(x) = (x-a)^k q(x)$ for some $q(x)$ with degree $n-k$. If $f(x)$ has no additional zeros then the Corollary holds since $f(x)$ has less than n zeros. Otherwise, $f(b) = 0$ for some $a \neq b$ hence $f(b) = (b-a)^k q(b) = 0$ and as $(b-a)^k \neq 0$ and F is an integral domain since it's a field it follows $q(b) = 0$. But, the $\deg(q(x)) = n-k < n$ hence by the inductive hypothesis $q(x)$ has at most $n-k$ zeros counting multiplicity thus $f(x) = (x-a)^k q(x)$ has at most $k + n - k = n$ zeros counting multiplicity. □

The argument above is great for you who are fans of formal induction, but, I am also fond of the simple argument, n is the degree of $f(x)$. Notice each zero a_1 generates a factor $(x - a_1)$ in the factorization of $f(x)$. Suppose there were $n+1$ zeros (possibly duplicate). Then

$$f(x) = (x-a_1)(x-a_2)\cdots(x-a_{n+1})g(x)$$

for some polynomial $g(x)$ and degree of $f(x)$ is at least $n+1$. This contradicts $\deg(f(x)) = n$ hence there cannot be more than n-zeros. I'm not usually a fan of contradiction, but, this argument resonates for me.

Example 3.4.15. *Consider $f(x) = x^n - 1 \in \mathbb{C}[x]$. Notice $\omega = exp(2\pi i/n)$ has $\omega^n = 1$ but $\omega^k \neq 1$ for $k = 1, 2, \ldots, n-1$. It follows that $1, \omega, \omega^2, \ldots, \omega^{n-1}$ are all solutions of $\omega^n = 1$. Furthermore,*

$$f(x) = x^n - 1 = (x-1)(x-\omega)(x-\omega^2)\cdots(x-\omega^{n-1})$$

The number $\omega = exp(2\pi i/n)$ is called the **primitive n-th root of unity** *in $\mathbb{C}$. To be pedantic, we really should say ω_n is the primitive n-th root. Then $\omega_2 = -1$, $\omega_3 = \cos(2\pi/3) + i\sin(2\pi/3)$ and $\omega_4 = i$ etc. For example,*

$$f(x) = x^4 - 1 = (x^2+1)(x^2-1) = (x+i)(x-i)(x-1)(x+1)$$

where $\omega_4 = i$ and $\omega_4^2 = -1$ and $\omega_4^3 = -i$ and $\omega_4^4 = 1$.

We have studied principal ideals a bit in previous lectures, we now give a name to a ring where every ideal is principal.

Definition 3.4.16. *A* **principal ideal domain** *or* **PID** *is an integral domain R in which every ideal has the form $\langle a \rangle = \{ra \mid r \in R\}$ for some $a \in R$.*

Many of our examples are PIDs, some are not. This much we can say:

Theorem 3.4.17. *If F is a field then $F[x]$ is a principal ideal domain.*

Proof: we know $F[x]$ is an integral domain. Suppose I is an ideal in $F[x]$. If $I = 0$ then $I = \langle 0 \rangle$ is principal. If $I \neq 0$ then the degree of polynomials in I is bounded below hence there must be an polynomial of least degree by the well-ordering-principal. Let $g(x)$ be a polynomial of least degree in I. If $f(x) \in I$ then note the division algorithm provides $q(x)$ with $f(x) = g(x)q(x) + r(x)$ with $r(x) = 0$ or $\deg(r(x)) < \deg(g(x))$. But, $g(x)$ is of minimal degree in I and $r(x) = f(x) - g(x)q(x) \in I$ hence $r(x) = 0$. Thus $f(x) = g(x)q(x)$ and $f(x) \in \langle g(x) \rangle$ and hence $I \subseteq \langle g(x) \rangle$. Conversely, it is easy to see $\langle g(x) \rangle \subseteq I$ thus $I = \langle g(x) \rangle$ and as I was arbitrary we've shown $F[x]$ is a PID. □

From the proof above we also obtain the following:

Theorem 3.4.18. *If F is a field and I a nonzero ideal in $F[x]$ and $g(x) \in F[x]$. Then, $I = \langle g(x) \rangle$ if and only if $g(x)$ is a nonzero polynomial of minimum degree in I.*

Example 3.4.19. *Consider $\phi : \mathbb{R}[x] \to \mathbb{C}$ given by $\phi(f(x)) = f(i)$. Observe $Ker(\phi)$ is an ideal in $\mathbb{R}[x]$ hence $Ker(\phi)$ is a principal ideal. Notice, no linear polynomial $f(x) = mx + b$ has $f(i) = 0$ since $mi + b = 0$ implies $b = -mi$ which is impossible as $m, b \in \mathbb{R}$. Consequently, $x^2 + 1 \in Ker(\phi)$ is an element of smallest degree in $Ker(\phi)$ which implies $Ker(\phi) = \langle x^2 + 1 \rangle$. If $a + ib \in \mathbb{C}$ then $\phi(a + bx) = a + bi$ hence $\phi(\mathbb{R}[x]) = \mathbb{C}$. Thus, by the first isomorphism theorem for rings, $\mathbb{R}[x]/\langle x^2 + 1 \rangle \approx \mathbb{C}$.*

Remark 3.4.20. Sorry about the timing this Semester. Ideally, I'd like to avoid having both the construction of the field of fractions and the formal polynomials in the same day. Too much construction. That said, I think the next Lecture will be less contructive (but, in a good way).

I recommend working on Exercises for Gallian Chapter 16 such as: (page 290-293)

$$\#1 - 43 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 26: (these are collected at the start of Lecture 28)

Problem 103: Prove that the addition and multiplication for the field of fractions construction has: (i.) an associative multiplication, (ii) the left distributive property. (this makes the proof of Theorem 3.3.15 a bit closer to being complete)

Problem 104: Prove Corollary 3.4.12.

Problem 105: Gallian number 20 from page 279.

Problem 106: Gallian number 40 from page 280.

Problem 107: Gallian number 63 from page 281.

Problem 108: Gallian number 12 from page 291.

3.5 Lecture 27: factorization of polynomials

What are the rules for factoring? How do we factor? We begin to answer these questions in certain special cases. We discover some suprising results about the interplay between $\mathbb{Z}$, $\mathbb{Z}_n$ and $\mathbb{Q}$.

Definition 3.5.1. *Let D be an integral domain. We say $f(x) \in D[x]$ which is neither zero nor a unit in $D[x]$ is* **irreducible over** *D if whenever $f(x) = g(x)h(x)$ with $g(x), h(x) \in D[x]$ then $g(x)$ or $h(x)$ is a unit in $D[x]$. A nonzero, nonunit, element of $D[x]$ that is not irreducible over D is known as a* **reducible polynomial over** *D.*

In other words, if a polynomials is not not reducible then it's reducible.

Example 3.5.2. *Consider $f(x) = x^2 + 1$. Note $f(x)$ is irreducible over $\mathbb{R}$ or $\mathbb{Q}$. However, $f(x)$ is reducible over $\mathbb{C}$ as $f(x) = (x + i)(x - i)$.*

Example 3.5.3. *If $f(x) = 2x + 4$ then $f(x) = 2(x + 2)$ thus $f(x)$ is reducible over $\mathbb{Z}$ as 2 is not a unit in $\mathbb{Z}$. On the other hand, $f(x)$ is irreducible over $\mathbb{Q}$ or $\mathbb{R}$ as $2x + 4 = g(x)h(x)$ implies one of these is a nonzero constant. In $\mathbb{Q}$ or $\mathbb{R}$ every nonzero element is a unit.*

Our main point in these examples is that context matters. Irreducibility depends both on the polynomial in question and the ring from which coefficients are taken.

Example 3.5.4. *Let $f(x) = x^2 - 7$ then $f(x) = (x - \sqrt{7})(x + \sqrt{7})$ hence $f(x)$ is reducible over $\mathbb{R}$ (it is obvious that the factors are not units in $\mathbb{R}[x]$, the units in $\mathbb{R}[x]$ are all in $\mathbb{R}^\times$). In contrast, $f(x)$ is irreducible over $\mathbb{Q}$ or $\mathbb{Z}$.*

Example 3.5.5. *Consider $f(x) = x^2 + 1$. We use Corollary 3.4.12 in what follows. Considering $f(x) \in \mathbb{Z}_3$ we calculate:*

$$f(0) = 1, \quad f(1) = 1 + 1 = 2, \quad f(2) = 4 + 1 = 5 = 2$$

thus $f(x)$ has no factor of the form $x - a$ in $\mathbb{Z}_3[x]$. That is, $x^2 + 1$ is irreducible in $\mathbb{Z}_3[x]$. In contrast, for $f(x) \in \mathbb{Z}_5[x]$ we have $f(2) = 4 + 1 = 5 = 0$ hence $(x - 2) \mid f(x)$. We seek a for which:

$$x^2 + 1 = (x - 2)(x + a) = x^2 + (a - 2)x - 2a$$

apparently, $a - 2 = 0$ whereas $-2a = 1$ which are simultaneously solved by $a = 2$ as $-4 = 1$ modulo 5. Indeed, this squares well with the following calculation: in $\mathbb{Z}_5[x]$ we find:

$$x^2 + 1 = x^2 - 4 = (x - 2)(x + 2)$$

As you can see, $f(x)$ is reducible over $\mathbb{Z}_5$.

The following theorem is very useful.

Theorem 3.5.6. *Suppose F is a field and $f(x) \in F[x]$ has degree 2 or 3 then $f(x)$ is reducible over F if and only if $f(x)$ has a zero in F.*

Proof: let F be a field and $f(x) \in F[x]$ with degree 2 or 3. If $f(x)$ is reducible then $f(x)$ has a factorization including a linear factor hence $f(x)$ has a zero[7] by Corollary 3.4.12. Conversely, if $f(x)$ has a zero c then $f(x) = (x - c)g(x)$ where either $g(x)$ is degree 1 or degree 2. Thus, $g(x)$ is not a unit and find $f(x)$ is reducible. □

I use the observation that units of $F[x]$ are simply the nonzero constant polynomials in $F[x]$ which we naturally identify with $F^\times$.

[7] hmmm, it seems half of the solution to Problem 104 is contained in the proof of Theorem 17.1

Example 3.5.7. *Consider $f(x) = x^4+14x^2+49 = (x^2+7)^2$ thus $f(x)$ is reducible over $\mathbb{R}$ yet $f(x)$ has no zeros in $\mathbb{R}$. At fourth order we lose the necessary connection between zeros and reducibility.*

The next few theorems we consider are probably new to most students in this course.

Definition 3.5.8. *The* **content** *of a nonzero polynomial $a_nx^n + \cdots + a_1x + a_0 \in \mathbb{Z}[x]$ is the $gcd(a_0, a_1, \ldots, a_n)$. If the content of $f(x) \in \mathbb{Z}[x]$ is 1 then we say $f(x)$ is a* **primitive polynomial**.

Gallian calls this Gauss's Lemma. That doesn't seem overly descriptive given Gauss's work.

Example 3.5.9. *Let $f(x) = 3x+6$ then the content of $f(x)$ is $gcd(3, 6) = 3$. Notice, $f(x) = 3(x+2)$ and $x+2$ is primitive as $gcd(1, 2) = 1$. Any monic polynomial is primitive, $g(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ has $gcd(1, a_{n-1}, \ldots, a_1, a_0) = 1$. The idea of the content is to find that integer which naturally factors out of a polynomial in $\mathbb{Z}[x]$. Of course, $3x^2 + 5x + 7$ is also primitive since its coefficients are relatively prime. We can't factor out an integer $n > 1$ from a primitive polynomial.*

Theorem 3.5.10. *The product of two primitive polynomials is primitive.*

Proof: we follow Gallian's argument on page 297. Suppose $f(x), g(x) \in \mathbb{Z}[x]$ are primitive and $f(x)g(x)$ is not primitive. If p is a prime divisor of the content of $f(x)g(x) = ph(x)$ then consider the polynomials $\overline{f(x)}, \overline{g(x)} \in \mathbb{Z}_p[x]$ formed by reducing the coefficients of $f(x), g(x)$ respective. Observe,

$$0 = \overline{ph(x)} = \overline{f(x)} \cdot \overline{g(x)}$$

Thus, as $\mathbb{Z}_p[x]$ is an integral domain, $\overline{f(x)} = 0$ or $\overline{g(x)} = 0$. It follows p divides $f(x)$ or $g(x)$ thus $f(x)$ or $g(x)$ is not primitive. Hence, by proof by contradiction, $f(x)g(x)$ is primitive. □

A concept is used in the proof above which merits some discussion. If $\phi : R \to S$ is a ring homomorphism then there is a natural homomorphism $\psi : R[x] \to S[x]$ induced by mapping the coefficients of R to corresponding coefficients of S. In particular,

$$\Psi(a_nx^n + \cdots + a_1x + a_0) = \phi(a_n)x^n + \cdots + \phi(a_1)x + \phi(a_0)$$

for each $a_nx^n + \cdots + a_1x + a_0 \in R[x]$. In the proof for the primitive product theorem we used the natural homomorphism $\phi(k) = [k]_p$ where $\phi : \mathbb{Z} \to \mathbb{Z}_p$ to induce $\psi : \mathbb{Z}[x] \to \mathbb{Z}_p[x]$. Our notation was $\overline{f(x)}$ for $\psi(f(x))$. We continue to use such induced homomorphisms of polynomials in many of the proofs and examples we soon consider, often without explicit mention.

Theorem 3.5.11. *Let $f(x) \in \mathbb{Z}[x]$. If $f(x)$ is reducible over $\mathbb{Q}$ then it is reducible over $\mathbb{Z}$.*

Proof: Let $f(x) \in \mathbb{Z}[x]$ be monic[8]. Also, suppose there exist $h(x), g(x) \in \mathbb{Q}[x]$ with $f(x) = h(x)g(x)$. Suppose a is the least common multiple of the denominators of the coefficients in $h(x)$ and let b be the least common multiple of the denominators in $g(x)$. It follows $ah(x), bg(x) \in \mathbb{Z}[x]$ and $abf(x) = ah(x) \cdot bg(x)$. If c_h is the content of $ah(x)$ and c_g is the content of $bg(x)$ then there are primitive polynomials $g_1(x), h_1(x)$ for which $bg(x) = c_gg_1(x)$ and $ah(x) = c_hh_1(x)$. Observe,

$$abf(x) = ah(x) \cdot bg(x) = c_gg_1(x) \cdot c_hh_1(x) = c_hc_gh_1(x)g_1(x)$$

note $h_1(x)g_1(x)$ is primitive as it is the product of primitive polynomials. Thus the content of $abf(x)$ precisely c_hc_g. But, $f(x)$ is monic thus ab is the content of $abf(x)$. Hence, $ab = c_hc_g$ and it follows

[8] a polynomial is monic if it has a leading coefficient of 1

$f(x) = h_1(x)g_1(x)$ where $h_1(x), g_1(x) \in \mathbb{Z}[x]$. In summary, for monic $f(x) \in \mathbb{Z}[x]$ if $f(x) = h(x)g(x)$ for some $h(x), g(x) \in \mathbb{Q}[x]$ then there exist $h_1(x), g_1(x) \in \mathbb{Z}[x]$ for which $f(x) = h_1(x)g_1(x)$ with $\deg(h(x)) = \deg(h_1(x))$ and $\deg(g(x)) = \deg(g_1(x))$. If $f(x) \in \mathbb{Z}[x]$ is not monic then we can factor out the content c of $f(x)$ to write $f(x) = cf_1(x)$ where $f_1(x)$ is primitive. If $f(x)$ is reducible over $\mathbb{Q}$ then it follows $f_1(x)$ is reducible hence by our argument for primitive polynomials $f_1(x)$ is reducible over $\mathbb{Z}$ and consequently $f(x) = cf_1(x)$ is reducible over $\mathbb{Z}$ as well. □

Example 3.5.12. *Consider,* $f(x) = 6x^2 + 19x - 7$ *notice*

$$f(x) = 6x^2 + 19x - 7 = 6(x^2 + (19/6)x - 7/6) = 6(x + 7/2)(x - 1/3)$$

hence $f(x) = (2x + 7)(3x - 1)$. *If we can reduce* $f(x) \in \mathbb{Z}[x]$ *using* $\mathbb{Q}$ *then the reduction transfers nicely back to* $\mathbb{Z}[x]$. *Pragmatically, in this example, it's way easier to just see that* $f(x) = (2x + 7)(3x - 1)$ *from the outset.*

Gauss taught us that modular arithmetic gives great insight into ordinary integer arithmetic. Here is a prime example of such indirect reasoning. Notice p could be any prime.

Theorem 3.5.13. *Let* $p \in \mathbb{Z}$ *be prime and suppose* $f(x) \in \mathbb{Z}[x]$ *has* $deg(f(x)) \geq 1$. *Consider* $\overline{f(x)}$ *the corresponding polynomial in* $\mathbb{Z}_p[x]$ *formed from* $f(x)$ *by reducing the coefficients of* $f(x)$ *modulo* p. *If* $\overline{f(x)}$ *is irreducible over* $\mathbb{Z}_p$ *and* $deg(f(x)) = deg(\overline{f(x)})$ *then* $f(x)$ *is irreducible over* $\mathbb{Q}$.

Proof: suppose $f(x) \in \mathbb{Z}[x]$ with $\deg(f(x)) \geq 1$. Furthermore, suppose $\overline{f(x)}$ is irreducible over $\mathbb{Z}_p$ and $\deg(f(x)) = \deg(\overline{f(x)})$ but $f(x)$ is reducible over $\mathbb{Q}$. Hence, by Theorem 3.5.11 there exist $g(x), h(x) \in \mathbb{Z}[x]$ with $f(x) = g(x)h(x)$ and $\deg(g(x)), \deg(h(x)) < \deg(f(x))$. Using the homomorphism of $\mathbb{Z}[x]$ and $\mathbb{Z}_p[x]$ given by $f(x) \mapsto \overline{f(x)}$ we find

$$\overline{f(x)} = \overline{g(x)} \cdot \overline{h(x)}$$

Note, since the leading coefficient might be divisible by p the degree of the induced polynomials could be smaller; $\deg(\overline{g(x)}) \leq \deg(g(x))$ and $\deg(\overline{h(x)}) \leq \deg(h(x))$. However, $\deg(f(x)) = \deg(\overline{f(x)})$ hence

$$\deg(\overline{g(x)}) \leq \deg(g(x)) < \deg(f(x)) = \deg(\overline{f(x)})$$

and

$$\deg(\overline{h(x)}) \leq \deg(h(x)) < \deg(f(x)) = \deg(\overline{f(x)})$$

hence $\overline{f(x)} = \overline{g(x)} \cdot \overline{h(x)}$ shows $\overline{f(x)}$ is reducible thus contradicting the irreducibility of $\overline{f(x)}$. Thus $f(x)$ must be irreducible given the conditions of the Theorem. □

Example 3.5.14. *Consider* $f(x) = 29x^3 + 5x^2 + 2x + 1$. *Modulo 2,* $\overline{f(x)} = x^3 + x^2 + 1$ *hence*

$$\overline{f}(0) = 1 \qquad \& \qquad \overline{f}(1) = 1 + 1 + 1 = 1$$

hence $\overline{f(x)}$ *is irreducible in* $\mathbb{Z}_2[x]$ *from which we find* $f(x)$ *is irreducible over* $\mathbb{Q}$.

I used a combination of Theorems 3.5.6 and 3.5.13 to guide my logic in the above Example. I'll use Gallian's example from the paragraph on page 299.

Example 3.5.15. *Consider* $f(x) = 21x^3 - 3x^2 + 2x + 8$. *Over* $\mathbb{Z}_2$ *we can factor* $\overline{f(x)} = x^3 + x^2 = x^2(x+1)$. *However, if we study the polynomial induced from* $f(x)$ *in* $\mathbb{Z}_5(x)$ *we can calculate modulo 5,* $\overline{f(x)} = x^3 + 2x^2 + 2x + 3$ *hence*

$$\overline{f}(0) = 3, \quad \overline{f}(1) = 1 + 2 + 2 + 3 = 3, \quad \overline{f}(2) = 8 + 2(4) + 2(2) + 3 = 23 = 3,$$

$$\overline{f}(3) = \overline{f}(-2) = -8 + 8 - 4 + 3 = -1, \quad \overline{f}(4) = \overline{f}(-1) = -1 + 2 - 2 + 3 = 2.$$

But, sometimes, no choice of p reveals the irreducibility. Theorem 3.5.13 only affirms irreducibility over $\mathbb{Q}$, it does not deny it.

Example 3.5.16. *Let $f(x) = x^4 + 1$. We can show that $\overline{f(x)}$ is reducible in $\mathbb{Z}_p[x]$ for* **any** *prime p. Yet, $f(x) = x^4 + 1$ is irreducible over $\mathbb{Q}$. (proof of these claims is the content of Exercise 29, which it seems likely I assign)*

There is an obvious way to trade a polynomial in $\mathbb{Q}[x]$ for a corresponding polynomial in $\mathbb{Z}[x]$. After making this correspondence we are free to use the tools at our disposal for irreducibility over $\mathbb{Q}$ for polynomials in $\mathbb{Z}[x]$.

Example 3.5.17. *Let $f(x) = (3/7)x^4 - (2/7)x^2 + (9/35)x + 3/5$ the contruct the corresponding $h(x) = 35f(x) = 15x^4 - 10x^2 + 9x + 21$. It should be clear that irreducibility of $h(x)$ over $\mathbb{Q}$ is naturally tied to irreducibility of $f(x)$. Working modulo 2, $\overline{h(x)} = x^4 + x + 1$ and $\overline{h}(0) = 1$ and $\overline{h}(1) = 1 + 1 + 1 = 1$ thus $\overline{h(x)}$ has no linear factors. To search for possible quadratic factors we need only consider $x^2, x^2 + 1$, $x^2 + x$ and $x^2 + x + 1$ as there are no other quadratic factors possible in $\mathbb{Z}_2[x]$. Since x^2 and $x^2 + 1$ and $x^2 + x$ have zeros in $\mathbb{Z}_2$ it follows they cannot be factors of $\overline{h(x)}$. To see why $x^2 + x + 1$ is not a factor consider the following:*

$$(x^2 + x + 1)(x^2 + ax + b) = x^4 + x + 1$$

then $x^4 + (a + 1)x^3 + (b + a + 1)x^2 + (a + b)x + b = x^4 + x + 1$ from which we would require

$$a + 1 = 0, \quad b + a + 1 = 0, \quad a + b = 0, \quad b = 1$$

these equations are inconsistent as the first two provide $b = 0$ whereas the last gives $b = 1$. Thus $x^2 + x + 1$ does not factor $\overline{h(x)}$ and we deduce $\overline{h(x)}$ is irreducible in $\mathbb{Z}_2[x]$ thus $h(x)$ is irreducible over $\mathbb{Q}$ and hence $f(x) = \frac{1}{35}h(x)$ is irreducible over $\mathbb{Q}$.

To decide irreduciblility of quartics in a given $\mathbb{Z}_p[x]$ we can enumerate the possible quadratics and test if they factor the given quartic via long-division or the algebraic technique I used in the Example above. This is illustrated for $p = 3$ in Example 8 of Gallian on page 299-300 and is motivation for Problems 15 and 16 on page 308. Given the effort required for such an example, the criterion below is amazing:

Theorem 3.5.18. Eisenstein's Criterion: *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]$. If there is a prime p such that $p \nmid a_n$ but $p \mid a_j$ for $j = n - 1, \ldots, 0$ and $p^2 \nmid a_0$ then $f(x)$ is irreducible over $\mathbb{Q}$.*

Proof: I'll postpone proof until a bit later, I found the argument given in Example 4 on page 321 of Gallian far more interesting than the proof by contradiction given on page 300. □

Example 3.5.19. *Consider $f(x) = 13x^7 + 2x^6 + 4x^3 + 18x + 2$. Observe $p = 2$ is such that $2 \nmid 13$ and $2^2 = 4 \nmid 2$ but 2 does divide $2, 4, 18$ and 2 (and the zero coefficients, note $p \mid 0$ for any p since $0 = p(0)$) thus by Eisenstein's Criterion with $p = 2$ we find $f(x)$ is irreducible over $\mathbb{Q}$.*

Consider, if $S = 1 + x + \cdots + x^{p-1}$ then $xS = x + x^2 + \cdots + x^p$ then

$$S - xS = (x + x^2 + \cdots + x^p) - (1 + x + \cdots + x^{p-1}) = x^p - 1$$

thus, formally, solving for S yields $1 + x + \cdots + x^{p-1} = \frac{x^p - 1}{1 - x}$. Perhaps you remember this algebra from the derivation of the geometric series. In any event, the polynomial $\Phi_p(x) = 1 + x + \cdots + x^{p-1}$ is **defined** to be the p-th *cyclotomic polynomial*.

Theorem 3.5.20. *For $p \in \mathbb{Z}$ prime $\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ is irreducible over $\mathbb{Q}$.*

Proof: the proof in Gallian you'll find in many books, and, my notes: let $f(x) = \Phi_p(x+1)$ thus

$$f(x) = \frac{(x+1)^p - 1}{(x+1) - 1} = \frac{1}{x}\left(x^p + \binom{p}{1}x^{p-1} + \binom{p}{2}x^{p-2} + \cdots + \binom{p}{p-1}x + 1 - 1\right)$$

Cleaning things up a bit,

$$f(x) = x^{p-1} + px^{p-2} + \cdots + p$$

where we may observe every coefficient except the leading coefficient is divided by p and the constant term is not divisible by p^2 hence $f(x)$ is irreducible by Eisenstein's Criterion. Suppose $\Phi_p(x)$ is reducible over $\mathbb{Q}$. In particular, suppose there exist $g(x), h(x) \in \mathbb{Q}[x]$ of degree less than $p-1$ where $\Phi_p(x) = g(x)h(x)$. Then $\Phi_p(x+1) = f(x) = g(x+1)h(x+1)$ shows $f(x)$ is reducible since $g(x+1), h(x+1) \in \mathbb{Q}[x]$ is easily seen with a little algebra. But, this contradicts the irreducibility of $f(x)$ hence $\Phi_p(x)$ is irreducible over $\mathbb{Q}$. □

Irreducible polynomials are useful for building new fields. This is seen in the Corollary to the Theorem below:

Theorem 3.5.21. *Let F be a field and suppose $p(x) \in F[x]$. Then $\langle p(x)\rangle$ is a maximal ideal in $F[x]$ if and only if $p(x)$ is irreducible over F.*

Proof: suppose that F is a field and $p(x) \in F[x]$. If $\langle p(x)\rangle$ is a maximal ideal in $F[x]$ then $\langle p(x)\rangle$ is a nonzero proper ideal hence $p(x) \neq 0$ and $p(x)$ is nonconstant. Suppose $p(x) = g(x)h(x)$ is a factorization of $p(x)$ over F. If $j(x) \in \langle p(x)\rangle$ then $j(x) = p(x)k(x) = g(x)h(x)k(x)$ thus $j(x) \in \langle g(x)\rangle$ and we find $\langle p(x)\rangle \subseteq \langle g(x)\rangle \subseteq F[x]$. Thus, by maximality, $\langle p(x)\rangle = \langle g(x)\rangle$ or $\langle g(x)\rangle = F[x]$. If $\langle p(x)\rangle = \langle g(x)\rangle$ then we have $g(x) \in \langle p(x)$ hence $g(x) = q(x)p(x)$ and $p(x) = g(x)h(x)$ so $\deg(g(x)) \geq \deg(p(x))$ and $\deg(p(x)) \geq \deg(g(x))$ from which we find $\deg(g(x)) = \deg(p(x))$. On the other hand, if $\langle g(x)\rangle = F[x]$ then each $f(x) = g(x)k(x)$ for some $k(x) \in F[x]$ for each $f(x) \in F[x]$. It follows that $g(x) \in F^\times$ hence $\deg(g(x)) = 0$. In summary, if $p(x) = g(x)h(x)$ then neither of the factors may have nontrivial degree smaller than that of $p(x)$. That is, $p(x)$ is irreducible over F.

Conversely, suppose $p(x)$ is irreducible. Suppose I is an ideal of $F[x]$ for which $\langle p(x)\rangle \subseteq I \subseteq F[x]$. Recall from Theorem 3.4.17 we know $F[x]$ is a PID hence $I = \langle g(x)\rangle$ for some $g(x) \in F[x]$. Note $p(x) = p(x)1 \in \langle p(x)\rangle \subseteq \langle g(x)\rangle$ hence there exists $k(x) \in F[x]$ for which $p(x) = k(x)g(x)$. However, irreducibilty of $p(x)$ implies either $\deg(k(x)) = 0$ or $\deg(g(x)) = 0$. If $\deg(k(x)) = 0$ then $\langle p(x)\rangle = \langle g(x)\rangle$. If $\deg(g(x)) = 0$ then $\langle p(x)\rangle = F[x]$. Thus $\langle p(x)\rangle$ is a maximal ideal in $F[x]$. □

Corollary 3.5.22. *Let F be a field and $p(x) \in F[x]$ is irreducible over F. Then $F[x]/\langle p(x)\rangle$ is a field.*

Proof: if F is a field and $p(x)$ is an irreducible polynomial then $\langle p(x)\rangle$ is maximal by Theorem 3.5.21. Thus $F[x]/\langle p(x)\rangle$ if a field by Theorem 3.2.25. □

Corollary 3.5.23. *Let F be a field and $p(x), a(x), b(x) \in F[x]$. If $p(x)$ is irreducible over F and $p(x) \mid a(x)b(x)$ then $p(x) \mid a(x)$ or $p(x) \mid b(x)$.*

Proof: suppose $p(x) \in F[x]$ is irreducible over a field F. Then $\langle p(x)\rangle$ is a maximal ideal hence a prime ideal as $F[x]/\langle p(x)\rangle$ is a field and thus an integral domain which implies primality of $\langle p(x)\rangle$ via Theorem 3.2.24. If $a(x), b(x) \in F[x]$ an $p(x) \mid a(x)b(x)$ then $a(x)b(x) = p(x)k(x)$ hence $a(x)b(x) \in \langle p(x)\rangle$ hence $a(x) \in \langle p(x)\rangle$ or $b(x) \in \langle p(x)\rangle$ as $\langle p(x)\rangle$ is a prime ideal. But, $a(x) \in \langle p(x)\rangle$ implies $p(x) \mid a(x)$ and $b(x) \in \langle p(x)\rangle$ implies $p(x) \mid b(x)$. The Corollary follows. □.

The Theorem above is important in the proof that $\mathbb{Z}[x]$ forms a *Unique Factorization Domain*. In particular, the uniqueness stems from this Theorem. Gallian's Example 10 illustrates the utility of the theory in building weird new fields.

Example 3.5.24. *Consider $F = \mathbb{Z}_2$ and the polynomial $x^3 + x + 1$. Notice $x^3 + x + 1 \neq 0$ for $x = 0, 1$ thus $x^3 + x + 1$ is irreducible over $\mathbb{Z}_2$ and hence $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$ is a field. See page 302-303 for further calculations in this field with eight elements. Another way we can understand this field is to work directly with indeterminants. The essential rule is that $x^3 = -x - 1 = x + 1$ in $\mathbb{Z}_2$. So, we can look at elements of the field as $a + bx + cx^2$ where $a, b, c \in \mathbb{Z}_2$ and we multiply as usual subject the interesting rule $x^3 = x + 1$. For example,*

$$x(e + fx + gx^2) = ex + fx^2 + gx^3 = ex + fx^2 + g(x + 1) = g + (e + 1)x + fx^2$$

Or, to focus on the interesting part,

$$x(x^2) = x^3 = x + 1 \qquad \& \qquad x^2(x^2) = xx^3 = x(x + 1) = x^2 + x$$

Consider, always working modulo 2,

$$(x + 1)(x^2 + x) = x^3 + x^2 + x^2 + x = x + 1 + x = 1$$

Of course this field is less fun if we write the coset and not just the representative. In practice, we just write the representative when we do a lot of calculation in a particular context. For example, $\mathbb{C} = \mathbb{R}[x]/\langle x^2+1\rangle$ has typical element $a+bx+\langle x^2+1\rangle$, but, we usually just write $a+bi$ where $i^2 = -1$.

I'll include another of Gallian's excellent examples here:

Example 3.5.25. *The polynomial $x^2 + 1 \in \mathbb{Z}_3$ can be shown to be irreducible. Thus*

$$\mathbb{Z}_3[x]/\langle x^2 + 1\rangle \approx \{a + bx + \langle x^2 + 1\rangle \mid a, b \in \mathbb{Z}_3\}$$

forms a field with nine elements. At the level of representatives, $(a+bx)(c+dx) = ac-bd+(ad+bc)x$ so you can see this is isomorphic to $\mathbb{Z}_3[i]$ which Gallian gave as Example 12 in Chapter 14.

We begin to understand the interplay between ideals in rings and the structure of polynomials. The next feature to explore is the polynomial analog of the prime factorization of integers. Any integer $z \in \mathbb{Z}$ can be expressed as $z = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$ where $p_1, p_2, \ldots, p_k$ are distinct primes. This decomposition is unique upto reordering of the primes.

Theorem 3.5.26. *Every nonzero, non-unit polynomial $f(x)$ in $\mathbb{Z}[x]$ can be written as:*

$$f(x) = b_1 b_2 \cdots b_s p_1(x) p_2(x) \cdots p_m(x)$$

where $b_1, b_2, \ldots, b_s$ are irreducible polynomials of degree 0 and $p_1(x), p_2(x), \ldots, p_m(x)$ are irreducible polynomials of positive degree. This decomposition is unique up to reordering in the sense that if

$$f(x) = c_1 c_2 \ldots c_t q_1(x) q_2(x) \cdots q_n(x)$$

then $t = s$ and $m = n$ and for each j there exists k such that $c_j = \pm b_k$ for $j = 1, \ldots, t$ for each $j' = 1, \ldots, n$ there exists k' such that $p_{j'}(x) = \pm q_{k'}(x)$.

Proof: I'll let you read the proof in Gallian. The argument has three stages. First, we peel off the content which is factored via the prime factorization of integers. This leaves a primitive polynomial which we are able to factor into irreducible factors using a simple induction argument. Finally, the unique factorization centers around the use of the analog of Gauss' lemma for polynomials paired with the fact that the units of $\mathbb{Z}$ are just ± 1. $\square$

Remark 3.5.27. How do we find the **units** in a given unital ring R? We have to solve $xy = 1$ for all possible $x, y \in R$. For $\mathbb{Z}$ a bit of common sense immediately reveals that $x, y = \pm 1$ is all that can be done since otherwise either x or y is forced outside $\mathbb{Z}$. For example, 2 needs $\frac{1}{2} \in \mathbb{Q}$ for a multiplicative inverse. We learn in the next Lecture that many interesting examples are paired with a **norm** and this new calculational tool allows us deeper insight into the structure of units.

I recommend working on Exercises for Gallian Chapter 17 such as: (page 307-311)

$$\#1 - 37 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 27: (these are collected at the start of Lecture 29)

Problem 109: Gallian number 28 from page 292.

Problem 110: Gallian number 38 from page 292.

Problem 111: Gallian number 39 from page 292.

Problem 112: Gallian number 8 from page 308.

Problem 113: Gallian number 10 from page 308.

Problem 114: Gallian number 29 from page 309.

3.6 Lecture 28: divisibility in integral domains I

In this Lecture and the next we study the material presented in Chapter 18 of Gallian. This Lecture is mostly focused on the interplay between the three concepts defined below:

Definition 3.6.1. *Let D be an integral domain. Let $a, b \in D$*

(i.) *a and b are* **associates** *if there exists a unit $u \in D$ for which $b = au$.*

(ii.) *a is an* **irreducible** *if a is not a unit and whenever $a = cd$ then c or d is a unit.*

(iii.) *a with $a \neq 0$ is* **prime** *if a is not a unit and $a \mid bc$ implies $a \mid b$ or $a \mid c$.*

The terms irreducible and prime have been interchanged at various points of your mathematical education. For example, some texts call the irreducible factors in a polynomial factorization the prime factors. It depends on which book you were taught from etc. In the integers every irreducible is prime. The definition of prime in $\mathbb{Z}$ is often given to be that $p \in \mathbb{Z}$ has only itself and 1 as positive divisors. Allowing for negative divisors we'd say p is prime only if $p, -p, 1, -1$ are its sole divisors. This is precisely the notiion of irreducibility defined above. In contrast, we recognize (iii.) as Euclid's Lemma for $\mathbb{Z}$. Of course, both hold for primes in $\mathbb{Z}$ so a prime in $\mathbb{Z}$ is both prime and irreducible as given by (ii.) and (iii.) of the above Definition. Prime and irreducible are not generally equivalent in rings. The example below taken from Gallian page 313 serves well to illustrate:

Example 3.6.2. *Consider $\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\}$ where d is* **square-free**. *To say d is square-free is to say that the prime factorization of d has no factor of the form p^2 for some prime p. For example, $35 = 5(7)$ is square free, but $d = 50 = 5^2(2)$ is not square free. Consider $d = -3$ and study $1 + \sqrt{-3} \in \mathbb{Z}[\sqrt{-3}]$ we can show*[9] *$1 + \sqrt{-3} = xy$ implies x or y is a unit thus $1 + \sqrt{-3}$ is irreducible. On the other hand, note:*

$$(1 + \sqrt{-3})(1 - \sqrt{-3}) = 1 - (-3) = 4 = (2)(2)$$

thus $1 + \sqrt{-3}$ divides $(2)(2)$ yet $1 + \sqrt{-3}$ does not divide 2. Why? Suppose $a, b \in \mathbb{Z}$ such that

$$(\ 1 + \sqrt{-3}\ )(\ a + b\sqrt{-3}\ ) = 2 \quad \Rightarrow \quad (a - 3b) + (b + a)\sqrt{-3} = 2$$

from which we find $a - 3b = 2$ and $a + b = 0$ hence $a = -b$ thus $4a = 2$ so $a = 2/4$ which is absurd as $a \in \mathbb{Z}$ thus $1 + \sqrt{-3}$ does not divide 2. Therefore, $1 + \sqrt{-3}$ is **not prime**, *but, $1 + \sqrt{-3}$* **is irreducible**.

To prove $1 + \sqrt{-3}$ is irreducible we best introduce a new concept: taken from Dummit and Foote page 270. I

Definition 3.6.3. *Let R be an integral domain. Any function $N : R \to \mathbb{N} \cup \{0\}$ with $N(0) = 0$ is called a* **norm** *on R. If $N(a) > 0$ for $a \neq 0$ then N is said to be a* **positive norm**.

In particular, if we study $\mathbb{Z}[\sqrt{d}]$ where d is square-free then I propose we define the norm by analogy to the square of the modulus in $\mathbb{C}$. Remember, $|x + iy|^2 = x^2 + y^2$ can be captured as $|z| = zz^*$ where $z^* = x - iy$. By the same token, if we define $(a + b\sqrt{d})^* = a - b\sqrt{d}$ then

$$(a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - db^2$$

This motivates the following convenient definition of norm:

[9] we'll use the concept of a norm to accomplish this a bit later in this Lecture, see Example 3.6.6

Definition 3.6.4. *Let $\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\}$ where d is square-free then define*

$$N(a + b\sqrt{d}) = |a^2 - db^2|$$

for each $a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$.

The fact that the formula above defines a norm is immediate from the fact $N(0) = 0$ and the fact that the absolute value is non-negative. If $d < 0$ then we can write $N(a + b\sqrt{d}) = a^2 + db^2$ as the sum of squares is automatically non-negative.

Theorem 3.6.5. *If d is square-free and $N(a + b\sqrt{d}) = |a^2 - db^2|$ for each $a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ then*

(i.) *$N(x) = 0$ if and only if $x = 0$*

(ii.) *$N(xy) = N(x)N(y)$ for all $x, y \in \mathbb{Z}[\sqrt{d}]$*

(iii.) *$x \in \mathbb{Z}[\sqrt{d}]$ is a unit if and only if $N(x) = 1$*

(iv.) *if $N(x)$ is prime then x is irreducible in $\mathbb{Z}[\sqrt{d}]$*

Proof: I leave (i.) and (ii.) this as a rather enjoyable exercises. To prove (iii.), suppose x is a unit then $xy = 1$ for some y and hence $N(1) = N(x)N(y)$ but $N(1) = |1^2 + d(0^2)| = 1$ hence $1 = N(x)N(y)$ but $N(x), N(y) \in \mathbb{Z}$ hence $N(x) = N(y) = 1$. Next, to prove (iv.) suppose $N(x)$ is prime and suppose $x = yz$ for some $y, z \in \mathbb{Z}[\sqrt{d}]$ then $N(x) = N(yz) = N(y)N(z)$. Now, $N(x)$ is prime thus either $N(y) = 1$ or $N(z) = 1$ and hence either y or z is a unit by (iii.). Therefore, x is irreducible. □

Example 3.6.6. *Let us see why $1 + \sqrt{-3}$ is irreducible. Suppose $1 + \sqrt{-3} = xy$. Observe*

$$N(1 + \sqrt{-3}) = 1^2 - (-3)1^2 = 4 = N(xy) = N(x)N(y)$$

if x, y are not units then we must have $N(x) = N(y) = 2$. Consider,

$$a^2 + 3b^2 = 2$$

there is no solution! Consequently, $1 + \sqrt{-3} = xy$ implies x or y is a unit. Thus $1 + \sqrt{-3}$ is irreducible.

Gallian warns us that proving things in $\mathbb{Z}[\sqrt{d}]$ is more trouble when $d > 1$. Let us work through his Example 2 on page 313.

Example 3.6.7. *Consider $7 \in \mathbb{Z}[\sqrt{5}]$. Suppose $7 = xy$ for some $x, y \in \mathbb{Z}[\sqrt{5}]$. We have*

$$N(7) = N(xy) = N(x)N(y) \quad \Rightarrow \quad 49 = N(x)N(y)$$

if x, y are not units we must have $N(x) = N(y) = 7$. Suppose $x = a + b\sqrt{7}$ with $N(x) = 7$ then

$$7 = |a^2 - 5b^2| \quad \text{or if you prefer} \quad a^2 - 5b^2 = \pm 7.$$

Any integer solution of $a^2 - 5b^2 = \pm 7$ is an $\mathbb{Z}_7$ solution of $a^2 - 5b^2$. Explicit checking of possible solutions shows the only solution is $a = b = 0$ modulo 7. Thus $7 \mid a$ and $7 \mid b$ which gives $|a^2 - 5b^2|$ is divisible by 49. Yet, $|a^2 - 5b^2| = 7$ which is clearly not divisible by 49 hence no solution of $a^2 - 5b^2 = \pm 7$ exists for $a, b \in \mathbb{Z}$.

Theorem 3.6.8. *In an integral domain every prime is an irreducible.*

Proof: suppose a is a prime in an integral domain. If $a = xy$ then as a is prime we have $a \mid x$ or $a \mid y$. Suppose $a \mid x$ then $x = ab$ for some b. Thus,

$$x(1) = x = ab = (xy)b = x(yb)$$

thus $1 = yb$ and we find y is a unit. Similar argument shows x is a unit in the case $a \mid y$ thus $a = xy$ implies x or y is a unit and we conclude that a is ireducible. □

The concept of associates is helpful for some calculations we have struggled with a bit in our recent work. Here is a Theorem that should help us with the task of identifying possible coset representatives in a given quotient of a unital ring R by an ideal I:

Theorem 3.6.9. *Let R be a commutative ring with identity 1. If a, b are associates then $\langle a \rangle = \langle b \rangle$. Furthermore, if R is an integral domain and $I = \langle a \rangle$ then any other generator of I is an associate of a.*

Proof: if a, b are associates then there exists a unit u in R for which $a = bu$ and $b = au^{-1}$. Let $x \in \langle a \rangle$ then $x = ar$ for some $r \in R$. Hence $x = bur$ and as $ur \in R$ this shows $x \in \langle b \rangle$ hence $\langle a \rangle \subseteq \langle b \rangle$. If $y \in \langle b \rangle$ then $y = br = au^{-1}r \in \langle a \rangle$ hence $\langle b \rangle \subseteq \langle a \rangle$ and thus $\langle a \rangle = \langle b \rangle$. Suppose $\langle c \rangle = \langle a \rangle$ for some $c \in R$. If $\langle a \rangle = \{0\}$ then $a = 0$ otherwise $a \neq 0$ implies $a(1) = a \in \langle a \rangle \neq \{0\}$ and $a = 0$ then implies $c = 0$ as well. The Theorem is trivially true for $a = 0$ since 0 is an associate of itself and there is no distinct associate of 0. Suppose $a \neq 0$ hence $c \neq 0$. Note $a, c \in \langle c \rangle$ and $a, c \in \langle a \rangle$ thus there exists $s, r \in R$ for which $a = rc$ and $c = sa$ hence $a = rc = (rs)a$. As $a \neq 0$ we deduce from the cancellation property of the integral domain R that $rs = 1$ hence r is a unit and $a = rc$ shows a, c are associates. □

What happens in general when R is not integral. Is it possible that $\langle a \rangle = \langle c \rangle$ and a, c are not associates? Consider, $R = \mathbb{Z}_6$ for then $\langle 2 \rangle = \langle 4 \rangle = \{0, 2, 4\}$. Are $2, 4$ associates? Well, can we find a unit $u \in U(\mathbb{Z}_6) = \{1, 5\}$ for which $4 = 2u$? There are two choices: $2(1) = 2 \neq 4$ and $2(5) = 10 = 4$. Yes, in this case, $2(5) = 4$ and 5 is a unit hence $2, 4$ are associates. This shows the second part of Theorem 3.6.9 **can** be true outside the context that R be an integral domain. For an non-example, see this mathstack Q and A.

Theorem 3.6.10. *In a principal ideal domain, an element is prime if and only if it is irreducible.*

Proof: Let D be a PID. Note D is an integral domain by assumption[10] thus Theorem 3.6.8 tells us that each prime is irreducible. Conversely, suppose a is irreducible. Suppose $a \mid bc$ for some $b, c \in D$. Define

$$I = \{ax + by \mid x, y \in D\}$$

we can show I is an ideal. Note $z, w \in I$ have the form $z = ax + by$ and $w = ax' + by'$ for some $x, y, x', y' \in D$. Thus,

$$z - w = ax + by - (ax' + by') = a(x - x') + b(y - y') \in I$$

and for $r \in D$,

$$rz = r(ax + by) = a(rx) + b(ry) \in I$$

[10] a PID is an integral domain in which every ideal is principal.

thus I is an ideal. Since D is a PID we know I is principal. Thus there exists $d \in D$ for which $I = \langle d \rangle$. Observe $a = a(1) + b(0) \in I$ thus $a = rd$ for some $r \in D$. Since a is irreducible we have r or d is a unit.

If d is a unit then $1 = dd'$ for some $d' \in D$ thus $1 \in I$. Therefore, $1 = ax + by$ for some $x, y \in D$. Multiply by c to see:

$$c = cax + cby = acx + (bc)y.$$

Naturally, $a \mid acx$ and we assumed $a \mid bc$ thus, by the equation above, $a \mid c$.

If r is a unit then $a = rd$ provides a and r are associates. Theorem 3.6.9 provides $\langle d \rangle = \langle a \rangle$ hence $I = \langle a \rangle$ and as $b = a(0) + b(1) \in I$ we find $b = a\lambda$ for some $\lambda \in D$. Therefore, $a \mid b$.

In summary, for an irreducible $a \in D$ we find $a \mid bc$ implies $a \mid b$ or $a \mid c$ which shows a is prime. □

In short, PIDs allow us to carelessly interchange the concepts of prime and irreducible. It's sort of like those new cars where they encourage you to ignore the road[11].

Example 3.6.11. *$\mathbb{Z}$ is a principal ideal domain. You can prove any ideal in $\mathbb{Z}$ has the form $\langle n \rangle = n\mathbb{Z}$. Likewise, if F is a field then we showed that $F[x]$ is a principal ideal domain in Theorem 3.4.17. Not all integral domains are principal. Gallian provides us the example $\mathbb{Z}[x]$ of $\langle 2, x \rangle$ which he defines a bit differently on page 314-315. Details can be found in his Example 3.*

I recommend working on Exercises for Gallian Chapter 17 such as: (page 307-311)

$$\#1 - 37 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 28: (these are collected at the start of Lecture 30)

Problem 115: Prove part (i.) of Theorem 3.6.5.

Problem 116: Prove part (ii.) of Theorem 3.6.5.

Problem 117: Gallian number 14 on page 308

Problem 118: Gallian number 4 from page 325.

Problem 119: Gallian number 17 from page 326.

Problem 120: Gallian number 18 from page 326.

[11] current commercials teach me it's cool to day dream in the car

3.7 Lecture 29: divisibility in integral domains II

In this Lecture we complete our study of Chapter 18 of Gallian. Here we explore the interplay between Euclidean Domains, Principal Ideal Domains and Unique Factorization Domains.

Definition 3.7.1. *Let D be an integral domain. Then D is said to be a* **Euclidean Domain** *if there is a norm N on D such that for any two elements $a, b \in D$ with $b \neq 0$ there exists elements $q, r \in D$ with*

$$a = qb + r$$

and $r = 0$ or $N(r) < N(b)$. We call q the **quotient** *and r the* **remainder** *of the division.*

You can contrast the definition above to that which is given in Gallian. In part, a Euclidean Domain is an integral domain D with a function $d : D \to \mathbb{N} \cup \{0\}$ such that $d(a) \leq d(ab)$ for all $a, b \neq 0$ in D. If we have a positive norm for which $N(xy) = N(x)N(y)$ then define $d(x) = N(x)$ and note:

$$d(ab) = N(ab) = N(a)N(b) = d(a)d(b)$$

and as $a, b \neq 0$ we have $d(a), d(b) \in \mathbb{N}$ thus $d(a) = d(ab)/d(b) \leq d(ab)$. In short, if we have a positive multiplicative norm then it provides a measure (in the langauge of Gallian page 321). I should caution, we do not assume all norms are multiplicative, see Example 3.7.3.

We should notice a Euclidean Domain does not generally come with a division algorithm which produces a unique quotient and remainder. Even the integers allow for non-unique quotient and remainder in a division. Notice Theorem 3.6.5 applies to norms for rings other than $\mathbb{Z}[\sqrt{d}]$ for d square-free. If N is a norm which is positive and multiplicative then we satisfy (i.) and (ii.) of Theorem 3.6.5 hence (iii.) and (iv.) follow since the proof of (iii.) and (iv.) simply require the verity of (i.) and (ii.).

Example 3.7.2. *Consider $D = \mathbb{Z}$ with $N(x) = |x|$. It is simple to see N defines a positive norm and $N(xy) = |xy| = |x||y| = N(x)N(y)$ for all $x, y \in \mathbb{Z}$. Notice $|u| = 1$ implies $u = \pm 1$. The units in $\mathbb{Z}$ are just $1, -1$. Let me give an explicit example to make the ambiguity of the division algorithm a bit more explicit. Consider $a = 54$ and $b = 8$ we have:*

$$54 = 6(8) + 6 \qquad or \qquad 54 = 7(8) - 2.$$

Now, in the context of the integers the use of a positive remainder is what is usually done.

I merely mean to indicate that even in $\mathbb{Z}$ the division algorithm may not be unique.

Example 3.7.3. *If F is a field then $D = F[x]$ is a Euclidean Domain where we define $N(f(x)) = deg(f(x))$. Since $deg(f(x)g(x)) = deg(f(x)) + deg(g(x))$ we don't have a multiplicative norm. The units of D are nonzero constant polynomials which have $N(f(x)) = N(c) = 0$.*

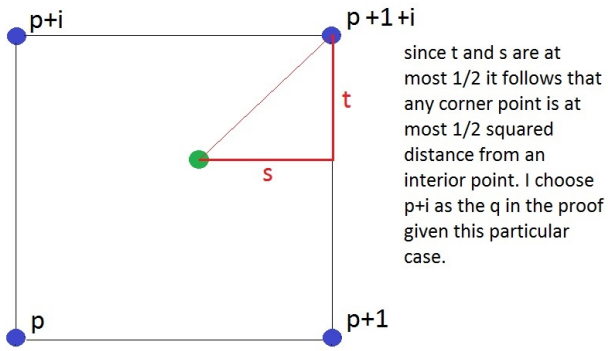
Example 3.7.4. *The Gaussian integers $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ form a Euclidean Domain with $N(a + ib) = a^2 + b^2$. It is easy to prove $N(zw) = N(z)N(w)$ and $N(z) = 0$ iff $z = 0$ hence $N : \mathbb{Z}[i] \to \mathbb{N} \cup \{0\}$ forms a multiplicative norm. The proof that $\mathbb{Z}[i]$ is a Euclidean Domain with respect to N is a bit involved. I'll let you read page 322-323 for Gallian's proof. I'll sketch a similar proof here. To divide $a+ib$ by $c+id$ we may accomplish this explicitly in $\mathbb{C}$ as $z = \frac{a+ib}{c+id}$ is a complex number. The Gaussian integers from a lattice of points and we simply pick one of the four points in $\mathbb{Z}[i]$ which are closest to z and call it q. Define $r = a+ib-q(c+id)$ then clearly $a+ib = q(c+id)+r$ and as*

$$\frac{a + ib}{c + id} = z = q + \frac{r}{c + id}$$

by the construction of q, worst case scenario we find z as the center point of a cell in the $\mathbb{Z}[i]$ lattice. Notice the center point is distance $1/\sqrt{2}$ from each of the closest 4 points. Thus:

$$\left|\frac{r}{c+id}\right| < \frac{1}{\sqrt{2}} \quad \Rightarrow \quad |r| < \frac{|c+id|}{\sqrt{2}} \quad \Rightarrow \quad N(r) < N(c+id)/2.$$

Perhaps the following picture helps explain the proof in the Example above:



No matter where $\frac{a+ib}{c+id}$ lands in the complex plane the closest point in $\mathbb{Z}[i]$ will be within $1/\sqrt{2}$ distance. When we study other $\mathbb{Z}[\sqrt{-d}]$ for $d > 0$ the geometry of this argument is spoiled. There is much to learn about Euclidean Domains which is not emphasized in Gallian. Familar algorithms and concepts in $\mathbb{Z}$ have natural generalizations to abstract Euclidean Domains. For example, we can execute the Euclidean Algorithm in $\mathbb{Z}[i]$ just as we do in $\mathbb{Z}$ by systematically removing first the divisor, then the remainder, then the remainder of the remainder's division etc...

Example 3.7.5. *Consider $\alpha = 11 + 3i = a + ib$ and $\beta = 3i + 2 = c + id$ (a, b, c, d notation in reference to the proof above). Let's walk through the Euclidean Algorithm in vector format: in each step I have to do side calculation (not shown) to decide which multiple of the previous remainder should be subtracted to make the difference minimal. If I don't see it by inspection then I follow the method of the proof.*

$$(11 + 3i, 3i + 2) = (\alpha, \beta)$$
$$(3i + 2, 1 + i) = (\beta, \alpha - (2 - 2i)\beta)$$
$$(1 + i, -i) = (\alpha - (2 - 2i)\beta, \beta - (3 + i)[\alpha - (2 - 2i)\beta])$$

at which point we stop since $-i$ is a unit in $\mathbb{Z}[i]$. Thus,

$$-i = \beta - (3 + i)\alpha + (3 + i)(2 - 2i)\beta$$

or

$$-i = (9 - 4i)\beta - (3 + i)\alpha$$

hence

$$1 = (4 + 9i)(3i + 2) + (1 - 3i)(11 + 3i).$$

This calculation shows the greatest common divisor of $11 + 3i$ and $3i + 2$ is 1, or, you could say $-1, i, -i$. In fact, to study this properly we need to embrace the concept that the gcd is an ideal. In this case,

$$\langle 11 + 3i \rangle + \langle 3i + 2 \rangle = \langle 1 \rangle = \mathbb{Z}[i]$$

The ideals $\langle 11+3i \rangle$ and $\langle 3i+2 \rangle$ are **comaximal** *since there sum is the entire ring. Comaximal ideals are the ideal version of relatively prime. Note, two integers a, b are relatively prime if $gcd(a, b) = 1$ which implies $ak + bl = 1$ hence $x = akx + blx$ for each $x \in \mathbb{Z}$ and thus $\langle a \rangle + \langle b \rangle = \mathbb{Z}$.*

The calculations and concepts we find in Euclidean Domains were largely pioneered by mathematicians such as Euler, Gauss and their students in the nineteenth century. The necessity of facing the existence of a unique factorization and/or how to deal with the absence of a unique factorization property took a bit longer to be appreciated. As Gallian describes on page 316, the assumption of unique factorization misled Gabriel Lamé to claim he had a proof of Fermat's last theorem (which is that $x^n + y^n = z^n$ has no integer solutions for $n \geq 3$). Unfortunately, Lamé was not familar with the work of Kummer which demonstrated the factorization into irreducibles was not unique in the natural sense which Lamé assumed.

It seems Gauss was aware of this issue when he basically avoided using abstract ring arguments. Gauss was aware of Euler's work and Euler and Lagrange used objects like $a + b\sqrt{-d}$ to prove various assertions about primes. Gauss likely realized the danger made explicit by Kummer. Stillwell explains this story in more depth in his text *Elements of Number Theory*. Basically, Gauss just brute-force[12] solved the problems which Euler and Lagrange had been working on in more elegant ways. In some sense, this was bad mathematics, it took some time for us to return to the elegance which Euler and Lagrange had partially understood. The fix to the ambiguity suffered by Lamé was given in part by Kummer with his introduction of **ideal numbers**. This program was fleshed out by Dedekind. Basically, ideals play the role that numbers previously held. The ambiguity is washed away in that there is a unique factorization property for ideals in a ring of algebraic integers[13]

Ultimately, the work of Dedekind brought questions to the mind of Emmy Noether who was one of the first true abstract algebraists. Her work was about structure much more than particular examples. She embraced the concept of abstraction as a means to solve many problems in an elegant fashion. I mention Noether here because the chain condition argument below is certainly due to her influence on our current understanding of abstract algebra.

Definition 3.7.6. *Let D be an integral domain. D is a* **Unique Factorization Domain** *if*

(i.) *every nonzero element of D can be written as a product of irreducible elements in D,*

(ii.) *the factorization of a given element in D into irreducibles is unique up to reordering and associates. In particular, if $x \in D$ has irreducible factorizations $x = x_1 x_2 \cdots x_n$ and $x = y_1 y_2 \cdots y_n$ then there exist units $u_1, u_2, \ldots, u_n$ for which*

$$\{y_1, y_2, \ldots, y_n\} = \{u_1 x_1, u_2 x_2, \ldots, u_n x_n\}$$

where we do not intend the above equality to imply an ordering.

The uniqueness up to associates is easy enough to see in the context of $\mathbb{Z}$ where the units are ± 1 or $F[x]$ where any nonzero scalar is a unit.

Theorem 3.7.7. Ascending Chain Condition in a PID: *In a principal ideal domain, any strictly increasing chain of ideals $I_1 \subset I_2 \subset \cdots$ must be finite in length.*

Proof: let $I_1 \subset I_2 \subset \cdots$ be a chain of strictly increasing ideals in an principlal ideal domain D. Note $\cup_{j \in \mathbb{N}} I_j$ forms an ideal thus $I = I_1 \cup I_2 \cup \cdots = \langle d \rangle$ for some $d \in D$. Note $d \in I$ implies $d \in I_n$ for some $n \in \mathbb{N}$. But, $I_i \subseteq I = \langle d \rangle \subseteq I_n$ for each $i \in \mathbb{N}$ hence I_n must be the terminal ideal in the chain. $\square$

[12] as in he solved congruence questions via explicit algebra in $\mathbb{Z}$ etc.

[13] see page 767, Corollary 16, of Dummit and Foote, this claim is quite a bit beyond our current course.

I didn't prove unique factorization of $\mathbb{Z}[x]$ (gory detail on page 304-305 of Gallian for the curious), but, if I had this still would not help as $\mathbb{Z}[x]$ is not a PID. That said, if F is a field then this proof gets us that $F[x]$, a PID by Theorem 3.4.17, is a unique factorization domain. The proof of this theorem is perhaps the most interesting proof we will study this semester:

Theorem 3.7.8. *Every principal ideal domain is a unique factorization domain.*

Proof: let D be a PID with set of units U. Let $a_0 \in D$ with $a_0 \neq 0$ and $a_0 \notin U$. Game plan:

(1.) show a factorization of a_0 contains at least one irreducible

(2.) show there is a factorization of a_0 into a product of irreducibles

(3.) show uniqueness up to associates

(1.) If a_0 is irreducible then we have shown a_0 contains an irreducible. Otherwise, $a_0 = a_1 b_1$ where a_1 is not a unit and $b_1 \neq 0$. If a_1 is irreducible then a contains an irreducible. Otherwise, suppose $a_1 = a_2 b_2$ where $b_2 \neq 0$ and a_2 is not a unit. Continue in this fashion to define a_{n+1} not a unit and $b_{n+1} \neq 0$ for which $a_n = a_{n+1} b_{n+1}$ for $n = 3, 4, \ldots$. Observe, $a_n = a_{n+1} b_{n+1}$ implies $\langle a_n \rangle \subset \langle a_{n+1} \rangle$ for $n = 0, 1, 2, \ldots$ thus by Theorem 3.7.7 there exists k for which this ascending chain of ideals terminates:

$$\langle a_0 \rangle \subset \langle a_1 \rangle \subset \langle a_2 \rangle \subset \cdots \subset \langle a_k \rangle.$$

But, the chain terminates when a_k does not permit a factorization into non-units. Hence a_k is irreducible hence $a_0 = r a_k$ shows a_0 contains an irreducible.

(2.) if a_0 is irreducible then we have a factoring of a_0 into irreducibles. Otherwise, by (1.) there exists an irreducible p_1 and a non-unit c_1 for which $a_0 = p_1 c_1$. If c_1 is an irreducible then we have factored a_0 into irreducibles. Otherwise, apply (1.) to the non-unit c_1 to find $c_1 = p_2 c_2$ where p_2 is irreducible and c_2 is not a unit. Notice we have another ascending chain of ideals:

$$\langle a_0 \rangle \subset \langle p_1 \rangle \subset \langle p_2 \rangle \subset \cdots$$

this must terminate, say at $\langle p_t \rangle$. By the construction of the chain, we find p_t is an irreducible and

$$a_0 = p_1 c_1 = p_1 p_2 c_2 = \cdots = p_1 p_2 \cdots p_t.$$

Therefore, a_0 is factored into a product of irreducibles.

(3.) Suppose a_0 has two factorizations into irreducibles:

$$a_0 = p_1 p_2 \cdots p_t = q_1 q_2 \cdots q_s$$

We prove the factorization is unique by induction on t. Suppose $t = 1$ then $a_0 = p_1 = q_1 q_2 \cdots q_s$ implies $s = 1$ as to say otherwise contradicts the irreducibilty of p_1. Next, suppose inductively, any factorization into less than t irreducibles is unique up to associates. Again, if

$$a_0 = p_1 p_2 \cdots p_t = q_1 q_2 \cdots q_s$$

then note $p_1 \mid q_1 q_2 \cdots q_s$ hence (by an exercise I might assign) there exists some q_j for which $p_1 \mid q_j$ and thus $p_1 = u_j q_j$ for some unit u_j. Then,

$$q_j u_j p_2 \cdots p_t = q_j q_2 \cdots q_s$$

and by the cancellation property for integral domains ($q_j \neq 0$)

$$u_j p_2 \cdots p_t = q_j q_2 \cdots q_s$$

and by the induction hypothesis we conclude that the remaining $t-1$ irreducibles $u_j p_2, \ldots, p_{j-1}$, $p_{j+1}, \ldots, p_t$ must be associated to $s-1 = t-1$ irreducibles $q_2, \ldots, q_{j-1}, q_{j+1}, \ldots, q_t$. Thus, the factorization of a_0 into irreducibles is unique up to associates and ordering. □

I tried to follow Gallian pretty closely here. Essentially the same proof is given on page 319-320.

Corollary 3.7.9. *Let F be a field. Then $F[x]$ is a unique factorization domain.*

Proof: we proved in Theorem 3.4.17 for F a field the polynomials $F[x]$ form a PID hence by Theorem 3.7.8 we find $F[x]$ is a UFD. □

I abbreviate to illustrate the utility of these abbreviations.

Theorem 3.7.10. *Every euclidean domain is a principal ideal domain.*

Proof: let D be a Euclidean Domain with norm N. If I is a nonzero ideal in D then notice $S = \{N(x) \mid x \in I\}$ is a nonempty subset of non-negative integers. Thus, by the Well-Ordering-Principle, S has a smallest member s_o. Let $x_o \in I$ be a member of I for which $N(x_o) = s_o$. If $z \in I$ then apply the division algorithm in D to obtain q and r for which

$$z = qx_o + r$$

Note $z \in I$ by assumption and $qx_o \in I$ by as $x_o \in I$ thus

$$r = z - qx_o \in I$$

Therefore, $r = 0$ as $r \neq 0$ would provide $r \in I$ for which $N(r) < N(x_o) = s_o$ which contradicts the minimality of s_o in S. In short, every element $z \in I$ is found in $\langle x_o \rangle$. But, I was arbitrary nonzero ideal hence every nonzero ideal is princpal. Moreover, $\langle 0 \rangle = \{0\}$ and the Theorem follows. □

I hope you see this proof is nearly identical in structure to that we gave for Theorem 3.4.17. In retrospect, we could have skipped that proof and simply applied this general result to the context of the norm on $F[x]$ being specified by the degree function.

Corollary 3.7.11. *Every euclidean domain is a unique factorization domain.*

Proof: note Theorem 3.7.10 gives that D Euclidean implies D is a PID. Then Theorem 3.7.8 provides that D a PID implies D is a UFD. □

Notice that $\mathbb{Z}[x]$ is a UFD, but, $\mathbb{Z}[x]$ is not a PID. The implications in the proof above are not reversible. An example of a PID which is not a Euclidean Domain is a bit harder to find. Gallian gives a reference. I'll add the following link: Tom Oldfield's Construction of PIDs which are not Euclidean Domains the other answer by Bill Dubuque is also useful. Both answers are a bit beyond this course. I expect you to be aware of these results, but, I don't expect you can actually produce a PID which is not a Euclidean Domain. In contrast, knowing that $\mathbb{Z}[x]$ is a UFD but not a PID is exactly the sort of thing you ought to know.

Theorem 3.7.12. *If D is a unique factorization domain then $D[x]$ is a unique factorization domain.*

Proof: I'll give exactly as much proof as Gallian on this one. □

Next, we study an elegant proof of Eisenstein's Criterion: (stated as Theorem 3.5.18 in these notes)

Proof: (Gallian credits Richard Singer for the proof we give here). Suppose $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]$ and the prime p is such that $p \nmid a_n$ but $p \mid a_j$ for $j = n-1, \ldots, 0$ and $p^2 \nmid a_0$. Suppose $f(x)$ is reducible over $\mathbb{Q}$. Then $f(x) = g(x)h(x)$ in $\mathbb{Z}[x]$ by Theorem 3.5.11. Notice modulo p the polynomial reduces to $\overline{f(x)} = a_n x^n$ hence $a_n x^n = \overline{g(x)}\ \overline{h(x)}$. But, x is an irreducible in $\mathbb{Z}_p[x]$ and as $\mathbb{Z}_p[x]$ is a UFD as $\mathbb{Z}_p$ is a field we deduce that $x \mid \overline{g(x)}$ and $x \mid \overline{h(x)}$ from which we deduce $\overline{g(0)} = 0$ and $\overline{h(0)} = 0$ thus $p \mid g(0)$ and $p \mid h(0)$ and $f(x) = h(x)g(x)$ gives $f(0) = a_0 = h(0)g(0)$ and we find $p^2 \mid a_0$ which is a contradiction. Consequently, $f(x)$ is irreducible over $\mathbb{Q}$. □.

Example 3.7.13. *A nice example where unique factorization fails is provided by* $\mathbb{Z}[\sqrt{-5}]$. *Note* $\mathbb{Z}[\sqrt{-5}]$ *forms a subring of* $\mathbb{C}$ *hence is commutative and has no zero divisors. Moreover,* $1 = 1 + 0\sqrt{-5} \in \mathbb{Z}[\sqrt{-5}]$ *hence* $\mathbb{Z}[\sqrt{-5}]$ *is an integral domain. We have multiplicative norm* $N(a+b\sqrt{-5}) = a^2 + 5b^2$. *Solving*

$$a^2 + 5b^2 = 1$$

we find just two solutions, $a = 1, b = 0$ *or* $a = -1, b = 0$. *There are just the units* $-1, 1$ *thus judging if a pair of elements are associates is quite easy. Observe,*

$$46 = (2)(23) \qquad \& \qquad 46 = (1 + 3\sqrt{-5})(1 - 3\sqrt{-5})$$

It is immediately clear these the factors 2, 23, $1 + 3\sqrt{-5}$ *and* $1 - 3\sqrt{-5}$ *are not associates. Furthermore, their irreducibility may be shown from the usual arguments involving the norm. Suppose* $2 = xy$ *for some* $x, y \in \mathbb{Z}[\sqrt{-5}]$ *then* $N(2) = 4 = N(x)N(y)$ *and if* x, y *are not units then we need* $N(x) = N(y) = 2$. *Yet,* $a^2 + 5b^2 = 2$ *clearly has no solution in* $\mathbb{Z}$. *Therefore,* 2 *is irreducible. Similarly, if* $23 = xy$ *then we would need to find a solution to* $a^2 + 5b^2 = 23$ *to give solution to* $23 = xy$ *where* x, y *are not units. Explicit trial of reasonable* $\mathbb{Z}$ *rules out hope of a solution to* $a^2 + 5b^2 = 23$. *Continuing, if* $1 + 3\sqrt{-5} = xy$ *then* $N(1 + 3\sqrt{-5}) = 1 + 5(9) = 46$ *we require* $N(x) = 2$ *and* $N(y) = 23$ *without loss of generality. Again, it is not possible to solve* $a^2 + 5b^2 = 2$ *over* $\mathbb{Z}$. *In summary, we have provided two factorizations of* 46 *into irreducibles and there is no hope these are equivalent up to associates and reordering.* $\mathbb{Z}[\sqrt{-5}]$ *is not a UFD.*

I recommend working on Exercises for Gallian Chapter 18 such as: (page 325-327)

$$\#1 - 37 \ (odds).$$

also, Supplemental Exercises for Chapters 15-18 (page 331-332)

$$\#1 - 35 \ (odds).$$

keep in mind the notation $\oplus_{Gallian} = \times_{me}$. Of course, I doubt anyone has time to do these all, but, the more you do, the more you know. (incidentally, this homework is worth 150hpts, the 4-problem assignments in the past are weighted 100hpts in contrast)

Problems for Lecture 29: (these are collected at the start of Lecture 31)

Problem 121: Calculate the $gcd(\alpha, \beta)$ for $\alpha = 12 + 3i$ and $\beta = 6 - 9i$. Use the vector Euclidean algorithm much as in Example 3.7.5.

Problem 122: Gallian number 10 on page 325

Problem 123: Gallian number 14 from page 325.

Problem 124: Gallian number 19 from page 326.

Problem 125: Gallian number 7 from page 331.

Problem 126: Gallian number 24 from page 332.

3.8 Lecture 30: extension fields

Here we follow Section 29 of Fraleigh's *Abstract Algebra*. We have discovered and studied many abstract fields in various lectures up to this point. This Lecture introduces the major elementary theorems of field theory[14].

Definition 3.8.1. *A field E is an* **extension field** *of F if $F \subseteq E$ and the operations of F are those of E restricted to F. We call F the* **base field** *of the extension.*

We already know several examples.

Example 3.8.2. $\mathbb{R}$ *is an extension of* $\mathbb{Q}$.

Example 3.8.3. $\mathbb{C}$ *is an extension of* $\mathbb{R}$.

We also may take note that:

Example 3.8.4. $\mathbb{C}$ *is an extension of* $\mathbb{Q}$.

We've also studied other less common cases.

Example 3.8.5. *The set $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ forms a subfield of $\mathbb{R}$. We see $\mathbb{Q}(\sqrt{2})$ as an extension field of $\mathbb{Q}$. Furthermore, we find $\mathbb{R}$ is an extension field of $\mathbb{Q}(\sqrt{2})$.*

Example 3.8.6. *In Problem 95 (Gallian's exercise 45 of page 262) we showed $\mathbb{Z}_2[x]/\langle x^2 + x + 1\rangle$ is a field. Noting that $\{I, 1 + I\}$ serves as an isomorphic copy of $\mathbb{Z}_2$ in $\mathbb{Z}_2[x]/\langle x^2 + x + 1\rangle$ we find $\mathbb{Z}_2[x]/\langle x^2 + x + 1\rangle$ is an extension of $\mathbb{Z}_2$. This field appears as the final example in Section 29 of Fraleigh.*

In the Example above we assume the reader is willing to identify a field which is isomorphic to $\mathbb{Z}_2$ as $\mathbb{Z}_2$. This slight abuse of language pervades this section. The field which is honestly extended is usually an isomorphic copy of the field we propose to extend.

Theorem 3.8.7. Fundamental Theorem of Field Theory (Kronecker, 1887): *let F be a field and $f(x) \in F[x]$ a nonconstant polynomial. Then there exists an extension field E of F in which $f(x)$ has a zero.*

Proof: if F is a field and $f(x) \in F[x]$ is a nonconstant polynomial then $f(x)$ is not a unit and hence there exists a factorization of $f(x)$ into irreducibles by Corollary 3.7.9. Suppose $p(x)$ is an irreducible in the factorization of $f(x)$; that is $f(x) = p(x)g(x)$ for $p(x)$ irreducible in $F[x]$. Suppose

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

We propose $E = F[x]/\langle p(x)\rangle$. Since $\langle p(x)\rangle$ is irreducible it follows E is a field (see Corollary 3.5.22). It remains to show $f(x)$ has a zero in E. Let $\alpha = x + \langle p(x)\rangle$. Observe,

$$\alpha^j = (x + \langle p(x)\rangle)^j = x^j + \langle p(x)\rangle$$

Hence,

$$p(\alpha) = \sum_{a_j=0}^{n} a_j\big(x^j + \langle p(x)\rangle\big) = \left(\sum_{a_j=0}^{n} a_j x^j\right) + \langle p(x)\rangle = p(x) + \langle p(x)\rangle = \langle p(x)\rangle = 0. \qquad \square$$

The proof above tells us how to create a field with a zero for a given polynomial.

[14] not to be confused with field theory in physics, which, means something rather different

Example 3.8.8. *Consider* $f(x) = x^2 + 4x + 5 \in \mathbb{R}[x]$. *Notice,*

$$f(x) = (x+2)^2 + 1$$

hence there is no real zero of $f(x)$ *and hence* $f(x)$ *is irreducible over* $\mathbb{R}$. *However,* $\alpha = x + \langle x^2 + 4x + 5\rangle$ *will serve as a zero of* $f(x)$ *in* $\mathbb{R}[x]/\langle x^2 + 4x + 5\rangle$. *Indeed,*

$$\begin{aligned} f(\alpha) &= \alpha^2 + 4\alpha + 5 \\ &= (x^2 + \langle f(x)\rangle) + (4x + \langle f(x)\rangle) + 5 + \langle f(x)\rangle \\ &= x^2 + 4x + 5 + \langle f(x)\rangle \\ &= \langle f(x)\rangle. \end{aligned}$$

Here $f(\alpha)$ *is understood to be* $f(x)$ *evaluated via the evaluation homomorphism. Furthermore, notice that* $5 \in \mathbb{R}$ *is replaced with* $5 + \langle x^2 + 4x + 5\rangle$ *in the evaluation. We cannot add* $5 \in \mathbb{R}$ *to cosets in* $\mathbb{R}[x]/\langle x^2 + 4x + 5\rangle$, *but, the coset represented by* 5 *is in natural correspondance to* 5. *In short, Fraleigh and other abstract algebra texts expect you to set* $5 = 5 + \langle f(x)\rangle$ *in such discussions. Admittedly, this is necessary, but, I'm not entirely pleased about the lack of discussion on this point.*

I follow Fraleigh's Example 29.5 next:

Example 3.8.9. *The polynomial* $f(x) = x^4 - 5x^2 + 6 = (x^2 - 2)(x^2 - 3)$ *is reducible over* $\mathbb{Q}$, *however,* $x^2 - 2$ *and* $x^2 - 3$ *are irreducible over* $\mathbb{Q}$. *It follows we can form field* $\mathbb{Q}[x]/\langle x^2 - 2\rangle$ *in which the element* $\alpha = x + \langle x^2 - 2\rangle$ *satisfies* $\alpha^2 - 2 = 0$. *Likewise,* $\beta = x + \langle x^2 - 3\rangle$ *satisfies* $\beta^2 - 3 = 0$ *in* $\mathbb{Q}[x]/\langle x^2 - 3\rangle$. *In fact,* $f(\alpha) = (\alpha^2 - 2)(\alpha^2 - 3) = (0)(\alpha^2 - 3) = 0$ *and* $f(\beta) = (\beta^2 - 2)(\beta^2 - 3) = (\beta^2 - 2)(0) = 0$. *We can calculate, viewing* $f(x) \in (\mathbb{Q}(\alpha))[x]$

$$f(x) = (x + \alpha)(x - \alpha)(x^2 - 3)$$

whereas if we view $f(x) \in (\mathbb{Q}(\beta))[x]$

$$f(x) = (x^2 - 2)(x + \beta)(x - \beta).$$

lSince $\mathbb{Q}(\alpha) = \mathbb{Q}[x]/\langle x^2 - 2\rangle$ *is a field we can form* $\mathbb{Q}(\alpha)[t]$ *and study the quotient by* $\langle t^2 - 3\rangle$

$$\mathbb{Q}(\alpha)[t]/\langle t^2 - 3\rangle$$

in this extension field we obtain $\beta = t + \langle t^2 - 3\rangle$ *in the sense that* $\beta^2 - 3 = 0$ *and viewing* $f(x) \in \mathbb{Q}(\alpha)[t]/\langle t^2 - 3\rangle$ *we find*

$$f(t) = (t + \alpha)(t - \alpha)(t + \beta)(t - \beta)$$

The notation $\mathbb{Q}(\alpha, \beta)$ *is also used to denote the smallest extension field of* $\mathbb{Q}$ *which contains* α, β.

Definition 3.8.10. *Let* E *be an extension field of a field* F. *An element* $\alpha \in E$ *is called* **algebraic over** F *if there exists a nonzero polynomial* $f(x) \in F[x]$ *for which* $f(\alpha) = 0$. *An element* $\alpha \in E$ *which is not algebraic is called* **transcendental over** F.

At times I forget to mention the base field over which a given element is algebraic or transcendental. This is not wise because a given number is both algebraic and transcendental depending on context.

Example 3.8.11. *Let $\alpha \in F$ a field then $f(x) = x - \alpha$ has $f(\alpha) = 0$ thus viewing $F = E$ we find $\alpha \in F$ is algebraic over F. This means π is algebraic over $\mathbb{R}$, i is algebraic over $\mathbb{C}$ etc... the examples are endless here.*

Example 3.8.12. *$\alpha = i \in \mathbb{C}$ is algebraic over $\mathbb{R}$ since $f(x) = x^2 + 1$ has $f(i) = i^2 + 1 = 0$.*

Example 3.8.13. *$\sqrt{2}$ is algebraic over $\mathbb{Q}(\sqrt{3})$ since $x^2 - 2 \in \mathbb{Q}(\sqrt{3})[x]$ has $\sqrt{2}$ as a zero in the extension field $\mathbb{Q}(\sqrt{2}, \sqrt{3})$.*

Example 3.8.14. *In fact if $d \in \mathbb{Z}$ then $\sqrt{d}$ is algebraic over $\mathbb{Q}$ since $x^2 - d \in \mathbb{Q}[x]$ takes $\sqrt{d}$ as a zero. Here $\mathbb{Q}(\sqrt{d})$ might be a real or complex extension field. Or, in the case $d = n^2$ we have $\sqrt{d} = n \in \mathbb{Q}$ so $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}$.*

Example 3.8.15. *The real number $\alpha = 2 + \sqrt{1 + \sqrt{3}}$ is algebraic over $\mathbb{Q}$ since*

$$(\alpha - 2)^2 = 1 + \sqrt{3} \quad \Rightarrow \quad (\alpha - 2)^2 - 1 = \sqrt{3} \quad \Rightarrow \quad [(\alpha - 2)^2 - 1]^2 = 3$$

thus α is a zero of the rational polynomial

$$f(x) = [(x - 2)^2 - 1]^2 - 3 = [x^2 - 4x + 3]^2 - 3 = x^4 - 8x^3 + 22x^2 - 24x + 6.$$

Given a number constructed from a finite sequence of arithmetic operations such as addition, subtraction, multiplication and positive roots will be an algebraic number since we can play the game we play here to systematically remove the radicals by successive squaring or cubing etc.

Proving the assertion of the next example would require significant effort on our part. However, there are exercises in some calculus II texts to provide a good part of the proof. See Salas, Hille, and Eitgen's text for instance.

Example 3.8.16. *$\pi, e \in \mathbb{R}$ are algebraic over $\mathbb{R}$ as $x - \pi$ and $x - e$ naturally take π and e as zeros. However, there do not exist $f(x) \in \mathbb{Q}[x]$ for which $f(\pi) = 0$ or $f(e) = 0$ thus π and e are transcendental over $\mathbb{Q}$.*

The common venacular for discussing number systems in number theory is given below.

Definition 3.8.17. *If $x \in \mathbb{C}$ is algebraic over $\mathbb{Q}$ then we say x is an* **algebraic number**. *An $x \in \mathbb{C}$ which is not algebraic over $\mathbb{Q}$ is known as a* **transcendental** *number.*

The following Theorem is helpful towards understanding the structure of transcendental numbers and how they behave in extension fields:

Theorem 3.8.18. *Let E be an extension field of F and $\alpha \in E$. Then α is transcendental over F if and only if ϕ_α gives an isomorphism of $F[x]$ with a subdomain of E. In particular, α is transcendental if and only if ϕ_α is injective.*

Proof: recall $\phi_\alpha : F[x] \to E$ is the evaluation homomorphism of $F[x]$ in the extension field E of F where we often denote $\phi_\alpha(f(x)) = f(\alpha)$. In particular, we define

$$\phi_\alpha(a_n x^n + \cdots + a_1 x + a_0) = a_n \alpha^n + \cdots + a_1 \alpha + a_0$$

We know ϕ_α is a ring homomorphism. We know ϕ_α is an injection if and only if

$$\mathrm{Ker}(\phi_\alpha) = \{f(x) \in F[x] \mid f(\alpha) = 0\}$$

In other words, ϕ_α is injective if and only if there is no polynomial $f(x) \in F[x]$ for which $f(\alpha) = 0$. Thus, ϕ_α is injective iff α is transcendental over F. By the first isomorphism theorem of rings, $F[x]/\mathrm{Ker}(\phi_\alpha) \approx \phi_\alpha(F[x])$ which provides $F[x] \approx \phi_\alpha(F[x])$. $\square$

The image of $F[x]$ under ϕ_α is not a field in the case that α is transcendental. Essentially, it just gives us polynomials in the transcendental.

Example 3.8.19. *Consider π as transcendental over $\mathbb{Q}$. If we denote the smallest field which contains π and $\mathbb{Q}$ by $\mathbb{Q}(\pi)$ then it is not the case that $\mathbb{Q}(\pi) = \phi_\pi(\mathbb{Q}[x])$. We could write $\phi_\pi(\mathbb{Q}[x]) = \mathbb{Q}[\pi]$ as*

$$\phi_\pi(\mathbb{Q}[x]) = \{a_0 + a_1\pi + a_2\pi^2 + \cdots + a_n\pi^n \mid a_0, \ldots, a_n \in \mathbb{Q},\ n \in \mathbb{N} \cup \{0\}\}.$$

In short, a transcendental number over $\mathbb{Q}$ behaves as an indeterminant. Incidentally, $\mathbb{Q}(\pi)$ is formed by the field of fractions of $\mathbb{Q}[\pi]$. You could think of $\mathbb{Q}(\pi)$ as rational functions in the variable π. Generically, for a field F the integral domain $F[x]$ is contained within the field of fractions $F(x)$ which is naturally associated with rational functions over F.

There is more to say about transcendental numbers, but, I think that's all we have for our current discussion. Let us return to the study of algebraic numbers. We've seen there is a natural interplay between the number α and the polynomial which takes α as its zero. It is useful to develop some notation to select a particular, most simple, polynomial corresponding to a given α. Consider:

Theorem 3.8.20. *Let E be an extension field of F with $\alpha \in E$ such that α is algebraic over F. Then there exists an irreducible polynomial $p(x) \in F[x]$ for which $p(\alpha) = 0$. Moreover, $p(x)$ is unique up to a multiplicative constant in F polynomial of least degree for which $p(\alpha) = 0$. Furthermore, if $f(\alpha) = 0$ for $f(x) \in F[x]$ and $f(x) \neq 0$ then $p(x)$ divides $f(x)$.*

Proof: suppose F is a field with extension field E and $\alpha \in E$ is algebraic over F. As usual, we use the evaluation homomorphism to define $\phi_\alpha(f(x)) = f(\alpha)$. Note, $\mathrm{Ker}(\phi_\alpha)$ is an ideal of the $F[x]$ hence, as we know $F[x]$ is a principal ideal domain, there exists $p(x) \in F[x]$ for which $\mathrm{Ker}(\phi_\alpha) = \langle p(x) \rangle$. By definition, $p(\alpha) = 0$. If $f(x) \in F[x]$ with $f(\alpha) = 0$ then $f(x) \in \mathrm{Ker}(\phi_\alpha)$ hence $f(x) = g(x)p(x)$ for some $g(x) \in F[x]$. Observe $p(x)$ is a polynomial of least degree which takes α as a zero and $p(x) \mid f(x)$.

To see $p(x)$ is irreducible, suppose otherwise; that is suppose $p(x) = g(x)h(x)$ with $g(x), h(x)$ noncontant. Hence $p(\alpha) = g(\alpha)h(\alpha) = 0$. It follows $g(x)$ has $g(\alpha) = 0$, but this contradicts our observation that $p(x)$ is a polynomial of least degree for which α is a zero. We find $p(x)$ is irreducible.

Uniqueness of $p(x)$? If $\langle p(x) \rangle = \langle q(x) \rangle$ then as $F[x]$ is an integral domain we know $p(x)$ and $q(x)$ are associaties (see Theorem 3.6.9) hence as the units of $F[x]$ are just nonzero constant polynomials we find $q(x) = cp(x)$ for some $c \in F^\times$. $\square$

Recall a polynomial is **monic** if it has a leading coefficient of 1. For example, $2x^2 + 1$ is not monic whereas $x^4 + 2x + 3$ is monic.

Definition 3.8.21. *Let E be an extension field of F and suppose $\alpha \in E$ is algebraic over F. The unique monic polynomial $p(x) \in F[x]$ of least degree for which $p(\alpha) = 0$ is known as the* **irreducible polynomial for α over F.** *Define $irr(\alpha, F) = p(x)$ and $deg(\alpha, F) = deg(irr(\alpha, F))$ is the* **degree of α over F.**

Example 3.8.22. *Note $\sqrt{2} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ and $irr(\sqrt{2}, \mathbb{Q}) = x^2 - 2$ hence $deg(\sqrt{2}, \mathbb{Q}) = 2$.*

Example 3.8.23. *Note $i \in \mathbb{C}$ is algebraic over $\mathbb{R}$ and $irr(i, \mathbb{R}) = x^2 + 1$ hence $deg(i, \mathbb{R}) = 2$.*

Example 3.8.24. *Observe number $\sqrt[5]{2}$ is a zero of $x^5 - 2$. Moreover, $x^5 - 2$ is irreducible by Eisenstein's Criterion with $p = 2$. Hence $irr(\sqrt[5]{2}, \mathbb{Q}) = x^5 - 2$ and we note $\sqrt[5]{2}$ has degree 5 over $\mathbb{Q}$*

Definition 3.8.25. *A field E is an* **simple extension** *of F if E is an extension field of F for which there exists $\alpha \in E$ with $F(\alpha) = E$. We define $F(\alpha)$ is the smallest field which contains F and α. Likewise, $F(\alpha_1, \ldots, \alpha_n)$ is the smallest field which contains F and $\alpha_1, \ldots, \alpha_n$.*

Operationally, we could define $F(\alpha)$ as the intersection of all fields which contain F and α. We say $F(\alpha)$ is the field F **with** α **adjoined**. Or, $F(\alpha_1, \ldots, \alpha_n)$ is F with $\alpha_1, \ldots, \alpha_n$ **adjoined**. I hope you can forgive me for using some of this language without formally defining it earlier. Better late than never I think.

Theorem 3.8.26. *Let E be a simple extension $F(\alpha)$ of a field F with α algebraic over F and $deg(\alpha, F) = n \geq 1$. Then for each $\beta \in E$ there exist unique $b_0, b_1, \ldots, b_{n-1} \in F$ for which:*

$$\beta = b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_{n-1}\alpha^{n-1}.$$

Proof: suppose $E = F(\alpha)$ where $\text{irr}(\alpha, F) = p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$. By construction, $p(\alpha) = 0$ hence

$$\alpha^n = -a_{n-1}\alpha^{n-1} - a_{n-2}\alpha^{n-2} - \cdots - a_1\alpha - a_0 = -\sum_{k=0}^{n-1} a_k\alpha^k \quad (\star).$$

The smallest field containing[15] α and F is given by the quotient $F[x]/\langle p(x)\rangle$ where we identify $\alpha = x + \langle p(x)\rangle$. Hence $E = F[x]/\langle p(x)\rangle$ has arbitrary elements of the form $b_0 + b_1\alpha + \cdots + b_m\alpha^m$ for $m \in \mathbb{N}$. We use $\star$ to reduce any expression with $m \geq n$ as follows: first, $\star$ shows how to reduce $m = n$. Suppose inductively there exist $c_j \in F$ for which $\alpha^m = \sum_{j=0}^{n-1} c_j\alpha^j$ for some $m \geq n$. Consider,

$$\alpha^{m+1} = \alpha\alpha^m = \alpha\sum_{j=0}^{n-1} c_j\alpha^j = c_{n-1}\alpha^n + \sum_{j=0}^{n-2} c_j\alpha^{j+1} = -c_{n-1}\sum_{k=0}^{n-1} a_k\alpha^k + \sum_{k=1}^{n-1} c_{k-1}\alpha^k$$

Thus, $\alpha^{m+1} = -c_{n-1}a_0 + \sum_{k=1}^{n-1}(c_{k-1} - c_{n-1}a_k)\alpha^k$ which verifies the induction step. Next, we consider uniqueness of the expansion. Suppose there exist $b_0, \ldots, b_{n-1} \in F$ and $b'_0, \ldots, b'_{n-1} \in F$ for which

$$\beta = b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_{n-1}\alpha^{n-1} = b'_0 + b'_1\alpha + b'_2\alpha^2 + \cdots + b'_{n-1}\alpha^{n-1}.$$

Define $g(x) = (b'_0 - b_0) + (b'_1 - b_1)x + \cdots + (b'_{n-1} - b_{n-1})x^{n-1}$ and notice by construction of b_j, b'_j we have $g(\alpha) = 0$. Yet, $\deg(g(x)) = n - 1$ and so we find $g(x) = 0$ as $p(x)$ is the polynomial of smallest positive degree for which $p(\alpha) = 0$. Note $g(x) = 0$ only if all its coefficients are zero hence $b'_j - b_j = 0$ for $j = 0, \ldots, n-1$ which proves $b'_0 = b_0, b'_1 = b_1, \ldots, b'_{n-1} = b_{n-1}$ hence the representation of β in terms of the F-linear combination of $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is unique. □

The Corollary below follows immediately from the proof above since we know linear independence of a set is equivalent to the equating coefficients property of a set of vectors.

[15] Fraleigh avoids this point by simply defining $F(\alpha)$ to be the quotient of $F[x]/\langle p(x)\rangle$, see Case I on page 270

Corollary 3.8.27. *Let E be a simple extension $F(\alpha)$ of a field F with α algebraic over F and $deg(\alpha, F) = n \geq 1$. Then $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is a basis for $F(\alpha)$ as a vector space over F. Furthermore, $F(\alpha)$ is a vector space of dimension n over F; $dim(F(\alpha)) = deg(\alpha, F)$.*

Example 3.8.28. *The complex numbers $\mathbb{C}$ can be seen as a simple extension of $\mathbb{R}$ by i. Note $\mathbb{R}(i) = \mathbb{C}$ has $irr(i, \mathbb{R}) = x^2 + 1$ and in fact $\mathbb{C}$ is a vector space of dimension 2 over $\mathbb{R}$ with basis $\{1, i\}$.*

Our focus in this Lecture and the next is primarily on fields and their extensions. However, I must say, many of the ideas we study here are available for application outside the context of polynomials with coefficients in a field. Also, quotients by reducible polynomials can be interesting. For example, $\mathbb{R}[x]/\langle x^2 - 1\rangle$ forms a set with $j = x + \langle x^2 - 1\rangle$ satisfying the property $j^2 = 1$. Numbers of the form $a + bj$ are known as **hyperbolic numbers**. Hyperbolic numbers are a little tricky since $(1 + j)(1 - j) = 1 - j^2 = 0$ yet $1 \pm j \neq 0$. Perhaps we'll study algebra constructions further once we have completed our study of field extensions.

Problems for Lecture 30: (these are collected 11-30-16)

Problem 127: Fraleigh page 272, number 5

Problem 128: Fraleigh page 272, number 7

Problem 129: Fraleigh page 272, number 12

Problem 130: Fraleigh page 272, number 16

Problem 131: Fraleigh page 273, number 23

Problem 132: Fraleigh page 273, number 25

3.9 Lecture 31: algebraic extensions

Here we follow Section 31 of Fraleigh's *Abstract Algebra*.

Definition 3.9.1. *An extension E of F is called an* **algebraic** *extension of F if every element of E is algebraic over F. If E is not an algebraic extension of F then it is called a* **transcendental** *extension of F.*

We used the notation $[G : H]$ to represent the number of H-cosets in G where G was a finite group and $H \leq G$. The notation introduced below shouldn't cause confusion as the meaning should be clear from the context.

Definition 3.9.2. *If an extension field E of a field F forms a vector space of finite dimension n over F then we say E is a* **finite extension of degree** *n and we write $[E : F] = n$*

Suppose $[E : F] = 1$ then we can argue $E = F$. Since F contains 1 it follows this forms a basis for E hence $F = E = \text{span}(1)$. In other words, if we have a vector space E over F which contains a copy of F and the vector space E has dimension 1 then $E = F$. Conversely, if $E = F$ then clearly $[E : F] = 1$. What follows is more interesting:

Theorem 3.9.3. *A finite extension field E over a field F is an algebraic extension of F.*

Proof: let $\alpha \in E$ where $\dim(E) = n$ over F. Observe the set $S = \{1, \alpha, \alpha^2, \ldots, \alpha^n\}$ has $(n+1)$-vectors in E. Therefore, S is a linearly dependent subset of E. It follows there exist $c_0, c_1, \ldots, c_{n-1} \in F$ (not all zero) for which

$$c_0 + c_1\alpha + c_2\alpha^2 + \cdots + c_n\alpha^n = 0.$$

Thus α is algebraic over F as the nonzero $f(x) = c_0 + c_1x + \cdots + c_nx^n \in F[x]$ has $f(\alpha) = 0$. □

Theorem 3.9.4. *If E is a finite extension field of a field F and K is a finite extension field of a field F then $[K : F] = [K : E][E : F]$.*

Proof: suppose $E = \text{span}_F\{\alpha_1, \ldots, \alpha_n\}$ and $K = \text{span}_E\{\beta_1, \ldots, \beta_m\}$ where $\{\alpha_i\}$ and $\{\beta_j\}$ form basis for E and K respective. We propose $S = \{\alpha_i\beta_j \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ forms a basis for K as a vector space over F. Let $\gamma \in K = \text{span}_E\{\beta_1, \ldots, \beta_m\}$ then there exist $b_j \in E$ for which $\gamma = \sum_{j=1}^m b_j\beta_j$. But, for each $j = 1, \ldots, m$ we have $b_j \in E = \text{span}_F\{\alpha_1, \ldots, \alpha_n\}$ hence there exist $c_{ij} \in F$ for which $b_j = \sum_{i=1}^n c_{ij}\alpha_i$. Substituting,

$$\gamma = \sum_{j=1}^m b_j\beta_j = \sum_{j=1}^m \left(\sum_{i=1}^n c_{ij}\alpha_i \right) \beta_j = \sum_{i=1}^n \sum_{j=1}^m c_{ij}\alpha_i\beta_j$$

thus $\gamma \in \text{span}_F(S)$ and it follows $K = \text{span}(S)$. Linear independence of S over F follows naturally from the linear independence of the bases $\{\alpha_i\}$ and $\{\beta_j\}$. In detail: if

$$\sum_{i=1}^n \sum_{j=1}^m c_{ij}\alpha_i\beta_j = 0$$

then

$$\sum_{i=1}^n \left(\sum_{j=1}^m c_{ij}\alpha_i \right) \beta_j = 0 \quad \Rightarrow \quad \sum_{j=1}^m c_{ij}\alpha_i = 0$$

for each $j = 1, \ldots, m$ by linear independence of $\{\beta_1, \ldots, \beta_m\}$. Then, for each $j = 1, \ldots, m$ we may argue $\sum_{j=1}^{m} c_{ij}\alpha_i = 0$ implies $c_{ij} = 0$ for each $i = 1, \ldots, n$ by linear independence of $\{\alpha_1, \ldots, \alpha_n\}$. In summary, $c_{ij} = 0$ for all i, j possible and we have established the linear independence of S. Note, $\dim_F(K) = mn$ whereas $\dim_F(E) = n$ and $\dim_E(K) = m$ in summary,

$$[K : F] = \dim_F(K) = mn = \dim_E(K)\dim_F(E) = [K : E][E : F]. \qquad \square$$

Induction naturally extends the Theorem above to multiple extensions:

Corollary 3.9.5. *If F_i is a field and F_{i+1} is a finite extension of F_i for each $i = 1, \ldots, r$ then F_r is a finite extension of F_1 where $[F_r : F_1] = [F_r : F_{r-1}][F_{r-1} : F_{r-2}] \cdots [F_2 : F_1]$.*

The next Corollary is a useful tool. It plays an important role in the study of contructible numbers. Roughly, numbers which are constructible follow from quadratic extensions. If a particular number amounts to a degree three extension then the fact that $2 \nmid 3$ forbids the construction of that particular number. For example, the trisection of an angle by compass-straight-edge operations.

Corollary 3.9.6. *If E is an extension field of F and $\alpha \in E$ is algebraic over F and $\beta \in F(\alpha)$ then $deg(\beta, F)$ divides $deg(\alpha, F)$.*

Proof: suppose E is an extension field of F and $\alpha \in E$ is algebraic. If $\beta \in F(\alpha)$ then observe $F \leq F(\beta) \leq F(\alpha)$. Thus, by Theorem 3.9.4,

$$[F(\alpha) : F(\beta)][F(\beta) : F] = [F(\alpha) : F] \quad \Rightarrow \quad [F(\alpha) : F(\beta)]\deg(\beta, F) = \deg(\alpha, F).$$

Thus, $\deg(\beta, F) \mid \deg(\alpha, F)$ and the Corollary follows. $\square$

Example 3.9.7. *Suppose $Q(\sqrt{2})$ has a zero β for $x^3 - 2$. We have $deg(\beta, \mathbb{Q}) = 3$ as $x^3 - 2$ is irreducible by Eisenstein's Criterion with $p = 2$. If $\beta \in \mathbb{Q}(\sqrt{2})$ then by the Corollary 3.9.9, $3 \mid 2$. Thus, no zero to $x^3 - 2$ exists within $\mathbb{Q}(\sqrt{2})$.*

Example 3.9.8. *Consider $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ then*

$$(\mathbb{Q}(\sqrt{2}))(\sqrt[3]{2}) = \{c_1 + c_2 2^{1/3} + c_3 2^{2/3} \mid c_1, c_2, c_3 \in \mathbb{Q}(\sqrt{2})\}$$

Observe $c_1 + c_2 2^{1/3} + c_3 2^{2/3} \in (\mathbb{Q}(\sqrt{2}))(\sqrt[3]{2})$ can be expressed in terms of rational coefficients $a_1, b_1, a_2, b_2, a_3, b_3 \in \mathbb{Q}$ as follows:

$$\begin{aligned} c_1 + c_2 2^{1/3} + c_3 2^{2/3} &= (a_1 + b_1\sqrt{2}) + (a_2 + b_2\sqrt{2})2^{1/3} + (a_3 + b_3\sqrt{2})2^{2/3} \\ &= a_1 + b_1 2^{3/6} + a_2 2^{2/6} + b_2 a^{5/6} + a_3 2^{4/6} + 2b_3 2^{1/6}. \end{aligned}$$

Thus $(\mathbb{Q}(\sqrt{2}))(\sqrt[3]{2}) = \mathbb{Q}(2^{1/6})$. In other words, $\mathbb{Q}(2^{1/2}, 2^{1/3}) = \mathbb{Q}(2^{1/6})$.

Theorem 3.9.9. *Let E be an algebraic extension field of F. Then there exist a finite number of elements $\alpha_1, \alpha_2, \ldots, \alpha_n \in E$ such that $E = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$ iff E is a finite extension of F*

Proof: see page 286 of Fraleigh. $\square$.

The **primitive element** in Steinitz' Theorem below is the element c. Generally, if an extension field $E = F(c)$ then we say c is a **primitive** element of E.

Theorem 3.9.10. Primitive Element Theorem (Steinitz, 1910) *If F is a field with $Char(F) = 0$ and a, b are algebraic over F then there exists $c \in F(a, b)$ such that $F(a, b) = F(c)$.*

Proof: see page 367 of Gallian. □

In the case $\text{Char}(F) = p$ then we still the more complicated result given by Theorem 3.9.9.

Problems for Lecture 31: (these are collected 11-30-16)

Problem 133: Fraleigh page 291, number 3

Problem 134: Fraleigh page 291, number 6

Problem 135: Fraleigh page 291, number 7

Problem 136: Fraleigh page 291, number 9

Problem 137: Fraleigh page 291, number 10

Problem 138: Fraleigh page 292, number 29

3.10 Lecture 32: algebraically closed fields

I intend to discuss the end of Section 31 in Fraleigh and tie up any loose ends from the previous Lectures on extension fields. I'll probably follow pages 286-291 of Fraleigh fairly closely, omitting some details as appropriate due to time constraints.