COMPUTATION OF SCHUBERT GALOIS GROUPS

A Dissertation

by

CHRISTOPHER JAMES BOTT

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Frank Sottile
Committee Members,	Maurice Rojas
	Patricia Klein
	Matthew Call
Head of Department,	Peter Howard

August 2025

Major Subject: Mathematics

ABSTRACT

Many know and love Galois groups, which help us understand the solvability of equations and the structure of field extensions. What many do not know is that in the first comprehensive treatise on Galois Theory, Jordan also considered Galois groups in the context of enumerative geometry. Enumerative geometry concerns counting the number of geometric objects that satisfy some geometric conditions, such as how many lines are on a nonsingular cubic surface, or how many plane conics are tangent to five general plane conics. For such problems, Galois groups describe the symmetries of the solution set. For “interesting" examples, the Galois groups are called enriched, because they additionally encode geometric structure inherent to the problem. We describe and build upon the current state of classifying and computing enriched Galois groups in Schubert calculus, a branch of enumerative geometry where all geometric objects and conditions involved are restricted to be linear spaces.

We develop the theory from first principles, starting from solving equations one might encounter in high school, and moving through solving more difficult equations and systems of equations, naturally developing algebraic structures along the way. We review the fundamentals of algebraic geometry, and define Galois groups in enumerative geometry as monodromy groups of branched covers.

With proper tools in hand, we describe the Schubert calculus framework. We explain how to formulate Schubert problems in Grassmannians and partial flag varieties of various types. We share the state of the art for classifying and computing Schubert Galois groups, which involves the work of Derksen, Vakil, and Sottile (with collaborators), as well as give current conjectures.

Finally, we develop a Macaulay2 package, which computes not just the number of solutions to a Schubert problem, but also the ideal defining the system of polynomial equations in local coordinates. Such ideals can then be used to further study Schubert problems further (including arithmetic and reality questions). We implement the Frobenius algorithm, which we then use to investigate Schubert Galois groups in many settings.

DEDICATION

To my wonderful wife, Noelle.

ACKNOWLEDGMENTS

I would like to take this special opportunity to thank many who have contributed in deep and meaningful ways to helping me improve myself and my future, both personally and professionally, while I have been in this PhD program at Texas A&M University. You have all made my life so much richer, and I cannot thank you enough for that!

First and foremost, I must thank my Heavenly Father for His infinite love and mercy. All of my successes in life I credit to Him. He has given me life, a great desire to learn and teach mathematics, formal opportunities to learn and grow, and my family, friends, and mentors who I could not have done this without. So praise be to God for this accomplishment in my life! May I use this PhD to better serve Him and His children.

I thank my dear wife, Noelle, who was willing to move from Utah to Texas, away from family, to go on this incredible adventure with me! Who has been my constant support, my best friend, my confidant, and partner. Who worked so hard to provide for our family day after day to ensure our future financial well-being. I cannot thank you enough Noelle, for your incredible sacrifice, so that I could fulfill my dream of being a professor. We did it! I love you, and I will always love you.

I thank my two children, Eleanor (6) and Soren (3), for their love and support as well! Eleanor was six weeks old when we first moved to College Station, and has brought such a light into our lives. We love her enthusiasm and kindness to others, and it is fun to see her budding interest in math as well! Soren came halfway through my PhD program, and I am so glad that we did! He makes us laugh when we are stressed, and is growing into such an active and good boy. These two have taught me how to love and teach at a deeper level, and have revealed to me many of my personal flaws that I have been able to work on during my program. I acknowledge that learning how to parent while doing my PhD was difficult and may have slowed things down, but I wouldn't change a thing. As much as it is my dream to teach mathematics as a career, being a father has always been a greater dream and reality.

I thank my parents, Dale and Anita Bott, who raised me to have a love of learning. My father

is a family pioneer, the first person on his side to go to college and receive a Bachelor's degree, which has been a great example to me. My mother passed away in 2015 from brain cancer at age 45, but was an avid learner and teacher during her lifetime. I was born about a month before she and my father got their Bachelor's degrees, and she stayed home after that to be with me and my soon-to-come three younger siblings. I was taught at home to love and explore all subjects. My mom never quite understood why I was so obsessed with math while she was passionate about English and reading, but she always pushed me to pursue higher education with a passion. One of my last memories of my mom was telling her that I would soon be graduating with my Bachelor's degree, to which in a confused state she said, "wait - but I want you to do more than that". She had just recently received her Master's degree in Education, taking one online class at a time while undergoing chemotherapy, and I knew she wanted me to "go all the way" and get a PhD. Most importantly, she knew that that is what I wanted, and I reassured her of that. And now here I am. Thank you Mom! For your example, for your love, and for pushing me to reach my potential.

I thank my parents-in-law, Matt and Julie Reid, who took a chance on me joining their family, but who have always welcomed me with open arms. I am so grateful for your unwavering support! I thank Noelle's sister, Bonnie, who watched Eleanor until she was two years old while I went to class, and thank Bonnie's husband (and one of my best friends), James Scott, for all our conversations while he was a graduate student here at A&M at the same time! Those two years together were so special! I thank my brother, Tyler, and his wife, Emma, who lived near us in Bryan for two years as well. Having family support while being away from home was priceless, and I'm grateful for the time we had together!

I thank my other family members who have always supported me in this journey. To my siblings, siblings-in-law, my grandparents, uncles, aunts, nieces, nephews, and cousins. We are a close family, and I am grateful for each of you! I thank my new mom, Mari Bott, who my dad married in 2016, and who has been such a strength to him and the rest of us. We love you! I thank Mari's five children, and their families, for their support and love as well.

I thank so many friends who have watched our kids and given emotional support over the

years. I simply cannot name everybody, but thank you to Kennedy and Kelsey Holloway, Tyler and Toria Atkinson, Jon and Jeniece Bennett, Omar and Paige Aranda, Sam and Angela Briggs, Dan and Jensen Mosman, Alec and Hannah Judd, Ian and Kaylee Scadden, and Jacob and Elizabeth Williams. You are amazing!

I thank my church family who have been our home away from home. Thank you to my leaders and those I have served with there - Nate Sharp, Nate Roeth, Aaron Hull, Todd Graham, Matt Call, Austin Rogers, David Klabunde, Omar Aranda, James Scott, Chris Sorensen, Tim Smart, Brad Brimley, Tim Lutz, Michael Johanson, Bryan Wilson, McKay Osborne, Brennan Young, Ian Lignell, Troy Kema, Andrew Carruth, Ryan Larkin, John Moeller, Dan Searcy, Matt Bludorn, Charley and Suzanne Todd, Tommy Stuart, Holly Rasmussen, and Shawn Larkin. A special thank you to Dallin Smith, who for over three years now has taught early morning Seminary with me, and has been a great example to me while going through his own PhD program at A&M. And thank you to our high school students who consistently woke up and spent hours learning with us how to be more like Jesus Christ, and how to better spread His light to the world! Those mornings studying the scriptures with incredible youth have changed my life forever, and have been a highlight of my time in Bryan/College Station!

I thank my many teachers over the years, especially my high school calculus teacher, Christy Hall, in whose class my obsession with mathematics began.

I thank my mentors and teachers from Brigham Young University, where I studied math for six years before my time at A&M, and where I learned that math went so much farther than calculus - Tyler Jarvis and his algebraic geometry group, Rachel Webb, Amanda Francis, Darrin Doud, Stephen Humphries, Lennard Bakker, Emily Evans, Jeff Humpherys, Denise Halverson, Mark Hughes, Curtis Kent, Gary Lawlor, Shue-Sum Chow, Rodney Forcade, Jared Whitehead, and Lonette Stoddard. A special thank you to my Masters Thesis advisor, Nathan Priddis, who prepared me well for my PhD program, and with whom I published for the first time. I am so excited and grateful to be joining the math faculty at BYU in the Fall, and look forward to helping other students “enter to learn; go forth to serve"!

I thank all of my teachers and mentors at A&M who took my graduate learning to the next level - Frank Sottile, J.M. Landsberg, Greg Pearlstein, Maurice Rojas, Tricia Klein, Laura Matusevich, and Eric Rowell.

I thank those who have helped me improve my teaching at A&M, starting with Peter Howard and David Manuel who took a chance on me and let me be instructor of record my first year in the program. I thank my course coordinators - Janice Epstein, Bradford Garcia, and Angie Allen. A special thank you to John Weeks, who not only was a course coordinator and graduate TA coordinator, but my mentor for the Academy for Future Faculty program. John has observed my teaching and given helpful feedback, coordinated teaching training that was incredibly impactful, and has critiqued all of my job application materials. Thank you John!

I thank those involved with the Masters of Quantitative Finance program at A&M, and for allowing me to teach the corresponding math boot camp - Andrea Bonito, Maurice Rojas, Igor Zelenko, Rebecca (Itz) Calloway, Donna McClure, and Misty Page. And thank you to the amazing MSQF students who inspire me and who have made me a better teacher!

I thank my fellow graduate students at A&M for their friendship and discussions, especially those in Frank's research group over the years - Taylor Brysiewicz, Elise Walker, Thomas Yahl, Matt Faust, Jordy Lopez-Garcia, Jonah Robinson, Nate Welty, Kelly Maluccio, Somak Dutta, Ruzho Sagarayj, Daniel Dale, and John Ajayi.

Last, but certainly not least, I thank my incredible advisor, Frank Sottile. Frank enthusiastically welcomed me into his research group during a very difficult time in my life, and I will be forever grateful for that. He genuinely cares about mentoring, and has inspired me with his strong work-ethic, love of mathematics, and generous nature. Frank has spent countless hours teaching me about algebraic geometry, computation, and the nature of the academy. He has pushed me to present my work at many conferences, connect with as many mathematicians as possible, and provides settings for mathematicians at all levels to learn from one another. Frank has welcomed my family into his home, and my children love him. Thank you Frank for all our discussions, your faith in me and my future, and for the time you have sacrificed on my behalf.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Frank Sottile [advisor], Professor Maurice Rojas, and Professor Patricia Klein of the Department of Mathematics and Professor Matthew Call of the Department of Management. The material in Sections 3 and 4 is joint work with Frank Sottile.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a graduate fellowship from Texas A&M University. No other outside source of funding was provided.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. GALOIS GROUPS IN ENUMERATIVE GEOMETRY

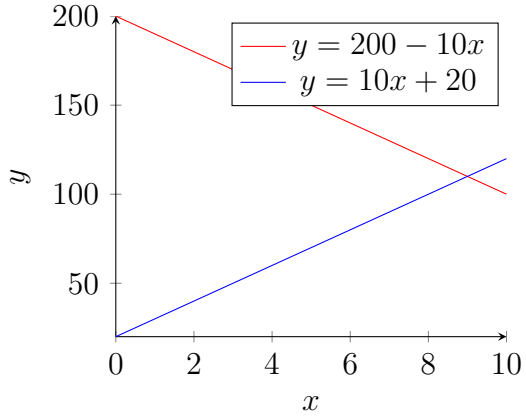
1.1 Equations and Algebraic Structures

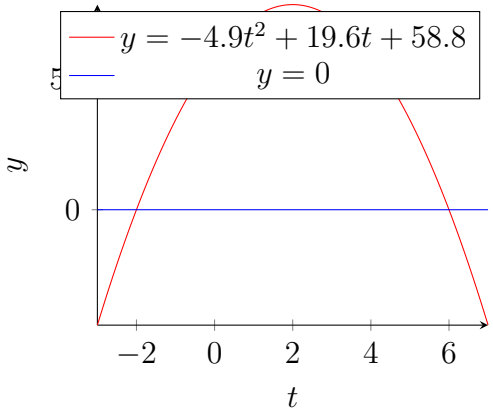
Fundamental to mathematics is the notion of an equation. For example, students in school consistently practice solving equations such as $10x+20 = 200-10x$, or $-4.9t^2+19.6t+58.8 = 0$.

The equation $10x + 20 = 200 - 10x$ is like a balanced scale, where the strategy is to perform a sequence of arithmetic operations to both sides of the equation, at each step resulting in a new equation that is equivalent to the one before. Adding $10x$ to both sides results in $20x + 20 = 200$, and then subtracting 20 from both sides yields $20x = 180$. Finally, dividing both sides by 20 gives the final answer of $x = 9$. Students get incredibly efficient at employing this algorithm!

On the other hand, to solve $-4.9t^2 + 19.6t + 58.8 = 0$, one must use a different strategy. One could use the quadratic formula $t = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$ with $a = -4.9, b = 19.6,$ and $c = 58.8$, obtaining the correct final answers of $t = -2$ and $t = 6$. Alternatively, one could divide both sides by -4.9 to get the equivalent, but simpler, equation $t^2 - 4t - 12 = 0$, which can be solved by factoring to get $(t + 2)(t - 6) = 0$, with the same solutions $t = -2$ and $t = 6$ as before.

These equations can also be interpreted geometrically. From this perspective, solving the equation $10x + 20 = 200 - 10x$ is the same as asking for the x-coordinate of the point where the two lines with equations $y = 10x + 20$ and $y = 200 - 10x$ intersect. Similarly, solving the equation $-4.9t^2 + 19.6t + 58.8 = 0$ is analogous to asking where the parabola $y = -4.9t^2 + 19.6t + 58.8$ crosses the x-axis (intersects the line $y = 0$).





While students get incredibly efficient at solving equations, and have some idea as to how to visualize them geometrically, one question often persists - “why?". Why should one even care about equations and solutions in the first place? Why spend so much time and energy in school drilling algorithms when there seem to be more interesting and useful pursuits?

What many do not understand is that equations describe and quantify relationships! For example, consider Eleanor (today age six) and Soren (today age three), whose ages will of course change over time. One could introduce variables to quantify these ages, defining Eleanor's age to be y and Soren's age to be x. Equations can then be used to describe how the ages y and x are related. The equation $y = 100x$ would not relate Eleanor's and Soren's ages since it is not currently true, and $y = 2x$ would also not be a good equation to relate the ages, because even though it is true in the moment, next year when Eleanor is seven and Soren is four, the equation will no longer hold. Instead, the equation $y = x+3$ is much more useful, and quantifies the relationship “Eleanor is three years older than Soren." Defining variables and using equations to relate those variables is called “modeling", and there is a famous statement in mathematics that “all models are wrong, but some are useful". The equation $y = x + 3$ is useful enough, but admittedly if Eleanor and Soren do not have the same birthday, there will be times (shorter than a year) when one age increases before the other one does (i.e. Eleanor turns seven before Soren turns four), invalidating the equation. One could then refine the model to account for a need for greater precision (changing units from years to months or days, for example), or one could be sufficiently pleased with the equation for how one desires to use it. Either way, the equation itself “does not remember" the real world situation that it was chosen to model, math “living in a world of its own". One can possibly use math algorithms to solve an equation, but it is essential for the user to interpret the solutions in the context of the original problem to make sure that the results make sense and are useful. Whether students like it or not, equations are regularly used by many professionals in their careers, because people care about objects/ideas/etc. which are related to one another in quantifiable ways.

One possible application for the equation $10x + 20 = 200 - 10x$ is to economics. Say there is an item being sold in the marketplace, and past market data reveals that $y = 10x + 20$ relates price

(y) to the supply (x) of that item, while $y = 200 - 10x$ relates price (y) to the demand (x) of that item. Solving the equation $10x + 20 = 200 - 10x$ then gives $x = 9$ as the equilibrium quantity, with $y = \$110$ as the equilibrium price for that item. In other words, when 9 items are made and sold for 110 apiece in the market, all the items are sold with no other items desired by consumers, resulting in no waste or lost opportunity cost for the market as a whole.

On the other hand, the equation $y = -4.9t^2 + 19.6t + 58.8 = 0$ could arise in the context of kinematics in physics, like a ball being launched directly upward with initial velocity of $19.6m/s^2$ from an initial height of $58.8m$. Here t is time (in seconds), y is the height of the ball, and the model is ignoring some factors such as wind resistance. In this case, the solution $t = 6$ to $y = 0$ reveals that it would take precisely 6 seconds for the ball to hit the ground under the given model, which one could experimentally verify. The other solution, $t = -2$ is extraneous, since time cannot be negative in the context of the given application.

While applications such as the examples given help describe why equations are useful in some careers, students rarely ask other "why" questions about equations, like why do the algorithms they drill so hard work? Why does performing certain arithmetic operations to both sides of an equation successfully cancel terms, allowing one to get closer to the answer with each step? Why do certain formulas and techniques give a solution? These questions concern the theory of mathematics, and are surprisingly deep. Perhaps surprisingly, abstraction is the key to understanding the answers to these kind of "why" questions.

1.2 Groups

Consider solving the simplest possible type of equation involving addition of real numbers: $a + x = b$. Here x is an unknown real number, while a and b are fixed real numbers (so for example, if $a = 2$ and $b = 1$, then we consider the equation $2 + x = 1$). To solve:

1. Add $-a$ to both sides: $-a + (a + x) = -a + b$
2. Move the parentheses: $(-a + a) + x = -a + x$
3. Cancel the $-a$ and a: $0 + x = -a + b$

4. Remove the 0 to get the final answer: $x = -a + b$

Similarly, consider solving the simplest possible type of equation involving the multiplication of real numbers: $a \cdot x = b$ (where this time, we need $a \neq 0$). Again, notice the process for solving:

1. Multiply $\frac{1}{a}$ to both sides: $\frac{1}{a} \cdot (a \cdot x) = \frac{1}{a} \cdot b$

2. Move the parentheses: $(\frac{1}{a} \cdot a) \cdot x = \frac{1}{a} \cdot b$

3. Cancel the $\frac{1}{a}$ and a: $1 \cdot x = \frac{1}{a} \cdot b$

4. Remove the 1 to get the final answer: $x = \frac{1}{a} \cdot b$

In both cases, we have an equation of the form $a * x = b$, where a, b, and x are elements in some set (real numbers in our first example, and nonzero real numbers in the second example), and where $*$ is some binary operation that combines two elements of the set to give another element of that set (in our examples, $*$ is $+$ and $\cdot$). To solve the equations, we first needed the existence of inverse elements ($-a$ for addition, and $\frac{1}{a}$ for multiplication). Second, we needed to be able to move parentheses, also known as the associative law for addition and multiplication. Third, the whole point of having inverse elements and moving parentheses is to cancel something, but what does that mean? By definition, the inverse of an element should combine with that element to become an identity element, and so we need the existence of such an identity element (0 for addition, and 1 for multiplication). Finally, the identity element by definition should combine with other elements in a way that leaves them alone. For the general setup $a * x = b$, if we denote the inverse of a as a^{-1} and the identity element as e, then the process for solving becomes:

1. Use the binary operation $*$ to combine a^{-1} with both sides: $a^{-1} * (a * x) = a^{-1} * b$

2. Move the parentheses: $(a^{-1} * a) * x = a^{-1} * b$

3. Cancel the a^{-1} and a: $e * x = a^{-1} * b$

4. Remove the e to get the final answer: $x = a^{-1} * b$.

Definition 1.2.1. *A **group** is a set G equipped with a binary operation $* : G \times G \to G$ such that*

1. *$*$ is associative: For all $g_1, g_2, g_3 \in G$, $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$*
2. *G has an identity element e: For all $g \in G$, $e * g = g$*
3. *Every element of G has an inverse: For all $g \in G$, there exists $g^{-1} \in G$ such that $g^{-1} * g = e$.*

Hence, as we have seen above, the set of real numbers $\mathbb{R}$ is a group under addition, and the set of nonzero real numbers $\mathbb{R} \setminus \{0\}$ is a group under multiplication. On the other hand, $\mathbb{R}$ is NOT a group under subtraction, since subtraction is not associative: $3 - (2 - 1) = 2 \neq 0 = (3 - 2) - 1$. Additionally, the natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$ is NOT a group under addition, since while addition is associative and $\mathbb{N}$ has 0 as an additive identity, no element (besides 0) has an additive inverse. Hence, one cannot solve all equations like $a + x = b$ purely in the world of the natural numbers. For example, $2 + x = 1$ has no solution in $\mathbb{N}$. One has to "extend" the world of the natural numbers to a group, like the integers $\mathbb{Z}$ (the minimal such extension) or real numbers, to solve such equations.

Now, not all groups arise in mathematics as sets of numbers. Additionally, the groups above are infinite (as sets, they have infinite cardinality), but from now on, we will primarily be focusing on finite groups. We call the cardinality of a finite group the **order** of that group. Perhaps the most important family of finite groups are the symmetric groups, which have actually have functions as elements. We denote the subset $\{1, \ldots, n\} \subseteq \mathbb{N}$ by $[n]$ throughout.

Definition 1.2.2. *The **symmetric group on n letters** is $S_n = \{$bijections $f : [n] \to [n]\}$, equipped with composition of functions $\circ : S_n \times S_n \to S_n$ as a binary operation.*

A bijection means an invertible function, and since we are considering such a function from $[n]$ to itself, a bijection is equivalent to a permutation of the numbers 1 through n. Thus, the order of S_n is $|S_n| = n!$, which while potentially quite large, is finite. For example, there are 6 permutations of the numbers 1, 2, and 3, so $S_3 = \{123, 132, 213, 231, 312, 321\}$, where here we are using "one-line notation", meaning if $f = 213$, $f(1) = 2$, $f(2) = 1$, and $f(3) = 3$. With this notation (somewhat unconventional, but important for our purposes), $Id_{[3]} = 123$.

Since the composition of bijections is a bijection, $\circ$ is a well defined binary operation on S_n. Furthermore, function composition is associative. The identity function, denoted here by $Id_{[n]}$, is an identity element for S_n, and by definition each bijection f has an inverse function f^{-1} such that $f \circ f^{-1} = f^{-1} \circ f = Id_{[n]}$. Therefore, the symmetric group S_n, really is a group as defined above! Hence, if $x : [n] \to [n]$ were an unknown bijection, but $f, g \in S_n$ were known, one could solve equations of the form $f \circ x = g$, using the fact that S_n is a group.

There are many situations where we want to consider some groups as "the same", even if they are not presented to us in precisely the same way. For example, the group $S_{\{a,b,c\}} = \{\text{bijections } f : \{a, b, c\} \to \{a, b, c\}\}$, equipped with function composition, should be considered "the same" as S_3 above. To make this notion precise, we introduce the notions of group homomorphisms and isomorphisms. Here we introduce the notation $(G, *)$ to mean the group G has $*$ as its binary operation.

Definition 1.2.3. *Let $(G, *)$ and $(H, \diamond)$ be groups. A **group homomorphism** $\varphi : G \to H$ is a function such that for all $g_1, g_2 \in G$, $\varphi(g_1 * g_2) = \varphi(g_1) \diamond \varphi(g_2)$. A **group isomorphism** is a bijective group homomorphism. If such a group isomorphism exists, we say that G and H are **isomorphic** as groups.*

The idea is that a group homomorphism "preserves" the group structure between the two sets. For example $\varphi : S_3 \to S_{\{a,b,c\}}$ given by $\varphi(f)(a) = f(1)$, $\varphi(f)(b) = f(2)$, and $\varphi(f)(c) = f(3)$ is an isomorphism. It is really just a relabeling, replacing every instance of 1 with a, 2 with b, and 3 with c, and so will satisfy the group homomorphism property trivially. Such a relabeling is not unique, but just having one establishes that S_3 and $S_{\{a,b,c\}}$ are isomorphic groups.

For a more interesting example, we again return to geometry! In fact, groups often arise as a set of symmetries (bijections preserving relevant structure) of geometric objects! For example, consider a fixed equilateral triangle in the plane. By a symmetry of the triangle, we mean a rigid motion (isometry) that preserves the triangle. In other words, the points of the triangle may individually move and change position, but the overall triangle shape is preserved. In this case, one can show that the only symmetries that arise are certain reflections and rotations, as shown in the figure

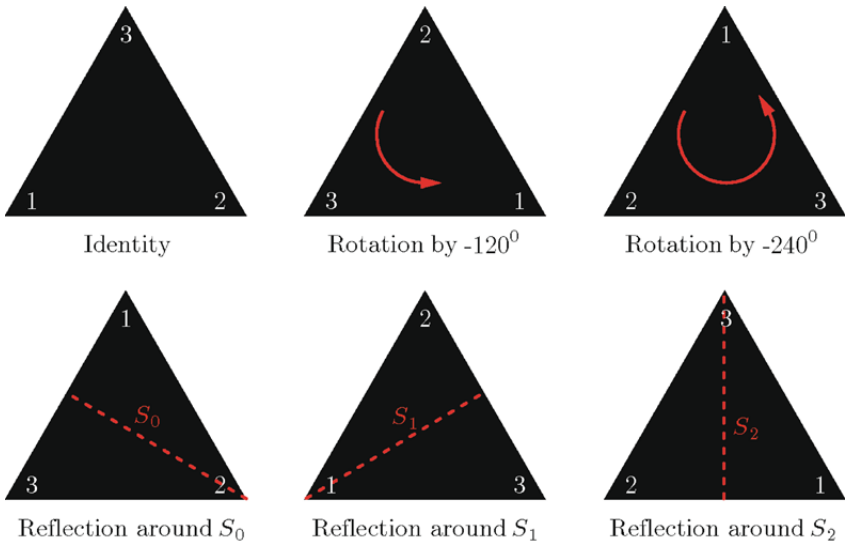


Figure 1.1: Symmetries of the Triangle

below. Notice that if we label the vertices of our fixed triangle as 1, 2, and 3, each symmetry results in a permutation of these vertices, and in fact all 6 permutations arise in this way! Here we see another way to view the group S_3, where each symmetry (rotation or reflection) is identified with the corresponding permutation of vertices which it causes. This is a homomorphism (and since bijective, an isomorphism), since each permutation associated to a composition of symmetries is the same as the composition of the associated permutations.

Another important concept that naturally arises from this example is that of a subgroup. In fact, it is the statement of Cayley's Theorem that every finite group is isomorphic to a subgroup of S_n for some n. Hence, subgroups allow us to view every finite group as a group of some, but perhaps not all, permutations.

Definition 1.2.4. *Let $(G, *)$ be a group. A subset $H \subseteq G$ is a **subgroup** if H is a group when equipped with the restriction of $*$ to H as a binary operation.*

In particular, a subgroup H must be closed under the binary operation $*$, contain the identity e of G (which will still be the identity in H), and be closed under inverses. The identity alone $\{e\}$ and G itself are always subgroups of G. From our example above of S_3, viewed as the rigid motion symmetries of an equilateral triangle, what if we only cared about rotational symmetries?

Then we only get 3 permutations of the vertices of the triangle instead of 6. Since rotating $120°$ counterclockwise twice is the same as rotating by $240°$, and three times ($360°$) is the same thing as doing nothing (the identity symmetry), the subgroup is said to be "generated" by the $120°$ rotation element, and we call the group "cyclic".

Definition 1.2.5. *A group G is **cyclic** if there exists $g \in G$, called a **generator** for G, such that $G = \{g^n | n \in \mathbb{Z}\}$, where here $g^n = g*g*\ldots*g$ (n times) if n is positive and $g^n = g^{-1}*g^{-1}*\ldots*g^{-1}$ ($-n$ times) if n is negative ($g^0 = e$).*

There are two possibilities for a cyclic group: (1) Each g^n is distinct, in which case the group is countably infinite, and in fact isomorphic to $\mathbb{Z}$ under addition (via k goes to g^k, since $g^k * g^l = g^{k+l}$ by definition), or (2) There exists a minimal $n \in \mathbb{N}$ such that $g^n = e$, so the cyclic group is finite of order n. In this second case, the cyclic group is isomorphic to $\mathbb{Z}/n\mathbb{Z}$, the integers modulo n under modular addition, since $g^k * g^l = g^{k+l} = g^{(k+l \mod n)}$. Hence, in the example of restricting to only rotational symmetries described above, we have a subgroup of S_3 that is isomorphic to $\mathbb{Z}/3\mathbb{Z}$. Similarly, we get three cyclic subgroups of S_3 isomorphic to $\mathbb{Z}/2\mathbb{Z}$, each generated by a single reflection, since reflecting twice across the same axis returns the triangle to its original position.

Cyclic groups have the additional property that elements commute with respect to the binary operation, since $g^k * g^l = g^{k+l} = g^{l+k} = g^l * g^k$. While this property is not required to solve simple equations like $a * x = b$ or $x * a = b$, it is very desirable, and so such groups are called "abelian" (named after the Norwegian mathematician, Niels Abel).

Definition 1.2.6. *A group $(G, *)$ is **abelian** if $*$ is commutative: for all $g, h \in G$, $g * h = h * g$*

S_3 is actually the smallest group that is NOT abelian: $213 \circ 231 = 132 \neq 321 = 231 \circ 213$, for example. In other words, reflections and rotations do not have to commute with one another. On the other hand $(\mathbb{R}, +)$ is abelian, but not cyclic (for any real number a, $\{na | n \in \mathbb{Z}$ is a cyclic subgroup, but not all of $\mathbb{R}$ itself).

In contrast to the symmetries of the triangle, consider the symmetry group of a square, this time labeling the vertices a, b, c, and d (see the figure below). Again, we can consider rotations and

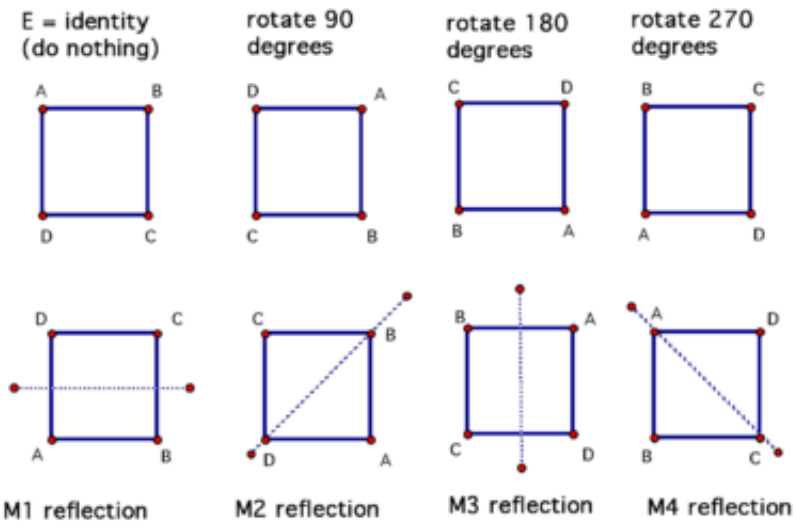


Figure 1.2: Symmetries of the Square

reflections that preserve the square, and the corresponding permutations of the vertices. However this time, not every permutation of the vertices arises (only 8 permutations instead of 24), resulting in a proper subgroup D_4 of S_4. D_4 is called the **dihedral group of order** 8. In fact, there is a geometric obstruction to obtaining all 24 possible permutations of the vertices, since diagonal pairs of vertices have to be preserved under reflections and rotations. Since the symmetry group, D_4, is not the entire symmetric group, S_4, we call such a symmetry group **enriched**, because it encodes interesting geometry that must be preserved by the transformations we are considering. We also again obtain the cyclic subgroup isomorphic to $\mathbb{Z}/4\mathbb{Z}$ generated by 90° clockwise rotation.

In general, for the rigid motion symmetries of the regular n-gon, we will always get a subgroup $D_n \subseteq S_n$, called the **dihedral group of order** $2n$. Restricting to rotational symmetries only, we obtain a subgroup isomorphic to $\mathbb{Z}/n\mathbb{Z} \subseteq S_n$.

Moving on to 3-dimensions, we can similarly ask for the rotational symmetry groups of the five Platonic solids: the tetrahedron, cube (hexahedron), octahedron, dodecahedron, and icosahedron.

This time, for each of the platonic solids, we will label the sides (think of them as dice), and again we will consider rotations permute this labeling. For the tetrahedron, there are four choices for which triangular side is the base, and then three rotations with that fixed base that preserve the

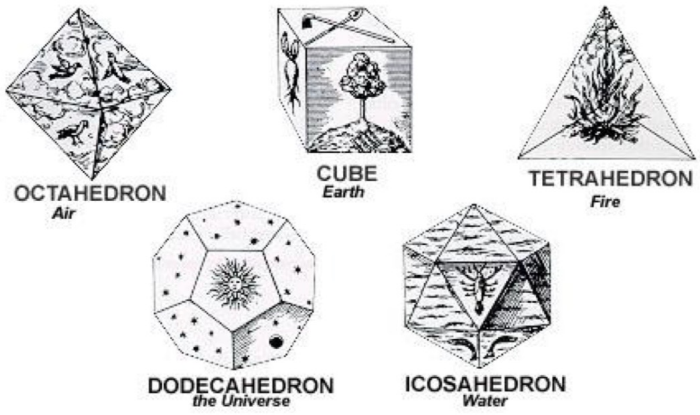


Figure 1.3: The Platonic Solids

shape of the tetrahedron. Hence, there are $(4)(3) = 12$ permutations (half of the 24 possible), and so again we get a proper subgroup A_4 of S_4, called the **alternating subgroup of** S_4. In general, a permutation in S_n can be assigned a "sign" (even or odd), and half of the permutations in S_n are even, and half are odd. Similarly to how multiplication of even natural numbers is even, the composition of even permutations is an even permutation. The identity permutation is also even, and the inverse of an even permutation is always even. Hence, for any n, there is the subgroup of all even permutations, $A_n \subseteq S_n$, called the **alternating subgroup of** S_n.

By a similar process, the cube has 24 permutations: there are six choices for which side is "on top" (what you would say you obtained by "rolling the dice"), and for each choice, four rotations that preserve the cube, so $24 = (6)(4)$. In fact, the rotational symmetry group of the cube is isomorphic to S_4, viewed as a subgroup of S_6.

Platonic solids come in dual pairs (see Figure 1.4), by taking the convex hull of the midpoint of every side, forming a new Platonic solid with number of vertices and number of sides swapped with those values of the Platonic solid you started with. This truly is a duality, since repeating the process twice gets one back to the Platonic solid one started with. Furthermore, rotations preserve both Platonic solids in precisely the same way (viewing one inscribed inside the other), and so

Figure 1.4: Duality of Platonic Solids

dual Platonic solids will have the same symmetry groups. The tetrahedron is self-dual, the cube and octahedron are dual, and the dodecahedron and icosahedron are dual. Therefore, the symmetry group of the octahedron is also isomorphic to S_4, but this time viewed as a subgroup of S_8.

By the same process as for the tetrahedron and the cube, the symmetry group of the dodecahedron has 60 permutations, since there are 12 choices of sides that could be "on top", and five rotations for each (since the sides are pentagons), so $60 = (12)(5)$. In fact, this symmetry group is isomorphic to A_5, but viewed as a subgroup of S_{12}. Finally, by duality, the symmetry group of the icosahedron is also isomorphic to A_5, but viewed as a subgroup of S_{20}. In all five cases, again we obtain enriched symmetry groups, revealing how "special" the geometric objects under consideration are.

Before moving on to more difficult equations that require more algebraic structure than that of a group, we conclude with a few advanced group-theoretic notions that we will need later. One should feel free to skip the rest of this section, and refer back to it as needed.

Definition 1.2.7. *Let $(G, *)$ be a group. A subgroup $N \subseteq G$ is said to be a **normal subgroup**, denoted $N \trianglelefteq G$, if for all $n \in N$ and $g \in G$, $gng^{-1} \in N$. A group G is said to be a **simple group** if it has no non-trivial, proper normal subgroups. If $N \subseteq G$ is a subgroup, one obtains an equivalence relation $\sim$ on G where $g \sim h$ if $g^{-1} * h \in N$, and the equivalence class of g is denoted by gN for all $g \in G$. The set of equivalence classes is denoted by G/N, and if further $N \trianglelefteq G$, G/N has the structure of a group, called a **quotient group**. Here the binary operation is given by $(gN) * (hN) = (g * h)N$ (which requires N to be normal to be well-defined), the identity is eN*

(where e denotes the identity in G), and the inverse of gN is $g^{-1}N$.

Definition 1.2.8. *Every finite group G has a **composition series**, which is a chain of normal subgroups (each normal in the next, not necessarily in G) $\{e\} = H_0 \trianglelefteq H_1 \trianglelefteq \cdots \trianglelefteq H_n = G$, with strict inclusions, such that each **factor group** H_{i+1}/H_i is simple. The factor groups are called **composition factors**, n is called the **composition length**. By the Jordan-H older Theorem, any two composition series of a group are equivalent: the have the same composition length and the same composition factors, up to permutation and isomorphism. For a finite group G, if the factor groups are all cyclic of prime order (isomorphic to $\mathbb{Z}/p\mathbb{Z}$ for some prime p), then we say that G is a **solvable group**. This is equivalent to the usual definition of solvable groups where the factor groups are required to be abelian, since the only abelian simple groups are cyclic of prime order.*

Definition 1.2.9. *Let p be a prime. A **p-group** is a finite group G with order a power of p: $|G| = p^k$ for some k. For any finite group G, a **Sylow p-subgroup** is a maximal p-subgroup of G, i.e. it is a p-group, and not contained in any other p-group contained in G. In other words, if $|G| = p^K m$ with $p \nmid m$, then a Sylow p-subgroup is a subgroup of G of order p^K.*

A p-group G has the special property that if $|G| = p^k$, then G has normal subgroups of order p^m for all $1 \leq m \leq k$. Additionally, there is a theorem, called the First Sylow Theorem (which we write here without proof), that asserts that for any prime $p \mid |G|$, for G a finite group, there exists a Sylow p-subgroup of G.

1.3 Fields

To make sense of solving equations of the form $a * x = b$, we needed the notion of a group. What if we moved beyond these simplest possible equations, to simple equations which involve two binary operations, like $ax + b = cx + d$? Here we could have written something abstract again, such as $a \diamond x * b = c \diamond x * d$, but for simplicity we will just write addition and multiplication in the standard simplified way, with the understanding that the two operations are abstract and that we are not necessarily discussing these familiar operations, or working over a number system like $\mathbb{R}$. Here notationally we use “brackets" interchangeably with parentheses for readability: so

$((a+b)+c) = [(a+b)+c]$, for example. To solve such an abstract system (where $a \neq c$, or else there would be nothing to solve):

1. We need an additive inverse for cx, say $-cx$. Add $-cx$ to both sides on the left: $-cx+(ax+b) = -cx+(cx+d)$.

2. We need $+$ to be associative, so we can move the parentheses on both sides: $(-cx+ax)+b = (-cx+cx)+d$

3. We need an additive identity, say 0: $(-cx+ax)+b = 0+d \implies (-cx+ax)+b = d$

4. Again, we need an additive inverse for b, say $-b$. Add $-b$ to both sides on the right: $[(-cx+ax)+b]+(-b) = d+(-b) \implies (-cx+ax)+[b+(-b)] = d+(-b) \implies (-cx+ax)+0 = d+(-b) \implies -cx+ax = d+(-b)$ (using our associativity and identity for $+$ as before)

5. We need our two operations to satisfy the distributive law with respect to one another: $(-c+a)x = d+(-b)$

6. We need a multiplicative inverse for $-c+a$, say $(-c+a)^{-1}$: $(-c+a)^{-1}[(-c+a)x] = (-c+a)^{-1}[d+(-b)]$

7. We need our multiplication to be associative: $[(-c+a)^{-1}(-c+a)]x = (-c+a)^{-1}[d+(-b)]$

8. We need a multiplicative identity, say 1: $1x = (-c+a)^{-1}[d+(-b)] \implies x = (-c+a)^{-1}[d+(-b)]$, which is our solution!

If we further want to solve similar equations such as $ax+b = d+cx$ (like the equation $10x+20 = 200-10x$ in the introductory section to this chapter), we would also need $+$ to be commutative. At the same time, equations like $ax+b = xc+d$ would also necessitate for our multiplication to be commutative as well. With all the properties listed so far, one could solve any linear equation in one variable, motivating our definition of a field:

Definition 1.3.1. *A **field** is a set F, equipped with two binary operations $+ : F \times F \to F$ and $\cdot : F \times F \to F$, such that $(F, +)$ and $(F \setminus \{0\}, \cdot)$ are abelian groups, satisfying the distributive law: For all $a, b, c \in F$, $a(b + c) = ab + ac$.*

Note that earlier we mentioned how $(\mathbb{R}, +)$ and $(\mathbb{R} \setminus \{0\}, \cdot)$ are groups. Since they are each abelian ("traditional" addition and multiplication are commutative), and the distributive law holds in $\mathbb{R}$, $\mathbb{R}$ is a field. However, the natural numbers $\mathbb{N}$ and integers $\mathbb{Z}$ are not fields: we already mentioned that $(\mathbb{N}, +)$ is not a group, and even though $(\mathbb{Z}, +)$ is an abelian group, $(\mathbb{Z} \setminus \{0\}, \cdot)$ is not (it doesn't have multiplicative inverses). As a result, in $\mathbb{Z}$ we cannot solve equations like $2x = 1$. To remedy this, we again "expand our world" and obtain the rational numbers $\mathbb{Q}$, the minimal field containing $\mathbb{Z}$, obtained by adding solutions to all equations $ax = b$ for $a, b \in \mathbb{Z}$ ($a \neq 0$).

All of the above mentioned examples of fields are infinite, but there are finite fields as well. As a family of examples, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is a field for every prime p, where the operations are modular addition and multiplication.

1.4 Polynomial Equations and Rings

As long as one is working over a field, linear equations can be solved. But what about polynomial equations, such as $-4.9t^2 + 19.6t + 58.8 = 0$, where the coefficients in this case come from $\mathbb{R}$? In general, we consider polynomial equations where the coefficients come from some field.

Definition 1.4.1. *A **single-variable polynomial** f over a field F is $f = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_1 x + a_0$ for some $d \in \mathbb{N}$ (called the **degree** of f), and for some $a_d, \ldots, a_0 \in F$. If the degree is not specified, we may write $f = \sum a_i x^i$, where it is understood that all $a_i = 0$ for i greater than the degree of f, whatever that may be. Two polynomials are defined to be equal, $\sum a_i x^i = \sum b_i x^i$, if and only if $a_i = b_i$ for all i. The set of all single-variable polynomials over F is denoted $F[x]$, called the **polynomial ring over** F.*

The set of polynomials $F[x]$ also has addition and multiplication binary operations: If $f = \sum a_i x^i$ and $g = \sum b_i x^i$, $f + g = \sum (a_i + b_i) x^i$, and $fg = \sum (\sum_{k+l=i} a_k b_l) x^i$. The set $(F[x], +)$

is in fact an abelian group, but $(F[x], \cdot)$ is not, with the only axiom not being satisfied being that polynomials do not in general have multiplicative inverses (unless the degree $d = 0$, in which case $f = a_0 \in F \setminus \{0\}$). However, this is fine, because we are not at this time interested in solving equations of the form $fx = g$, where $x \in F[x]$ is unknown, but instead are trying to solve equations of the form $f(x) = 0$, where $f \in F[x]$.

Instead, we say that sets like $\mathbb{Z}$ and $F[x]$, which are additive abelian groups with a compatible multiplication, are "rings" (in fact, integral domains), hence the name for $F[x]$.

Definition 1.4.2. *A **ring** is a set R, equipped with two binary operations $+ : R \times R \to R$ and $\cdot : R \times R \to R$, such that $(R, +)$ is an abelian group, and for all $r, s, t \in R$,*

1. *Multiplication is associative: $r(st) = (rs)t$*
2. *The distributive laws are satisfied: $r(s + t) = rs + rt$, and $(r + s)t = rt + st$*

Note that by definition, fields are a special type of ring. In general, rings can exhibit behavior different than that of $\mathbb{Z}$ and $F[x]$. For example, with modular addition and multiplication $\mathbb{Z}/n\mathbb{Z}$ is a ring for all n, but if n is composite, say $n = ab$ (so $a, b \neq 1$), then $ab = 0 \mod n$, but $a, b \neq 0 \mod n$. In this case, we call $a, b \mod n \in \mathbb{Z}/n\mathbb{Z}$ zero-divisors. Nonzero zero-divisors, or two nonzero elements multiplying together to be 0 (the additive inverse), cannot occur in fields, $\mathbb{Z}$, or $F[x]$, which are "integral domains".

Definition 1.4.3. *An **integral domain** is a ring $(R, +, \cdot)$ such that for all $r, s \in R$:*

1. *Multiplication is commutative: $rs = sr$*
2. *There exists a multiplicative identity $1 \in R$: $1r = r$*
3. *There are no non-zero zero divisors: If $rs = 0$, then $r = 0$ or $s = 0$*

Like with groups **subring** is a subset that is a ring with the restricted binary operations. An example of a ring that doesn't satisfy any of the additional axioms of an integral domain is $R = M_2(2\mathbb{Z})$, 2×2 matrices with even integers as entries (matrix multiplication doesn't have to

commute, the identity matrix doesn't have even entries, and two non-zero matrices can multiply to be the zero matrix). R is a subring of $M_2(\mathbb{R})$, all 2×2 matrices with real entries, which does have the identity matrix as multiplicative identity. Note that some authors prefer to have the definition of a ring include the existence of a multiplicative identity to avoid examples like $M_2(2\mathbb{Z})$, in which case they would consider what we call a ring without multiplicative identity a "rng" (pronounced "wrong", since the "i", the identity, is missing).

If one has a ring R, it is also possible to consider all single-variable polynomials over R, with the same binary operations as before, since inverses were not involved. In this case we denote the set of all such polynomials as $R[x]$.

Like groups, there are also homomorphisms and isomorphisms of rings. Below we use the same convention for addition and multiplication as our abstract binary operations for two rings at the same time, with the understanding that these operations are likely different across the two rings.

Definition 1.4.4. *Let R and S be rings. Then, a **ring homomorphism** $\varphi : R \to S$ is a function such that for all $r_1, r_2 \in R$:*

1. $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$
2. $\varphi(r_1 r_2) = \varphi(r_1)\varphi(r_2)$
3. *If R and S further have multiplicative identities, $\varphi(1) = 1$.*

*A **ring isomorphism** is a bijective ring homomorphism, and we say two rings R and S are **isomorphic** if there is an isomorphism between them. A **field homomorphism (isomorphism)** is a ring homomorphism (isomorphism) $\varphi : R \to S$, where both R and S are fields.*

Since every field F has a multiplicative identity 1, one can consider the sequence (1, $1 + 1$, $1 + 1 + 1$, ...) in F. There are two possibilities for this sequence of elements of F: either every element is distinct, or at some point the sequence repeats an element. If there is a repeat, one can solve the equation $1 + 1 + \cdots + 1$ (n times) = $1 + 1 + \cdots + 1$ (m times) (say $m > n$) to obtain $1 + 1 + \cdots + 1$ ($m - n$ times) $= 0$. This gives rise to the following definition.

Definition 1.4.5. *The **characteristic** of a field F is the minimal number n such that $1+1+\cdots+1 = 0$. If no such n exists, we say that F has characteristic 0.*

If F has characteristic 0, F contains an isomorphic copy of the rational numbers $\mathbb{Q}$, given by the isomorphism (onto its image) $\varphi : \mathbb{Q} \to \varphi(\mathbb{Q}) \subseteq F$, given by $\varphi(\frac{a}{b}) = (1+1+\cdots+1)^{-1}(1+1+\cdots+1)$, where the first sum is b times and the second sum is a times (if either a or b are negative, replace the corresponding 1's with -1's). On the other hand, if the characteristic of F is $n \neq 0$, then $n = p$ for some prime p. This is because if n were composite, say $n = kl$, then $0 = (1+1+\cdots+1)$ (n times) = $(1+1+\cdots+1)$ (k times) $\cdot$ $(1+1+\cdots+1)$ (l times), which implies $0 = (1+1+\cdots+1)$ (k times) or $0 = (1+1+\cdots+1)$ (l times), either contradicting the minimality of n. Hence a field F either has characteristic 0 or prime characteristic. When F has prime characteristic p, we again have a field isomorphism $\varphi : \mathbb{Z}/p\mathbb{Z} \to \varphi(\mathbb{Z}/p\mathbb{Z}) \subseteq F$ via $\varphi(n \mod p) = 1+1+\cdots 1$ (n times). Hence, every field F contains a subfield isomorphic to either $\mathbb{Q}$ or $\mathbb{Z}/p\mathbb{Z}$, called the **prime subfield** of F.

We now return to solving equations of the form $f(x) = 0$, where $f \in F[x]$ (F a field). Solutions to $f(x) = 0$ are also called the **roots** of f. For example, consider a right triangle with both base and height of length 1 and unknown hypotenuse c (see Figure 1.5). By the Pythagorean Theorem, $1^2 + 1^2 = c^2 \implies c^2 - 2 = 0$. Here $f = c^2 - 2 \in F[c]$. If one were trying to solve this equation over the rational numbers $\mathbb{Q}$, there would be no solution! The two solutions (over, say the real numbers $\mathbb{R}$), $c = -\sqrt{2}$ and $c = \sqrt{2}$ require us again to "expand our world", leading to the notion of a "field extension". In fact, even though the solutions exist in $\mathbb{R}$, the smallest field extension where the solutions exist is $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$, which lies between the fields $\mathbb{Q}$ and $\mathbb{R}$. Note $-\sqrt{2} = (0) + (-1)\sqrt{2} \in \mathbb{Q}(\sqrt{2})$, so $\mathbb{Q}(\sqrt{2})$ contains both roots.

Definition 1.4.6. *Let F be a field. A **field extension** K/F is a field K containing F. If one has such a field extension K/F, an element $a \in K$ is **algebraic over** F if it is the root of a non-zero polynomial $f \in F[x]$. An **algebraic extension** is a field extension K/F where every element of K is algebraic over F. A field F is **algebraically closed** if every $f \in F[x]$ has its roots in F. An **algebraic closure** of a field F is an algebraic extension K/F, where K is algebraically closed.*

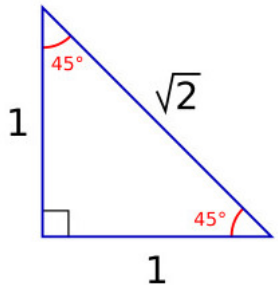


Figure 1.5: The square root of two appearing naturally as a length

In our example above, $\mathbb{Q}(\sqrt{2})$ is an algebraic extension of $\mathbb{Q}$ and the splitting field of $x^2 - 2 \in \mathbb{Q}[x]$, but is not algebraically closed, since for example, it does not contain the roots of $x^2 - 3$. While $\mathbb{R}$ has solutions to all the equations we have considered so far, it does not contain the roots $-i = -\sqrt{-1}$ and $i = \sqrt{-1}$ of $x^2 + 1 \in \mathbb{Q}[x]$, and so $\mathbb{R}$ is not algebraically closed. Instead, $\mathbb{C} = \mathbb{R}(i) = \{a + bi \mid a, b \in \mathbb{R}\}$ is algebraically closed and is the algebraic closure of $\mathbb{R}$ (this is the Fundamental Theorem of Algebra, proven later in section 1.6). On the other hand, $\mathbb{Q}(i) = \{a+bi \mid a, b \in \mathbb{Q}\}$ is the splitting field of $x^2 + 1 \in \mathbb{Q}[x]$, and is countable, unlike the real numbers. The algebraic closure of $\mathbb{Q}$ is denoted by $\overline{\mathbb{Q}}$, and is called the **field of algebraic numbers**. Elements of $\mathbb{R}$ like $e = \lim_{n\to\infty}(1 + \frac{1}{n})^n$ and $\pi = \tan^{-1}(1) = 4\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{2k-1}$ are not algebraic over $\mathbb{Q}$, and so are called **transcendental** over $\mathbb{Q}$. Instead, these transcendental elements in $\mathbb{R}$ are defined by limiting processes, and in fact this is how one constructs $\mathbb{R}$ as a set of equivalence classes of Cauchy sequences with terms in $\mathbb{Q}$, involving analysis instead of algebra.

So far in our examples we have only considered quadratic equations (where extensions just involve adding the square root of some element), but in general constructing the splitting field of $f \in F[x]$ involves the notions of ideals and quotient rings.

Definition 1.4.7. *Let R be a ring. An **ideal** $I \trianglelefteq R$ is a non-empty subset I of R such that for all $f, g \in I$ and $r \in R$, $f + g \in I$ and $rf \in I$. For any $f \in R$, one obtains **the ideal generated by** f, $(f) = \{rf \mid r \in R\}$. For an ideal $I \trianglelefteq R$, if there exists $f \in I$ such that $I = (f)$, then we say I is a **principal ideal**.*

In other words, an ideal $I \trianglelefteq R$ is an abelian subgroup that is closed under multiplication from its ambient ring R. One motivation for this definition is that when considering solving an equation $f(x) = 0$ for $f \in F[x]$, any solution will also be a solution of $r(x)f(x) = 0$ for any $r(x) \in F[x]$ (since F is a field, and thus has no non-zero zero divisors). In other words, finding the roots of f is equivalent to finding the common roots to all polynomials in the ideal (f).

One can also use ideals to get equivalence relations on a ring R: Given an ideal $I \trianglelefteq R$, define the relation $\sim$ on R by $r_1 \sim r_2$ if $r_1 - r_2 \in I$ (put another way, $r_1 = r_2 + f$ for some $f \in I$). One can show that this indeed gives an equivalence relation on R, at which point we denote the equivalence class of $r \in R$ by $r + I$. As will be shown, this is more than mere notation.

Definition 1.4.8. *Let R be a ring, and $I \trianglelefteq R$ an ideal. The **quotient ring** R/I is the set of equivalence classes $\{r + I \mid r \in R\}$, equipped with binary operations $(r_1 + I) + (r_2 + I) = (r_1 + r_2) + I$ and $(r_1 + I)(r_2 + I) = (r_1 r_2) + I$.*

One can show that these binary operations are well-defined and satisfy the axioms of a ring, where the additive identity is $I = 0 + I$, the additive inverse of $r + I$ is $(-r) + I$, and if R has multiplicative identity 1, the multiplicative identity of R/I is $1 + I$. Note that the notation R/I for a quotient ring is identical to the notation K/F for a field extension, which is unfortunate, but should still be clear from context. However, the the two concepts, while certainly not the same, are related in a special case relevant to our current discussion. In fact, if $f \in F[x]$ is an irreducible polynomial (meaning it cannot be expressed as $f = gh$ for non-constant $g, h \in F[x]$), then $K = F[x]/(f)$ is actually a field containing (an isomorphic copy of) the base field F. Hence, K/F is a field extension, and is how we construct the splitting field of f (a minimal algebraic field extension containing all the roots of f).

In the examples we considered earlier, $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[x]/(x^2 - 2)$, $\mathbb{Q}(i) = \mathbb{Q}[x]/(x^2 + 1)$, and $\mathbb{C} = \mathbb{R}(i) = \mathbb{R}[x]/(x^2 + 1)$. For a higher degree example, consider the cubic polynomial $f = 2x^3 - 9x^2 - 6x + 3 \in \mathbb{Q}[x]$, which is irreducible by the Rational Root Test. By Cardano's cubic formula, roots of f are $x = \alpha_k = \frac{3 - \omega^k \sqrt[3]{39 - 26i} - \omega^{2k} \sqrt[3]{39 + 26i}}{2}$ $(k = 1, 2, 3)$, where $\omega - \frac{1}{2} + \frac{\sqrt{3}}{2} i$ is a primitive 3rd root of unity (so $\omega^3 = 1$). The splitting field of f is $K = \mathbb{Q}(\alpha_1, \alpha_2, \alpha_3) = \mathbb{Q}(\sqrt[3]{2}, \omega)$.

In general, one can compute a splitting field K for $f \in F[x]$ of degree n by constructing a chain of fields $F = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_{r-1} \subseteq K_r = K$ such that each K_i is an extension of K_{i-1} containing a new root of f. Since f has at most n roots, the construction will require at most n extensions. The algorithm for computing each K_i is given by:

- Factor $f = f_1 \cdots f_k$ into irreducible factors over K_{i-1}
- Choose any non-linear irreducible factor f_i in $\{f_1, \ldots, f_k\}$
- Construct $K_i = K_{i-1}[x]/(f_i)$
- Repeat this process until f splits completely into linear factors.

This process for each extension was dependent on our choice of a non-linear irreducible factor of f, but in the end K can be shown to be unique up to isomorphism.

We already have discussed that fields are the correct setting for solving general linear equations, but now we can solve any polynomial equation $f(x) = 0$ for $f \in F[x]$ (F a field) by again "expanding our world" and passing to the splitting field of f, the smallest field where the equation can be solved. One can also pass to the algebraic closure of F, $\overline{F}$,where all polynomial equations $f(x) = 0$ for $f \in F[x]$ can be solved. In the other direction, if you have an element $\alpha \in \overline{F}$, one can find a minimal polynomial equation $f(x) = 0$ with α as a solution.

Definition 1.4.9. *Let K/F be a field extension, and let $\alpha \in K$ be algebraic over F. Then, the* ***minimal polynomial*** *$m_\alpha \in F[x]$ is the monic (leading coefficient is 1) polynomial of least degree among all $f \in F[x]$ having α as a root.*

One can show that the minimal polynomial is unique and irreducible. For example, even though $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ is a root of $x^3 - 1 \in \mathbb{Q}[x]$, its minimal polynomial is $m_\omega = x^2 + x + 1$.

1.5 Galois Extensions and Galois Groups

We are now ready to discuss Galois groups - the central object considered in this work! Galois groups are groups associated to certain field extensions, called Galois extensions. We build towards this goal by first introducing important types of field extensions.

Definition 1.5.1. *An algebraic field extension K/F is a **normal extension** if every irreducible $f \in F[x]$ with a root in K splits into linear factors in $K[x]$. The **normal closure** of an algebraic extension K/F is the smallest field M such that $F \subseteq K \subseteq M$ and and M/F is a normal extension.*

The algebraic field extensions that we have considered so far are all splitting fields of a polynomial f, and so will be normal since they contain all the roots of f. To help illustrate what can go wrong, consider $f = x^3 - 2$, which has as roots $x = \sqrt[3]{2}$, $\omega\sqrt[3]{2}$, and $\omega^2\sqrt[3]{2}$ (where again $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ is the primitive 3rd root of unity from before). Then, $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ is an algebraic field extension containing $\sqrt[3]{2}$, but it is not a normal extension, since it does not contain the other roots of f, meaning f does not split completely into linear factors over $\mathbb{Q}(\sqrt[3]{2})$. However, the splitting field $K/\mathbb{Q}$ is still a normal extension, and is the normal closure of the extension $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$.

Normal extensions are critical for solving polynomial equations of the form $f(x) = 0$ for $f \in F$ (F a field), since they contain all of the solutions, not just some of them. At the same time, for our purposes we want distinct solutions, not repeated solutions. This leads to the notion of separability.

Definition 1.5.2. *A **separable polynomial** is a polynomial $f \in F[x]$ (F a field) such that over the algebraic closure $\overline{F}$ of F, the roots of f are distinct (so the number of distinct roots of f is equal to the degree of f). An algebraic extension K/F is a **separable extension** if for all $\alpha \in E$, the minimal polynomial m_α is a separable polynomial (so has no repeated roots in any extension field).*

Separability will not actually be a concern for our extensions, since over fields of characteristic 0 or finite fields (which are all that we consider), all field extensions are separable. However, we include this definition for completeness. Together with normality, separability gives the type of field extensions best suited for solving equations - Galois extensions. A field **automorphism** is a field isomorphism $\varphi : F \to F$ from a field F to itself.

Definition 1.5.3. *A **Galois extension** is a field extension K/F that is both normal and separable. Equivalently, K is the splitting field of some separable $f \in F[x]$. The **Galois group** associated*

to a Galois extension is $\mathrm{Gal}(K/F) = \{$*field automorphisms* $\sigma : K \to K \mid \sigma|_F = Id_F\}$*. In other words, an automorphism* $\sigma \in \mathrm{Gal}(K/F)$ *has the property that* $\sigma(a) = a$ *for all* $a \in F$ *(it "fixes" the base field* F*). Galois groups are also defined for separable extensions* K/F *as* $\mathrm{Gal}(M/F)$*, where* M *is the normal closure of* K*. The Galois group of a polynomial* $f \in F[x]$ *is* $\mathrm{Gal}(f) = \mathrm{Gal}(K/F)$*, where* K *is the splitting field of* f*.*

From our previous examples, $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q}) \cong \mathrm{Gal}(\mathbb{Q}(i)/\mathbb{Q}) \cong \mathrm{Gal}(\mathbb{C}/\mathbb{R}) \cong \mathbb{Z}/2\mathbb{Z}$, where in each case the only field automorphisms fixing the base field are the identity map, and conjugation ($\sqrt{2} \leftrightarrow -\sqrt{2}, i \leftrightarrow -i$). In the $\mathbb{C}/\mathbb{R}$ case, this automorphism is also known as complex-conjugation. Notice how in these examples, the elements of the Galois group permute the roots. This generalizes to higher degree examples.

Lemma 1.5.4. *For any Galois extension* K/F *and separable polynomial* $f \in F[x]$ *with roots in* K*,* $G = \mathrm{Gal}(K/F)$ *permutes the roots of* f*.*

Proof. If $\alpha \in K$ is a root of $f = a_n x^n + \cdots + a_1 x + a_0 \in F[x]$, and $\sigma \in G$, then since σ is a field automorphism of K fixing F, $f(\sigma(\alpha)) = a_n(\sigma(\alpha))^n + \cdots + a_1\sigma(\alpha) + a_0 = \sigma(a_n\alpha^n + \cdots + a_1\alpha + a_0) = \sigma(f(\alpha)) = \sigma(0) = 0$. Hence, $\sigma(\alpha)$ is also a root of f. Since f is separable, and σ is a bijection, with $\sigma \in G$ arbitrary, G permutes the roots of f. □

In other words, the elements of the Galois group of a polynomial $\mathrm{Gal}(f)$ fix the coefficients of the polynomial, but permute its roots, the solutions to $f(x) = 0$ in its splitting field. Hence, the Galois group is a subgroup of the permutation group $\mathrm{Gal}(f) \subseteq S_d$, where d is the degree of f, and so the order of the group is $|\mathrm{Gal}(f)| \leq d!$. In particular, the Galois group of a polynomial is a finite group. One can view this Galois group as the "symmetries" of the solution set to the polynomial equation $f(x) = 0$, like how we originally viewed some groups as symmetries of certain geometric objects (here the objects being points on the affine line F.

For more examples, consider $f = x^3 - 1 \in \mathbb{Q}[x]$ and $g = x^3 - 2 \in \mathbb{Q}[x]$. Then, the splitting field of f is $\mathbb{Q}(\omega)$, and the splitting field of g is $\mathbb{Q}(\sqrt[3]{2}, \omega)$, where again $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. Both f and g have degree 3, and so after fixing an ordering of their roots, their Galois groups can be viewed

as subgroups of S_3 (different orderings of the roots give different, though isomorphic, subgroups). Note that $\mathrm{Gal}(f) = \mathrm{Gal}(\mathbb{Q}(\omega)/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$ since the roots of f are 1, ω, and ω^2, and $1 \in \mathbb{Q}$ must be fixed by any automorphism. Hence, only ω and ω^2 can be permuted. On the other hand, $\mathrm{Gal}(g) \cong S_3$ is the full symmetric group. The roots of g are $\sqrt[3]{2}$, $\omega\sqrt[3]{2}$, $\omega^2\sqrt[3]{2}$ (fix this order to get permutations), and every automorphism of $\mathbb{Q}(\sqrt[3]{2}, \omega)$ fixing $\mathbb{Q}$ is determined by the images of $\sqrt[3]{2}$ and ω, making sure that the third power of these images are still 2 and 1, respectively. One gets the following automorphisms, extended polynomially to the entire field:

1. $\sqrt[3]{2} \to \sqrt[3]{2}$ and $\omega \to \omega$ (induces the identity permutation 123)
2. $\sqrt[3]{2} \to \sqrt[3]{2}$ and $\omega \to \omega^2$ (induces the permutation 132)
3. $\sqrt[3]{2} \to \omega\sqrt[3]{2}$ and $\omega \to \omega$ (induces the permutation 231)
4. $\sqrt[3]{2} \to \omega^2\sqrt[3]{2}$ and $\omega \to \omega$ (induces the permutation 312)
5. $\sqrt[3]{2} \to \omega\sqrt[3]{2}$ and $\omega \to \omega^2$ (induces the permutation 213)
6. $\sqrt[3]{2} \to \omega^2\sqrt[3]{2}$ and $\omega \to \omega^2$ (induces the permutation 321)

Note that in the examples above, there was a strong relationship between the Galois extensions and the corresponding Galois groups. This is the case in general. Note that by an intermediate field E of a field extension K/F, we mean that E is a field such that $F \subseteq E \subseteq K$.

Theorem 1.5.5. *(The Fundamental Theorem of Galois Theory) Given a field extension K/F that is finite and Galois, there is a one-to-one correspondence between its intermediate fields and the subgroups of its corresponding Galois group $\mathrm{Gal}(K/F)$. Explicitly,*

- *For any intermediate field E of K/F, we obtain a subgroup $\mathrm{Gal}(K/E) \subseteq \mathrm{Gal}(K/F)$*
- *For any subgroup $H \subseteq \mathrm{Gal}(K/F)$, we obtain an intermediate field $K^H = \{\alpha \in K \mid \sigma(\alpha) = \alpha \text{ for all } \sigma \in H\}$, called the **fixed field of** H*
- *This correspondence is one-to-one, so $K^{\mathrm{Gal}(K/E)} = E$ and $\mathrm{Gal}(K/K^H) = H$.*

- *The correspondence is inclusion-reversing, meaning that subgroups H_1 and H_2 of $\mathrm{Gal}(K/F)$ satisfy $H_1 \subseteq H_2$ if and only if $K^{H_2} \subseteq K^{H_1}$*

- *In particular, K corresponds to the trivial subgroup $\{0\} \subseteq \mathrm{Gal}(K/F)$, and F corresponds to the entire Galois group $\mathrm{Gal}(K/F)$.*

Returning to our previous example, $\mathbb{Q}(\omega)$ and $\mathbb{Q}(\sqrt[3]{2})$ are intermediate fields of the Galois extension $\mathbb{Q}(\sqrt[3]{2}, \omega)/\mathbb{Q})$, with Galois groups $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, \omega)/\mathbb{Q}(\omega)) \cong \mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Q}(\sqrt[3]{2}, \omega)/\mathbb{Q}(\sqrt[3]{2}) \cong \mathbb{Z}/2\mathbb{Z}$, both viewed as subgroups of S_3. These intermediate fields are the fixed fields of these subgroups.

To conclude this section, along with our discussion to solving polynomial equations $f(x) = 0$ over a field, we discuss formulas for solving such equations. When f has degree 2, solutions exist over an extension field where certain square roots exist, using the quadratic formula. When f has degree 3, similarly one can use Cardano's cubic formula over a field extension containing $i = \sqrt{-1}$, a primitive cube root of unity ω, and relevant cubic roots. We say that such formulas allow one to solve quadratic and cubic equations algebraically (using only addition, subtraction, multiplication, division, exponents, and radicals). One also says that such equations are solvable by radicals. Though much more complicated of a formula, quartic (degree 4 polynomials) are also solvable by radicals. On the other hand, using the Fundamental Theorem of Galois Theory, we will show that there is no such formula for quintics (degree 5 polynomials), or polynomials of higher degree.

Definition 1.5.6. *A polynomial $f \in F[x]$ (F a field) is **solvable by radicals** if there exists a chain of field extensions $F = F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_m = K$ such that*

1. *$F_i = F_{i-1}[\alpha_i]$, where $\alpha_i^{m_i} \in F_{i-1}$ (so α_i is a solution to the polynomial equation $x^{m_i} - a$ for some $a \in F_{i-1}$)*

2. *K contains the splitting field of f*

If F is a field of characteristic 0 (like $\mathbb{Q}$ and its algebraic extensions, $\mathbb{R}$, or $\mathbb{C}$), then by the Fundamental Theorem of Galois Theory, a polynomial $f \in F[x]$ is solvable by radicals if and only

if $\mathrm{Gal}(f)$ is a solvable group (equating the subgroups in the composition series of the finite Galois group of f with the intermediate fields in the chain of field extensions by adding radicals). Since S_2, S_3, and S_4 are each solvable groups, any polynomial of those degrees is solvable by radicals. On the other hand, for $n \geq 5$, $A_n \trianglelefteq S_n$ is the only non-trivial, proper normal subgroup (A_n is simple), so since there are polynomials $f \in F[x]$ with Galois group isomorphic to S_n, in general polynomials of degree 5 or greater are not solvable by radicals. Thus, there is no general quintic equation or equation for solving higher degree polynomial equations. However, if $f \in F[x]$ has degree $d \geq 5$ with solvable Galois group $\mathrm{Gal}(f) \subsetneq S_d$, then the specific polynomial equation $f(x) = 0$ is still solvable by radicals.

1.6 Systems of Linear Equations and Abstract Vector Spaces

Now that we've fully developed the theory of fields to solve linear equations, and field extensions to solve polynomial equations, both in one variable, we now consider systems of linear equations in several variables. Using the methods of elimination and/or substitution, students in high school learn how to solve systems such as: (the solution is $x = \frac{3}{2}$, $y = -1$)

$$\begin{cases} 2x - 3y = 6 \\ 4x + 5y = 1 \end{cases} \tag{1.1}$$

In general, a system of linear equations (with m equations in n unknowns, called an $m \times n$ system) is

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

In matrix notation, if $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$, then we can succinctly write the system of linear equations as $A\vec{x} = \vec{b}$. In linear algebra, where one usually encounters such systems for the first time, each a_{ij}, x_k, and b_l are considered to be in $\mathbb{R}$ or $\mathbb{C}$, which are fields. Of course, just like single-variable linear equations, these systems make sense over any field, and so we will consider this more abstract approach to linear systems. Again, continuing the theme of this chapter, we will ask "what operations and properties will be required to solve such systems?" This time, the answer is encapsulated in the notion of a vector space.

Definition 1.6.1. *A **vector space over a field** F is an abelian group $(V, +)$ equipped with a linear field action $\cdot : F \times V \to V$. Elements of V are called **vectors** and $+$ is called **vector addition**. On the other hand, elements of F are called **scalars**, the action is called **scalar multiplication**, and as before for $a \in F$ and $v \in V$, we write av for $a \cdot v$ (even though a and v belong to potentially different sets). By a linear field action, we mean that scalar multiplication satisfies for all $v, v_1, v_2 \in V$ and $a, a_1, a_2 \in F$:*

1. *Scalar addition distributes over the action:* $(a_1 + a_2)v = a_1v + a_2v$
2. *The action distributes over vector addition:* $a(v_1 + v_2) = av_1 + av_2$
3. *The multiplicative identity acts as the identity:* $1v = v$
4. *The action is compatible with field multiplication (a type of associativity):* $(a_1a_2)v = a_1(a_2v)$

Even considering solving a 2×2 system (where say $a_{11} \neq 0$, and label the first equation R_1 and the second equation R_2)

$$\begin{cases} R_1 : a_{11}x + a_{12}y = b_1 \\ R_2 : a_{21}x + a_{22}y = b_2 \end{cases} \tag{1.2}$$

to use the method of elimination (also known as the Gauss-Jordan elimination algorithm, and which is equivalent to performing row operations on the corresponding augmented matrix $(A \mid \vec{b})$, the first two steps would be to:

1. Multiply both sides of the first equation by $\frac{1}{a_{11}}$, replacing the first equation with this new equation in our system. $(R_1 \to \frac{1}{a_{11}} R_1)$

2. Multiply both sides of the new first equation by $-a_{21}$, and then add this new left hand side of the first equation to the left hand side of the second equation (and same thing for the right hand sides). Replace the second equation with the result. $(R_2 \to -a_{21} R_1 + R_2)$

One can check that like balancing single variable equations, the new system has the same solutions as the previous system. Additionally, one can check that by performing these two steps, one uses all of the properties of a vector space to obtain the new system

$$\begin{cases} x + \frac{a_{12}}{a_{11}} y = \frac{b_1}{a_{11}} \\ (\frac{-a_{12}a_{21}}{a_{11}} + a_{22}) y = \frac{-a_{21}b_1}{a_{11}} + b_2 \end{cases} \tag{1.3}$$

The second linear equation can then be solved for y using the properties of the field F, substituted for y in the first equation, and then the first equation can also be solved for x using the properties of the field F. Hence, no additional properties are needed, and a vector space is precisely the necessary algebraic structure for solving the system. However, unlike equations considered previously in this chapter, there is still the possibility that no solution exists (if $\frac{-a_{12}a_{21}}{a_1 1} + a_{22} = 0$, but $\frac{-a_{21}b_1}{a_{11}} + b_2 \neq 0$) or that there are infinitely many solutions (if $\frac{-a_{12}a_{21}}{a_1 1} + a_{22} = \frac{-a_{21}b_1}{a_{11}} + b_2 = 0$). Geometrically, if this system were over $\mathbb{R}$, one could visualize a unique solution as the intersection point of two lines, and so no solutions corresponds to parallel lines, which never intersect. On the other hand, the case with infinitely many solutions corresponds to both lines being the same line, resulting in a “double line", every point of which is a solution. However, the vector space properties are required for computing a unique solution, if it exists, and still can be utilized to determine whether the system has no solutions or infinitely many solutions (since the matrix corresponding

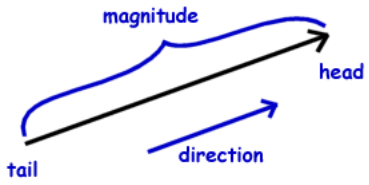


Figure 1.6: A vector in the plane

to our new system after our two steps is in row echelon form). As before, whenever solutions do not exist, we “expand our world" so that solutions do exist, which is the topic of our next section.

We now consider examples of vector spaces, and start with the namesake of vector spaces: vectors in the plane. This allows us to view geometric objects in a coordinate-free and coordinate-dependent manner, motivating the upcoming algebraic geometry.

Consider a plane - not with grid lines or axes like the Cartesian plane high school students are used to, but more like a blank whiteboard. In that plane, one can consider basic geometric objects like lines, circles, and triangles. The point is that no equations or variables are necessary - these geometric objects that we interact with in life exist independent of coordinates. Similarly, one can consider vectors in the plane, which are arrows with a tail, head, magnitude direction (see Figure 1.6). One confusing thing about vectors is that they are represented in the plane in the same way that another geometric object is drawn - rays. However, rays are considered to be “one-sided" lines, where the arrow indicates that the object continues progressing, but this is not the case for vectors. The head of the vector simply indicates direction, and does not continue - the entire geometric object is there as pictured.

Now, vectors in physics represent forces - one can imagine an object being “pushed" from the tail to the head of the vector. The magnitude of the vector measures the strength of the force, and the direction of the vector gives the direction of the force. The physical location of where the vector lies in the plane is inconsequential - one translate that vector (without changing its magnitude or direction), and it is still considered the same vector. With this interpretation in mind, forces can

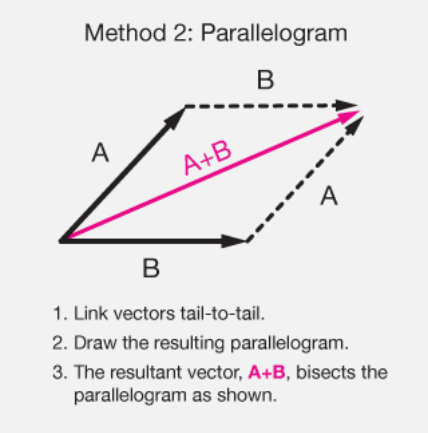


Figure 1.7: Parallelogram Law for Adding Vectors

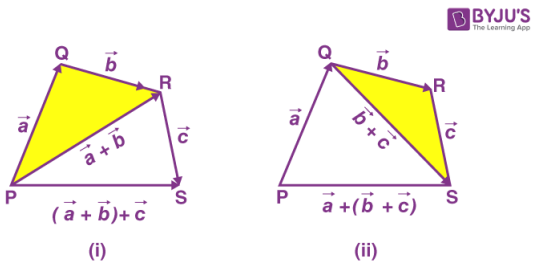


Figure 1.8: Vector addition is associative

be composed. For example, throwing a Wiffle ball outdoors can represent one force, but if there is wind present, that represents another force. As a result, the Wiffle ball thrown in one direction will not continue in that direction, nor completely in the direction of the wind either, but somewhere in between. Hence, we define the addition of vectors to be the vector representing the composition of forces, given for example by the parallelogram law.

From this definition, one can see readily that vector addition is commutative. Similarly, one can see that vector addition is associative.

Furthermore, there is a "0-vector" $\vec{0}$, which looks like a point and has 0 magnitude, which is an additive identity among vectors in the plane. For any vector $\vec{v}$ in the plane, one can rotate that vector $180°$ to get $-\vec{v}$ so that it has the same tail and magnitude, but is pointing in the opposite direction. This $-\vec{v}$ is an additive inverse for $\vec{v}$, since $(-\vec{v}) + \vec{v} = \vec{0}$. Hence, the set of vectors in

the plane form an abelian group - but this is not all! One can also scale a vector in the plane by a real scalar $c \in \mathbb{R}$: If that scalar is positive, $c\vec{v}$ points in the same direction, but its magnitude is multiplied by c. If that scalar is negative, $c\vec{v}$ points in the opposite direction, and its magnitude is multiplied by $|c|$. In particular, $1\vec{v} = \vec{v}$, $(ab)\vec{v} = a(b\vec{v})$, and scalar multiplication and addition satisfy the distributive laws. Therefore, the set of vectors in the plane form a vector space! Note that there is no natural multiplication operation for vectors, so this is truly all of the algebraic structure that we obtain. As a result of being a vector space, one can solve systems of linear equations involving vectors in the plane.

Now linear algebra is not just the study of vector spaces, but also the study of linear transformations between vector spaces. Since a vector space has vector addition and scalar multiplication as operations, linear transformations are functions that respect this structure.

Definition 1.6.2. *Let V and W be vector spaces over the same field F. Then, a **linear transformation** $L : V \to W$ is a function such that for all $v, v_1, v_2 \in V$ and $c \in F$,*

1. $L(v_1 + v_2) = L(v_1) + L(v_2)$
2. $L(cv) = cL(v)$

*If $L : V \to W$ is a bijective linear transformation, then we call it an **isomorphism of vector spaces**, and that V and W are **isomorphic as vector spaces**, denoted $V \cong W$.*

For vectors in the plane, some possible linear transformations are rotations and reflections. One can see that from the parallelogram law definition of vector addition, it does not matter if one adds vectors first, and then transforms the result, or if one transforms the vectors separately first, and then add them together - both give the same result. Similarly, scaling before or after a rotation or reflection does not matter - the result is the same. This is the flavor of how linear transformations work in general.

Now, while vectors in the plane and linear transformations between them can be fully described without coordinates, it is often useful to introduce coordinates to do hands on calculations. For example, it is easier to do vector calculations on a computer using coordinates rather than having

the computer draw vectors, move them around, and such. We just caution that coordinates give a representation of vectors in the plane, and are not inherently what the vectors are. This is similar to how numbers can be conveniently represented in the decimal (base 10) system, but there are many other representations of numbers that could have alternatively been used (like binary or Roman numerals for example). Regardless of how one represents numbers, what they really represent is quantity, and similarly vectors in the plane represent quantities with magnitude and direction. As another geometric example, circles are completely defined as the set of points a fixed distance (radius r) away from a given point (its center). This does not require coordinates, and circles can be physically constructed from a ruler and compass, for example. However, if one introduces Cartesian coordinates in the plane, the center now is given by some point (h, k), and the circle can be (via the Pythagorean Theorem) described as the solutions to the equation $(x-h)^2+(y-k)^2 = r^2$. But this equation is not what the circle is, just a way to represent such a circle. Changing the coordinate system would change the equation, and the same is true for vectors.

Choosing a unit of measurement (like inches, for example), there are two distinguished vectors in the plane, the vector $\vec{i}$ which has magnitude 1 and is perfectly horizontal (with arrow pointing to the right), and the vector $\vec{j}$ which has magnitude 1 and is perfectly vertical (with arrow pointing up). Then, every vector $\vec{v}$ in the plane can be uniquely written as a linear combination of these two distinguished vectors, which we now describe. If $\vec{v}$ is completely horizontal or vertical, $\vec{v}$ is just a (could be positive or negative) scaling of $\vec{i}$ or $\vec{j}$, and so that case is simple. Otherwise, $\vec{v}$ is the hypotenuse of a right triangle, with base b and height h, and so $\vec{v} = \epsilon_1 b\vec{i} + \epsilon_2 h\vec{j}$, where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$ depends on the direction $\vec{v}$ is facing. We call $\{\vec{i}, \vec{j}\}$ the **standard basis** for vectors in the plane.

Fixing the vectors $\vec{i}$ and $\vec{j}$ into a specific position in the plane with their tails together forms a coordinate system. The origin is where these tails are, and represents the zero vector $\vec{0}$. We get axes with tick marks as usual, representing integer multiples of $\vec{i}$ and $\vec{j}$. Any vector $\vec{v} = \epsilon_1 b\vec{i} + \epsilon_2 h\vec{j} = x\vec{i} + y\vec{j}$ $(x, y \in \mathbb{R})$ is then represented by the point where its head is $(\epsilon_1 b, \epsilon_2 h) = (x, y)$. The four quadrants are determined by the signs of ϵ_1 and ϵ_2 (the positive quadrant, or quadrant I being

where both are positive, for example). Note that there were many choices in how we came up with this standard, or Cartesian, coordinate system, given by a unit of measurement and a location for $\vec{0}$. Other choices would have given a different coordinate system. Furthermore, we could have started with any two vectors not facing the same or opposite directions, which would have resulted in a coordinate system with a parallelogram-shaped grid system instead of the usual squares.

Now that we have our standard coordinate system, we call the set of vectors in the plane represented in these coordinates $\mathbb{R}^2 = \{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \in \mathbb{R}\}$. Here we use the matrix $\begin{bmatrix} x \\ y \end{bmatrix}$ to distinguish the vector $\vec{v} = x\vec{i} + y\vec{j}$ from the point where its head lies (x, y), to remind us that the geometric object has magnitude and direction, and that these vectors lie in a real vector space (and so can be added together or multiplied by scalars in $\mathbb{R}$). In this coordinate system, $\vec{i}$ is represented by the matrix $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{j}$ is represented by the matrix $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. With respect to the coordinate system which we have fixed, we can also represent our operations in coordinates, and see that $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$ and $c \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ cy \end{bmatrix}$. Hence, these operations on matrices are not arbitrarily chosen when learning about vectors in linear algebra - they represent the coordinate-free operations of vector addition and scalar multiplication of vectors in a particular coordinate system.

Just like vectors themselves, linear transformations can be represented in coordinates as matrices as well. Consider first 90° counter-clockwise rotation $L = R_{90^\circ}$ of vectors in the plane (which does not require coordinates to define). Since any $\vec{v} = x\vec{i} + y\vec{j}$ and R_{90° is a linear transformation, $R_{90^\circ}(\vec{v}) = xR_{90^\circ}(\vec{i}) + yR_{90^\circ}(\vec{j})$, so to understand what R_{90° does to vectors in general, it suffices to understand how R_{90° transforms the standard basis vectors $\vec{i}$ and $\vec{j}$. In coordinates, R_{90° sends $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ to $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. On the other hand R_{90° sends $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $-\vec{e_1} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. Hence, since $R_{90^\circ}(\vec{v}) = xR_{90^\circ}(\vec{i}) + yR_{90^\circ}(\vec{j})$, R_{90° sends in coordinates $\begin{bmatrix} x \\ y \end{bmatrix}$ to $x \begin{bmatrix} 0 \\ 1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$.

Thus in coordinates, R_{90° is represented by the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (the images of $\vec{i}$ and $\vec{j}$ in coordinates as columns), and this motivates the matrix-vector multiplication $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$.

In general for a linear transformation L such that in coordinates $L(\vec{e_1}) = \begin{bmatrix} a \\ c \end{bmatrix}$ and $L(\vec{e_2}) = \begin{bmatrix} b \\ d \end{bmatrix}$, the matrix representing L is $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and if $\vec{v} = x\vec{i} + y\vec{j}$, $L(\vec{v}) = xL(\vec{e_1}) + yL(\vec{e_2})$ is given in coordinates by the matrix-vector multiplication $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$.

Using the operations in coordinates as motivation, the example of vectors in the plane and their linear transformations generalizes beyond visualization to the vector space (over the field F) $F^n = \{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mid \text{each } x_i \in F\}$, which has coordinate-wise addition and scalar multiplication as before. Again, we have a **standard basis** $\{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}\}$, where $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\ldots$, and $\vec{e_n} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$. Hence, any vector $\vec{v}$ in F^n can be uniquely written as $\vec{v} = x_1\vec{e_1} + x_2\vec{e_2} + \cdots + x_n\vec{e_n}$

In particular, any linear transformation $L : F^n \to F^m$ can be represented by a $m \times n$ matrix,

whose n columns are given by $L(\vec{e_1}), L(\vec{e_2}), \ldots, L(\vec{e_n})$, each $L(\vec{e_i})$ expressed as a column matrix with respect to the standard basis for F^m. In this way, $L(\vec{v})$ can be computed in coordinates as a matrix-vector product.

In fact, the set of all linear transformations $L : V \to W$ itself forms a vector space! We denote by this space $\mathcal{L}(V, W)$, with addition and scalar multiplication of linear transformations defined point-wise: $(L_1 + L_2)(\vec{v}) = L_1(\vec{v}) + L_2(\vec{v})$, and $(cL)(\vec{v}) = cL(\vec{v})$ (both of which can be verified to satisfy the properties of linear transformations). For $V = F^n$ and $W = F^m$, with respect to the standard bases the linear transformations in $\mathcal{L}(F^n, F^m)$ are in a one-to-one correspondence with the set of $m \times n$ matrices representing the transformations, denoted by $M_{m\times n}(F)$. Hence, we can see that the matrices representing the addition of transformations is the component-wise addition of the corresponding matrices, as taught in linear algebra! Similarly, the matrix representing a scalar multiplied by a transformation is just multiplying that scalar to every entry of the matrix representing that transformation. Thus, the set of matrices $M_{m\times n}(F)$ is also a vector space, isomorphic to $\mathcal{L}(F^n, F^m)$.

As we have seen, the usual formulas for matrix algebra are not arbitrary, but represent the operations of linear transformations. But what about the somewhat unusual definition of matrix multiplication? As it turns out, the product of two matrices represents the composition of the corresponding linear transformations, which one can check is also a linear transformation. Specifically, if $L_1 : F^n \to F^m$ and $L_2 : F^m \to F^p$ are linear transformations represented by the matrices A_1 (size $m \times n$) and A_2 (size $p \times m$), respectively, then the matrix A_2A_1 (size $p \times n$) represents the composition of the linear transformations $L_2 \circ L_1 : F^n \to F^p$.

In the case that two linear transformations are represented by square matrices, i.e. $L_1, L_2 : F^n \to F^n$, their compositions $L_1 \circ L_2$ and $L_2 \circ L_1$ always well-defined. Hence, $\mathcal{L}(F^n, F^n)$ has composition as a natural multiplication operation. Furthermore, there is a multiplicative identity, the identity map, which in the standard basis is represented by the identity matrix, so $\mathcal{L}(F^n, F^n)$ has the structure of a non-commutative, associative F-algebra with identity. Finally, if a linear transformation $L : F^n \to F^n$ is an isomorphism (which is certainly not always the case), then there

exists an inverse linear transformation $L^{-1} : F^n \to F^n$, whose corresponding matrix is called the inverse of the matrix representing L (they multiply both ways to give the identity matrix). Focusing on invertible linear transformations only, since the composition of invertible linear transformations is an invertible linear transformation, we thus obtain a group, which when represented by matrices is called the **general linear group**, and denoted by $\mathrm{GL}(n, F)$

Central to our discussion of vectors and linear transformations (as well as their operations) in coordinates was the standard basis $\{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}\}$ in F^n. As it turns out, every vector space V has a basis (in fact many bases), which (if the vector space is finite-dimensional), allows one to do linear algebra in coordinates:

Definition 1.6.3. *Let V be a vector space over a field F. A subset of vectors $S \subseteq V$ is **linearly dependent** if there exists a linear relation among them, i.e. there exist $a_1, \ldots, a_k \in F$ (not all 0) and $v_1, \ldots, v_k \in V$ such that $a_1 v_1 + \cdots + a_k v_k = 0$. A subset $S \subseteq V$ is **linearly independent** if it is not linearly independent, i.e. if $a_1 v_1 + \cdots + a_k v_k = 0 \implies a_1 = a_2 = \cdots = a_k = 0$ for all $v_1, \ldots, v_k \in S$. A subset $S \subseteq V$ **spans** V if for all $v \in V$, there exist $c_1, \ldots, c_k \in F$ and $v_1, \ldots, v_k \in V$ such that $v = c_1 v_1 + \cdots + c_k v_k$. A subset $S \subseteq V$ is a **basis** for V if it is linearly independent and spans V. If $S = \{v_1, \ldots, v_n\} \subseteq V$ is a finite basis, then we say that V is **finite-dimensional**, and in particular that V has **dimension** n (the size of the basis). We say that V is **infinite-dimensional** if no finite basis for V exists.*

One can show that the dimension of a finite-dimensional vector space is well-defined, or in other words that if V has dimension n, every basis of V has cardinality n. In fact, if V is a vector space over a field F of dimension n, then $V \cong F^n$. In this sense, choosing a basis for a finite-dimensional vector space is the same thing as choosing a coordinate system, resulting in all vectors, linear transformations, and their operations being able to be represented by matrices like F^n.

Linear independence of n vectors in F^n (and hence whether the vectors form a basis or not) can be established by the determinant.

Definition 1.6.4. *A k-**multilinear map** from a product of vector spaces $V_1, \ldots, V_k$ to a vector space W is a function $L : V_1 \times \cdots \times V_k \to W$ that is linear in each component separately. In*

other words, fixing elements $v_2 \in V_2, \ldots, v_k \in V_k$, the map $L_1 : V_1 \to W$ given by $L_1(v) = L(v, v_2, \ldots, v_n)$ is a linear transformation (and the same is true for the other coordinates). A k-multilinear map $L : V^k \to W$ from the product of a vector space V with itself k times, denoted V^k, to a vector space W is said to be ***alternating*** *if for any permutation $\sigma \in S_k$, $L(v_1, \ldots, v_n) = \mathrm{Sign}(\sigma) L(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$. The determinant is the unique alternating n-multilinear map $\det : (F^n)^n \to F$ that takes the standard basis for F^n (in the usual ordering) to 1.*

In coordinates, $(F^n)^n$ is the same thing as n column vectors put side to side into an $n \times n$ matrix, and so in coordinates we can consider the determinant of a matrix $A = (a_{ij})$ to be $\det(A) = \sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$. While this formula quickly becomes unwieldy for large n, and so there are many other algorithms for computing the determinant of a matrix, this formulation of the determinant reveals that it is a polynomial in the entries of A. Furthermore, a matrix is invertible if and only if its determinant is non-zero. The determinant is additionally a group homomorphism $GL(n, F) \to F^*$, since for any two matrices A and B, $\det(AB) = \det(A)\det(B)$, so if both $\det(A), \det(B) \neq 0$, $\det(AB) \neq 0$.

Geometrically over $\mathbb{R}$, the determinant of two vectors in the plane (together, a 2×2 matrix), is the signed (dependent on orientation) area of the parallelogram that the vectors form. Hence, the determinant being 0 corresponds to the area being 0, which only can occur when the vectors are scalar multiples of one another, or in other words they are linearly dependent. In 3-dimensions, the determinant of three vectors is the signed volume of the parallelepiped that they form. Again, if all three vectors are scalar multiples of one another (together they only span a 1-dimensional subspace), then again the determinant is 0 showing linear dependence, but this also occurs for example if the third vector lies in the plane spanned by the first two vectors (together they only span a 2-dimensional subspace), since volume is a 3-dimensional concept and so the volume of a parallelogram is still considered to be 0 (even though it may have finite area). This generalizes to determinants being hypervolumes of generalized parallelepipeds in $\mathbb{R}^n$, and is the motivation for the formula for determinants over any field.

While we do not focus on infinite-dimensional vector spaces, there are many important such

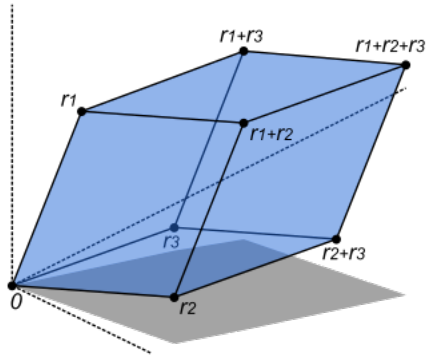


Figure 1.9: The Determinant in $\mathbb{R}^3$ is the Volume of a Parallelepiped

spaces. For example, if we ignore polynomial multiplication and instead only allow multiplication by elements of F, the abelian group $F[x]$ is a vector space with basis $\{1, x, x^2, \ldots\}$. If we consider all three operations at the same time, $F[x]$ is called an F-algebra. Additionally, the sets $C^0(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid \text{f is continuous}\}$ and $C^1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid \text{f is continuously differentiable}\}$ are essential real infinite-dimensional vector spaces (operations defined point-wise) in analysis. With this terminology, the derivative $\frac{d}{dx} : C^1(\mathbb{R}) \to C^0(\mathbb{R})$ and definite integral $\int_a^b : C^0(\mathbb{R}) \to \mathbb{R}$ are familiar examples of linear transformations from calculus, which cannot be represented as matrices due to infinite-dimensionality.

As with groups, rings, and fields, we again have the notion of a sub-vector space of a vector space V, called a **subspace** of V, which is a subset $U \subseteq V$ that is itself a vector space (under the restricted operations from V). Note that $\{0\}$ and V are always subspaces of V. If V has dimension n, then any subspace has dimension less than or equal to n (with equality only achieved for the subspace V of itself). For example, in $\mathbb{R}^3$, the origin is the trivial subspace, the 1-dimensional subspaces are lines through the origin (each an isomorphic copy of $\mathbb{R}$), and the 2-dimensional subspaces are planes through the origin (each an isomorphic copy of $\mathbb{R}^2$), and the only 3-dimensional subspace is $\mathbb{R}^3$ itself. We will often refer to a k-dimensional subspace of a vector space V a **k-plane**.

Let V be a vector space over F, and let $W_1, W_2 \subseteq V$ be subspaces. Then, some important subspaces we can obtain are $W_1 \cap W_2 \subseteq V$ and $W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\} \subseteq V$.

If $W_1 \cap W_2 = \{\vec{0}\}$, we further write $W_1 \bigoplus W_2$ for $W_1 + W_2$, and call this the **(internal)** direct sum of W_1 and W_2. Note that $W_1 \cup W_2 \subseteq V$ is not a subspace. Now, if we have a linear transformation $L : V \to W$, we also get important subspaces $N(L) = \{v \in V \mid L(v) = \vec{0}\} \subseteq V$, called the **null space** or **kernel** of L, and $R(L) = L(V) \subseteq W$ is the **image** of L (also called the **range** of L in elementary linear algebra courses). The dimensions $\dim(R(L))$ and $\dim(N(L))$ are called the **rank** and **nullity** of L, respectively, and these terms are also applied to any matrix representing L in coordinates.

In addition to subspaces, we have quotients of vector spaces by subspaces, which is very similar to the quotients of rings by ideals we considered earlier. Using these quotients, we get the Isomorphism Theorems. Often one encounters such theorems for the first time in the context of groups or rings, but for vector spaces, by taking dimensions we recover some important facts about the dimensions of certain subspaces.

Definition 1.6.5. *Let V be a vector space over a field F, and let $M \subseteq V$ be a subspace. We define an equivalence relation $\sim$ on V by $v_1 \sim v_2$ if $v_1 - v_2 \in M$, and denote by $v + M$ the equivalence class of $v \in V$. The **quotient space** V/M is the set of equivalence classes $\{v + M \mid v \in V\}$, which is a vector space over F, equipped with operations $(v_1 + M) + (v_2 + M) = (v_1 + v_2) + M$ and $c(v + M) = (cv) + M$ for all $v, v_1, v_2 \in V$ and $c \in F$. In this quotient vector space, the zero vector is $\vec{0} + M$, and the additive inverse of $v + M$ is $(-v) + M$. If V and M are finite-dimensional, then the dimension $\dim(V/M) = \dim(V) - \dim(M)$.*

One way to visualize a quotient space V/M is as the set of translates of the subspace M. For example, if M is a 1-dimensional subspace of $\mathbb{R}^2$ (so a line through the origin), the elements of V/M can be visualized as all of the lines in $\mathbb{R}^2$ parallel to M. Adding two lines together amounts to adding any vectors from each of the lines together, with the resulting vector lying on the line that we call the sum of the two lines. The same logic applies to scalar multiplication. For this example, $\dim(V/M) = \dim(\mathbb{R}^2) - \dim(M) = 2 - 1 = 1$, so again we can think of the quotient as isomorphic to a line in $\mathbb{R}^2$ itself. One can find such an isomorphism by choosing any 1-dimensional subspace W of $\mathbb{R}^2$ that is not M (line through the origin that isn't M), and identifying every line

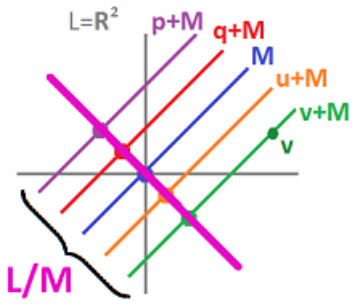


Figure 1.10: Visualizing a Quotient Space V/M in $\mathbb{R}^2$

in V/M with its point of intersection with W. This idea is deep, because what we are saying is that we have a geometric object with structure (a vector space), and partitioning that object into sub-geometric objects of interest (lines) yields together another geometric object with the same structure (the quotient space). This is the beginnings of what are called **parameter spaces**, or **moduli spaces**, with W being the parameter space in our example. We will see other essential moduli spaces throughout this dissertation, such as projective spaces, Grassmannians, and flag manifolds.

Theorem 1.6.6. *(The First Isomorphism Theorem and the Rank-Nullity Theorem) Let $L : V \to W$ be a linear transformation of vector spaces over a field F. Then, $V/N(L) \cong R(L)$. Furthermore,* $\dim(V) = \dim(N(L)) + \dim(R(L))$

Proof. Define $\tilde{L} : V/N(L) \to R(L)$ by $\tilde{L}(v + N(L)) = L(v)$. One can show that this is a well-defined bijective linear transformation, giving the desired isomorphism. The last result follows from $\dim(V/N(L)) = \dim(V) - \dim(N(L))$, and that vector space isomorphisms preserve dimension. □

One important consequence of Theorem 1.6.6 occurs when solving a differential equation (or system of equations) $D(f) = g$, D being the linear differential operator, and f and g being functions (or vectors of functions). One can first solve the corresponding homogeneous differential equation $D(f) = \vec{0}$, which is by definition finding the null space $N(D)$. Then, since

$V/N(D) \cong R(D)$, as long as $g \in R(D)$, a general solution is of the form $h + N(D)$, where h is a particular solution, but any $\tilde{h}$ with $\tilde{h} - h \in N(D)$ will do.

Theorem 1.6.7. *(The Second Isomorphism Theorem and Intersections of Subspaces) Let $W_1, W_2 \subseteq V$ be subspaces of a vector space V over a field F. Then, viewing $W_2 = \{\vec{0} + w_2 \mid w_2 \in W_2\} \subseteq W_1 + W_2$, we have $(W_1 + W_2)/W_2 \cong W_1/(W_1 \cap W_2)$ As a special case, $\dim(W_1 \bigoplus W_2) = \dim(W_1) + \dim(W_2)$. Furthermore, $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$. In particular, $\dim(W_1 \cap W_2) \geq \dim(W_1) + \dim(W_2) - \dim(V)$.*

Proof. Define $L : W_1 + W_2 \to W_1/(W_1 \cap W_2)$ by $L(w_1 + w_2) = w_1 + (W_1 \cap W_2)$. One can show that this is a surjective linear transformation (in particular $R(L) = W_1/(W_1 \cap W_2)$), and that the null space $N(L) = W_2$, so by the first isomorphism theorem, $(W_1 + W_2)/W_2 \cong W_1/(W_1 \cap W_2)$. The other results involving dimension come from $\dim((W_1 + W_2)/W_2) = \dim(W_1/(W_1 \cap W_2)) \implies \dim(W_1 + W_2) - \dim(W_2) = \dim(W_1) - \dim(W_1 \cap W_2)$, and rearranging. The last statement uses $\dim(W_1 + W_2) \leq \dim(V)$, so $\dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 + W_2) \geq \dim(W_1) + \dim(W_2) - \dim(V)$. □

Recall that a linear system might not have solutions due to the corresponding linear spaces not intersecting (like parallel lines in the plane). One can guarantee non-trivial intersections of subspaces by making sure that the intersection subspace has dimension greater than 0. For example (and an example that is crucial at the beginning of the next section), consider two planes H_1 and H_2 through the origin in $\mathbb{C}^3$ (2-dimensional subspaces of $\mathbb{C}^3$). These planes must intersect non-trivially, since by our theorem, $\dim(H_1 \cap H_2) \geq \dim(H_1) + \dim(H_2) - \dim(V) = 2 + 2 - 3 = 1 > 0$.

Now, we connect abstract vector spaces to our discussion of field extensions and Galois groups from the previous section. Since every field is a 1-dimensional vector space over itself (F^n with $n = 1$), every field can be viewed as a subspace of its field extensions K/F, where K is viewed as a vector space over F (since one just restricts which scalars are permitted for scalar multiplication). For example, the field extension $\mathbb{C}/\mathbb{R}$ allows us to view $\mathbb{C}$ as a 2-dimensional real vector space, and in fact our definition $\mathbb{C} = \mathbb{R}(i) = \{a + bi \mid a, b \in \mathbb{R}\}$ describes $\mathbb{C}$ in this way

with $\{1, i\}$ as the real basis. Similarly from earlier examples, $\mathbb{Q}(\omega)$ has $\{1, \omega\}$ and $\mathbb{Q}(\sqrt[3]{2}, \omega)$ has $\{1, \omega, w^2, \sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}\}$ as bases over $\mathbb{Q}$. Note that the dimensions of these fields over $\mathbb{Q}$ are equal to the orders of their Galois groups $\mathbb{Z}/2\mathbb{Z}$ and S_3, respectively. This is not a coincidence!

Definition 1.6.8. *Let K/F be a field extension. The **degree** of the extension is the dimension of the vector space K over F, and is denoted $[K : F]$. If $[K : F] < \infty$, we say that K/F is a **finite field extension**. Note that $[K : F] = 1$ if and only if $K = F$, so if $[K : F] > 1$ is finite, we say K/F is a **proper** finite extension. If E is an intermediate field of K/F, then $[K : F] = [K : E][E : F]$.*

Proposition 1.6.9. *Finite field extensions K/F are always algebraic*

Proof. Let $\alpha \in K$, and consider the set $S = \{1, \alpha, \alpha^2, \ldots, \alpha^n\}$, where $[K : F] = n < \infty$ (by assumption since a finite extension). Since $S \subseteq K$ has $n + 1$ elements, it must be a linearly dependent set, so there exist $c_0, \ldots, c_n \in F$ (not all 0) such that $c_n\alpha^n + \cdots + c_1\alpha + c_0 = 0$. Hence, $\alpha \in K$ is a root of the polynomial $c_nx^n + \cdots + c_1x + c_0 \in F[x]$, so α is algebraic over F. Since $\alpha \in K$, was arbitrary, K/F is an algebraic extension. □

For a field extension, $\mathrm{Aut}(K/F) = \{\text{field isomorphisms } \varphi : K \to K \mid \varphi(x) = x \text{ for all } x \in F\}$ always makes sense (and is called the **automorphism group** of K/F), regardless of whether the extension is Galois or not. However, one of the many reasons that this group contains information about the intermediate fields of K/F (the Fundamental Theorem of Galois Theory) is because $[K : F] = |\mathrm{Aut}(K/F)|$ if and only if K/F is a Galois extension. In fact, we could have defined Galois extensions equivalently this way. Additionally we have a relation between degrees of sub-extensions and orders of subgroups as well, consistent with the inclusion-reversing property. Specifically, in the statement of the Fundamental Theorem of Galois Theory, we can now add that if K/F is a Galois extension, and if $H \subseteq G$ is a subgroup, then $|H| = [K : K^H]$ and $\frac{|Gal(K/F)|}{|H|} = [K^H : F]$. We can now use the theorem to prove that $\mathbb{C}$ is an algebraically closed field. Here we also need a few basic facts from analysis, since the construction of $\mathbb{C}$ is as an algebraic extension of $\mathbb{R}$, which was not a purely algebraic extension of $\mathbb{Q}$.

Theorem 1.6.10. *(The Fundamental Theorem of Algebra) The set $\mathbb{C}$ of complex numbers is an algebraically closed field (meaning $\overline{\mathbb{C}} = \mathbb{C}$).*

Proof. It suffices to show that $\mathbb{C}$ has no proper finite field extension, since if there were any proper algebraic extension $L/\mathbb{C}$, choosing any $\alpha \in L$, α is algebraic so $\mathbb{C}(\alpha)/\mathbb{C}$ would be a finite extension (with degree the degree of the minimal polynomial of α). Hence, assume $K/\mathbb{C}$ is a finite extension (we show $[K : \mathbb{C}] = 1$, so $K = \mathbb{C}$. Since the normal closure of K over $\mathbb{R}$ still has a finite degree over $\mathbb{C}$ (or $\mathbb{R}$), we may further assume without loss of generality that $K/\mathbb{R}$ is normal (and so a Galois extension, since extensions over characteristic 0 fields are always separable).

Let $G = \mathrm{Gal}(K/\mathbb{R})$. Since $K \subseteq \mathbb{C} \subseteq \mathbb{R}$, $|G| = [K : \mathbb{R}] = [K : \mathbb{C}][\mathbb{C} : \mathbb{R}] = 2[K : \mathbb{C}]$. By the First Sylow Theorem, G has a Sylow 2-subgroup H (a maximal subgroup with order a power of 2), and so $\frac{|G|}{|H|}$ is odd. By the Fundamental Theorem of Galois Theory, there is an intermediate field E of $K/\mathbb{R}$ such that $\mathrm{Gal}(K/E) = H$. As $[E : \mathbb{R}] = \frac{|G|}{|H|}$ is odd, and by the Intermediate Value Theorem every odd degree polynomial has at least one real root (so is not irreducible unless degree 1), it must be that $E = \mathbb{R}$ ($[E : \mathbb{R}] = 1$). Thus, $[K : \mathbb{R}]$ and $[K : \mathbb{C}]$ must be powers of 2. By way of contradiction, assume $[K : \mathbb{C}] > 1$. Then, G is a 2-group and so must contain a subgroup H' such that $\frac{|G|}{|H'|} = 2$, giving by the Fundamental Theorem of Galois Theory an intermediate field E' of $K/\mathbb{C}$ with $[E' : \mathbb{C}] = 2$, which is a contradiction since every element $z \in \mathbb{C}$ has a complex square root $\sqrt{z} \in \mathbb{C}$ (so quadratics are all reducible by the quadratic formula). Therefore, it must be that $[K : \mathbb{C}] = 1$, so $K = \mathbb{C}$, and $\mathbb{C}$ has no proper finite extensions (hence no algebraic extensions by our above discussion), so $\mathbb{C}$ is algebraically closed. □

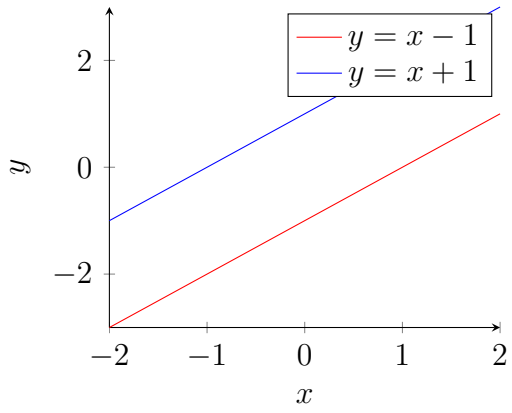
1.7 Projective Space and Polynomial Equations

Returning to our discussion of “expanding our world" to solve increasingly difficult equations, recall that the axioms of a vector space allow one to fully solve systems of linear equations under a very strong condition - the solutions have to exist! Geometrically, over $\mathbb{R}$, even a 2×2 system might have no solution, corresponding to parallel lines that do not intersect. The existence of parallel lines (the “parallel postulate") is one of the five (and the most controversial) axiom of

Euclidean geometry in the plane, as presented in Euclid's The Elements. However, experience tells us that we do not live in a flat plane, but on the surface of an approximate sphere that has curvature (sorry Flat Earthers!). Hence, seemingly parallel lines (look closely at the border of Canadian prairie provinces) might begin to converge, and the shortest distance between two locations might not be the "straight line" between them (hence why pilots use great circle routes to optimize travel times). The theme here is that while the ground beneath us might look like a plane, if we zoom out and have a higher perspective, we notice that reality involves more complicated and interesting geometry. This is the central notion of a manifold - a geometric object which locally looks like Euclidean space $\mathbb{R}^n$, but which likely has more interesting geometry globally. We first introduce another example of a manifold, the projective plane, where parallel lines intersect (different than the spherical geometry informally described above), and then describe manifolds and projective space in general.

Consider the system of equations corresponding to the parallel lines $y = x - 1$ and $y = x + 1$ in $\mathbb{C}^2$ (note that we work over the algebraically closed field $\mathbb{C}$, but our pictures are all over $\mathbb{R}$):

$$\begin{cases} x - y - 1 = 0 \\ x - y + 1 = 0 \end{cases} \tag{1.4}$$



But what if these weren't *actually* lines? What if in our limited perspective, we don't see the whole picture, but only a projection of a larger reality? This is the fundamental idea of projective

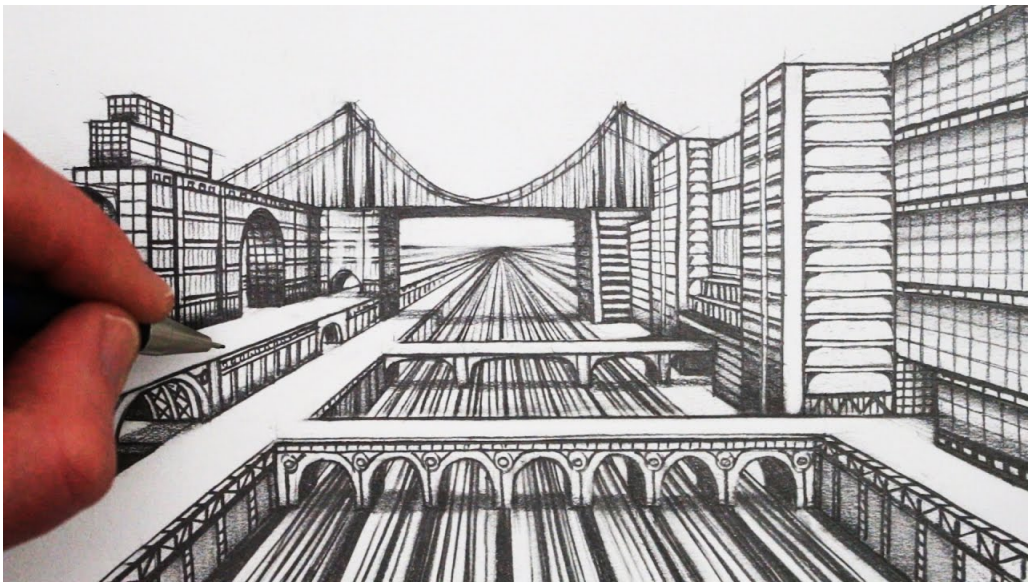


Figure 1.11: Example of Perspective Art

geometry, inspired by perspective art, where lines appear to intersect at a “line at infinity", visually seen as a horizon.

To make this precise, consider that the lines $x - y - 1 = 0$ and $x - y + 1 = 0$ in the plane $\mathbb{C}^2$ are actually the projections of the planes $x - y - z = 0$ and $x - y + z = 0$, respectively, onto the $z = 1$ plane in space $\mathbb{C}^3$. Note that these planes pass through the origin (i.e. are 2-dimensional subspaces of $\mathbb{C}^3$), and so their intersection must be a subspace of dimension 1 - the line through the origin in $\mathbb{C}^3$ given implicitly by the two equations $z = 0, x - y = 0$ (see Figure 1.10). This line does not intersect the plane we originally viewed the two lines as being parallel in, and so while the planes do intersect, the projection that we started with “missed" the intersection. However, if we take another projection, say to the $y = 1$ plane in $\mathbb{C}^3$, the equation of our planes become the lines $z = x - 1$ and $z = -x + 1$, which intersect at the point $(1, 1, 0) \in \mathbb{C}^3$. In other words, we can interpret this as having looked at the seemingly parallel lines from the wrong perspective, but in changing our view we see where they in fact meet. The “world" where such solutions to 2×2 linear systems always exist is called the (complex) projective plane $\mathbb{P}^2$.

Definition 1.7.1. *The **complex projective plane** $\mathbb{P}^2 = \{$lines through the origin in $\mathbb{C}^3\}$. In other*

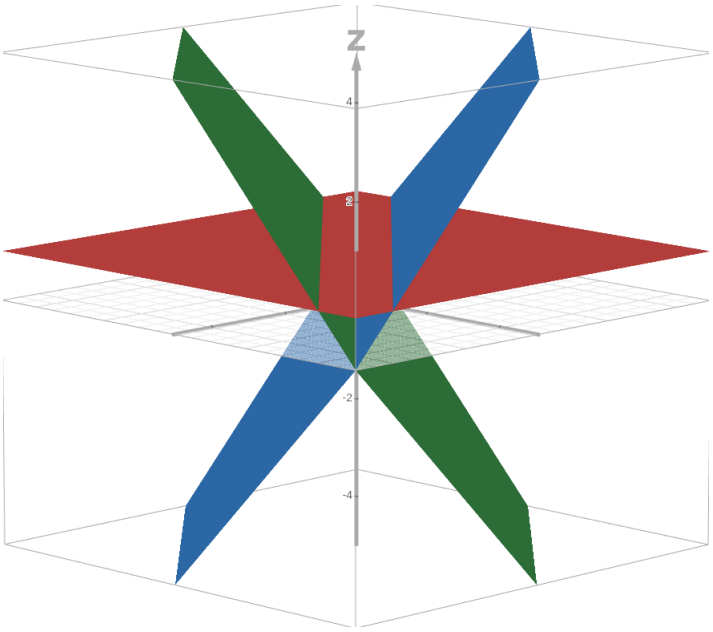


Figure 1.12: Parallel lines are actually a projection of intersecting planes

words, $\mathbb{P}^2$ *is the set of* 1*-dimensional linear subspaces of* $\mathbb{C}^3$.

Hence, points in $\mathbb{P}^2$ are what we traditionally view as lines through the origin in $\mathbb{C}^3$, and so lines in $\mathbb{P}^2$ are what we traditionally view as planes through the origin (lines of lines) in $\mathbb{C}^3$. As a result, any two distinct lines in $\mathbb{P}^2$ (planes through the origin in $\mathbb{C}^3$) intersect in a point in $\mathbb{P}^2$ (line through the origin in $\mathbb{C}^3$), so 2×2 linear systems always have solutions in $\mathbb{P}^2$!

So far we've discussed $\mathbb{P}^2$ without coordinates, but for any point in $\mathbb{P}^2$, we can choose a basis for the corresponding line through the origin in $\mathbb{C}^3$, giving a representative for that point. Since 1-dimensional, any other choice of basis element (representative for the point) for the line will only differ by a non-zero scalar multiple. Hence, we obtain an equivalent definition of $\mathbb{P}^2$ as a set of equivalence classes of non-zero vectors in $\mathbb{C}^3$: $\mathbb{P}^2 = (\mathbb{C}^3 \setminus \{\vec{0}\})/\sim$, where $\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \sim \lambda \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ for all $\lambda \in \mathbb{C} \setminus \{0\}$. In other words, we represent the elements of $\mathbb{P}^2$ by vectors in $\mathbb{C}^3$, understanding that the point in $\mathbb{P}^2$ is really the span of the vector representing it, and we call such a representation **homogeneous coordinates** for that point.

Whenever one is working with equivalence classes, it is often helpful to have certain representatives of the equivalence class that are easier to work with, and these are often called “normal forms" or “canonical forms". The rational numbers $\mathbb{Q}$ are such a set of equivalence classes, since while $\frac{3}{7}$ and $\frac{18}{42}$ are the same number (have the same equivalence class), the reduced fraction $\frac{3}{7}$ (the normal form for the class) is often preferred. However, it is beneficial to not always have to work with normal forms, since $\frac{18}{42}$ would be a better representation to use when computing $\frac{3}{7} + \frac{1}{42}$, for example.

Similar to working with other sets of equivalence classes, such as the rational numbers, for the projective plane we have choices of local coordinates for points that are often preferable over homogeneous coordinates to work with. For a point $\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{P}^2$, we know that at least one of $X, Y, Z \in \mathbb{C}$ are non-zero (since otherwise you have the 0-vector, which does not span a 1-dimensional subspace). Hence, if $Z \neq 0$, one can obtain another representative (different choice of basis for the line) by multiplying by $\lambda = \frac{1}{Z}$ to obtain local coordinates $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$, where $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$. The set of such points is called the **affine chart** $\mathbb{A}^2_Z = \{ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{P}^2 \mid Z \neq 0\}$, and is the set of lines in $\mathbb{C}^3$ not contained in the xy-plane. Such lines must intersect the plane $z = 1$ in $\mathbb{C}^3$ and they can be identified with this point of intersection $(x, y, 1)$. In other words, we want to say that $\mathbb{A}^2_Z$ is isomorphic (and we will soon make this precise) to $\mathbb{C}^2$. Similarly, we get **affine charts** $\mathbb{A}^2_X$ (points where $X \neq 0$) and $\mathbb{A}^2_Y$ (points where $Y \neq 0$), both isomorphic to $\mathbb{C}^2$, and $\mathbb{P}^2 = \mathbb{A}^2_X \cup \mathbb{A}^2_Y \cup \mathbb{A}^2_Z$ (with great overlap between these charts). In our motivating example, the parallel lines we considered were in $\mathbb{A}^2_Z \subseteq \mathbb{P}^2$, but their intersection, while in $\mathbb{P}^2$, was not in the

affine chart $\mathbb{A}^2_Z$. Instead, the intersection was the point $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{P}^2$ (notice $Z = 0$), which does belong to the other affine charts $\mathbb{A}^2_X$ and $\mathbb{A}^2_Y$ in $\mathbb{P}^2$. The affine charts endow $\mathbb{P}^2$ with the structure of a manifold, which is a special type of topological space. We define both now, and use the notation $\mathcal{P}(X)$ for the power set of a set X (the set of all subsets of X).

Definition 1.7.2. *A **topological space** is a set X, equipped with a collection of subsets $\tau \subseteq \mathcal{P}(X)$ satisfying:*

1. $\emptyset, X \in \tau$
2. *τ is closed under finite intersections: If $U_1, \ldots, U_n \in \tau$, then $U_1 \cap \cdots \cap U_n \in \tau$*
3. *τ is closed under arbitrary unions: If Λ is an arbitrary index set, and $\{U_\lambda \mid \lambda \in \Lambda\} \subseteq \tau$, then $\bigcup_{\lambda \in \Lambda} U_\lambda \in \tau$*

*We call τ a **topology** on X, and sometimes write (X, τ) to indicate that X is endowed with the specific topological structure given by τ. Elements of τ are called the **open sets** in X, and the complements of elements of τ are called the **closed sets** in X (both with respect to τ). Note that $\emptyset^c = X$ and vice-versa, so sets like these that are open and closed are referred to as **clopen**.*

A standard example of a topological space is $\mathbb{R}$ with the Euclidean topology: open sets are unions of open intervals $(c - r, c + r)$ (where $c \in \mathbb{R}$ and $r \in \mathbb{R}_{>0}$). This example generalizes to $\mathbb{R}^n$, $\mathbb{C}^n$ or any metric space (X, d) where the open sets are unions of open balls $B_r(x) = \{p \in X \mid d(x, p) < r\}$ (where $r \in \mathbb{R}_{>0}$ and $x \in X$). Since any open set is defined to be a union of these open balls, and for any point in the intersection of two open balls, there is a third open ball containing that point and contained in the other two, we say that they form a **basis** for the the topology and are called **basic open sets**. Note that this definition is different than a basis for a vector space, but this should always be clear from context.

Like the algebraic structures we've introduced so far, we also are interested in the functions between topological spaces that preserve the structure. In topology, these are continuous functions.

Definition 1.7.3. *Let (X_1, τ_1) and (X_2, τ_2) be topological spaces. A **continuous map** is a function $f : X_1 \to X_2$ such that for every open set $U \in \tau_2$ in X_2, its preimage $f^{-1}(U) \in \tau_1$ is open in X_1. Similarly, we say f is continuous at a point $c \in X_1$ if for every open set $U \in \tau_2$ containing $f(c)$, there exists an open set $V \in \tau_1$ containing c such that $f^{-1}(U) \subseteq V$. The statement that f is continuous is equivalent to being continuous at every point $c \in X_1$. If the topologies have bases B_1 and B_2, respectively, continuity can be established just by considering the basic open sets. A bijective continuous map $f : X_1 \to X_2$, whose inverse $f^{-1} : X_2 \to X_1$ is also continuous, is called a **homeomorphism**, in which case we say that the topological spaces (X_1, τ_1) and (X_2, τ_2) are **homeomorphic**.*

Considering functions $f : \mathbb{R} \to \mathbb{R}$ and basic open sets (in this case, intervals), the statement $f^{-1}((c - \delta, c + \delta)) \subseteq (f(c) - \epsilon, f(c) + \epsilon)$ is equivalent to the usual definition from calculus: f is continuous at $c \in \mathbb{R}$ if for all $\epsilon > 0$, there exists $\delta > 0$ such that whenever $x \in \mathbb{R}$ satisfies $0 \leq |x - c| < \delta$, $|f(x) - f(c)| < \epsilon$.

Note that topological spaces were defined in terms of open sets, but equivalently they could have been defined in terms of closed sets via De Morgan's laws: $(\bigcup_{\lambda \in \Lambda} U_\lambda)^c = \bigcap_{\lambda \in \Lambda} U_\lambda^c$ and $(\bigcap_{\lambda \in \Lambda} U_\lambda)^c = \bigcup_{\lambda \in \Lambda} U_\lambda^c$. In other words, a topological space could also have been defined as a set X with a collection of closed sets containing $\emptyset$ and X, that is closed under finite unions and arbitrary intersections. Furthermore, a continuous map between topological spaces can equivalently be defined as a function where the preimages of closed sets are closed.

The topology on $\mathbb{P}^2$, called the **Zariski topology**, is best defined in terms of closed sets, which are defined to be intersections of projective plane curves. A **projective plane curve** is the zero set to a homogeneous (every term has the same total degree d) three-variable polynomial $F(X, Y, Z) = 0$, which $F : \mathbb{P}^2 \to \mathbb{C}$ is not well-defined (since by homogeneity $F(\lambda X, \lambda Y, \lambda Z) = \lambda^d F(X, Y, Z) \neq F(X, Y, Z)$ for all $\lambda \in \mathbb{C} \setminus \{0\}$), but its zero set $V(F) = \{ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{P}^2 \mid F(X, Y, Z) = 0 \}$ is well defined (since multiplication by λ^d doesn't change the solution set).

Hence, a basis for the open sets of the Zariski topology on $\mathbb{P}^2$ are $\{D_F \mid F(X,Y,Z) \text{ is homogeneous}\}$, where $D_F = V(F)^c = \{ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{P}^2 \mid F(X,Y,Z) \neq 0\}$. One can see that the three affine charts are special cases of such basic open sets: $\mathbb{A}^2_X = D_X$, $\mathbb{A}^2_Y = D_Y$, and $\mathbb{A}^2_Z = D_Z$.

Beyond that of a topological space, the affine charts make $\mathbb{P}^2$ into a manifold, which will be crucial for our future discussion.

Definition 1.7.4. *Let X be a topological space, and fix $n \in \mathbb{N}_{>0}$. A **local coordinate chart** on X is a homeomorphism $\varphi : U \to V$, where $U \subseteq X$ and $V \subseteq \mathbb{R}^n$ are each open in their respective topologies. In other words, φ^{-1} gives a parametrization of U by the open subset $V \subseteq \mathbb{R}^n$. Two local coordinate charts $\varphi : U_\alpha \to V_\alpha$ and $\varphi_\beta : U_\beta \to V_\beta$ are **compatible** if the map $\varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$ is a homeomorphism. An **atlas** on X is an open covering $X = \bigcup_{\lambda \in \Lambda} U_\lambda$ by mutually compatible local coordinate charts $U_\lambda \subseteq X$. There is an equivalence relation on the set of atlases on a topological space X, where two such atlases are considered equivalent if their union is an atlas as well (so the local coordinate charts in each are compatible with one another). A n-**dimensional topological manifold** is a topological space X, equipped with an equivalence class of atlases.*

One can obtain different types of manifolds by changing the compatibility condition in the definition above, leaving everything else the same. For example, if each $\varphi_\alpha \circ \varphi_\beta^{-1}$ is further required to be a diffeomorphism (it and its inverse are smooth, meaning derivatives of all orders exist), we say the topological manifold X is an n**-dimensional smooth manifold**. If one considers local coordinate charts where $V \subseteq \mathbb{C}^\kappa$ (topologically, homeomorphic to $\mathbb{R}^{2n}$), and if each $\varphi_\alpha \circ \varphi_\beta^{-1}$ is further required to be a bi-holomorphism (it and its inverse are holomorphic, meaning complex analytic), we say the topological manifold X is an n**-dimensional complex manifold** (and would also be a $2n$-dimensional smooth manifold).

For the projective plane $\mathbb{P}^2$, $\{\varphi_X : \mathbb{A}^2_X \to \mathbb{C}^2, \varphi_Y : \mathbb{A}^2_Y \to \mathbb{C}^2, \varphi_Z : \mathbb{A}^2_Z \to \mathbb{C}^2\}$ is an atlas making $\mathbb{P}^2$ into a 4-dimensional topological and smooth manifold, and a 2-dimensional complex

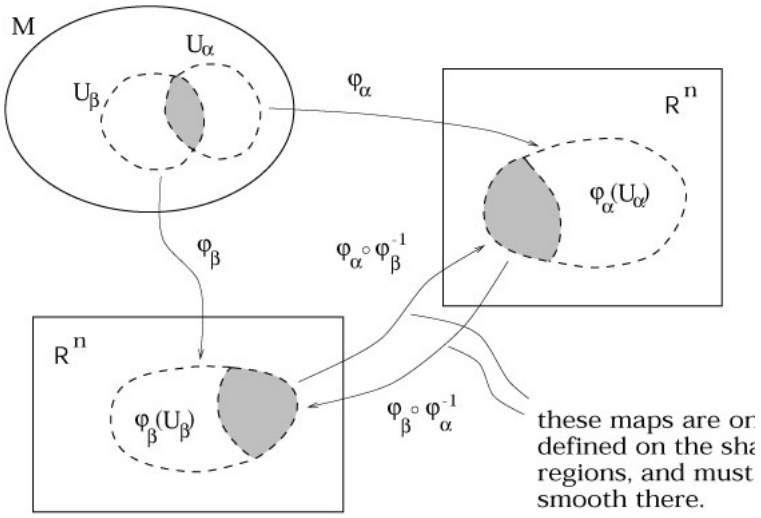


Figure 1.13: Compatible charts on a manifold M

manifold, where for example $\varphi_X(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}) = \begin{bmatrix} \frac{Y}{X} \\ \frac{Z}{X} \end{bmatrix}$, $\varphi_Y^{-1}(\begin{bmatrix} x \\ z \end{bmatrix}) = \begin{bmatrix} x \\ 1 \\ z \end{bmatrix}$, and so $\varphi_X \circ \varphi_Y^{-1}(\begin{bmatrix} x \\ z \end{bmatrix}) = \begin{bmatrix} \frac{1}{x} \\ \frac{z}{x} \end{bmatrix}$ (the other charts and compositions are similar). Note that $\varphi_X \circ \varphi_Y^{-1}$ is a rational function (quotients of polynomials) defined everywhere on the overlap, and so is a diffeomorphism and bi-holomorphism. We already demonstrated the usefulness of moving from coordinate chart to coordinate chart, allowing for solutions to 2×2 systems of linear equations when there are parallel lines. We now move to more general systems of linear, and in fact polynomial equations - which solutions are best found in what is called projective space. We will also describe how the projective plane contains all solutions to intersections of projective plane curves (a 2×2 polynomial system in 3 variables). First, we make the notion of multivariable polynomials precise.

Definition 1.7.5. *A **monomial** in the n variables $x_1, \ldots, x_n$ is $x_1^{a_1} \cdots x_n^{a_n}$, where $a_1, \ldots, a_n \in \mathbb{N}$. Note that the set of monomials in n variables is in one-to-one correspondence with $\mathbb{N}^n$, with $\alpha = (a_1, \ldots, a_n)$ being the **multi-degree** of the monomial, and often write x^α as a shorthand for the monomial. We say that the monomial has **total degree** $|\alpha| = a_1 + \cdots + a_n$. A n **variable***

polynomial over a field *F is a finite linear combination of monomials $f = \sum_{\alpha \in I} c_\alpha x^\alpha$, where I is a finite index set and each $c_\alpha \in F$. We denote by $F[x_1, \ldots, x_n]$ the set of all n variable polynomials over F, which is by construction the infinite-dimensional vector space over F with basis $\{x^\alpha \mid \alpha \in \mathbb{N}^n\}$. However, it has more structure as an F-algebra, and we often call $F[x_1, \ldots, x_n]$ a* ***polynomial ring*** *since polynomials can be multiplied by linearly extending the multiplication on monomials $x^\alpha x^\beta = x^{\alpha+\beta}$, where the addition in $\mathbb{N}^n$ is done coordinate-wise. If $S \subseteq F[x_1, \ldots, x_n]$ is a set of polynomials, the* ***ideal generated by*** *S is $I = (S) = \{$finite sums $\sum_\alpha h_\alpha f_\alpha \mid h_\alpha \in F[x_1, \ldots, x_n], f_\alpha \in S\}$.*

Definition 1.7.6. *Let F be a field.* ***Affine n-space over*** *F is $\mathbb{A}^n(F)$, which is the set F^n, ignoring the vector space structure, and with a certain topology, called the Zariski topology, which we now describe. For any subset $S \subseteq F[x_1, \ldots, x_n]$, an* ***affine variety*** *is the solution to the polynomial system $V(S) = \{p \in \mathbb{A}^n(F) \mid f(p) = 0 \text{ for all } f \in S\}$. One can see that if $I = (S)$ is the ideal generated by S, then $V(S) = V(I)$, so it suffices to consider systems of ideals of polynomial equations. It turns out that $F[x_1, \ldots, x_n]$ is a* ***Noetherian ring****, meaning that all of its ideals are finitely generated, so for all ideals $I \trianglelefteq F[x_1, \ldots, x_n]$ there exists finite $S' = \{f_1, \ldots, f_k\}$ such that $V(I) = V(S')$, and so we write $V(f_1, \ldots, f_k)$ for the variety of the polynomial system with k equations (which solves the entire ideal of equations). A variety generated by a single polynomial equation $V(f)$ is called a* ***hypersurface****, and note that $V(f_1, \ldots, f_k) = V(f_1) \cap \cdots V(f_k)$, so every affine variety is the intersection of finitely many hypersurfaces. The* ***Zariski topology*** *on $\mathbb{A}^n(F)$ is given by defining closed sets to be affine varieties, with basic open sets then given to be complements of hypersurfaces.*

Definition 1.7.7. *Let F be a field.* ***Projective n-space over*** *F is*
$\mathbb{P}^n(F) = \{$the set of 1-dimensional subspaces of $F^{n+1}\}$, equipped with its own Zariski topology. In ***homogeneous coordinates****, $\mathbb{P}^n(F) = (F^{n+1} \setminus \{\vec{0}\})/\sim$, where two vectors in F^{n+1} are equivalent if they are non-zero scalar multiples of one another (so they are both bases for the same 1-dimensional subspaces, just as for $\mathbb{P}^2$ earlier). A polynomial $f \in F[x_0, \ldots, x_n]$ is* ***homogeneous*** *of degree d if each of its monomials has the same total degree d. A* ***projective variety*** *is*

$V(S) = V(I) = V(f_1, \ldots, f_n)$ *as in the definition of an affine variety, but where all polynomials in* S *are required to be homogeneous. With hypersurfaces defined just as for affine space (but with a single homogeneous polynomial equation), the* ***Zariski topology*** *on* $\mathbb{P}^n(F)$ *is given by defining closed sets to be projective varieties, with basic open sets then given to be complements of hypersurfaces.*

Definition 1.7.8. *For either an affine or projective variety (which for short, we just call a* ***variety****)* $X = V(I)$*, a* ***subvariety*** *is a variety* Y *(of the same type) that is a subset* $Y \subset X$*. An* ***irreducible variety*** *is a variety such that if* $X = Y \cup Z$ *for subvarieties* $Y, Z \subseteq X$*, then* $X = Y$ *or* $X = Z$ *(it can't be written non-trivially as the union of two varieties). A variety that is not irreducible is called* ***reducible****. The* ***dimension*** *of a variety is the maximal length of chains of* $X_0 \subseteq X_1 \subseteq \cdots \subseteq X_d$ *of non-empty, distinct, and irreducible subvarieties of* X*. In this case, if a chain of length* d *is maximal, we say that* X *is* d*-dimensional. If* $Y \subseteq X$*, we call* $\dim_X(Y) = \dim(X) - \dim(Y)$ *the* ***codimension*** *of* Y *in* X*.*

Note that $\mathbb{P}^2 = \mathbb{P}^2(\mathbb{C})$, and in general we write $\mathbb{P}^n$ for $\mathbb{P}^n(\mathbb{C})$. In general, $\mathbb{P}^n$ with homogeneous coordinates $x_0, \ldots, x_n$ is an n-dimensional complex manifold with affine charts $\mathbb{A}^n_{x_i}$ where the x_i-coordinate is non-zero (and so has a normal form, where there is a 1 in that coordinate after re-scaling). The local coordinate charts $\varphi_{x_i} : \mathbb{A}^n_{x_i} \to \mathbb{C}^n$ are just the extension of those we defined for $\mathbb{P}^2$. Note that these maps still make sense from $\mathbb{P}^n(F)$ to $\mathbb{A}^n(F)$ for any field F, and that the transition functions $\varphi_{x_i} \circ \varphi_{x_j}^{-1}$ are still given by rational polynomial functions defined everywhere on the overlap, and so in general we call $\mathbb{P}^n(F)$ an **algebraic n-dimensional manifold over** F, since it is also an n-dimensional projective variety (the two notions of dimension coincide).

If one has more experience with analysis of $\mathbb{R}$, $\mathbb{R}^n$, or general metric spaces, they are familiar with the idea of compact subsets. In $\mathbb{R}^n$ a subset is compact if and only if it is closed (complement of an open set) and bounded (contained in some $B_x(r)$), and in metric spaces compact subsets must be complete and totally bounded. One can see that inherently, compactness is a topological property, and generalizes to topological spaces in general.

Definition 1.7.9. *Let* (X, τ) *be a topological space. A subset* $K \subseteq X$ *is* ***compact*** *if for every open*

cover of K (a collection of open subsets $\{U_\lambda \mid \lambda \in \Lambda\}$ such that $K \subseteq \bigcup_{\lambda\in\Lambda} U_\lambda$), there exists a finite subcover (a sub-collection $\{U_{\lambda_1}, \ldots, U_{\lambda_n}\}$ (each $\lambda_i \in \Lambda$) such that $K \subseteq \bigcup_{i=1}^{n} U_{\lambda_i}$.

Additionally, connectedness of a space is a property defined in terms of open sets: a topological space is **disconnected** if it is the union of two disjoint non-empty open sets, and is **connected** otherwise.

Compactness and connectedness are purely topological properties of a space, and so are preserved by continuous functions: If $f : X_1 \to X_2$ is continuous, and $K \subseteq X_1$ is a compact subset, then $f(K) \subseteq X_2$ is compact as well. If $f : X_1 \to X_2$ is continuous, and $C \subseteq X_1$ is a connected subset, then $f(C) \subseteq X_2$ is connected as well. Points are closed in $\mathbb{C}^n$ (topologically the same as $\mathbb{R}^{2n}$), and polynomial functions are continuous, so if $f : \mathbb{C}^{n+1} \to \mathbb{C}$ is $f(z_0, \ldots, z_n) = z_0^2 + \cdots + z_n^2$, $\mathbb{S}^n = f^{-1}(\{1\}) \subseteq \mathbb{C}^{n+1}$ is a closed subset, called the **complex n-sphere**. Since it is bounded by definition, it is a compact subset (note that it is also connected). There is a map $\mathbb{S}^n \to \mathbb{P}^n$ that takes a vector in $\mathbb{C}^{n+1}$ to the line it spans. This is surjective and continuous, showing that projective space is a compact manifold (and connected as well). In fact, projective space $\mathbb{P}^n$ can be viewed as a compactification of any of its affine charts $\mathbb{A}^n$, this compactification being a key reason why it has such great intersection theoretic properties.

In addition to being a compact manifold, projective space $\mathbb{P}^n$ additionally has the property that it is **paved (or stratified) by affines**. We call this the Schubert cell decomposition of $\mathbb{P}^n$, and such a decomposition holds true for any $\mathbb{P}^n(F)$, and will be central to future discussion. To define the Schubert decomposition, we must first introduce the notion of complete flags. The inspiration behind the name flag is a point on a line, contained in a plane in $\mathbb{R}^3$: the point on a line is the ball at the top of a flagpole, and attached to the pole is the flag itself, which is like a plane (see Figure 1.12).

Definition 1.7.10. *Let V be an n-dimensional vector space over a field F. Then, a **complete flag** in V, denoted $F_\bullet$, is a saturated chain of distinct subspaces of V, one for each dimension from 0 to n: $F_\bullet$ is the chain $\{0\} \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n = V$, where each $F_i \subseteq V$ is a subspace with $\dim(F_i) = i$. We denote by $\mathrm{Fl}(n)$ the set of all complete flags in $V = F^n$. Similarly,*

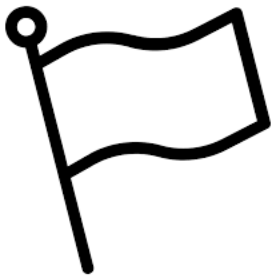


Figure 1.14: Cartoon of a Complete Flag

*a **partial flag in** V **of shape** $(a_1, \ldots, a_s)$, also denoted as $F_\bullet$, is a chain of distinct subspaces $\{0\} \subseteq F_{a_1} \subseteq F_{a_2} \subseteq \cdots \subseteq F_{a_s} = V$, where $\dim(F_{a_i}) = a_i$. We denote by $\mathrm{Fl}(a_1, \ldots, a_s; n)$ the set of partial flags in V of shape $(a_1, \ldots, a_s)$.*

If $\{\vec{e_1}, \ldots, \vec{e_n}\}$ is the standard basis for F^n, then one can obtain the **identity flag** $I_\bullet$ by defining $I_i = \mathrm{Span}\{\vec{e_1}, \ldots, \vec{e_i}\}$. Similarly, one can obtain the **opposite flag** $O_\bullet$ by defining $O_i = \mathrm{Span}\{\vec{e_n}, \vec{e}_{n-1}, \ldots, \vec{e}_{n-i+1}\}$. Once a flag $F_\bullet$ in a vector space V is fixed, that flag partitions $\mathbb{P}^n(F)$ into disjoint sets, called Schubert cells, by grouping together the lines in F^{n+1} that have the same **attitude** with respect to the flag, or in other words, lines that intersect the vector spaces F_i in the same dimensions.

Definition 1.7.11. *Let $F_\bullet$ be a complete flag in F^{n+1}. For each $i = 1, \ldots, n$, the corresponding **Schubert cell** in $\mathbb{P}^n(F)$ is $\Omega^\circ_{\{i\}}F_\bullet = \{x \in \mathbb{P}^n(F) \mid x \subseteq F_{n+2-i}$ and $x \not\subseteq F_{n+1-i}\}$, and $\Omega^\circ_{\{n+1\}}F_\bullet = F_1$. Again for each $i = 2, \ldots, n+1$, the corresponding **Schubert variety** in $\mathbb{P}^n(F)$ is $\Omega_{\{i\}}F_\bullet = \overline{\Omega^\circ_{\{i\}}F_\bullet} = \{x \in \mathbb{P}^n(F) \mid x \subseteq F_{n+2-i}\}$ (and again $\Omega_{\{n+1\}}F_\bullet = F_1$), where the closure is in the Zariski topology.*

To illustrate Schubert cells and Schubert varieties, fix the opposite flag $O_\bullet$ in $\mathbb{P}^2$. Then, $O_1 = \mathrm{Span}\{\vec{e_3}\}$, $O_2 = \mathrm{Span}\{\vec{e_3}, \vec{e_2}\}$, and $O_3 = \mathrm{Span}\{\vec{e_3}, \vec{e_2}, \vec{e_1}\}$. In this case, $\Omega^\circ_1 O_\bullet = \{\begin{bmatrix} 1 \\ * \\ * \end{bmatrix}\}$, $\Omega^\circ_2 O_\bullet =$

$\{\begin{bmatrix}0\\1\\*\end{bmatrix}\}$, and $\Omega_3^\circ O_\bullet = \{\begin{bmatrix}0\\0\\1\end{bmatrix}\}$ (where $*$ here means any complex number). In other words, $\Omega_1^\circ O_\bullet = \mathbb{A}_x^2$ is the affine chart (what we will refer to as the "big Schubert cell" or "big cell") containing all lines that intersect the $x = 1$ plane. This contains all points of $\mathbb{P}^2$ except for those representing lines in the yz-plane (which is O_2) in $\mathbb{C}^3$. At the same time, $\Omega_2^\circ O_\bullet$ represents all lines which are in the yz-plane (O_2), except the z-axis (O_1). Then, $\Omega_3^\circ O_\bullet$ is the z-axis itself (O_1). Note that we have a disjoint union $\mathbb{P}^2 = \Omega_1^\circ O_\bullet \cup \Omega_2^\circ O_\bullet \cup \Omega_3^\circ O_\bullet$, with $\Omega_1^\circ O_\bullet \cong \mathbb{A}^2$, $\Omega_2^\circ O_\bullet \cong \mathbb{A}^1$, and $\Omega_2^\circ O_\bullet \cong \mathbb{A}^0$ (a single point), so this is what we mean by $\mathbb{P}^2$ being paved by affines. Taking closures, we have that $\Omega_1 O_\bullet = \{\begin{bmatrix}*\\*\\*\end{bmatrix}\} \cong \mathbb{P}^2$, $\Omega_2 O_\bullet = \{\begin{bmatrix}0\\*\\*\end{bmatrix}\} \cong \mathbb{P}^1$, and $\Omega_3 O_\bullet = \{\begin{bmatrix}0\\0\\*\end{bmatrix}\} \cong \mathbb{P}^0$ (a single point), and so each Schubert variety is a projective space. One can easily generalize computing the Schubert decomposition with respect to the opposite flag for any projective space $\mathbb{P}^n(F)$.

Schubert varieties are key for understanding why projective space $\mathbb{P}^n(F)$ is the ideal setting for solving systems of polynomial equations in $F[x_0, \ldots, x_n]$ (which polynomials we can make homogeneous like in our motivating example going from lines to planes). This is because Schubert varieties are normal forms for equivalence classes in an F-algebra called the Chow ring of $\mathbb{P}^n(F)$, called Schubert classes, which form a finite-dimensional basis for the F-vector space structure of the ring. Furthermore, the intersections of varieties correspond to multiplication in the ring, revealing that understanding intersections of Schubert varieties is sufficient for understanding all the intersection theory in $\mathbb{P}^n(F)$. The rest of this section will be less rigorous as intersection theory in algebraic geometry is quite technical, but we hope to illustrate the main ideas with examples. Note below by a free abelian group, we mean the analogue of a vector space over the integers $\mathbb{Z}$. Though $\mathbb{Z}^n$ is not a vector space, it does have a standard basis over $\mathbb{Z}$, which is a ring. Such objects are called free abelian groups, and are an example of a free module (modules over a ring having the same axioms as a vector space, just with scalars from a ring instead of a field).

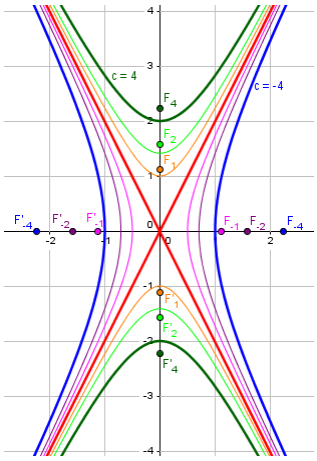
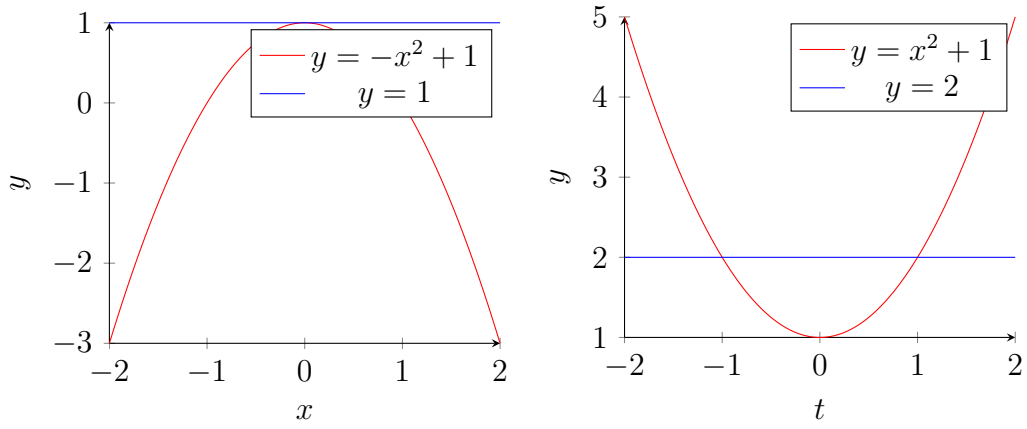


Figure 1.15: Hyperbola and Two Disjoint Lines Are Rationally Equivalent

Definition 1.7.12. *Let F be a field. The **set of k-cycles on** $\mathbb{P}^n(F)$, denoted by $Z_k(\mathbb{P}^n(F))$, is the free abelian group with basis every codimension k irreducible subvariety of $\mathbb{P}^n(F)$. In other words, a k-cycle C is a formal linear combination of irreducible subvarieties with integer coefficients: $C = n_1Y_1 + \cdots + n_kY_k$. We informally say that two cycles C_1 and C_2 are **rationally equivalent**, denoted $C_1 \sim C_2$, if one can rationally deform C_1 into C_2 (meaning there exists a flat family of k-cycles, rationally parametrized by $\mathbb{P}^1$, with C_1 and C_2 both in the family). To get the idea of rational equivalence, see Figure 1.13, which shows how a family of hyperbolas can deform rationally onto two disjoint lines, showing that the hyperbola is rationally equivalent to two disjoint lines. The k-th Chow group is the set of equivalence classes $A^k(\mathbb{P}^n(F)) = Z_k(\mathbb{P}^n)/\sim$. The **Chow ring** $A(\mathbb{P}^n(F))$ is the free abelian group graded by codimension $A(\mathbb{P}^n(F)) = \bigoplus_{k=0}^{n} A^k(\mathbb{P}^n(F))$, so multiplication takes $A^k(\mathbb{P}^n(F)) \times A^l(\mathbb{P}^n(F)) \to A^{k+l}(\mathbb{P}^n(F))$ (and $A^k(\mathbb{P}^n(F))$ is understood to be $\{\vec{0}\}$ for $k < 0$ or $k > n$. The multiplication in general is quite technical to describe, but for any two irreducible subvarieties $Y, Z \subseteq X$, if the intersection $Y \cap Z$ is transverse, $[Y][Z] = [Y \cap Z]$, and this is extended linearly over $\mathbb{Z}$.*



Note a transverse intersection at a point means that there are enough tangent vectors to both subvarieties to span the entire space (the equations for the tangent space are in direct sum). Looking at the two graphs pictured above, on the left one has a non-transverse intersection of the line and the parabola at their unique point of intersection since they share a tangent line, but on the right for either point of intersection of the line and parabola, there are two distinct tangent lines, so the intersection is transverse.

The Chow ring can be defined for any smooth projective variety in much the same way, and by a theorem of Totaro, the Chow ring of such a variety that has an affine stratification into cells has a basis given by distinct equivalence classes of the closures of those cells. For projective space $\mathbb{P}^n(F)$, this means for any complete flag $F_\bullet$, $A(\mathbb{P}^n(F))$ has its set of rational equivalence classes of Schubert varieties (which are distinct) $\{[\Omega_i F_\bullet] \mid i = 1, \ldots, n+1\}$ as a basis. Hence, since the product of vectors will still be a vector and can be expressed in terms of this basis, it suffices (after knowing transversality holds) to know the intersections of Schubert varieties to know how multiplication works in the ring (the intersection theory of $\mathbb{P}^n(F)$ in general).

For $\mathbb{P}^2$, the Chow ring $A(\mathbb{P}^2) \cong \mathbb{Z}[\zeta]/(\zeta^3)$, where $\zeta = [\Omega_2 O_\bullet]$. Here, a basis is given by the Schubert varieties $\{1, \zeta, \zeta^2\}$, where the multiplicative identity $1 = [\Omega_1 O_\bullet]$ represents all of $\mathbb{P}^2$, ζ represents a line (copy of $\mathbb{P}^1$) in $\mathbb{P}^2$, and $\zeta^2 = [\Omega_3 O_\bullet]$ represents a point in $\mathbb{P}^2$. The fact that $\zeta^2 = [\Omega_3 O_\bullet]$ represents that every two lines in $\mathbb{P}^2$ intersect in a point, which we observed earlier. Note that the degree of ζ corresponds to codimension. Furthermore, it can be shown that the rational equivalence class of every irreducible projective plane curve $V(f)$ of degree d in $A(\mathbb{P}^2)$ is

$d\zeta$. This leads to understanding general intersections of curves (solutions to polynomial systems) in $\mathbb{P}^2$.

Theorem 1.7.13. *(Bézout's Theorem) Let X and Y be irreducible plane curves in $\mathbb{P}^2$ of degrees d and e, respectively. If $X \cap Y$ is a transverse intersection, then $X \cap Y$ consists of de distinct points.*

Proof. Since transverse $[X \cap Y] = [X][Y] = (d\zeta)(e\zeta) = (de)\zeta^2$, the class in $A(\mathbb{P}^2)$ representing de points. □

This construction generalizes to $\mathbb{P}^n$, as $A(\mathbb{P}^n) \cong \mathbb{Z}[H]/(H^{n+1})$, which has the Schubert varieties $\{1, H, \ldots, H^n\}$ as basis. Here again irreducible hypersurfaces of degree d have rational equivalence class dH, where $H = [\Omega_2 O_\bullet]$ is the class of a hyperplane (copy of $\mathbb{P}^{n-1}$ in $\mathbb{P}^n$). There is a more general notion of degree for a subvariety of any codimension (not just hypersurfaces), and an irreducible subvariety of degree d and codimension k in general has rational equivalence class dH^k. Hence, we get the generalized versions of Bézout's Theorem:

Theorem 1.7.14. *If X and Y are irreducible subvarieties in $\mathbb{P}^n$ of complementary codimensions k and l (meaning $k + l = n$), and degrees d and e, respectively, then if $X \cap Y$ is transverse, $X \cap Y$ consists of de distinct points. Furthermore, if $X_1, \ldots, X_n$ are irreducible hypersurfaces of degrees $d_1, \ldots, d_n$ that intersect generically transversely, then $X_1 \cap \cdots \cap X_n$ consists of $d_1 \cdots d_n$ distinct points.*

Proof. Just as for Bézout's Theorem above. □

1.8 Monodromy Groups of Branched Covers

We first give some final preliminaries from topology and algebraic geometry.

Definition 1.8.1. *Let B be a connected topological space. A **covering space** of B is a topological space X, equipped with a continuous map $\pi : X \to B$, with the property that for all $b \in B$, there exists an open set $U_b \subseteq B$ containing b and a discrete space D_b such that $\pi^{-1}(U_b) = \bigcup_{x \in D_b} V_x$ and $\pi|_{V_x} : V_x \to U_b$ is a homeomorphism for all $x \in D_b$. The map π is called a **covering**, the open*

*sets V_x are called **sheets**, which are uniquely defined up to homeomorphism, and the discrete set $\pi^{-1}(\{b\})$ is called the **fiber** of b for all $b \in B$. It can be shown that π is connected and that the cardinality of the discrete set D_b is the same for all $b \in B$, called the **degree** of the covering.*

One reason that covering spaces are important in topology is that they satisfy the lifting property: Let B be a connected and locally connected based topological space with base point $b \in B$, and let $\pi : X \to B$ be a degree d covering with fiber $S = \pi^{-1}(\{b\}) = \{x_1, \ldots, x_d\}$. For a loop $\gamma : [0,1] \to B$ based at b (i.e. $\gamma(0) = \gamma(1) = b$), for each x_i we obtain a lift to the covering space $\tilde{\gamma}_i : [0,1] \to X$, which is a path (not necessarily a loop) with $\tilde{\gamma}_i(x_i) \in S$. In other words, for each base point $b \in B$ and loop γ based at b, we obtain a permutation of S, i.e. an element of S_d. Considering all such base points and loops, we get a subgroup of S_d, called the **monodromy group** of the covering $\pi : X \to B$.

Now we move from topology to algebraic geometry.

Definition 1.8.2. *A **morphism of projective varieties** $X \in \mathbb{P}^n(F)$ and $Y \subseteq \mathbb{P}^m(F)$ is a function $f : X \to Y$ whose component functions are all homogeneous polynomials of the same degree, who do not vanish simultaneously on Y. A **rational function** on a projective variety $X \subseteq \mathbb{P}^n(F)$ is any rational polynomial $h = \frac{P}{Q}$ ($P, Q \in F[x_0, \ldots, x_n]$) equivalent to $\frac{P'}{Q'}$ where $P', Q' \in F[x_0, \ldots, x_n]$ are homogeneous of the same degree, and Q' does not vanish identically on X (there exists some $p \in X$ with $Q'(p) \neq 0$). Here by equivalent we mean that $PQ' - P'Q$ vanishes at every point of X. The set of rational functions on X forms a field, called the **field of rational functions on** X, denoted by $F(X)$. For any morphism of projective varieties $f : X \to Y$, there is an induced field homomorphism $f^* : F(Y) \to F(X)$, given by $f^*(h) = h \circ f$. If the image of f is **dominant** (meaning its image is dense, surjective being a special case), then f^* is injective, giving a field extension $F(X)/F(Y)$.*

Definition 1.8.3. *A **degree** d **branched covering** is a morphism of projective varieties $f : X \to B$ such that there exists a dense open subset $U \subseteq B$ with $f : f^{-1}(U) \to U$ a degree d covering space map. Since such a map is dominant by definition, it induces a field extension on the fields of rational functions $F(X)/F(B)$.*

A classic example of monodromy is for the branched covering given by $\pi : \mathbb{C} \to \mathbb{C}$, $\pi(z) = z^2$. Away from $z = 0$, this is a degree 2 covering space. Taking a non-zero base point (say $z = 1 \in \mathbb{C}$), that base point will have two distinct square roots, and lifting a loop based at the base point will either send the roots to themselves, or will permute the roots (depending on what is called the winding number of the loop). The case where the roots are permuted is depicted in Figure 1.16.

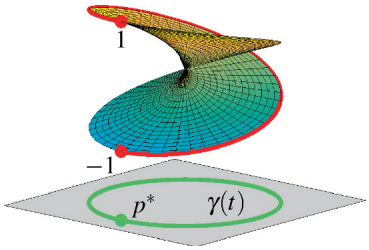


Figure 1.16: Monodromy for a degree 2 branched cover

Now that we've covered the preliminaries and motivation for projective algebraic geometry, having introduced and showed the usefulness of many types of algebraic structures (groups, rings, fields, and vector spaces) and geometric structures (topological spaces and manifolds), we combine everything together to define Galois groups in enumerative geometry. Enumerative geometry considers counting the number of geometric objects that interact with other geometric objects in some prescribed way, and these Galois groups will encapsulate the symmetries of the solution set.

A classical question in enumerative geometry is "how many lines are on a nonsingular cubic surface?" Here by a nonsingular cubic surface, we mean the projective hypersurface $V(f) \in \mathbb{P}^3$ given as the solution to a homogeneous degree 3 polynomial equation in 4 variables: $f(x, y, z, w) = a_1x^3+a_2x^2y+a_3x^2z+a_4x^2w+a_5xy^2+a_6xyz+a_7xyw+a_8xz^2+a_9xzw+a_{10}xw^2+a_{11}y^3+a_{12}y^2z+a_{13}y^2w + a_{14}yz^2 + a_{15}yzw + a_{16}yw^2 + a_{17}z^3 + a_{18}z^2w + a_{19}zw^2 + a_{20}w^3 = 0$ (here the $a_i \in \mathbb{C}$. By nonsingular here we mean that the gradient vector (from multivariable calculus) $\nabla f = \vec{0}$ has no solutions in $\mathbb{P}^3$. The answer for the number of lines is the well-known Cayley-Salmon Theorem (two papers in 1849), which is that they all contain 27 lines, independent of which nonsingular

cubic surface you choose. (See APPENDIX A. for a proof after more preliminaries from Chapter 2).

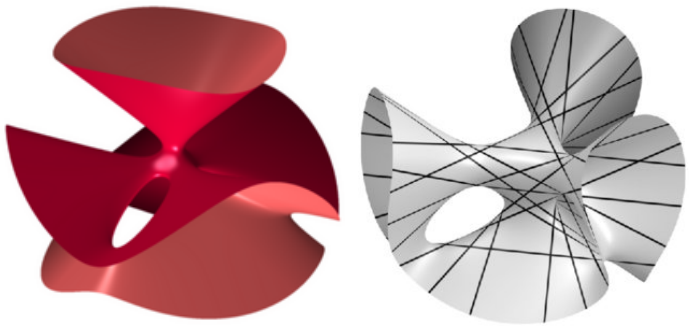


Figure 1.17: Two nonsingular cubic surfaces, the second with 27 lines revealed

While this answer of 27 lines is already remarkable, one can ask further ask “what are the symmetries of these lines?", in the same way that we ask about the symmetries of an equilateral triangle (the permutation group S_3) or the square (the dihedral group of order 8). Recall that while the triangle with labeled vertices can achieve all possible permutations of its vertices via rigid motions, the square with labeled vertices cannot, due to pairs of diagonal vertices being preserved under reflections and rotations.

For a nonsingular cubic surface, we similarly consider a labeling of its 27 lines, but we will be considering monodromy permutations as our geometric transformations rather than reflections and rotations. Recall that the homogeneous polynomial defining a nonsingular cubic surface has 20 coefficients. As one varies these coefficients in a loop, at each step one has a new cubic surface with 27 lines moving along the loop. Since we are varying in loop, eventually we get back to the coefficients that we started with, and so arrive at the same nonsingular cubic surface with its original set of 27 lines. However, the labeling that we imposed on these lines may have permuted, depending on the loop. Considering all possible loops, we obtain the our desired symmetry group, the monodromy group of our cubic surface. This is made precise by considering the incidence variety of pairs of cubic surfaces with their lines, and its projection onto the cubics, which is

a branched covering of degree 27. Away from the singular cubic surfaces, this is a degree 27 covering space map, which has a monodromy group as described earlier. As a note, by $\mathbb{P}^{19}_{\text{cubics}}$ we mean that the set of all cubics are given by the 20 coefficients ($\mathbb{C}^{20}$) up to scaling, since multiplying $f(x, y, z, w) = 0$ by a non-zero scalar does not change the solution set. Also by $\mathbb{G}(1, \mathbb{P}^3)$, we mean the set of lines in $\mathbb{P}^3$, and will make this notation precise in the next chapter.

$$\begin{array}{c} \Gamma := \{(\ell, F) \in \mathbb{G}(1, \mathbb{P}^3) \times \mathbb{P}^{19}_{\text{cubics}} \,:\, F|_\ell \equiv 0\} \\ \Big\downarrow \\ \mathbb{P}^{19}_{\text{cubics}} \end{array}$$

Schläfli (1858) showed that 27 lines on a nonsingular cubic surface have a remarkable incidence configuration with symmetry group having $58140 \ll 27!$ permutations. Further, this group encodes how every line intersects exactly 10 other lines and that every pair of disjoint lines intersects exactly 5 other lines, giving geometric obstructions to obtaining all possible permutations of the lines. Gradually recognized by Cartan (1896), Coble (1915–17), and du Val (1936) as the Weyl group of type E_6, and so we will refer to this enriched Galois group as E_6 from now on.

Further, this monodromy group is also a Galois group in the traditional sense. Our above branched covering is a map of algebraic varieties, and the induced algebraic extension $\mathbb{C}(\Gamma)/\mathbb{C}(\mathbb{P}^{19})$ of function fields has degree 27. If K is the normal closure of $\mathbb{C}(\Gamma)/\mathbb{C}(\mathbb{P}^{19})$, then $\mathrm{Gal}(K/\mathbb{C}(\mathbb{P}^{19})) = E_6$ (many proofs of this given in the 20th century).

As the problem of 27 lines on a nonsingular cubic surface was found to have a remarkable symmetry group, we call its Galois group E_6 enriched (with additional structure) because it was not the full symmetric group S_{27}. In general if an enumerative geometry problem has d solutions, we say that its Galois group is **full-symmetric** if it is S_d, and it is **enriched** otherwise. We also use these terms to describe the problem as full-symmetric or enriched as well. One can then ask, besides 27 lines on a nonsingular cubic surface, what other enriched enumerative geometry problems are there? Given an enumerative geometry problem, we have the general framework, with monodromy and Galois groups defined in the same way as for cubic surfaces.

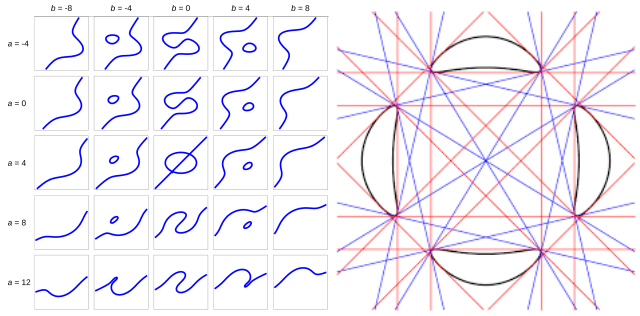


Figure 1.18: 9 Flexes on a Cubic and 28 Bitangents to a Quartic

$$\Gamma \subseteq \{\text{Solution Space}\} \times \{\text{Parameter Space}\}\,.$$
$$\downarrow$$
$$\{\text{Parameter Space}\}$$

The equivalence of monodromy groups and Galois groups in the context of enumerative geometry (not just for cubic surfaces) was shown by Harris over $\mathbb{C}$ using tools from complex analysis, but the ideas trace back to Hermite. See APPENDIX B for a modern proof given by Sottile and Yahl, with ideas originally presented by Vakil (which uses scheme theory and has the appeal of being true over any field).

By 1979, in addition to 27 lines on a nonsingular cubic surface, there were only two other known enumerative geometry problems to have enriched Galois groups - the problems of 9 flexes on a nonsingular cubic (degree 3) plane curve, and 28 bitangents to a nonsingular quartic (degree 4) plane curve.

In his 1979 landmark paper “Galois Groups of Enumerative Problems", Harris naturally generalized the three known enriched problems, and showed that in each case, the generalizations to higher dimensions had full-symmetric Galois groups. Additionally, he showed that the famous problem of 3264 conics tangent to five general conics had a Galois group that was fully-symmetric, again revealing how rare enriched Galois groups are. This stalled progress in the field until 2003, when Derksen and Vakil found an enriched Galois problem involving six 4-dimensional subspaces in $\mathbb{C}^8$ intersecting four general 4-dimensional subspaces, each in dimension at least 2. This type

of enumerative geometry problem, involving only linear spaces, is called a Schubert problem. In 2006, Vakil was able to generalize this Schubert problem to an infinite family of Schubert problems with enriched Galois groups, leading to the program of classifying all possible enriched Schubert problems, which is the focus of the rest of this dissertation.

2. SCHUBERT PROBLEMS

In the last section, we discussed Galois groups in the context of enumerative geometry, and the historical search for enriched Galois groups. The first first infinite family of enriched Galois groups (by Vakil) was found for enumerative geometry problems only involving linear spaces. Such problems are called Schubert problems, and the study of Schubert problems is called Schubert calculus, each named after the mathematician Hermann Schubert (1841-1911), not the famous Austrian composer, Franz Schubert. We introduce Schubert problems with an important problem called the Problem of Four Lines, which while insightful, does not have an enriched Galois group. We then give the framework for Schubert problems, which have three main components made precise by the notions of Grassmannians, flags, and Schubert Varieties. We then dive into the details on Schubert Galois groups. Finally, we introduce Schubert calculus for partial flag varieties of various types, and describe how Schubert problems can be solved using cohomology.

Figure 2.1: Hermann and Franz Schubert

2.1 The Problem of Four Lines

The Problem of Four Lines is an important motivating example for Schubert calculus. Given four general lines ℓ_1, ℓ_2, ℓ_3, and ℓ_4 in $\mathbb{P}^3$ (so planes through the origin in $\mathbb{C}^4$), how many other lines in $\mathbb{P}^3$ intersect each of the ℓ_i simultaneously? Here by general lines, we mean that the ℓ_i do not intersect one another, that no three of them are coplanar (lie on a plane in $\mathbb{P}^3$), and a bit more. Imagining these lines in $\mathbb{R}^3$ (the space that we seemingly live in), it is not obvious at first glance that there should be any lines that intersect four general lines in space. Or if there were any, why would there not be infinitely many such solutions? Incredibly, there turns out to be a finite number of solutions, and there is a geometric reason for this being the case (all of this typical for Schubert problems in general).

Focusing on the first three lines, ℓ_1, ℓ_2, and ℓ_3, it is a classical result that they determine a unique hyperboloid of one sheet in $\mathbb{P}^3$. This hyperboloid is a quadric surface (a projective hypersurface given as the vanishing of a homogeneous degree 2 polynomial), with two rulings of lines on it (can be written as the union of infinitely many non-intersecting lines in two different ways). The three lines ℓ_1, ℓ_2, and ℓ_3 belong to the first ruling, and the second ruling consists of the infinitely many lines that intersect each of ℓ_1, ℓ_2, ℓ_3 simultaneously (and is the grid of lines depicted in Figure 2.2). Hence, if this were the Problem of Three Lines, there would be infinitely many solutions - but it is not! We must also consider ℓ_4, which intersects the hyperboloid in two points (since it is degree 2), and each of these points is on a unique line in the second ruling of the hyperboloid. Calling these lines m_1 and m_2, they are the solutions to our Problem of Four Lines - there are exactly 2 lines that intersect four general lines in $\mathbb{P}^3$.

Furthermore, one can investigate the Galois group of the Problem of Four Lines by looking at the monodromy group of a branched cover, where a point of the base space (parameter space) is a choice of four general lines, and the fiber above that point is the two solution lines. Since there are only two solutions, the monodromy group is a subgroup of S_2. By fixing ℓ_1, ℓ_2, and ℓ_3 (which fixes our hyperboloid), we can get a loop in our parameter space by varying our fourth line ℓ_4 in a loop. One such loop is rotating ℓ_4 180° about the point p in Figure 2.2, which swaps the intersection

points of ℓ_4 with the hyperboloid, therefore permuting the solution lines m_1 and m_2. Thus, the Galois group for this problem is all of S_2, and so not an enriched Galois group.

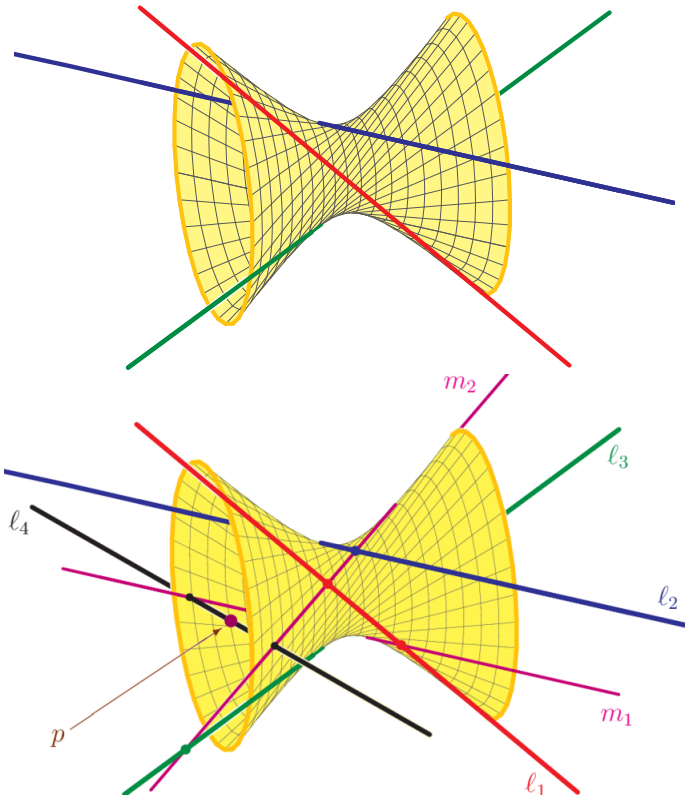


Figure 2.2: Two Solutions to the Problem of Four Lines

As a fun fact, the classic result that three general lines in space determine a unique hyperboloid of one sheet is credited to Christopher Wren, the English architect responsible for many famous landmarks, including St. Paul's Cathedral in London.

What is so remarkable about obtaining 2 solutions is that the answer didn't depend on the particular four lines we started out with. Each choice of four general lines, the parameters for the Problem of Four Lines, determines a hyperboloid (dependent on the first three lines) with two rulings, and the two solutions lines are in the second ruling of that hyperboloid. This theme was originally coined as "conservation of number" by Schubert, and will be true of all the Schubert

Figure 2.3: Christopher Wren and St. Paul's Cathedral

problems that we consider.

Informally, a **Schubert problem** involves a vector space V over a field F, and asks how many k-dimensional subspaces of V intersect some fixed (general) subspaces of V in certain prescribed dimensions. Since lines in $\mathbb{P}^3$ are 2-dimensional subspaces (2-planes) of $\mathbb{C}^4$, the Problem of Four Lines can be re-described in this language as "how many 2-planes in $\mathbb{C}^4$ intersect four general 2-planes in $\mathbb{C}^4$, each in dimension 1 (a line). Hence the answer is 2 2-planes in $\mathbb{C}^4$. To describe a general Schubert problem, there are three essential components that we must formalize:

1. Our solutions, which k-dimensional subspaces of V (Formalized by Grassmannians)

2. The fixed general subspaces we want our solutions to intersect. These will be the parameters for our problem (Formalized by Complete Flags)

3. How our solutions intersect the fixed general subspaces in certain prescribed dimensions (Formalized by Schubert Varieties)

The next two sections give the desired formalizations, allowing us to rigorously understand all that goes into a Schubert problem.

2.2 Grassmannians

Grassmannians naturally generalize projective spaces, and are ubiquitous in differential and algebraic geometry. We define them, and describe their structure as algebraic manifolds and projective varieties. The solutions to Schubert problems will be a finite subset of a Grassmannian.

Definition 2.2.1. *Let V be a vector space over a field F. The **Grassmannian** of k-planes in V is* $\mathrm{Gr}(k, V) = \{$*subspaces* $H \subseteq V \mid \dim(H) = k\}$. *If $V = F^n$, we write* $\mathrm{Gr}(k, n)$ *for* $\mathrm{Gr}(k, F^n)$. *Projecting down a dimension, even though $\mathbb{P}^n$ is not a vector space, we abuse notation and write* $\mathbb{G}(k, \mathbb{P}^n) = \{$*linear subsets* $H \subseteq \mathbb{P}^n \mid \dim(H) = k\} = \mathrm{Gr}(k+1, n+1)$

Note that projective spaces are Grassmannians: $\mathbb{P}^n(F) = \mathrm{Gr}(1, \mathbb{F}^{n+1})$. In **Stiefel coordinates** (a generalization of homogeneous coordinates), we will be viewing the Grassmannian $\mathrm{Gr}(k, n)$ of k-planes in $\mathbb{F}^n$ as the set $Mat^k_{n\times k}(\mathbb{C})$ of rank k $n \times k$ matrices under an equivalence relation, where two $n \times k$ matrices A and C (both of full rank k) are equivalent if there is an invertible matrix $P \in \mathrm{GL}(k, F)$ such that $AP = C$. Like how vectors in $\mathbb{P}^n$ represent the line that they span, and so are equivalent up to non-zero scalar multiple, the k-plane represented by a full-rank $n \times k$ matrix is recovered by taking the span of the columns of the matrix, and two matrices equivalently represent the same k-plane if their columns are different bases for the same k-dimensional subspace.

Note since elements of $\mathrm{Gr}(k, n)$ represented as matrices have full rank k, by definition that matrix has a maximal minor ($k \times k$ sub-matrix) with non-zero determinant. We will be using symbols α and β to represent k-tuples of row indices, i.e. for example $\alpha = (\alpha_1, \ldots, \alpha_k)$ for some $1 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_k \leq n$. Using this notation, a matrix $M \in Gr(k, n)$ has $\binom{n}{k}$ maximal minors, and at least one, say M_α, the sub-matrix of M with rows $\alpha_1, \ldots, \alpha_k$, has $det(M_\alpha) \neq 0$ (like how at least one entry of a vector in projective space must be non-zero).

The algebraic manifold structure on $\mathrm{Gr}(k, n)$ is then given as follows:

1. For each of the $\binom{n}{k}$ α's, we have that the coordinate charts of $\mathrm{Gr}(k, n)$ are U_α, the set of rank k, $n \times k$ matrices (under equivalence) M with minor M_α having non-zero determinant. Like projective space, each equivalence class in U_α has a unique representative with rows

α being the identity matrix. (since $det(M_\alpha) \neq 0$, just multiply on the right by M_α^{-1} to get this representative). Thus we will see that $\mathrm{Gr}(k,n)$ is $k(n-k)$ dimensional (removing the k rows giving the identity matrix, there are $k(n-k)$ coordinates remaining).

2. The coordinate map of U_α is $\varphi_\alpha : U_\alpha \to \mathbb{A}^{k(n-k)}(F)$, where $\varphi(M) = MM_\alpha^{-1}$, with rows α (they will form an identity matrix) then omitted. Note here that $\mathbb{A}^{k(n-k)}(F)$ and the set of $(n-k)\times k$ matrices are being identified. Again, on the above unique representatives, this will just be a projection map away from the rows given by α.

3. $\varphi_\alpha^{-1} : \mathbb{A}^{n(n-k)}(F) \to U_\alpha$ takes an $(n-k)\times k$ matrix B, and makes an $n \times k$ matrix M out of it with rows α being the identity, other rows being filled in by those of B (in the same order).

4. The change of coordinates map $\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$ then takes a $(n-k)\times k$ matrix B and applies the following sequence of transformations to it:

   (1) Makes an $n \times k$ matrix M out of it with rows α being the identity, other rows being filled in by those of B (in the same order)

   (2) Multiply MM_β^{-1}

   (3) Omit rows of the product coming from β Since the entries of the resulting matrix are all rational polynomials (by Cramer's rule for inverses of matrices) defined everywhere on $U_\alpha \cap U_\beta$, this endows $\mathrm{Gr}(k,n)$ with the structure of a $k(n-k)$-dimensional algebraic manifold.

On the other hand, the injective **Plücker embedding** $\mathrm{Gr}(k,n) \to \mathbb{P}^{\binom{n}{k}-1}(F)$ sending M to $[\det(M_\alpha)]$ (ordered in some consistent way) gives $\mathrm{Gr}(k,n)$ the structure of a smooth projective variety. This map is well defined since MP for any $P \in GL_k(\mathbb{F})$ would get sent to $\det(P)[\det(M_\alpha)] = [\det(M_\alpha)]$ (since in projective space). More importantly, the determinants of minors of a matrix are viewed as its coordinates, and these determinants (as coordinates) are precisely the set of solutions to a set of homogeneous polynomial equations (called **Grassman-Plücker relations**): $\sum_{l=1}^{k+1}(-1)^l W_{\alpha_1,\ldots,\alpha_{k-1},\beta_l} W_{\beta_1,\ldots,\beta_{l+1}} = 0$ (where in the 2nd W, β_l is omitted).

Here the indices of the W's refer to the rows of the minor.

For the Problem of Four Lines considered earlier, the relevant Grassmannian is $\mathbb{G}(1, \mathbb{P}^{\text{⊮}}) = \mathrm{Gr}(2, 4)$.

The Plücker embedding takes a matrix $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix}$ to the point $\begin{bmatrix} W_{1,2} \\ W_{1,3} \\ W_{1,4} \\ W_{2,3} \\ W_{2,4} \\ W_{3,4} \end{bmatrix} \in \mathbb{P}^5$ ($\binom{4}{2} - 1 = 5$), where $W_{i,j}$ is the determinant $x_{i1}x_{j2} - x_{i2}x_{j1}$ of the 2×2 minor $\begin{bmatrix} x_{i1} & x_{i2} \\ x_{j1} & x_{j2} \end{bmatrix}$. One can then verify that the coordinates $W_{i,j}$ satisfy the polynomial equation $W_{1,2}W_{3,4} - W_{1,3}W_{2,4} + W_{1,4}W_{2,3} = 0$, and in fact the image of the embedding is defined by this single equation, so $\mathrm{Gr}(2, 4)$ is a quadric hypersurface in $\mathbb{P}^5$ (so degree 2, codimension 1, and so dimension 4.

One can also see the 4-dimensional manifold structure of $\mathrm{Gr}(2, 4)$ from the Stiefel coordinates, since for example $U_{1,2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ y_{31} & y_{32} \\ y_{41} & y_{42} \end{bmatrix}$, which as 4 coordinates outside of the 2×2 identity submatrix. Applying the Plücker embedding to just this affine chart gives $\begin{bmatrix} 1 \\ y_{32} \\ y_{42} \\ -y_{31} \\ -y_{41} \\ y_{31}y_{42} - y_{32}y_{41} \end{bmatrix} \in \mathbb{A}^5_{W_{1,2}} \subseteq \mathbb{P}^5(F)$. This generalizes, revealing that the Plücker embedding takes affine charts of $\mathrm{Gr}(k, n)$ to the corresponding affine charts of $\mathbb{P}^{\binom{n}{k}-1}(F)$ (indexed by the same minor).

For those familiar with wedge products of vector spaces, he injective Plücker embedding can be made coordinate free (and is often given this way) by $\mathrm{Gr}(k, V) \to \mathbb{P}(\Lambda^k(V))$, sending

$\mathrm{Span}(w_1, \ldots, w_k)$ to $[w_1 \wedge \ldots \wedge w_k]$. With this perspective, the Grassmannian is identified with its image and is the set of totally decomposable $\omega \in \mathbb{P}(\Lambda^k(V))$.

2.3 Flags and Schubert Varieties

Now that we are familiar with Grassmannians (where the solutions to our Schubert problems will lie), we need to formalize the subspaces our solutions will be required to intersect, and how our solutions will do so in prescribed dimensions. This need for formalization leads us to generalize the Schubert decomposition of $\mathbb{P}^n(F)$ into Schubert cells and Schubert varieties to Grassmannians. Recall from Section 1.7 that a complete flag $F_\bullet$ is a saturated chain of subspaces of a vector space V. Given such a flag, we can consider how k-dimensional subspaces of V interact with that flag (such as considering all subspaces that intersect the F_i subspace in our flag in dimension at least $a_i \in \mathbb{N}$). This partitions the Grassmannian $\mathrm{Gr}(k, V)$ into Schubert cells, the closures of which will be our Schubert varieties for the Grassmannian (closures in the Zariski topology, inherited as a projective subvariety of $\mathbb{P}^n(F)$) via the Plücker embedding).

Definition 2.3.1. *A **Schubert condition** on $\mathrm{Gr}(k, n)$ is an increasing sequence α of k integers in $\{1, \ldots, n\}$, i.e. $\alpha = (\alpha_1, \ldots, \alpha_k)$ with $\alpha_1 < \cdots < \alpha_k$. We write $\alpha = \alpha_1\alpha_2 \cdots \alpha_k$ for the Schubert condition in one-line notation. The set of all Schubert conditions for $\mathrm{Gr}(k, n)$ is denoted $\binom{[n]}{k}$. The **(Coxeter) length** of a Schubert condition $\alpha \in \binom{[n]}{k}$ is $|\alpha| = \sum_{i=1}^{k}(\alpha_i - i)$. For any $\alpha \in \binom{[n]}{k}$ and flag $F_\bullet \in \mathrm{Fl}(n)$, the corresponding **Schubert variety** is $\Omega_\alpha F_\bullet = \{H \in \mathrm{Gr}(k, V) | \dim(H \cap F_{n+1-\alpha_{k+1-i}}) \geq i \text{ for } i = 1, \ldots, k\}$. Note that $|\alpha|$ gives the codimension of the Schubert variety $\Omega_\alpha F_\bullet$ as a subvariety of $\mathrm{Gr}(k, n)$ (i.e. $\dim(\Omega_\alpha F_\bullet) = k(n-k) - |\alpha|$. The **Bruhat order** on $\binom{[n]}{k}$ is the partial order where $\alpha \leq \beta$ if $\alpha_i \leq \beta_i$ for all $i = 1, \ldots k$. For any $\alpha \in \binom{[n]}{k}$ and flag $F_\bullet \in \mathrm{Fl}(n)$, the corresponding **Schubert cell** is $\Omega^\circ_\alpha F_\bullet = \Omega_\alpha F_\bullet \setminus \bigcup_{\beta \lneq \alpha} \Omega_\beta F_\bullet$.*

The above definition uses what is called the “codimension convention", since the length of a Schubert condition gives the codimension of the corresponding Schubert variety. This is our standing convention (unless mentioned otherwise, which we will do in a moment), since many authors alternatively index Schubert varieties using partitions, the length of the partition also giving

the codimension of the Schubert variety. On the other hand there is a "dimension convention", which is easier to write down, where $\Omega_\alpha F_\bullet = \{H \in \mathrm{Gr}(k,V) | \dim(H \cap F_i) \geq i \text{ for } i = 1, \ldots, k\}$. To simplify calculations, we only use the dimension convention for the following example (getting the corresponding Schubert conditions in codimension notation via $\alpha - i \to n + 1 - \alpha_{k+1-i}$):

We find all the Schubert cells and Schubert varieties in the Schubert decomposition of $\mathbb{G}(1, \mathbb{P}^{\not 4})$ (again the Grassmannian relevant for the Problem of Four Lines). By a flag in $\mathbb{P}^3$ (instead of $\mathbb{C}^4$), we mean a point $F_0 = p \in \mathbb{P}^3$ on a line $F_1 = \ell \subseteq \mathbb{P}^3$, contained in a plane $F_2 = H \subseteq \mathbb{P}^3$, where $F_3 = \mathbb{P}^3$. Hence in dimension notation, we have $\Omega_\alpha F_\bullet = \{\text{lines } \mu \in \mathbb{G}(1, \mathbb{P}^3) | \dim(\mu \cap F_{\alpha_i - 1}) \geq i - 1 \text{ for } i = 1, 2, 3\}$. Some of these dimension of intersection conditions are already guaranteed to be satisfied, by the formula $\dim(W_1 \cap W_2) \geq \dim(W_1) + \dim(W_2) - \dim(V)$ from Section 1.6 (and then projectivized). In this case, $\dim(\mu \cap \mathbb{P}^3) \geq 1$ and $\dim(\mu \cap H) \geq 0$ are always satisfied. Some dimension of intersection conditions are also implied by others. Now, there are 6 Schubert conditions in $\mathbb{G}(1, \mathbb{P}^3)$: $\{34, 24, 23, 14, 13, 12\}$, and the corresponding Schubert varieties (using the dimension convention) are:

- $\Omega_{34} F_\bullet = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \dim(\mu \cap H) \geq 0 \text{ and } \dim(\mu \cap \mathbb{P}^3) \geq 1\} = \mathbb{G}(1, \mathbb{P}^3)$
- $\Omega_{24} F_\bullet = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \dim(\mu \cap \ell) \geq 0 \text{ and } \dim(\mu \cap \mathbb{P}^3) \geq 1\} = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \mu \cap \ell \neq \emptyset\}$
- $\Omega_{23} F_\bullet = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \dim(\mu \cap \ell) \geq 0 \text{ and } \dim(\mu \cap H) \geq 1\} = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \mu \subseteq H\}$
- $\Omega_{14} F_\bullet = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \dim(\mu \cap p) \geq 0 \text{ and } \dim(\mu \cap \mathbb{P}^3) \geq 1\} = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid p \in \mu\}$
- $\Omega_{13} F_\bullet = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \dim(\mu \cap p) \geq 0 \text{ and } \dim(\mu \cap H) \geq 1\} = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid p \in \mu \subseteq H\}$
- $\Omega_{12} F_\bullet = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \dim(\mu \cap p) \geq 0 \text{ and } \dim(\mu \cap \ell) \geq 1\} = \{\mu \in \mathbb{G}(1, \mathbb{P}^3) \mid \mu = \ell\}$

Unlike the Schubert subvarieties of $\mathbb{G}(1, \mathbb{P}^3)$, the Schubert cells are disjoint, and can be described as follows:

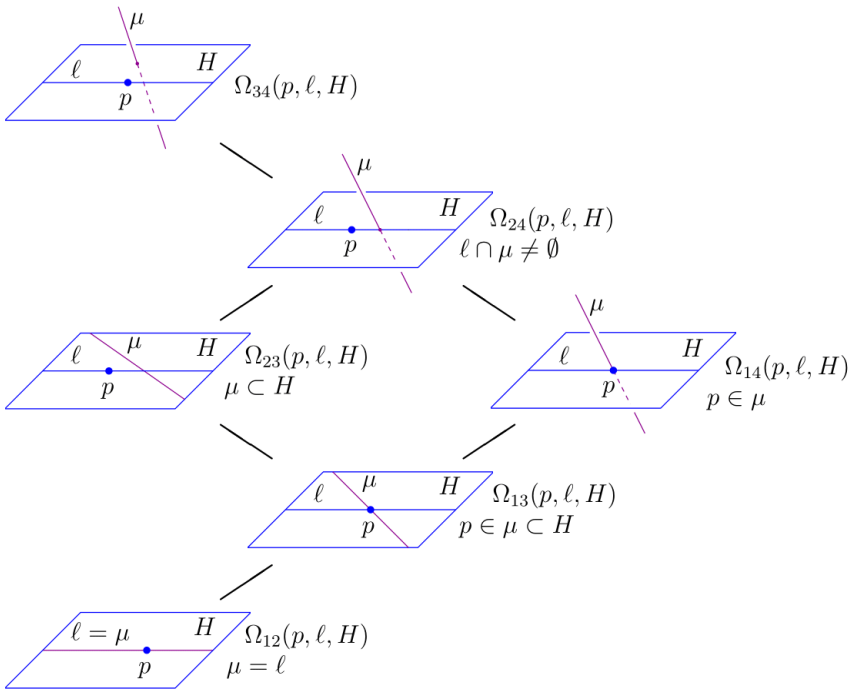


Figure 2.4: The Schubert Decomposition of $\mathbb{G}(1, \mathbb{P}^3)$ (Dimension Convention)

- Ω°_{34}: Lines that intersect the plane H, but do not intersect the line ℓ (this is the general case for a line in $\mathbb{P}^3$, and the other conditions below are specializations)
- Ω°_{24}: Lines that intersect the line ℓ, but do not contain the point p and are not contained in the plane H
- Ω°_{23}: Lines that are contained in the plane H (and so must intersect the line ℓ as a consequence), but do not contain the point p
- Ω°_{14}: Lines that contain the point p (and so must intersect the line ℓ as a consequence), but are not contained in the plane H
- Ω°_{13}: Lines that are contained in the plane H and contain the point p, but are distinct from the line ℓ
- Ω°_{12}: The line ℓ itself

2.4 Schubert Problems and Their Galois Groups

Now that we have formalized the separate components of a Schubert problem (Grassmannians, flags, and Schubert varieties), we are ready to precisely define Schubert problems in this language, as well as their corresponding Galois groups.

Definition 2.4.1. *Given a list of Schubert conditions* $\{\alpha^1, \ldots, \alpha^r\} \in \binom{[n]}{k}$*, with* $|\alpha^1| + \cdots + |\alpha^r| = k(n-k)$ *(the dimension of* $\mathrm{Gr}(k,n)$*), a* ***Schubert problem*** *is to compute* $|\Omega_{\alpha^1} F^1_\bullet \cap \ldots \cap \Omega_{\alpha^r} F^r_\bullet|$ *in* $\mathrm{Gr}(k,n)$*. Choosing specific, but general, complete flags* $\{F^1_\bullet, \ldots, F^r_\bullet\}$ *yields an* ***instance of the Schubert problem****.*

This intersection gives the number of k-dimensional subspaces of F^n that intersect the subspaces given by $\{F^1_\bullet, \ldots, F^r_\bullet\}$ in dimensions prescribed by $\{\alpha^1, \ldots, \alpha^r\}$. Note that for each $\mathrm{Gr}(k,n)$, there are finitely many Schubert problems, each of which can be computed as the intersection of Schubert varieties. Since the codimensions of the Schubert varieties in the intersection add up to the dimension of $\mathrm{Gr}(k,n)$, one expects this intersection to be a finite number of points, giving the finitely many solutions to the Schubert problem. Running over all n and $1 \leq k \leq n$, this gives a wealth of enumerative geometry problems, tens of millions of which can be realistically (in one's lifetime) computed using computer software (see the next section for details).

We now return to the Problem of Four Lines, which be formalized as a Schubert problem as follows (using the dimension convention to match our previous discussion of $\mathbb{G}(1, \mathbb{P}^{\text{IV}})$): We fix four general lines in space as before ℓ_1, ℓ_2, ℓ_3, and ℓ_4. For each ℓ_i ($i = 1, 2, 3, 4$), we complete it to a flag $F^i_\bullet$ by choosing a point $F^i_0 = p_i \in \ell_i = F^i_1$, and by choosing a plane $F^i_2 = H_i \subseteq \mathbb{P}^3 = F^i_3$. Now, we have four general flags $F^1_\bullet, F^2_\bullet, F^3_\bullet, F^4_\bullet \in \mathrm{Fl}(\mathbb{P}^{\text{IV}})$. The condition that a line $\mu \in \mathrm{Gr}(1, \mathbb{P}^{\text{IV}})$ intersects ℓ_i ($\mu \cap \ell_i \neq \emptyset$) is then given by the Schubert variety $\Omega_{24} F^i_\bullet \subseteq \mathrm{Gr}(1, \mathbb{P}^3)$. Thus, the condition that a line $\mu \in \mathrm{Gr}(1, \mathbb{P}^{\text{IV}})$ intersects all four of the ℓ_i simultaneously is if and only if $\mu \in \Omega_{24} F^1_\bullet \cap \Omega_{24} F^2_\bullet \cap \Omega_{24} F^3_\bullet \cap \Omega_{24} F^4_\bullet$. Note that these Schubert varieties are each distinct, even though they all have the same Schubert condition, since the flags used to define them are distinct. Therefore, the fact that there are 2 solutions to the Problem of Four Lines is equivalent

to the statement $|\Omega_{24}F_\bullet^1 \cap \Omega_{24}F_\bullet^2 \cap \Omega_{24}F_\bullet^3 \cap \Omega_{24}F_\bullet^4| = 2$, which can be computed using algebraic geometry (which we describe in a few different ways later on). Furthermore, it should be noted that while the Problem of Four Lines was able to be solved using a geometric argument involving a hyperboloid of one sheet, in general Schubert problems have no such argument, and so interpreting the problem as understanding the intersection of Schubert varieties is crucial to finding solutions, and studying them further.

Given a Schubert problem defined by the Schubert conditions $\alpha^1, \ldots, \alpha^r \in \binom{[n]}{k}$,

$$\begin{array}{l} \Gamma := \{(\mu, F_\bullet^1, \ldots, F_\bullet^r) \mid \mu \in \Omega_{\alpha^i} F_\bullet^i \text{ for all } i = 1, \ldots, r\} \\ \Big\downarrow \\ \mathrm{Fl}(n)^r \end{array}$$

is the corresponding branch cover, where $\mu \in \mathrm{Gr}(k, n)$. There is a dense open subset of $\mathrm{Fl}(n)$ where the projection is a degree d covering map, and we say that an r-tuple of flags in that open subset is a **general choice of flags**. Fixing general $(F_\bullet^1, \ldots, F_\bullet^r) \in \mathrm{Fl}(n)^r$, its fiber is an instance of the Schubert problem with its d solutions. As for the Problem of 27 Lines on a Nonsingular Cubic Surface, we have a monodromy group for this branched cover, which is isomorphic to the Galois group of the induced inclusion of function fields $G = \mathrm{Gal}(F(\Gamma)/F(\mathrm{Fl}(n)^r))$. Since this group permutes the solutions to the instance of the Schubert problem, it is a subgroup of S_d. If $G = S_d$, we again say that the Galois group is **full-symmetric**, and otherwise that the Galois group G is **enriched** (and also say that the Schubert problem is enriched).

In 2006, Derksen and Vakil discovered the enriched Schubert problem (using codimension notation) $|\Omega_{1256}F_\bullet^1 \cap \Omega_{1256}F_\bullet^2 \cap \Omega_{1256}F_\bullet^3 \cap \Omega_{1256}F_\bullet^4| = 6$ in $\mathrm{Gr}(4, 8)$. This problem is a generalization of the Problem of Four Lines, and states that given four general 4-planes $H_1, \ldots, H_4 \in \mathrm{Gr}(4, 8)$, there are exactly 6 other 4-planes $\mu \in \mathrm{Gr}(4, 8)$ such that $\dim(\mu \cap H_i) \geq 2$ for all $i = 1, 2, 3, 4$. Hence, the Galois group is a transitive subgroup of S_6, and in fact is an isomorphic copy of S_4 in S_6, and so has $24 < 720$ elements and is enriched.

Vakil later generalized this example into an infinite family of enriched Schubert problems, with problems in every Grassmannian $\mathrm{Gr}(k, n)$ for $4 \leq k \leq n - 4$. But this does not account for every

enriched Schubert problem in Grassmannians! Later, Martín del Campo, Sottile, and Williams classified all enriched Schubert problems in $\mathrm{Gr}(4,8)$ and $\mathrm{Gr}(4,9)$ (2019). Their classification found that all enriched Galois groups in $\mathrm{Gr}(4,8)$ and $\mathrm{Gr}(4,9)$ have the structure of being iterated wreath products of symmetric groups, and so the current conjecture is that for Grassmannians, all enriched Galois groups will be iterated wreath products of symmetric groups - the Inverse Galois Problem for Schubert problems in Grassmannians still open.

2.5 Partial Flag Varieties

As it turns out, the Schubert decomposition and Schubert problems make sense in more general settings than Grassmannians. In this section we introduce partial flag varieties and their Schubert problems, and then in the next section introduce variants of Grassmannians and more generally partial flag varieties.

Recall that the **partial flag variety (or partial flag manifold) of shape** $(a_1, \ldots, a_s; n)$ is the set of partial flags

$$\mathrm{Fl}(a_1, \ldots, a_s; n) = \{(F_{a_1}, \ldots, F_{a_s}) \in \mathrm{Gr}(a_1, n) \times \cdots \times \mathrm{Gr}(a_s, n) \mid F_{a_1} \subseteq \cdots \subseteq F_{a_s}\},$$

which is a subvariety of $\mathrm{Gr}(a_1, n) \times \cdots \times \mathrm{Gr}(a_s, n)$ (considered so after using the s-fold Segre embedding on the products of the Plücker embeddings of the relevant Grassmannians), since inclusions of vector spaces can be encoded by rank conditions of matrices, which are polynomials). Note that Grassmannians are a special type of partial flag varieties: $\mathrm{Gr}(k,n) = \mathrm{Fl}(k;n)$, also called a one-step partial flag variety. Additionally, the **(full, or complete) flag variety** $\mathrm{Fl}(n) = \mathrm{Fl}(1, 2, \ldots, n-1; n)$ is a special case of partial flag variety.

Theorem 2.5.1. *Let $\pi : X \times Y \to Y$ be a surjective morphism of projective varieties, and let $d \in \mathbb{N}$. If $\dim(\pi^{-1}(\{y\}) = d$ for all $y \in Y$, then $\dim(X) = d + \dim(Y)$.*

To compute the dimension of the variety $\mathrm{Fl}(a_1, \ldots, a_s; n)$, we first consider the projection of a **two-step partial flag variety** $\pi : \mathrm{Fl}(a_1, a_2; s) \to \mathrm{Gr}(a_2, s)$. For any a_2-plane $H \in \mathrm{Gr}(a_2; s)$, it fiber $\pi^{-1}(\{H\}) = \{(F_{a_1}, H) \mid F_{a_1} \subseteq H\} \cong \mathrm{Gr}(a_1, a_2)$. Hence by the Theorem 2.5.1,

$\dim(\mathrm{Fl}(a_1, a_2; n)) = \dim(\mathrm{Gr}(a_2, n)) + \dim(\mathrm{Gr}(a_1, a_2)) = a_2(n - a_2) + a_1(a_2 - a_1)$. Similarly, considering the projection $\pi : \mathrm{Fl}(a_1, \ldots, a_s; n) \to \mathrm{Gr}(a_s, n)$, and any a_s-plane $H \in \mathrm{Gr}(a_s, n)$, its fiber $\pi^{-1}(\{H\}) \cong \mathrm{Fl}(a_1, \ldots, a_{s-1}; a_s)$, so using induction on the Theorem on the Dimension of Fibers, $\dim(\mathrm{Fl}(a_1, \ldots, a_s; n) = (\sum_{k=1}^{s-1} a_k(a_{k+1} - a_k)) + a_s(n - a_s)$. In particular, $\dim(\mathrm{Fl}(n)) = 1(2-1)+2(3-2)+\cdots+(n-1)[n-(n-1)] = \sum_{k=1}^{n-1} k = \frac{n(n-1)}{2} = \binom{n}{2}$. We can also see the dimension for partial flag varieties arising from the construction of $\mathrm{Fl}(a_1, \ldots, a_s; n)$ as an algebraic manifold, but first we need the following index sets:

Definition 2.5.2. *A permutation $\sigma \in S_n$ has a **descent** in position $i \in [n-1]$ if $\sigma(i) > \sigma(i+1)$, and the **descent set** of a permutation is $D(\sigma) = \{i \in [n-1] \mid \sigma(i) > \sigma(i+1)\}$. The set of **permutations with possible descents only in positions** $a_1, \ldots, a_s$ is $D(a_1, \ldots, a_s; n) = \{\sigma \in S_n \mid D(\sigma) \subseteq \{a_1, \ldots, a_n\}\}$, which will be our index set for our local coordinate charts and Schubert decomposition of $\mathrm{Fl}(a_1, \ldots, a_s; n)$. Note that $|D(a_1, \ldots, a_s; n)| = \binom{n}{a_1}\binom{n-a_1}{a_2-a_1}\binom{n-a_2}{a_3-a_2}\cdots\binom{n-a_{s-1}}{a_s-a_{s-1}}$.*

The index set $\binom{[n]}{k}$ for the local coordinate charts and Schubert decomposition of $\mathrm{Gr}(k, n)$ can be naturally identified with $D(k; n)$, where in one-line notation, only the first k images of the permutation σ are written to obtain α. This truncated representation α of a permutation σ is called a **partial permutation**, and the full permutation σ can be uniquely recovered by concatenating the integers $[n] \setminus \{\alpha_1, \ldots, \alpha_k\}$, ordered from least to greatest, to the end of α. This works because no descents except in position k guarantee that the first k integers of $\sigma \in D(k; n)$ are increasing, and then a potential descent, and then increasing after that. Similarly, any $\sigma \in D(a_1, \ldots, a_s; n)$ can be represented uniquely by a partial permutation α, by only writing the first a_s images of σ in one-line notation, with the full permutation σ recovered by concatenating with the remaining elements of $[n]$ in increasing order at the end of α. Our convention is to represent elements of $D(a_1, \ldots, a_s; n)$ by their partial permutation representatives α, but it is sometimes useful to complete the permutation to $\sigma \in S_n$. For example, we can define the **length** of permutation to be $|\sigma| = |\{(i, j) \in [n] \times [n] \mid i < j \text{ and } \sigma(i) > \sigma(j)\}$, and this agrees with our definition of length for $\alpha \in \binom{[n]}{k}$. Finally, note that $D(1, 2, \ldots, n-1; n) = S_n$, so the full flag manifold is simply indexed by permutations σ (with no restrictions on descents).

For any $\alpha \in D(a_1, \ldots, a_s; n)$, we obtain a local coordinate chart $U_\alpha \subseteq \mathrm{Fl}(a_1, \ldots, a_s; n)$ as the set of rank-a_s, $n \times a_s$ matrices. Here, F_{a_i} is recovered as the span of the first a_i columns of the matrix for all $i = 1, \ldots, s$. As for the Stiefel coordinates for the Grassmannian, these matrices are only unique up to change of basis, but this change of basis must preserve the chain of vector subspaces. Hence, two such representations are equivalent if there exists an invertible block upper triangular $a_s \times a_s$ matrix, with blocks of size $a_1 \times a_1$, $a_2 \times (a_2 - a_1)$, ..., $a_{s-1} \times (a_s - a_{s-1})$, and $n \times (n - a_s)$. Again there is a normal form for such matrices, with an $a_1 \times a_1$ identity matrix in rows indexed by $\alpha_1, \ldots, \alpha_{a_1}$ (with zeros to the right of this matrix), and an $(a_i - a_{i-1}) \times (a_i - a_{i-1})$ identity matrix in rows indexed by $\alpha_{a_{i-1}+1}, \ldots, \alpha_{a_i}$ (with zeros to the right of this matrix), for all $i = 2, \ldots, s$. The remaining coordinates are the affine charts for $\mathrm{Fl}(a_1, \ldots, a_s; n)$, with transition functions just like the Grassmannian, but with invertible block upper triangular matrices as described above.

For each $\alpha \in D(a_1, \ldots, a_s; n)$, let $\beta^i \in \binom{[n]}{a_i}$ be given by $\beta^i = \mathrm{Sort}(\alpha_1, \ldots, \alpha_{a_i})$ for all $i = 1, \ldots, s$, where here "Sort" means order least to greatest. Given such an $\alpha \in D(a_1, \ldots, a_s; n)$ and complete flag $F_\bullet \in \mathrm{Fl}(n)$, we obtain a Schubert variety $\Omega_\alpha F_\bullet = \{(H_{a_1}, \ldots, H_{a_s}) \in \mathrm{Fl}(a_1, \ldots, a_s; n) \mid H_{a_i} \in \Omega_{\beta^i} F_\bullet \in \mathrm{Gr}(a_i, n) \text{ for all } i = 1, \ldots, s\} \subseteq \mathrm{Fl}(a_1, \ldots, a_s; n)$. We continue to have the Bruhat order on $D(a_1, \ldots, a_s; n)$, and so the corresponding Schubert cells of $\mathrm{Fl}(a_1, \ldots, a_s; n)$ with respect to $\alpha \in D(a_1, \ldots, a_s; n)$ and $F_\bullet \in \mathrm{Fl}(n)$ are $\Omega^\circ_\alpha F_\bullet = \Omega_\alpha F_\bullet \setminus \bigcup_{\gamma \not\leq \alpha} \Omega_\gamma F_\bullet$.

With all of these notions generalized, we again get Schubert problems as the 0-dimensional intersections of Schubert varieties, and corresponding Galois groups as for Grassmannians. Enriched Galois groups for Schubert problems in partial flag varieties are known, but little has been done to understand them in general. We give experimental data in Chapter 4 for all such Galois groups where $n = 6$ and the number of solutions is less than or equal to 250.

2.6 Schubert Problems in Types B, C, and D

So far we have considered Grassmannians, and more generally partial flag varieties, which are chains of arbitrary subspaces of a vector space of specified dimensions. However, if our vector space has additional structure, in this case a non-degenerate bilinear form, we will want our

subspaces to respect this additional structure.

Definition 2.6.1. *Let V be a finite-dimensional vector space over a field F. A **non-degenerate bilinear form** is a 2-multilinear form $\langle \cdot, \cdot \rangle : V \times V \to F$ such that $\langle x, y \rangle = 0$ for all $y \in V \implies x = \vec{0} \in V$. A non-degenerate bilinear form is **symmetric** if $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in V$, and it is **skew-symmetric** if $\langle x, y \rangle = -\langle y, x \rangle$ for all $x, y \in V$.*

In coordinates, after fixing an ordered basis $\{b_1, \ldots, b_n\}$ of V, a non-degenerate bilinear form on V is represented by an invertible matrix M, such that for all $x, y \in V$ (expressed as column vectors in the usual fashion), $\langle x, y \rangle = x^T M y$ (where x^T is the transpose of x, so a row vector). The matrix M is found by $M_{ij} = \langle e_i, e_j \rangle$. Expressed this way, a symmetric non-degenerate bilinear form is one represented by a symmetric matrix $M^T = M$, and a skew-symmetric non-degenerate bilinear form is one represented by a skew-symmetric matrix $M^T = -M$. Inner products are special cases of non-degenerate symmetric bilinear forms over real or complex vector spaces, where the form is further required to be positive-definite, i.e. $\langle x, x \rangle \geq 0$ for all $x \in V$, with equality achieved if and only if $x = \vec{0} \in V$. Once we have such a non-degenerate bilinear form (symmetric or skew-symmetric), we can define analogues of partial flag varieties and their Schubert problems.

Definition 2.6.2. *Let V be an n-dimensional vector space over a field F, with symmetric or skew-symmetric non-degenerate bilinear form $\langle \cdot, \cdot \rangle$. Then, for a subspace $W \subseteq V$, we define its **annihilator** to be $W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}$. A subspace $W \subseteq V$ is **isotropic** (with respect to $\langle \cdot, \cdot \rangle$) if $W \subseteq W^\perp$. In other words, $W \subseteq V$ is isotropic if the bilinear form restricted to W $\langle \cdot, \cdot \rangle|_W : W \times W \to F$ is identically the zero-map: $\langle w_1, w_2 \rangle = 0$ for all $w_1, w_2 \in W$. A partial flag $F_{a_1} \subseteq \cdots \subseteq F_{a_s} \in \mathrm{Fl}(a_1, \ldots, a_s; n)$ is also said to be **isotropic** (with respect to $\langle \cdot, \cdot \rangle$) if $F_{a_s} \subseteq F^n$ is isotropic. We classify the corresponding sets of isotropic flags as follows (where for Types B, C, and D, we require $2a_s < n$):*

- *Type A partial flag varieties $\mathrm{Fl}_A(a_1, \ldots, a_s; n)$: The usual partial flag variety $\mathrm{Fl}(a_1, \ldots, a_s; n)$ that we studied in the previous section.*

- *Type B partial flag varieties $\mathrm{Fl}_B(a_1, \ldots, a_s; n)$: the set of isotropic flags of shape $(a_1, \ldots, a_s; n)$ with respect to a symmetric bilinear form $\langle \cdot, \cdot \rangle$, and where n is odd.*

- *Type C partial flag varieties $\mathrm{Fl}_C(a_1, \ldots, a_s; n)$: the set of isotropic flags of shape $(a_1, \ldots, a_s; n)$ with respect to a skew-symmetric bilinear form $\langle \cdot, \cdot \rangle$. Skew-symmetric bilinear forms only can be defined for even-dimensional vector spaces, and so n is even.*

- *Type D partial flag varieties $\mathrm{Fl}_D(a_1, \ldots, a_s; n)$: the set of isotropic flags of shape $(a_1, \ldots, a_s; n)$ with respect to a symmetric bilinear form $\langle \cdot, \cdot \rangle$, and where n is even.*

*For Type C, since $n = 2k$ is even, we call a maximal isotropic subspace $W \subseteq V$ a **Lagrangian subspace**. It turns out that an isotropic subspace $W \subseteq V$ is Lagrangian if and only if $\dim(W) = k$ (half of the dimension of V). We thus define the **Lagrangian Grassmannian** to be $\mathrm{LG}(k) = \mathrm{Fl}_C(k; 2k)$.*

These partial flag varieties of types B, C, and D also have analogues of Schubert varieties, but this time indexed by signed permutations instead of traditional permutations. If $\dim(V) = 2k$ or $2k + 1$, a **signed permutation** σ is like a permutation in S_k, written in one-line notation, but where each $i \in [k]$ can have a bar over it or not. For example, $2\overline{1}3$ and $\overline{1}3\overline{2}$ are signed permutations when $k = 3$, but $1\overline{1}2$ is not. Hence, there are $2^k \cdot k!$ such signed permutations. For types B and C, we index using signed permutations with no other restrictions, but for type D we index via signed permutations with an even number of bars over the $i \in [k]$ in the one-line notation for the permutations.

Signed permutations can be viewed as traditional partial permutations in S_{2k} or S_{2k+1} by replacing $\overline{i}$ with $k + i$ for each $i \in [k]$ with a bar on top. Thus, completing this partial permutation, we can also view signed permutations as special types of full permutations. In this way, we obtain affine local coordinate charts, Schubert cells, and Schubert varieties with the same definition as for type A partial flag varieties (descent restrictions and all), just requiring our subspaces to be isotropic with respect to the corresponding symmetric or skew-symmetric bilinear form, and for the full flags $F_\bullet$ used to be isotropic to this form as well.

Having such a Schubert cell decomposition (paving by affines), again the classes of the Schubert varieties (closures of Schubert cells) give a basis for the Chow ring of $\mathrm{Fl}_S(a_1, \ldots, a_s; n)$ (where $S \in \{A, B, C, D\}$). Intersections of Schubert varieties can always be shown to be transverse, and so multiplication in this ring corresponds to these intersections of Schubert varieties, governing all intersections of subvarieties in the partial flag varieties. We discuss how to work computationally with these Chow groups in the first section of Chapter 3. Intersections that are 0-dimensional again give Schubert problems, the solutions being the number of points in the intersection. These Schubert problems in types B, C, and D all have corresponding Galois groups as well (in the same way as for type A), and these groups can be shown to be enriched or not. The classification (and even computation) of such Schubert Galois groups is in the very early stages, and is currently an open problem. In Chapter 3, we present a software package that will find the number of solutions to Schubert problems in partial flag varieties of types A, B, C, and D. In Chapter 4, we use this package to investigate enriched Galois groups for such problems in types A and C.

3.1 Cohomology Computations

In the previous chapter, we introduced Grassmannians and the partial flag varieties of types A, B, C, and D. In each case, given a complete flag $F_\bullet$, we gain a partition of the variety into Schubert cells, called its Schubert decomposition. The closures of these cells are called Schubert varieties, and a 0-dimensional intersection of some Schubert varieties (each with respect to a different flag) constitutes a Schubert problem. Additionally, the rational equivalence classes of the Schubert varieties form a basis for the Chow ring of the variety, which encapsulates all intersection-theoretic information of the variety.

But how does one actually solve a Schubert problem (finding the number of points of intersection of Schubert varieties)? One way to do this is to first focus on the complete flag varieties $\mathrm{Fl}_S(n)$, where $S \in \{A, B, C, D\}$. For any partial flag variety $\mathrm{Fl}_S(a_1, \ldots, a_s; n)$, we consider the projection $\pi : \mathrm{Fl}_S(n) \to \mathrm{Fl}_S(a_1, \ldots, a_s; n)$ by forgetting some of the flags. Given any Schubert variety $\Omega_\alpha F_\bullet \subseteq \mathrm{Fl}_S(a_1, \ldots, a_s; n)$, indexed by a partial (possibly signed) permutation α and with respect to a specific complete flag $F_\bullet$, note that its fiber under the projection $\pi^{-1}(\{\Omega_\alpha F_\bullet\}) = \Omega_{\tilde{\alpha}} F_\bullet \subseteq \mathrm{Fl}_S(n)$ is a Schubert variety in the complete flag variety, where $\tilde{\alpha}$ is α completed to a full (possibly signed) permutation. While by construction the codimensions of the varieties $\Omega_\alpha F_\bullet$ and $\Omega_{\tilde{\alpha}} F_\bullet$ are the same, the dimensions are not, since they live in different spaces. However, there is a distinguished Schubert variety $\Omega_{a_1,\ldots,a_s} \subseteq \mathrm{Fl}_S(n)$ such that a dense open subset of $\Omega_{a_1,\ldots,a_s}$ is isomorphic to a dense open subset of $\mathrm{Fl}_S(a_1, \ldots, a_s; n)$, and so a dense open subset of the original Schubert variety $\Omega_\alpha F_\bullet$ can be viewed as the intersection of dense open subsets of $\Omega_{a_1,\ldots,a_s}$ and $\Omega_{\tilde{\alpha}} F_\bullet \in \mathrm{Fl}_S(n)$. This construction works for intersections as well, and so in other words we can consider any Schubert problem $\Omega_{\alpha^1} F_\bullet^1 \cap \cdots \cap \Omega_{\alpha^r} F_\bullet^r \subseteq \mathrm{Fl}_S(a_1, \ldots, a_s; n)$ as the Schubert problem in the complete flag variety $\Omega_{a_1,\ldots,a_s} \cap \Omega_{\tilde{\alpha}^1} F_\bullet^1 \cap \cdots \cap \Omega_{\tilde{\alpha}^r} F_\bullet^r \subseteq \mathrm{Fl}_S(n)$, at least in local coordinates. Hence, for the rest of this section, we only consider the complete flag varieties

$\mathrm{Fl}_S(n)$, since all Schubert problems in partial flag varieties can be interpreted in this setting.

Recall that the i**th elementary symmetric polynomial** $e_i \in \mathbb{Z}[x_1, \ldots, x_n]$ is $e_i(x_1, \ldots, x_n) = \sum_{1 \leq k_1 \leq \cdots \leq k_i \leq n} x_{k_1} \cdots x_{k_i}$. Then, the Chow rings of the complete flag varieties are each isomorphic to the quotient of a polynomial ring by an ideal involving elementary symmetric polynomials:

- $A(\mathrm{Fl}_A(n)) \cong \mathbb{Z}[x_1, \ldots, x_n]/(e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n))$
- $A(\mathrm{Fl}_B(n)) \cong \mathbb{Z}[x_1, \ldots, x_n]/((e_1(x_1^2, \ldots, x_n^2), \ldots, e_n(x_1^2, \ldots, x_n^2))$
- $A(\mathrm{Fl}_C(n)) \cong \mathbb{Z}[x_1, \ldots, x_n]/(e_1(x_1^2, \ldots, x_n^2), \ldots, e_n(x_1^2, \ldots, x_n^2))$
- $A(\mathrm{Fl}_D(n)) \cong \mathbb{Z}[x_1, \ldots, x_n]/(e_1(x_1^2, \ldots, x_n^2), \ldots, e_{n-1}(x_1^2, \ldots, x_n^2), e_n(x_1, \ldots, x_n))$

The images of the Schubert classes in $A(\mathrm{Fl}_S(n))$ form an integer basis for the corresponding polynomial quotient ring, and there are polynomials in $\mathbb{Z}[x_1, \ldots, x_n]$, called **Schubert polynomials of Type** S, whose images in the quotient ring represent the images of the Schubert classes. These Schubert polynomials will be indexed by (possibly signed) permutations as were the Schubert varieties, and the degree of a Schubert polynomial corresponds to the codimension of the corresponding Schubert variety. In particular, the above isomorphisms are of graded rings, so the top graded part of the Chow ring being 1-dimensional implies that the top graded part of the polynomial quotient ring is also 1-dimensional. This top part has the class of a single Schubert polynomial as its basis element, corresponding to the unique top codimension (dimension of $\mathrm{Fl}_S(n)$) Schubert variety (the class of a point). Since multiplication of Schubert varieties in the Chow ring corresponds to their intersection, the same is true of the multiplication of the corresponding Schubert polynomials. Hence, for a Schubert problem in $\mathrm{Fl}_S(n)$, one can solve the problem (number of points of intersection) by multiplying the corresponding Schubert polynomials together. Since the degrees will add to be the top codimension (since the intersection is 0-dimensional), after quotienting out by the ideal the product will simply be a positive integer multiplied by the unique Schubert polynomial corresponding to the class of a point, the integer being the desired solution to the Schubert problem.

In practice, polynomial multiplication in a quotient ring is best done using a computer algebra system, like Macaulay2. We have developed a software package, called SchubertIdeals.m2, in Macaulay2, which is dedicated to studying Schubert problems in partial flag varieties. The first routines from our package grant a user the ability to generate Schubert polynomials, and multiply them together in the relevant quotient ring to solve Schubert problems. We first describe how to construct Schubert polynomials, and give examples along the way. We do this first for $\mathrm{Fl}(n) = \mathrm{Fl}_A(n)$, where Schubert polynomials will be indexed by usual permutations in S_n, and then adapt for the other types indexed by signed permutations. For the remainder of this section, for readability we write $\omega \in S_n$ in an adapted one-line notation with brackets as $[\omega_1, \ldots, \omega_n]$ rather than $\omega(1) \cdots \omega(n)$.

For $i = 1, \ldots, n-1$, let $\sigma_i = [1, \ldots, i+1, i, \ldots, n]$ be the simple transposition which swaps entry i and entry $i+1$ when multiplying on the right (or left) of any permutation: If $\omega = [\omega_1, \ldots, \omega_n] \in S_n$, $\omega\sigma_i = [\omega_1, \ldots, \omega_{i+1}, \omega_i, \ldots, \omega_n]$. Then, it is well known that the set of simple transpositions $\{\sigma_i \mid i = 1, \ldots, n\}$ generates S_n, and satisfy

$$\sigma_i^2 = 1, \text{ if } |i-j| > 1, \text{ then } \sigma_i\sigma_j = \sigma_j\sigma_i, \text{ and } \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$$

. Recall that a permutation $\omega \in S_n$ has an inversion (i, j) if $i < j$ but $\omega_i > \omega_j$, and that the length of the permutation $\ell(\omega)$ is the number of inversions in ω. If $\omega \in S_n$ can be written as a product of simple transpositions, $\omega = \sigma_{a_1} \cdots \sigma_{a_l}$, where $l = \ell(\omega)$, we say that the sequence $a_1, \ldots, a_l$ is a **reduced word** for ω. Reduced words for a permutation are not unique, but they all have the same length. The permutation with the longest length is $\omega_0 = [n, n-1, \ldots, 1]$, and has length $\ell(\omega_0) = \binom{n}{2}$, which is the dimension of $\mathrm{Fl}(n)$. The corresponding Schubert variety (and later Schubert polynomial) will have codimension (degree) $\binom{n}{2}$, representing the class of a point (our unique top degree basis element discussed earlier). On the other hand, the identity permutation $e = [1, 2, \ldots, n]$ has length $\ell(e) = 0$, and so corresponds to all of $\mathrm{Fl}(n)$.

We define an action of S_n on $\mathbb{Z}[x_1, \ldots, x_n]$ as follows: for $\omega = [\omega_1, \ldots, \omega_n] \in S_n$ and $f \in$

$\mathbb{Z}[x_1, \ldots, x_n]$, $\omega f(x_1, \ldots, x_n) = f(x_{\omega_1}, \ldots, x_{\omega_n})$. From this action, for each $i = 1, \ldots, n-1$, we define the **divided difference operator** $\partial_i : \mathbb{Z}[x_1, \ldots, x_n] \to \mathbb{Z}[x_1, \ldots, x_n]$ by $\partial_i f(x_1, \ldots, x_n) = \frac{f(x_1,\ldots,x_n) - \sigma_i f(x_1,\ldots,x_n)}{x_i - x_{i+1}}$, where again σ_i is the ith simple transposition. Since $\sigma_i(f - \sigma_i(f)) = -(f - \sigma_i(f))$, it is divisible by $x_i - x_{i+1}$, and so $\partial_i(f)$ is a polynomial of degree 1 less than the degree of f. As an example, $\partial_2(x_1^2 x_2) = \frac{x_1^2 x_2 - x_1^2 x_3}{x_2 - x_3} = x_1^2$. Furthermore, $\partial_i^2 = 0$, if $|i - j| > 1$ then $\partial_i \partial_j = \partial_j \partial_i$, and $\partial_i \partial_{i+1} \partial_i = \partial_{i+1} \partial_i \partial_{i+1}$ (similar to the properties observed for the σ_i). Using these properties, it can be shown that independent of choice of reduced word for $\omega = \sigma_{a_1} \cdots \sigma_{a_l}$, that $\partial_\omega = \partial_{a_1} \cdots \partial_{a_l}$ is well-defined.

Definition 3.1.1. *For each $\omega \in S_n$, the **Schubert polynomial** indexed by ω is given by $\mathfrak{S}_\omega = \partial_{\omega^{-1}\omega_0}(x_1^{n-1} x_2^{n-2} \cdots x_{n-1})$, where $\partial_{\omega^{-1}\omega_0} = \partial_{a_1} \cdots \partial_{a_l}$ for any reduced word $a_1, \ldots, a_l$ for $\omega^{-1}\omega_0 \in S_n$, and where $\omega_0 = [n, n-1, \ldots, 1]$ is the longest length element of S_n.*

For some examples, note that

- $\mathfrak{S}_{\omega_0} = x_n^{n-1} x_{n-1}^{n-2} \cdots x_1$ (the **staircase monomial**, which is our unique top degree Schubert polynomial representing the class of a point)
- $\mathfrak{S}_{\sigma_i} = x_1 + \cdots + x_i$
- $\mathfrak{S}_e = 1$ (representing all of $\mathrm{Fl}(n)$)

We now share our Macaulay2 code for working with the Chow ring

$$A(\mathrm{Fl}_A(n)) \cong \mathbb{Z}[x_1, \ldots, x_n]/(e_1(x_1, \ldots, x_n), \ldots, e_n(x_1, \ldots, x_n))$$

and its basis of Schubert polynomials, along with sample output from tests showing how the code works in practice. In the Macaulay2 code, "- -" at the beginning of a line indicates a comment about the functionality of the forthcoming function. Each function in the package is called a "method" in Macaulay2, and the input types are given (like "List", or "ZZ" for $\mathbb{Z}$, etc). In the examples after each function, values like $i132$ correspond to user input, and $o132$ correspond to

Macaulay2's output. Semicolons are placed occasionally at the end of input to suppress long output for readability.

```
-- Completes a partial permutation into a full one.
completePermutation = method(TypicalValue=>List)
completePermutation(List,ZZ):=(w,n) ->(
      wcomplete := w;
      for i from 1 to n do(
            if isSubset({i},wcomplete)==false then wcomplete=append(wcomplete,i));
      return(wcomplete))

-- EXAMPLE of completePermutation:
i132 : completePermutation({7,8,4,1,2,3},9)

o132 = {7, 8, 4, 1, 2, 3, 5, 6, 9}
o132 : List

-- Gives the length of a partial permutation.
typeALength = method()
typeALength(List,ZZ) := (w,n) -> (
      wcomp := completePermutation(w,n);
      count := 0;
      for i from 1 to n do
            for j from i+1 to n do
                  if wcomp_(i-1) > wcomp_(j-1) then count = count+1;
      return(count))

-- EXAMPLE of typeALength:
i133 : typeALength({7,8,4,1,2,3},9)

o133 = 15

-- Computes a reduced word for a permutation.
bubbleSort = method()
bubbleSort(List) := (L) -> (
      n := length(L);
      sorted := reverse(sort(L));
      swaps := {};
      while (L != sorted) do(
            for i from 0 to (n-2) do(
                  if L_(i) < L_(i+1) then(
```

```
                        swaps = prepend(i,swaps);
                        L = switch(i,i+1,L);
                        break)));
      return(swaps))

-- EXAMPLES of bubbleSort:

i134 : bubbleSort({1,2,3,4})

o134 = {0, 1, 2, 0, 1, 0}

o134 : List

i135 : bubbleSort({4,3,2,1})

o135 = {}

o135 : List

-- Applies a divided difference operator to a polynomial.
deltaSwapA = method()
deltaSwapA(Thing,Ring,ZZ) := (f,R,k) -> (
      ringvars := gens R;
      fnew := sub((f-sub(f,{ringvars_(k)=>ringvars_(k+1),ringvars_(k+1)=>ringvars_(k)}))/
      (ringvars_(k)-ringvars_(k+1)),R);
      return(fnew))

-- EXAMPLES of deltaSwapA:

i136 : R = QQ[a,b,c,d]

o136 = R

o136 : PolynomialRing

i137 : f = a^3*b^2*c

        3 2
o137 = a b c

o137 : R
```

```
i138 : deltaSwapA(f,R,0)

        2 2
o138 = a b c

o138 : R

i139 : deltaSwapA(f,R,1)

        3
o139 = a b*c

o139 : R

i140 : deltaSwapA(f,R,2)

        3 2
o140 = a b

o140 : R

-- Computes the Schubert polynomial given a reduced word for a permutation.
polyRepA = method();
polyRepA(List,Ring) := (w,R) -> (
      ringvars := gens R;
      n := length(ringvars);
      pointclass := 1;
      for i from 1 to (n-1) do(
            pointclass = pointclass*(ringvars_(i-1))^(n-i));
      polyrep := pointclass;
      for i in w do(
            polyrep = deltaSwapA(polyrep,R,i));
      return(polyrep))

-- EXAMPLES of polyRepA:

i141 : polyRepA(bubbleSort({1,4,3,2}),QQ[a,b,c,d])

        2       2    2             2
o141 = a b + a*b  + a c + a*b*c + b c
```

```
o141 : QQ[a..d]

i142 : for perm in permutations({1,2,3,4}) do(
          print(perm);
          print(polyRepA(bubbleSort(perm),QQ[a,b,c,d])))
{1, 2, 3, 4}
1
{1, 2, 4, 3}
a + b + c
{1, 3, 2, 4}
a + b
{1, 3, 4, 2}
a*b + a*c + b*c
{1, 4, 2, 3}
 2          2
a  + a*b + b
{1, 4, 3, 2}
 2       2    2             2
a b + a*b  + a c + a*b*c + b c
{2, 1, 3, 4}
a
{2, 1, 4, 3}
 2
a  + a*b + a*c
{2, 3, 1, 4}
a*b
{2, 3, 4, 1}
a*b*c
{2, 4, 1, 3}
 2       2
a b + a*b
{2, 4, 3, 1}
 2         2
a b*c + a*b c
{3, 1, 2, 4}
 2
a
{3, 1, 4, 2}
 2     2
a b + a c
```

```
{3, 2, 1, 4}
 2
a b
{3, 2, 4, 1}
 2
a b*c
{3, 4, 1, 2}
 2 2
a b
{3, 4, 2, 1}
 2 2
a b c
{4, 1, 2, 3}
 3
a
{4, 1, 3, 2}
 3     3
a b + a c
{4, 2, 1, 3}
 3
a b
{4, 2, 3, 1}
 3
a b*c
{4, 3, 1, 2}
 3 2
a b
{4, 3, 2, 1}
 3 2
a b c

-- Computes a polynomial ring in n variables and the ideal inside generated by the elementary
-- symmetric polynomials in that many variables.
elementarySymmetricIdeal = method()
elementarySymmetricIdeal(ZZ) := (n) -> (
      R := QQ[y_(1)..y_(n)][t];
      f :=1_R;
      for i from 1 to n do(
      f = f*(y_(i)+t));
      coeffs := (coefficients (f-t^n))_(1);
      S := QQ[y_(1)..y_(n)];
```

```
      I := sub(ideal(coeffs),S);
      return(S,I))

-- EXAMPLES of elementarySymmetricIdeal:

i143 : elementarySymmetricIdeal(3)

o143 = (QQ[y ..y ], ideal (y  + y  + y , y y  + y y  + y y , y y y ))
            1   3           1    2    3   1 2    1 3    2 3   1 2 3

o143 : Sequence

i144 : (S,I) = elementarySymmetricIdeal(4);

o144 : Sequence

-- Computes the intersection number of a Schubert problem using cohomology in the full Type A flag
-- manifold.
intA = method()
intA(List,Ring,Ideal) := (alphas,S,I) -> (
      n := numgens S;
      f := 1_S;
      for alpha in alphas do(
            f = f*polyRepA(bubbleSort(alpha),S));
      f = f % I;
      g := polyRepA(bubbleSort(reverse(toList(1..n))),S) % I;
      numsols := f / g;
      return(numsols))

-- EXAMPLES of intA:

i145 : intA({{2,1,3,4},{3,4,2,1}},S,I)

o145 = 1

o145 : frac S

i146 : intA({{2,1,3,4},{4,3,1,2}},S,I)

o146 = 0
```

```
o146 : frac S

i147 : intA({{2,1,3,4},{1,3,2,4},{1,3,2,4},{1,3,2,4},{1,3,2,4},{1,2,4,3}},S,I)

o147 = 2

o147 : frac S

-- Computes the intersection number of a Schubert problem using cohomology in a general Type A
-- partial flag manifold.
partialIntA = method()
partialIntA(List,List,Ring,Ideal) := (flagshape,alphas,S,I) -> (
      l := length(flagshape);
      n := flagshape_(-1);
      newalphas := {};
      for alpha in alphas do(
            newalpha := completePermutation(alpha,n);
            newalphas = append(newalphas,newalpha));
      dualclass := {};
      for k from 1 to (l-1) do(
            for j from (flagshape_(-(k+1)) + 1) to flagshape_(-k) do(
          dualclass = prepend(j,dualclass)));
      for i from 1 to flagshape_(0) do(
            dualclass = prepend(i,dualclass));
      newalphas = append(newalphas,dualclass);
      numsols := intA(newalphas,S,I);
      return(numsols))

-- EXAMPLES of partialIntA:

i148 : (S,I) = elementarySymmetricIdeal(5);

o148 : Sequence

i149 : partialIntA({1,2,5},{{2,1},{2,1},{1,3},{1,3},{1,3},{1,3},{1,3}},S,I)

o149 = 5

o149 : frac S

i150 : partialIntA({1,2,5},{{2,1},{2,1},{2,1},{1,3},{1,3},{1,3},{1,3}},S,I)
```

```
o150 = 3

o150 : frac S

i151 : (S,I) = elementarySymmetricIdeal(6);

o151 : Sequence

i152 : partialIntA({2,4,6},{{1,4,2,5},{1,4,2,5},{1,4,2,5},{1,4,2,5}},S,I)

o152 = 6

o152 : frac S
```

To obtain the analogues of Schubert polynomials for types B, C, and D, we need general definitions of root systems and Weyl groups for the classical groups of those types. For the following discussion, let V be a vector space over $\mathbb{Q}$ with a positive definite symmetric bilinear form (α, β). Each vector $\alpha \in V$ determines a reflection σ_α found by fixing the hyperplane perpendicular to α and sending α to $-\alpha$.

Definition 3.1.2. *A **root system** is a subset $R \subseteq V$ such that*

1. *R is finite and $0 \notin R$*
2. *If $\alpha \in R$, then σ_α leaves R invariant*
3. *If $\alpha, \beta \in R$, then $(\alpha, \beta) \in \mathbb{Z}$*

The root systems of types A, B, C, and D (described shortly) will additionally all be reduced ($\alpha \in R \implies -\alpha$ is the only scalar multiple of α in R) and irreducible (cannot be expressed as the union of two proper root systems). Hence, we will assume all root systems discussed from now on to have these additional properties.

Definition 3.1.3. *The **Weyl group** W associated to a root system R is the group generated by $\{\sigma_\alpha \mid \alpha \in R\}$.*

Let $\{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{Q}^n$. Then, S_n is generated by the reflections $\sigma_i = \sigma_{e_{i+1}-e_i}$, which swaps e_i and e_{i+1} (corresponding to our simple transpositions earlier). Thus, the Weyl group corresponding to the root system $R = A_{n-1}$ is S_n. In general, like vector spaces, root systems have bases, but the definition differs slightly than for that of a vector space.

Definition 3.1.4. *A **basis** of a root system R is a subset $B \subseteq R$ such that*

1. *B is linearly independent*
2. *For each $\alpha \in R$, there exists a collection $\{a_\beta \in \mathbb{Z} \mid \beta \in B\}$, with all $a_\beta \geq 0$ or with all $a_\beta \leq 0$, such that $\alpha = \sum_{\beta \in B} a_\beta \beta$*

Proposition 3.1.5. *If B is a basis for a root system R, and if W is the Weyl group associated to R, then $\{\sigma_\beta \mid \beta \in B\}$ generates W, and is a minimal generating set for W.*

Reduced and irreducible root systems can be completely classified into one of nine types: four infinite families A_n, B_n, C_n, and D_n (for any $n \in \mathbb{Z}_{>0}$, and five exceptional root systems E_6, E_7, E_8, F_4, G_2. We only consider the four infinite families here, but all we discuss can also be obtained for the exceptional groups. Here are the bases for the root systems that we will consider (subscripts corresponding to the number of basis elements):

- $A_{n-1} : B = \{e_{i+1} - e_i \mid i = 1, \ldots, n-1\}$
- $B_n : B = \{e_1\} \cup \{e_{i+1} - e_i \mid i = 1, \ldots, n-1\}$
- $C_n : B = \{2e_1\} \cup \{e_{i+1} - e_i \mid i = 1, \ldots, n-1\}$
- $D_n : B = \{e_1 + e_2\} \cup \{e_{i+1} - e_i \mid i = 1, \ldots, n-1\}$

From Proposition 3.1.5, we can then compute generators for the corresponding Weyl groups of each type (which will be our indexing sets for the Schubert varieties/polynomials in our full flag manifolds). For example, the linear transformation taking $e_{i+1} - e_i$ to its negative and fixing all other basis elements is the simple transposition σ_i, again showing that $W_{A_{n-1}} = S_n$.

Each of the other types of root systems B_n, C_n, and D_n contain the basis for A_{n-1}, and so the simple transpositions $\{\sigma_i \mid i = 1, \ldots, n-1\}$ are generators for their Weyl groups as well, but in each case there is one additional generator. For B_n, e_1 is also a basis element, and the linear transformation taking e_1 to $-e_1$ can be viewed as a new operation on a permutation ω by sending ω_1 to $-\omega_1$. We call this new transformation σ_0, so that $\omega\sigma_0 = [\omega_1, \omega_2, \ldots, \omega_n]\sigma_0 = [-\omega_1, \omega_2, \ldots, \omega_n]$. The Weyl group W_{B_n} is then the group generated by $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$, also called the **hyperoctahedral group on n letters**, or the **group of signed permutations**. One can represent W_{B_n} as the group of permutation matrices, but where the 1's can be $+1$ or -1 (so $|W_{B_n}| = 2^n \cdot n!$). Another way to represent W_{B_n} is like the bracketed one-line notation we've been using this section for S_n, but where there can be bars over elements (thought of as a negative sign, so applying the bar twice cancels the operation). Like for S_n, we now want to define a notion of length for elements of W_{B_n}, and we do so for all general Weyl groups.

Definition 3.1.6. *Let $B = \{\beta_1, \ldots, \beta_n\}$ be an ordered basis for a root system R, and let W be the corresponding Weyl group. Hence, W is generated by $\{\sigma_{\beta_1}, \ldots, \sigma_{\beta_n}\}$. If $\omega \in W$, then $\omega = \sigma_{\beta_{i_1}} \cdots \sigma_{\beta_{i_l}}$, and if l is the minimal number of generators required to write ω, then we say $l = \ell(\omega)$ is the (Coxeter) **length** of ω, and that $\sigma_{\beta_{i_1}}, \ldots, \sigma_{\beta_{i_l}}$ is a **reduced word** for ω.*

The longest length element of W_{B_n} is $\omega_0 = [\overline{1}, \ldots, \overline{n}]$, which has length n^2 (the dimension of $\mathrm{Fl}_B(n)$). For C_n, everything carries over as for B_n, since the new basis element is $2e_1$, which also gets sent to its negative via σ_0. Therefore, $W_{C_n} = W_{B_n}$.

For D_n, we have the reflection sending $e_1 + e_2$ to $-e_1 - e_2$ that sends e_1 to $-e_2$ and e_2 to $-e_1$ (since this fixes the perpendicular hyperplane). Let $\sigma_{\overline{1}}$ act on S_n by $\omega\sigma_{\overline{1}} = [\omega_1, \omega_2, \omega_3, \ldots, \omega_n]\sigma_{\overline{1}} = [\overline{\omega_2}, \overline{\omega_1}, \omega_3, \ldots, \omega_n]$. Since $\sigma_{\overline{1}} = \sigma_0\sigma_1\sigma_0$, $W_{D_n} \subseteq W_{B_n}$ is the subgroup of signed permutations with a number of -1's equivalent to $n \mod 2$. The longest length element of W_{D_n} is then the signed permutation $\omega_0 = [\pm 1, \overline{2}, \ldots, \overline{n}]$, where $+1$ or -1 is chosen to give an even number of sign changes depending on whether n is odd or even. Either way, its length is $n^2 - n$ (the dimension of $\mathrm{Fl}_D(n)$). To summarize:

- $W_{A_{n-1}} = (\sigma_i \mid i = 1, \ldots, n-1)$

- $W_{B_n} = (\sigma_0) \cup (\sigma_i \mid i = 1, \ldots, n-1)$
- $W_{C_n} = (\sigma_0) \cup (\sigma_i \mid i = 1, \ldots, n-1)$
- $W_{D_n} = (\sigma_{\overline{1}}) \cup (\sigma_i \mid i = 1, \ldots, n-1)$

Just like permutations, Weyl group generators act on polynomials $f \in \mathbb{Z}[x_1, \ldots, x_n]$. Our new generators for types B, C, and D act as

$\sigma_0 f(x_1, x_2 \ldots, x_n) = f(-x_1, x_2, \ldots, x_n)$, and $\sigma_{\overline{1}} f(x_1, x_2, x_3, \ldots, x_n) = f(-x_2, -x_1, x_3, \ldots, x_n)$

. Additionally, we associate to every root $\alpha \in R$ the equation of its perpendicular hyperplane $\gamma(\alpha)$, where in general $\gamma(\alpha)$ is obtained from α by replacing e_i with x_i. For example, $\gamma(e_{i+1} - e_i) = x_{i+1} - x_i$ for all $i = 1, \ldots, n-1$.

Definition 3.1.7. *Let R be a root system with basis B. For each root $\alpha \in B$, define the **divided difference operator** $\partial_\alpha : \mathbb{Z}[x_1, \ldots, x_n] \to \mathbb{Z}[x_1, \ldots, x_n]$ by $\partial_\alpha f = \frac{f - \sigma_\alpha f}{-\gamma(\alpha)}$.*

Proposition 3.1.8. *$f - \sigma_\alpha f$ is divisible by $\gamma(\alpha)$, and so $\partial_\alpha f$ really is a polynomial.*

Proof. Every point in the hyperplane perpendicular to α is fixed by σ_α. Therefore, $f - \sigma_\alpha f$ is 0 whenever $\gamma(\alpha) = 0$. From commutative algebra, this implies that the ideal generated by $f - \sigma_\alpha f$ is contained in the ideal generated by $\gamma(\alpha)$. Hence, $f - \sigma_\alpha f = g\gamma(\alpha)$ for some polynomial g. □

For the various types (classified by their corresponding root systems), we have the following divided difference operators:

- A, B, C, D: $\partial_i f = \frac{f - \sigma_i f}{x_i - x_{i+1}}$ $(i = 1, \ldots, n-1)$
- $B : \partial_0^B f = \frac{f - \sigma_0 f}{-x_1}$
- $C : \partial_0^C f = \frac{f - \sigma_0 f}{-2x_1}$
- $D : \partial_{\overline{1}} f = \frac{f - \sigma_{\overline{1}} f}{-x_1 - x_2}$

We then obtain Schubert polynomials for types B, C, and D as we did for type A, but using the analogous longest length element ω_0 and notion of reduced word (with the new generators). We give our Macaulay2 code for these types, starting with Type C, but since this is similar to that for Type A, we omit comments and examples.

```
completeSignedPermutation = method()
completeSignedPermutation(List,ZZ) := (w,n) -> (
      wnew := w;
      for i from 1 to n do(
if (isSubset({i},wnew)==false and isSubset({-i},wnew)==false) then wnew=append(wnew,i));
      return(wnew))

typeCLength = method()
typeCLength(List) := (w) -> (
      n := length(w);
      count := 0;
      for i from 1 to n do
            for j from i+1 to n do
                  if w_(i-1) > w_(j-1) then count = count+1;
      for i from 1 to n do
            for j from i to n do
                  if w_(i-1) + w_(j-1) > 2*n+1 then count = count+1;
      return(count))

signedToNot = method()
signedToNot(List) := (perm) -> (
      n := length(perm);
      wnew := {};
      for i from 1 to n do(
            if perm_(i-1) > 0 then wnew = append(wnew,perm_(i-1));
    if perm_(i-1) < 0 then wnew = append(wnew,2*n+1+perm_(i-1)));
      return wnew)

signedBubbleSort = method()
signedBubbleSort(List) := (L) -> (
      n := length(L);
      eventual := {};
      for i from 1 to n do(
            eventual = append(eventual,2*n+1-i));
      swaps := {};
```

```
      while (L != eventual) do(
            if L == reverse(sort(L)) then(
          smallest := L_(-1);
          L = drop(L,-1);
  L = append(L,2*n+1-smallest);
  swaps = prepend(n-1,swaps));
            for i from 0 to (n-2) do(
                  if L_(i) < L_(i+1) then(
                        swaps = prepend(i,swaps);
                        L = switch(i,i+1,L);
                        break)));
      return(swaps))

deltaSwapC = method()
deltaSwapC(Thing,Ring) := (f,R) -> (
      ringVars := gens R;
      return sub((f-sub(f,{ringVars_(-1)=>(-1)*ringVars_(-1)}))/(2*ringVars_(-1)),R))

elementarySchurDeterminantC = method()
elementarySchurDeterminantC(List,ZZ) := (lambda,n) -> (
      Rt := QQ[y_(1)..y_(n)][t];
      f := 1_(Rt);
      for i from 1 to n do(
            f = f*(y_(i)+t));
      elempolys := ((coefficients (f-t^n))_(1))_(0);
      S := QQ[y_(1)..y_(n)];
      fixedElemPolys := {};
      for i from 1 to n do fixedElemPolys = append(fixedElemPolys,sub(elempolys_(i-1),S));
      M := mutableMatrix(S,n,n);
      for i from 1 to n do(
            for j from 1 to n do(
                  if (lambda_(i-1)+j-i) == 0 then M_(i-1,j-1) = 1;
                  if ((lambda_(i-1)+j-i) > 0 and (lambda_(i-1)+j-i) <= n) then M_(i-1,j-1) =
                  fixedElemPolys_(lambda_(i-1)+j-i-1)));
      return determinant(matrix M))

polyRepC = method()
polyRepC(List,Ring) := (w,R) -> (
      ringVars := gens R;
      n := length(ringVars);
      pointclass := 1;
```

```
      for i from 1 to (n-1) do(
            pointclass = pointclass*(ringVars_(i-1))^(n-i));
      lambda := {};
      for i from 1 to n do(
    lambda = prepend(i,lambda));
      delta := sub(elementarySchurDeterminantC(lambda,n),R);
      polyrep := pointclass*delta;
      for i in w do(
    if (i != n-1) then(
                polyrep = deltaSwapA(polyrep,R,i));
            if (i == n-1) then(
  polyrep = deltaSwapC(polyrep,R)));
      return(polyrep))

elementarySymmetricSquaresIdeal = method()
elementarySymmetricSquaresIdeal(ZZ) := (n) -> (
      R := QQ[y_(1)..y_(n)][t];
      f := 1_R;
      for i from 1 to n do(
            f = f*(y_(i)^2+t));
      coeffs := (coefficients (f-t^n))_(1);
      S := QQ[y_(1)..y_(n)];
      I := sub(ideal(coeffs),S);
      return(S,I,S/I))

intC = method()
intC(List,Ring,Ideal) := (alphas,S,I) -> (
      f := 1_S;
      for alpha in alphas do(
            f = f*polyRepC(signedBubbleSort(signedToNot(alpha)),S));
      f = f % I;
      return (((coefficients f)_(1))_(0))_(0))

partialIntC = method()
partialIntC(List,List,Ring,Ideal) := (flagType,alphas,S,I) -> (
      l := length(flagType);
      n := flagType_(-1);
      newAlphas := {};
      for alpha in alphas do(
            newAlpha := completeSignedPermutation(alpha,n);
            newAlphas = append(newAlphas,newAlpha));
```

```
      dualClass := {};
      for k from 2 to (l-1) do(
            for j from (flagType_(-(k+1)) + 1) to flagType_(-k) do(
          dualClass = prepend(j,dualClass)));
      for i from 1 to flagType_(0) do(
            dualClass = prepend(i,dualClass));
      for i from 1 to (n-flagType_(-2)) do(
            dualClass = append(dualClass,-i));
      newAlphas = append(newAlphas,dualClass);
      return(intC(newAlphas,S,I)))

-------------------------------------------------

-- TYPE B CODE

-------------------------------------------------

typeBLength = method()
typeBLength(List) := (w) -> (
      n := length(w);
      count := 0;
      for i from 1 to n do
            for j from i+1 to n do
                  if w_(i-1) > w_(j-1) then count = count+1;
      for i from 1 to n do
            for j from i to n do
                  if w_(i-1) + w_(j-1) > 2*n+2 then count = count+1;
      return(count))

deltaSwapB = method()
deltaSwapB(Thing,Ring) := (f,R) -> (
      ringVars := gens R;
      return sub((f-sub(f,{ringVars_(-1)=>(-1)*ringVars_(-1)}))/(ringVars_(-1)),R))

elementarySchurDeterminantB = method()
elementarySchurDeterminantB(List,ZZ) := (lambda,n) -> (
      Rt := QQ[y_(1)..y_(n)][t];
      f := 1_(Rt);
      for i from 1 to n do(
            f = f*(y_(i)+t));
      elempolys := ((coefficients (f-t^n))_(1))_(0);
```

```
      S := QQ[y_(1)..y_(n)];
      fixedElemPolys := {};
      for i from 1 to n do fixedElemPolys = append(fixedElemPolys,(1/2)*(sub(elempolys_(i-1),S)));
      M := mutableMatrix(S,n,n);
      for i from 1 to n do(
            for j from 1 to n do(
                  if (lambda_(i-1)+j-i) == 0 then M_(i-1,j-1) = 1;
                  if ((lambda_(i-1)+j-i) > 0 and (lambda_(i-1)+j-i) <= n) then M_(i-1,j-1) =
                  fixedElemPolys_(lambda_(i-1)+j-i-1)));
      return determinant(matrix M))

polyRepB = method()
polyRepB(List,Ring) := (w,R) -> (
      ringVars := gens R;
      n := length(ringVars);
      pointclass := 1;
      for i from 1 to (n-1) do(
            pointclass = pointclass*(ringVars_(i-1))^(n-i));
      lambda := {};
      for i from 1 to n do(
    lambda = prepend(i,lambda));
      delta := sub(elementarySchurDeterminantB(lambda,n),R);
      polyrep := pointclass*delta;
      for i in w do(
    if (i != n-1) then(
                  polyrep = deltaSwapA(polyrep,R,i));
            if (i == n-1) then(
  polyrep = deltaSwapB(polyrep,R)));
      return(polyrep))

intB = method()
intB(List,Ring,Ideal) := (alphas,S,I) -> (
      f := 1_S;
      for alpha in alphas do(
            f = f*polyRepB(signedBubbleSort(signedToNot(alpha)),S));
      f = f % I;
      return (((coefficients f)_(1))_(0))_(0))

partialIntB = method()
partialIntB(List,List,Ring,Ideal) := (flagType,alphas,S,I) -> (
      l := length(flagType);
```

```
      n := flagType_(-1);
      newAlphas := {};
      for alpha in alphas do(
            newAlpha := completeSignedPermutation(alpha,n);
            newAlphas = append(newAlphas,newAlpha));
      dualClass := {};
      for k from 2 to (l-1) do(
            for j from (flagType_(-(k+1)) + 1) to flagType_(-k) do(
          dualClass = prepend(j,dualClass)));
      for i from 1 to flagType_(0) do(
            dualClass = prepend(i,dualClass));
      for i from 1 to (n-flagType_(-2)) do(
            dualClass = append(dualClass,-i));
      newAlphas = append(newAlphas,dualClass);
      return(intB(newAlphas,S,I)))

-------------------------------------------------

-- TYPE D CODE

-------------------------------------------------

typeDLength = method()
typeDLength(List) := (w) -> (
      n := length(w);
      count := 0;
      for i from 1 to n do
            for j from i+1 to n do
                  if w_(i-1) > w_(j-1) then count = count+1;
      for i from 1 to n do
            for j from i+1 to n do
                  if w_(i-1) + w_(j-1) > 2*n+1 then count = count+1;
      return(count))

signedBubbleSortD = method()
signedBubbleSortD(List) := (L) -> (
      n := length(L);
      eventual := {};
      for i from 1 to (n-1) do(
            eventual = append(eventual,2*n+1-i));
      if (n % 2 == 0) then eventual = append(eventual,n+1);
```

```
      if (n % 2 != 0) then eventual = append(eventual,n);
      swaps := {};
      while (L != eventual) do(
            if L == reverse(sort(L)) then(
          smallest := L_(-1);
  nextSmallest := L_(-2);
          L = drop(L,-1);
  L = drop(L,-1);
  L = append(L,2*n+1-smallest);
  L = append(L,2*n+1-nextSmallest);
  swaps = prepend(n-1,swaps));
            for i from 0 to (n-2) do(
                  if L_(i) < L_(i+1) then(
                        swaps = prepend(i,swaps);
                        L = switch(i,i+1,L);
                        break)));
      return(swaps))

deltaSwapD = method()
deltaSwapD(Thing,Ring) := (f,R) -> (
      ringVars := gens R;
      return sub((f-sub(f,{ringVars_(-2)=>(-1)*ringVars_(-1),ringVars_(-1)=>(-1)*ringVars_(-2)}))/
      (ringVars_(-2)+ringVars_(-1)),R))

elementarySchurDeterminantD = method()
elementarySchurDeterminantD(List,ZZ) := (lambda,n) -> (
      Rt := QQ[y_(1)..y_(n)][t];
      f := 1_(Rt);
      for i from 1 to n do(
            f = f*(y_(i)+t));
      elempolys := ((coefficients (f-t^n))_(1))_(0);
      S := QQ[y_(1)..y_(n)];
      fixedElemPolys := {};
      for i from 1 to n do fixedElemPolys = append(fixedElemPolys,(1/2)*(sub(elempolys_(i-1),S)));
      M := mutableMatrix(S,n-1,n-1);
      for i from 1 to (n-1) do(
            for j from 1 to (n-1) do(
                  if (lambda_(i-1)+j-i) == 0 then M_(i-1,j-1) = 1;
                  if ((lambda_(i-1)+j-i) > 0 and (lambda_(i-1)+j-i) <= n) then M_(i-1,j-1) =
                  fixedElemPolys_(lambda_(i-1)+j-i-1)));
      return determinant(matrix M))
```

```
polyRepD = method()
polyRepD(List,Ring) := (w,R) -> (
      ringVars := gens R;
      n := length(ringVars);
      pointclass := 1;
      for i from 1 to (n-1) do(
            pointclass = pointclass*(ringVars_(i-1))^(n-i));
      lambda := {};
      for i from 1 to (n-1) do(
    lambda = prepend(i,lambda));
      delta := sub(elementarySchurDeterminantD(lambda,n),R);
      polyrep := pointclass*delta;
      for i in w do(
    if (i != n-1) then(
                  polyrep = deltaSwapA(polyrep,R,i));
            if (i == n-1) then(
  polyrep = deltaSwapD(polyrep,R)));
      return(polyrep))

elementarySymmetricDIdeal = method()
elementarySymmetricDIdeal(ZZ) := (n) -> (
      Rt := QQ[y_(1)..y_(n)][t];
      f := 1_(Rt);
      squareProd := 1_(Rt);
      prod := 1_(Rt);
      for i from 1 to n do(
            f = f*(y_(i)^2+t);
    squareProd = squareProd*((y_(i))^2);
    prod = prod*(y_(i)));
      coeffs := (coefficients (f-t^n-squareProd))_(1);
      S := QQ[y_(1)..y_(n)];
      I := sub(ideal(coeffs,prod),S);
      return(S,I,S/I))

intD = method()
intD(List,Ring,Ideal) := (alphas,S,I) -> (
      f := 1_S;
      for alpha in alphas do(
            f = f*polyRepD(signedBubbleSortD(signedToNot(alpha)),S));
      f = f % I;
```

```
      return (((coefficients f)_(1))_(0))_(0))


partialIntD = method()
partialIntD(List,List,Ring,Ideal) := (flagType,alphas,S,I) -> (
      l := length(flagType);
      n := flagType_(-1);
      newAlphas := {};
      for alpha in alphas do(
            newAlpha := completeSignedPermutation(alpha,n);
            newAlphas = append(newAlphas,newAlpha));
      dualClass := {};
      for k from 2 to (l-1) do(
            for j from (flagType_(-(k+1)) + 1) to flagType_(-k) do(
          dualClass = prepend(j,dualClass)));
      for i from 1 to flagType_(0) do(
            dualClass = prepend(i,dualClass));
      for i from 1 to (n-flagType_(-2)-1) do(
            dualClass = append(dualClass,-i));
      if (n-flagType_(-2) % 2 == 0) then dualClass = append(dualClass,n-flagType_(-2));
      if (n-flagType_(-2) % 2 != 0) then dualClass = append(dualClass,flagType_(-2)-n);
      newAlphas = append(newAlphas,dualClass);
      return(intC(newAlphas,S,I)))
```

3.2 An Aside: Partial Flag Varieties Abstractly as Lie Groups

Before continuing our discussion of the functionality of our Macualay2 package, SchubertIdeals.m2, and its usefulness in solving Schubert problems, we take an aside to give another coordinate-free definition of the complete flag varieties and their Schubert varieties in types A, B, C, and D. To do this, we need the notion of a Lie group.

Definition 3.2.1. *Let N be an n-dimensional smooth manifold, and let M be an N-dimensional smooth manifold. A map $F : N \to M$ is **smooth** if for every $p \in N$, there is a local coordinate chart $\varphi : U \to \mathbb{R}^n$ on N with $p \in U$ and a local coordinate chart $\psi : V \to \mathbb{R}^m$ on M with $F(p) \in V$ such that $F(U) \subseteq V$, and the map $\psi \circ F \circ \varphi : \varphi(U) \subseteq \mathbb{R}^\kappa \to \psi(V) \subseteq \mathbb{R}^m$ is smooth (the partial derivatives of all orders of each of its coordinate functions exist). A **Lie group** G is a set that is simultaneously a smooth manifold and group, such that the binary operation*

$* : G \times G \to G$ and inversion map (taking an element $g \in G$ to g^{-1}) $\mu : G \to G$ are smooth maps. Note here the product of two manifolds $M \times N$ is a manifold with local coordinate charts being products of the local coordinate charts on M with those of N.

To each of the four types of root systems considered in the previous section, we assign a classical Lie group:

- $A_{n-1} : \mathrm{SL}(n, \mathbb{C})$
- $B_n : \mathrm{SO}(2n+1, \mathbb{C})$
- $C_n : \mathrm{Sp}(2n, \mathbb{C})$
- $D_n : \mathrm{SO}(2n, \mathbb{C})$

Here, $\mathrm{SL}(n, \mathbb{C})$ is the **special linear group**, consisting of invertible matrices with determinant 1, and the other groups are the **special orthogonal group** $\mathrm{SO}(k, \mathbb{C})$ ($k = 2n$ or $k = 2n+1$) and the **symplectic group** $\mathrm{Sp}(k, \mathbb{C})$ ($k = 2n$), which are subgroups of $\mathrm{SL}(k, \mathbb{C})$ that preserve some non-degenerate bilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^k$. For $\mathrm{SO}(k, \mathbb{C})$, the form is symmetric, and for $\mathrm{Sp}(k, \mathbb{C})$, the form is skew-symmetric. In other words, if we let R_k be the reverse $k \times k$ identity matrix with 1's along the anti-diagonal and 0's elsewhere, and if we let $J = \begin{bmatrix} 0 & R_n \\ -R_n & 0 \end{bmatrix}$, then $\mathrm{SO}(k, \mathbb{C})$ is the subset of matrices P in $\mathrm{SL}(k, \mathbb{C})$ with $P^T R_k P = R_k$, and $\mathrm{Sp}(k, \mathbb{C})$ is the subset of matrices P in $\mathrm{SL}(k, \mathbb{C})$ with $P^T J P = J$.

Let R, W, and G be the fixed root system, Weyl group, and Lie group associated to one of the four classical types $S \in \{A_{n-1}, B_n, C_n, D_n\}$ listed above, everything indexed by n. Further, let $B \subseteq G$ in each case be the **Borel subgroup** consisting of upper triangular matrices in G. Then abstractly, the **complete flag variety** is the quotient $\mathrm{Fl}_S(n) = G/B$, which is a smooth complex projective variety. The classical Lie groups have a **Bruhat decomposition** into double cosets, $G = \sqcup_{\omega \in W} B\omega B$, from which we can define the **Schubert cells in** G/B to be $\Omega_\omega^\circ = \frac{B\omega B}{B}$, and the corresponding **Schubert subvarieties** to be the closures of Schubert cells $\Omega_\omega = \overline{\Omega_\omega^\circ}$ (in the

Zariski topology). Here, there the complete flag $F_\bullet$ is omitted from the notation since the flag is considered to be the span of the columns of the identity matrix - the **identity flag**. Additionally, we have that $\Omega_\omega = \bigcup_{\nu \leq \omega} \Omega_\nu^\circ$, where $\nu \leq \omega$ is in the **strong Bruhat order**, meaning that if $a_1, \ldots, a_l$ is a reduced word for ω, then there exists a subsequence $b_1, \ldots, b_q$ such that $\sigma_{b_1} \cdots \sigma_{b_q} = \nu$.

As for the partial flags $\mathrm{Fl}_S(a_1, \ldots, a_s; n)$, there is a related abstract construction. Using the same notation as before, an intermediate group $B \subseteq P_J \subseteq G$ is called a **parabolic subgroup**. Furthermore, in the theory of Lie (or algebraic) groups, G has a **maximal torus** $T \subseteq B$, where a **torus** is an abelian, connected, and compact subgroup of G, so T is maximal (under inclusion) with respect to these properties. We then define $W_J = \{\omega \in W \mid \omega T \subseteq P_J\}$ and $W^J = \{\text{maximal (Coxeter) length coset representatives } u \text{ of cosets in } W/W_J\}$ (which we've been using partial permutations to represent), which allow us to obtain another Bruhat decomposition of G into double cosets: $G = \sqcup_{u \in W^J} BuP_J$. As a result we get the abstract definition of partial flag varieties $\mathrm{Fl}_S(a_1, \ldots, a_s; n) = G/P_J$ (J related to the shape $(a_1, \ldots, a_s; n)$), as well as Schubert cells $\Omega_u^\circ = \frac{BuP_J}{P_J}$ and Schubert varieties $\Omega_u = \overline{\Omega_u}$.

If X is any partial flag variety of any type, then its cohomology ring $H^*(X)$ is the same as its Chow ring. Therefore, each subvariety V of X determines an element $[V] \in H^*(X)$. In particular, each Schubert subvariety corresponds to a cohomology class $[\overline{\Omega_\omega^\circ}]$. The cup product in $H^*(X)$ corresponds to the intersection of subvarieties of X (as long as they are in general position, so the intersection is transverse). The decomposition of X into Schubert cells Ω_ω°, which are of even real dimension and whose boundaries are unions of smaller Schubert cells, implies that the cohomology ring $H^*(X, \mathbb{Z})$ is concentrated in even dimensions (hence commutative), and induces a corresponding $\mathbb{Z}$-basis for $H^*(X, \mathbb{Z})$ of Schubert classes C_ω.

While these abstract definitions are good to keep in mind, and we presented them here for completeness, we will be more interested in computation and therefore how to represent such objects in local coordinates.

3.3 Partial Flag Varieties, Flags, and Schubert Varieties in Coordinates

We now describe how to represent the various parts of a Schubert problem in coordinates, and share our code from SchubertIdeals.m2 with slight explanations of the methods, as well as examples. We only give coordinates for Type A and Type C, but this can be extended to Types B and D as well. Recall that for a Grassmannian $\mathrm{Gr}(k,n)$, we represent $\mathrm{Gr}(k,n)$ in coordinates by the Schubert variety $\Omega_{\{1,2,\dots,k\}}O_\bullet$, where $O_\bullet$ is the opposite flag with subspaces the spans of the columns of J_n (notation introduced last section), and where $\alpha = \{1,\dots,k\} \in \binom{[n]}{k}$ is the identity partial permutation (still the identity permutation when completed). In this representation, $\mathrm{Gr}(k,n)$ is represented by its affine chart $U_{1,\dots,k}$, which is the $n \times k$ matrix $\begin{bmatrix} I_k \\ (x_{ij}) \end{bmatrix}$, with x_{ij} being the $(n-k)\times k$ matrix of $k(n-k)$ local coordinates for $\mathrm{Gr}(k,n)$. Similarly, the partial flag variety $\mathrm{Fl}(a_1,\dots,a_s;n)$ is also the Schubert variety indexed by the identity partial permutation $[1,\dots,a_s]$, $\Omega_{a_1,\dots,a_s}O_\bullet$, which is represented in coordinates by an $n \times a_s$ matrix with identity matrices $I_{a_i - a_{i-1}}$ ($i = 1,\dots,a_s$, with $a_0 = 0$) in rows $1,\dots,a_1$, $a_1+1,\dots,a_2$, …, $a_{s-1}+1,\dots,a_s$, respectively. In particular, the Schubert variety representing all of the complete flag variety $\mathrm{Fl}(n)$ is given in local coordinates (with respect to the opposite flag) as a lower triangular matrix (with last column omitted), with 1's along the diagonal. Importantly, we use the codimension convention when using partial permutations to index Schubert varieties. Some other sources prefer to use the dimension convention, and so we also have functionality that can convert between the two. Here are some examples, using our package's Macaulay2 method typeAStiefelCoords:

```
    -- Converts a partial permutation in dimension notation to the corresponding
    one in codimension notation. (Equivalently, gives the dual class).
dimToCodim = method()
dimToCodim(List,List) := (flagshape,alpha) -> (
      s := length(flagshape) - 1;
      n := flagshape_(-1);
      breaks := prepend(0,flagshape);
      alphadual := {};
      for b from 1 to s do(
            k := breaks_(b) - breaks_(b-1);
            for i from 1 to k do(
```

```
                  alphadual = append(alphadual,n+1-alpha_(breaks_(b-1)+k-i))));
      return(alphadual))

-- EXAMPLES of dimToCodim:

i168 : dimToCodim({2,4},{3,4})

o168 = {1, 2}

o168 : List

i169 : dimToCodim({2,4},{1,4})

o169 = {1, 4}

o169 : List

i170 : dimToCodim({3,17,21},{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17})

o170 = {19, 20, 21, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18}

o170 : List

-- A helper function for typeAStiefelCoords
splitPermutation = method()
splitPermutation(List,List) := (flagshape,alpha) -> (
      gaps := {flagshape_(0)};
      for i from 1 to (length(flagshape)-2) do(
            gaps = append(gaps,flagshape_(i)-flagshape_(i-1)));
      splitperm := {};
      copyalpha := alpha;
      for gap in gaps do(
            subalpha := {};
            for i from 0 to (gap-1) do(
                  subalpha = append(subalpha,copyalpha_(0));
                  copyalpha = delete(copyalpha_(0),copyalpha));
            splitperm = append(splitperm,subalpha));
      return(splitperm))

-- EXAMPLE of splitPermutation:
```

```
i159 : splitPermutation({2,3,6,9},{7,8,4,1,2,3})

o159 = {{7, 8}, {4}, {1, 2, 3}}

o159 : List

-- Gives the Stiefel Coordinates for a Type A Schubert Variety
typeAStiefelCoords = method()
typeAStiefelCoords(List,List,Ring) := (flagshape,alpha,K) -> (
      n := flagshape_(-1);
      as := flagshape_(-2);
      S := K[x_(1,1)..x_(n,as)];
      alphalist := splitPermutation(flagshape,alpha);
      firstalpha := alphalist_(0);
      l := length(firstalpha);
-- Define matrix of correct size (and over the correct ring) that we can manipulate for the first
subalpha
      M := mutableMatrix(S,n,l);
-- Set leading ones in lxl identity submatrix with rows indexed by alpha
      for i from 1 to l do M_(firstalpha_(i-1)-1,i-1) = 1;
-- Set variables below the leading 1's
      for j from 1 to l do
      for i from firstalpha_(j-1)+1 to n do M_(i-1,j-1) = x_(i,j);
-- Set to 0 all entries above and to the left of leading 1's
      for i from 1 to l do
      for j from 1 to i-1 do M_(firstalpha_(i-1)-1,j-1) = 0;
-- Make matrix non-mutable
      M = matrix M;
-- Remove firstalpha from alphalist
      alphalist = delete(firstalpha,alphalist);
-- Now repeat and concatenate the matrices
      indexshift := length(firstalpha);
      for subalpha in alphalist do(
            l = length(subalpha);
            N := mutableMatrix(S,n,l);
            for i from 1 to l do N_(subalpha_(i-1)-1,i-1) = 1;
            for j from 1 to l do
                  for i from subalpha_(j-1)+1 to n do N_(i-1,j-1) = x_(i,j+indexshift);
            for i from 1 to l do
                  for j from 1 to i-1 do N_(subalpha_(i-1)-1,j-1) = 0;
            N = matrix N;
```

```
            M = M | N;
            indexshift = indexshift + l);
      M = mutableMatrix M;
      for i from 1 to as do(
            for j from i to as-1 do(
                  M_(alpha_(i-1)-1,j) = 0));
      M = matrix M;
-- Create a new ring with variables only those that show up in the matrix M
      R := K[support M];
-- Make it so that M is a matrix over the new ring
      M = sub(M,R);
-- Return Stiefel coordinates and new ring
      return({M, R}))
```

EXAMPLES of typeAStiefelCoords:

```
i155 : typeAStiefelCoords({1,2,3,4,5},{4,5,1,2},QQ)

o155 = {| 0       0 1       0       |, QQ[x   , x   ..x   , x   ]}
        | 0       0 x_(2,3) 1       |      2,3   3,3   3,4   5,1
        | 0       0 x_(3,3) x_(3,4) |
        | 1       0 0       0       |
        | x_(5,1) 1 0       0       |

o155 : List

i156 : typeAStiefelCoords({2,4},{1,2},QQ)

o156 = {| 1       0       |, QQ[x   ..x   ]}
        | 0       1       |      3,1   4,2
        | x_(3,1) x_(3,2) |
        | x_(4,1) x_(4,2) |

o156 : List

i157 : typeAStiefelCoords({2,4,6},{1,2,3,4},QQ)

o157 = {| 1       0       0       0       |, QQ[x   ..x   , x   ..x   ]}
        | 0       1       0       0       |      3,1   4,2   5,1   6,4
        | x_(3,1) x_(3,2) 1       0       |
        | x_(4,1) x_(4,2) 0       1       |
```

```
        | x_(5,1) x_(5,2) x_(5,3) x_(5,4) |
        | x_(6,1) x_(6,2) x_(6,3) x_(6,4) |

o157 : List

i158 : typeAStiefelCoords({1,2,3,4,5,6},{1,2,3,4,5},QQ)

o158 = {| 1       0       0       0       0       |, (ring omitted to fit on page)
        | x_(2,1) 1       0       0       0       |
        | x_(3,1) x_(3,2) 1       0       0       |
        | x_(4,1) x_(4,2) x_(4,3) 1       0       |
        | x_(5,1) x_(5,2) x_(5,3) x_(5,4) 1       |
        | x_(6,1) x_(6,2) x_(6,3) x_(6,4) x_(6,5) |

o158 : List

i159 : splitPermutation({2,3,6,9},{7,8,4,1,2,3})

o159 = {{7, 8}, {4}, {1, 2, 3}}

o159 : List

i160 : typeAStiefelCoords({1,4},{1},QQ)

o160 = {| 1       |, QQ[x   , x   , x   ]}
        | x_(2,1) |      2,1   3,1   4,1
        | x_(3,1) |
        | x_(4,1) |

o160 : List

i161 : typeAStiefelCoords({2,4},{1,2},QQ)

o161 = {| 1       0       |, QQ[x   ..x   ]}
        | 0       1       |      3,1   4,2
        | x_(3,1) x_(3,2) |
        | x_(4,1) x_(4,2) |

o161 : List

i162 : typeAStiefelCoords({1,2,4},{1,2},QQ)
```

```
o162 = {| 1       0       |, QQ[x   , x   ..x   ]}
        | x_(2,1) 1       |      2,1   3,1   4,2
        | x_(3,1) x_(3,2) |
        | x_(4,1) x_(4,2) |

o162 : List

i163 : typeAStiefelCoords({3,4},{1,2,3},QQ)

o163 = {| 1       0       0       |, QQ[x   ..x   ]}
        | 0       1       0       |      4,1   4,3
        | 0       0       1       |
        | x_(4,1) x_(4,2) x_(4,3) |

o163 : List

i164 : typeAStiefelCoords({1,3,4},{1,2,3},QQ)

o164 = {| 1       0       0       |, QQ[x   , x   , x   ..x   ]}
        | x_(2,1) 1       0       |      2,1   3,1   4,1   4,3
        | x_(3,1) 0       1       |
        | x_(4,1) x_(4,2) x_(4,3) |

o164 : List

i165 : typeAStiefelCoords({2,3,4},{1,2,3},QQ)

o165 = {| 1       0       0       |, QQ[x   ..x   , x   ..x   ]}
        | 0       1       0       |      3,1   3,2   4,1   4,3
        | x_(3,1) x_(3,2) 1       |
        | x_(4,1) x_(4,2) x_(4,3) |

o165 : List

i166 : typeAStiefelCoords({1,2,3,4},{1,2,3},QQ)

o166 = {| 1       0       0       |, QQ[x   , x   ..x   , x   ..x   ]}
        | x_(2,1) 1       0       |      2,1   3,1   3,2   4,1   4,3
        | x_(3,1) x_(3,2) 1       |
        | x_(4,1) x_(4,2) x_(4,3) |
```

```
o166 : List

i167 : typeAStiefelCoords({4,4},{1,2,3,4},QQ)

o167 = {| 1 0 0 0 |, QQ[]}
        | 0 1 0 0 |
        | 0 0 1 0 |
        | 0 0 0 1 |

o167 : List
```

Similarly, we have functionality for computing the partial flag varieties of Type C, as well as the corresponding Schubert varieties. We give our code here without examples.

```
typeCStiefelCoords = method()
typeCStiefelCoords(List,List,Ring) := (flagshape,alpha,K) -> (
      n := flagshape_(-1);
      as := flagshape_(-2);
      S := K[x_(1,1)..x_(n,as)];
      alphalist := splitPermutation(flagshape,alpha);
      firstalpha := alphalist_(0);
      l := length(firstalpha);
-- Define matrix of correct size (and over the correct ring) that we can manipulate for the first
subalpha
      M := mutableMatrix(S,n,l);
-- Set leading ones in lxl identity submatrix with rows indexed by alpha
      for i from 1 to l do M_(firstalpha_(i-1)-1,i-1) = 1;
-- Set variables below the leading 1's
      for j from 1 to l do
      for i from firstalpha_(j-1)+1 to n do M_(i-1,j-1) = x_(i,j);
-- Set to 0 all entries above and to the left of leading 1's
      for i from 1 to l do
      for j from 1 to i-1 do M_(firstalpha_(i-1)-1,j-1) = 0;
-- Make matrix non-mutable
      M = matrix M;
-- Remove firstalpha from alphalist
      alphalist = delete(firstalpha,alphalist);
-- Now repeat and concatenate the matrices
      indexshift := length(firstalpha);
```

```
      for subalpha in alphalist do(
            l = length(subalpha);
            N := mutableMatrix(S,n,l);
            for i from 1 to l do N_(subalpha_(i-1)-1,i-1) = 1;
            for j from 1 to l do
                  for i from subalpha_(j-1)+1 to n do N_(i-1,j-1) = x_(i,j+indexshift);
            for i from 1 to l do
                  for j from 1 to i-1 do N_(subalpha_(i-1)-1,j-1) = 0;
            N = matrix N;
            M = M | N;
            indexshift = indexshift + l);
      M = mutableMatrix M;
      for i from 1 to as do(
            for j from i to as-1 do(
                  M_(alpha_(i-1)-1,j) = 0));
      M = matrix M;
-- Create a new ring with variables only those that show up in the matrix M
      R := K[support M];
-- Make it so that M is a matrix over the new ring
      M = sub(M,R);
-- Create the symplectic form J
     J := mutableMatrix(R,n,n);
     halfn := sub(n/2,ZZ);
     for i from 1 to halfn do J_(n-i,i-1) = 1;
     for j from halfn+1 to n do J_(n-j,j-1) = -1;
     J = matrix J;
-- Create ideal of symplectic relations
     rels := ideal(0_R);
     for i from 1 to as do
     for j from i to as do rels = rels + (transpose(submatrix(M,{i-1}))*J*submatrix(M,
     {j-1}))_(0,0);
-- Return Stiefel coordinates and new ring, along with the ideal of relations among the variables
and the dimension of that ideal
     {M, R, rels, dim(rels)})
```

3.4 Computing Ideals of Schubert Problems

Once one has coordinates for Schubert varieties, one can solve and study Schubert problems. While in the first section of this chapter, we solved Schubert problems using the Chow ring and Schubert polynomials, this only gave the number of solutions to a Schubert problem. However,

what if one wanted the actual solutions to an instance of the Schubert problem (meaning flags have been specified)? In other words, what if one wanted the ideal (in local coordinates) of the polynomial equations whose solutions are the solutions to the Schubert problem? Having such an ideal is advantageous, as one can use the ideal to recover the number of solutions (computing the degree of the ideal, as long as the ideal is 0-dimensional), and more than that! With the ideal in hand, one can study the arithmetic and reality of solutions, as well as the Galois group corresponding to the Schubert problem (the symmetries of the polynomial system). The SchubertIdeals.m2 package is specifically designed to compute such ideals, which we will now explain. Furthermore, we use these ideals to study Galois groups, which is explained in detail in Chapter 4.

We first illustrate our use of efficient equations for the ideal of a Schubert problem in local coordinates for Derksen's problem (which we will again revisit when considering Galois groups): Fixing four general flags $F_\bullet^1, F_\bullet^2, F_\bullet^3, F_\bullet^4 \in \mathrm{Fl}(n)$, how many $H \in \mathrm{Gr}(4,8)$ satisfy $\dim(H \cap F_4^i) \geq 2$ for all $i = 1, \ldots, 4$. In other words, how many 4-planes intersect four general 4-planes in $\mathbb{C}^8$? Since $H \in \Omega_{1256} F_\bullet^i \iff \dim(H \cap F_4^i) \geq 2$ for all $i = 1, \ldots, 4$, this problem can be restated as finding the number of points in the intersection of Schubert varieties $\Omega_{1256}F_\bullet^1 \cap \Omega_{1256}F_\bullet^2 \cap \Omega_{1256}F_\bullet^3 \cap \Omega_{1256}F_\bullet^4$.

A naive way of finding the system of equations defining the ideal of this intersection is for each $i = 1, \ldots, 4$, to form an augmented 8×8 matrix $\begin{bmatrix} H \mid F_4^i \end{bmatrix}$, where H is an 8×4 rank-4 matrix of indeterminates, and each F_4^i is a fixed 8×4 rank-4 matrix of constants. Then, $\dim(H \cap F_4^i) \geq 2$ is equivalent to the rank of $\begin{bmatrix} H \mid F_4^i \end{bmatrix}$ being less than or equal to 6, which means that the 7×7 non-maximal minors of $\begin{bmatrix} H \mid F_4^i \end{bmatrix}$ (the determinant of the matrix after deleting 1 row and 1 column) all vanish simultaneously. Since there are 64 such minors (8 choices for each deleted row, and independently 8 choices for each deleted column), this gives 32 homogeneous cubic equations and 32 homogeneous quartic equations for the ideal of each $\Omega_{1256}F_\bullet^i$ (so $64 \cdot 4 = 256$ total equations for the ideal of the intersection). However, it turns out that the ideal for each $\Omega_{1256}F_\bullet^i$ is generated by only 16 cubic minors, but it is not clear a priori which 16 suffice. We present a much more efficient formulation of this problem, which involves only 17 quartic equations for each of the

$\Omega_{1256}F_\bullet^i$ (so $17 \cdot 4 = 68$ total for the intersection). This is the minimal number of such equations. Furthermore, in our formulation we use local coordinates rather than homogeneous coordinates for $\mathrm{Gr}(4,8)$ and its Schubert varieties, so since $\dim(\Omega_{1256}F_\bullet^i) = 12$, our efficient formulation uses only 12 indeterminates (the minimal possible number). Here in local coordinates, we have that the matrix representation of $\Omega_{1256}O_\bullet$ is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ x_{31} & x_{32} & 0 & 0 \\ x_{41} & x_{42} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ x_{71} & x_{72} & x_{73} & x_{74} \\ x_{81} & x_{82} & x_{83} & x_{84} \end{bmatrix}.$$

Recall that the Plücker embedding $\mathrm{Gr}(k,n) \to \mathbb{P}^{\binom{n}{k}-1}(F)$ takes each $H \in \mathrm{Gr}(k,n)$ to its vector of $k \times k$ minors $(p_\alpha(H) \mid \alpha \in \binom{[n]}{k})$, called the **Plücker coordinates** of H. Here, $p_\alpha(H)$ is the determinant of the $k \times k$ submatrix with rows indexed by α. In particular, each Schubert variety $\Omega_\alpha O_\bullet$ is cut out from $\mathrm{Gr}(k,n)$ by a subset of Plücker coordinates. Specifically, $H \in \Omega_\alpha O_\bullet \iff p_\beta(H) = 0$ for all $\beta \in \binom{[n]}{k}$ with $\beta \not\geq \alpha$, because given a general matrix $H \in \Omega_\alpha O_\bullet$, the rank of the $k \times k$ submatrix with rows $\beta_1, \ldots, \beta_k$ is k unless $\beta_i > \alpha_i$ for some i. This again uses the Bruhat order on $\binom{[n]}{k}$, where $\beta \geq \alpha \iff \beta_i \geq \alpha_i$ for all $i = 1, \ldots, k$. Returning to Derksen's problem in $\mathrm{Gr}(4,8)$, there are precisely 17 indicies $\beta \in \binom{[8]}{4}$ with $\beta \not\geq 1256$.

Now, $H \in \Omega_\alpha F_\bullet \iff F^{-1}H \in \Omega_\alpha O_\bullet \iff p_\beta(H) = 0$ for all $\beta \not\geq \alpha$. Using the Cauchy-Binet formula, we can write $p_\beta(F^{-1}H) = \sum_{\gamma \in \binom{[n]}{k}} p_{\beta,\gamma}(F^{-1}) p_\gamma(H)$, where $p_{\beta,\gamma}(F^{-1}) = \det((F^{-1})_{\beta_i,\gamma_j})_{i,j=1}^k$ is the (β,γ)-th entry in the matrix $\wedge^k(F^{-1})$. From this discussion, we get the Theorem used for efficient equations for the ideal of a Schubert problem:

Theorem 3.4.1. *Let $\mathcal{Y}$ be Stiefel (local) coordinates for $Y \subseteq \mathrm{Gr}(k,n)$, and compute the Plücker vector $P(\mathcal{Y}) = (p_\beta(\mathcal{Y}) \mid \beta \in \binom{[n]}{k})$ for $\mathcal{Y}$. Compute the rectangular matrix $P(\alpha)(F^{-1}) =$*

$(p_{\beta,\gamma}(F^{-1}) \mid \beta \not\geq \alpha, \gamma \in \binom{[n]}{k})$. *Then, the entries in the matrix-vector product* $P(\alpha)(F^{-1}) \cdot P(\mathcal{Y})$ *cut out* $Y \cap \Omega_\alpha F_\bullet$.

In our implementation, given a Schubert problem with Schubert conditions $\alpha^1, \ldots, \alpha^r \in \binom{[n]}{k}$ and flags $F^2_\bullet, \ldots, F^r_\bullet$ (we will always have $F^1_\bullet = O_\bullet$), we choose $\mathcal{Y}$ to be the local coordinates of $Y = \Omega_{\alpha^1} O_\bullet$, and get the equations of $\Omega_{\alpha^1} O_\bullet \cap \Omega_{\alpha^2} F^2_\bullet \cap \cdots \cap \Omega_{\alpha^r} F^r_\bullet$ by adjoining the sets of equations defining each $Y \cap \Omega_{\alpha^i} F^i_\bullet$ for each $i = 2, \ldots, r$, using the Theorem above, and obtaining the ideal generated by all such equations. To compare with the naive algorithm, the original reported implementation of this method computed the ideal for the 6 solutions to Derksen's problem $\Omega_{1256} F^1_\bullet \cap \Omega_{1256} F^2_\bullet \cap \Omega_{1256} F^3_\bullet \cap \Omega_{1256} F^4_\bullet$ in 20 seconds, compared to 20 minutes the naive way. In contrast, using the typeASchubertIdeal method in our package SchubertIdeals.m2, the computation takes 0.23767 seconds.

As for generalizing the above method for Grassmannians to Schubert problems in partial flag varieties $\mathrm{Fl}(a_1, \ldots, a_s; n) \subseteq \mathrm{Gr}(a_1, n) \times \cdots \times \mathrm{Gr}(a_s, n)$, we consider the projections $\pi : \mathrm{Fl}(a_1, \ldots, a_s; n) \to \mathrm{Gr}(a_i, n)$ for each $i = 1, \ldots, n$, which give Schubert problems in each Grassmannian $\mathrm{Gr}(a_i, n)$ by sorting the partial permutation with possible descents α in positions 1 through a_i to get another $\alpha \in \binom{[n]}{a_i}$. Doing this for each relevant Schubert variety in the Schubert problem, and for each projection to a Grassmannian, we simply adjoin all the equations obtained to form one large ideal giving the equations in the partial flag variety.

We give the code and many examples now, suppressing many of the flags and ideals for larger examples due to their size:

```
-- Returns whether a partial permutation is not greater than or equal to another in the Bruhat
order.
notGreaterThan = method(TypicalValue=>Boolean)
notGreaterThan(List,List) := (beta,alpha) -> (
      notgreaterthan := false;
      for i from 1 to length(beta) do
            if beta_(i-1) < alpha_(i-1) then notgreaterthan = true;
      return(notgreaterthan))

-- EXAMPLES of notGreaterThan:
```

```
i246 : notGreaterThan({1,2},{3,4})

o246 = true

i247 : notGreaterThan({3,4},{1,2})

o247 = false

i248 : notGreaterThan({1,4},{2,3})

o248 = true

i249 : notGreaterThan({1,2},{1,2})

o249 = fals

-- Gives all partial permutations not greater than or equal to a fixed one.
allNotGreaterThan = method()
allNotGreaterThan(List,ZZ) := (alpha, n) -> (
      L := {};
      for beta in subsets(splice {1..n},length(alpha)) do
            if notGreaterThan(beta,alpha) then L = append(L,beta);
      return(L))

-- EXAMPLES of allNoteGreatherThan:

i250 : allNotGreaterThan({1,2},4)

o250 = {}

o250 : List

i251 : allNotGreaterThan({1,3},4)

o251 = {{1, 2}}

o251 : List

i252 : allNotGreaterThan({1,4},4)
```

```
o252 = {{1, 2}, {1, 3}, {2, 3}}

o252 : List

i253 : allNotGreaterThan({2,3},4)

o253 = {{1, 2}, {1, 3}, {1, 4}}

o253 : List

i254 : allNotGreaterThan({2,4},4)

o254 = {{1, 2}, {1, 3}, {2, 3}, {1, 4}}

o254 : List

i255 : allNotGreaterThan({3,4},4)

o255 = {{1, 2}, {1, 3}, {2, 3}, {1, 4}, {2, 4}}

o255 : List

-- Computes the P(alpha)(F^{-1}) matrix that is essential in finding a minimal number of
generators for the ideal of a Schubert problem.
cauchyBinetCoefficients = method()
cauchyBinetCoefficients(List,List,Matrix,Ring) := (grassmannianshape,betas,F,K) -> (
      k := grassmannianshape_(0);
      n := grassmannianshape_(1);
      Finv := inverse F;
      M := mutableMatrix(K,length(betas),binomial(n,k));
      subs := subsets(splice {1..n},k);
      kones := splice{k:1};
      for i from 0 to length(betas)-1 do(
            for j from 0 to binomial(n,k)-1 do(
                  M_(i,j) = det(submatrix(Finv,betas_(i)-kones,subs_(j)-kones))));
      M = matrix M;
      return(M))

-- EXAMPLE of cauchyBinetCoefficients:

i256 : cauchyBinetCoefficients({2,4},allNotGreaterThan({2,4},4),id_(QQ^4),QQ)
```

```
o256 = | 1 0 0 0 0 0 |
       | 0 1 0 0 0 0 |
       | 0 0 1 0 0 0 |
       | 0 0 0 1 0 0 |

               4        6
o256 : Matrix QQ  <--- QQ

-- Computes the ideal for a Type A Schubert problem.
typeASchubertIdeal = method()
----- NOTE: There should be m alphas and m-1 flags (first flag will be assumed to be the
reverse identity and not given as input)
----- NOTE: The flags should be general and the alpha's codimensions should add up to k(n-k) to
give an actual Schubert problem
typeASchubertIdeal(List,List,List,Ring) := (flagshape,alphas,flags,K) -> (
      n := last(flagshape);
      q := length(flags);
      subspaces := delete(n,flagshape);
      bigcoords := (typeAStiefelCoords(flagshape,alphas_(0),K))_(0);
      bigring := (typeAStiefelCoords(flagshape,alphas_(0),K))_(1);
      eqns := ideal(0_bigring);
      for a in subspaces do(
           coords := submatrix(bigcoords,{0..(a-1)});
   PY := exteriorPower(a,coords);
           conds := {sort(take(alphas_(0),a))};
           for i from 1 to q do(
                conds = append(conds,sort(take(alphas_(i),a))));
           for i from 1 to length(conds)-1 do(
                 eqns = eqns +
                 sub(ideal(cauchyBinetCoefficients({a,n},allNotGreaterThan(conds_(i),n),flags_(i-1
                 ),K)*PY),bigring)));
      return(eqns))

-- EXAMPLES of typeASchubertIdeal:

i257 : F1 = random(QQ^4,QQ^4)

o257 = | 3   1/4 5/4 1/5 |
       | 1/3 1/5 2   8   |
       | 3   3/5 1/8 7   |
```

```
       | 8/7 1/4 1/2 2/3 |

                4        4
o257 : Matrix QQ  <--- QQ

i258 : F2 = random(QQ^4,QQ^4)

o258 = | 2/7  8/9 1/2 1/3  |
       | 4/7  2/5 3/5 1/7  |
       | 1/10 7/2 5/4 10/9 |
       | 7/3  1/6 7/9 7/8  |

                4        4
o258 : Matrix QQ  <--- QQ

i259 : F3 = random(QQ^4,QQ^4)

o259 = | 10/7 1   1/3 2   |
       | 3/4  2   2/3 1/5 |
       | 3/2  5/3 3   1/7 |
       | 5/2  2   1/4 1   |

                4        4
o259 : Matrix QQ  <--- QQ

i260 : I = typeASchubertIdeal({2,4},{{3,4},{1,2}},{F1},QQ)

o260 = ideal 0

o260 : Ideal of QQ[]

i261 : (dim I, degree I)

o261 = (0, 1)

o261 : Sequence

i262 : I = typeASchubertIdeal({2,4},{{1,2},{3,4}},{F1},QQ);

o262 : Ideal of QQ[x   ..x   ]
                  3,1   4,2
```

```
i263 : (dim I, degree I)

o263 = (0, 1)

o263 : Sequence

i264 : I = typeASchubertIdeal({2,4},{{1,3},{1,3},{1,3},{1,3}},{F1,F2,F3},QQ);

o264 : Ideal of QQ[x   , x   ..x   ]
                  2,1   4,1   4,2

i265 : (dim I, degree I)

o265 = (0, 2)

o265 : Sequence

i266 : I = typeASchubertIdeal({2,4},{{1,3},{1,3}},{F1},QQ)

                  439740            36960       483840       655200       134400
o266 = ideal (0, - ------x   x    + ------x    - ------x    + ------x    + ------)
                  436537 2,1 4,2   436537 2,1   436537 4,1   436537 4,2   436537

o266 : Ideal of QQ[x   , x   ..x   ]
                  2,1   4,1   4,2

i267 : (dim I, degree I)

o267 = (2, 2)

o267 : Sequence

i268 : I = typeASchubertIdeal({2,4},{{1,4},{2,3}},{F1},QQ);

o268 : Ideal of QQ[x   , x   ]
                  2,1   3,1

i269 : (dim I, degree I)

o269 = (-1, 0)
```

```
o269 : Sequence

i270 :
       F1 = random(QQ^5,QQ^5);

                5        5
o270 : Matrix QQ  <--- QQ

i271 : F2 = random(QQ^5,QQ^5);

                5        5
o271 : Matrix QQ  <--- QQ

i272 : F3 = random(QQ^5,QQ^5);

                5        5
o272 : Matrix QQ  <--- QQ

i273 : F4 = random(QQ^5,QQ^5);

                5        5
o273 : Matrix QQ  <--- QQ

i274 : F5 = random(QQ^5,QQ^5);

                5        5
o274 : Matrix QQ  <--- QQ

i275 : F6 = random(QQ^5,QQ^5);

                5        5
o275 : Matrix QQ  <--- QQ

i276 :
       I = typeASchubertIdeal({1,2,5},{{2,1},{2,1},{1,3},{1,3},{1,3},{1,3},{1,3}},
       {F1,F2,F3,F4,F5,F6},QQ);

o276 : Ideal of QQ[x   ..x   ]
                   3,1   5,2
```

```
i277 : (dim I, degree I)

o277 = (0, 5)

o277 : Sequence

i278 : I = typeASchubertIdeal({1,2,5},{{2,1},{2,1},{2,1},{1,3},{1,3},{1,3},{1,3}},
{F1,F2,F3,F4,F5,F6},QQ);

o278 : Ideal of QQ[x   ..x   ]
                  3,1   5,2

i279 : (dim I, degree I)

o279 = (0, 3)

o279 : Sequence

i280 :
       F1 = random(QQ^6,QQ^6);

               6        6
o280 : Matrix QQ  <--- QQ

i281 : F2 = random(QQ^6,QQ^6);

               6        6
o281 : Matrix QQ  <--- QQ

i282 : F3 = random(QQ^6,QQ^6);

               6        6
o282 : Matrix QQ  <--- QQ

i283 : I = typeASchubertIdeal({2,4,6},{{1,4,2,5},{1,4,2,5},{1,4,2,5},{1,4,2,5}},{F1,F2,F3},QQ);

o283 : Ideal of QQ[x   , x   , x   , x   ..x   , x   ..x   ]
                  2,1   3,1   3,3   5,1   5,2   6,1   6,4

i284 : (dim I, degree I)
```

```
o284 = (0, 6)

o284 : Sequence

i285 :
       F = random(QQ^4,QQ^4)

o285 = | 2/3 1/2 7/5 5/3 |
       | 3/2 7/3 5/2 3/5 |
       | 1   2/3 1/3 5/3 |
       | 1/4 7/6 1/5 2   |

                4        4
o285 : Matrix QQ  <--- QQ

i286 : I = typeASchubertIdeal({1,4},{{1},{4}},{F},QQ);

o286 : Ideal of QQ[x   , x   , x   ]
                  2,1   3,1   4,1

i287 : typeAStiefelCoords({1,4},{1},QQ)

o287 = {| 1       |, QQ[x   , x   , x   ]}
        | x_(2,1) |      2,1   3,1   4,1
        | x_(3,1) |
        | x_(4,1) |

o287 : List

i288 : gens gb I

o288 = | 5x_(4,1)-6 x_(3,1)-1 25x_(2,1)-9 |

                                    1                            3
o288 : Matrix (QQ[x   , x   , x   ])  <--- (QQ[x   , x   , x   ])
                 2,1   3,1   4,1              2,1   3,1   4,1

i289 :
       F = matrix{{1/2,4/5,1/2,7/8},{1/2,3/2,1/3,6/5},{8/9,10/9,2,8/5},{1,1/3,2,7/9}}

o289 = | 1/2 4/5  1/2 7/8 |
```

```
       | 1/2 3/2  1/3 6/5 |
       | 8/9 10/9 2   8/5 |
       | 1   1/3  2   7/9 |

                4        4
o289 : Matrix QQ  <--- QQ

i290 : J = typeASchubertIdeal({2,4},{{1,2},{3,4}},{F},QQ);

o290 : Ideal of QQ[x   ..x   ]
                   3,1   4,2

i291 : typeAStiefelCoords({2,4},{1,2},QQ)

o291 = {| 1       0       |, QQ[x   ..x   ]}
        | 0       1       |      3,1   4,2
        | x_(3,1) x_(3,2) |
        | x_(4,1) x_(4,2) |

o291 : List

i292 : gens gb J

o292 = | 111x_(4,2)+490 333x_(4,1)-2312 37x_(3,2)+114 37x_(3,1)-224 |

                              1                      4
o292 : Matrix (QQ[x   ..x   ])  <--- (QQ[x   ..x   ])
                  3,1   4,2              3,1   4,2

i293 : M = matrix{{1,0},{0,1},{224/37,-114/37},{2312/333,-490/111}}

o293 = | 1        0        |
       | 0        1        |
       | 224/37   -114/37  |
       | 2312/333 -490/111 |

                4        2
o293 : Matrix QQ  <--- QQ

i294 : F12 = submatrix(F,{0,1})
```

```
o294 = | 1/2 4/5  |
       | 1/2 3/2  |
       | 8/9 10/9 |
       | 1   1/3  |

                4        2
o294 : Matrix QQ  <--- QQ

i295 : F13 = submatrix(F,{0,2})

o295 = | 1/2 1/2 |
       | 1/2 1/3 |
       | 8/9 2   |
       | 1   2   |

                4        2
o295 : Matrix QQ  <--- QQ

i296 : F14 = submatrix(F,{0,3})

o296 = | 1/2 7/8 |
       | 1/2 6/5 |
       | 8/9 8/5 |
       | 1   7/9 |

                4        2
o296 : Matrix QQ  <--- QQ

i297 : F23 = submatrix(F,{1,2})

o297 = | 4/5  1/2 |
       | 3/2  1/3 |
       | 10/9 2   |
       | 1/3  2   |

                4        2
o297 : Matrix QQ  <--- QQ

i298 : F24 = submatrix(F,{1,3})

o298 = | 4/5  7/8 |
```

```
       | 3/2  6/5 |
       | 10/9 8/5 |
       | 1/3  7/9 |

                4        2
o298 : Matrix QQ  <--- QQ

i299 : F34 = submatrix(F,{2,3})

o299 = | 1/2 7/8 |
       | 1/3 6/5 |
       | 2   8/5 |
       | 2   7/9 |

                4        2
o299 : Matrix QQ  <--- QQ

i300 : rank(M | F12)

o300 = 4

i301 : rank(M | F13)

o301 = 3

i302 : rank(M | F14)

o302 = 3

i303 : rank(M | F23)

o303 = 3

i304 : rank(M | F24)

o304 = 3

i305 : rank(M | F34)

o305 = 2
```

Once one has the ideal of a Schubert problem, this ideal can be used to study the solutions to the problem further. Again, this will be key for studying the Galois group of a Schubert problem in the next chapter, but in the meanwhile we present an application to studying the reality of the solutions to Derksen's problem. Using the ideal, we can use another package in Macaulay2, RealRoots.m2 to compute an eliminant for the ideal, which is a single-variable polynomial in the ideal of degree 6 (the number of solutions to the problem). The functionality of RealRoots.m2 then allows us to count the number of real solutions out of the 6 solutions computed over the complex numbers. Here is the input and output of the relevant code, showing that out of the 6 solutions, for a particular choice of general flags, there are 2 real solutions (again the ideal and eliminant suppressed do to their size):

```
i3 : loadPackage("SchubertIdeals", Reload => true)

o3 = SchubertIdeals

o3 : Package

i4 : loadPackage("RealRoots", Reload => true)

o4 = RealRoots

o4 : Package

i5 : typeAStiefelCoords({4,8},{1,2,5,6},QQ)

o5 = {| 1       0       0       0       |, QQ[x   ..x   , x   ..x   ]}
      | 0       1       0       0       |      3,1   4,2   7,1   8,4
      | x_(3,1) x_(3,2) 0       0       |
      | x_(4,1) x_(4,2) 0       0       |
      | 0       0       1       0       |
      | 0       0       0       1       |
      | x_(7,1) x_(7,2) x_(7,3) x_(7,4) |
      | x_(8,1) x_(8,2) x_(8,3) x_(8,4) |

o5 : List

i6 : typeALength({1,2,5,6},8)
```

```
o6 = 4

i7 :
     F1 = random(QQ^8,QQ^8)

o7 = | 3/2  2    5/6 6/5 9   4/9 9/10 10/3 |
     | 1/10 3/2  4/5 5   3/4 1/4 1/3  7/4  |
     | 5/6  7/3  2/3 1/2 3/7 2   9/5  1/10 |
     | 5/3  8/7  1/4 2   2   3/2 5/2  1/2  |
     | 8/3  3/8  9/5 6/7 1/2 1   9/4  1/9  |
     | 1    1/3  3/5 1/5 9/4 1/2 6/5  4/3  |
     | 3    2    7/3 4   1   1/2 2/3  4    |
     | 5    9/10 8   4   2   8/5 3/10 5/3  |

               8        8
o7 : Matrix QQ  <--- QQ

i8 : F2 = random(QQ^8,QQ^8)

o8 = | 1/8 5/9 5/7  2/5 9/5  3/2  3/4 1/3 |
     | 1/8 8   1/3  6   5/2  5/3  1/2 1/2 |
     | 7/9 1/5 7/6  1/7 4/3  4/3  9/8 8/9 |
     | 1/9 5   5/6  1/2 1    5/7  3/4 1/3 |
     | 2   1/3 3/10 5   2    7/2  2/7 1/8 |
     | 1/7 7/5 7/5  9/5 5/2  10/3 1/2 3/8 |
     | 3   5/4 7/10 1/3 9/10 5    1/4 9   |
     | 7/4 8/3 1/2  3   3    5    2/5 8/5 |

               8        8
o8 : Matrix QQ  <--- QQ

i9 : F3 = random(QQ^8,QQ^8)

o9 = | 1/8 3    4    1/5  3   9/7 1   1   |
     | 5/3 5/3  9/10 3/2  1   2   2   7/3 |
     | 3/2 8    4/5  8/3  7/4 2   5   5/8 |
     | 4/9 3/7  6/5  1/7  1/4 9/7 3/4 1   |
     | 5/2 10/3 1/7  1    3   1/2 2   8   |
     | 5/3 5/2  1    2    10  4   6   1   |
     | 9/2 5/2  8/9  10/7 3/8 1/3 7/9 1/4 |
```

```
     | 7/9 1/2  8/5  6    3/5 3   4/7 1/8 |

              8        8
o9 : Matrix QQ  <--- QQ

i10 : I = typeASchubertIdeal({4,8},{{1,2,5,6},{1,2,5,6},{1,2,5,6},{1,2,5,6}},{F1,F2,F3},QQ);

o10 : Ideal of QQ[x   ..x   , x   ..x   ]
                  3,1   4,2   7,1   8,4

i11 : dim I

o11 = 0

i12 : degree I

o12 = 6

i13 : f = univariateEliminant(sum(gens ring(I)),I);

i14 : SturmCount(f)

o14 = 2
```

Finally, we mention that we also implemented code that will compute the ideals for Schubert problems of Type C. We give our code here for completeness, but again leave out explanations and examples. We also share our implementation here for creating special types of flags, namely secant flags, osculating flags, and the parametrized Symplectic flags relevant for computing Schubert problems in Type C:

```
secantFlag = method()
secantFlag(List,Ring) := (L,R) -> (
      n := length(L);
      secantflag := mutableMatrix(R,n,n);
      for i from 1 to n do
            for j from 1 to n do secantflag_(i-1,j-1) = L_(j-1)^i;
      secantflag = matrix secantflag;
      return(secantflag))
```

```
randomSecantFlag = method()
randomSecantFlag(ZZ,Ring) := (n,R) -> (
      L := {};
      for i from 1 to n do
            L = append(L,random(R)*random(R));
      return(secantFlag(L,R)))

-- Osculating Flags

osculatingFlag = method()
osculatingFlag(QQ, ZZ) := (t,n) -> (
      F := mutableMatrix(QQ,n,n);
      for i from 0 to n-1 do
            for j from i to n-1 do
                  F_(j,i) = t^(j-i)/((j-i)!);
      F = matrix F;
      return F)

parametrizedSymplecticFlag = method()
parametrizedSymplecticFlag(QQ, ZZ) := (t, n) -> (
      F := mutableMatrix(QQ,n,n);
      for i from 0 to n-1 do F_(i,0) = t^(i)/(i!);
            for j from 1 to n-1 do
                  for k from j to n-1 do
                        F_(k,j) = F_(k-1,j-1);
      for l from sub(n/2,ZZ) to n-1 do if odd l then
            for m from 0 to n-1 do
                  F_(l,m) = -1*F_(l,m);
      F = matrix F;
      return F)

typeCGrassmannianSchubertIdeal = method()
typeCGrassmannianSchubertIdeal(List,List,List,Ring) := (grassmannianshape,alphas,flags,K) -> (
      k := grassmannianshape_(0);
      n := grassmannianshape_(1);
      coords := typeCStiefelCoords(grassmannianshape,alphas_(0),K);
      R := coords_(1);
      I := coords_(2);
      PY := exteriorPower(k,coords_(0));
      for i from 1 to length(alphas)-1 do
            I = I +
```

```
            ideal(cauchyBinetCoefficients(grassmannianshape,allNotGreaterThan(alphas_(i),n),flags_
            (i-1),K)*PY);
      return(I))

typeCSchubertIdeal = method()
typeCSchubertIdeal(List,List,List,Ring) := (flagshape,conditions,flags,K) -> (
      n := last(flagshape);
      s := length(flags);
      subspaces := delete(n,flagshape);
      bigRing := (typeCStiefelCoords(flagshape,conditions#0,K))#1;
      eqns := ideal(0_bigRing);
      for a in subspaces do(
           conds := {take(conditions#0,a)};
           for i from 1 to s do(
                conds = append(conds,sort(take(conditions#i,a))));
           eqns = eqns + sub(typeCGrassmannianSchubertIdeal({a,n},conds,flags,K),bigRing));
      return(eqns))
```

4.1 The Frobenius Map and Cycle Types

Fix a prime p, and let $\overline{\mathbb{F}}_p$ be the algebraic closure of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, the finite field with p elements.

Lemma 4.1.1. *Every finite extension of $\mathbb{F}_p$ of a given degree is unique up to isomorphism.*

Proof. Let K be a finite extension of $\mathbb{F}_p$ of degree n (K is an n-dimensional $\mathbb{F}_p$-vector space). Then as an $\mathbb{F}_p$-vector space, $K \cong (\mathbb{F}_p)^n$, so as a set, K has p^n elements. Since K is a field, its non-zero elements K^* form a multiplicative group of order $p^n - 1$, so satisfy the polynomial equation $x^{p^n-1} = 1$ by Lagrange's Theorem. Hence, all p^n elements of K satisfy $x^{p^n} = x$ (after multiplying the previous equation by x), which has at most p^n roots over $\mathbb{F}_p$ (since $\mathbb{F}_p$ is a field). Therefore, K is isomorphic to the splitting field of the polynomial $x^{p^n} - x$ over $\mathbb{F}_p$, and splitting fields are unique up to isomorphism. □

Since it is unique up to isomorphism, we denote by $\mathbb{F}_{p^n}$ any degree n extension of $\mathbb{F}_p$. As the algebraic closure of a field is the union of all finite extensions, we can thus write $\overline{\mathbb{F}}_p = \bigcup_{n=1}^{\infty} \mathbb{F}_{p^n}$ (after embedding all such extensions into $\overline{\mathbb{F}_p}$.

Now we are ready to introduce the Frobenius map, which is our main topic of study for this section, and the key ingredient to the Frobenius algorithm.

Definition 4.1.2. *The **Frobenius map** is* $\mathrm{Frob} : \overline{\mathbb{F}}_p \to \overline{\mathbb{F}}_p$*, defined by* $\mathrm{Frob}(\alpha) = \alpha^p$. Frob *is also referred to as the **Frobenius endomorphism** or **Frobenius automorphism**, the latter justified below.*

Proposition 4.1.3. Frob *is a field automorphism with $\mathbb{F}_p$ as its fixed field.*

Proof. Recall that a field automorphism is a bijective homomorphism from a field to itself. Frob is a field homomorphism, since $\mathrm{Frob}(1) = 1^p = 1 \neq 0$, and $\mathrm{Frob}(\alpha\beta) = (\alpha\beta)^p = \alpha^p\beta^p = \mathrm{Frob}(\alpha)\,\mathrm{Frob}(\beta)$ and $\mathrm{Frob}(\alpha+\beta) = (\alpha+\beta)^p = \sum_{k=0}^{p} \binom{p}{k}\alpha^{p-k}\beta^k$ by the Binomial Theorem. By

definition, $\binom{p}{k} = \frac{p!}{k!(p-k)!}$, so if $1 \leq k \leq p-1$, $p!$ is divisible by p, but neither $k!$ nor $(p-k)!$ are divisible by p (since their prime factors are all smaller than p, and p is prime). Hence in $\overline{\mathbb{F}}_p$, $\binom{p}{k} = 0$ for $1 \leq k \leq p-1$, and $\binom{p}{0} = \binom{p}{p} = 1$, so $\mathrm{Frob}(\alpha+\beta) = (\alpha+\beta)^p = \alpha^p + \beta^p = \mathrm{Frob}(\alpha) + \mathrm{Frob}(\beta)$.

Note that every homomorphism of fields is injective, since the kernel of a field homomorphism is an ideal, and a field has only itself and $\{0\}$ as ideals. Thus since Frob is a field homomorphism, it is injective.

Let F denote the restriction of Frob to $\mathbb{F}_{p^n}$. Note $\mathbb{F}_{p^n}$ is contained in the image of F, since $\mathbb{F}_{p^n}$ is a field (closed under taking pth powers). Since F is injective and $\mathbb{F}_{p^n}$ is a finite set, F is thus surjective as well. Therefore, $\overline{\mathbb{F}}_p = \bigcup_{n=1}^{\infty} \mathbb{F}_{p^n}$ implies that Frob is surjective.

From the proof of Lemma 4.1.1, $F^n(x) = x^{p^n} = x$ for all $x \in \mathbb{F}_{p^n}$, so by degree consideration, $\mathbb{F}_{p^n}$ is the fixed field of F^n. In particular, $\mathbb{F}_p$ is the fixed field of Frob. □

We now move on to discussing the cycle type of the Frobenius map as a Galois group element, viewed as an element of a permutation group. The following lemmas help build towards this. Recall that a **separable** polynomial is one with no repeated roots.

Lemma 4.1.4. *Let L be any field. A monic polynomial $f \in L[x]$ is separable if and only if the greatest common divisor $(f, f') = 1$, where f' is the formal derivative of f.*

Proof. ($\Rightarrow$) Let $a_1, \ldots, a_n \in \overline{L}$ be the distinct roots of f. If f is separable, then it is of the form $f = (x-a_1)\cdots(x-a_n)$ in $\overline{L}[x]$, with $a_i \neq a_j$ for $i \neq j$. Hence, the formal derivative (using the product rule many times) is

$$f' = (x-a_2)\cdots(x-a_n) + (x-a_1)(x-a_3)\cdots(x-a_n) + \cdots + (x-a_1)\cdots(x-a_{n-1})$$

. Since we're computing the greatest common divisor over $\overline{L}$, $(f, f') = 1$ is equivalent to f and f' not having any common roots. Since the roots $a_1, \ldots a_n$ of f are not roots of f' (evaluating f' at

a_i, only one term doesn't cancel, and is a product of non-zero elements of $\overline{L}$), $(f, f') = 1$.
($\Leftarrow$ via contrapositive) If f is inseparable, then it is of the form $f = (x-a)^2 h(x)$ in $\overline{L}[x]$. Hence, the formal derivative (using the product rule) is $f' = 2(x-a)h(x) + (x-a)^2 h'(x)$. Since $(x-a)$ divides both f and f' (even if the characteristic of L is 2, $h'(x) = 0$, or both), $(f, f') \neq 1$. □

Lemma 4.1.5. *An irreducible polynomial $g \in \mathbb{F}_p[x]$ is separable.*

Proof. Dividing by the leading coefficient, we may assume that g is monic. Then, either $(g, g') = 1$ (by degree considerations and g being irreducible), or $g' = 0$, causing $(g, g') = g$. However, $g' = 0$ implies that every exponent of g is divisible by p (by the power rule when computing g'. Hence, $g = a_r x^{pb_r} + \cdots + a_1 x^{pb_1} + a_0 x^{pb_0} = (a_r x^{b_r})^p + \cdots + (a_1 x^{b_1})^p + (a_0 x^{b_0})^p = (a_r x^{b_r} + \cdots + a_1 x^{b_1} + a_0 x^{b_0})^p$ (using $a_i^p = a_i$ for all i by Fermat's Little Theorem and induction on $a^p + b^p = (a+b)^p$, proved earlier). Hence $g' = 0$ implies g is a pth power and thus not irreducible, a contradiction. Therefore, $(g, g') = 1$, so by the previous Lemma, g is separable. □

Consider an irreducible polynomial $g \in \mathbb{F}_p[x]$ of degree n and its splitting field K. Since g is separable, $K/\mathbb{F}_p$ is a Galois extension. Thus, $G = \mathrm{Gal}(K/\mathbb{F}_p)$ permutes the n roots of g. Choosing an ordering for the roots of g, this gives an isomorphism of G with a subgroup of S_n, and one can consider cycle types of elements of G under this isomorphism. Again denoting by F the restriction of Frob to K, $F \in G$, so one can ask for the cycle type of F. The answer is remarkable, and nowhere near true over $\mathbb{Q}$ (already breaking over $\mathbb{Q}$ for irreducible cubic polynomials).

Proposition 4.1.6. *With the setup as above, G is always cyclic of order n, with F as generator. Hence, F has cycle type n (also stated as F is an n-cycle).*

Proof. First, we show that K^* (same as $K \setminus \{0\}$) is cyclic. We know that K^* is a finite abelian group (of order $p^n - 1$, where $n = |G|$ is the degree of the extension). Hence, by the Fundamental Theorem of Finite Abelian Groups, $K^* \cong \mathbb{Z}/u_1\mathbb{Z} \times \ldots \times \mathbb{Z}/u_k\mathbb{Z}$, where $u_1 | \ldots | u_k$, and $|K| = p^n - 1 = u_1 \ldots u_k$. By this isomorphism, we know that $x^{u_k} = 1$ for all $x \in K^*$, but since the polynomial $x^{u_k} - 1$ has at most u_k many roots and K^* has $p^n - 1$ many elements, this forces

$u_k = p^n - 1$ (since we already know $x^{p^n-1} = 1$ for all $x \in K^*$). Thus, $K^* \cong \mathbb{Z}/(p^n-1)\mathbb{Z}$, and so is cyclic.

Now, since K^* is cyclic, there exists some x that generates K^*. In other words, every element of K^* is uniquely of the form x^l for $1 \leq l \leq p^n - 1$. Let $\sigma \in G$. Then, since σ is by definition an automorphism of K, $\sigma(0) = 0$, $\sigma(x) \in K^* \implies \sigma(x) = x^k$ for some k, and in general, $\sigma(x^l) = (\sigma(x))^l = (x^k)^l = x^{kl} = (x^l)^k$. With $0^k = 0$ as well, we can then say that each $\sigma \in G$ is defined by sending every element y in K to y^k, for some k (i.e. it is a "kth-power map). Note that every element of G must have this property, but not every kth power map will be in G.

Let $k = p^m r$, where $p \nmid r$. Then, $(1+x)^k = (1+x)^{p^m r} = [(1+x)^{p^m}]^r = (1+x^{p^m})^r = \sum_{i=0}^{r} \binom{r}{i} x^{r^m i}$ (by induction on $(1+x)^p = 1 + x^p$ and the Binomial Theorem). Since $p \nmid r$, $\binom{r}{1} = r \neq 0 \in \mathbb{F}_p$, and so if $r \neq 1$, r is the coefficient of a term in $(1+x)^k$ besides 1 and x^k. Thus if $r \neq 1$, the polynomial $(1+x)^k - (1+x^k)$ is not the zero polynomial, and has degree less than $k \leq p^n - 1$, but K has p^n elements, so it cannot be that $(1+x)^k = 1 + x^k$ for all $x \in K$. Hence, if $r \neq 1$, the kth power map $\sigma(y) = y^k$ is not a field homomorphism, and so cannot be in G. Therefore, for kth power maps $\sigma \in G$, it must be that $r = 1$, so k is a power of p. We show in fact that all such maps are in G, and thus are precisely the elements of G.

Now, note that the map $\sigma(x) = x^{p^m} = F^m(x)$, which is a composition of $F \in G$ with itself m times. G is a group, so the only possible elements of G are of the form F^m for $1 \leq m \leq n$. Hence, $|G| \leq n$. Further, $x^{p^m} - x = 0$ for all $x \in K$ is only possible due to degree constraints if $m = n$, so the elements $\{F, F^2, \ldots, F^n\}$ are all distinct, with $F^n = Id_K$. Therefore, $G = \langle F \rangle$, so is cyclic of order n, and viewing F as permuting the roots of $g = f \mod p$, F is an n-cycle. $\square$

The general case (g not necessarily irreducible) follows as a corollary.

Corollary 4.1.7. *Let $g \in \mathbb{F}_p[x]$ with splitting field K, and denote by F and G the Frobenius restriction and Galois group as before. Then, the cycle type of F matches the decomposition type*

of g over $\mathbb{F}_p$ (the list of the degrees of the factors of g, sorted as a partition).

Proof. $\mathbb{F}_p[x]$ is a unique factorization domain, so let $g = g_1 \dots g_k$ be a decomposition of g as the product of irreducible polynomials. Since F permutes the roots of each irreducible factor g_i of g separately, by Proposition 4.1.6, F acts as a $\deg(g_i)$ cycle on the roots of g_i, so the cycle type of F in G is $(\deg(g_1), \dots, \deg(g_k))$, which matches the decomposition type of g by definition. □

4.2 An Aside: Frobenius as a Natural Transformation

Since the Frobenius map can be generalized to be over any ring of characteristic p, its universality can be extended (not necessary to future discussion, but interesting nonetheless).

Proposition 4.2.1. *Let A be a characteristic p ring, and let $\mathrm{Frob}_A : A \to A$ be the generalized Frobenius map sending each $a \in A$ to $a^p \in A$. If A has no nonzero nilpotent elements, then Frob_A is injective. Also Frob_A is not in general surjective, even if A is a field.*

Proof. Let A have no nonzero nilpotent elements, and say $\mathrm{Frob}(a) = 0$. Then $a^p = 0$, so by definition, a is a nilpotent, so it must be that $a = 0$, so Frob is injective. For a counterexample to surjectivity, let $A = \mathbb{F}_p(t)$. Then, the image of Frob_A does not contain t, since if it did, there would be a rational function $\frac{f(t)}{g(t)} \in A$ such that $\frac{f(t)^p}{g(t)^p} = t$. However, the degree of $t = \frac{f(t)^p}{g(t)^p}$ is $p \cdot \deg(f) - p \cdot \deg(g)$, which is a multiple of p, contradicting that $\deg(t) = 1$. □

Since Frob is not necessarily surjective, we avoid the term “Frobenius automorphism" and only say “Frobenius endomorphism" in general. Let K be a field of characteristic p. Then by above, $\mathrm{Frob} : K \to K$ is injective. If Frob is surjective on K (so is an automorphism), we say that the field K is **perfect**. Hence, we have seen that $\overline{\mathbb{F}_p}$ and finite extensions of $\mathbb{F}_p$ are perfect (in fact so are all reduced Artinian algebras over $\mathbb{F}_p$, which are products of finite field extensions), but $\mathbb{F}_p(t)$ is not.

We now recall some category theory:

Definition 4.2.2. *A (small) **category** $\mathcal{C}$ consists of the following data:*

1. *A collection (class) of objects.*

2. *For every two (not necessarily distinct) objects A and B in $\mathcal{C}$, a set of morphisms, denoted by* $\mathrm{Hom}_{\mathcal{C}}(A, B)$.

3. *For every three (not necessarily distinct) objects A, B, and C in $\mathcal{C}$, a function* $\circ : \mathrm{Hom}_{\mathcal{C}}(B, C) \times \mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{C}}(A, C)$*, called composition. Here $\circ(g, f)$ is denoted by $g \circ f$, and the composition function is required to be associative, i.e. for every four (not necessarily distinct) objects A, B, C, and D in $\mathcal{C}$, if $h \in \mathrm{Hom}_{\mathcal{C}}(C, D)$, $g \in \mathrm{Hom}_{\mathcal{C}}(B, C)$, and $f \in \mathrm{Hom}_{\mathcal{C}}(A, B)$, then $h \circ (g \circ f) = (h \circ g) \circ f$ in* $\mathrm{Hom}_{\mathcal{C}}(A, D)$.

4. *For each object A in $\mathcal{C}$, an identity morphism $1_A \in \mathrm{Hom}_{\mathcal{C}}(A, A)$ satisfying for any $f \in \mathrm{Hom}_{\mathcal{C}}(A, B)$ and $\tilde{f} \in \mathrm{Hom}_{\mathcal{C}}(B, A)$, $f \circ 1_A = f$ and $1_A \circ \tilde{f} = \tilde{f}$.*

For example, some categories first encountered are the category of sets and set functions, the category of vector spaces and linear transformations, and the categories of the real numbers $\mathbb{R}$ (so only one object) with different sets of morphisms, such as continuous, differentiable, or Riemann-integrable functions. For our purposes, we will consider the category of all characteristic p rings with morphisms being ring homomorphisms between such rings, and denote this category as Ring_p.

Definition 4.2.3. *A (covariant) functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ between categories $\mathcal{C}$ and $\mathcal{D}$ assigns to*

1. *Every object A in $\mathcal{C}$, an object $\mathcal{F}(A)$ in $\mathcal{D}$, and to*

2. *Every morphism $f \in \mathrm{Hom}_{\mathcal{C}}(A, B)$, a morphism $\mathcal{F}(f) \in \mathrm{Hom}_{\mathcal{D}}(\mathcal{F}(A), \mathcal{F}(B))$, such that*

3. $\mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f)$ *($\mathcal{F}$ preserves composition), and*

4. $\mathcal{F}(1_A) = 1_{\mathcal{F}(A)}$ *($\mathcal{F}$ preserves identity).*

A first example of a functor is the “forgetful functor" that sends a vector space to its underlying set (forgetting the addition and scalar multiplication structures), and a linear transformation to its underlying set function (forgetting that the map is linear). For our purposes, we consider an even more basic functor, the identity functor $\mathrm{Id} : \mathrm{Ring}_p \to \mathrm{Ring}_p$ that just sends characteristic p rings and homomorphisms to themselves.

Definition 4.2.4. *If $\mathcal{F}$ and $\mathcal{G}$ are functors from a category $\mathcal{C}$ to a category $\mathcal{D}$, a **natural transformation** $\varphi : \mathcal{F} \to \mathcal{G}$ is a collection of morphisms in $\mathcal{D}$ $\{\varphi_A : \mathcal{F}(A) \to \mathcal{G}(A)\}$, one for each object A in $\mathcal{C}$, such that for each morphism $f : A \to B$ in $\mathcal{C}$ the diagram commutes:*

$$\begin{array}{ccc} \mathcal{F}(A) & \xrightarrow{\varphi_A} & \mathcal{G}(A) \\ {\scriptstyle \mathcal{F}(f)}\downarrow & & \downarrow{\scriptstyle \mathcal{G}(f)} \\ \mathcal{F}(B) & \xrightarrow{\varphi_B} & \mathcal{G}(B) \end{array}$$

An important first example of a natural transformation is the determinant of an $n \times n$ matrix for some fixed n. Here $\mathcal{C}$ is the category of commutative rings and ring homomorphisms, $\mathcal{D}$ is the category of groups and group homomorphisms, $\mathcal{F}$ is the functor that takes a ring R to the general linear group $\mathrm{GL}_n(R)$ of invertible $n \times n$ matrices over R, and $\mathcal{G}$ is the functor that takes a ring R to its group of units (invertible elements) R^*. Then, for any ring R, the determinant is a group homomorphism $\det : \mathrm{GL}_n(R) \to R^*$ that takes a matrix to its determinant. It is a homomorphism since for matrices $A, B \in \mathrm{GL}_n(R)$, $\det(AB) = \det(A)\det(B)$, and is a natural transformation since the determinant of a matrix is a polynomial in the entries of the matrix. Hence, given a ring homomorphism $f : R \to S$, it does not matter if one applies f to all the entries of a matrix A, and then one takes the determinant of that matrix, or if one takes the determinant of the matrix first, and then applies f to the result (since f is a ring homomorphism). The determinant being a natural transformation encapsulates that it is a map that is universally defined, independent of the ring or specific matrix one is working with.

For our purposes, $\mathcal{F} = \mathcal{G} = \mathrm{Id}$, and $\mathrm{Frob} : \mathrm{Ring}_p \to \mathrm{Ring}_p$ is the collection of Frobenius maps $\{\mathrm{Frob}_A : A \to A\}$ running over each characteristic p ring A.

Proposition 4.2.5. *The Frobenius map is a natural transformation from the identity functor on the category of characteristic p rings to itself.*

Proof. All we must show is that for every two characteristic p rings A and B, and ring homomor-

phism $f : A \to B$ between them, that the diagram below commutes:

$$\begin{array}{ccc} A & \xrightarrow{\text{Frob}_A} & A \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ B & \xrightarrow{\text{Frob}_B} & B \end{array}$$

In other words, we must show that $\text{Frob}_B \circ f = f \circ \text{Frob}_A$, but this is just that for every $a \in A$, $f(a)^p = f(a^p)$, which is always true since f is a ring homomorphism. $\square$

Basically, what we've shown is that Frobenius is a "natural" map that arises whenever you have a characteristic p ring, defined "universally" and independently of the particular ring. Furthermore, it respects the morphisms in that category. The Frobenius automorphism is even more general than this, and can be extended to schemes over rings of characteristic p.

4.3 Lifting Frobenius to Characteristic 0

Now, we consider a separable polynomial $f \in \mathbb{Z}[x] \subset \mathbb{Q}[x]$ and its splitting field E over $\mathbb{Q}$. We want to understand the Galois group $M = \text{Gal}(E/\mathbb{Q})$, but first take a detour by reducing everything modulo a prime. For any prime $p \in \mathbb{Z}$ (not dividing the discriminant $\Delta(f)$), we get the induced polynomial $g = (f \mod p) \in \mathbb{F}_p[x]$ (where each coefficient of f in $\mathbb{Z}$ is reduced modulo p). The polynomial g has a splitting field K and Galois group G over $\mathbb{F}_p$, and the Frobenius map $F \in G$ acts on the roots of g as a permutation with cycle type determined by the decomposition type (degrees of irreducible factors) of g. The goal of this section is to lift $F \in G$ to a Frobenius substitution $\sigma_p \in M$ that has the same cycle type as F. This σ_p will not be unique, but will be unique up to conjugation, and since conjugacy classes in a permutation group correspond to cycle types, we will obtain the cycle type of an element of M. In other words, we gain information about M by reducing f modulo a prime to get g, with the decomposition type of g corresponding to the cycle type of an element of M.

To understand the Frobenius substitution σ_p, we must first understand the mathematical notion of a place.

Definition 4.3.1. *Let $E/\mathbb{Q}$ be a finite extension. A **place** of E over p is a map $\psi : E \to \overline{\mathbb{F}_p} \cup \{\infty\}$ for which*

1. *$\psi^{-1}(\overline{\mathbb{F}_p})$ is a subring of E, and $\psi : \psi^{-1}(\overline{\mathbb{F}_p}) \to \overline{\mathbb{F}_p}$ is a ring homomorphism, and*
2. *For nonzero $x \in E$, $\psi(x) = \infty \iff \psi(x^{-1}) = 0$*

Note that the symbol ∞ is required if we want to take elements of E mod p, since if we desire $p \mod p = 0$, then $\frac{1}{p} \mod p = \frac{1}{0} \mod p = \infty$. Here are the basic facts about places:

Proposition 4.3.2. *With the setup above, let $p \nmid \Delta(f)$.*

1. *A place of E over p exists.*
2. *If ψ and ψ' are places of E over p, then $\psi = \psi' \circ \tau$ for some $\tau \in M = \mathrm{Gal}(E/\mathbb{Q})$.*
3. *The element $\tau \in M$ in (2) is determined by ψ and ψ'.*
4. *If ψ is a place of E over p, and $\alpha_1, \ldots, \alpha_n$ are the roots of f, with $g = f \mod p$, then $\psi(\alpha_1), \ldots, \psi(\alpha_n)$ are the roots of g in $\overline{\mathbb{F}_p}$.*

Proof. Let $\alpha_1, \ldots, \alpha_n$ be the roots of f (monic and irreducible, with integer coefficients), let p be a prime not dividing the discriminant of f, and let $\beta_1, \ldots, \beta_n$ be the roots of $g = f \mod p$. Hence, both f and g are separable, and $E = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$.

We first show that no place $\psi : E \to \overline{\mathbb{F}_p} \cup \{\infty\}$ of E over p sends any root α_i to ∞. By way of contradiction, and without loss of generality, assume $\psi(\alpha_1) = \infty$. Then by the definition of a place (as f is irreducible, 0 is not a root of f), $\psi(\alpha_1^{-1}) = 0$, so if $R = \psi^{-1}(\overline{\mathbb{F}_p})$, $\alpha_1^{-1} \in R$, but $\alpha_1 \notin R$. Now, since $\mathbb{Q}$ is a field, $\mathbb{Q}(\alpha_1) = \mathbb{Q}(\alpha_1^{-1}) = \mathbb{Q}[\alpha_1^{-1}]$, so there exists $h \in \mathbb{Q}[x]$ such that $\alpha_1 = h(\alpha_1^{-1})$. In fact, we can construct h from f as follows: if $f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ for some $c_i \in \mathbb{Z}$, α_1 a root of f implies $f(\alpha_1) = \alpha^n + c_{n-1}\alpha_1^{n-1} + \cdots + c_1\alpha_1 + c_0 = 0$, so $(\alpha_1^{-1})^{n-1} f(\alpha_1) = \alpha_1 + c_{n-1} + \cdots + c_1(\alpha_1^{-1})^{n-2} + c_0(\alpha_1^{-1})^{n-1} = 0$. Solving for α_1, we obtain

$\alpha_1 = -[c_{n-1} + c_{n-2}\alpha_1^{-1} + \cdots + c_1(\alpha_1^{-1})^{n-2} + c_0(\alpha_1^{-1})^{n-1}]$. This right hand side is our h, with coefficients in $\mathbb{Z} \subseteq R$, so $\alpha_1 = h(\alpha_1^{-1}) \in R$ (since R is a ring), a contradiction. Hence, no $\psi(\alpha_i) = \infty$.

Next, we show that $R = \mathbb{Z}_{(p)}[\alpha_1, \ldots, \alpha_n]$, where $\mathbb{Z}_{(p)} = \{\frac{a}{b} \in \mathbb{Q} \mid p \nmid b\}$. Since $R \subseteq E$ must be a subring, $0, 1 \in R$, and so $\mathbb{Z} \subseteq R$. Since $\overline{\mathbb{F}_p}$ has characteristic p and $\psi|_R$ is a ring homomorphism, ψ maps $\mathbb{Z}$ to the prime subring of $\overline{\mathbb{F}}_p$, which is $\mathbb{F}_p$. Hence, by the universal property of localization, since $\mathbb{Z} \setminus p\mathbb{Z} \subseteq \mathbb{Z}$ is multiplicatively closed, contains 1, and avoids the kernel (is in fact the complement of the kernel in this case), the map $\psi|_{\mathbb{Z}}$ extends to a well-defined ring homomorphism on $\mathbb{Z}_{(p)} = (\mathbb{Z} \setminus p\mathbb{Z})^{-1}\mathbb{Z} \subseteq E$. In other words, for $\frac{a}{b} \in \mathbb{Q}$ reduced, we have $\psi(\frac{a}{b}) = \begin{cases} (a \mod p)(b \mod p)^{-1}, \text{ if } p \nmid b \text{ (i.e. } \frac{a}{b} \in \mathbb{Z}_{(p)}), \\ \infty, \text{ if } p \mid b \text{ (i.e. } \frac{a}{b} \notin \mathbb{Z}_{(p)}) \end{cases}$, the last part coming from the definition of a place, since if $p \mid b$, $(\frac{a}{b})^{-1} = \frac{b}{a} \in \mathbb{Z}_{(p)}$ (since reduced) is sent via ψ to 0. Then by our previous argument, each $\alpha_i \in R$, so $R = \mathbb{Z}_{(p)}[\alpha_1, \ldots, \alpha_n]$ (since R contains $\mathbb{Z}_{(p)}$ and the α_i, and no other elements of $\mathbb{Q}$ by above).

Now, we show that each $\psi(\alpha_i)$ is a root of g. We do this by proving the commutativity of the diagram

$$\begin{array}{ccc} R & \xrightarrow{f|_R} & R \\ {\scriptstyle \psi|_R}\downarrow & & \downarrow{\scriptstyle \psi|_R} \\ \overline{\mathbb{F}_p} & \xrightarrow{g} & \overline{\mathbb{F}_p} \end{array}$$

, i.e. that $g \circ \psi|_R = \psi|_R \circ f|_R$. For any $x \in R$, if $f(x) = \sum c_i x^i$, $\psi|_R$ is a ring homomorphism, so $\psi|_R(f(x)) = \sum \psi(c_i)\psi(x)^i = (c_i \mod p)\psi(x)^i = g(\psi(x))$, since $g = f \mod p$ by definition. Hence, the diagram commutes. In particular, evaluating x at $\alpha_i \in R$ yields $g(\psi(\alpha_i)) = \psi(f(\alpha_i)) = \psi(0) = 0$, again since $\psi|_R$ is a ring homomorphism. Thus, each $\psi(\alpha_i)$ is a root of g. Furthermore, if $\psi(\alpha_i) = \psi(\alpha_j)$ for some $i \neq j$, then $\psi|_R$ a ring homomorphism and $\Delta(f) = \prod_{k<l}(\alpha_k - \alpha_l)^2$ gives $0 = \psi(\Delta(f)) = (\Delta(f) \mod p)$, but $p \nmid \Delta(f)$, a contradiction. Hence, the $\psi(\alpha_i)$ are distinct, and so since $\deg(f) = \deg(g) = n$, $\{\psi(\alpha_1), \ldots, \psi(\alpha_n)\}$ is the

collection of roots of g.

Finally, we give a construction of all the places of E over p, revealing that places exist and are related to one another via composition with Galois group elements. We do this by reviewing the algorithm for (non-uniquely) constructing E as the splitting field of f. We already know how ψ must act on $\mathbb{Q}$ by our above argument, so we just have to determine $\psi(\alpha_i)$ for each root α_i, and that this gives a ring homomorphism. First, without loss of generality $\mathbb{Q} \subsetneq \mathbb{Q}[\alpha_1] = \mathbb{Q}[x]/(f) \subseteq E$, so we define $\psi(\alpha_1) = \beta_1$. If $\mathbb{Q}[\alpha_1] = E$, each α_i can be written as a polynomial in α_1, so extend ψ accordingly so that it is a ring homomorphism, and we are done. Note that in this case, our Galois group M must be isomorphic to $\mathbb{Z}/n\mathbb{Z}$. If $\mathbb{Q}[\alpha_i] \neq E$, still extend ψ so it is a ring homomorphism on each $\alpha_i \in \mathbb{Q}[\alpha_1]$ as before. After possibly relabeling, say $\alpha_2 \notin \mathbb{Q}[\alpha_1]$. Hence, $\mathbb{Q} \subsetneq \mathbb{Q}[\alpha_1] \subsetneq \mathbb{Q}[\alpha_1, \alpha_2] \subseteq E$, so define (again after possibly relabeling so that β_2 is not $\psi(\alpha_i)$ for any $\alpha_i \in \mathbb{Q}[\alpha_1]$), define $\psi(\alpha_2) = \beta_2$. Again, if $\mathbb{Q}[\alpha_1, \alpha_2] = E$, we can write the remaining α_i as polynomials in α_1 and α_2, so extend ψ accordingly so that it is a ring homomorphism. If $\mathbb{Q}[\alpha_1, \alpha_2] \neq E$, we repeat this process, which is guaranteed to terminate since by definition $E = \mathbb{Q}[\alpha_1, \ldots, \alpha_n]$. Note that if we had to do the process the maximum of n times, this implies that our Galois group M must be isomorphic to S_n.

With this process, we have $\psi(\alpha_1) = \beta_1$, and for $i \neq 1$, there exists β_j with $\psi(\alpha_i) = \beta_j$, so we get a permutation of $\{\alpha_1, \ldots, \alpha_n\}$ based on the indices of their images. Due to the choices made in our construction, any other permutation from a place ψ' with $\psi(\alpha_1) = \beta_1$ is only achievable if there is a Galois group element $\tau \in M$ such that $\psi = \psi' \circ \tau$, and this τ is unique since we are viewing the field automorphism based on how it permutes the roots of f. Furthermore, the choice $\psi(\alpha_1) = \beta_1$ was arbitrary, so by considering the other choices $\psi(\alpha_i) = \beta_1$, and compositions with elements of our Galois group M, we get all possible constructions of E (up to isomorphism), as well as the corresponding permutations of the roots of f, which define corresponding places. Note note all permutations are necessarily achievable in this way, but the permutations that do arise are

precisely those given by M acting on the roots of f, completing our proof. $\square$

Definition 4.3.3. *Let $p \nmid \Delta(f)$ and let ψ be a place of E over p (exists by Proposition 4.3.2). Then, $\mathrm{Frob} \circ \psi$ is also a place of E over p (after defining $\mathrm{Frob}(\infty) = \infty$), and so by Proposition 4.3.2, there is a unique element, $\mathrm{Frob}_\psi \in M$, called the **Frobenius substitution**, for which $\psi \circ \mathrm{Frob}_\psi = \mathrm{Frob} \circ \psi$.*

Proposition 4.3.4. *$\mathrm{Frob} \in G$ and $\mathrm{Frob}_\psi \in M$, viewed as permutations, have the same cycle type.*

Proof. By definition, Frob_ψ is characterized by $\psi(\mathrm{Frob}_\psi(x)) = \mathrm{Frob}(\psi(x))$ for all $x \in E$. Hence, Frob_ψ permutes the roots $\alpha_1, \ldots, \alpha_n$ of f in the same way as Frob permutes the roots $\psi(\alpha_1), \ldots, \psi(\alpha_n)$ of g, and so Frob and Frob_ψ have the same cycle type. $\square$

Proposition 4.3.5. *While Frob_ψ depends on the choice of place ψ, it is unique up to conjugation.*

Proof. By the previous proposition, any other place of E over p is of the form $\psi \circ \tau$. So by the definition of the Frobenius substitution (with $\tau(x)$ replacing x), $\psi(\mathrm{Frob}_\psi(\tau(x))) = \mathrm{Frob}(\psi(\tau(x)))$. Inserting a $\tau \circ \tau^{-1}$ in a clever way on the left hand side, we obtain, $(\psi \circ \tau)((\tau^{-1} \circ \mathrm{Frob}_\psi \circ \tau)(x)) = \mathrm{Frob}((\psi \circ \tau)(x))$. However, this is the characterizing property of the Frobenius substitution $\mathrm{Frob}_{\psi\circ\tau}$, and so $\mathrm{Frob}_{\psi\circ\tau} = \tau^{-1} \circ \mathrm{Frob}_\psi \circ \tau$. Hence, as ψ varies over the places of E over a fixed prime p, $\mathrm{Frob}\,\psi$ ranges over a conjugacy class in M. $\square$

We denote by σ_p a general member of the above conjugacy class, which is our desired Frobenius substitution as discussed at the beginning of this section.

Everything that has been discussed in this section can be extended in a natural way to the case where $f \in L[x]$, where L is a finite extension of $\mathbb{Q}$. See Lang VII for more details.

4.4 Using Frobenius to Study Galois Groups over $\mathbb{Q}$

For us, we will be fixing a "large" prime p (think 10009 for example). We consider enumerative geometry problems with a parameter space, where the number of solutions is constant for a dense open subset of parameters. For each choice of parameters (an "instance" of the problem), we obtain a zero-dimensional ideal I, which via proper choice of Gröbner basis has an eliminant $f \in \mathbb{Q}[x]$,

which is the f we consider as in the previous section. The roots of this eliminant f completely determine the solutions to our instance of our problem. Importantly, we can reduce f modulo p to obtain a polynomial $g \in \mathbb{F}_p[x]$, and note by the previous sections that its decomposition type is the same as the cycle type of our lifted Frobenius substitution σ_p in our Galois group M. In this way, we obtain information (a cycle type) about one element of M. By choosing random parameters for the problem again, we again gain information about the cylcle type of an element of M. In fact, for p large enough, we are sampling $\sigma_p \in M$ uniformly, and so the frequencies of the cycle types obtained via this Frobenius algorithm approach the distribution of the cycle types in M. By sampling many elements and seeing which cycle types arise and at what frequencies, one can often uniquely identify the group M. Since for a problem with d solutions, $M \subseteq S_d$ is a transitive subgroup, understanding the structure of the transitive subgroups of S_d is critical in identifying the group M.

We demonstrate this again using the SchubertIdeals.m2 package, as well as our additional coding up of the Frobenius algorithm. Below our example is again Derksen's problem that 6 4-planes in $\mathrm{Gr}(4,8)$ intersect four general 4-planes in $\mathrm{Gr}(4,8)$, each in dimension at least two.

```
loadPackage("SchubertIdeals")
loadPackage("RealRoots")

smartFactor = method()
smartFactor(RingElement) := (F) ->(
    P:=factor(F);
    L1:=new List from(P);
    L2:=for p in L1 list(new List from p);
    return(L2)
    )

-- Heuristic: numiterations = 6*numsols
frobeniusAlgorithm = method()
frobeniusAlgorithm(List,ZZ,ZZ,ZZ) := (L,p,numsols,numiterations) -> (
    flagtype := L_(0);
    n := last(flagtype);
    conditions := L_(1);
    l := length(conditions);
```

```
    P := ZZ/p;
    datastuff := {};
    fullcycle := false;
    fullminusonecycle := false;
    primecycle := false;
    for i from 1 to numiterations do(
flags := {};
for j from 1 to (l-1) do(
    flags = append(flags,random(P^n,P^n)));
if det(product(flags)) == 0 then continue;
I := typeASchubertIdeal(flagtype,conditions,flags,P);
f := smartFactor univariateEliminant(sum(gens ring(I)),I);
degreecyclelist := sort(flatten for fac in f list(degree(fac#0)));
if sum(degreecyclelist) != numsols then continue;
        if sum(degreecyclelist) == numsols then datastuff = append(datastuff,degreecyclelist);
if degreecyclelist == {numsols} then fullcycle = true;
if degreecyclelist == {1,numsols-1} then fullminusonecycle = true;
for k in degreecyclelist do(
     if (k > numsols/2) and (k <= numsols-2) and (isPrime(k)==true) then primecycle = true);
if ((fullcycle == true) and (fullminusonecycle == true) and (primecycle == true)) then break;
   );
    frequencytable := {};
    for cycle in unique(datastuff) do(
        frequencytable = append(frequencytable,(cycle,number(datastuff,i->i==cycle))));
    return(fullcycle,fullminusonecycle,primecycle,frequencytable)
 )

frobeniusDegreeThree = method()
frobeniusDegreeThree(List,ZZ,ZZ,ZZ) := (L,p,numsols,numiterations) -> (
    flagtype := L_(0);
    n := last(flagtype);
    conditions := L_(1);
    l := length(conditions);
    P := ZZ/p;
    datastuff := {};
    twocycle := false;
    for i from 1 to numiterations do(
flags := {};
for j from 1 to (l-1) do(
    flags = append(flags,random(P^n,P^n)));
if det(product(flags)) == 0 then continue;
```

```
I := typeASchubertIdeal(flagtype,conditions,flags,P);
f := smartFactor univariateEliminant(sum(gens ring(I)),I);
degreecyclelist := sort(flatten for fac in f list(degree(fac#0)));
if sum(degreecyclelist) != numsols then continue;
        if sum(degreecyclelist) == numsols then datastuff = append(datastuff,degreecyclelist);
if degreecyclelist == {1,2} then twocycle = true;
if (twocycle == true) then break;
   );
    frequencytable := {};
    for cycle in unique(datastuff) do(
        frequencytable = append(frequencytable,(cycle,number(datastuff,i->i==cycle))));
    return(twocycle,frequencytable)
 )

frobeniusDegreeFour = method()
frobeniusDegreeFour(List,ZZ,ZZ,ZZ) := (L,p,numsols,numiterations) -> (
    flagtype := L_(0);
    n := last(flagtype);
    conditions := L_(1);
    l := length(conditions);
    P := ZZ/p;
    datastuff := {};
    threecycle := false;
    fourcycle := false;
    for i from 1 to numiterations do(
flags := {};
for j from 1 to (l-1) do(
    flags = append(flags,random(P^n,P^n)));
if det(product(flags)) == 0 then continue;
I := typeASchubertIdeal(flagtype,conditions,flags,P);
f := smartFactor univariateEliminant(sum(gens ring(I)),I);
degreecyclelist := sort(flatten for fac in f list(degree(fac#0)));
if sum(degreecyclelist) != numsols then continue;
        if sum(degreecyclelist) == numsols then datastuff = append(datastuff,degreecyclelist);
if degreecyclelist == {1,3} then threecycle = true;
if degreecyclelist == {4} then fourcycle = true;
if ((threecycle == true) and (fourcycle == true)) then break;
   );
    frequencytable := {};
    for cycle in unique(datastuff) do(
        frequencytable = append(frequencytable,(cycle,number(datastuff,i->i==cycle))));
```

```
    return(threecycle,fourcycle,frequencytable)
 )

frobeniusDegreeFive = method()
frobeniusDegreeFive(List,ZZ,ZZ,ZZ) := (L,p,numsols,numiterations) -> (
    flagtype := L_(0);
    n := last(flagtype);
    conditions := L_(1);
    l := length(conditions);
    P := ZZ/p;
    datastuff := {};
    twothreecycle := false;
    for i from 1 to numiterations do(
flags := {};
for j from 1 to (l-1) do(
    flags = append(flags,random(P^n,P^n)));
if det(product(flags)) == 0 then continue;
I := typeASchubertIdeal(flagtype,conditions,flags,P);
f := smartFactor univariateEliminant(sum(gens ring(I)),I);
degreecyclelist := sort(flatten for fac in f list(degree(fac#0)));
if sum(degreecyclelist) != numsols then continue;
        if sum(degreecyclelist) == numsols then datastuff = append(datastuff,degreecyclelist);
if degreecyclelist == {2,3} then twothreecycle = true;
if (twothreecycle == true) then break;
   );
    frequencytable := {};
    for cycle in unique(datastuff) do(
        frequencytable = append(frequencytable,(cycle,number(datastuff,i->i==cycle))));
    return(twothreecycle,frequencytable)
 )

frobeniusDegreeSix = method()
frobeniusDegreeSix(List,ZZ,ZZ,ZZ) := (L,p,numsols,numiterations) -> (
    flagtype := L_(0);
    n := last(flagtype);
    conditions := L_(1);
    l := length(conditions);
    P := ZZ/p;
    datastuff := {};
    twothreecycle := false;
    fivecycle := false;
```

```
    for i from 1 to numiterations do(
flags := {};
for j from 1 to (l-1) do(
    flags = append(flags,random(P^n,P^n)));
if det(product(flags)) == 0 then continue;
I := typeASchubertIdeal(flagtype,conditions,flags,P);
f := smartFactor univariateEliminant(sum(gens ring(I)),I);
degreecyclelist := sort(flatten for fac in f list(degree(fac#0)));
if sum(degreecyclelist) != numsols then continue;
        if sum(degreecyclelist) == numsols then datastuff = append(datastuff,degreecyclelist);
if degreecyclelist == {1,2,3} then twothreecycle = true;
if degreecyclelist == {1,5} then fivecycle = true;
if ((twothreecycle == true) and (fivecycle == true)) then break;
   );
    frequencytable := {};
    for cycle in unique(datastuff) do(
        frequencytable = append(frequencytable,(cycle,number(datastuff,i->i==cycle))));
    return(twothreecycle,fivecycle,frequencytable)
 )

-- EXAMPLE of Computing Frobenius Elements for the Galois group to Derksen's Problem:

i3 : loadPackage("SchubertIdeals", Reload => true)

o3 = SchubertIdeals

o3 : Package

i4 : loadPackage("RealRoots", Reload => true)

o4 = RealRoots

o4 : Package

i6 : F1 = random((ZZ/10009)^8,(ZZ/10009)^8)

o6 = | -2926 -3311 2520  -1653 -3648 -2745 3186  4037  |
     | -3427 -6    1411  4873  3358  3954  -301  4208  |
     | 3638  -1963 3419  3139  3769  765   -109  -1570 |
     | -3682 3763  1830  2797  152   1123  2298  2878  |
     | -2291 2776  -295  -2869 -1482 714   -4604 2581  |
```

```
     | 17    -1528 -4273 -4041 -3346 2855  1648  4041  |
     | 2467  1438  1529  -4985 -4484 -1338 -116  4615  |
     | -399  -1372 2658  4400  -4998 983   -3871 -398  |

               ZZ  8         ZZ  8
o6 : Matrix (-----)  <--- (-----)
             10009         10009

i7 : F2 = random((ZZ/10009)^8,(ZZ/10009)^8)

o7 = | -4007 4394  3212  5000  1320  4499  419   1872  |
     | -4479 3422  -1330 3639  453   2740  65    -1030 |
     | -1465 -1134 -4371 -572  -2050 1706  -2566 2959  |
     | 698   1600  366   3820  -785  -2493 -322  3530  |
     | 3459  -1341 -2347 -4912 2857  457   -2397 -3723 |
     | -1769 1892  1223  415   -2977 2458  -2397 3705  |
     | -760  -3231 232   2721  -2527 -2362 -1187 -1235 |
     | 3360  -139  1562  -4920 -2764 4007  -3354 -4031 |

               ZZ  8         ZZ  8
o7 : Matrix (-----)  <--- (-----)
             10009         10009

i8 : F3 = random((ZZ/10009)^8,(ZZ/10009)^8)

o8 = | 1381  -1970 -597  1998  -2628 -2523 1868  3761  |
     | 3800  4624  2963  3687  -1470 -2936 -4126 -4840 |
     | -2190 1428  3314  -4063 4072  -3064 1695  -294  |
     | -681  3514  -4748 -1250 606   2980  -4581 1719  |
     | -3800 1107  -1204 1715  -1917 -2615 -3781 3177  |
     | -2221 135   4041  -4391 4678  3540  31    -953  |
     | -1929 1394  -2979 -4580 3695  -4196 -3347 2736  |
     | -1569 -94   3216  -4032 -4238 -3734 -4181 36    |

               ZZ  8         ZZ  8
o8 : Matrix (-----)  <--- (-----)
             10009         10009

i9 :
     I = typeASchubertIdeal({4,8},{{1,2,5,6},{1,2,5,6},{1,2,5,6},{1,2,5,6}},{F1,F2,F3},ZZ/10009);
```

```
                ZZ
o9 : Ideal of -----[x   ..x   , x   ..x   ]
              10009  3,1   4,2   7,1   8,4

i10 : dim I

o10 = 0

i11 : degree I

o11 = 6

i12 : eliminant = smartFactor univariateEliminant(sum(gens ring(I)),I)

o12 = {{Z - 141, 1}, {Z - 466, 1}, {Z - 713, 1}, {Z - 1170, 1}, {Z - 2066, 1}, {Z - 3906, 1}}

o12 : List

i13 : cycletype = sort(flatten for fac in eliminant list(degree(fac#0)))

o13 = {1, 1, 1, 1, 1, 1}

o13 : List

i14 : F1 = random((ZZ/10009)^8,(ZZ/10009)^8)

o14 = | 1581  552   -2882 4485  -1381 4605  4978  -3424 |
      | 3800  2474  -3307 -3423 -2775 -524  -2446 -3526 |
      | -1892 748   4333  -1690 1865  3346  -3020 -3247 |
      | 2201  476   4156  -3199 663   -4577 1040  -2275 |
      | -4437 935   -2180 -3067 -2022 -2732 -859  2859  |
      | -2688 -2709 -1498 2894  3293  1078  -288  1873  |
      | 583   -3810 4132  2528  -1686 -4772 2189  2556  |
      | 4133  -1368 -2615 732   2834  -2933 -1329 -1855 |

               ZZ  8         ZZ  8
o14 : Matrix (-----)  <--- (-----)
              10009         10009

i15 : F2 = random((ZZ/10009)^8,(ZZ/10009)^8)
```

```
o15 = | -1297 4666  3548  -3192 -1487 -3572 1204  2505  |
      | -4931 3012  2371  1187  1787  4850  -3818 -116  |
      | -1390 -1976 3919  3582  -3597 458   4784  -4154 |
      | 2271  36    472   -2763 93    1456  -4452 1404  |
      | -2165 -4532 4678  -3836 -3674 2267  -3159 -4749 |
      | -2081 95    -2112 -4318 1993  -4144 2632  -2962 |
      | -2824 -2831 3915  -1599 3271  -3335 -4127 -1883 |
      | -1857 -4492 -647  -3964 4507  159   -3050 -4585 |

               ZZ  8         ZZ  8
o15 : Matrix (-----)  <--- (-----)
             10009         10009

i16 : F3 = random((ZZ/10009)^8,(ZZ/10009)^8)

o16 = | -2820 -845  1149  -4871 -3936 1561  -4751 -2405 |
      | 10    4929  -960  2738  2560  166   2201  3206  |
      | -2973 2637  -337  -2025 4378  1548  444   -254  |
      | 588   -1712 1263  -1901 -1807 4419  -3163 1360  |
      | 4405  -983  -3934 -767  -3737 -3412 4099  646   |
      | -2578 4764  933   -4600 -4186 2794  4480  56    |
      | 4923  3711  589   -3667 4811  4070  3592  342   |
      | 1576  -1568 359   1651  -3510 3175  -2423 3039  |

               ZZ  8         ZZ  8
o16 : Matrix (-----)  <--- (-----)
             10009         10009

i17 : I = typeASchubertIdeal({4,8},{{1,2,5,6},{1,2,5,6},{1,2,5,6},{1,2,5,6}},{F1,F2,F3},ZZ/10009);

               ZZ
o17 : Ideal of -----[x   ..x   , x   ..x   ]
              10009  3,1   4,2   7,1   8,4

i18 : eliminant = smartFactor univariateEliminant(sum(gens ring(I)),I)

         3       2                     3        2
o18 = {{Z  - 779Z  - 802Z - 460, 1}, {Z  - 1132Z  + 4797Z - 2208, 1}}

o18 : List
```

```
i19 : cycletype = sort(flatten for fac in eliminant list(degree(fac#0)))

o19 = {3, 3}

o19 : List

i20 : for i from 1 to 4 do(
      print(frobeniusDegreeSix({{4,8},{{1,2,5,6},{1,2,5,6},{1,2,5,6},{1,2,5,6}}},10009,6,25)))
(false, false, {({1, 1, 2, 2}, 7), ({2, 4}, 10), ({3, 3}, 6), ({1, 1, 1, 1, 1, 1}, 2)})
(false, false, {({1, 1, 2, 2}, 11), ({3, 3}, 8), ({1, 1, 1, 1, 1, 1}, 1), ({2, 4}, 5)})
(false, false, {({1, 1, 2, 2}, 11), ({2, 4}, 6), ({1, 1, 1, 1, 1, 1}, 2), ({3, 3}, 6)})
(false, false, {({3, 3}, 12), ({1, 1, 2, 2}, 10), ({2, 4}, 3)})
```

In the above example, for the first choice of general flags we computed the identity permutation as a Frobenius element of our Galois group, given by $\{1, 1, 1, 1, 1, 1\}$. We then chose another set of general flags, and this time obtained a $(3, 3)$-cycle as a Frobenius element of our Galois group, given by $\{3, 3\}$. Note that we do not know the exact element, just its cycle type. We then iterate and compute four batches of 25 Frobenius elements each, obtaining a frequency table in each case of cycle types observed. In each batch, the "false, false" in the output means that no $(2, 3)$-cycles or 5-cycles were observed. This is because there are 6 solutions to Derksen's problem, and so the Galois group is a subgroup of S_6, which is generated by any $(2, 3)$-cylce and 5-cycle, the obtaining of which would have allowed us to instantly determine that the Galois group was full-symmetric. However, no such cycles were found in our batches, but instead in the first batch of 25 Frobenius elements, 7 $(2, 2)$-cycles were found, 10 $(2, 4)$-cycles were obtained, and the rest were 6 $(3, 3)$-cycles and the identity twice. Looking at the other batches (all together, we computed here 102 Frobenius element cycle types), no other cycle types than those mentioned were observed, even though there are many more cycle types in S_6. All of this data indicates that the Galois group for Derksen's problem is enriched (though itself is not a proof), and in fact Derksen showed that the Galois group is a copy of S_4 in S_6, and so has $24 < 720$ elements. In Figure 4.1, it can be observed that the only transitive subgroup of S_6 with only the observed cycle types is labeled S_4 (b), matching Derksen's finding.

- For degree 6, there are 16 transitive subgroups of S_6, with generators and cycle types as follows:

#	Order	!	Name	Generators	1	2	2,2	2,3	2,4	2,2,2	3	3,3	4	5	6
6T1	6		C_6	(123456)	1					1		2			2
6T2	6		S_3	(135)(246), (14)(23)(56)	1					3		2			
6T3	12		$S_3 \times C_2$	(123456), (14)(23)(56)	1		3			4		2			2
6T4	12	□	A_4	(14)(25), (135)(246)	1		3					8			
6T5	18		F_{18}	(246), (14)(25)(36)	1					3	4	4			6
6T6	24		$A_4 \times C_2$	(36), (135)(246)	1	3	3			1		8			8
6T7	24	□	S_4 (a)	(14)(25), (135)(246), (15)(24)	1		9		6			8			
6T8	24		S_4 (b)	(14)(25), (135)(246), (15)(24)(36)	1		3			6		8	6		
6T9	36		$S_3 \times S_3$	(246), (15)(24), (14)(25)(36)	1		9			6	4	4			12
6T10	36	□	F_{36}	(246), (15)(24), (1452)(36)	1		9		18		4	4			
6T11	48		$S_4 \times C_2$	(36), (135)(246), (15)(24)	1	3	9		6	7		8	6		8
6T12	60	□*	A_5	(12346), (14)(56)	1		15					20		24	
6T13	72		$F_{36} \rtimes C_2$	(246), (24), (14)(25)(36)	1	6	9	12	18	6	4	4			12
6T14	120	*	S_5	(12346), (12)(34)(56)	1		15			10		20	30	24	20
6T15	360	□*	A_6	(12)(3456), (123)	1		45		90		40	40		144	
6T16	720	*	S_6	(123456), (12)	1	15	45	120	90	15	40	40	90	144	120

Figure 4.1: Cycle Types of Transitive Subgroups of S_6

4.5 Computation of Some Schubert Galois Groups and Future Work

All Galois groups for Schubert problems in $\mathrm{Fl}(a_1, \ldots, a_s; n)$ with $n \leq 5$ are known to be full-symmetric, but there is a known enriched Schubert problem in $\mathrm{Fl}(2, 4; 6)$ with 6 solutions that has an isomorphic copy of S_3 in S_6 as its Galois group. At the same time, there is not a full classification of enriched Galois groups in $\mathrm{Fl}(a_1, \ldots, a_s; 6)$, and so we have done large-scale experimentation to gather data as to what groups might appear. Our design was as follows:

1. For each shape $(a_1, \ldots, a_s; 6)$, we used Maple code to obtain the list of Schubert conditions for every possible Schubert problem in $\mathrm{Fl}(a_1, \ldots, a_s; 6)$ with 250 solutions or fewer. Our Maple code was refined to omit certain lists of Schubert conditions that give Schubert problems with already understood enriched Galois groups, called triangular Schubert problems, since their groups are products or wreath products of groups already computed. With this modification, there were still a total of $1,812,629$ Schubert problems in the various $\mathrm{Fl}(a_1, \ldots, a_s; 6)$ whose Galois groups we gathered data for.

2. After sufficient testing of running the Frobenius algorithm on various problems, we used multivariate regression to estimate how long it would take to run Frobenius on each Schubert problem based on past performance. As a result, we were able to create batches of Schubert

problems, each of which we expected to take about 24 hours to run.

3. Using the Whistler cluster through the mathematics department at Texas A&M University, we were able to run 32 batches at a time, compared to only 1 batch at a time on a personal laptop computer. Performing this procedurally required a Bash script to interact with the cluster, and a Python file that would run our Macaulay2 implementation of the Frobenius algorithm multiple times per Schubert problem.

4. For each problem with d solutions, we ran the Frobenius algorithm on that problem to compute up to $6d$ cycle types, searching for specific cycle types based on the number of solutions that would force the Galois group to be full-symmetric. Based on the cycle types observed, if the group was proven to be full-symmetric, the algorithm would move on to the next problem. If the full $6d$ cycle types were computed and did not guarantee that the group was full-symmetric, that problem as well as the cycle types computed with their frequencies were stored in an output file.

Future work will entail analyzing the data from the large-scale computations, which took months to perform. Each Schubert problem that wasn't proven to be full-symmetric will be studied by computing more Frobenius elements, and then using the frequency data to determine which transitive subgroup of S_d we believe the group to be. At this point, a geometric argument will be required to prove what the Galois group of the problem is, ideally also revealing what geometric obstructions prohibit the group from being full-symmetric.

Additionally, there are many more Schubert problems to compute beyond those in the partial flag varieties $\mathrm{Fl}(a_1, \ldots, a_s; 6)$. Already our Maple code (over months) has computed all Schubert problems in $\mathrm{Fl}(a_1, 7) = \mathrm{Gr}(a_1, 7)$ and $\mathrm{Fl}(a_1, a_2; 7)$, but beyond Grassmannians and two-step flag varieties, determining all possible Schubert problems is computationally infeasible. As a result, we will have to be more selective in computing data for Schubert Galois groups in the future.

Another avenue of research is to focus on Schubert problems in Type C. Sottile has written a companion package to SchubertIdeals.m2 in Singular, and has used this package to compute cycle types of Frobenius elements for Galois group computations for Schubert problems in Lagrangian Grassmannians $\mathrm{Fl}_C(n, 2n)$. From the data he has computed (which can be found at

https://franksottile.github.io/research/stories/LGalois/index.html

), he has provided some geometric arguments to prove that there are enriched Galois groups beyond the iterated wreath products in Type A Grassmannians. There is still much work to be done in Type C, particularly beyond Lagrangian Grassmannians to other isotropic Type C Grassmannians and partial flag varieties in general. At the same time, we will be comparing the speed and functionality of SchubertIdeals.m2 and the companion Singular package, and will submit the results with the packages to the Journal of Software for Algebra and Geometry.

As for Types B and D (and the exceptional Lie types), nothing is known about Schubert Galois groups. The goal will first be to write software that can find ideals of Schubert problems in this setting, but already there will have to be adaptations in Type D due to the need for pfaffians in the equations.

Already in the analysis of Sottile's Type C computations, we have observed a limitation to the method of computing Frobenius elements. If two transitive subgroups of S_d have the same order and cycle types, with the same frequencies of those cycle types (but the subgroups are not the same), then the Frobenius elements that we compute cannot be used to distinguish between the two. Thus, more techniques are required, and in particular, there is hope that computing monodromy via numerical homotopy methods to compute specific elements of the Galois group (not just the cycle type) will suffice. For Type C this requires more difficult homotopies than those usually employed in computing monodromy, but efforts have been made to overcome this.

REFERENCES

[1] C.J. Brooks, A. Martín del Campo, and F. Sottile, Galois groups of Schubert problems of lines are at least alternating, *Trans. Amer. Math. Soc.* 367 (2015), no. 6, 4183–4206

[2] J.D. Hauenstein, J.I. Rodriguez, and F. Sottile, Numerical Computation of Galois Groups, Found. *Comput. Math.* 18 (2018), no. 4, 867–890.

[3] T. Brysiewicz, J. I. Rodriguez, F. Sottile, and T. Yahl. Decomposable sparse polynomial systems. *J. Softw. Algebra Geom.*, 11(1):53–59, 2021.

[4] T. Brysiewicz, J. I. Rodriguez, F. Sottile, and T. Yahl. Solving decomposable sparse systems. *Numer. Algorithms*, 88(1):453–474, 2021.

[5] B. Huber, F. Sottile, and B. Sturmfels, Numerical Schubert calculus, *Journal of Symbolic Computation* 26 (1998), no. 6, 767–788.

[6] T. Duff, C. Hill, A. Jensen, K. Lee, A. Leykin, and J. Sommars. Solving polynomial systems via homotopy continuation and monodromy. *IMA Journal of Numerical Analysis*, 39(3):1421–1446, 2019.

[7] T. Duff, V. Korotynskiy, T. Pajdla, and M. Regan. Galois/monodromy groups for decomposing minimal problems in 3D reconstruction, 2021. arXiv:2105.04460.

[8] T. Duff, S. Telen, E. Walker, and T. Yahl. Parameter homotopies in Cox coordinates. 2020. arXiv:2012.04255.

[9] A. Esterov. Galois theory for general systems of polynomial equations. *Compos. Math.*, 155(2):229–245, 2019

[10] A. Esterov. Permuting the roots of univariate polynomials whose coefficients depend on parameters. 2022. arXiv:2204.14235.

[11] J. Harris. Galois groups of enumerative problems. *Duke Math. Journal*, 46(4):685–724, 1979.

[12] C. Hermite. Sur les fonctions algébriques. *CR Acad. Sci.(Paris)*, 32:458–461, 1851.

[13] C. Jordan. *Traité des Substitutions et des Équations algébriques.* Gauthier-Villars, Paris, 1870.

[14] A. Leykin and F. Sottile, Galois groups of Schubert problems via homotopy computation, *Math. Comp.* 78 (2009), no. 267, 1749–1765.

[15] A. Martín del Campo and F. Sottile, Experimentation in the Schubert calculus, Schubert Calculus, Osaka 2012 (H. Naruse, T. Ikeda, M. Masuda, and T. Tanisaki, eds.), *Advanced Studies in Pure Mathematics*, vol. 71, Mathematical Society of Japan, 2016, pp. 295–336.

[16] F. Sottile and T. Yahl. Galois groups in enumerative geometry and applications, 2021. arXiv:2108.07905.

[17] T. Yahl. Computing Galois groups of finite Fano problems. 2022. arXiv:2209.07010.

[18] A. Martín del Campo, F. Sottile, and R. Williams, Classification of Schubert Galois groups in Gr(4, 9), Arnold Math J. 9, 393–433 (2023). https://doi.org/10.1007/s40598-022-00221-2.

[19] J. Ruffo, Y. Sivan, E. Soprunova, and F. Sottile, Experimentation and conjectures in the real Schubert calculus for flag manifolds,*Experiment. Math.* 15 (2006), no. 2, 199–221.

[20] F. Sottile and J. White, Double transitivity of Galois groups in Schubert calculus of Grassmannians, *Algebr. Geom.* 2 (2015), no. 4, 422–445.

[21] F. Sottile, J. White, and R. Williams, Galois groups in simple Schubert problems are at least alternating, in preparation.

[22] F. Sottile, R. Williams, and L. Ying, Galois groups of composed Schubert problems, Facets of Algebraic Geometry: A Collection in Honor of William Fulton's 80th Birthday (P. Aluffi, D. Anderson, M. Hering, M. Mustată, and S & Payne, eds.), London Mathematical Society Lecture Note Series, Cambridge University Press, 2022.

APPENDIX A

PROOF OF 27 LINES ON A CUBIC SURFACE

Even though the problem of 27 lines on a cubic surface is not inherently a Schubert problem (since cubic surfaces are not linear spaces), it can still be solved using the Chow ring of the Grassmannian $A(\mathrm{Gr}(2,4))$. In general, the Chow ring $A(\mathrm{Gr}(k,V))$ (equivalently the cohomology ring) can be described by using the Chern classes of two natural vector bundles over $\mathrm{Gr}(k,V)$: $0 \to T \to \underline{V} \to Q \to 0$, where T is the tautological bundle whose fiber over any $H \in \mathrm{Gr}(k,V)$ is the subspace $H \subseteq V$ itself, $\underline{V} = \mathrm{Gr}(k,V) \times V$ is the trivial vector bundle of rank n, with V as fiber, and Q is the quotient vector bundle of rank $n-k$, with V/H as fiber. The Chern classes of the bundles T and Q are $c_i(T) = (-1)^i \sigma_{(1)^i}$ (where $(1)^i$ is the partition whose Young diagram consists of a single column of length i), and $c_i(Q) = \sigma_i$. The tautological sequence then gives the presentation of the Chow ring as $A(\mathrm{Gr}(k,V)) = \frac{\mathbb{Z}[c_1(T),\ldots,c_k(T),c_1(Q),\ldots,c_{n-k}(Q)]}{(c(T)c(Q)-1)}$

For $\mathrm{Gr}(2,4)$, the Chow ring $A(\mathrm{Gr}(2,4))$ has the presentation $A(\mathrm{Gr}(2,4)) = \frac{\mathbb{Z}[\sigma_1,\sigma_{1,1},\sigma_2]}{((1-\sigma_1-\sigma_{1,1})(1+\sigma_1+\sigma_2)-1)}$. Here (in dimension convention), σ_1 is the class of Ω_{24}, $\sigma_{1,1}$ is the class of Ω_{23}, and σ_2 is the class of Ω_{14}. As a graded abelian group, we have that:

- $A^0(\mathrm{Gr}(2,4)) = \mathbb{Z} \cdot 1$ (the class of Ω_{34})
- $A^2(\mathrm{Gr}(2,4)) = \mathbb{Z} \cdot \sigma_1$
- $A^4(\mathrm{Gr}(2,4)) = \mathbb{Z} \cdot \sigma_{1,1} \bigoplus \mathbb{Z} \cdot \sigma_2$
- $A^6(\mathrm{Gr}(2,4)) = \mathbb{Z} \cdot \sigma_{2,1}$ (the class of Ω_{13})
- $A^8(\mathrm{Gr}(2,4)) = \mathbb{Z} \cdot \sigma_{2,2}$ (the class of Ω_{12})

Recall that $\mathrm{Gr}(2,4) \cong \mathbb{G}(1,3)$, and note that the equation of a line can be given as a section of $\Gamma(\mathbb{G}(1,3), T^*)$. Since a cubic surface $X \subseteq \mathbb{P}^3$ is given as a generic homogeneous cubic polynomial, it is also given as a generic section $s \in \Gamma(\mathbb{G}(1,3), \mathrm{Sym}^3(T^*))$. A line $L \subseteq \mathbb{P}^3$ is a subvariety

of X if and only if the section vanishes on $[L] \in \mathbb{G}(1,3)$. Therefore, the Euler class of $\mathrm{Sym}^3(T^*)$ can be integrated over $\mathbb{G}(1,3)$ to get the number of points where the generic section vanishes on $\mathbb{G}(1,3)$. In order to get the Euler class, the total Chern class of T^* must be computed, which is given as $c(T^*) = 1 + \sigma_1 + \sigma_{1,1}$.

The splitting formula then reads as the formal equation $c(T^*) = (1+\alpha)(1+\beta) = 1+\alpha+\beta+\alpha\cdot\beta$, where $c(\mathcal{L}) = 1 + \alpha$ and $c(\mathcal{M}) = 1 + \beta$ for formal line bundles $\mathcal{L}, \mathcal{M}$. The splitting relation gives the relations $\sigma_1 = \alpha + \beta$ and $\sigma_{1,1} = \alpha \cdot \beta$.

Since $\mathrm{Sym}^3(T^*)$ can be viewed as the direct sum of formal line bundles $\mathrm{Sym}^3(T^*) = \mathcal{L}^{\otimes 3} \bigoplus (\mathcal{L}^{\otimes 2} \bigotimes \mathcal{M}) \bigoplus (\mathcal{L} \bigotimes \mathcal{M}^{\otimes 2}) \bigoplus \mathcal{M}^{\otimes 3}$ whose total Chern class is $c(\mathrm{Sym}^3(T^*)) = (1 + 3\alpha)(1 + 2\alpha + \beta)(1 + \alpha + 2\beta)(1 + 3\beta)$, it follows that $c_4(\mathrm{Sym}^3(T^*)) = (3\alpha)(2\alpha + \beta)(\alpha + 2\beta)(3\beta) = 9\alpha\beta[2(\alpha + \beta)^2 + \alpha\beta] = 9\sigma_{1,1}(2\sigma_1^2 + \sigma_{1,1} = 27\sigma_{2,2}$. Above, we used the fact that in the Chow ring, $\sigma_{1,1} \cdot \sigma_1^2 = \sigma_{2,1} \cdot \sigma_1 = \sigma_{2,2}$ and $\sigma_{1,1}^2 = \sigma_{2,2}$.

Since $\sigma_{2,2}$ is the top class (the class of a point), the integral is then $\int_{\mathbb{G}(1,3)} 27\sigma_{2,2} = 27$. $\square$

APPENDIX B

PROOF THAT MONODROMY GROUPS AND GALOIS GROUPS ARE EQUIVALENT

Let F be a field. Let $\pi : X \to Z$ be a branched cover of irreducible varieties over F. As π by definition is dominant, the function field $F(Z)$ of Z embeds as a subfield of the function field $F(X)$ of X. This realizes $F(X)/F(Z)$ as a finite extension of degree d, where d is the degree of π. Let K be the normal closure of this extension, and so $K/F(Z)$ is a Galois extension with corresponding Galois group Gal_π, also called the Galois group of the branched cover π. Note that Gal_π is a transitive subgroup of S_d that is well-defined up to conjugation.

There is also a geometric construction of Gal_π. For $s = 1, \ldots, d$, let X_Z^s be the s-fold iterated fiber product of $\pi : X \to Z$, so $X_Z^s = X \times_Z X \times_Z \cdots \times_Z X$ (with s factors in the product). The fiber of $\pi^s : X_Z^s \to Z$ over a point $z \in Z$ is the s-fold Cartesian product $(\pi^{-1}(\{z\}))^s$ of the fiber of π over z.

The fiber product has many irreducible components when $s > 1$, possibly of different dimensions. Let $U \subseteq Z$ be the maximal dense open subset over which π is étale - fibers $\pi^{-1}(\{z\})$ for $z \in U$ are zero-dimensional reduced schemes of degree d. The complement of U is called the branch locus B of π. The big diagonal of X_Z^s is the closed subscheme consisting of s-tuples with a repeated coordinate. Let $X_Z^{(s)}$ be the closure in X_Z^s of the complement of the big diagonal in $(\pi^s)^{-1}(U)$. The fiber of $X_Z^{(s)}$ over a point $z \in U(\overline{F})$ consists of s-tuples of distinct points of the fiber $\pi^{-1}(\{z\})$

When $s = d$, the symmetric group S_d acts on $X_Z^{(d)}$, permuting each d-tuple. It permutes the irreducible components and acts simply transitively on the fiber above a point $z \in U(\overline{F})$. Let $X' \subseteq X_Z^{(d)}$ be an irreducible component (they are all isomorphic when $s = d$).

We compare this to the construction of the splitting field of a single-variable polynomial. Replacing X and Z by appropriate affine open subsets, we may embed X as a hypersurface in $Z \times \mathbb{A}_t^1$ with $\pi : X \to Z$ the projection. Writing $F[X]$ and $F[Z]$ for their coordinate rings, there is a monic

irreducible polynomial $f \in (F[Z])[t]$ of degree d such that $F[X] = (F(Z))[t]/(f) = (F(Z))(\alpha)$, where α is the image of t in $F[X]$. If X' is an irreducible component of $X_Z^{(d)}$, then $F(X') = (F(Z))(\alpha_1, \ldots, \alpha_d)$, where $\alpha_i \in F[X']$ is given by the composition of inclusion $X' \subseteq X_Z^{(d)}$, the i-th coordinate projection $X_Z^{(d)} \to X$, and the function α. As $i \neq j \implies \alpha_i \neq \alpha_j$ (X' does not lie on the big diagonal), we see that $\alpha_1, \ldots, \alpha_d$ are the roots of f in $F(X')$. Thus, $F(X')$ is the splitting field of f and hence is Galois over $F(Z)$.

The monodromy group Mon_π of the branched cover is the subgroup of S_d that preserves X'. Elements of Mon_π are automorphisms of the extension $F(X')/F(Z)$, so that $\mathrm{Mon}_\pi \subseteq \mathrm{Gal}((F(X')/F(Z))$, the Galois group of $F(X')/F(Z)$. Since Mon_π acts simply transitively on fibers of $X' \to Z$ above points in $U(\overline{F})$, its order is the degree of the map $X' \to Z$, which is the order of the field extension $F(X')/F(Z)$. Hence, we arrive at the desired result that $\mathrm{Mon}_\pi = \mathrm{Gal}((F(X')/F(Z)) = \mathrm{Gal}_\pi$. $\square$