Number Theory in Function Fields - Notes (by Evan Dummit, 2021, v. 1.00)

Contents

0 Number Theory in Function Fields 1

0.1 (Sep 9) Overview + Fermat's Last Theorem for Polynomials . . . 2

0.2 (Sep 13) Quotients of $\mathbb{F}_q[t]$ + Prime-Counting Part 1 . . . 4

0.3 (Sep 16) Prime-Counting Part 2 + The Zeta Function . . . 8

0.4 (Sep 20) Dirichlet Series + Multiplicative Functions . . . 11

0.5 (Sep 23) dth Powers and the Reciprocity Law . . . 14

0.6 (Sep 27) Generalizations + Applications of the Reciprocity Law . . . 17

0.7 (Sep 30) Group Characters + Primes in Arithmetic Progressions Part 1 . . . 21

0.8 (Oct 4) Primes in Arithmetic Progressions, Part 2 . . . 25

0.9 (Oct 7) Homework #1 Discussion . . . 28

0.10 (Oct 14) Preliminaries: Transcendence and Localization . . . 28

0.11 (Oct 18) Discrete Valuations and Divisors in Function Fields . . . 33

0.12 (Oct 21) The Riemann-Roch Theorem and Applications . . . 38

0.13 (Oct 25) Riemann-Roch Over $\mathbb{C}$ + Counting Divisors . . . 41

0.14 (Oct 28) Zeta Functions and the Weil Conjectures . . . 46

0.15 (Nov 1) Affine Varieties + Hilbert's Nullstellensatz . . . 50

0.16 (Nov 4) Affine Varieties, Curves, and Projective Space . . . 53

0.17 (Nov 8) Homework #2 Discussion . . . 57

0.18 (Nov 15) Projective Space and Projective Curves . . . 57

0.19 (Nov 18) Rational Maps and Extensions of Function Fields . . . 62

0.20 (Nov 22) Primes in Extensions . . . 67

0.21 (Nov 29) Ramification and the Riemann-Hurwitz Theorem . . . 71

0.22 (Dec 2) The abc Conjecture and the S-Unit Equation . . . 74

0.23 (Dec 6) Proof of the Riemann Hypothesis . . . 78

0.24 (Dec 9) Homework #3 Discussion . . . 82

0 Number Theory in Function Fields

These are lecture notes for the graduate course Math 7315: Number Theory in Function Fields, taught at Northeastern in Fall 2021.

0.1 (Sep 9) Overview + Fermat's Last Theorem for Polynomials

- The goal of this course is to elucidate some of the many analogies between number theory in number fields and number theory in function fields.
  - Some things from classical number theory: primes, factorizations, congruences and modular arithmetic, Fermat's and Euler's theorems, the prime number theorem, quadratic reciprocity (and higher reciprocity), Dirichlet's theorem on primes in arithmetic progressions, zeta functions.
  - Some things from the more modern take on algebraic and analytic number theory: algebraic number fields and their rings of integers, Galois theory and its interplay with number fields, discriminants, class groups, Dirichlet's unit theorem, cyclotomic fields, ramification, L-functions, the Riemann hypothesis.
  - Our goal is to do as much of these things as possible in the context of function fields, where many of the results are more approachable, because the function-field setting has a major kit of additional tools (namely, algebraic geometry).
  - Though do note: number theory in function fields is a beautiful subject in its own right, and not just because it has so many similarities to algebraic number theory.
- To start, let $q = p^f$ be a prime power, and let $\mathbb{F}_q$ be the finite field with q elements. The story begins with the polynomial ring $A = \mathbb{F}_q[t]$.
  - We have the degree map on A: explicitly, for coefficients $a_i \in \mathbb{F}_q$ and an element $f = a_0 + a_1 t + \cdots + a_n t^n$ with $a_n \neq 0$, we define $\deg(f) = n$ and $\text{sgn}(f) = a_n$. (We also set $\deg(0) = -\infty$ and $\text{sgn}(0) = 0$.)
  - Exercises (trivial): $\deg(fg) = \deg(f) + \deg(g)$, $\text{sgn}(fg) = \text{sgn}(f)\text{sgn}(g)$, and $\deg(f+g) \leq \max(\deg f, \deg g)$ with equality whenever $\deg f \neq \deg g$.
  - The polynomials with sign 1 (i.e., monic polynomials) behave analogously to the integers with positive sign (i.e., the positive integers).
  - We also note that the degree properties easily give a characterization of the units of A: they are the nonzero constant polynomials.
- Our first basic result is the standard division-with-remainder algorithm for polynomials, which we record over arbitrary fields for no extra cost:
- Proposition (Polynomial Division): If F is any field, then for any $f, g \in F[t]$ with $g \neq 0$, there exist unique $q, r \in F[t]$ such that $f = qg + r$ and $\deg r < \deg g$.
  - The idea is simply to prove that the usual long-division algorithm works.
  - Proof: For existence, fix g and induct on $\deg f$. If $\deg f < \deg g$ then just take $q = 0$ and $r = f$.
  - For the inductive step, suppose the result holds for polynomials of degree less than $n \geq \deg g$.
  - Write $f = a_n t^n + \cdots + a_0$ and $g = b_m t^m + \cdots + b_0$, where $b_m \neq 0$ since $g \neq 0$.
  - Then $f^\dagger = f - \dfrac{a_n}{b_m} t^{n-m} g$ has degree less than n, since we have cancelled the leading term of f. (Note that here is where we are using the fact that F is a field, so that $\dfrac{a_n}{b_m}$ also lies in F.)
  - By induction $f^\dagger = q^\dagger g + r^\dagger$ for some $q^\dagger$, $r^\dagger$ with $\deg(r^\dagger) < \deg(b)$.
  - Then $f = \left[q^\dagger + \dfrac{a_n}{b_m} t^{n-m}\right] g + r^\dagger(x)$, so we can take $q = q^\dagger + \dfrac{a_n}{b_m} t^{n-m}$ and $r = r^\dagger$.
  - For uniqueness, if we have $f = qg + r = q'g + r'$, then $r - r' = (q' - q)g$ has degree less than $\deg g$ but is also divisible by g, hence must be zero.
- So, $F[x]$ is a Euclidean domain, meaning that it is also a PID (all ideals are principal) and a UFD (every element can be factored uniquely into a product of irreducibles up to reordering and unit factors).
- As it turns out, unique factorization is essentially all we need to prove Fermat's Last Theorem for polynomials.
  - We would like to show that the equation $f^n + g^n = h^n$ has no nontrivial solutions in polynomials f, g, h. Aside from the case $n = 4$, it is enough to treat the situation where n is a prime.

- But we do need to be a little bit careful to write down exactly what the trivial solutions look like, beyond the obvious ones where one of f, g, h is zero.
- For example, if f, g, h are all constants, we can certainly have lots of solutions to $f^n + g^n = h^n$, depending on the field and on n (e.g., $1^5 + 1^5 = 2^5$ inside $\mathbb{F}_3$).
- We need to avoid the situation where n is divisible by $p = \text{char}(\mathbb{F}_q)$, since $f^p + g^p = (f+g)^p$ for any polynomials $f, g \in \mathbb{F}_q[t]$.
- Also, since the equation is homogeneous, we can scale solutions to get new solutions.
- To avoid all of these situations, we can consider only the case where f, g, h are relatively prime (since if they are not, then any common divisor of two of them also divides the third, so we could cancel it) and where the exponent n is not divisible by the characteristic p.

• Theorem (FLT for Polynomials): Suppose that $f, g, h \in F[t]$ are pairwise relatively prime and that $p \geq 3$ is prime with $p \neq \text{char}(F)$. Then the only solutions to $f^p + g^p = h^p$ are when f, g, h are all constants.

- We will remark that $p \geq 3$ is needed, since the usual parametrization of Pythagorean triples also works for polynomials: if we take $f = a^2 - b^2$, $g = 2ab$, $h = a^2 + b^2$ for any polynomials $a, b \in F[t]$, then $f^2 + g^2 = h^2$.

• We will give two different proofs: the first uses a classical-style infinite descent argument, while the second uses a more function-field type of argument.

- Proof 1: Without loss of generality, we may assume that F is algebraically closed, since any solution to $f^p + g^p = h^p$ over F is still a solution over the algebraic closure $\overline{F}$.
- We show the result by inducting on $d = \deg f + \deg g$. The base case $d = 0$ is trivial, since there is nothing to prove. So now suppose we have a solution with $d > 0$.
- By the assumption that $p \neq \text{char}(F)$, there are p distinct pth roots of unity in F: say, $1, \zeta_p, \zeta_p^2, \dots, \zeta_p^{p-1}$, and we can factor $f^p + g^p = (f+g)(f+\zeta_p g)(f+\zeta_p^2 g)\cdots(f+\zeta_p^{p-1} g)$.
- We then note that all of the terms $f + \zeta_p^i g$ are relatively prime: if d divides both $f + \zeta_p^i g$ and $f + \zeta_p^j g$, then d also divides the difference $(\zeta_p^i - \zeta_p^j) g$ hence divides g, hence also divides $(f + \zeta_p^i g) - \zeta_p^i g = f$, but f and g are relatively prime by assumption.
- Then by unique factorization inside $F[t]$, since all of the terms in the product $(f+g)(f+\zeta_p g)(f+\zeta_p^2 g)\cdots(f+\zeta_p^{p-1} g)$ are relatively prime and their product is a pth power (namely, h^p), each term must be a pth power up to a unit factor. But since F is algebraically closed, everything in F has a pth root in F, so the unit factor is also a pth power.
- Thus, in particular, we see that $f + g = a^p$, $f + \zeta_p g = b^p$, and $f + \zeta_p^2 g = c^p$ are all pth powers.
- Using basic linear algebra to eliminate f and g yields the relation $-\zeta_p a^p + (1+\zeta_p) b^p = c^p$, so if we set $a' = (-\zeta_p)^{1/p} a$, $b' = (1+\zeta_p)^{1/p} b$, and $c' = c$, then we have $(a')^p + (b')^p = (c')^p$.
- Note that a', b' cannot both be constant, since then f, g would have been constant. But we also have $\deg(a') + \deg(b') = \deg(f+g)/p + \deg(f+\zeta_p g)/p \leq 2\max(\deg f, \deg g)/p < d$, so we have constructed a solution with smaller positive degree, but this contradicts the induction hypothesis. Therefore, there are no nonconstant solutions.

• Before giving the second proof, we need a few preliminary results.

- First, if f has prime factorization $f = \prod_i p_i^{a_i}$, define $\text{rad}(f) = \prod_i p_i$, the product of the monic irreducible polynomials dividing f.

• Lemma: We have $\deg \gcd(f, f') \geq \deg f - \deg \text{rad} f$, where f' is the derivative of f.

- Proof: Suppose $f = p^a q$ where p is irreducible and doesn't divide q. Then $f' = ap^{a-1}p'q + p^a q' = p^{a-1}(ap'q + pq')$ is divisible by p^{a-1}. Therefore, $\gcd(f, f')$ is divisible by p^{a-1}.
- Taking the product over all primes dividing f shows that $\prod_i p_i^{a_i - 1}$ divides $\gcd(f, f')$, so $\gcd(f, f') \cdot \text{rad}(f)$ is divisible by $\prod_i p_i^{a_i-1} \prod_i p_i = \prod_i p_i^{a_i} = f$, so taking degrees yields the inequality.
- Exercise: determine when equality holds in the Lemma.

- Next, we show a result due independently to Mason and Stothers:

- Proposition (Mason-Stothers): Suppose that $f, g, h \in F[t]$ are nonconstant, relatively prime, that $f + g = h$, and that not all of f', g', h' are zero. Then $\max(\deg f, \deg g, \deg h) \le \deg \operatorname{rad}(fgh) - 1$.

  - Proof: If $f + g = h$ then $f' + g' = h'$, and then $fg' - f'g = (f + g)g' - (f' + g')g = hg' - h'g$.
  - Note also that $fg' - f'g$ is nonzero: $fg' = f'g$ yields $f'/f = g'/g$ and then integrat1ing yields $\ln(f) = \ln(g) + C$ yielding $f = Dg$ (as long as f', g' are not zero), contradicting the assumption that f, g are relatively prime.
  - Alternatively, if $fg' = f'g$ then f must divide $f'g$ hence that f must divide f' since f, g are relatively prime. It is not hard to see that $f | f'$ is equivalent to saying $f' = 0$, but by the same argument we would also have $g' = 0$ and $h' = 0$, contradicting the assumption above.
  - Now let $d_f = \gcd(f, f')$, $d_g = \gcd(g, g')$, $d_h = \gcd(h, h')$. Then d_f, d_g, d_h all divide $fg' - f'g = hg' - h'g$, and they are all relatively prime since they are divisors of the relatively prime polynomials f, g, h.
  - This means $d_f d_g d_h$ divides $fg' - f'g$, so taking degrees yields $\deg(d_f d_g d_h) \le \deg(fg' - f'g) \le \deg(f) + \deg(g) - 1$.
  - By the Lemma, we have $\deg(d_f) \ge \deg(f) - \deg \operatorname{rad} f$, $\deg(d_g) \ge \deg(g) - \deg \operatorname{rad} g$, $\deg(d_h) \ge \deg(h) - \deg \operatorname{rad} h$, so summing yields $\deg(f) + \deg(h) + \deg(h) - \deg \operatorname{rad}(fgh) \le \deg(d_f d_g d_h) \le \deg(f) + \deg(g) - 1$, and therefore $\deg(h) \le \deg \operatorname{rad}(fgh) - 1$.
  - By rearranging we obtain the same bounds on $\deg(f)$ and $\deg(g)$, and so we are done.

- At last, we can finish the second proof of Fermat's Last Theorem for polynomials:

  - Proof 2: Suppose $f^p + g^p = h^p$. By the assumption on the characteristic, we have $(f^p)'$, $(g^p)'$, $(h^p)'$ not all zero.
  - Then by Mason-Stothers, we see $\max(\deg f^p, \deg g^p, \deg h^p) \le \deg \operatorname{rad}(f^p g^p h^p) - 1$, which is equivalent to $p \cdot \max(\deg f, \deg g, \deg h) \le \deg \operatorname{rad}(fgh) - 1 \le \deg f + \deg g + \deg h - 1$ since the radical ignores powers.
  - Now apply the simple observation that $\max(a, b, c) \ge (a + b + c)/3$ and set $d = \deg f + \deg g + \deg h$ to see that $p \cdot d/3 \le d - 1$, which is impossible, since $d \le p \cdot d/3$ by the hypothesis that $p \ge 3$.

0.2 (Sep 13) Quotients of $\mathbb{F}_q[t]$ + Prime-Counting Part 1

- We now return to study the structure of quotient rings of $A = \mathbb{F}_q[t]$, which (re-posed) is simply studying modular arithmetic in this ring.

  - In particular, we will recover almost identical versions of Fermat's little theorem, Euler's theorem, and Wilson's theorem.
  - We will also take some time to look at the structure of the unit group of A/gA, which turns out to be a bit more complicated to write down than the unit group of $\mathbb{Z}/m\mathbb{Z}$.

- As noted last lecture, A is a Euclidean domain, so it is a PID and also a UFD. Since every ideal is principal, if we want to understand the structure of the quotient rings of A, we only have the quotients of the form A/gA to consider.

  - We can also assume A is monic by replacing it with its unique monic associate, which does not change the quotient ring A/gA.

- Using the division algorithm, we can write down the residue classes in A/gA (and thereby compute its cardinality) quite nicely:

- Proposition: Let $g \in \mathbb{F}_q[t]$ be nonzero. Then the residue classes in A/gA are uniquely represented by the polynomials of degree less than $\deg(g)$. In particular, $\#(A/gA) = q^{\deg g}$.

- Proof: If $f \in \mathbb{F}_q[t]$ is any polynomial, then by the division algorithm we can write $f = qg + r$, and so inside A/gA we see $\overline{f} = \overline{r}$. So the possible remainders give a complete set of residue class representatives – but by the uniqueness of the quotient and remainder, no two remainders are equivalent mod g, so in fact they give all of the residue classes exactly once.
- For the counting, if $\deg(g) = n$, then the remainders are of the form $c_0 + c_1 t + \cdots + c_{n-1} t^{n-1}$ with $c_i \in \mathbb{F}_q$. Since there are n coefficients each of which has q possible values, there are $q^n = q^{\deg g}$ possible ways to select a remainder.

- The size of the quotient ring gives a convenient way of measuring the "size" of a polynomial that behaves pleasantly under multiplication:

- Definition: For $g \in \mathbb{F}_q[t]$, we define $|g|$, the norm of g, to be $q^{\deg g}$. By the calculation above, $|g| = \#(A/gA)$ when $g \neq 0$.
  - Exercises (easy): $|fg| = |f| \cdot |g|$, $|f + g| \leq \max(|f|, |g|)$ with equality whenever $|f| \neq |g|$.

- Our next goal is to understand the units of A/gA, since this is the context in which to pose Fermat's and Euler's theorems.
  - Regardless of the polynomial g, the units of A/gA will contain an isomorphic copy of the constant polynomials (i.e., the units of A), which is the multiplicative group $\mathbb{F}_q^*$.
  - As is well-known, the multiplicative group of a finite field is cyclic. We record a few proofs of this fact, for completeness:

- Proposition (Multiplicative Group of $\mathbb{F}_q$): If G is a finite multiplicative subgroup of a field F, then G is cyclic.
  - All known proofs of this fact are essentially nonconstructive, to varying degrees: there does not seem to be a nice algorithm for writing down a multiplicative generator of a finite field that is appreciably better than a brute-force search.
  - Proof 1: Let G be a finite multiplicative subgroup of F. By the fundamental theorem of finite(ly generated) abelian groups, G is isomorphic to a direct product of cyclic groups.
  - Let m be the lcm of the orders of these cyclic groups: then $x^m = 1$ for all $x \in G$. Since $F[t]$ has unique factorization, the polynomial $t^m - 1 \in F[t]$ has at most m roots in F, so $\#G \leq m$. On the other hand, by Lagrange's theorem, the order of every element in G divides $\#G$, so m divides $\#G$. We must therefore have $m = \#G$.
  - But since $\#G$ is equal to the product of the orders of the cyclic groups, we see that the product of these orders equals their lcm, so the orders are all relatively prime. This means G is cyclic, as claimed.
  - Proof 2: Let M be the maximal order among all elements in G: we claim that the order of every element in G divides M. To see this, suppose g has order M, and let h be any other element of order k. If k does not divide M, then there is some prime q which occurs to a higher power q^f in the factorization of k than the corresponding power q^e dividing M.
  - By properties of orders, the element g^{q^f} has order M/q^f, and the element h^{k/q^e} has order q^e. Since these two orders are relatively prime and $gh = hg$ (since these are elements in a field), we see that the element $g^{q^f} \cdot h^{k/q^e}$ has order $M \cdot q^{f-e}$. This is a contradiction because this element's order is larger than M. Thus, k divides M as claimed.
  - For the second claim, any element of order M generates a subgroup of G having M elements, so $M \leq \#G$.
  - Furthermore, by the above, we know that all elements in G have order dividing M, so the polynomial $t^M - 1$ has $\#G$ roots in $F[t]$. By unique factorization, this requires $M \geq \#G$, and so we have $M = \#G$. Now select any element of order M: it generates G.
  - Proof 3: Observe by Lagrange's theorem that $t^{\#G} - 1$ factors as the product $\prod_{d | \#G} \Phi_d(t)$, where $\Phi_d(t) = \prod_{\mathrm{order(g)=d}} (t - g)$ is the dth cyclotomic polynomial. By an inductive argument, or by observing invariance under the Galois action, all of the polynomials $\Phi_d(t)$ have coefficients in $F[t]$. By induction on d using the fact that $t^d - 1$ has at most (hence exactly) d roots in F and in G, one has that $\deg(\Phi_d) = \varphi(d)$. In particular, $\deg(\Phi_{\#G}) = \varphi(\#G) > 0$, so there is an element of order $\#G$ in G.

- Now we tackle the question of the units of A/gA.
  - We can simplify the problem first: if we factor $g = p_1^{a_1} \cdots p_d^{a_d}$ where the p_i are distinct monic irreducible polynomials, then all of the ideals $(p_i^{a_i})$ are pairwise comaximal, so by the Chinese remainder theorem, we see $A/gA \cong (A/p_1^{a_1}A) \times (A/p_2^{a_2}A) \times \cdots \times (A/p_d^{a_d}A)$.
  - Taking units on both sides then gives $(A/gA)^* \cong (A/p_1^{a_1}A)^* \times (A/p_2^{a_2}A)^* \times \cdots \times (A/p_d^{a_d}A)^*$. So it is enough to study the structure of the ring A/p^aA where p is irreducible.
- Proposition (Structure of A/p^aA): For $A = \mathbb{F}_q[t]$ and p a monic irreducible polynomial, we have the following:
  1. The cardinality of $(A/p^aA)^*$ is $\#(A/p^aA)^* = |p|^{a-1}(|p|-1) = |p^a|(1 - 1/|p|)$.
     - Exercise: A commutative ring R with 1 has a unique maximal ideal M if and only if the set of nonunits in R forms an ideal (which is then a unique maximal ideal M). Note that a ring with this property is called a local ring.
     - Proof: The ring A/p^aA has a unique maximal ideal, namely pA/p^aA, and is therefore a local ring, because the quotient $(A/p^aA)/(pA/p^aA) \cong A/pA$ is a field by the third isomorphism theorem.
     - By the exercise above, evvery element not in the maximal ideal is a unit, and the cardinality of the maximal ideal is $1/|p|$ times the cardinality of the entire ring (since the elements in the ideal are just the multiples of p). The formula follows.
  2. $(A/p^aA)^* \cong$ [cyclic group of order $|p|-1$] $\times$ [an abelian $\tilde{p}$-group].
     - Proof: The reduction-mod-p map is a surjective group homomorphism from $(A/p^aA)^* \to (A/pA)^*$, and the latter is the multiplicative group of the field A/pA hence is cyclic of order $|p|-1$.
     - Pulling back a generator yields that $(A/p^aA)^*$ contains a cyclic subgroup of order $|p|-1$. By the cardinality calculation in (1), the remaining piece has order $|p|^{a-1}$ and is therefore a $\tilde{p}$-group (and it is clearly abelian).
     - Remark: the direct product decomposition writes each element modulo p^a as [its residue modulo p] times [an element congruent to 1 modulo p].
  3. The $\tilde{p}$-part of $(A/p^aA)^*$ has exponent at most $\tilde{p}^s$ where $\tilde{p}^s \geq a$.
     - Proof: By the above, the elements in the $\tilde{p}$-part are of the form $1 + bp$ for some $b \in \mathbb{F}_q[t]$.
     - Since we are in characteristic $\tilde{p}$, we then have $(1+bp)^{\tilde{p}^s} = 1 + (bp)^{\tilde{p}^s}$, and since $p^{\tilde{p}^s}$ is divisible by p^a by assumption, we see $(1+bp)^{\tilde{p}^s} \equiv 1 \pmod{p^a}$, which is to say, the element $1+bp$ modulo p^a has order dividing $\tilde{p}^s$ (as required).
  4. As $a \to \infty$, the number of cyclic factors in the $\tilde{p}$-part of $(A/p^aA)^*$ goes to infinity.
     - The point here is that we get a different kind of behavior than over $\mathbb{Z}$: over $\mathbb{Z}$, we see that $(\mathbb{Z}/p^a\mathbb{Z}) \cong \begin{cases} \mathbb{Z}/(p^a - p^{a-1})\mathbb{Z} & \text{for odd primes } p \\ (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2^{a-3}\mathbb{Z}) & \text{for } p = 2 \end{cases}$, and so even for large prime powers, the quotient is either cyclic or basically cyclic.
     - For polynomials, we end up getting a large number of cyclic factors when we take a large power, regardless of the prime.
     - Proof: Since the exponent of the $\tilde{p}$-part is at most $\tilde{p}^s$, if we have a total of j cyclic factors then the order of the group is at most $\tilde{p}^{sj}$. So we need $\tilde{p}^{sj} \geq |p|^{a-1} = q^{\deg(p)\cdot(a-1)} = \tilde{p}^{f\cdot\deg(p)\cdot(a-1)}$ and so $j \geq f \cdot \deg(p) \cdot (a-1)/s$.
     - Since $s \sim \log_p a$, we see that for a fixed field $\mathbb{F}_q$ (i.e., fixed f) and fixed prime p (i.e., fixed $\deg p$), we have $j \sim C(a-1)/\log_p a \to \infty$ as $a \to \infty$.
- Now that we have established some basic things about the unit group of A/p^aA, we can establish the analogues of Fermat's little theorem, Euler's theorem, and Wilson's theorem.
  - First, we need the analogue of the Euler phi-function. We define $\Phi(f) = \#(A/fA)^*$ to be the number of polynomials of degree less than $\deg f$ that are relatively prime to f.
  - By our calculations with the unit group earlier, we have the usual formula $\Phi(f) = |f| \prod_{p|f\text{prime}}(1 - 1/|p|)$, which is the analogue of $\varphi(n) = n \prod_{p|n \text{ prime}}(1 - 1/p)$ for the phi-function over $\mathbb{Z}$.

- Proposition ("Euler"): If $f \in \mathbb{F}_q[t]$ is nonzero and g is relatively prime to f, then $g^{\Phi(f)} \equiv 1 \pmod{f}$.
  - Proof 1: Apply Lagrange's theorem to $\overline{g}$ in $(A/fA)^*$.
  - Proof 2: Multiplication by $\overline{g}$ is a bijection on the cosets in $(A/fA)^*$. Thus, $\prod_{u \in (A/fA)^*} u = \prod_{u \in (A/fA)^*}(ug) = g^{\Phi(f)} \prod_{u \in (A/fA)^*} u$ inside $(A/fA)^*$, and cancelling the unit factor $\prod_{u \in (A/fA)^*} u$ yields $g^{\Phi(f)} = 1$ inside $(A/fA)^*$.
- Proposition ("Fermat"): If $p \in \mathbb{F}_q[t]$ is irreducible, then $a^{|p|} \equiv a \pmod{p}$ for any $a \in \mathbb{F}_q[t]$.
  - Proof: If $p|a$ the result is trivial. Otherwise, a is a unit modulo p and the result follows from Euler above.
- We can use Fermat's to prove an analogue of Wilson's theorem:
- Proposition (Factoring, 1): Suppose If $p \in \mathbb{F}_q[t]$ is irreducible of degree d. Then $x^{|p|} - x \equiv \prod_{\deg f < d}(x - f)$ mod p.
  - Proof: As we have noted, in A/p the polynomials of degree $< d$ represent all of the residue classes modulo p.
  - By Fermat, each of these polynomials is a root of $x^{|p|} - x$. But by unique factorization, this polynomial has at most $|p|$ distinct roots, and we have just exhibited $|p|$ roots, so these are all of the roots, and the factorization follows.
- Corollary ("Wilson"): If $p \in \mathbb{F}_q[t]$ is irreducible, then $\prod_{\deg f < d, f \neq 0} f \equiv -1 \pmod{p}$.
  - Proof 1: Dividing the result above by x yields $x^{|p|-1} - 1 \equiv \prod_{\deg f < d, f \neq 0}(x - f)$ mod p.
  - Now set $x = 0$: if the characteristic is odd, then the number of minus signs on the RHS is even and the result follows, while if the characteristic is even, then $1 = -1$ so the result still follows.
  - Proof 2: If $\overline{f}$ does not have order 2 in A/pA, then $\overline{f} \neq \overline{f}^{-1}$ and so we can pair up and discard $(\overline{f}, \overline{f}^{-1})$ without affecting the product.
  - When we have done this for all possible pairs, the only elements left are the elements of order dividing 2 (i.e., the solutions to $x^2 = 1$), which are $x = \pm 1$. In characteristic not 2, the product is -1, while in characteristic 2, the product is $1 = -1$.
  - Exercise: Generalize proof 2 to show that if G is a finite abelian group, then the product of all elements in g is the unique element in G of order 2 (if there is one), or is otherwise 1.
- We also record a useful result about roots of unity:
- Proposition (Roots of Unity): If d divides $|p| - 1$, then there are d dth roots of unity in A/pA: in other words, $x^d \equiv 1 \pmod{p}$ has d solutions.
  - Exercise: For positive integers a, b, $\gcd(x^a - 1, x^b - 1) = x^{\gcd(a,b)} - 1$.
  - Proof: Note that $x^{|p|-1} - 1$ splits completely mod p as shown above. By the exercise, $x^d - 1$ divides $x^{|p|-1} - 1$ when d divides $|p| - 1$, and so $x^d - 1$ also splits completely, which is to say, it has d roots mod p.
  - Exercise: Prove the converse of this proposition: if there are d dth roots of unity in A/pA, then d divides $|p| - 1$.
- Now that we have established most of the classical results for modular arithmetic, we move to our next item: counting primes.
  - We will do things in a more ad hoc manner first, and then give a more general approach using zeta functions that will allow us to go further.
- Our first step is to write down a generalization of the fact we used to establish Wilson's theorem above:
- Theorem (Factoring, II): For a positive integer m, the polynomial $t^{q^m} - t$ factors over $\mathbb{F}_q$ as the product of all monic irreducible polynomials of degree dividing m.

- Proof 1 ("Elementary"): We will show that $t^{q^m} - t$ has no repeated factors, that each of the claimed polynomials does divide it, and that no other polynomials divide it.
- Exercise (easy): A polynomial in $F[x]$ has no repeated factors if and only if it is relatively prime to its derivative.
- Since $(t^{q^m} - t)' = q^m t^{q^m-1} - 1 = -1$ in characteristic p, the polynomial is relatively prime to its derivative, so it has no repeated factors by the exercise.
- Exercise: For positive integers q, a, b, $\gcd(q^a - 1, q^b - 1) = q^{\gcd(a,b)} - 1$.
- Next, suppose p is irreducible of degree dividing m. If $p = t$ the result is trivial, and otherwise, in A/p we have $t^{q^m-1} \equiv 1 \bmod p$ because $q^m - 1$ is a multiple of $|p| - 1 = q^{\deg p} - 1$ by the exercise above along with Euler's theorem. This means $t^{q^m-1} - 1$ is divisible by p as required.
- Finally, suppose p is irreducible of degree not dividing m. Then in A/p we have $t^{q^m-1} \equiv t^{q^{\gcd(m,\deg p)}} \neq 1 \bmod p$ by the exercise above along with Euler's theorem and the fact that $q^{\gcd(m,\deg p)} < q^{\deg p}$. This means $t^{q^m-1} - 1$ is not divisible by p as required.
- We have shown that $t^{q^m} - t$ has no repeated factors, that each of the claimed polynomials does divide it, and that no other polynomials divide it. Since the polynomial is monic, its factorization must therefore be as claimed.

0.3 (Sep 16) Prime-Counting Part 2 + The Zeta Function

- We now give another proof of the counting result from last time.
  - Proof 2 ("Galois"): By basic Galois theory, $\mathrm{Gal}(\mathbb{F}_{q^m}/\mathbb{F}_q)$ is a cyclic group of order m generated by the Frobenius map $x \mapsto x^q$[1].
  - By the Galois correspondence, the intermediate fields of $\mathbb{F}_{q^m}/\mathbb{F}_q$ are $\mathbb{F}_{q^d}$ for $d|m$. Therefore, p is irreducible of degree dividing $d \iff \mathbb{F}_q[t]/(p)$ is (isomorphic to) an intermediate field of $\mathbb{F}_{q^m}/\mathbb{F}_q \iff p$ divides $x^{q^m} - x$.
  - Since $x^{q^m} - x$ is separable, its factorization must therefore be as claimed.
- Corollary: If a_d is the number of irreducible monic polynomials in $A = \mathbb{F}_q[t]$ of degree d, then $\sum_{d|n} da_d = q^n$.
  - Proof: Count degrees in the theorem above.
- We can use this recurrence to write down an exact formula for a_d using Mobius inversion.
- Definition: The Mobius μ-function is defined as $\mu(n) = \begin{cases} 0 & \text{if } n \text{is not squarefree} \\ (-1)^r & \text{if } n \text{is the product of } r \text{distinct primes} \end{cases}$. Note $\mu(1) = 1$.
  - Exercise: Show that $\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases}$.
- Proposition (Mobius Inversion): If f, n are integer functions such that $g(n) = \sum_{d|n} f(d)$, then $f(n) = \sum_{d|n} \mu(d) g(n/d)$.
  - Proof: Induct on n. The base case $n = 1$ is trivial.
  - For the inductive step, we have $\sum_{d|n} \mu(d) g(n/d) = \sum_{d|n} \mu(d) \cdot \sum_{d'|n/d} f(d') = \sum_{dd'|n} \mu(d) f(d') = \sum_{d'|n} f(d') \sum_{d|(n/d')} \mu(d) = f(n)$ because the last inner sum is zero except for when $n/d' = 1$.
- By using Mobius inversion to the sequence $\{da_d\}$, we can write down formulas for the number of monic irreducible polynomials of degree d.

[1]This follows by noting that $\mathbb{F}_{q^m}$ is the splitting field of $x^{q^m} - x$ over $\mathbb{F}_q$ and since this polynomial is separable as noted in proof 1, the order of the Galois group is m. The Frobenius map is an injective field map from $\mathbb{F}_{q^m}$ to itself, hence an automorphism by finiteness, and its order is clearly at least m (since $x^{q^d} - x$ has at most q^d solutions) and at most m (by Lagrange).

- Proposition (Prime Counting): If a_n is the number of monic irreducible polynomials in $\mathbb{F}_q[t]$ of degree n, then $a_n = \frac{1}{n}\sum_{d|n}\mu(d)q^{n/d}$.
    - The first few values are $a_1 = q$, $a_2 = \frac{1}{2}(q^2-q)$, $a_3 = \frac{1}{3}(q^3-q)$, $a_4 = \frac{1}{4}(q^4-q^2)$, $a_5 = \frac{1}{5}(q^5-q)$, $a_6 = \frac{1}{6}(q^6-q^3-q^2+q)$, ...
    - Proof: Immediate from applying Mobius inversion to the sequence $\{na_n\}$.
- We can also do some basic asymptotic analysis using the formula above.
    - The main term is $\frac{1}{n}q^n$, and then the next biggest possible term is $\frac{1}{n}q^{n/2}$, so we see that $a_n = \frac{1}{n}q^n + O(q^{n/2}/n)$.
    - If we write $X = q^n$ (which is the total number of monic polynomials of degree n), we see that the number of "primes" in A of "size" $\sim X$ is $a_n = \frac{X}{\log_q X} + O(\frac{\sqrt{X}}{\log_q X})$.
    - This is quite in the spirit of the prime number theorem over $\mathbb{Z}$, which says that the number of primes $\leq X$ is $\Pi(X) = \frac{X}{\log X} + O(\frac{X}{(\log X)^2})$. If we replace $X/\log X$ with the logarithmic integral $\text{li}(x) = \int_2^x \frac{dt}{\log t}$, then as shown by von Koch, the Riemann hypothesis is equivalent to the error estimate $\Pi(X) = \text{li}(x) + O(\sqrt{X}\log x)$.
    - Qualitatively, then, we have already obtained a prime-counting result that is closely analogous to the best possible one predicted by the Riemann hypothesis.
- Up until this point, our approach has been purely algebraic. However, by introducing analytic methods, we can give even easier solutions to these (and other) counting problems. The necessary object of study is the zeta function, which we now define:
- Definition: For $A = \mathbb{F}_q[t]$, the zeta function of A is $\zeta_A(s) = \sum_{f\in A \text{ monic}} \frac{1}{|f|^s}$ for $s \in \mathbb{C}$.
    - Compare to the definition of the Riemann zeta function $\zeta(s) = \sum_{n>0}\frac{1}{n^s}$ for $s \in \mathbb{C}$.
    - Unlike the Riemann zeta function, however, we can actually just evaluate the zeta function for A: since there are q^d monic polynomials of degree d, we see that $\sum_{\deg(f)\leq d \text{ monic}} \frac{1}{|f|^s} = 1 + \frac{q}{q^s} + \frac{q^2}{q^{2s}} + \cdots + \frac{q^d}{q^{ds}} = \frac{1-q^{(d+1)(1-s)}}{1-q^{1-s}}$, and so taking $d \to \infty$ we see that $\zeta_A(s) = \frac{1}{1-q^{1-s}}$ whenever $\text{Re}(s) > 1$ (to ensure convergence).
    - We have an obvious meromorphic continuation for $\zeta_A(s)$ to the complex plane (i.e., via the formula above), and it is clear that ζ is analytic everywhere except for a simple pole at $s = 1$.
    - Exercise: Show that the residue of $\zeta_A(s)$ at $s = 1$ (which is to say, the value of $\lim_{s\to 1}(s-1)\zeta_A(s)$) is $1/\log q$.
    - We also have a functional equation for $\zeta_A(s)$: if we set $\xi_A(s) = q^{-s}(1-q^{-s})^{-1}\zeta_A(s)$, then $\xi_A(s) = \xi_A(1-s)$.
    - Exercise: Do the algebra to establish the functional equation.
- We can also represent $\zeta_A(s)$ as an Euler product, just as with the Riemann zeta function.
    - Explicitly, by the uniqueness of prime factorization, we can formally write $\zeta_A(s) = \sum_{f\in A \text{ monic}} \frac{1}{|f|^s} = \prod_{p \text{monic irred}}(1 + \frac{1}{|p|^s} + \frac{1}{|p|^{2s}} + \cdots) = \prod_{p \text{monic irred}}(1 - 1/|p|^s)^{-1}$, and both sides are absolutely convergent for $\text{Re}(s) > 1$.
    - To prove this equality rigorously, we need to do some estimations on tails of the respective series, but since everything converges absolutely, this is not so difficult; we leave the precise details as an exercise.

- We can use the Euler product for the zeta function to obtain the same prime counts that we got earlier.

- Proposition (Prime Counting, Again): If a_d is the number of irreducible monic polynomials in $A = \mathbb{F}_q[t]$ of degree d, then $\sum_{d|n} da_d = q^n$, and so by Mobius inversion as before, we see $a_n = \frac{1}{n}\sum_{d|n}\mu(d)q^{n/d}$.
  - Proof: Group the terms in the Euler product together by degree: if $\deg p = d$ then $|p|^s = q^{ds}$.
  - Thus, since there are a_d monic irreducibles of degree d by definition, we see that $\zeta_A(s) = \prod_{p \text{monic irred}}(1 - 1/|p|^s)^{-1} = \prod_{d=1}^{\infty}(1-q^{-ds})^{-a_d}$.
  - Noting from earlier that $\zeta_A(s) = \frac{1}{1-q^{1-s}}$, if we substitute $u = q^{-s}$, we obtain the equality $\frac{1}{1-qu} = \prod_{d=1}^{\infty}(1-u^d)^{-a_d}$.
  - Taking the log-derivative of both sides yields $\frac{q}{1-qu} = \sum_{d=1}^{\infty}\frac{da_d u^{d-1}}{1-u^d}$. These expressions are equal as power series in u, and thus corresponding coefficients must also be equal.
  - The LHS is $\frac{q}{1-qu} = q\sum_{k=0}^{\infty}(qu)^k$ while the RHS is $\sum_{d=1}^{\infty} da_d u^{d-1}\sum_{l=0}^{\infty}u^{dl} = \sum_{d=1}^{\infty}\sum_{l=0}^{\infty} da_d u^{d(l+1)-1}$. So the coefficient of u^{n-1} on the LHS is $q \cdot q^{n-1} = q^n$, while the coefficient of u^{n-1} on the RHS is $\sum_{d(l+1)=n} da_d = \sum_{d|n} da_d$.
  - Thus, $q^n = \sum_{d|n} da_d$ as claimed.

- Of course, we have already proven this result by counting irreducible polynomials algebraically. However, this approach using the zeta function also extends to solve other counting problems quite conveniently.

- Proposition (Squarefree Counting): The number of monic squarefree polynomials of degree n over $\mathbb{F}_q[t]$ is equal to $b_n := q^n - q^{n-1}$. Equivalently, a randomly-chosen degree-n polynomial is squarefree with probability $1 - 1/q = 1/\zeta_A(2)$.
  - Compare this result to the corresponding fact about integers (which is a little harder to pose because we have to phrase it over a range): if α_n is the probability that a randomly-chosen integer in $[1, n]$ is squarefree, then $\lim_{n\to\infty}\alpha_n = 6/\pi^2 = 1/\zeta(2)$.
  - Proof: Consider the product $\pi = \prod_{p \text{ monic irred}}(1 + |p|^{-s})$.
  - By multiplying out the terms, we see that for $\text{Re}(s) > 1$, we have $\pi = \sum_{f \text{ monic}}\frac{\delta(f)}{|f|^s}$ where $\delta(f) = \begin{cases} 1 & \text{if } f \text{is squarefree} \\ 0 & \text{if } f \text{is not squarefree}\end{cases}$, since the denominators in the Euler product only include prime factors of exponents 0 and 1.
  - Now, since $1 + |p|^{-s} = \frac{1-|p|^{-2s}}{1-|p|^{-s}}$, taking the product over monic irreducibles and using the fact that the resulting numerator and denominator products converge absolutely allows us to write $\pi = \prod_{p \text{ monic irred}}\frac{1-|p|^{-2s}}{1-|p|^{-s}} = \frac{\prod_{p \text{ monic irred}} 1-|p|^{-2s}}{\prod_{p \text{ monic irred}} 1-|p|^{-s}} = \frac{\zeta_A(2s)}{\zeta_A(s)}$.
  - Setting $u = q^{-s}$ yields $\frac{1-qu^2}{1-qu} = \frac{\zeta_A(2s)}{\zeta_A(s)} = \pi = \sum_{f \text{ monic}}\frac{\delta(f)}{|f|^s} = \sum_{n=0}^{\infty} b_n u^n$.
  - But as a power series in u, we have $\frac{1-qu^2}{1-qu} = (1-qu^2)(1+qu+q^2u^2+\cdots)$, and so comparing coefficients yields $b_n = q^n - q^{n-1}$ as claimed.

- In a similar way, we can use the zeta function to write down formulas for the number of monic kth-powerfree polynomials of a given degree over $\mathbb{F}_q[t]$.
  - Specifically, these values are packaged as the coefficients in the Euler product $\prod_{p \text{ monic irred}}(1 + |p|^{-s} + |p|^{-2s} + \cdots + |p|^{-(k-1)s}) = \frac{\zeta_A(ks)}{\zeta_A(s)}$, and then by doing a calculation like the one above, one can write down an explicit formula.

- Exercise: Finish this calculation and give the actual formula for the number of cubefree polynomials of degree n.
- It is also worthwhile interpreting this Euler product calculation heuristically in terms of probabilities.
- Explicitly, we would expect (under suitable probability assumptions) that the probability of a given polynomial not being divisible by f is $(1 - 1/|f|)$.
- So, assuming independence (which can be made rigorous by appealing to the Chinese remainder theorem), the probability that a given polynomial is not divisible by any prime power p^k for all monic irreducible p is $\prod_{p \text{ monic irred}}(1 - 1/|p|^k) = 1/\zeta_A(k)$: this is why the 1/zeta factor shows up in the answer.

0.4 (Sep 20) Dirichlet Series + Multiplicative Functions

- Another classical object of study in elementary number theory over $\mathbb{Z}$ are arithmetic functions related to divisors, such as the Euler φ-function, the divisor-counting function, and the sum-of-divisors function.
  - All of these are examples of multiplicative functions, which have the property that $f(ab) = f(a)f(b)$ whenever a, b are relatively prime. (Note the infelicitous terminology: if $f(ab) = f(a)f(b)$ for all a, b, f is instead called completely multiplicative.)
  - In particular, if n has prime factorization $n = \prod_i p_i^{a_i}$ and f is multiplicative, then $f(n) = \prod_i f(p_i^{a_i})$.
  - We will briefly review some results about multiplicative functions in the classical setting, and then redo them in the function-field setting.
- It is a standard combinatorial principle that if we want to understand a function with domain $\mathbb{N}$, we should look at its generating function.
  - A natural first guess would be to use the standard power series $\sum_{n=0}^{\infty} f(n)x^n$.
  - However, this type of generating function is useful primarily for functions that behave additively. For number-theoretic functions, we instead want to use a Dirichlet series.
- Definition: If $h : \mathbb{N} \to \mathbb{C}$ is a complex-valued function defined on positive integers, then its associated Dirichlet series is $D_h(s) = \sum_{n=1}^{\infty} \dfrac{h(n)}{n^s}$.
  - Example: If $h(n) = 1$ for all n, then $D_h(s) = \zeta(s)$, the Riemann zeta function.
  - In order for this series to converge, we need h not to grow too fast. One may check that if $h(n) = O(n^\alpha)$ then $D_h(s)$ is absolutely convergent for $\mathrm{Re}(s) > 1 + \alpha$. (We will mostly ignore issues of convergence, since our functions will grow polynomially at worst, and so we may manipulate the series as if they were formal power series.)
  - If h is multiplicative, then it is a straightforward calculation to see that $D_h(s)$ has an Euler product expansion: $D_h(s) = \prod_{p \text{ prime}}(1 + \dfrac{h(p)}{p} + \dfrac{h(p^2)}{p^2} + \cdots)$, on the appropriate domain of convergence.
- The key property of Dirichlet series is that they reproduce desired behaviors under multiplication:
- Proposition (Dirichlet Multiplication): If $f, g : \mathbb{N} \to \mathbb{C}$ are functions, then $D_f(s) \cdot D_g(s) = D_{f*g}(s)$ where $f * g$ is the Dirichlet convolution defined via $(f * g)(n) = \sum_{d|n} f(d)g(n/d)$.
  - Proof: $D_f(s)D_g(s) = \sum_{a=1}^{\infty} \sum_{b=1}^{\infty} \dfrac{f(a)g(b)}{(ab)^s} = \sum_{n=1}^{\infty} \dfrac{1}{n^s} \sum_{ab=n} f(a)g(b) = \sum_{n=1}^{\infty} \dfrac{(f * g)(n)}{n^s} = D_{f*g}(s)$.
- The Dirichlet convolution, owing to the fact that it is merely multiplication of the underlying Dirichlet series, has various nice properties.
  - Exercise: Show that Dirichlet convolution is commutative and associative, and has an identity element given by $I(n) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases}$.

- Exercise: Show that f has an inverse under Dirichlet convolution if and only if $f(1) \neq 0$.
- Exercise: If $f(1) \neq 0$ and f is multiplicative, then its Dirichlet inverse f^{-1} is also multiplicative.
- Exercise: Show that if two of f, g, and $f * g$ are multiplicative, then the third is also.

- By exploiting Dirichlet convolution, we can find the Dirichlet series for many basic multiplicative functions in terms of the Riemann zeta function.

  - Recall $I(n) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases}$ and the Mobius function $\mu(n) = \begin{cases} 0 & \text{if } n \text{is not squarefree} \\ (-1)^r & \text{if } n \text{is the product of } r \text{distinct primes} \end{cases}$.
  - Also define $N(n) = n$ and $1(n) = 1$ (for all n).
  - Exercise: Show that $D_I(s) = 1$, $D_1(s) = \zeta(s)$, and $D_N(s) = \zeta(s-1)$.
  - First, we note that $\mu * 1 = I$, since $(\mu * 1)(n) = \sum_{d|n} \mu(d)1(n/d) = \sum_{d|n} \mu(d) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1 \end{cases}$ as noted in an exercise previously. Therefore, by multiplicativity of the Dirichlet series, we see that $D_\mu(s)D_1(s) = D_I(s)$, so that $D_\mu(s) = \dfrac{1}{\zeta(s)}$.
  - Exercise: Use $\mu * 1 = I$ to establish Mobius inversion: if $g(n) = \sum_{d|n} f(n)$ then $f(n) = \sum_{d|n} \mu(d)g(n/d)$.
  - Exercise: For the Euler φ-function, show that $\sum_{d|n} \varphi(d) = n$.
  - The previous exercise says that $\varphi * 1 = N$, and so by composing with μ and using associativity, we see that $\varphi = \mu * N$. Then we have $D_\varphi(s) = D_\mu(s)D_N(s) = \dfrac{\zeta(s-1)}{\zeta(s)}$.
  - In principle, we could have established this fornula for $D_\varphi(s)$ by manipulating the zeta function directly, but this method is both more difficult and requires knowing the actual (non-obvious) formula for the answer ahead of time.
  - We can also find the Dirichlet series for the divisor-counting function $d(n) = \#\{d \in \mathbb{N} : d|n\}$ quite easily by noting that $d(n) = \sum_{d|n} 1(d)1(d/n)$: this means $d = 1 * 1$, so $D_d(s) = D_1(s)^2 = \zeta(s)^2$.
  - Exercise: If σ is the sum-of-divisors function $\sigma(n) = \sum_{d|n} d$, show that $D_\sigma(s) = \zeta(s)\zeta(s-1)$.
  - Exercise: If σ_k is the sum-of-kth-powers-of-divisors function $\sigma_k(n) = \sum_{d|n} d^k$, find and prove a formula for $D_{\sigma_k}(s)$ in terms of the Riemann zeta function.

- One of the main applications of computing the Dirichlet series for these various arithmetic functions is that we can extract information about average growth rates from them.

  - In the classical case, obtaining average-growth results is moderately delicate, so we will instead just focus on the function-field case.

- Here are the function-field analogues of these classical multiplicative functions, which are now complex-valued functions on monic polynomials rather than positive integers:

  - The identity: $I(f) = \begin{cases} 1 & \text{for } f = 1 \\ 0 & \text{for } f \neq 1 \end{cases}$.
  - The norm: $N(f) = |f|$.
  - The Mobius μ-function: $\mu(f) = \begin{cases} 0 & \text{if } f \text{is not squarefree} \\ (-1)^r & \text{if } f \text{is the product of } r \text{distinct primes} \end{cases}$.
  - The Euler Φ-function: $\Phi(f) = \#(A/fA)^* = |f| \prod_{p|f \text{ prime}} (1 - 1/|p|)$.
  - The divisor-counting function: $d(f) = \#\{\text{monic } d|f\}$.
  - The sum-of-divisors function: $\sigma(f) = \sum_{d|f \text{ monic}} |d|$, or more generally the sum-of-kth-powers-of-divisors function $\sigma_k(f) = \sum_{d|f \text{ monic}} |d|^k$. (Note here that we take the norm of the divisors, since we want a $\mathbb{C}$-valued function.)

- It is easy to check that all of these functions are multiplicative, and to write down formulas for all of them in terms of the prime factorization of $f = p_1^{a_1} \cdots p_k^{a_k}$.
- Exercise: Verify that $d(f) = (a_1+1)\cdots(a_k+1)$ and $\sigma(f) = \dfrac{|p_1|^{a_1+1}-1}{|p_1|-1} \cdots \dfrac{|p_k|^{a_k+1}-1}{|p_k|-1}$.

- We have essentially the same definition for the Dirichlet series in the function-field case:
- Definition: If $h : \{\text{monics}\} \to \mathbb{C}$ is a complex-valued function defined on monic polynomials in $\mathbb{F}_q[t]$, then its associated Dirichlet series is $D_h(s) = \sum_{f \text{ monic}} \dfrac{h(f)}{|f|^s}$.
  - As before, we will mostly ignore issues of convergence, but just as in the classical case, one may check that if $h(f) = O(|f|^\alpha)$ then $D_h(s)$ converges absolutely for $\mathrm{Re}(s) > 1+\alpha$.
  - We also have the same Dirichlet convolution operator: if $g, h : \{\text{monics}\} \to \mathbb{C}$ are functions, then $D_g(s) \cdot D_h(s) = D_{g \star h}(s)$ where $(g * h)(f) = \sum_{d|f \text{ monic}} g(d)h(f/d)$.
  - Dirichlet convolution is commutative, associative, and has the identity element $I(f) = \begin{cases} 1 & \text{for } f = 1 \\ 0 & \text{for } f \neq 1 \end{cases}$.
  - All of the same formulas for our arithmetic functions in terms of the zeta function follow through just as before. Here, however, we can actually write out the expressions explicitly, since we have a formula $\zeta_A(s) = \dfrac{1}{1-q^{1-s}}$.
- Proposition (Some Dirichlet Series): For $u = q^{-s}$, we have the following formulas: $D_I(s) = 1$, $D_N(s) = \zeta_A(s-1) = \dfrac{1}{1-u}$, $D_1(s) = \zeta_A(s) = \dfrac{1}{1-qu}$, $D_\mu(s) = \dfrac{1}{\zeta_A(s)} = 1 - qu$, $D_\Phi(s) = \dfrac{\zeta_A(s-1)}{\zeta_A(s)} = \dfrac{1-qu}{1-q^2u}$, $D_d(s) = \zeta_A(s)^2 = \dfrac{1}{(1-qu)^2}$, and $D_\sigma(s) = \zeta_A(s)\zeta_A(s-1) = \dfrac{1}{(1-qu)(1-q^2u)}$.
  - Proof: Exercise.
- Using these formulas we can recover average-value results quite easily.
- Definition: If $h : \{\text{monics}\} \to \mathbb{C}$ is a function, the average value of h on degree-n polynomials is $\mathrm{Avg}_n(h) = \dfrac{1}{q^n} \sum_{\deg(f)=n \text{ monic}} h(f)$. If the limit $\lim_{n\to\infty} \mathrm{Avg}_n(h)$ exists, we call it the "average value" of h.
  - We can also easily average h on polynomials of degree $\leq n$: the desired sum is instead $\dfrac{1}{1+q+\cdots+q^n} \sum_{\deg(f)\leq n} h(f)$.
  - Exercise: Show that if $\lim_{n\to\infty} \mathrm{Avg}_n(h) = \alpha$, then $\lim_{n\to\infty} \dfrac{1}{1+q+\cdots+q^n} \sum_{\deg(f)\leq n} h(f) = \alpha$ as well, so it is irrelevant whether we average over degree exactly n or $\leq n$.
  - The nice result here is that we can read off the value of $\mathrm{Avg}_n(h)$ from the coefficients of the Dirichlet series for h: explicitly, we have $D_h(s) = \sum_{n=1}^\infty \dfrac{\sum_{\deg(f)=n} h(f)}{q^{ns}} = \sum_{n=1}^\infty \dfrac{q^n \mathrm{Avg}_n(h)}{q^{ns}} = \sum_{n=1}^\infty q^n \mathrm{Avg}_n(h) u^n$ for $u = q^{-s}$.
  - So we can calculate these averages simply expanding out the Dirichlet series calculated above as power series in $u = q^{-s}$ and then dividing by q^n.
  - For example, $D_\mu(s) = 1 - qu$, so the average value of μ is 1 on degree-0 polynomials, -1 on degree-1 polynomials, and 0 on higher-degree polynomials.
  - Similarly, $D_d(s) = \dfrac{1}{(1-qu)^2} = (1 + qu + q^2u^2 + \cdots)^2 = 1 + 2qu^2 + 3q^2u^3 + \cdots$, so the average value of d on degree-n polynomials is $n+1$.
  - Likewise, $D_\Phi(s) = \dfrac{1-qu}{1-q^2u} = (1-qu)(1 + q^2u + q^4u^2 + q^6u^3 + \cdots) = 1 + (q^2-q)u + (q^4-q^3)u^2 + \cdots$, so the average value of Φ on degree-n polynomials is $(q^{2n} - q^{2n-1})/q^n = q^n - q^{n-1}$.
  - Exercise: Show that the average value of σ on degree-n polynomials is $(q^{n+1}-1)/(q-1)$.

0.5 (Sep 23) dth Powers and the Reciprocity Law

- Our next task is to discuss the analogue of another famous result from elementary number theory: Gauss's celebrated law of quadratic reciprocity, along with its higher-order generalizations. A brief recap of the story over $\mathbb{Z}$:
  - If $a \in (\mathbb{Z}/p\mathbb{Z})^*$, we say a is a quadratic residue if $a \equiv b^2 \pmod p$ for some b, and otherwise we say a is a quadratic nonresidue.
  - Since the quadratic residues are simply the image of the squaring map on $(\mathbb{Z}/p\mathbb{Z})^*$, by the first isomorphism theorem there are $(p-1)/2$ of them. (One may also simply enumerate them as $1^2, 2^2, \dots, [(p-1)/2]^2$.)
  - The Legendre symbol $\left(\frac{a}{p}\right)$ is defined to be $+1$ on quadratic residues and -1 on quadratic nonresidues. By writing a as a power of the generator of $(\mathbb{Z}/p\mathbb{Z})^*$, one then obtains Euler's criterion: $a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) \pmod p$, from which one sees that the Legendre symbol is multiplicative. Equivalently, it is a group homomorphism from $(\mathbb{Z}/p\mathbb{Z})^*$ to $\{\pm 1\}$.
  - Exercise: Another group homomorphism from $(\mathbb{Z}/p\mathbb{Z})^*$ to $\{\pm 1\}$ is obtained by calculating the signature of the permutation associated to multiplication by a, as an element of the symmetric group S_{p-1}. Prove Zolotarev's lemma: this homomorphism is the same as the Legendre symbol.

- The law of quadratic reciprocity gives an unexpected relation between the Legendre symbols $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$ for distinct odd primes p and q.
  - Explicitly, as first proven by Gauss, we have $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}$. Equivalently, $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ if p or q is 1 mod 4, and otherwise $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$ if both p, q are 3 mod 4.
  - A priori, it would seem that there is no reason for the values of $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$ to be related to one another, since they are discussing seemingly independent questions (whether p is a square mod q and whether q is a square mod p).
  - But in fact, these questions are related: for $p^* = (-1)^{(p-1)/2}$, the value of $\left(\frac{p^*}{q}\right)$ determines whether the ideal (p) splits in the ring of integers of the quadratic extension $\mathbb{Q}(\sqrt{q^*})$ while the value of $\left(\frac{q^*}{p}\right)$ determines whether the ideal (q) splits in the ring of integers of the quadratic extension $\mathbb{Q}(\sqrt{p^*})$.
  - These two questions are related because there are several ways to understand the splitting of (q) in $\mathcal{O}_{\sqrt{p^*}}$.
  - First, from basic algebraic number theory, to determine whether (q) splits in $\mathcal{O}_{\sqrt{p^*}}$, one can study the splitting of the minimal polynomial $x^2 - x + \frac{1-p^*}{2}$ modulo q, which splits precisely when its discriminant p^* is a square: in other words, when $\left(\frac{p^*}{q}\right) = 1$.
  - Alternatively, one may look at the action of the local qth-power Frobenius map inside the Galois group of the cyclotomic field $\mathbb{Q}(\zeta_p)$, whose unique quadratic subfield is $\mathbb{Q}(\sqrt{p^*})$. Since the Galois group is cyclic, the Frobenius element Frob_q fixes $\mathbb{Q}(\sqrt{p^*})$ if and only if $q \in (\mathbb{Z}/p\mathbb{Z})^\times$ lies in $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/K)$. But this group is the unique index-2 subgroup of $(\mathbb{Z}/p\mathbb{Z})^*$, which is simply the quadratic residues, so this means (q) splits precisely when $\left(\frac{q}{p}\right) = 1$.
  - Comparing these two statements yields that $\left(\frac{p^*}{q}\right) = 1$ if and only if $\left(\frac{q}{p}\right) = 1$, and this can be shown to be equivalent to the usual version of quadratic reciprocity.
  - Exercise: For odd primes p, q, show that $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right)$ is equivalent to $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}$.

- There are very many other proofs of quadratic reciprocity, many of which involve lengthy formal manipulations of various sums and (generally) yield little to no intuition about why the result is actually true. There is a fairly nice proof using Gauss sums that, suitably interpreted, is really the same as the one given above.

- In $\mathbb{F}_q[t]$, we can generalize the reciprocity law to encompass general dth powers in a fairly convenient way.

- Definition: If $f \in \mathbb{F}_q[t]$ is nonconstant and a is relatively prime to f, we say that a is a dth-power residue modulo f if $x^d \equiv a \pmod{f}$ has a solution for x. (In other words, when a is the dth power of something mod f.)
  - Example: Over $\mathbb{F}_2[t]$, we see $t+1$ is a quadratic residue modulo t^3+t+1 since $t+1 \equiv (t^2+t+1)^2 \pmod{t^3+t+1}$.
  - Example: Over $\mathbb{F}_5[t]$, we see $3t^2+3t+4$ is a cubic residue modulo t^3+t+1 since $3t^2+3t+4 \equiv (t^2+2t)^3 \pmod{t^3+t+1}$.
  - By the Chinese remainder theorem, $x^d \equiv a \pmod{f}$ has a solution if and only if $x^d \equiv a \pmod{p^d}$ has a solution for each prime power p^d in the factorization of f.
  - Thus, we need only consider the case where the modulus is a prime power, and we can handle this case fairly easily using our earlier analysis of the structure of $(A/p^dA)^*$.

- We can start by looking at the prime-modulus case, since it is the simplest.
  - As we have mentioned previously, $(A/pA)^*$ is the multiplicative group of the finite field A/pA, so this group has order $q^{\deg p} - 1 = \tilde{p}^{f \deg p} - 1$.
  - If d does not divide $|p|-1$, then the dth power map on $(A/pA)^*$ is injective by Lagrange's theorem, so it is a bijection, and so everything in $(A/pA)^*$ is a dth power.
  - This means we can ignore divisors of d that aren't factors of $|p|-1$, and so the only interesting case is when d divides $|p|-1$.
  - By analogy with Euler's criterion in $\mathbb{Z}$, we would expect that the value of $a^{(|p|-1)/d}$ will identify whether or not a is a dth power. This is indeed the case:

- Proposition (dth Roots Mod p): If $p \in \mathbb{F}_q[t]$ is irreducible, a is not divisible by p, and d is a divisor of $|p|-1$, then $x^d \equiv a \pmod{p}$ is solvable if and only if $a^{(|p|-1)/d} \equiv 1 \pmod{p}$.
  - Proof 1: First, if $x^d \equiv a \pmod{p}$ then $a^{(|p|-1)/d} \equiv x^{|p|-1} \equiv 1 \pmod{p}$ by Euler.
  - For the converse, recall that we showed previously that $x^d \equiv 1 \pmod{p}$ has d solutions mod p whenever d divides $|p|-1$.
  - Therefore, the kernel of the dth-power map on $(A/pA)^*$ has size d, so by the first isomorphism theorem, the image, which is precisely the set of dth powers, has size $(|p|-1)/d$.
  - But by the same observation, there are exactly $(|p|-1)/d$ solutions to the equation $x^{(|p|-1)/d} \equiv 1 \pmod{p}$, so by the above, these must be exactly the dth powers.
  - Proof 2: As shown previously, $(A/pA)^*$ is cyclic of order $|p|-1$. Let u be a generator.
  - Since every element in $(A/pA)^*$ is a power of u, it is easy to see that for any d dividing $|p|-1$, the dth powers in $(A/pA)^*$ are precisely $\{u^d, u^{2d}, u^{3d}, \dots, u^{d(|p|-1)d} = 1\}$. All of these elements clearly satisfy $x^{(|p|-1)/d} \equiv 1 \pmod{p}$.
  - Conversely, if $a = u^k$ has $a^{(|p|-1)/d} \equiv 1 \pmod{p}$, then $u^{k(|p|-1)/d} \equiv 1 \pmod{p}$ so since u has order $|p|-1$, d must divide k.

- Now that we have analyzed the prime case, the prime-power case follows by "lifting" the solutions from the prime case.
  - This is a consequence of a much more general result known as Hensel's lemma, which we might as well do in general.

- <u>Proposition</u> (Hensel's Lemma): If $p \in \mathbb{F}_q[t]$ is irreducible, $a \in \mathbb{F}_q[t]$, and $r(x)$ is any polynomial such that $r(a) \equiv 0 \pmod{p^d}$ and $r'(a) \not\equiv 0 \pmod{p}$, then there is a unique k modulo p such that $r(a + kp^d) \equiv 0 \pmod{p^{d+1}}$. Explicitly, if $u = f'(a)^{-1} \pmod{p}$, then $k = -\frac{uf(a)}{p^d}$.
  - By repeatedly applying Hensel's lemma, we can lift a solution of $r(a) \equiv 0 \pmod{p}$ to a solution modulo p^2, and then lift that to a solution modulo p^3, and so on and so forth, until we have a solution to the equation modulo any power of p.
  - This iteration process yields a sequence of solutions $x \equiv a_j \pmod{p^j}$ for each j, where $a_{j+1} = a_j - \frac{1}{r'(a)} r(a_j)$, which one may recognize as the iteration procedure from Newton's root-finding method. In fact, if we instead think of solving the polynomial $r(x) = 0$ p-adically (which amounts to taking the inverse limit $\varprojlim(A/p^d A)$), this lifting procedure is precisely Newton's method with starting point $x = a$.
  - <u>Proof</u>: First, by the binomial theorem we have $(a + p^d k)^n = a^n + na^{n-1} p^d k + [\text{terms divisible by } p^{2d}] \equiv a^n + na^{n-1} p^d k \pmod{p^{d+1}}$.
  - Then if $r(t) = \sum c_n t^n$ we see that $r(a + p^d k) \equiv \sum c_n (a^n + na^{n-1} p^d k) \equiv \sum c_n a^n + p^d k \sum nc_n a^{n-1} \equiv r(a) + p^d k \cdot r'(a) \pmod{p^{d+1}}$.
  - By hypothesis, $r(a) + p^d k \cdot r'(a)$ is divisible by p^d. So dividing the congruence $r(a + kp^d) \equiv 0 \pmod{p^{d+1}}$ by p^d yields $\frac{r(a)}{p^d} + kr'(a) \equiv 0 \pmod{p}$, which has the unique solution $k \equiv -\frac{uf(a)}{p^d} \pmod{p}$, as claimed.
- This version of Hensel's lemma is quite a bit more than we really need here, but it will be helpful to have it available later.
- <u>Corollary</u> (dth Roots Mod p^e): If $p \in \mathbb{F}_q[t]$ is irreducible, d divides $|p| - 1$, and p does not divide a, then $x^d \equiv a \pmod{p}$ has a root if and only if $x^d \equiv a \pmod{p^e}$ has a root for every $e \geq 1$.
  - Proof: If there is a solution to $x^d \equiv a \pmod{p^e}$ then clearly there is a solution mod p.
  - Conversely, if there is a solution mod p, then we claim we may lift the solution mod p^e using Hensel's lemma.
  - We just need to check that the derivative is not zero: for $r(x) = x^d$ we have $r'(a) = da^{d-1}$. Then $d \neq 0$ mod p because d divides $|p| - 1 = \tilde{p}^{f \deg p} - 1$ and so d cannot be divisible by the characteristic $\tilde{p}$, and also $a \neq 0$ mod p because p does not divide a. Thus, Hensel's lemma applies, and we are done.
- <u>Corollary</u> (Counting dth Powers): If $p \in \mathbb{F}_q[t]$ is irreducible and d divides $|p| - 1$, then there are $\Phi(p^e)/d$ total dth-power residues modulo p^e.
  - <u>Proof 1</u>: Count residue classes: as shown earlier there are $(|p| - 1)/d = \Phi(p)/d$ total dth-power residue classes modulo p. By the corollary above, the dth-power residue classes modulo p^e are precisely those that reduce to a dth power modulo p. So the probability of selecting one is $\Phi(p)/(d\,|p|)$, and thus the total number is $|p|^e \cdot \Phi(p)/(d\,|p|) = \Phi(p^e)/d$.
  - <u>Proof 2</u> (sketch): The dth-power homomorphism commutes with reduction modulo p. Then just count the sizes of the various kernels and images and use the first isomorphism theorem.
  - <u>Exercise</u>: Show that for any monic polynomial m, there are $\Phi(m)/d^{\lambda(m)}$ total dth powers modulo m, where $\lambda(m)$ is the number of distinct monic irreducible factors of m.
- Returning back to the prime case, in the particular case where d divides $q - 1$, then the dth roots of unity in $(A/pA)^*$ actually lie inside $\mathbb{F}_q$, because $x^d = 1$ already has d solutions inside $\mathbb{F}_q$ (since $\mathbb{F}_q^*$ is cyclic of order $q - 1$).
  - We have shown above that a is a dth power modulo p if and only if $a^{(|p|-1)/d} \equiv 1 \pmod{p}$.
  - We can use this as the basis for our definition of the dth-power residue symbol, in analogy with Euler's criterion over $\mathbb{Z}$.
- <u>Definition</u>: If $p \in \mathbb{F}_q[t]$ is irreducible and d divides $q - 1$, then we define the <u>dth-power residue symbol</u> $\left(\frac{a}{p}\right)_d$ to be the unique element of $\mathbb{F}_q$ congruent to $a^{(|p|-1)/d}$ modulo p.

- Example: For $d = 2$ over $\mathbb{F}_3[t]$, we calculate $\left(\frac{t}{t^2+t+2}\right)_2 \equiv t^4 \equiv 2 \pmod{t^2+t+2}$.
- Example: For $d = 3$ over $\mathbb{F}_7[t]$, we calculate $\left(\frac{t}{t^2+2t+2}\right)_7 \equiv t^{16} \equiv 4 \pmod{t^2+2t+2}$.
- Example: For $d = 3$ over $\mathbb{F}_7[t]$, we calculate $\left(\frac{t}{t^2+t+6}\right)_7 \equiv t^{16} \equiv 1 \pmod{t^2+t+6}$, which means t is a cube modulo t^2+t+6.

- Proposition (Properties of Residue Symbols): If $p \in \mathbb{F}_q[t]$ is irreducible and d divides $q-1$, the following hold:

1. $\left(\frac{a}{p}\right)_d = 0$ if and only if p divides a.
2. If $a \equiv b \pmod{p}$ then $\left(\frac{a}{p}\right)_d = \left(\frac{b}{p}\right)_d$.
3. The residue symbol is multiplicative: for any a, b, $\left(\frac{ab}{p}\right)_d = \left(\frac{a}{p}\right)_d \left(\frac{b}{p}\right)_d$.
4. $\left(\frac{a}{p}\right)_d = 1$ if and only if a is a dth-power residue modulo p.
5. If ζ is any dth root of unity in $\mathbb{F}_q$, then there exists $a \in \mathbb{F}_q[t]$ with $\left(\frac{a}{p}\right)_d = \zeta$.
6. The residue symbol is a surjective group homomorphism from $(A/pA)^*$ to μ_d, the group of dth roots of unity in $\mathbb{F}_q$.
7. If $d|d'$ then $\left(\frac{a}{p}\right)_d = \left(\frac{a}{p}\right)_{d'}^{d'/d}$.
8. If $\alpha \in \mathbb{F}_q$ then $\left(\frac{\alpha}{p}\right)_d = \alpha^{(q-1)/d \cdot \deg p}$.

   - Proofs: (1)-(4) are trivial from the definition or results previously shown. (5) follows by the first isomorphism theorem, since the kernel of the $(|p|-1)/d$th-power map has size $(|p|-1)/d$ hence the image has size d. (6) is a rephrasing of (3) and (5).
   - (7) follows by noting $\left(\frac{a}{p}\right)_{d'}^{d'/d} \equiv (a^{(|p|-1)/d'})^{d'/d} = a^{(|p|-1)/d} \equiv \left(\frac{a}{p}\right)_d \pmod{p}$, and then observing that since the residue symbols are both elements of $\mathbb{F}_q$, the congruence mod p forces actual equality.
   - For (8), first note that $\frac{|p|-1}{d} = \frac{q^{\deg p}-1}{d} = (1+q+q^2+\cdots+q^{\deg p-1})(q-1)/d$. Then since $\alpha^q = \alpha$ by Fermat's little theorem in $\mathbb{F}_q$, we have $\left(\frac{\alpha}{p}\right)_d \equiv \alpha^{(|p|-1)/d} = (\alpha \cdot \alpha^q \cdot \alpha^{q^2} \cdot \cdots \cdot \alpha^{q^{\deg p-1}})^{(q-1)/d} = \alpha^{\deg p \cdot (q-1)/d} \pmod{p}$. Then as in (7), the congruence modulo p forces equality.

- We can now state the dth-power reciprocity law, which we will prove next time:
- Theorem (dth-Power Reciprocity): If d divides $q-1$ and P, Q are monic irreducible polynomials in $\mathbb{F}_q[t]$, then $\left(\frac{Q}{P}\right)_d = (-1)^{(q-1)(\deg P)(\deg Q)/d}\left(\frac{P}{Q}\right)_d$.

0.6 (Sep 27) Generalizations + Applications of the Reciprocity Law

- To prove the reciprocity law, we first need a reciprocity result about roots of polynomials known as Weil reciprocity:
- Lemma (Weil Reciprocity): If $P(t) = (t-r_1)\cdots(t-r_n)$ and $Q(t) = (t-s_1)\cdots(t-s_n)$ are monic polynomials over a field F, with the $r_i, s_j \in F$, then $\prod_{i=1}^n Q(r_i) = (-1)^{(\deg P)(\deg Q)} \prod_{j=1}^m P(s_j)$.

- Proof: Note that $Q(r_i) = \prod_{j=1}^m (r_i - s_j)$ so $\prod_{i=1}^n Q(r_i) = \prod_{i=1}^n \prod_{j=1}^m (r_i - s_j)$. In the same way, $\prod_{j=1}^m P(s_j) = \prod_{j=1}^m \prod_{i=1}^n (s_j - r_i)$.
- These expressions are the same up to switching the order of the products and scaling each of the $mn = (\deg P)(\deg Q)$ terms by -1, so the result follows.

- We can now state and prove the dth-power reciprocity law:
- Theorem (dth-Power Reciprocity): If d divides $q-1$ and P, Q are monic irreducible polynomials in $\mathbb{F}_q[t]$, then $\left(\frac{Q}{P}\right)_d = (-1)^{(q-1)(\deg P)(\deg Q)/d} \left(\frac{P}{Q}\right)_d$.
  - The main idea of the proof is to exploit properties of the Frobenius map on the roots of P and Q in their splitting field over $\mathbb{F}_q$, and then use Weil reciprocity.
  - Proof: From property (7) of the residue symbol, we have $\left(\frac{a}{p}\right)_d = \left(\frac{a}{p}\right)_{d'}^{d'/d}$, so it is enough to prove the reciprocity law when $d = q-1$.
  - Now let α be a root of P and β be a root of Q in a splitting field $E/\mathbb{F}_q$ for the polynomial PQ.
  - Since $E/\mathbb{F}_q$ is a finite-degree extension of a finite field, its Galois group is cyclic and generated by the qth-power Frobenius map.
  - Also, since P and Q are irreducible over $\mathbb{F}_q$, we must have the factorizations

$$P(t) = (t-\alpha)(t-\alpha^q)(t-\alpha^{q^2})\cdots(t-\alpha^{q^{\deg P-1}})$$
$$Q(t) = (t-\beta)(t-\beta^q)(t-\beta^{q^2})\cdots(t-\beta^{q^{\deg Q-1}})$$

    since α, α^q, α^{q^2}, ... are all the Galois conjugates of α and P is irreducible (with the same logic applying to β and Q).
  - Inside $E[t]$, we have $\left(\frac{Q}{P}\right)_{q-1} \equiv [Q(t)]^{(q^{\deg P}-1)/(q-1)} = [Q(t)]^{1+q+q^2+\cdots+q^{\deg P-1}} = Q(t)Q(t)^q Q(t)^{q^2}\cdots Q(t)^{q^{\deg P-1}} \equiv Q(t)Q(t^q)Q(t^{q^2})\cdots Q(t^{q^{\deg P-1}}) \pmod{P}$ since $Q(t^q) = Q(t)^q$ in characteristic q.
  - Reducing both sides modulo the factor $t-\alpha$ of P (equivalently, evaluating both sides at $t=\alpha$) then yields $\left(\frac{Q}{P}\right)_{q-1} \equiv Q(\alpha)Q(\alpha^q)\cdots Q(\alpha^{q^{\deg P}-1}) \pmod{t-\alpha}$. Since the right-hand side of this expression is the product of the values of Q evaluated at the roots of P, it is the same for any other root of P we choose in place of α.
  - So by the Chinese remainder theorem, in fact $\left(\frac{Q}{P}\right)_{q-1} \equiv Q(\alpha)Q(\alpha^q)\cdots Q(\alpha^{q^{\deg P}-1}) \pmod{P}$. But the right-hand side is an element of E, and since it is a $(q-1)$st root of unity (or alternatively, since it is Galois-invariant), it must actually be in $\mathbb{F}_q$. So since these quantities are congruent modulo P, they must actually be equal as elements of $\mathbb{F}_q$.
  - This means $\left(\frac{Q}{P}\right)_{q-1} = Q(\alpha)Q(\alpha^q)\cdots Q(\alpha^{q^{\deg P}-1})$. In the same way, $\left(\frac{P}{Q}\right)_{q-1} = P(\beta)P(\beta^q)\cdots P(\beta^{q^{\deg Q}-1})$.
  - Weil reciprocity then says $Q(\alpha)Q(\alpha^q)\cdots Q(\alpha^{q^{\deg P}-1}) = (-1)^{(\deg P)(\deg Q)} P(\beta)P(\beta^q)\cdots P(\beta^{q^{\deg Q}-1})$, so we see $\left(\frac{Q}{P}\right)_{q-1} = (-1)^{(\deg P)(\deg Q)} \left(\frac{P}{Q}\right)_{q-1}$, which establishes the case $d = q-1$.
  - The case where d divides $q-1$ follows immediately and gives the general statement above.

- Just as in the case of $\mathbb{Q}$, to give a convenient method for calculating residue symbols, we can extend the definition to include nonprime moduli (i.e., generalizing the Jacobi symbol):
- Definition: If $b \in \mathbb{F}_q[t]$ has prime factorization $b = uq_1^{b_1}\cdots q_n^{b_n}$ for distinct monic irreducible q_i and $u \in \mathbb{F}_q^\times$, then we define the general residue symbol as $\left(\frac{a}{b}\right)_d = \prod_{j=1}^n \left(\frac{a}{q_i}\right)^{b_i}$.

- Proposition (Properties of Residue Symbols, II): If $b \in \mathbb{F}_q[t]$ is nonzero and d divides $q-1$, the following hold:

1. $\left(\frac{a}{b}\right)_d$ is either 0 or a dth root of unity, and $\left(\frac{a}{b}\right)_d \neq 0$ if and only if a, b are relatively prime.
2. If $a_1 \equiv a_2 \pmod{b}$ then $\left(\frac{a_1}{b}\right)_d = \left(\frac{a_2}{b}\right)_d$.
3. The residue symbol is multiplicative on the top: $\left(\frac{a_1 a_2}{b}\right)_d = \left(\frac{a_1}{b}\right)_d \left(\frac{a_2}{b}\right)_d$.
4. The residue symbol is multiplicative on the bottom: $\left(\frac{a}{b_1 b_2}\right)_d = \left(\frac{a}{b_1}\right)_d \left(\frac{a}{b_2}\right)_d$.
5. If $\gcd(a, b) = 1$ and a is a dth-power residue modulo b, then $\left(\frac{a}{b}\right)_d = 1$. (The converse need not hold.)
6. If $d | d'$ then $\left(\frac{a}{b}\right)_d = \left(\frac{a}{b}\right)_{d'}^{d'/d}$.
7. If $\alpha \in \mathbb{F}_q$ then $\left(\frac{\alpha}{b}\right)_d = \alpha^{(q-1)/d \cdot \deg b}$.

  - Proofs: (1)-(4) follow straightforwardly from the definition, while (6) and (7) follow the same way as for the residue symbol with prime modulus. For (5), if $a \equiv c^d \pmod{p}$ then $\left(\frac{a}{b}\right)_d = \left(\frac{c^d}{b}\right)_d = \left(\frac{c}{b}\right)_d = 1$ since $\left(\frac{c}{b}\right)$ is a dth root of unity (since it is not zero since a, b are relatively prime).
  - We will remark that the residue symbol $\left(\frac{\star}{b}\right)_d : (A/bA)^* \to \mu_d$ is still a group homomorphism since it is multiplicative by (3), but it is not necessarily surjective when b is not prime. For example, if $b = p^d$ is a dth power, then by (4) we see that $\left(\frac{a}{b}\right)_d = \left(\frac{a}{p}\right)_d^d = 1$ for all $a \in (A/bA)^*$. (This also shows that the converse of (5) is false, as noted above.)

- We can write down the reciprocity law for general dth-power residue symbols:
- Theorem (General Reciprocity Law): If d divides $q-1$ and a, b are any nonzero polynomials in $\mathbb{F}_q[t]$, then $\left(\frac{a}{b}\right)_d = (-1)^{(q-1)(\deg a)(\deg b)/d} [\text{sgn} a]^{(q-1)/d \cdot \deg b} [\text{sgn} b]^{-(q-1)/d \cdot \deg a} \left(\frac{b}{a}\right)_d$.

  - Proof (sketch): As in the prime case, reduce to the case $d = q-1$. Then pull out the leading coefficients of a, b (these are where the sgn$a$ and sgn$b$ terms come from) and then apply the definition of the general residue symbol to write $\left(\frac{a}{b}\right)_{q-1}$ and $\left(\frac{b}{a}\right)_{q-1}$ as products of residue symbols with prime moduli, apply the prime-modulus reciprocity law, and tally up the results. The full details are left as an exercise.

- A standard application of quadratic reciprocity over $\mathbb{Z}$ is to characterize all of the prime moduli for which a given integer m is a quadratic residue.

  - Typical examples of such statements: -1 is a quadratic residue mod p when $p \equiv 1 \pmod{4}$, 3 is a quadratic residue mod p when $p \equiv 1, 11 \pmod{12}$, 5 is a quadratic residue mod p when $p \equiv 1, 4 \pmod{5}$, and so forth.
  - Aside from the special cases of -1 and 2, one may answer this question simply by factoring m as a product of primes $m = q_1 \cdots q_k$, so that $\left(\frac{m}{p}\right) = \left(\frac{q_1}{p}\right) \cdots \left(\frac{q_k}{p}\right)$, and then applying quadratic reciprocity to flip each of the quadratic residue symbols. The end result is that the statement $\left(\frac{m}{p}\right) = +1$ is equivalent to a congruence condition for p modulo $4m$, which one may calculate explicitly if desired.

- We can use this same type of argument to solve the analogous problem in function fields:

- Theorem (Criterion for dth-Power Residues): Let $m \in \mathbb{F}_q[t]$ be monic and $d|(q-1)$, and let $\{a_1, \dots, a_k\}$ be coset representatives for the residue classes in $(A/mA)^*$ with $\left(\dfrac{a}{m}\right)_d = +1$ and $\{b_1, \dots, b_k\}$ be coset representatives for the residue classes in $(A/mA)^*$ with $\left(\dfrac{b}{m}\right)_d = -1$ (if there are any). Then the following hold:

  1. If $\deg(m)$, $(q-1)/d$, or $\mathrm{char}(\mathbb{F}_q)$ is even, then m is a dth power modulo an irreducible monic polynomial p if and only if $p \equiv a_i \pmod{m}$ for some i.
  2. If $\deg(m)$, $(q-1)/d$, and $\mathrm{char}(\mathbb{F}_q)$ are all odd, then m is a dth power modulo an irreducible monic polynomial p if and only if either $\deg(p)$ is even and $p \equiv a_i \pmod{m}$ for some i, or $\deg(p)$ is odd and $p \equiv b_i \pmod{m}$ for some i.

  - Proof: Note that $p \equiv a_i \pmod{m}$ is equivalent to saying $\left(\dfrac{p}{m}\right)_d = 1$, while $p \equiv b_i \pmod{m}$ is equivalent to saying $\left(\dfrac{p}{m}\right)_d = -1$.
  - Since p and m are monic, by the reciprocity law we see $\left(\dfrac{m}{p}\right)_d = (-1)^{(q-1)/d \cdot \deg(m)\deg(p)} \left(\dfrac{p}{m}\right)_d$.
  - First, if q is even, then $\mathrm{char}(\mathbb{F}_q) = 2$: then $-1 = 1$ over $\mathbb{F}_q$, so $\left(\dfrac{m}{p}\right)_d = \left(\dfrac{p}{m}\right)_d$. Likewise, if $\deg(m)$ or $(q-1)/d$ is even, then the exponent of -1 is even, so again we see $\left(\dfrac{m}{p}\right)_d = \left(\dfrac{p}{m}\right)_d$. Together with the observation above, (1) follows.
  - For (2), if $\deg(m)$, $(q-1)/d$, and $\mathrm{char}(\mathbb{F}_q)$ are all odd, then $-1 \neq 1$ and $(-1)^{(q-1)/d \cdot \deg(m)\deg(p)} = (-1)^{\deg p}$. So $\left(\dfrac{m}{p}\right)_d = \left(\dfrac{p}{m}\right)_d$ if $\deg(p)$ is even while $\left(\dfrac{m}{p}\right)_d = -\left(\dfrac{p}{m}\right)_d$ if $\deg(p)$ is odd. This yields (2).

- Example: Identify all monic irreducibles $p \in \mathbb{F}_3[t]$ such that t is a square modulo p.
  - There are two residue classes in $(A/tA)^*$, namely 1 and 2, and we see $\left(\dfrac{1}{t}\right)_2 = 1$ while $\left(\dfrac{2}{t}\right)_2 = -1$.
  - Since $\deg(m) = 1$, $(q-1)/d = 1$, and $\mathrm{char}(\mathbb{F}_q) = 3$, we are in case (2). Thus, m is a quadratic residue modulo the monic irreducible polynomial p precisely when $\deg(p)$ is odd and $p \equiv 2 \pmod{t}$, or when $\deg(p)$ is even and $p \equiv 1 \pmod{t}$.
  - For example, we see that t is a square modulo the irreducible polynomial $t^3 + 2t + 2 \in \mathbb{F}_3[t]$, and indeed with some more work, one may calculate $t \equiv (t^2 + t + 2)^2 \pmod{t^3 + 2t + 2}$.
  - Exercise: Extend this example to describe all monic irreducibles $p \in \mathbb{F}_q[t]$ such that t is a square modulo p for arbitrary finite fields $\mathbb{F}_q$.

- Another interesting application of the dth-power reciprocity law is to establish a "Hasse principle"-type result for dth powers.
  - Obviously, if a polynomial with integer coefficients has a solution in $\mathbb{Z}$, then it also has solutions modulo p^k for all prime powers p^k (equivalently, it has a p-adic solution for each p) and it also has a real solution.
  - The Hasse principle asks when the converse of this observation is valid: if a polynomial has a p-adic root and a real root, does it necessarily have a rational root? The general idea is that one may try to piece together information modulo the prime powers for many primes p using the Chinese remainder theorem, but it is not clear when this actually forces the existence of a global solution.
  - As first proven by Minkowski for integer coefficients (and then later extended by Hasse for number-field coefficients), for quadratic polynomials this local-global principle holds: if a quadratic polynomial has a p-adic root and a real root, it necessarily has a rational root.
  - The result is known to be false for cubic forms: Selmer's famous counterexample is the cubic equation $3x^3 + 4y^3 + 5z^3 = 0$, which has no rational solution but does have real solutions and p-adic solutions for all p.

- Even in the absence of a literal Hasse-principle statement, in many cases one can analyze the precise obstructions to lifting local solutions to global solutions. (An example of this sort of obstruction can be found in the statement of the Grunwald-Wang theorem.)

- Theorem (Hasse Principle for dth Powers): Let $m \in \mathbb{F}_q[t]$ have positive degree and $d|(q-1)$. If $x^d \equiv m$ (mod p) is solvable for all but finitely many irreducible polynomials p, then $x^d = m$ has a solution in $\mathbb{F}_q[t]$ (i.e., m is globally a dth power).
  - Proof: Let $m = \beta q_1^{d_1} \cdots q_k^{d_k}$ where the q_i are distinct monic irreducibles and β is a constant. We first show that if any d_i is not divisible by d, then there are infinitely many irreducibles p such that $\left(\frac{m}{p}\right)_d \neq 1$.
  - To show this, suppose without loss of generality that d_1 is not divisible by d. We inductively construct an infinite set of irreducibles $\{r_i\}$ with $\left(\frac{m}{r_i}\right)_d \neq 1$, so suppose we have a set (possibly empty to start) $\{r_1, \dots, r_s\}$ of monic irreducibles not dividing m with $\left(\frac{m}{r_i}\right)_d \neq 1$ for all i.
  - Select any primitive dth root of unity ζ_d: then there exists an element $c \in \mathbb{F}_q[t]$ with $\left(\frac{c}{q_i}\right)_d = \zeta_d$ by our properties of the dth-power residue symbol.
  - By the Chinese remainder theorem, there exist solutions a to the system of congruences $a \equiv c$ (mod q_1), $a \equiv 1$ (mod $q_2 \cdots q_k$), $a \equiv 1$ (mod $r_1 \cdots r_s$). Select any such solution that is monic and has degree divisible by $2d$.
  - For this a, we have $\left(\frac{a}{m}\right)_d = \prod_{i=1}^k \left(\frac{a}{q_i}\right)_d^{d_i} = \zeta_d^{d_1} \neq 1$ since d_1 is not divisible by d.
  - Then by the reciprocity law, we then have $\left(\frac{m}{a}\right)_d = (-1)^{(q-1)/d\cdot(\deg m)(\deg a)} \left(\frac{a}{m}\right)_d = \left(\frac{a}{m}\right)_d \neq 1$, since the exponent of -1 has a factor of 2 from $\deg a$.
  - Since the general dth-power residue symbol is multiplicative on the bottom, there must be some monic irreducible factor r_{s+1} of a such that $\left(\frac{m}{r_{s+1}}\right) \neq 1$ since $\left(\frac{a}{m}\right)_d \neq 1$. This monic irreducible factor is relatively prime to $r_1 \cdots r_s$ since $a \equiv 1$ (mod $r_1 \cdots r_s$), so we have found another monic irreducible to add to our list.
  - By induction, we can construct infinitely many such irreducibles.
  - Now, if $x^d \equiv m$ (mod p) is solvable for all but finitely many irreducible polynomials p, then by the above, each of the exponents d_i must be divisible by d. This means $m = \beta \cdot \tilde{m}^d$ for some monic polynomial $\tilde{m}$, so all that remains is to show that β is a dth power.
  - For any irreducible p not dividing m, we have $\left(\frac{m}{p}\right)_d = \left(\frac{\beta}{p}\right)_d = \beta^{(q-1)/d\cdot\deg p}$ as we have previously shown. Since there are irreducibles of any desired degree in $\mathbb{F}_q[t]$, select p to be one of degree relatively prime to d with $\left(\frac{m}{p}\right)_d = 1$: then $\beta^{(q-1)/d\cdot\deg p} = 1$ implies $\beta^{(q-1)/d} = 1$, which is equivalent to saying that β is a dth power. Then m itself is a dth power, as claimed.

0.7 (Sep 30) Group Characters + Primes in Arithmetic Progressions Part 1

- Our next task is to prove the function-field analogue of Dirichlet's theorem on primes in arithmetic progressions.
  - Over $\mathbb{Q}$, Dirichlet's theorem says that for any positive integer m and any a relatively prime to m, there exist infinitely many primes in the arithmetic progression $\{a, a+m, a+2m, a+3m, \dots\}$: in other words, congruent to a modulo m.
  - Exercise (easy): Show that if a is not relatively prime to m, then there are only finitely many primes congruent to a modulo m.

- There are $\varphi(m)$ residue classes modulo m that contain infinitely many primes, so one can ask more precisely about how the primes are distributed among these residue classes.
  - In fact, the primes are asymptotically uniformly distributed among these residue classes: the proportion of primes congruent to a modulo m approaches $1/\varphi(m)$ upon taking an appropriate limit.
  - Explicitly, define the natural density of a set S of primes to be $\lim_{n\to\infty} \frac{S \cap \{1,2,\dots,n\}}{\{\text{primes}\} \cap \{1,2,\dots,n\}}$, provided the limit exists.
  - Then, as first proven by de la Vallée Poussin, the natural density of the primes congruent to a modulo m is $1/\varphi(m)$ when a is relatively prime to m.
- However, the natural density is somewhat difficult to handle with analytic methods. From the standpoint of zeta functions, a more natural choice is the Dirichlet density:
- Definition: If S is a set of primes, the Dirichlet density of S is the value $\delta_S = \lim_{s\to 1+} \frac{\sum_{\text{primes } p\in S} p^{-s}}{\sum_{\text{primes } p} p^{-s}}$, assuming the limit exists.
  - Note that the sum in the numerator is always finite for $\text{Re}(s) > 1$ by comparison to the sum for the zeta function.
  - Exercise: If S is finite, show that its Dirichlet density is 0.
  - One may prove that if a set has natural density δ, then its Dirichlet density is also δ. The converse is not true, however: a simple counterexample due to Serre is the set S of primes whose leading digit is 1 in base 10.
  - Exercise (hard): Show that the set of primes whose leading digit is 1 in base 10 has undefined natural density, but has Dirichlet density $\log_{10} 2$. (The answer works out the same if you use integers with leading digit 1.)
- The corresponding definition for function fields is as follows:
- Definition: If T is a set of monic irreducibles in $\mathbb{F}_q[t]$, its Dirichlet density is $\delta_T = \lim_{s\to 1+} \frac{\sum_{p\in T} |p|^{-s}}{\sum_p |p|^{-s}}$, assuming the limit exists.
  - We note that both the numerator and denominator sums converge for $\text{Re}(s) > 1$.
- Our main result is the following:
- Theorem (Analogue of Dirichlet's Theorem): Let $m \in \mathbb{F}_q[t]$ have positive degree and let a be relatively prime to m. Then the Dirichlet density of the set of primes congruent to a (mod m) exists and is $1/\Phi(m)$. In particular, there are infinitely many such primes.
  - The fundamentally hard part of proving this theorem is to establish the nonvanishing of the L-functions for nontrivial characters at $s = 1$.
  - In order to explain what this means (and then do it), we will begin with a brisk discussion of Dirichlet characters and their properties.
- Definition: Let G be a finite abelian group. A group character χ of G is a homomorphism $\chi : G \to \mathbb{C}^\times$.
  - Note that $\chi(1) = 1$ for every character, and also if $g \in G$ has order d, then $1 = \chi(1) = \chi(g^d) = \chi(g)^d$, so $\chi(g)$ is a dth root of unity. Thus in general, χ is a map from G to the group of complex $|G|$th roots of unity.
  - Example: For any G, the trivial character χ_{triv} has $\chi_{\text{triv}}(g) = 1$ for all $g \in G$.
  - Example: If $G = (\mathbb{Z}/p\mathbb{Z})^\times$, the quadratic residue symbol $\chi(a) = \left(\frac{a}{p}\right)$ is a group character.

- Example: If $G = (A/pA)^\times$ for $A = \mathbb{F}_q[t]$ and d divides $q-1$, the dth-power residue symbol $\chi(a) = \left(\frac{a}{p}\right)_d$ gives a group character, provided we identify the dth roots of unity in $\mathbb{F}_q$ with the dth roots of unity in $\mathbb{C}$ (simply choose any fixed isomorphism).

- We will be interested in the case where G is the group of units $(\mathbb{Z}/m\mathbb{Z})^\times$ or $(A/fA)^\times$, in which case we call χ a Dirichlet character.
  - In some situations it is slightly more convenient to work with extended Dirichlet characters, which we extend to have domain $\mathbb{Z}/m\mathbb{Z}$ or A/fA by setting $\chi(a) = 0$ whenever a is not relatively prime to the modulus.
  - Exercise: Extended Dirichlet characters modulo m are the same as functions $\chi : \mathbb{Z} \to \mathbb{C}$ (or $A \to \mathbb{C}$) such that (i) $\chi(a+bm) = \chi(a)$ for all a, b, (ii) $\chi(ab) = \chi(a)\chi(b)$ for all a, b, and (iii) $\chi(a) \neq 0$ iff a is relatively prime to m.
- We can multiply two group characters on G pointwise, and this operation makes them into a group:
- Proposition (Dual Group of G): The set of group characters on G forms a group under pointwise multiplication. The identity is the trivial character and the inverse of χ is its complex conjugate $\overline{\chi}$. This group is called the dual group of G and is denoted $\hat{G}$.
  - Proof: These properties can be checked directly (exercise), or one may simply note that $\hat{G} = \text{Hom}(G, \mathbb{C}^\times)$.
- The dual group $\hat{G}$ is also an abelian group, so it is natural to wonder how its structure relates to G. In fact, it is isomorphic to G:
- Proposition (Dual Group, II): If G is a finite abelian group, its dual group $\hat{G}$ is isomorphic to G.
  - Proof: First consider the special case where G is a cyclic group of order n generated by g. Then $\chi(g^d) = \chi(g)^d$ for all d, so any group character χ is uniquely determined by the value of $\chi(g)$, which must be some nth root of unity.
  - Conversely, any such selection $e^{2\pi ia/n}$ for $\chi(g)$ yields a valid group character χ_a, namely with $\chi_a(g^d) = e^{2\pi iad/n}$. Since $\chi_a\chi_b = \chi_{a+b}$ and χ_1^n is the trivial character, we see that the dual group $\hat{G}$ is cyclic of order n (the map $a \mapsto \chi_a$ is an isomorphism of $\hat{G}$ with $\mathbb{Z}/n\mathbb{Z}$).
  - Now suppose $G = H \times K$ is a direct product. If $\chi : H \times K \to \mathbb{C}^\times$ is a homomorphism, let $\chi_H : H \to \mathbb{C}^\times$ and $\chi_K : K \to \mathbb{C}^\times$ be the projections $\chi_H(h) = \chi(h, 1)$ and $\chi_K(k) = \chi(1, k)$. Then χ_H is a group character of H, χ_K is a group character of K, and $\chi = \chi_H\chi_K$.
  - Conversely, any pair $(\chi_H, \chi_K) \in (\hat{H}, \hat{K})$ yields a character $\chi = \chi_H\chi_K \in \hat{G}$, so we see $\hat{G} \cong \hat{H} \times \hat{K}$.
  - Since every finite abelian group is a direct product of cyclic groups, and the result holds for cyclic groups and direct products, we are done.
- Exercise: If H is a subgroup of the finite abelian group G, define $H^\perp = \{\chi \in \hat{G} : \chi(H) = 1\}$. Show that $H^\perp \cong \widehat{G/H}$ and that $\hat{G}/H^\perp \cong \hat{H}$. Use these results along with $\hat{G} \cong G$ to conclude that the subgroup lattice of G is the same when turned upside down.
- The isomorphism between $\hat{G}$ and G above is non-canonical (i.e., it is not "coordinate-free" in the sense that we must pick specific generators for G and $\hat{G}$ to obtain the isomorphism).
  - However, there is a canonical isomorphism between $\hat{\hat{G}}$ (the double dual) and G given by the "evaluation map" φ, which maps an element $g \in G$ to the "evaluation-at-g" map e_g on characters $\chi \in \hat{G}$, defined by $e_g(\chi) = \chi(g)$.
  - Exercise: Verify that the evaluation map $\varphi : G \to \hat{\hat{G}}$ with $\varphi(g) = \{\chi \mapsto \chi(g)\}$ is an isomorphism from $\hat{\hat{G}}$ to G.
  - This result is a special case of Pontryagin duality, and has an analogous statement for duals of finite-dimensional vector spaces.

* In fact, it is really the algebraic analogue of Fourier inversion (the reason being that Fourier analysis on finite abelian groups involves sums over group characters in lieu of integrals). For a brief taste of the analogy, the main idea is to note that the map $e^{inx} : \mathbb{R} \to \mathbb{C}^{\times}$ is a group homomorphism, and thus is an "$\mathbb{R}$"-character.

- We can also put the structure of an inner product on group characters. To establish this we first show some simple orthogonality relations:
- Proposition (Orthogonality Relations): If G is a finite abelian group and χ is a group character, the following hold:

1. The sum $\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi \text{is trivial} \\ 0 & \text{otherwise} \end{cases}$.
   * Proof: If χ is trivial the sum is clearly $|G|$. If χ is not trivial, say with $\chi(h) \neq 1$, then $\sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(gh) = \chi(h) \sum_{g \in G} \chi(g)$ by reindexing (since $G = Gh$), and so $\sum_{g \in G} \chi(g) = 0$.
2. The sum $\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |G| & \text{if } g = 1 \\ 0 & \text{otherwise} \end{cases}$.
   * Proof: Apply Pontryagin duality to (1).
3. (Orthogonality 1) For any characters χ_1 and χ_2, $\sum_{g \in G} \chi_1(g)\overline{\chi_2(g)} = \begin{cases} |G| & \text{if } \chi_1 = \chi_2 \\ 0 & \text{otherwise} \end{cases}$.
   * Proof: Apply (1) to $\chi = \chi_1 \overline{\chi_2}$.
4. (Orthogonality 2) For any elements g_1 and g_2, $\sum_{\chi \in \hat{G}} \chi(g_1)\overline{\chi(g_2)} = \begin{cases} |G| & \text{if } g_1 = g_2 \\ 0 & \text{otherwise} \end{cases}$.
   * Proof: Apply (2) to $g = g_1 g_2^{-1}$, or apply Pontryagin duality to (3).
5. The pairing $\langle f_1, f_2 \rangle_G = \frac{1}{|G|} \sum_{g \in G} f_1(g)\overline{f_2(g)}$ is a complex inner product on functions $f : G \to \mathbb{C}$, and the elements of the dual group $\hat{G}$ are an orthonormal basis with respect to this inner product.
   * Proof: The inner product axioms are straightforward, and the fact that $\hat{G}$ yields an orthonormal basis follows from (3).
6. The pairing $\left\langle \hat{f}_1, \hat{f}_2 \right\rangle_{\hat{G}} = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \hat{f}_1(\chi)\overline{\hat{f}_2(\chi)}$ is a complex inner product on functions $\hat{f} : \hat{G} \to \mathbb{C}$, and the elements of G are an orthonormal basis with respect to this inner product.
   * Proof: The inner product axioms are straightforward, and the fact that $G \cong \hat{\hat{G}}$ yields an orthonormal basis follows from (4), or apply Pontryagin duality to (5).
7. (Fourier Inversion) For any function $f : G \to \mathbb{C}$, with the Fourier transform $\hat{f} : \hat{G} \to \mathbb{C}$ defined by $\hat{f}(\chi) = \langle f, \chi \rangle_G = \frac{1}{|G|} \sum_{g \in G} f(g)\overline{\chi(g)}$, we have $f(g) = \sum_{\chi \in \hat{G}} \hat{f}(\chi)\chi(g)$ for all $g \in G$.
   * Proof: This follows immediately from (5), since the elements of $\hat{G}$ are an orthonormal basis.

- Exercise: Prove Plancherel's theorem $\langle f_1, f_2 \rangle_G = \frac{1}{|G|} \left\langle \hat{f}_1, \hat{f}_2 \right\rangle_{\hat{G}}$ and deduce Parseval's theorem $\sum_{g \in G} |f(g)|^2 = \frac{1}{|G|} \sum_{\chi \in \hat{G}} \left| \hat{f}(\chi) \right|^2$.
- With the fundamentals taken care of, we can now focus on Dirichlet characters.
  * Studying primes congruent to a modulo m naturally leads to a question about Dirichlet characters via Fourier inversion, since we may decompose the characteristic function of [primes congruent to a modulo m] as a sum over Dirichlet characters for the group $G = (A/mA)^*$.

- Explicitly, if $\delta_a(p)$ is 1 when $p \equiv a \pmod{m}$ and 0 otherwise, then $\hat{\delta_a}(\chi) = \frac{1}{\Phi(m)} \sum_{g \in G} \delta_a(g)\overline{\chi(g)} = \frac{1}{\Phi(m)}\overline{\chi(a)}$, since the only nonzero value of $\delta_a(g)$ occurs when $g \equiv a \pmod{m}$.
- Then by Fourier inversion we have $\delta_a(p) = \sum_{\chi \in \hat{G}} \hat{\delta_a}(\chi)\chi(p) = \sum_{\chi \in \hat{G}} \frac{1}{\Phi(m)}\overline{\chi(a)}\chi(p)$. So the numerator for the Dirichlet density is $\sum_{p \equiv a \pmod{m}} |p|^{-s} = \sum_p \delta_a(p)\,|p|^{-s} = \frac{1}{\Phi(m)} \sum_{\chi \in \hat{G}} \left[\overline{\chi(a)} \sum_p \chi(p)\,|p|^{-s}\right]$.
- This is a bit complicated, but the point is that we have a sum over the Dirichlet characters of constants (namely $\overline{\chi(a)}$) times $\sum_p \frac{\chi(p)}{|p|^s}$, which is quite close to the Dirichlet series for the character χ (the only difference is that we are only summing over primes, rather than all monic polynomials).
- As we will see, we will be able to extract this sum over primes from the full Dirichlet series, which we now examine more closely.
- The main reason we go to this effort to use Fourier inversion is that the Dirichlet series for Dirichlet characters behave very nicely (far more nicely than the original series over primes congruent to a modulo m) because Dirichlet characters are multiplicative.

- Definition: If χ is a Dirichlet character modulo m, we define its associated Dirichlet L-series $L(s,\chi) = \sum_{f \text{ monic}} \frac{\chi(f)}{|f|^s}$.
  - Note that this is just the Dirichlet series for $\chi(f)$, as we defined it previously. It is traditional to denote these series with the letter L (which was the letter Dirichlet used for such functions).
  - As usual, the series converges absolutely for $\text{Re}(s) > 1$, since $|\chi(f)| \le 1$ for all f.
  - Furthermore, because Dirichlet characters are completely multiplicative, the L-series has a very simple Euler product: explicitly, $L(s,\chi) = \prod_{p \text{ prime}} \left[1 - \frac{\chi(p)}{p^s}\right]^{-1}$, for $\text{Re}(s) > 1$.
  - The Euler product is the key to calculating the Dirichlet density we wanted earlier: taking the logarithm of the Euler product gives $\log L(s,\chi) = -\sum_{p \text{ prime}} \log(1 - \chi(p)/p^s) \approx \sum_{p \text{ prime}} \frac{\chi(p)}{p^s}$ using the Taylor approximation $-\log(1-x) \approx x$ which is accurate for small $|x|$.
  - So our main task is to determine what happens to $\log L(s,\chi)$ as $s \to 1$, since this is the required input for calculating the Dirichlet density of the primes congruent to a modulo m.

0.8 (Oct 4) Primes in Arithmetic Progressions, Part 2

- Our main task is to determine what happens to $\log L(s,\chi)$ as $s \to 1$, since this is the required input for calculating the Dirichlet density of the primes congruent to a modulo m.
- Example: For the trivial character χ_{triv}, we have $L(s,\chi_{\text{triv}}) = \prod_{p|m \text{ prime}}(1 - |p|^{-s}) \cdot \zeta_A(s)$, since the terms with $p|m$ are missing from the Euler product for $L(s,\chi)$.
  - In particular, we see that $L(s,\chi_{\text{triv}})$ has an analytic continuation (since $\zeta_A(s)$ does) and a single simple pole at $s = 1$.
- For other characters, the L-series is essentially finite.
- Proposition (L-Series for Nontrivial Characters): Let m be a monic polynomial of positive degree and χ be a nontrivial Dirichlet character modulo m. Then $L(s,\chi)$ is a polynomial in q^{-s} of degree at most $\deg m - 1$, and in particular has an analytic continuation.
  - Proof: Let $A(n,\chi) = \sum_{\deg f = n} \chi(f)$ and note, as we have previously done in working out average-value results, that $L(s,\chi) = \sum_{n=0}^{\infty} A(n,\chi) q^{-ns}$. The claimed result is then equivalent to saying $A(n,\chi) = 0$ for $n \ge \deg m$.

- For this, suppose $\deg f = n \geq m$ and write $f = hm + r$ with $\deg r < \deg m$, where $\deg h = \deg f - \deg m$ and $\text{sgn}(h) = 1/\text{sgn}(m)$. Conversely, given such an h and r, we get a unique $f = hm + r$. Note that $\chi(f) = \chi(r)$, and also that there are $q^{n-\deg m}$ possible h.
- Then $A(n,\chi) = \sum_{\deg f = n} \chi(f) = \sum_{\deg f = n} \chi(r) = q^{n-\deg m} \sum_{\deg r < \deg m} \chi(r) = 0$ where the last sum is zero by the orthogonality relation (1).
- The observation about the analytic continuation is immediate (simply take the analytic continuation as the given polynomial in q^{-s}).

- As a consequence, we see that $L(s,\chi)$ has no pole at $s = 1$ when $\chi \neq \chi_{\text{triv}}$. Our next major goal is to prove that $L(1,\chi) \neq 0$ for $\chi \neq \chi_{\text{triv}}$.
- Lemma: Let χ be any Dirichlet character modulo m. Then for each monic irreducible p not dividing m, there exist $f_p, g_p > 0$ with $f_p g_p = \Phi(m)$ such that $\prod_{\chi \in \hat{G}} L(s,\chi) = \prod_{p \nmid m} (1 - |p|^{-f_p s})^{-g_p}$.
  - Proof: For a fixed monic irreducible $p \nmid m$, as we have previously noted the evaluation-at-p map $\chi \mapsto \chi(p)$ is a homomorphism from $\hat{G}$ to $\mathbb{C}^\times$.
  - Let the image be a cyclic group of order f_p and the kernel have size g_p: then $f_p g_p = \#\hat{G} = \#G = \Phi(m)$ by the first isomorphism theorem.
  - For this p, by grouping the fibers of the evaluation-at-p map together, for $\zeta = e^{2\pi i/f_p}$ we have $\prod_{\chi \in \hat{G}} (1 - \chi(p)/|p|^s)^{-1} = \prod_{j=0}^{f_p - 1} (1 - \zeta^j/|p|^s)^{-g_p}$, and this last product equals $(1 - |p|^{-f_p s})^{-g_p}$ since it is the evaluation of the polynomial $(1-t)(1 - e^{2\pi i/f_p} t) \cdots (1 - \zeta_{f_p}^{f_p - 1} t) = 1 - t^{f_p}$ at $t = |p|^{-s}$.
  - Thus, taking the product over all monic irreducibles $p \nmid m$ yields the claimed $\prod_{\chi \in \hat{G}} L(s,\chi) = \prod_{\chi \in \hat{G}} \prod_{p \nmid m} (1 - \chi(p)/|p|^s)^{-1} = \prod_{p \nmid m} (1 - |p|^{-f_p s})^{-g_p}$ after reversing the order of the products.

- We next show that $L(1,\chi) \neq 0$ for nonreal Dirichlet characters χ:
- Lemma (Nonvanishing, I): Let χ be any Dirichlet character modulo m such that $\chi \neq \overline{\chi}$. Then $L(1,\chi) \neq 0$.
  - Proof: If we expand the product $\prod_{\chi \in \hat{G}} L(s,\chi) = \prod_{p \nmid m} (1 - |p|^{-f_p s})^{-g_p}$ from the Lemma above, it yields a Dirichlet series with nonnegative coefficients and constant term 1.
  - Thus, if s is real and greater than 1 (so that the product converges), the value of the product is real and greater than 1.
  - If $\chi \neq \overline{\chi}$, then $\prod_{\chi \in \hat{G}} L(s,\chi) = L(s,\chi_{\text{triv}}) L(s,\chi) L(s,\overline{\chi}) \cdot$ [other terms].
  - Now suppose $L(1,\chi) = 0$: then we would have $L(1,\overline{\chi}) = 0$ also. But this would mean the product $\prod_{\chi \in \hat{G}} L(s,\chi)$ vanishes at $s = 1$, because the only term that has a pole at $s = 1$ is $L(s,\chi_{\text{triv}})$ and that pole has order 1, but we have two zeroes at $s = 1$ arising from $L(s,\chi)$ and $L(s,\overline{\chi})$.
  - But this is impossible because the value of the product is real and greater than 1 for $s > 1$. Thus, $L(1,\chi) \neq 0$.

- The case where $\chi = \overline{\chi}$ and $\chi \neq \chi_{\text{triv}}$ (i.e., when χ has order 2 in $\hat{G}$) is quite a bit trickier, since we cannot get away with such a simple order-of-vanishing argument.
- Lemma (Nonvanishing, II): Let χ be any Dirichlet character of order 2 modulo m (i.e., such that $\chi = \overline{\chi}$ but $\chi \neq \chi_{\text{triv}}$). Then $L(1,\chi) \neq 0$.
  - Proof: Suppose that $\chi = \overline{\chi}$ but $\chi \neq \chi_{\text{triv}}$, so that $\chi(p) \in \{\pm 1\}$ for $p \nmid m$, and define the function $G(s) = \dfrac{L(s,\chi_{\text{triv}}) L(s,\chi)}{L(2s,\chi_{\text{triv}})} = \prod_{p \nmid m} \dfrac{(1 - |p|^{-s})^{-1} (1 - \chi(p)\,|p|^{-s})^{-1}}{(1 - |p|^{-2s})^{-1}} = \prod_{p \nmid m} \dfrac{1 + |p|^{-s}}{1 - \chi(p)\,|p|^{-s}} = \prod_{p \nmid m, \chi(p) = 1} \dfrac{1 + |p|^{-s}}{1 - |p|^{-s}} = \prod_{p \nmid m, \chi(p) = 1} [1 + \sum_{k=1}^{\infty} |p|^{-ks}]$.
  - By expanding this last expression for G, we can see that its Dirichlet series has all coefficients nonnegative.
  - We also have $\dfrac{L(s,\chi_{\text{triv}})}{L(2s,\chi_{\text{triv}})} = \dfrac{\zeta_A(s)}{\zeta_A(2s)} \cdot \prod_{p \mid m} \dfrac{1 - |p|^{-s}}{1 - |p|^{-2s}} = \dfrac{1 - q^{1-2s}}{1 - q^{1-s}} \prod_{p \mid m} (1 + |p|^{-s})^{-1}$. Substituting this into the expression for G yields that $\dfrac{1 - q^{1-2s}}{1 - q^{1-s}} L(s,\chi) = \dfrac{L(s,\chi_{\text{triv}}) L(s,\chi)}{L(2s,\chi_{\text{triv}})} \prod_{p \mid m} (1 + |p|^{-s})^{-1} = G(s) \prod_{p \mid m} (1 + |p|^{-s})^{-1}$ is a Dirichlet series with all coefficients nonnegative.

- Suppose $G(s)\prod_{p|m}(1+|p|^{-s})^{-1} = \sum_{f \text{ monic}} \dfrac{h(f)}{|f|^s}$.
- Rewriting in terms of $u = q^{-s}$, and noting that $L^*(u,\chi) = L(s,\chi)$ is a polynomial in u as we proved earlier, we obtain the equality $\dfrac{1-qu^2}{1-qu}L^*(u,\chi) = \sum_{d=0}^{\infty}[\sum_{\deg(f)=d} h(f)]u^d$.
- Now suppose that $L(1,\chi) = L^*(q^{-1},\chi)$ is equal to zero. Then $1-qu$ would divide $L^*(u,\chi)$, which would mean that $\dfrac{1-qu^2}{1-qu}L^*(u,\chi)$ is a polynomial in u. But then the right-hand side would also be a polynomial in u. All of its coefficients are nonnegative (as noted above), which means it cannot have a positive root for u.
- But, finally, notice that $\dfrac{1-qu^2}{1-qu}L^*(u,\chi)$ is zero when $u = 1/\sqrt{q}$. This is a contradiction, and so $L^*(q^{-1},\chi) = L(1,\chi)$ must be nonzero.

* Now that we know $L(1,\chi)$ vanishes for nontrivial characters χ, we can prove Dirichlet's theorem:
* Theorem (Analogue of Dirichlet's Theorem): Let $m \in \mathbb{F}_q[t]$ have positive degree and let a be relatively prime to m. Then the Dirichlet density of the set of primes congruent to a (mod m) exists and is $1/\Phi(m)$. In particular, there are infinitely many such primes.
  - We have already obtained all of the necessary ingredients, so the proof is mostly a matter of putting them all together.
  - Proof: Recall the power series $-\log(1-x) = \sum_{k=1}^{\infty} x^k/k$, valid for $|x| < 1$.
  - Then for any Dirichlet character χ, we have $\log L(s,\chi) = \sum_p -\log(1-\dfrac{\chi(p)}{|p|^s}) = \sum_p \left[\sum_{k=1}^{\infty} \dfrac{\chi(p)^k}{k} |p|^{-ks}\right] = \sum_p \dfrac{\chi(p)}{|p|^s} + \sum_p \sum_{k=2}^{\infty} \dfrac{\chi(p)^k}{k} |p|^{-ks}$. The absolute value of the second term is bounded by $\sum_p \sum_{k=2}^{\infty} \dfrac{1}{k} |p|^{-ks} \le \sum_{k=2}^{\infty}\sum_{d=1}^{\infty} q^d q^{-kds} \le \sum_{n=1}^{\infty}(n+1)q^{-ns}$, which is bounded as $s \to 1+$.
  - Therefore, as $s \to 1+$, we have $\log L(s,\chi) = \sum_p \dfrac{\chi(p)}{|p|^s} + O(1)$. In particular, we see that $\sum_p |p|^{-s} = \log(s-1) + O(1)$ as $s \to 1+$, since $L(s,\chi_{\text{triv}})$ has a simple pole at $s = 1$.
  - Now, by Fourier inversion (as we previously worked out) we have $\sum_{p \equiv a \pmod{m}} |p|^{-s} = \sum_p \delta_a(p)|p|^{-s} = \dfrac{1}{\Phi(m)} \sum_{\chi \in \hat{G}} \left[\overline{\chi(a)} \sum_p \dfrac{\chi(p)}{|p|^s}\right]$.
  - So, the quotient for the Dirichlet density is $\dfrac{\sum_{p \equiv a \pmod{m}} |p|^{-s}}{\sum_p |p|^{-s}} = \dfrac{\frac{1}{\Phi(m)} \sum_{\chi\in\hat{G}} \left[\overline{\chi(a)} \sum_p \frac{\chi(p)}{|p|^s}\right]}{\sum_p |p|^{-s}} = \dfrac{1}{\Phi(m)}\left[\dfrac{\sum_{p\nmid m}|p|^{-s}}{\sum_p |p|^{-s}} + \dfrac{\sum_{\chi\ne\chi_{\text{triv}}} \overline{\chi(a)}\sum_p \frac{\chi(p)}{|p|^s}}{\sum_p |p|^{-s}}\right] = \dfrac{1}{\Phi(m)}\left[1 - \dfrac{\sum_{p|m}|p|^{-s}}{\log(s-1)+O(1)} + \dfrac{\sum_{\chi\ne\chi_{\text{triv}}} \log L(s,\chi) + O(1)}{\log(s-1)+O(1)}\right]$.
  - Now, taking the limit as $s \to 1+$ makes the second term go to zero (since the numerator is finite) and the third term go to zero (since $L(1,\chi) \ne 0$ for $\chi \ne \chi_{\text{triv}}$), and so the value of the limit is just $1/\Phi(m)$, as claimed.
* We can, in fact, improve this argument to show that the natural density of the primes congruent to a modulo m is equal to $1/\Phi(m)$, not just the Dirichlet density.
  - To do this requires showing that $L(s,\chi)$ is zero-free on a larger region: specifically, we need it to be zero-free for $\text{Re}(s) = 1$, rather than just $s = 1$.
  - The L-function is in fact zero-free on a much larger region: as we will eventually prove, the only zeroes of $L(s,\chi)$ are on the line $\text{Re}(s) = 1/2$; this is the Riemann hypothesis for function fields.

- Taking this zero-free result for granted, we again need to manipulate the series expressions for the $L(s,\chi)$. This time, we will use in a more substantial way the fact that the $L(s,\chi)$ for $\chi \neq \chi_{\text{triv}}$ are polynomials in $u = q^{-s}$ and compare the Euler products with their factorizations.

- Theorem (Strengthened Dirichlet Analogue): Let $m \in \mathbb{F}_q[t]$ have positive degree and let a be relatively prime to m. Then the number of primes congruent to a (mod m) having degree N is equal to $\frac{1}{\Phi(m)}\frac{q^N}{N} + O(\frac{q^{N/2}}{N})$, where the implied constant is independent of q and N. In particular, the proportion of primes of degree N congruent to a (mod m) is $\frac{1}{\Phi(m)} + O(q^{-N/2})$.
  - If we only know that the L-function is zero free for $\text{Re}(s) > \theta$ for some $\theta \in (1/2, 1)$, we instead get an error term of $O(\frac{q^{\theta N}}{N})$, which is still good enough to establish that the natural density of primes congruent to a (mod m) equals $1/\Phi(m)$.
  - Proof: For convenience, we first note the identity (*) $u\frac{\partial}{\partial u}\log(1-\alpha u^d)^{-1} = \sum_{N=1}^{\infty} d\alpha^k u^{dN}$.
  - As we showed previously, if $\chi \neq \chi_{\text{triv}}$ then $L^*(u,\chi) = L(q^{-s},\chi)$ is a polynomial in $u = q^{-s}$ of degree at most $m-1$. Since its constant term is 1, we obtain a factorization of the form $L^*(u,\chi) = \prod_{i=1}^{m-1}(1 - \alpha_i(\chi)u)$ for some constants $\alpha_i(\chi) \in \mathbb{C}$.
  - From the Euler product, we also have $L^*(u,\chi) = \prod_{p\nmid m}(1-\chi(p)u^{\deg p})^{-1} = \prod_{d=1}^{\infty}\prod_{p\nmid m,\deg p = d}(1 - \chi(p)u^d)^{-1}$.
  - Now apply the operator $u\frac{\partial}{\partial u}\log$ to the equality $\prod_{i=1}^{m-1}(1-\alpha_i(\chi)u) = \prod_{d=1}^{\infty}\prod_{p\nmid m,\deg p=d}(1-\chi(p)u^d)^{-1}$ and compare coefficients of u on both sides.
  - For the LHS, using the identity (*) with $d = 1$ yields $u\frac{\partial}{\partial u}\log L^*(u,\chi) = -\sum_{i=1}^{m-1}\sum_{N=1}^{\infty}\alpha_i(\chi)^N u^N = -\sum_{N=1}^{\infty}\left[\sum_{i=1}^{m-1}\alpha_i(\chi)^N\right]u^N$.
  - Letting $c_N(\chi) = -\sum_{i=1}^{m-1}\alpha_i(\chi)^N$ yields the expansion $u\frac{\partial}{\partial u}\log\prod_{i=1}^{m-1}(1-\alpha_i(\chi)u) = \sum_{N=1}^{\infty}c_N(\chi)u^N$. For $\chi = \chi_{\text{triv}}$, we have $c_N(\chi) = q^N + O(1)$, while for $\chi \neq \chi_{\text{triv}}$, by the Riemann hypothesis we have $|\alpha_i(\chi)| \in \{q^0, q^{1/2}\}$ for each i, and so $c_N(\chi) = O(q^{N/2})$.
  - For the RHS, we have $u\frac{\partial}{\partial u}\log L^*(u,\chi) = \sum_{d=1}^{\infty}\sum_{p\nmid m,\deg p=d}u\frac{\partial}{\partial u}\log(1-\chi(p)u^d)^{-1} = \sum_{d=1}^{\infty}\sum_{p\nmid m,\deg p=d}\sum_{k=1}^{\infty}d\,\chi(p)^k u$ $\sum_{N=1}^{\infty}\left[\sum_{d\mid N}\sum_{\deg p=N/d}d\,\chi(p)^{N/d}\right]u^N$ by applying the identity (*) and then grouping together all of the terms of the same degree. This means $c_N(\chi) = \sum_{d\mid N}\sum_{\deg p=d}d\,\chi(p)^d$.
  - Now, by separating out the terms with $d = 1$ from the others, we see $c_N(\chi) = \sum_{d\mid N}\sum_{\deg p=N/d}d\,\chi(p)^{N/d} = N\sum_{\deg p=N}\chi(p) + \sum_{d\mid N,d\geq 2}\sum_{\deg p=N/d}d\,\chi(p)^d$. The absolute value of the second term is at most $\sum_{d\mid N,d\geq 2}\sum_{\deg p=N/d}d \leq \sum_{d\mid N,d\geq 2}\frac{q^{N/d}}{N/d} = O(q^{N/2})$.
  - Therefore, we see $c_N(\chi) = N\sum_{\deg p=N}\chi(p) + O(q^{N/2})$.
  - Now we use our Fourier decomposition from earlier: we have $\frac{1}{\Phi(m)}\sum_{\chi\in\hat{G}}\overline{\chi(a)}c_N(\chi) = N\cdot\#\{\text{primes } p \equiv a \text{ (mod } m)\} + O(q^{N/2})$ using the expression we just computed.
  - Also, we have $\sum_{\chi\in\hat{G}}\overline{\chi(a)}c_N(\chi) = q^N + O(q^{N/2})$ by directly summing over characters: $\chi = \chi_{\text{triv}}$ contributes the q^N term and the other characters each contribute $O(q^{N/2})$.
  - Setting these two equal to one another yields $\#\{\text{primes } p \equiv a \text{ (mod } m)\} = \frac{1}{\Phi(m)}\cdot\frac{q^N}{N} + O(\frac{q^{N/2}}{N})$, as claimed.

0.9 (Oct 7) Homework #1 Discussion

0.10 (Oct 14) Preliminaries: Transcendence and Localization

- We now move into the second major part of the course, which deals with algebraic function fields: these are function fields of transcendence degree 1 over a general constant field F.

- Later, we will specialize to function fields over $\mathbb{F}_q$, which along with algebraic number fields constitute the global fields.
- Global fields (to be considered in contrast with local fields) share a number of common properties that we will elucidate and study.

- We begin by reviewing some basic facts about transcendental extensions.

- Definition: Let K/F be a field extension. We say a subset S of K is algebraically dependent over F if there exists a finite subset $\{s_1, \dots, s_n\} \in S$ and a nonzero polynomial $p \in F[x_1, \dots, x_n]$ such that $p(s_1, \dots, s_n) = 0$. If there exists no such p for any finite subset of S, we say S is algebraically independent.
  - The general idea here is that a set of elements is algebraically dependent if they satisfy some algebraic (i.e., polynomial) relation over F.
  - Example: Over $\mathbb{Q}$, the set $\{\pi, \pi^2\}$ is algebraically dependent, since $p(x, y) = x^2 - y$ has $p(\pi, \pi^2) = 0$.
  - Example: Over $\mathbb{Q}$, the set $\{\sqrt[3]{2}\}$ is algebraically dependent, since $p(x) = x^3 - 2$ has $p(\sqrt[3]{2}) = 0$.
  - More generally, the set $\{\alpha\}$ is algebraically independent over F if and only if α is transcendental over F.
  - Example: Over $\mathbb{R}$, the set $\{x + y, x^2 + y^2\}$ is algebraically independent. (Exercise: prove this.)
  - Example: Over $\mathbb{R}$, the set $\{x+y, x^2+y^2, x^3+y^3\}$ is algebraically dependent, since $p(a, b, c) = a^3 - 3ab + 2c$ has $p(x + y, x^2 + y^2, x^3 + y^3) = 0$.
  - Example: If $x_1, \dots, x_n$ are indeterminates inside $F(x_1, \dots, x_n)$, the function field in n variables, then the set $\{x_1, \dots, x_n\}$ is algebraically independent over F.

- The notion of algebraic independence generalizes the notion of linear independence, and as such the two concepts are related in various ways.
  - It is easy to see that any subset of an algebraically independent set is algebraically independent, while any set containing an algebraically dependent set is algebraically dependent.
  - Since having a basis of a vector space is very convenient for calculations, we might therefore hope to define an analogous "transcendence basis" to be an algebraically independent set that generates the extension K/F.
  - Unfortunately, such a set need not exist: for example, $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ has no such set, because there are no transcendental elements at all.
  - The correct analogy is instead to observe that a basis for a vector space is a maximal linearly independent set:

- Definition: Let K/F be a field extension. A transcendence base for K/F is an algebraically independent subset S of K that is maximal in the set of all algebraically independent subsets of K.
  - Remark: The term "transcendence basis" is also used occasionally.
  - By a straightforward Zorn's lemma argument, every extension has a transcendence base.
  - Example: The empty set $\emptyset$ is a transcendence base for $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$. More generally, K/F is algebraic if and only if $\emptyset$ is a transcendence base.
  - Example: The set $\{x\}$ is a transcendence base for $F(x)$ over F.

- Here are some of the fundamental properties of transcendence bases, many of which are analogous to properties of vector spaces:

- Proposition (Transcendence Bases): Suppose K/F is a field extension and S is a subset of K.
  1. If S is algebraically independent and $\alpha \in K$, then $S \cup \{\alpha\}$ is algebraically independent over F if and only if α is transcendental over $F(S)$.
     - This is the algebraic analogue of the statement that if S is linearly independent, then $S \cup \{\alpha\}$ is linearly independent if and only if α is not in the span of S.

   - Proof: Suppose $S \cup \{\alpha\}$ is algebraically dependent. Then there exists $s_i \in S$ and $p \in F[x]$ with $p(\alpha, s_1, \ldots, s_n) = 0$ and $p \neq 0$. View p as a polynomial in its first variable with coefficients in $F[s_1, \ldots, s_n]$: there must be at least one term involving α, as otherwise p would give an algebraic dependence in S. Then α is the root of a nonzero polynomial with coefficients in $F[s_1, \ldots, s_n] \subseteq F(s_1, \ldots, s_n) \subseteq F(S)$, so it is algebraic over $F(S)$.
   - Conversely, suppose that α is algebraic over $F(S)$. Then α is the root of some nonzero polynomial with coefficients in $F(S)$. Each coefficient of this polynomial is an element of $F(S)$; clearing denominators yields a nonzero polynomial p with coefficients in $F[s_1, \ldots, s_n]$ for the elements $s_i \in S$ that appear in these coefficients. This polynomial yields an algebraic dependence in $S \cup \{\alpha\}$.
2. S is a transcendence base of K/F if and only if K is algebraic over $F(S)$.
   - Proof: This follows from (1) and the maximality of transcendence bases: S is a transcendence base if and only if no elements in K can be adjoined to S while preserving algebraic independence, and by (1) this is equivalent to saying that all elements in K are algebraic over $F(S)$.
3. If T is a subset of K such that $K/F(T)$ is algebraic, then T contains a transcendence base of K/F.
   - Proof: Apply Zorn's lemma to the collection of all algebraically independent subsets of T, partially ordered by inclusion.
   - A maximal element M in this collection must then be a transcendence base for K/F: if $\beta \in K$ then β must be algebraic over $K/F(M)$ by the maximality of M, and then M is a transcendence base by (2).
4. If T is an algebraically independent subset of K, then T can be extended to a transcendence base of K/F.
   - Proof: This is the analogue of the fact that every linearly independent subset can be extended to a basis, and the proof follows from a similar Zorn's lemma argument.
5. If $S = \{s_1, \ldots, s_n\}$ is a transcendence base for K/F and $T = \{t_1, \ldots, t_m\}$ is any algebraically independent set, then there is a reordering of S, say $\{a_1, \ldots, a_n\}$, such that for each $1 \leq k \leq m$, the set $\{t_1, t_2, \ldots, t_k, a_{k+1}, \ldots, a_n\}$ is a transcendence base for K/F.
   - Proof: This is the analogue of the replacement theorem for linearly independent sets, and the proof proceeds inductively in essentially the same way. (We will omit the details.)
6. If S is a (finite) transcendence base for K/F, then any subset T of K having larger cardinality than S must be algebraically dependent.
   - Proof: If $S = \{s_1, \ldots, s_n\}$ is finite, apply the replacement theorem (5) to S and T. At the end of the replacement, the result is that $\{t_1, \ldots, t_n\}$ is a transcendence base. But then by (2), any additional element of T would be algebraic over $\{t_1, \ldots, t_n\}$, contradicting the algebraic independence of T.
7. Any two transcendence bases S and T for K/F have the same cardinality.
   - Proof: If the bases are infinite the result is immediate. If S has finite cardinality n, then the result follows by applying (6), since then T's cardinality m must satisfy $m \leq n$ (since T is algebraically independent and S is a transcendence base) and also $n \leq m$ (since S is algebraically independent and T is a transcendence base).

- The result of the last part of the proposition shows that any two transcendence bases have the same cardinality, and in analogy with the situation for vector spaces, this cardinality behaves somewhat like an extension degree:
- Definition: Let K/F be a field extension. The transcendence degree of K/F, denoted $\text{trdeg}(K/F)$, is the cardinality of any transcendence base of K/F.
- The key property of transcendence degree is that it is additive in towers:
- Theorem (Transcendence in Towers): If $L/K/F$ is a tower of extensions, then $\text{trdeg}(L/F) = \text{trdeg}(L/K) + \text{trdeg}(K/F)$.
   - The idea here is quite simple: we want to show that the union of transcendence bases for K/F and L/K gives a transcendence base for L/F.

- Proof: First suppose that both $\mathrm{trdeg}(K/F)$ and $\mathrm{trdeg}(L/K)$ are finite, and let $S = \{s_1, \ldots, s_n\}$ and $T = \{t_1, \ldots, t_m\}$ be transcendence bases for K/F and L/K. Then $S \cap T = \emptyset$ since each t_i is transcendental over K.
- Furthermore, K is algebraic over $F(S)$, so $K(T)$ is algebraic over $F(T)(S) = F(S \cup T)$ by our results on algebraic extensions.
- Then since L is algebraic over $K(T)$, we deduce that L is algebraic over $F(S \cup T)$, also by our results on algebraic extensions.
- Thus, by property (3) above, $S \cup T$ contains a transcendence base of L/F.
- Finally, we claim $S \cup T$ is algebraically independent over F, so suppose that $p(s_1, \ldots, s_n, t_1, \ldots, t_m) = 0$ for some $p \in F[x_1, \ldots, x_n, y_1, \ldots, y_m]$.
- Separate monomial terms to write $p(s_1, \ldots, s_n, t_1, \ldots, t_m) = 0$ as a sum $\sum f_i(s_1, \ldots, s_n) g_i(t_1, \ldots, t_m) = 0$ with $f_i \in F[x_1, \ldots, x_n]$ and $g_i \in F[y_1, \ldots, y_m]$.
- Now, since T is algebraically independent over $F(S) \subseteq K$, all of the $f_i(s_1, \ldots, s_n)$ must be zero (as elements of K). But since S is algebraically independent over F, that means all of the polynomials $f_i(x_1, \ldots, x_n)$ must be zero (as polynomials).
- This means p is the zero polynomial, and so $S \cup T$ is algebraically independent.

- Fields that are generated by a transcendence base are particularly convenient:
- Definition: The extension K/F is purely transcendental if $K = F(S)$ for some transcendence base S of K/F.
  - Equivalently, K/F is purely transcendental when it is generated (as a field extension) by an algebraically independent set.
  - If $S = \{s_1, \ldots, s_n\}$, then the purely transcendental extension $K = F(S)$ is ring-isomorphic to the function field $F(x_1, \ldots, x_n)$ in n variables: it is not hard to check that the map sending s_i to x_i is an isomorphism.
  - If K/F has transcendence degree 1 or 2 and E/F is an intermediate extension, then in fact E is also purely transcendental: the degree-1 case is a theorem of Lüroth that we will prove later, while the degree-2 case is a theorem of Castelnuovo. In higher degrees, there do exist extensions that are not purely transcendental, but it is not easy to verify this fact.
- Now let F be a field and K be an extension of F of transcendence degree 1.
  - By the results above, there exists $x \in K$ such that $K/F(x)$ has transcendence degree 0, which is to say, it is algebraic.
  - Since we do not want to worry for the moment about infinite-degree algebraic extensions, we will make the further assumption that this extension $K/F(x)$ has finite degree.
- Definition: We say K is an (algebraic) function field over F if there exists $x \in K$ such that x is transcendental over F and $K/F(x)$ is finite.
  - Example: $\mathbb{Q}(x)$ is an algebraic function field over $\mathbb{Q}$.
  - Example: $\mathbb{C}(x, \sqrt{x^2 - 1})$ is an algebraic function field over $\mathbb{C}$.
  - Note that the algebraic closure of F inside K has finite degree over F: this follows by noting that if E/F is algebraic inside K, then $[E : F] = [E(x) : F(x)] \leq [K : F(x)] < \infty$.
  - So, without loss of generality, we may replace F by its algebraic closure inside K. In this case we call F the constant field of K.
  - If F is the constant field of K, then there are no elements of K that are algebraic over F other than the elements of F themselves. Equivalently, every element of $K \backslash F$ is transcendental over F.
  - Finally, since the transcendence degree of K/F is 1, for any two $a, b \in K \backslash F$, there is some nonzero polynomial $g \in F[x, y]$ such that $F[a, b] = 0$.
- Now that we have some very basic facts about function fields, our next goal is to do number theory.
  - In order to do this, however, we need to know how to define primes in the function field context.

- Over $\mathbb{Q}$, the primes arise as the prime ideals of the ring of integers $\mathbb{Z}$, which we can define starting from $\mathbb{Q}$ purely in terms of integral closures. For other number fields, we also define their primes using integral closures.
- However, this approach will not work for function fields, because (as noted above) everything in K not in F is transcendental over F, so there is no sensible way to define a "ring of integers" inside K using integrality.
- Instead, we need to use a different sort of construction to give a sensible notion of a prime: that of a discrete valuation on K.

- In order to develop all of this properly, we also need to review some facts about localization.
- Proposition (Localization): Let R be a commutative ring with 1 and D be a multiplicatively closed subset of R containing 1. Then there exists a commutative ring $D^{-1}R$, the localization of R at D, and a ring homomorphism $\pi : R \to D^{-1}R$ such that any for any ring homomorphism $\psi : R \to S$ sending 1 to 1 and such that $\psi(d)$ is a unit in S for every $d \in D$, there exists a unique homomorphism $\Psi : D^{-1}R \to S$ such that $\Psi \circ \pi = \psi$.
  - More succinctly, any homomorphism $\psi : R \to S$ such that ψ maps all of the elements of D into units necessarily extends to a homomorphism $\Psi : D^{-1}R \to S$.
  - The main idea is simply to define "fractions" r/d with $r \in R$ and $d \in D$ via an appropriate equivalence relation, and then to write down the usual rules of fraction arithmetic and verify that all of the definitions are well posed.
  - Proof (outline): Define an equivalence relation on elements of $R \times D$ by setting $(r, d) \sim (s, e)$ whenever there exists $y \in D$ such that $y(ds - er) = 0$; it is straightforward to check that $\sim$ is an equivalence relation.
  - Denote the equivalence class of (r, d) by the symbol r/d and the set of all equivalence classes by $D^{-1}R$, and define the two operations $r/d + s/e = (re + ds)/(de)$ and $r/d \cdot s/e = (rs)/(de)$ on $D^{-1}R$. It is a tedious but straightforward check to see that these operations make $D^{-1}R$ into a commutative ring with 1.
  - Now define $\pi(r) = r/1$ and suppose $\Psi : D^{-1}R \to S$ is a homomorphism with $\Psi \circ \pi = \psi$.
  - Then we must have $\Psi(r/1) = (\Psi \circ \pi)(r) = \psi(r)$, and also $1 = \Psi(1/1) = \Psi(1/d)\Psi(d/1)$, meaning that $\Psi(1/d) = \psi(d)^{-1}$. Then $\Psi(r/d) = \Psi(r/1)\Psi(1/d) = \psi(r)\psi(d)^{-1}$.
  - But it is easy to see that this choice of Ψ does work, so it is the only such homomorphism.
- The point here is that $D^{-1}R$ is the smallest ring in which all elements of D become units.
  - When D contains no zero divisors (which is automatically the case if R is a domain and D does not contain zero), then R injects into $D^{-1}R$ via $r \mapsto r/1$.
  - A particular useful case of localization is to construct $\mathbb{Q}$ from $\mathbb{Z}$ (we take $D = \mathbb{Z}\backslash\{0\}$ and $R = \mathbb{Z}$) or more generally to construct the field of fractions of an integral domain R (take $D = R\backslash\{0\}$).
- We also note in passing that we can localize any R-module M in the same way: one simply writes down the same construction using pairs (m, d) with $m \in M$ and $d \in D$ in place of pairs (r, d).
  - Alternatively, one can obtain the localization of an R-module using tensor products: $D^{-1}M \cong M \otimes_R D^{-1}R$. (Morally, this tensor product just extends scalars from R to $D^{-1}R$, which is exactly what $D^{-1}M$ is.)
  - Exercise: Show that localization commutes with sums, intersections, quotients, finite direct sums, and is exact.
  - Exercise: Show that if I is an ideal of R, then $D = R\backslash I$ is multiplicatively closed if and only if I is prime.
- Our main situation of interest is that of localizing at a prime: this is the case where R is an integral domain and $D = R\backslash P$ is the complement of a prime ideal P of R.
  - Exercise: Show that if P is a prime ideal and $D = R\backslash P$, then $D^{-1}R$ is a local ring with unique maximal ideal $D^{-1}P = \pi(P) = e_P$, the extension of the ideal P to $D^{-1}R$.

- ∘ The utility of localizing at a prime is that it isolates the ring's behavior at that prime.
  - ∘ Example: The localization of $\mathbb{Z}$ at the prime ideal (p) is the ring $\mathbb{Z}_{(p)} = \{\frac{a}{b} \in \mathbb{Q} : p \nmid b\}$ of rational numbers whose denominator is not divisible by p. Its unique maximal ideal is $p\mathbb{Z}_{(p)}$, the set of multiples of p. The quotient ring $\mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)}$ is isomorphic to $\mathbb{Z}/p\mathbb{Z}$.
  - ∘ Note that $\mathbb{Z}_{(p)}$ is not the ring of p-adic integers $\mathbb{Z}_p$: the p-adic integers are obtained by taking a completion of the localization $\mathbb{Z}_{(p)}$ under the p-adic metric (which we will define later).
  - ∘ Example: Let k be a field and take R to be the ring of k-valued functions on a set S. If we let M_a be the set of functions vanishing at a point $a \in S$, then M_a is a maximal ideal of R. The localization $R_{M_a} = \{\frac{f}{g} \in R : g(a) \neq 0\}$ is the ring of k-valued rational functions defined at a. The unique maximal ideal of M_a is the ideal of all k-valued rational functions vanshing at a.
- This second example illustrates the utility of localizing at a prime, because it allows us to study the local behavior of a rational function near the point a.
  - ∘ For example, the elements of M_a are precisely those rational functions vanishing at a, while the elements of M_a^2 are the rational functions that vanish to order 2 at a (i.e., have a double root), and so forth.
  - ∘ More generally, if we localize a domain at a principal prime ideal, by looking at powers of the maximal ideal, we can measure what power of a prime a given element is divisible by.
- We will formalize all of this using discrete valuations, which provide us a way to identify primes using only the field structure, next time.

0.11 (Oct 18) Discrete Valuations and Divisors in Function Fields

- Definition: Let F be a field. A discrete valuation on F is a surjective function $v : F^\times \to \mathbb{Z}$ such that $v(ab) = v(a) + v(b)$ for all $a, b \in F$ and $v(a + b) \geq \min(v(a), v(b))$ for all $a, b \in F^\times$ with $a + b \neq 0$. The set $R = \{r \in F^\times : v(r) \geq 0\} \cup \{0\}$ is called the valuation ring of v.
  - ∘ For convenience, if v is a discrete valuation we often also write $v(0) = \infty$, in which case we can ignore the various exceptions in the definition above (e.g., $v(a + b) \geq \min(v(a), v(b))$ now holds for all a, b).
  - ∘ In general, we say an integral domain R is a discrete valuation ring (DVR) if it is the valuation ring for some discrete valuation on its field of fractions.
  - ∘ Example: For a fixed prime p, the p-adic valuation on $\mathbb{Q}$, which has $v_p(p^n \frac{r}{s}) = n$ for $p \nmid r, s$, is a discrete valuation. (For example, $v_2(4) = 2$, $v_2(\frac{1}{3}) = 0$, and $v_2(\frac{3}{4}) = -2$: the valuation simply gives the power of p in a rational number.) The associated valuation ring is the set of rational numbers whose denominator is not divisible by p: this is $\mathbb{Z}_{(p)}$, the localization of $\mathbb{Z}$ at (p).
  - ∘ Example: For a fixed irreducible polynomial p, the p-adic valuation on $\mathbb{F}_q(t)$, which has $v_p(p^n \frac{r}{s}) = n$ for $p \nmid r, s$, is a discrete valuation. (For example, $v_t(t^3) = 3$, $v_t(\frac{t}{t+1}) = 1$, and $v_{t+1}(\frac{t}{t+1}) = -1$.) The associated valuation ring is the set of rational functions whose denominator is not divisible by p: this is $A_{(p)}$, the localization of $A = \mathbb{F}_q[t]$ at (p).
  - ∘ In the two examples above, the valuation rings are both obtained as localizations. We can in fact construct DVRs by localizing in more generality.
  - ∘ Exercise (Corollary 8 from Section 16.2 of Dummit/Foote): If R is a Noetherian integrally-closed domain and P is a minimal nonzero prime ideal of R, then R_P is a DVR.
- Proposition (Properties of DVRs): Let R be a discrete valuation ring with field of fractions F and valuation v. Also $t \in R$ be any element with $v(t) = 1$ (such an element is called a uniformizer). Then the following hold:

  1. For any $r \in F^\times$, either r or $1/r$ is in R.

2. An element $u \in R$ is a unit of R if and only if $v(u) = 0$. In particular, if $\zeta \in F$ is any root of unity, then $v(\zeta) = 0$.
3. If $x \in R$ is nonzero and $v(x) = n$, then x can be written uniquely in the form $x = ut^n$ for some unit $u \in R$.
4. Every nonzero ideal of R is of the form (t^n) for some $n \geq 0$.
5. The ring R is a Euclidean domain (hence also a PID and a UFD) and also a local ring.
6. The ring S is a DVR if and only if it is a PID and a local ring but not a field.
   - Proofs: Exercises.

- We will also remark that a discrete valuation v on a field F naturally makes F into a metric space using the non-Archimedean metric $d_v(a, b) = 2^{-v(a-b)}$. Explicitly:

1. We clearly have $d_v(a, b) \geq 0$ with equality if and only if $a = b$.
2. Since $v(-1) = 0$ by (2) in the proposition above, we have $v(a - b) = v(b - a)$ and thus $d_v(a, b) = 2^{-v(a-b)} = 2^{-v(b-a)} = d_v(b, a)$.
3. From $v(x+y) \geq \min(v(x), v(y))$ we have $v(a-b) \geq \min(v(a-c), v(c-b))$, so negating yields $-v(a-b) \leq \max(-v(a-c), -v(c-b))$. Then $d_v(a, b) = 2^{-v(a-b)} \leq \max(2^{-v(a-c)}, 2^{-v(c-b)}) = \max(d_v(a, c), d_v(c, b))$.

   - We could also replace 2 by any real number greater than 1 in the definition of the metric without affecting anything.
   - With this metric, we can then speak fruitfully of Cauchy sequences, write down the metric topology on F, and take completions. (Completing $\mathbb{Q}$ under the p-adic metric yields the p-adic field $\mathbb{Q}_p$, while the completion of its valuation ring $\mathbb{Z}$ yields the ring of p-adic integers $\mathbb{Z}_p$.)

- We now have enough background to discuss primes in function fields. The point of all of these preliminaries is that there is a natural interplay between discrete valuations on F and the primes associated to F, at least in the case of $F = \mathbb{Q}$.

- Definition: If K is a function field over F, a prime P of K is the maximal ideal of a discrete valuation ring R containing F whose field of fractions is K. The associated valuation on K is denoted ord_P.

   - Explicitly, the idea is that if we have a discrete valuation on K, then the valuation ring R is a local ring whose unique maximal ideal represents a prime of K.

- It is worth going through why this definition is (up to some haziness) really the same as the usual one in the case $K = \mathbb{Q}$ that we already understand.

   - If we have a discrete valuation v on $\mathbb{Q}$, then $v(-1) = v(1) = 0$ and so $v(n) \geq 0$ for all integers n. This means that the valuation ring R contains $\mathbb{Z}$.
   - Exercise: If v is a discrete valuation on $\mathbb{Q}$, the set $P = \{n \in \mathbb{Z} : v(n) > 0\}$ is a prime ideal of $\mathbb{Z}$.
   - By the exercise, $P = (p)$ for some prime p. Then $v(a) = 0$ for $p \nmid a$, so if $v(p) = r$ we see $v(p^n \frac{a}{b}) = rn$ for $p \nmid a, b$. Since discrete valuations are onto and $v(p) > 0$, we must have $n = 1$, and so v is the usual p-adic valuation on $\mathbb{Q}$.
   - Therefore, the only discrete valuations on $\mathbb{Q}$ are the p-adic valuations. The corresponding valuation ring is then $\mathbb{Z}_{(p)}$ with unique maximal ideal $p\mathbb{Z}_{(p)}$.
   - We see that for each integer prime p, we obtain a unique prime ideal $P = p\mathbb{Z}_{(p)}$ inside the associated a valuation ring of $\mathbb{Q}$. The collection of valuations v_p, evaluated on a rational number $\alpha \in \mathbb{Q}$, measures "how divisible" the element α is by each of the primes p.
   - Note also that in this case, the quotient of the valuation ring $R = \mathbb{Z}_{(p)}$ by its maximal ideal $P = p\mathbb{Z}_{(p)}$ is isomorphic to $\mathbb{Z}/p\mathbb{Z}$, which has cardinality p: the size of this quotient R/P naturally gives us a way to measure the size of the prime P.

- We can do something quite similar in the function field case:

- Proposition (Degrees of Primes): If K is a function field over F and P is a prime with valuation ring R, then the quotient R/P is a finite-dimensional F-vector space. We define the degree of P to be the dimension of this vector space.
  - Proof: Since P is a maximal ideal of R and contains F, R/P is a field extension of F, so we just need to show its degree over F is finite.
  - Suppose $y \in P \backslash F$. As noted earlier, y is transcendental over F and $K/F(y)$ is a finite-degree extension. We claim that $[R/P : F] \leq [K : F(y)]$.
  - To see this, suppose that $x_1, x_2, \ldots, x_m \in R$ have the property that their reductions $\overline{x_1}, \overline{x_2}, \ldots, \overline{x_m} \in R/P$ are F-linearly independent, and suppose there is a linear dependence over $F(y)$: say $f_1(y)x_1 + f_2(y)x_2 + \cdots + f_m(y)x_m = 0$ for some $f_i(y) \in F[y]$.
  - If we cancel any common factors of y from the $f_i(y)$, and then reduce modulo P, we obtain a linear dependence of the $\overline{x_i}$ in R/P, contradiction (the point is that not all of the f_i are divisible by y, so at least one of them has a nonzero reduction modulo P).
  - Thus, any linearly independent set in R/P lifts to a linearly independent set in K, so we obtain the claimed inequality.
- Let's work all of this out in the case we mostly understand already: the purely transcendental extension $K = F(t)$.
  - From our discussion, if p is any irreducible polynomial in $A = F[t]$, we obtain a discrete valuation ring associated to the prime p as the localization $R = A_{(p)}$ and its unique maximal ideal is $P = pA_{(p)}$.
  - Then $R/P \cong A/(p)$, in which case the dimension of R/P as an F-vector space is the same as the dimension of $A/(p)$ as an F-vector space, and this is simply $\deg(p)$, since $\{\overline{1}, \overline{t}, \ldots, \overline{t^{\deg p - 1}}\}$ is a basis for $A/(p)$.
  - Thus, the degree as defined above agrees perfectly with our normal sense of the degree of a polynomial.
  - The associated p-adic valuation v_p is the same as the one we discussed earlier: $v_p(p^n \frac{r}{s}) = n$ for $p \nmid r, s$: it pick out the power of the prime p that divides a rational function $f = p^n \frac{r}{s} \in F(t)$.
  - Localizing A at a prime ideal yields almost all of the possible discrete valuation rings attached to $F[t]$.
  - But there is, in fact, one more: the valuation $v_\infty(f/g) = \deg(g)/\deg(f)$, whose associated valuation ring is obtained by localizing $A' = F[t^{-1}]$ at the prime ideal (t^{-1}). The resulting prime is known as the prime at infinity, and its degree is 1.
  - Exercise: Prove that the p-adic valuations v_p along with v_∞ are the only discrete valuations on $F(t)/F$. (Use a similar argument to the one for $\mathbb{Q}$ by identifying all possible uniformizers.)
  - The general philosophy is that there will be a few "infinite primes", and the rest are "finite primes" that arise from localizing at a prime ideal. (Over $\mathbb{Q}$, the infinite prime corresponds to the usual absolute value $|\cdot|$ resulting in the completion $\mathbb{R}$, but this is not a discrete valuation.)
- We can also give a brief explanation of some of the terminology (e.g., "function field").
  - Suppose P is a prime of K/F where F is algebraically closed (e.g., $F = \mathbb{C}$). Then the quotient R/P is a finite-degree field extension of F hence is simply (isomorphic to) F itself.
  - For any element $a \in K$, we can then simply "read off" the values of a at the various primes P by interpreting $a(P)$ as the image of a inside the quotient $R/P \cong F$.
  - This is why K/F is called a function field, since we may think of the actual elements of K as F-valued functions on the primes P. The elements of F correspond to constant functions, which is why we refer to F as the constant field of K.
  - Furthermore, as we will discuss later, we can think of the primes of P geometrically as "places" or "points".
  - For $K = \mathbb{C}(t)$, for example, we obtain a finite prime P_r corresponding to each element $t - r$ for $r \in \mathbb{C}$, along with the infinite prime. Explicitly, P_r is the collection of rational functions vanishing at r (which is the unique maximal ideal of the ring R of all rational functions defined at r), and the evaluation-at-r map yields an explicit isomorphism $R/P_r \cong \mathbb{C}$.

- Together, the finite primes P_r along with the infinite prime form the complex projective line $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$, which we may view analytically as being the Riemann sphere, and the field K consists of all of the $\mathbb{C}$-valued rational functions on the Riemann sphere.

- Our main goal now is to state, show, and use the Riemann-Roch theorem, which is unquestionably the most fundamental theorem about function fields. The first ingredient is divisors:

- Definition: The divisor group of K, written D_K, is the additive free abelian group generated by the primes of K.

  - The elements of D_K are of the form $D = \sum_P n_P P$ for $n_P \in \mathbb{Z}$, where all but finitely many of the n_P are zero. We will write $\text{ord}_P(D) = n_P$.

- Definition: The degree of a divisor $D = \sum_P n_P P$ is $\deg(D) = \sum_P n_P \deg(P)$. The degree map is a homomorphism from D_K to $\mathbb{Z}$; its kernel is the set of degree-0 divisors.

  - Note that this sum is well defined since all but finitely many n_P are zero.

- If $a \in K^\times$ is nonzero, we can attach a divisor to it by calculating its order at each prime P of K.

- Definition: We define the divisor of an element $a \in K^\times$ as $\text{div}(a) = \sum_P v_P(a) P$. The divisors of the form $\text{div}(a)$ for some $a \in K^\times$ are called principal divisors.

  - We will often also write $\text{ord}_P(a)$ (the order of a at P) interchangeably with $v_P(a)$ (the P-adic valuation of a).
  - Remark: In many sources, the divisor of a is often written (a). In our context, this can lead to ambiguities, since the same notation is also used for the ideal generated by a.
  - A priori, it is not clear we have actually given a well-defined divisor: to show this we need to establish that $v_P(a) = 0$ for all but finitely many primes P.
  - Assuming this for the moment, since $\text{ord}_P(a/b) = \text{ord}_P(a) - \text{ord}_P(b)$, summing over all primes shows that $\text{div}(a/b) = \text{div}(a) - \text{div}(b)$, so the principal divisors are a subgroup of the divisor group D_K.

- We must still show that the divisor of an element is actually well defined:

- Proposition (Divisors of Elements): For any $a \in K^\times$, we have $v_P(a) = 0$ for all but finitely many primes P of K.

  - Proof: First, if $a \in F^\times$, then for any prime P the associated valuation ring R contains F. In particular, since $a \in F^\times$ this means a is a unit in R hence has valuation 0. This means $v_P(a) = 0$ for all P, and so $\text{div}(a) = 0$.
  - Now suppose $a \notin F^\times$, so a is transcendental over F and $K/F(a)$ is finite.
  - If P is a prime of K and $v_P(a) > 0$, then by definition there is a discrete valuation ring R such that a is not a unit. Then R contains $F[a]$, and since R is integral over $F[a]$, it embeds into the integral closure of $F[a]$ inside K.
  - We lose nothing by enlarging R, so now assume R is the integral closure of $F[a]$ in K: then R is a Dedekind domain[2] since it is Noetherian, integrally closed, and the localization of R at any nonzero prime is a discrete valuation ring.
  - Since R is a Dedekind domain, every nonzero ideal can be factored uniquely as a product of prime ideals. So write $Ra = \mathfrak{p}_1^{b_1} \mathfrak{p}_2^{b_2} \cdots \mathfrak{p}_k^{b_k}$ for distinct prime ideals $\mathfrak{p}_1, \mathfrak{p}_2, \dots, \mathfrak{p}_k$. Localizing at the prime $\mathfrak{p}_i$ yields a unique prime P_i of K, and since all of the other ideals become invertible, we see $\text{ord}_{P_i}(a) = b_i$, and $\text{ord}_Q(a) \le 0$ for any other prime Q.
  - In particular, we see that there are only finitely many primes for which $v_P(a) > 0$.
  - In the same way, if $v_P(a) < 0$, by doing the same calculation for a^{-1} (i.e., by taking R' to be the integral closure of $F[a^{-1}]$ inside K and factoring $R' a^{-1} = \mathfrak{q}_1^{c_1} \cdots \mathfrak{q}_l^{c_l}$ as a product of prime ideals) we see that there are also only finitely many primes for which $v_Q(a) < 0$.

[2]For additional reference about Dedekind domains, including the factorization result we quote here, see section 16.3 of Dummit/Foote.

- Thus, the divisor $\text{div}(a) = \sum_P v_P(a)P$ is well defined, as claimed.

- Note that the proof given above gives an algorithm for computing divisors of elements: inside the integral closure of $F[a]$ (and $F[a^{-1}]$) inside K, we compute the prime ideal factorization of the ideal (a) (or (a^{-1})); then the exponent of the various primes give the corresponding valuations.

  - Let's work this out in the case of $K = \mathbb{C}(t)$: suppose we have a nonconstant rational function $a = \dfrac{f}{g}$, which we factor as $a = u\dfrac{(t-r_1)^{a_1}\cdots(t-r_k)^{a_k}}{(t-s_1)^{b_1}\cdots(t-s_l)^{b_l}}$ for distinct $r_1, \ldots, r_k, s_1, \ldots, s_l \in \mathbb{C}$ and some unit factor $u \in \mathbb{C}^\times$.
  - Each of the monic irreducibles $t - r_i$ and $t - s_j$ yields a unique prime of K. We have $v_{P_{t-r_i}}(a) = a_i$ and $v_{P_{t-s_j}}(a) = -b_j$ for each i, j, and the valuation at every other finite prime is 0. Also, we have $v_\infty(a) = \deg(g) - \deg(f) = \sum b_j - \sum_i a_i$.
  - Therefore, $\text{div}(a) = a_1 P_{t-r_1} + \cdots + a_k P_{t-r_k} - b_1 P_{t-s_1} - \cdots - b_l P_{t-s_l} + [\sum_j b_j - \sum_i a_i]\infty$.
  - Notice in particular that $\deg(\text{div}(a)) = \sum_i a_i - \sum b_j + [\sum_j b_j - \sum_i a_i] = 0$, which is to say, the divisor of any element of $K^\times$ has degree zero.
  - Furthermore, we see that the primes with positive order at a correspond precisely to the zeroes of $a = f/g$ (and the order of a at that prime is the order of vanishing of a there), while the primes with negative order correspond to poles (and the order of a at that prime is the order of the pole there).
  - We can also give a similar calculation for $K = F(t)$ where F is not algebraically closed in terms of the monic irreducible factors of $a = f/g$. The resulting divisor decomposition is essentially just the prime factorization of the rational function f/g:
  - Exercise: For $K = F(t)$, if $a = u\dfrac{p_1^{a_1}\cdots p_k^{a_k}}{q_1^{b_1}\cdots q_l^{b_l}}$ for $u \in F^\times$ and distinct monic irreducibles $p_1, \ldots, p_k, q_1, \ldots, q_l$ having associated primes $P_1, \ldots, P_k, Q_1, \ldots, Q_l$, show that $\text{div}(a) = a_1 P_1 + \cdots + a_k P_k - b_1 Q_1 - \cdots - b_l Q_l + [\sum_j b_j \deg(q_j) - \sum_i a_i \deg(p_i)]\infty$.

- Motivated by the calculations for $K = \mathbb{C}(t)$, we can also pick out the zeroes (respectively, poles) of an element by extracting only the portion of its divisor with positive (respectively, negative) coefficients:

- Definition: If $a \in K^\times$ has divisor $\text{div}(a) = \sum_P n_P P$, we define $\text{div}_+(a) = \sum_P \max(0, n_P)P = \sum_{P:n_P>0} n_P P$ and $\text{div}_-(a) = \sum_P \min(0, n_P)P = \sum_{P:n_P<0} n_P P$.

  - Notice that $\text{div}(a) = \text{div}_+(a) - \text{div}_-(a)$ for any element a.
  - Remark: There are various other notations for these quantities that are often used, such as $(a)_0$ for div_+ and $(a)_\infty$ for div_-, which are intended to evoke the idea of picking out the zeroes and poles of a.
  - For $a = \dfrac{f}{g} = u\dfrac{(t-r_1)^{a_1}\cdots(t-r_k)^{a_k}}{(t-s_1)^{b_1}\cdots(t-s_l)^{b_l}}$ in $\mathbb{C}(t)$, we have $\text{div}_+(a) = a_1P_1 + \cdots + a_kP_k$ (plus $[\deg g - \deg f]\infty$ if $\deg g - \deg f > 0$) and $\text{div}_-(a) = b_1Q_1 + \cdots + b_lQ_l$ (plus $[\deg g - \deg f]\infty$ if $\deg g - \deg f < 0$).
  - Exercise: For any field F, if $f(t), g(t) \in F[t]$ are relatively prime, show that $[F(t) : F(\frac{f(t)}{g(t)})] = \max(\deg f, \deg g)$. [Hint: Use Gauss's lemma to show that $q(y) = f(y) - \frac{f(t)}{g(t)}g(y) \in F(\frac{f(t)}{g(t)})[y]$ is the minimal polynomial of t over $F(\frac{f(t)}{g(t)})$.]
  - In the example above, we can also compute that $\deg(\text{div}_+(a)) = \deg(\text{div}_-(a)) = \max(\deg f, \deg g)$, and by the exercise above, this quantity is equal to the extension degree $[K : F(a)]$. In fact, this result is true in general:

- Theorem (Divisor Degrees): For any $a \in K^\times$, we have $\deg(\text{div}_+(a)) = \deg(\text{div}_-(a)) = [K : F(a)]$. As a consequence, $\deg(\text{div}(a)) = 0$.

  - We will defer the proof for $F = \mathbb{F}_q$ until later, since it requires a number of ingredients we have not developed yet. The general case we will skip (the result is not that difficult, but it is not especially enlightening for what we will be doing).
  - Our main observation here is that the divisor of an element $a \in K^\times$ always has degree 0, which is to say, the principal divisors are actually a subgroup of the group of degree-0 divisors.

0.12 (Oct 21) The Riemann-Roch Theorem and Applications

- Definition: We say two divisors D_1 and D_2 are linearly equivalent (and write $D_1 \sim D_2$) if $D_1 - D_2$ is principal. The resulting equivalence classes (i.e., divisors modulo principal divisors) form a group called the class group, or the Picard group, of K.
    - Exercise: Verify that this relation is an equivalence relation and that the equivalence classes are the elements in the quotient group of divisors modulo principal divisors.
    - Some notation for all of these various groups: $\mathrm{Div}(K) = D_K$ is the group of all divisors on K, $\mathrm{Div}^0(K)$ is the group of degree-0 divisors on K, $\mathrm{Cl}(K) = \mathrm{Pic}(K) = \mathrm{Div}(K)/[\text{principal divisors}]$ is the class group of K.
    - Since principal divisors all have degree zero, we can also form the reduced Picard group $\mathrm{Pic}^0(K) = \mathrm{Div}(K)/[\text{principal divisors}]$.
- For $K = F(t)$, the reduced Picard group is trivial:
- Proposition (Reduced Picard Group of $F(t)$): If $K = F(t)$, then $\mathrm{Pic}^0(K) = \mathrm{Div}(K)/[\text{principal divisors}]$ is the trivial group, and $\mathrm{Pic}(K) \cong \mathbb{Z}$.
    - Remark: It can be shown that the case $K = F(t)$ is essentially the only situation where the reduced Picard group is trivial.
    - Proof: The result is equivalent to showing that every divisor of degree 0 is principal, so suppose $D = \sum_P b_P P$ has degree 0.
    - Let $a = \prod_{P \neq \infty} p(t)^{b_p}$, where $p(t)$ is the monic irreducible polynomial associated to the finite prime P of K.
    - Then $\mathrm{ord}_P(a) = b_p$ for each prime $P \neq \infty$. But since $\sum_P b_P \deg(P) = 0$ by the assumption on D, and $\deg(\mathrm{div}(a)) = 0$ as well, we must have $\mathrm{ord}_\infty(a) = b_\infty$ also.
    - Then $\mathrm{ord}_P(a) = b_p$ for all primes P, meaning that $\mathrm{div}(a) = D$ and so D is principal as claimed.
    - The statement that $\mathrm{Pic}(K) \cong \mathbb{Z}$ follows immediately from $\mathrm{Div}(K)/\mathrm{Div}^0(K) \cong \mathbb{Z}$.
- To finish the discussion here, we remark on the analogy with the case of algebraic number fields.
    - If $K/\mathbb{Q}$ is an algebraic number field, we have an exact sequence
      $1 \to [\text{units of } \mathcal{O}_K] \to K^* \to [\text{fractional ideals of } \mathcal{O}_K] \to [\text{ideal class group of } K] \to 1$.
    - If K/F is an algebraic function field, the analogous exact sequence is
      $1 \to F^* \to K^* \to \mathrm{Div}^0(K) \to \mathrm{Pic}^0(K) \to 1$.
    - The constant field of K plays the role of the units of an algebraic number field, the group of degree-0 divisors plays the role of the fractional ideals in the ring of integers, and the reduced Picard group plays the role of the ideal class group.
- We now put a partial ordering on divisors that is motivated by the idea of divisibility for integers and rational functions.
    - The idea is that if we look at p-adic valuations of elements of $\mathbb{Q}$, we can identify the elements of $\mathbb{Z}$ as those whose valuations are nonnegative at every prime p.
- Definition: If a divisor $D = \sum_P n_P P$ has $n_P \geq 0$ for all primes P, we say D is effective and we write $D \geq 0$. We extend this notion to a partial ordering on divisors by writing $D_1 \leq D_2$ if and only if $D_2 - D_1$ is effective.
    - Exercise (easy): Check that the relation $D_1 \leq D_2$ is a partial ordering on divisors.
    - The partial ordering on divisors allows us to specify the order of zeroes and poles: to illustrate, for $K = \mathbb{C}(t)$, saying that f has a pole of order at most 2 at $z = 0$ and a zero of order at least 3 at $z = 1$ is equivalent to saying $\mathrm{div}(f) \geq -2P_{z-0} + 3P_{z-1}$.
- Definition: If D is a divisor, the Riemann-Roch space associated to D is the set $L(D) = \{a \in K^\times : \mathrm{div}(a) \geq -D\} \cup \{0\}$. Equivalently, an element $a \in K$ is in $L(D)$ if and only if $v_P(a) \geq -v_P(D)$ for all primes P of K.

- When D is an effective divisor, $L(D)$ represents all rational functions whose poles are "no worse" than D.
- More generally, if $D = \sum_P n_P P - \sum_Q m_Q Q$ with $n_i, m_i > 0$, then $L(D)$ consists of all $a \in K$ such that a has a zero of order at least m_Q at each prime Q, and may have poles only at the primes P, of order at most n_P at P.
- It is not hard to see that $L(D)$ is an F-vector space: if $a, b \in L(D)$, then $a+b \in L(D)$ because $v_P(a+b) \geq \min(v_P(a), v_P(b))$ for each prime P, and $ca \in L(D)$ for all $c \in F$ since $v_P(ca) = v_P(c) + v_P(a) = v_P(a)$ since $v_P(c) = 0$ for all primes P.
- Example: For $K = F(t)$ and $D = P_t$, we can see that $L(D) = \text{span}(1, t^{-1})$, since the only possible poles of an element $f/g \in L(D)$ function occur at $t = 0$ (of order 1) and the function must also have $\deg g \geq \deg f$ since there is no pole at the infinite prime P_∞.
- Example: For $K = F(t)$ and $D = 3P_\infty$, we can see that $L(D) = \text{span}(1, t, t^2, t^3)$ since the function f/g has no poles except a pole of order at most 3 at P_∞ (meaning that $\deg g \leq \deg f + 1$), which is to say, f/g is a polynomial of degree at most 3.
- Example: For $K = F(t)$ and $D = -P_t$, we can see that $L(D) = \{0\}$, since any nonzero element $f/g \in L(D)$ would need to be zero at $t = 0$ and defined at all other primes, but this cannot occur because g would have to be constant, but then $\deg f > \deg g$ would force f/g to have a pole at P_∞.
- Example: For arbitrary K, we have $L(0) = F$, since $\text{div}(a) = 0$ for all $a \in F^\times$, but any element $x \in K^\times \backslash F$ necessarily has at least one pole (at any prime associated to a prime occurring in the prime factorization of the ideal generated by x^{-1} in the integral closure of $F[x^{-1}]$ inside K).
- Exercise: Determine $L(D)$ when $K = F(t)$ for $D = P_t - P_\infty$, $P_t + P_\infty$, and $P_t + P_{t-1}$.

- Definition: If D is a divisor, we define $\ell(D) = \dim_F L(D)$.
  - Examples: By the examples worked out above, for $K = F(t)$ we have $l(P_t) = 2$, $l(3P_\infty) = 4$, and $l(-P_t) = 0$.
  - Example: For an arbitrary K, we have $\ell(0) = 1$, since $L(0) = F$.

- Proposition (Properties of $l(D)$): Let K be a function field over F and D be a divisor of K.

1. If $D_1 \leq D_2$, then $\ell(D_1) \leq \ell(D_2)$.
   - Proof: This follows immediately from the definition, since $D_1 \leq D_2$ clearly implies that $L(D_1)$ is a subspace of $L(D_2)$.
2. If $D_1 \sim D_2$, then $L(D_1) \cong L(D_2)$ and so $\ell(D_1) = \ell(D_2)$.
   - Proof: If $D_1 = D_2 + \text{div}(g)$, then the map from $L(D_1)$ to $L(D_2)$ sending $f \mapsto fg$ is easily seen to be an isomorphism of vector spaces since it has an inverse map $h \mapsto h/g$.
3. If $\deg(D) \leq 0$, then $L(D) = \{0\}$ and $l(D) = 0$ except when $D = \text{div}(a)$ is principal, in which case $L(D) = \text{span}(a)$ and $l(D) = 1$.
   - Proof: Suppose $f \in L(D)$ and $f \neq 0$. Then $0 = \deg(\text{div}(f)) \geq \deg(-D) = -\deg(D)$.
   - Furthermore, equality can hold only if $D = -\text{div}(f)$ for some $f \in K^\times$, in which case D is principal.
   - If D is principal, then $\ell(D) = \ell(0) = 1$ by (2), and $L(D) = Fa = \text{span}(a)$ by the same calculation.
4. If D_1 and D_2 are divisors with $D_1 \leq D_2$, then $\dim_F(L(D_2)/L(D_1)) \leq \deg(D_2) - \deg(D_1)$.
   - Proof: Induct on the sum of the coefficients of the primes in the effective divisor $B - A$. The base case $B - A = 0$ is trivial.
   - For the inductive step, suppose that $D_2 = D_1 + P$ for some prime P, and choose $x \in K$ such that $v_P(x) = v_P(D_2) = v_P(D_1) + 1$.
   - Then for any $y \in L(D_2)$, we have $v_P(xy) = v_P(x) + v_P(y) \geq v_P(D_2) - v_P(D_2) \geq 0$, so $xy \in R_P$ where R is the valuation ring associated to the prime P.
   - By composing with the evaluation map at P (i.e., taking the quotient of R_P by PR_P and then viewing this as isomorphic to R/P), we obtain an F-linear transformation $\varphi : L(D_2) \to R_P/PR_P \cong R/P$ with $\varphi(y) = (xy)(P)$.

- Then $y \in \ker(\varphi)$ if and only if $(xy)(P) = 0$ if and only if $v_P(xy) \geq 1$ if and only if $v_P(y) \geq 1 - v_P(D_2) = -v_P(D_1)$, and this last statement is equivalent to $y \in L(D_1)$.
- Thus, by the first isomorphism theorem, we have an injection from $L(D_2)/L(D_1)$ to R/P. Taking dimensions yields $\dim_F(L(D_2)/L(D_1)) \leq \dim_F(R/P) = \deg(P)$.
- This establishes the inductive step, so the general result follows.

5. For any effective divisor D, we have $\ell(D) \leq \deg(D) + 1$. In fact, this inequality holds for any divisor D of degree ≥ 0.
   - Proof: For effective divisors, this follows immediately by induction on the degree of D using (4), starting with the base case $l(0) = 1$.
   - For general divisors, the result is trivial if $\ell(D) = 0$, so suppose otherwise that $\ell(D) \geq 1$ and let $a \in L(D)$ be nonzero. Then $\text{div}(a) \geq -D$ which is equivalent to $D - \text{div}(a^{-1}) \geq 0$.
   - Then for $D' = D - \text{div}(a^{-1})$, we see that D is equivalent to the effective divisor D', and so by (2) we have $\ell(D) = \ell(D') \leq \deg(D') + 1 = \deg(D) + 1$, as required.
6. For any divisor D, the quantity $\ell(D)$ is finite.
   - Proof: If $\deg(D) < 0$ then (3) gives $\ell(D) = 0$, while if $\deg(D) \geq 0$ then (5) gives $\ell(D) \leq \deg(D)+1$.

- What we would like to be able to do now is to calculate the actual dimension $\ell(D)$ for arbitrary divisors D. Rather than delaying the point, we will now get right to our main result:
- Theorem (Riemann-Roch): For any algebraic function field K/F, there exists an integer $g \geq 0$, called the genus of K, and a divisor class $\mathcal{C}$, called the canonical class of K, such that for any divisor $C \in \mathcal{C}$ and any divisor $A \in \text{Div}(K)$, we have $\ell(A) = \deg(A) - g + 1 + \ell(C - A)$.
  - Remarks: The divisor class $\mathcal{C}$, as we will explain at length later in the case of Riemann surfaces, is the divisor class associated with the Weil differentials of K.
- We will not prove the general function-field version of the Riemann-Roch theorem, since it requires a fair bit of background to develop the necessary results about Weil differentials.
  - Instead, we will go through the proof of the analytic version of Riemann-Roch for Riemann surfaces, which contains most of the main ideas but is more accessible since the complex-analytic notion of a differential is quite natural.
- For now, we will run through some consequences of the Riemann-Roch theorem.
- Proposition (Corollaries of Riemann-Roch): Let K/F be an algebraic function field.

1. For any divisor A with $\deg(A) \geq 0$, we have $\deg(A) - g + 1 \leq \ell(A) \leq \deg(A) + 1$.
   - Proof: We showed the upper bound earlier using an inductive argument. The lower bound follows immediately from Riemann-Roch since $\ell(C - A) \geq 0$.
2. For $C \in \mathcal{C}$ we have $\ell(C) = g$ and $\deg(C) = 2g - 2$.
   - Proof: First set $A = 0$ in Riemann-Roch: this yields $\ell(0) = \deg(0) - g + 1 + \ell(C)$, so since $\ell(0) = 1$ and $\deg(0) = 0$, we get $\ell(C) = g$.
   - Now set $A = C$ in Riemann-Roch: this yields $\ell(C) = \deg(C) - g + 1 + \ell(0)$, and so $\deg(C) = \ell(C) + g - 1 - \ell(0) = 2g - 2$.
3. If $\deg(A) \geq 2g - 2$, then $\ell(A) = \deg(A) - g + 1$ except when $A \in \mathcal{C}$ (in which case $\ell(A) = g$).
   - Proof: If $\deg(A) \geq 2g-2$, then $\deg(C-A) \leq 0$, and so $\ell(C-A) = 0$ except when $C-A$ is principal (i.e., when $A \in \mathcal{C}$).
   - When $\ell(C - A) = 0$ Riemann-Roch immediately gives $\ell(A) = \deg(A) - g + 1$, and when $A \in \mathcal{C}$ we have $\ell(A) = g$ by (2).
4. The genus g is unique, as is the equivalence class $\mathcal{C}$.
   - Proof: Pick A of sufficiently large degree: then $\deg(A) - \ell(A) + 1 = g$ by (3), so g is uniquely determined.

- For $\mathcal{C}$, if $\ell(A) = \deg(A) - g + 1 + \ell(C - A) = \deg(A) - g + 1 + \ell(D - A)$ for some other divisor D, then $\ell(C - A) = \ell(D - A)$ for all A.
- Setting $A = C$ yields $\ell(D - C) = 1$ and setting $A = D$ yields $\ell(C - D) = 1$, and these are contradictory unless $D - C$ is principal, which is to say, $D \sim C$.

- Let's use Riemann-Roch to examine function fields of small genus. We start with the simplest genus $g = 0$.
  - By Riemann-Roch, we have $\ell(A) = \deg(A) + 1 + \ell(C - A)$ for any divisor A, and also $\deg(C) = -2$.
  - Also, by (3), if $\deg(A) \geq -1$ then $\ell(A) = \deg(A) + 1$. In particular, since $\deg(-C) = 2$, we have $\ell(-C) = 3$.
  - Now, for any prime P, we have $\ell(P) \leq \deg(P) + 1$. So, if P is any prime with $P \leq C$ (there must be at least one since $\deg(-C)$ is positive), we see $\ell(P) \leq \ell(-C) = 3$. Thus, $\deg(P)$ must be either 1 or 2.
  - First suppose that there is a prime P of degree 1. Then $\ell(P) = 2$. Since F is a subspace of $L(P)$, there is a basis of $L(P)$ of the form $\{1, x\}$ for some $x \notin F$.
  - Then since $\deg(\operatorname{div}(x) + P) = 1$ and $\operatorname{div}(x) + P \geq 0$, we must have $\operatorname{div}(x) + P = Q$ for some prime Q (necessarily of degree 1). Then $\operatorname{div}(x) = P - Q$, and so $[K : F(x)] = \deg(\operatorname{div}_+(x)) = \deg(P) = 1$, which means $K = F(x)$.
  - <u>Exercise</u>: Show in this case that the canonical class contains every divisor of K of degree -2.
  - Now suppose there is no prime P of degree 1: per earlier, we have a prime $P \leq C$ of degree 2.
  - Then $\ell(P) = 3$, so again since $L(P)$ contains F, we may take a basis for $L(P)$ of the form $\{1, x, y\}$ for some F-linearly independent $x, y \notin F$.
  - In the same way as above, we see that $\operatorname{div}(x) = P - Q$ and $\operatorname{div}(y) = P - R$ for some (necessarily distinct) primes Q and R of degree 2.
  - Then $[K : F(x)] = \deg(\operatorname{div}_+(x)) = 2$ and $[K : F(y)] = \deg(\operatorname{div}_+(y)) = 2$ also. Since $F(x) \neq F(y)$ (by linear independence and the fact that K is a degree-2 extension of both), we see $K = F(x, y)$.
  - Furthermore, Riemann-Roch says that $\ell(2P) = 1 + \deg(2P) = 5$, but we can find six different elements in $L(2P)$, namely $\{1, x, y, x^2, xy, y^2\}$. They must therefore be F-linearly dependent, so we see that x and y satisfy some quadratic relation $ax^2 + bxy + cy^2 + dx + ey = f$.
  - Geometrically, this case corresponds to a conic, while the case $K = F(x)$ corresponds to a line (since we can think of $F(x) = F(x, y)$ where y is a linear function of x).

0.13 (Oct 25) Riemann-Roch Over $\mathbb{C}$ + Counting Divisors

- We can use similar ideas to study the case where the genus g is equal to 1.
  - In this case, for $g = 1$ Riemann-Roch and its corollaries say that $\ell(A) = \deg(A) + \ell(C - A)$, that $\deg(C) = 0$ and $\ell(C) = 1$, and that if $\deg(A) \geq 1$ then $\ell(A) = \deg(A)$.
  - Unlike the case $g = 0$, we are not necessarily guaranteed to have a prime of any given degree any more, since we cannot use C to construct a prime of small degree – indeed, since $\deg(C) = 0$ and $\ell(C) = 1$, in fact C is principal (and $C \sim 0$).
  - So let us instead merely suppose that we do have a prime P of degree 1. Then $\ell(2P) = 2$, so choose a basis $\{1, x\}$ for $L(2P)$, where we necessarily must have $v_P(x) = 2$ since $x \notin L(P)$. Then $\ell(3P) = 3$, so choose a basis $\{1, x, y\}$ for $L(3P)$, where we must necessarily have $v_P(y) = 3$ since $y \notin L(2P)$.
  - Then, as above, $[K : F(x)] = \deg(\operatorname{div}_+(x)) = 2$ and $[K : F(y)] = \deg(\operatorname{div}_+(y)) = 3$, so since 2 and 3 are relatively prime, we see $K = F(x, y)$.
  - Now we would like to identify what kind of algebraic relation x, y must satisfy (they are, after all, algebraically dependent), which we can do by looking at the spaces $L(kP)$ for larger values of k, since the various monomials $x^i y^j$ will all only have poles at P.
  - We have $\ell(4P) = 4$, but we can only identify 4 elements that must lie in this space: $\{1, x, y, x^2\}$. (In fact, they are all linearly independent since they all have different valuations at P.)
  - Likewise, $\ell(5P) = 5$, but we only have 5 elements in this space: $\{1, x, y, x^2, xy\}$, but again, these elements are all linearly independent since they have different valuations at P.

  - But now consider $\ell(6P) = 6$: we can generate 7 elements in this space: $\{1, x, y, x^2, xy, x^3, y^2\}$. We must therefore have a linear dependence among these elements, and in fact since x^3 and y^2 are the only elements with valuation 6 at P, they must both occur with nonzero coefficients.
  - By rescaling x, y appropriately, we obtain an algebraic relation of the form $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ for some $a_1, a_2, a_3, a_4, a_6 \in F$: this is an elliptic curve, and the corresponding function field $K = F(x, y)$ is called an elliptic function field. (The indices on the coefficients a_i are listed that way because they are giving the "missing" pole valuation at P for the corresponding monomial term.)
  - When the characteristic of F is not 2 or 3, we may complete the square in y and the cube in x to obtain a simpler equation $y^2 = x^3 + Ax + B$.

- Our analysis, even in these comparatively simple situations, illustrates that there is a correspondence between algebraic function fields and algebraic plane curves.
  - As we will discuss later at length, there is an equivalence of the following two categories:
    1. (Objects) Algebraic function fields K/F of transcendence degree 1 where $K \cap \overline{F} = F$
       (Morphisms) Field injections fixing 1 (up to isomorphism)
    2. (Objects) Smooth projective curves defined over F
       (Morphisms) Non-constant rational maps defined over F (up to isomorphism)
  - The correspondence is obtained by associating a smooth curve C with the field of rational functions defined on C.

- We now discuss the details of Riemann-Roch in the case $F = \mathbb{C}$, where we can give an essentially complete argument (aside from some reliance on a few facts from complex analysis and differential topology).
  - Under the correspondence of curves and function fields, we are analyzing smooth complex projective curves, which are the same as 1-dimensional complex differentiable manifolds.
  - If we instead work over the reals, we can equivalently think of a 1-dimensional complex differentiable manifold as a compact Riemann surface X.
  - In this situation, the genus g also represents the topological genus of X (i.e., the number of "holes" in the surface, also equal to $1 - \chi/2$ where χ is the Euler characteristic).
  - The primes of the function field K are then simply the points P in X (since we are over $\mathbb{C}$, all of the primes have degree 1).
  - The elements of the function field K are then the meromorphic functions on X (i.e., the functions that are complex-differentiable except at a finite set of poles).
  - For $f \in K^\times$, the divisor $\text{div}(f) = \sum_{P \in X} v_P(f)P$ tabulates the zeroes and poles of f: if $v_P(f) = k > 0$ then f has a zero of order k at P, while if $v_P(f) = -k < 0$ then f has a pole of order k at P.
  - Two divisors D_1 and D_2 are equivalent when $D_1 - D_2$ is principal, which is to say, when they differ by the divisor of a meromorphic function.
  - We can also deduce a few facts about divisors of functions analytically (rather than algebraically as we did earlier).
  - For example, suppose $\text{div}(f) = 0$: then f has no poles and is therefore holomorphic, but since X is compact this means $|f|$ is bounded and so by Liouville's theorem, f is constant. (This also shows that the only functions holomorphic on all of X are the constants.)
  - Exercise: For any nonzero meromorphic f on X, show that $\deg(\text{div}(f)) = 0$. [Hint: Use Cauchy's argument principle: for any contour C, $\frac{1}{2\pi i}\int_C \frac{f'}{f}\,dz = Z - P$ is the number of zeroes minus the number of poles in C.]

- So far all of the basic theory is the same. However, on a differentiable manifold, we also have a natural notion of a meromorphic differential ω.
  - Specficially, a meromorphic 1-form (also called a meromorphic differential) is a differential that may locally be written as $\omega = f(z)\,dz$ for some meromorphic function f, where z is the local coordinate. (Being more precise requires a careful discussion of local coordinates and charts.)

- Example: If X is the Riemann sphere with its usual coordinate z on $\mathbb{C}$ and $1/z$ on $\mathbb{C} \cup \{\infty\} \backslash \{0\}$, some examples of meromorphic differentials are $z\,dz$ (it has a zero at 0 and a pole at ∞) and $\frac{z}{z+1}\,dz$ (it has a zero at 0 and a pole at -1).
- We can then define the divisor of a meromorphic differential $\omega = f\,dz$ as $\text{div}(\omega) = \sum_{P \in X} v_P(f)P$.
- If ω_1 and ω_2 are two meromorphic differentials, then $\omega_1/\omega_2 = f_1/f_2$ is locally a ratio of meromorphic functions hence is itself a meromorphic function.
- This means all meromorphic differentials share the same divisor class: this is the canonical class $\mathcal{C}$.
- We also have the natural notion of a holomorphic differential, which is a meromorphic differential having no poles. (This is the differential analogue of an effective divisor.)
- From differential topology, we have the following fundamental fact: the dimension of the vector space of holomorphic differentials on X is equal to the genus g. (Very roughly speaking, we can obtain independent holomorphic differentials by integrating around non-contractible paths on X.)
- Exercise: Explain why this fact is equivalent to saying $\ell(C) = g$.

- Our goal now is to give a concrete way to understand the dimension $\ell(C - A)$ for a divisor A.
  - We can do this by defining a space of differentials that mimics the Riemann-Roch space $L(D) = \{a \in K : \text{div}(a) \geq -D\}$.

- Definition: If D is a divisor on a compact Riemann surface X, we define $\Omega(D)$ to be the space of differentials ζ such that $\text{div}(\zeta) \geq -D$.
  - In the same way as for $L(D)$, it is easy to see that $\Omega(D)$ is an F-vector space.
  - Note that $\Omega(0)$ is the space of holomorphic differentials on X, which has dimension g as we noted earlier.

- Proposition: For any divisor D on a compact Riemann surface X, if $C = \text{div}(\omega)$ for a meromorphic differential ω, then $\Omega(D) \cong L(C - D)$.
  - Proof: Suppose that $\zeta \in \Omega(D)$ and consider ζ/ω. it is a meromorphic function, and we have $\text{div}(\zeta/\omega) = \text{div}(\zeta) - \text{div}(\omega) \geq D - C$, so $\omega/\zeta \in L(C - D)$.
  - Thus, the map $\zeta \mapsto \zeta/\omega$ is a linear transformation from $\Omega(D)$ to $L(C - D)$, and since clearly the map $f \mapsto f\omega$ is an inverse, it is an isomorphism.

- We now have most of the necessary ingredients for Riemann-Roch. The key additional piece is to introduce the calculation of residues of functions and differentials at a point P.
  - Given a nonzero rational function f, we may write f as a Laurent series centered at P: i.e., as $f = \sum_{n=k}^{\infty} a_k z^k$ where $k = v_P(f)$ (which may be positive or negative) and z is the local uniformizer at P. We define the residue of f at P to be the coefficient a_{-1}.
  - By Cauchy's residue theorem, we can also calculate residues via integration: $\int_C f(z)\,dz = 2\pi i \sum_P \text{Res}_P(f)$, where the sum is over all points P inside the contour C. In particular, by reversing the orientation of the curve and summing the results, we can see that the sum over all P of the residues of f is zero. (This is essentially just Stokes's theorem.)
  - In particular, if we have an effective divisor $D = P_1 + P_2 + \cdots + P_k$ for distinct points P_i, we obtain a map $\varphi : L(D) \to \mathbb{C}^k$ by taking $\varphi(D) = (\text{Res}_{P_1}(f), \text{Res}_{P_2}(f), \ldots, \text{Res}_{P_k}(f))$. The kernel of this map is the set of functions $g \in L(D)$ whose residue is zero at each P_i, but this would mean g is holomorphic on all of X, hence constant.
  - Thus, we obtain an exact sequence $0 \to \mathbb{C} \to L(D) \xrightarrow{\varphi} \mathbb{C}^k$.

- Intuitively, the statement of Riemann-Roch now comes from trying to answer the question: how close is the map φ to being surjective? In other words, what conditions are there on the values of the residues of a meromorphic function in $L(D)$ at the points P_i?
  - We can answer this question by looking at the residues of holomorphic and meromorphic differentials.

- If $\omega \in \Omega(0)$ is holomorphic, we define the residue of ω at P as the residue of the ratio $\frac{\omega}{dz}$ at P where dz is the local uniformizer at P (this is well-defined because $\frac{\omega}{dz}$ is a meromorphic function).
- In the same way as for functions, the sum of the residues of any meromorphic differential over all points must be zero: thus, for each $\omega \in \Omega(0)$ and $f \in L(D)$, we see that the sum of the residues of $f\omega$ must be zero. This means each differential imposes a linear condition on the possible choices of residues for f.
- More precisely, if $D = P_1 + P_2 + \cdots + P_k$ for distinct points P_i, we obtain a map $\psi : \Omega(0) \to \mathbb{C}^k$ by taking $\psi(D) = (\mathrm{Res}_{P_1}(\omega), \mathrm{Res}_{P_2}(\omega), \dots, \mathrm{Res}_{P_k}(\omega))$. The kernel of this map is the set of differentials $\omega \in \Omega(0)$ whose residue is zero at each P_i, which is to say, $\omega \in \Omega(D)$.
- Thus, we obtain another exact sequence $0 \to \Omega(D) \to \Omega(0) \xrightarrow{\psi} \mathbb{C}^k$.
- The images of the two maps φ and ψ are orthogonal by the observation made above: for any $f \in L(D)$ and any $\omega \in \Omega(0)$, the inner product of $\varphi(f)$ and $\psi(\omega)$ is $\sum_{i=1}^k \mathrm{Res}_{P_i}(f) \cdot \mathrm{Res}_{P_i}(\omega) = \sum_{i=1}^k \mathrm{Res}_{P_i}(f\omega) = 0$ by Stokes's theorem.
- So, since the images of φ and ψ are orthogonal, we see that $\dim(\mathrm{im}\varphi) + \dim(\mathrm{im}\psi) \leq k = \deg(D)$.
- By the nullity-rank theorem, since $\ker(\varphi) = \mathbb{C}$ we get $\dim(\mathrm{im}\varphi) = \dim(L(D)) - 1 = \ell(D) - 1$.
- Likewise, since $\ker(\psi) = \Omega(D)$ we get $\dim(\mathrm{im}\psi) = \dim(\Omega(0)) - \dim(\Omega(D)) = g - \ell(C - D)$.
- Thus, we obtain the inequality $\ell(D) - 1 + g - \ell(C - D) \leq \deg(D)$.

- If we had equality everywhere (i.e., if the images of φ and ψ were actually orthogonal complements) then we would get the Riemann-Roch theorem!
  - As it is, we only have the weaker statement that $\ell(D) - 1 + g - \ell(C - D) \leq \deg(D)$, which is known as Riemann's inequality (and only in the case where D is effective).
  - One can in fact establish that the images of φ and ψ are orthogonal complements with quite a bit more work.
  - In the event that $C - D$ is also effective, however, we can extract the desired result just from Riemann's inequality: in such a case, we have $\ell(D) - 1 + g - \ell(C - D) \leq \deg(D)$ and also $\ell(C - D) - 1 + g - \ell(D) \leq \deg(C - D) = \deg(C) - \deg(D)$, so adding the two inequalities yields $2g - 2 \leq \deg(C)$.
  - But since $\deg(C) = 2g - 2$ (another calculation we take for granted), we must have equality in both cases.
  - This establishes Riemann-Roch for divisors D where both D and $C - D$ are effective divisors (or equivalent to effective divisors, since as we showed, $\ell(D_1) = \ell(D_2)$ when $D_1 \sim D_2$).
  - In fact, this is nearly enough to get the general result, since as we showed, if $L(D) \neq 0$ then D is equivalent to an effective divisor. In general, one needs to verify that when $\ell(C - D) = 0$, one has $\deg(D) \geq \ell(D) - 1 + g$.
  - Assuming the inequality $\deg(D) \geq \ell(D) - 1 + g$, one obtains the general statement of Riemann-Roch: if both D and $C - D$ are equivalent to effective divisors, the result is as above, and if D is but $C - D$ is not, the result follows from $\deg(D) \geq \ell(D) - 1 + g$, and if $C - D$ is but D is not, the result is equivalent by interchanging D and $C - D$.
  - Finally, if neither D nor $C - D$ is equivalent to an effective divisor (i.e., if $\ell(D) = \ell(C - D) = 0$), then by the inequality above we must have $\deg(D) \geq g - 1$ and $\deg(C - D) \geq g - 1$. But since $\deg(C) = 2g - 2$ this forces $\deg(D) = g - 1$, in which case we do get $\deg(D) = \ell(D) - 1 + g - \ell(C - D)$, as required.

- To summarize, the main tools used in proving Riemann-Roch involve studying the relationships between divisors and differentials, and using structural statements about residues of functions and differentials.

- We will now use Riemann-Roch to study zeta functions of function fields, so we now assume that our base field is $F = \mathbb{F}_q$. Our first ingredient is that there necessarily exist divisors of all degrees:

- Proposition (Existence of Degree-1 Divisors): If K is a function field over $\mathbb{F}_q$, then there exists a divisor D of degree 1 over K, and hence there exist degrees of all degrees over K.
  - Proof (sketch): Let P be a prime of K and let $\sigma = \mathrm{Frob}_q$ be the q-power Frobenius automorphism of K.

- Exercise: If R is the valuation ring of P, show that σR is also a valuation ring with maximal ideal σP, and that σ gives an isomorphism of R/P with $\sigma R/\sigma P$. Show also that for any $a \in K$, $v_{\sigma P}(a) = v_P(\sigma^{-1}a)$.
- By the exercise, σP is also a prime of K and it has the same degree as P, so $\sigma P - P$ has degree 0.
- It can be shown that $\sigma P - P$ is equal to $\sigma D - D$ for some degree-0 divisor D (this is essentially Hilbert's theorem 90), which means $\sigma(P - D) = P - D$.
- Then $P - D$ is a divisor that is fixed by the Frobenius map, which (one may show) necessarily implies that $P - D$ has degree 1. (The principle is the same as the observation that the elements of $\overline{\mathbb{F}_q}$ fixed by Frobenius are precisely the elements of $\mathbb{F}_q$, which generate extensions of degree 1.)

- As an immediate consequence of the proposition above, we obtain an exact sequence $0 \to \text{Pic}^0(K) \to \text{Pic}(K) \to \mathbb{Z} \to 0$.
  - Our next goal is to show that the reduced Picard group is finite.

- Proposition (Finiteness of the Class Group): Let K be a function field over $\mathbb{F}_q$. Then the following hold:
  1. For any $n \geq 1$, the number of primes of K having degree n is finite.
     - Proof: Let $x \in K \backslash F$, so that $[K : F(x)]$ is finite. If P is a prime of K having degree n, then P lies over some prime of $F(x)$ of degree $\leq n$.
     - Since there are only finitely many primes in $F(x)$ of degree $\leq n$, and there are only finitely many different primes in a finite-degree extension $[K : F(x)]$ that lie above a particular prime in $F(x)$ (specifically, this number is bounded by the extension degree), we see that there are only finitely many primes of K.
     - Exercise: Show that the number of primes of degree $\leq n$ in K is at most $[K : F(x)]q^n$ for any $x \in K \backslash F$.
  2. For any $n \geq 0$, the number of effective divisors of K having degree n is finite.
     - Proof: Suppose $D = \sum_P n_P P$ is effective and has degree n. Then $\deg(P) \leq \deg(D) = n$ for each prime P appearing with a positive coefficient.
     - By (1), there are only finitely many possible primes P of degree at most n. For each such prime P, the coefficient n_P is at most $n/\deg(P)$, so there are finitely many possible choices for each n_P.
     - Thus, there are only finitely many possible terms $n_P P$ that can appear in D, and so the number of effective divisors of degree n is finite.
     - Exercise: Give an explicit upper bound in terms of $[K : F(x)]$, q, and n for the number of effective divisors of degree n in K.
  3. The reduced Picard group $\text{Pic}^0(K)$ is finite.
     - Proof: Let D be a divisor of degree 1. If A is any divisor of degree 0, then $\deg(gD + A) = g$, so $\ell(gD + A) \geq \deg(gD + A) - g + 1 = 1$ by Riemann-Roch.
     - Pick any nonzero $f \in L(gD + A)$: then $\text{div}(f) + gD + A \geq 0$ is some effective divisor B.
     - Then $A \sim B - gD$, so since there are finitely many possible B by (2), and gD is fixed, there are finitely many possible classes for A.
  4. If h_K is the class number of K (the cardinality of $\text{Pic}^0(K)$), then there are exactly h_K divisor classes of each possible degree.
     - Proof: Let D be a divisor of degree 1. Then for any divisor A of degree n, we see $A - nD$ has degree 0, and so by (3) there are h_K possible classes for $A - nD$ up to equivalence.
     - Since nD is fixed, this means there are h_K possible classes for A up to equivalence, as required.

- We conclude by writing down the construction of the zeta function for K.
  - We use essentially the same definition as for $\mathbb{F}_q[t]$, using the fact that effective divisors of K are the natural analogues of the monic polynomials in $\mathbb{F}_q[t]$, aside from some small considerations about the prime at ∞.

- Definition: If K is a function field over $\mathbb{F}_q$ and $A \geq 0$ is an effective divisor, the norm of A is $NA = q^{\deg A}$.

- Exercise (easy): If $A, B \geq 0$, show that $N(A+B) = NA \cdot NB$.

- Definition: If K is a function field over $\mathbb{F}_q$, the zeta function of K is $\zeta_K(s) = \sum_{A \geq 0} (NA)^{-s} = \sum_{A \geq 0} q^{-s \deg A}$.
  - Next time, we will establish the important properties of the zeta function.

0.14 (Oct 28) Zeta Functions and the Weil Conjectures

- Definition: If K is a function field over $\mathbb{F}_q$, the zeta function of K is $\zeta_K(s) = \sum_{A \geq 0} (NA)^{-s} = \sum_{A \geq 0} q^{-s \deg A}$.
  - As follows from the estimates in the exercises earlier (which we will improve later), the zeta function converges absolutely for $\mathrm{Re}(s) > 1$.
  - By grouping together the effective divisors by degree, we see that $\zeta_K(s) = \sum_{n=1}^{\infty} \dfrac{b_n}{q^{ns}}$ where b_n is the number of effective divisors of degree n.
  - For $K = \mathbb{F}_q(t)$, the zeta function of K is almost identical to that of the subring $A = \mathbb{F}_q[t]$, aside from the behavior of the prime at infinity.
  - Explicitly, any monic polynomial $p \in A$ yields a unique class of effective divisors of the form $\mathrm{div}(p) + nP_\infty$ for $n \geq \deg(p)$, and conversely, for any effective divisor $D = \sum_P n_P P$ of K, we get a unique associated monic polynomial $p \in A$ as $p = \prod_{P \neq \infty} p_i^{v_{P_i}(D)}$.
  - Exercise: Show that $\zeta_{\mathbb{F}_q(t)}(s) = (1 - q^{-s})^{-1} \zeta_{\mathbb{F}_q[t]}(s) = \dfrac{1}{(1 - q^{1-s})(1 - q^{-s})}$.
  - Also, since the norm is multiplicative, we get an Euler product for the zeta function: $\zeta_K(s) = \prod_P (1 - NP^{-s})^{-1}$. It is also absolutely convergent for $\mathrm{Re}(s) > 1$.
  - By grouping the primes together by degree, we see that $\zeta_K(s) = \prod_{d=1}^{\infty} (1 - q^{-ns})^{-a_d}$ where a_d is the number of primes of K of degree d.

- To go further, we need to improve our estimates on divisor-counting.

- Proposition (Divisor-Counting): Let K be a function field over $\mathbb{F}_q$.
  1. For any divisor A of K, the number of effective divisors equivalent to A is $(q^{\ell(A)} - 1)/(q-1)$.
     - Proof: As we have previously noted several times, that there exists an effective divisor B equivalent to A if and only if there exists $f \in K^\times$ such that $B = \mathrm{div}(f) + A \geq 0$, if and only if $f \in L(A)$.
     - Thus, if $\ell(A) = 0$, the formula is correct, since there are no effective divisors equivalent to A.
     - If $\ell(A) > 0$, then consider the function from $L(A) \backslash \{0\}$ to the effective divisors equivalent to A given by $f \mapsto \mathrm{div}(f) + A$. This map is surjective, and also f, g will have the same image precisely when $\mathrm{div}(f) = \mathrm{div}(g)$, which is to say, when $\mathrm{div}(f/g) = 0$, i.e., when $f/g \in \mathbb{F}_q^\times$.
     - Thus, since the function is surjective, the cardinality of the domain is $\#[L(A) \backslash \{0\}] = q^{\ell(A)} - 1$, and each fiber has size $q - 1$, the cardinality of the image is $(q^{\ell(A)} - 1)/(q-1)$, as claimed.
  2. If the h_K divisor classes of degree n are represented by $A_1, \dots, A_h$, then the number of effective divisors of degree n is $\sum_{i=1}^{h} \dfrac{q^{\ell(A_i)} - 1}{q - 1}$.
     - Proof: As noted in our earlier proposition, there are exactly h_K divisor classes of degree n.
     - By the well-definedness of divisor classes, each effective divisor class of degree n is equivalent to exactly one of $A_1, \dots, A_h$, so summing over the divisor classes and applying (1) yields the result.
  3. The zeta function $\zeta_K(s) = \sum_{A \geq 0} (NA)^{-s}$, and also its Euler product $\zeta_K(s) = \prod_P (1 - NP^{-s})^{-1}$, both converge absolutely for $\mathrm{Re}(s) > 1$.
     - Proof: By Riemann-Roch, if $\deg(A) > 2g - 2$ then $\ell(A) = \deg(A) - g + 1$.
     - Applying (2) shows that for $n > 2g - 2$, the number b_n of effective divisors of degree n is $h \cdot \dfrac{q^{\deg(A) - g + 1} - 1}{q - 1} = O(q^n)$.

- Thus, the zeta function sum is bounded in absolute value by $\sum_{n=1}^{\infty} O(q^n) q^{-ns}$, which converges absolutely for $\text{Re}(s) > 1$.
- Similarly, for the Euler product, by the usual results about convergence of products, $\zeta_K(s) = \prod_{d=1}^{\infty} (1 - q^{-ds})^{-a_d}$ will converge provided $\sum_{d=1}^{\infty} a_d \left|q^{ds}\right|$ converges, but $a_d \le b_d$, so it does converge for $\text{Re}(s) > 1$.

- Our next goal is to prove the Weil conjectures for K. The general statement of the Weil conjectures for projective varieties over $\mathbb{F}_q$ (equivalently, for function fields of arbitrary finite transcendence degree) are as follows:

- Theorem (Weil Conjectures): Let K be an algebraic function field of transcendence degree n over its constant field $\mathbb{F}_q$, with associated zeta function $\zeta_K(s)$. (Equivalently, let Y be a nonsingular n-dimensional projective variety defined over $\mathbb{F}_q$.) Then the following properties hold:

  1. (Rationality) The zeta function $\zeta_K(s)$ is a rational function of $u = q^{-s}$. More specifically, $\zeta_K(s) = \prod_{i=0}^{2n} p_i(u)^{(-1)^{i+1}} = \dfrac{p_1(u)p_3(u)\cdots p_{2n-1}(u)}{p_0(u)p_2(u)\cdots p_{2n}(u)}$ for appropriate polynomials $p_i(u) \in 1 + t\mathbb{Z}[t]$, where $p_0(u) = 1 - u$, $p_{2n}(u) = 1 - q^n u$, and $p_i(u) = \prod_j (1 - \alpha_{i,j} u)$ for some $\alpha_{i,j} \in \mathbb{C}$.
  2. (Functional Equation / Poincaré Duality) The zeta function has a functional equation $\zeta_K(n - s) = \pm q^{nE/2} u^E \zeta_K(s)$, where $E = 2 - 2g$ is the Euler characteristic of K. In particular, the map $\alpha \mapsto q^n/\alpha$ maps the zeroes of p_i to the zeroes of p_{2n-i}.
  3. (Riemann Hypothesis) For each i, j, the inverse zeroes $\alpha_{i,j}$ of p_i have $|\alpha_{i,j}| = q^{i/2}$. Equivalently, all of the zeroes of $p_k(u)$ lie on the line $\text{Re}(u) = k/2$.
  4. (Betti Numbers) If K is the function field of a nonsingular variety X defined over an algebraic number field with good reduction modulo $\tilde{p} = \text{char}(\mathbb{F}_q)$, then the degree of p_i is the ith Betti number of the space $X(\mathbb{C})$ of complex points on X.

- In our situation, we have $n = 1$ (i.e., for curves), in which case the Weil conjectures read as follows:

  1. $\zeta_K(s)$ is a rational function of $u = q^{-s}$: specifically, $\zeta_K(s) = \dfrac{L_K(u)}{(1-u)(1-qu)}$ for some polynomial $L_K(u) = \prod_j (1 - \alpha_j u)$.
  2. For $\xi_K(s) = q^{(g-1)s} \zeta_K(s)$, we have $\xi_K(1 - s) = \xi_K(s)$.
  3. The roots of L_K all lie on the line with $\text{Re}(s) = 1/2$.
  4. The degree of L_K is $2g$.

- Exercise: Using the explicit formula $\zeta_{\mathbb{F}_q(t)}(s) = \dfrac{1}{(1 - q^{-s})(1 - q^{1-s})}$, verify the Weil conjectures for $K = \mathbb{F}_q(t)$.

- The Weil conjectures have a long history. Here is a brief summary of some of it:
  - In the early 1800s, Gauss identified some components of these general results in particular examples for certain curves, in the context of counting points on elliptic curves modulo p.
  - In 1924, Artin conjectured the general results for curves and Hasse independently proved the results for elliptic curves.
  - In 1949, Weil formulated the general statement of the Weil conjectures (he had previously established Artin's conjectured statements in the case of curves).
  - Establishing the Weil conjectures in full took the development of about 20 more years' worth of algebraic geometry machinery: Dwork proved (1) in 1960, while Grothendieck proved (1), (2), and (4) in the 1960s, and Deligne finished (3) in 1973.

- At this point, we can prove parts (1), (2), and (4) of the Weil conjectures in our setting using the Riemann-Roch theorem:

- The main idea in the proof of (1) is to use the simple estimate given by Riemann-Roch for the number of effective divisors of large degree. To establish (2) requires using the full statement for divisors of low degree. Once (1) and (2) are established, (4) is quite simple, requiring only a calculation of the degree of a polynomial.
- Proof (1): Write $u = q^{-s}$ and set $Z_K(u) = \zeta_K(s) = \sum_{n=0}^{\infty} b_n u^n$ where b_n is the number of effective divisors of degree n over K.
- As we showed earlier, for $n > 2g-2$ we have $b_n = h_K \cdot \dfrac{q^{n-g+1}-1}{q-1}$ where h_K is the class number of K and g is the genus of K. Thus,

$$\begin{aligned} Z_K(u) &= \sum_{n=0}^{2g-2} b_n u^n + \sum_{n=2g-1}^{\infty} h_K \frac{q^{n-g+1}-1}{q-1} u^n \\ &= \sum_{n=0}^{2g-2} b_n u^n + \frac{h_K}{(q-1)} \left[\sum_{n=2g-1}^{\infty} q^{n-g+1} u^n - \sum_{n=2g-1}^{\infty} u^n \right] \\ &= \sum_{n=0}^{2g-2} b_n u^n + \frac{h_K}{(q-1)} \left[\frac{q^g}{1-qu} - \frac{1}{1-u} \right] u^{2g-1}. \end{aligned}$$

- Therefore, we have

$$\begin{aligned} (1-u)(1-qu)Z_K(u) &= (1-u)(1-qu) \sum_{n=0}^{2g-2} b_n u^n + \frac{h_K}{(q-1)} \left[q^g(1-u) + (1-qu) \right] u^{2g-1} \\ \text{so } L_K(u) &= \sum_{n=0}^{2g-2} b_n u^n (1-u)(1-qu) + h_K \frac{q^g-1}{q-1} u^{2g-1} + h_K \frac{q-q^g}{q-1} u^{2g}. \end{aligned}$$

  Each term is a polynomial with integral coefficients (since $q-1$ divides q^g-1 and $q-q^q$), and the total degree is clearly at most $2g$, as required.
- Furthermore, setting $u=0$ yields $L_K(0)=1$, so $L_K(u) \in 1+t\mathbb{Z}[t]$. Also, setting $u=1$ yields $L_K(1)=h_K$ and setting $u=1/q$ yields $L_K(1/q) = q^{g-1} h_K$. In particular, we see that L_K is nonzero at 1 and $1/q$, so $Z_K(u) = \zeta_K(s)$ does indeed have poles at $s=0$ and $s=1$, as required.
- Proof (2): As calculated in (1), we have $Z_K(u) = \sum_{n=0}^{2g-2} b_n u^n + \dfrac{h_K}{(q-1)} \left[\dfrac{q^g}{1-qu} - \dfrac{1}{1-u} \right] u^{2g-1}$.
- Also, from our proposition earlier, we have $b_n = \sum_{\deg(\overline{A})=n} \dfrac{q^{\ell(\overline{A})}-1}{q-1}$, where the sum is over the h_K divisor classes $\overline{A}$ of degree n.
- Plugging this in and multiplying by $q-1$ to clear denominators yields

$$\begin{aligned} (q-1)Z_K(u) &= \sum_{n=0}^{2g-2} \left[\sum_{\deg(\overline{A})=n} (q^{\ell(\overline{A})}-1) \right] u^n + h_K \frac{q^g u^{2g-1}}{1-qu} - h_K \frac{u^{2g-1}}{1-u} \\ &= \sum_{0 \le \deg(\overline{A}) \le 2g-2} q^{\ell(\overline{A})} u^{\deg(\overline{A})} + h_K \frac{q^g u^{2g-1}}{1-qu} - h_K \frac{1}{1-u}. \end{aligned}$$

- Therefore,

$$\begin{aligned} (q-1)\xi_K(s) = (q-1)u^{1-g} Z_K(u) &= u^{1-g} \sum_{0 \le \deg(\overline{A}) \le 2g-2} q^{\ell(\overline{A})} u^{\deg(\overline{A})} + h_K \frac{q^g u^g}{1-qu} - h_K \frac{u^{1-g}}{1-u} \\ &= \sum_{0 \le \deg(\overline{A}) \le 2g-2} q^{\ell(\overline{A})} u^{\deg(\overline{A})-g+1} + h_K \frac{(qu)^g}{1-qu} - h_K \frac{u^{-g}}{1-u^{-1}}. \end{aligned}$$

- We claim that this last expression is invariant under the substitution $u \mapsto q^{-1}u^{-1}$. This is clearly the case for the last two terms (since they are interchanged and each get a minus sign under the substitution), so we need only check the result for the sum.
- Substituting $u \mapsto q^{-1}u^{-1}$ in the sum yields $\sum_{0 \le \deg(\overline{A}) \le 2g-2} q^{\ell(\overline{A})}(qu)^{-\deg(\overline{A})+g-1} = \sum_{0 \le \deg(\overline{A}) \le 2g-2} q^{\ell(C-\overline{A})}u^{\deg(C-\overline{A})}$ using Riemann-Roch and the fact that $\deg(C) = 2g-2$.
- But again, since $\deg(C) = 2g-2$, this sum is the same as the original since the map $A \mapsto C - A$ merely reverses the order of the terms in the summation.
- Therefore, we see that $(q-1)\xi_K(1-s) = (q-1)\xi_K(s)$, as claimed.
- Proof (4): As calculated in (1), we see that L_K has degree at most $2g$.
- But as also calculated in (1), we have $L_K(0) = 1$, and by (2) since $q^{-g}u^{-2g}L_K(u) = L_K(q^{-1}u^{-1})$, taking $u \to \infty$ shows that $L_K(u)/u^{2g} \to q^g$ is nonzero, which means $L_K(u)$ has degree exactly $2g$.

- We do not have all of the necessary tools to establish the Riemann hypothesis part of the Weil conjectures yet.
  - However, we can do half of it: by (2), since $q^{-g}u^{-2g}L_K(u) = L_K(q^{-1}u^{-1})$, the map $u \mapsto q^{-1}u^{-1}$ must permute the roots $\pi_1, \pi_2, \dots, \pi_{2g}$ of L_K, which means that (suitably ordered) we must have $\pi_i = q/\pi_{2g-i}$.
  - Equivalently, this says that the roots of $\zeta_K(s)$ come in pairs, reflected across the line $\mathrm{Re}(s) = 1/2$. The Riemann hypothesis is that all of these roots actually lie on the line $\mathrm{Re}(s) = 1/2$ itself.

- Even though we cannot actually prove the Riemann hypothesis right now, we can still give some of its applications.

- Proposition (Hasse-Weil Bound): If K is a function field of genus g over $\mathbb{F}_q$, then the number a_1 of prime divisors of K of degree 1 satisfies the inequality $|a_1 - q - 1| \le 2g\sqrt{q}$.
  - Equivalently, if one phrases this in terms of algebraic curves, it says that the number of $\mathbb{F}_q$-points a_1 of a smooth projective curve defined over $\mathbb{F}_q$ satisfies $|a_1 - q - 1| \le 2g\sqrt{q}$.
  - This result was shown for genus $g = 1$ (i.e., for elliptic curves) by Hasse in 1933 and subsequently generalized by Weil to larger genus in 1949.
  - Proof: We have $Z_K(u) = \sum_{n=0}^{\infty} b_n u^n = \dfrac{\prod_{i=1}^{2g}(1-\pi_i u)}{(1-u)(1-qu)}$.
  - Thus, $\log Z_K(u) = \sum_{i=1}^{2g} \log(1-\pi_i u) - \log(1-u) - \log(1-qu)$, so

$$\begin{aligned} u\frac{d}{du}[\log Z_K(u)] &= \frac{u}{1-u} + \frac{qu}{1-qu} + \sum_{i=1}^{2g} \frac{\pi_i u}{1-\pi_i u} \\ &= \sum_{n=1}^{\infty}(1 + q^n - \sum_{i=1}^{2g} \pi_i^n)u^n. \end{aligned}$$

  - In particular, the coefficient of u^1 of $u\dfrac{d}{du}[\log Z_K(u)]$ is $1 + q - \sum_{i=1}^{2g} \pi_i$.
  - But since $Z_K(u) = \prod_{d=1}^{\infty}(1-u^d)^{-a_d}$, we also have $u\dfrac{d}{du}[\log Z_K(u)] = \sum_{d=1}^{\infty} a_d \dfrac{du^d}{1-u^d} = \sum_{n=1}^{\infty}[\sum_{d|n} da_d]u^n$.
  - In particular, the coefficient of u^1 of $u\dfrac{d}{du}[\log Z_K(u)]$ is also a_1.
  - So, by the triangle inequality, we have $a_1 = q + 1 - \sum_{i=1}^{2g} \pi_i$, so $|a_1 - q - 1| = \left|\sum_{i=1}^{2g} \pi_i\right| \le \sum_{i=1}^{2g} |\pi_i| = 2g\sqrt{q}$, where the last equality follows by the Riemann hypothesis $|\pi_i| = q^{1/2}$.

- Exercise: Show that if $q \ge 4g^2$, then there must exist primes of degree 1 in K.

0.15 (Nov 1) Affine Varieties + Hilbert's Nullstellensatz

- Proposition (Class Number Bounds): If K is a function field of genus g over $\mathbb{F}_q$, then the class number h_K satisfies $(\sqrt{q}-1)^{2g} \le h_K \le (\sqrt{q}+1)^{2g}$.
  - Proof: As we noted previously, $h_K = L_K(1) = \prod_{i=1}^{2g}(1-\pi_i) = \prod_{i=1}^{2g}(\pi_i - 1)$.
  - Since $|\pi_i| = q^{1/2}$, the absolute value of the product is bounded below by $(\sqrt{q}-1)^{2g}$ and above by $(\sqrt{q}+1)^{2g}$, as required. Since h_K is positive, we obtain the stated bounds.
- Exercise: Show that if $q > 4$ and $g > 0$, then the class number of K is greater than 1.
- We can also use a similar calculation to establish the analogue of the prime number theorem for K.
- Theorem (Prime Number Theorem for Function Fields): If K is a function field of genus g over $\mathbb{F}_q$ and a_n is the number of primes of K having degree n, then $a_n = \dfrac{q^n}{n} + O(\dfrac{q^{n/2}}{n})$.
  - Proof: As we calculated above, $\sum_{n=1}^{\infty}(1+q^n-\sum_{i=1}^{2g}\pi_i^n)u^n = u\dfrac{d}{du}[\log Z_K(u)] = \sum_{n=1}^{\infty}[\sum_{d|n} da_d]u^n$.
  - Thus, we have $\sum_{d|n} da_d = 1+q^n-\sum_{i=1}^{2g}\pi_i^n$, so by Mobius inversion and the fact that $|\pi_i| = q^{1/2}$, we obtain the formula

$$\begin{aligned} a_n &= \frac{1}{n}\left[\sum_{d|n}\mu(d)[1+q^{n/d}-\sum_{i=1}^{2g}\pi_i^{n/d}]\right] \\ &= \frac{1}{n}\left[(1+q^n-\sum_{i=1}^{2g}\pi_i^n) + \sum_{d|n,d<n}\mu(d)[1+q^{n/2}-\sum_{i=1}^{2g}\pi^{n/d}]\right] \\ &= \frac{1}{n}\left[q^n + 2gO(q^{n/2}) + O(q^{n/2})\right] \\ &= \frac{q^n}{n} + O(\frac{q^{n/2}}{n}) \end{aligned}$$

    as claimed.
- In the proof above, if we replace the Riemann hypothesis assumption $|\pi_i| = q^{1/2}$ by the weaker estimate $|\pi_i| \le q^{\alpha}$ then we instead get an error term of $O(q^{\alpha}/n)$.
  - For completeness, we can give a fairly simple argument now that a bound of this nature does hold for some $\alpha < 1$.
- Proposition (Zero-Free Region for Zeta): If K is a function field over $\mathbb{F}_q$, then $\zeta_K(s)$ has no zeroes on the line $\mathrm{Re}(s) = 1$, and thus there exists an $\alpha < 1$ such that $|\pi_i| \le q^{\alpha}$ for each of the inverse zeroes $\pi_1, \dots \pi_{2g}$ of $\zeta_K(s)$.
  - Our proof that there are no zeroes on $\mathrm{Re}(s) = 1$ mimics the proof in the classical case over $\mathbb{Q}$.
  - In the classical case, zero-free regions to the left of $\mathrm{Re}(s) = 1$ have been established, but they approach $\mathrm{Re}(s) = 1$ as $\mathrm{Im}(s) \to \infty$. We are able to obtain an improvement in the function field case (i.e., the second part, which allows us to move a uniform distance away from $\mathrm{Re}(s) = 1$) because the zeta function is periodic in the imaginary direction, so we need only identify its zeroes on a compact region.
  - Proof: First, note that $3 + 4\cos\theta + \cos 2\theta = 2(1+\cos\theta)^2 \ge 0$ for real θ.
  - If we write $s = \sigma + it$ with $\sigma > 1$, then the real part of $\log\zeta_K(s)$ is

$$\begin{aligned} \mathrm{Re}[\log\zeta_K(s)] &= \mathrm{Re}[\sum_P -\log(1-NP^{-s})] \\ &= \mathrm{Re}[\sum_P \sum_{m=1}^{\infty}\frac{1}{m}NP^{-sm}] \\ &= \sum_P \sum_{m=1}^{\infty}\frac{1}{m}NP^{-m\sigma}\cos(t\cdot\log(NP)^m). \end{aligned}$$

- Now, by replacing t with 0, t, and $2t$, using the trigonometric identity above for $\theta = \log(NP)^m$, and summing over all P and all m, we see that $3\text{Re}[\log \zeta_K(\sigma)] + 4\text{Re}[\log \zeta_K(\sigma+it)] + \text{Re}[\log \zeta_K(\sigma+2it)] \geq 0$.
- Exponentiating therefore yields $|\zeta_K(\sigma)|^3 \cdot |\zeta_K(\sigma+it)|^4 \cdot |\zeta_K(\sigma+2it)| \geq 1$ for all $\sigma > 1$ and all real t.
- Recall that we have shown $\zeta_K(s)$ is a rational function of $u = q^{-s}$ and its only poles are simple poles at $s = 1$ and $s = 1/q$. In particular, it is periodic with period $2\pi i/\log q$, and so the poles on the line $\text{Re}(s) = 1$ are located at the multiples of $2\pi i/\log q$.
- Now suppose that $\zeta_K(1+it) = 0$; note that we must have $t \neq 0$, since $\zeta_K(s)$ has a single simple pole at $s = 1$.
- Then $\dfrac{\zeta_K(\sigma+it)}{\sigma-1}$ is bounded as $\sigma \to 1$. Also, $(\sigma-1)\zeta_K(\sigma)$ is bounded as $\sigma \to 1$ again since $\zeta_K(s)$ has a simple pole at $s = 1$. Finally, $\zeta_K(\sigma+2it)$ is bounded as $\sigma \to 1$ as long as t is not an odd multiple of $\pi/\log q$.
- In such a case, we see that $(\sigma-1) \cdot |(\sigma-1)\zeta_K(\sigma)|^3 \cdot \left|\dfrac{\zeta_K(\sigma+it)}{\sigma-1}\right|^4 \cdot |\zeta_K(\sigma+2it)|$ is equal to $(\sigma-1)$ times a product of three bounded quantities as $\sigma \to 1$, so the limit as $\sigma \to 1$ is zero. This is impossible, since the limit must be ≥ 1 by the inequality above.
- If t is an odd multiple of $2\pi/\log q$, which is to say, when $q^{-(1+it)} = -1/q$, we require a different approach. As we will prove later, we have $L_K(-1) = h_{K_2}/h_K$, where $K_2 = K\mathbb{F}_{q^2}$ (i.e., the field obtained by extending the constant field of K from $\mathbb{F}_q$ to $\mathbb{F}_{q^2}$).
- Then $\zeta_K(1+it) = Z_K(-1/q)$, which is nonzero by the functional equation and the fact that $L_K(-1) \neq 0$.
- We conclude that $\zeta_K(1+it) \neq 0$ for all real t, as required.
- For the second part, we have just shown that $\zeta_K(s)$ does not vanish for $\text{Re}(s) = 1$. Furthermore, since $\zeta_K(s)$ is represented by an absolutely-convergent Euler product for $\text{Re}(s) > 1$, it does not vanish there, and so by the functional equation, $\zeta_K(s)$ also does not vanish for $\text{Re}(s) \leq 0$.
- Therefore, for $\pi_i = q^{\beta_i}$, we have $0 < \text{Re}(\beta_i) < 1$. If we then take $\alpha = \max(\text{Re}(\beta_i))$, then $\alpha < 1$ since there are only $2g$ total β_i. Then $|\pi_i| \leq q^\alpha$ as required.

- So far, our development and study of function fields has been analogous to number theory (with a little analysis thrown in).
  - However, a key part of the story is the correspondence between function fields and projective curves, and in many cases, thinking in terms of curves using algebraic geometry is more natural.
  - For example, there is another very important interpretation of the zeta function in terms of counting points on curves over $\mathbb{F}_q$ and its finite-degree extensions.
  - In order to exploit this correspondence as fully as possible (which is necessary in order to motivate the proof of the Riemann hypothesis for curves), we will now give a brisk treatment of algebraic varieties, with a particular focus on algebraic curves.

- First, the basics of affine space:
  - Definition: For a field k, we define affine n-space $\mathbb{A}^n(k) = \{(x_1, x_2, \dots, x_n) : x_i \in k\}$ to be the set of n-tuples of elements of k. The elements of $\mathbb{A}^n(k)$ are called points.
  - Definition: For $f \in k[x_1, \dots, x_n]$, we define the vanishing locus of f to be $V(f) = \{P \in \mathbb{A}^n(k) : f(P) = 0\}$, the set of points $P \in \mathbb{A}^n(k)$ where f vanishes. We extend this definition to subsets $T \subseteq k[x_1, \dots, x_n]$ by setting $V(T) = \cap_{f \in T} V(f) = \{P \in \mathbb{A}^n(k) : f(P) = 0 \text{ for all } f \in T\}$.
  - Exercise: Draw $V(x)$, $V(x^2)$, $V(y-x)$, $V(y-x^2)$, $V(xy)$, $V(x,y)$, and $V(y^2-x^3)$ in $\mathbb{A}^2(\mathbb{R})$.
  - Definition: For a subset $S \subseteq \mathbb{A}^n(k)$, we define the ideal of functions vanishing on S to be $I(S) = \{f \in k[x_1, \dots, x_n] : f(P) = 0 \text{ for all } P \in S\}$. It is easy to see that $I(S)$ is an ideal of $k[x_1, \dots, x_n]$ for any set S.
  - Exercise: Identify $I(S)$ in $\mathbb{R}[x,y]$ for $S = \{(t,0) : t \in \mathbb{R}\}$, $\{(t^2,t) : t \in \mathbb{R}\}$, $\{(1,1)\}$, $\{(0,0),(1,1)\}$, $\{(\cos t, \sin t) : t \in \mathbb{R}\}$, and $\{(t, \sin t) : t \in \mathbb{R}\}$.

- We have various properties of the maps V and I:

1. If I is the ideal generated by $T \subseteq k[x_1, \dots, x_n]$, then $V(T) = V(I)$. Thus, we need only consider the behavior of V on ideals, meaning that we will only consider I and V as maps I : [sets] $\to$ [ideals] and V : [ideals] $\to$ [sets].
2. $V(0) = \mathbb{A}^n(k)$, $V(1) = \emptyset$, and $V(x_1 - a_1, \dots, x_n - a_n) = \{(a_1, \dots, a_n)\}$.
3. $I(\emptyset) = k[x_1, \dots, x_n]$, $I(\mathbb{A}^n) = 0$ when k is infinite, and $I(\{(a_1, \dots, a_n)\}) = (x_1 - a_1, \dots, x_n - a_n)$.
4. $V(\cup_i I_i) = \cap_i V(I_i)$ and $V(IJ) = V(I) \cup V(J)$.
5. For ideals I and J, if $I \subseteq J$ then $V(I) \supseteq V(J)$, and for sets X and Y, if $X \subseteq Y$ then $I(X) \supseteq I(Y)$. (Thus, both I and V are inclusion-reversing.)
6. For any subset S of $k[x_1, \dots, x_n]$, $S \subseteq I(V(S))$ and $V(S) = V(I(V(S)))$.
7. For any subset X of $\mathbb{A}^n(k)$, $X \subseteq V(I(X))$ and $I(X) = I(V(I(X)))$. Furthermore, $I(X)$ is a radical[3] ideal.
   - Proofs: Exercises.

- Definition: For a field k, an affine algebraic set in $\mathbb{A}^n(k)$ is a subset of $\mathbb{A}^n(k)$ of the form $V(I)$ for some ideal I.
  - Examples: Single points $\{(a_1, \dots, a_n)\} = V(x_1 - a_1, \dots, x_n - a_n)$ are affine algebraic sets by (2) above. The sets $\{(t, 0) : t \in k\} = V(y)$ and $\{(t^2, t^3) : t \in k\} = V(y^2 - x^3)$ are affine algebraic sets.
  - By (4), we see that affine algebraic sets are closed under finite unions and arbitrary intersections, and (3) shows that $\mathbb{A}^n$ and $\emptyset$ are affine algebraic sets.
  - Thus, if we consider affine algebraic sets to be closed (with the open sets therefore being their complements), we obtain a topology on $\mathbb{A}^n(k)$. This topology is known as the Zariski topology.
  - By Hilbert's basis theorem, every ideal of $k[x_1, \dots, x_n]$ is finitely generated, so by (4) above, we see that every affine algebraic set is of the form $V(f_1) \cap V(f_2) \cap \dots \cap V(f_i)$ for some polynomials $f_1, \dots, f_i$. (Equivalently, the complements of the sets $V(f_i)$ form a base for the Zariski topology.)
  - It is natural to seek "minimal" elements under the Zariski topology.

- Definition: An affine algebraic set V is reducible if it can be written as $V = V_1 \cup V_2$ where $V_1, V_2 \neq V$, and it is irreducible otherwise.

- We have a few more properties:

8. V is irreducible if and only if $I(V)$ is a prime ideal of $k[x_1, \dots, x_n]$.
   - Proof: If $V = V_1 \cup V_2$ with $V_1, V_2 \neq V$, then $I(V_1)$ and $I(V_2)$ both properly contain V: if $f \in I(V_1) \backslash V$ and $g \in I(V_2) \backslash V$ then $fg \in I(V_1) \cap I(V_2) = I(V)$, meaning that $I(V)$ is not prime.
   - Conversely, if $fg \in I(V)$ with $f, g \notin I(V)$, we can take $V_1 = V \cap V(f)$ and $V_2 = V \cap V(g)$: then $V_1 \cup V_2 = V$ and $V_1, V_2 \neq V$ so V is reducible.

9. Any affine algebraic set V can be written uniquely as a union of irreducible affine algebraic sets $V_1 \cup V_2 \cup \dots \cup V_n$ such that $V_i \not\subseteq V_j$ for any $i \neq j$. (These sets V_i are the irreducible components of V.)
   - This is the geometric version of primary decomposition (generalizing the notion of prime factorization of elements).
   - Proof: For existence, let $\mathcal{F}$ be the collection of all V that cannot be written as a union of irreducible affine algebraic sets and consider the collection $\mathcal{I} = \{I(V) : V \in \mathcal{F}\}$, and suppose $\mathcal{F}$ is nonempty.
   - Since $\mathcal{I}$ is a collection of ideals in the Noetherian ring $k[x_1, \dots, x_n]$, it has a maximal element. The corresponding set V is then a minimal element of $\mathcal{F}$. If V is irreducible we obviously have a contradiction, and if V is not irreducible then it can be written as a proper union $V = V_1 \cup V_2$, but by minimality, V_1 and V_2 can both be written as a union of irreducible affine algebraic sets. In either case we get a contradiction, so $\mathcal{F}$ is empty.
   - We may freely assume that $V_i \not\subseteq V_j$ for any $i \neq j$ by throwing away any V_i that is a subset of another V_j.

[3] Recall that if I is an ideal of a commutative ring R, then the radical $\text{rad}(I) = \{r \in R : r^n \in I \text{ for some } n \geq 1\}$, and I is a radical ideal if $I = \text{rad}(I)$. (Note that $\text{rad}(I)$ is an ideal, as is easily seen via an application of the binomial theorem.)

- For uniqueness, suppose $V = W_1 \cup \cdots \cup W_k$ is another decomposition. Then $V_i = \cup_{j=1}^{k}(V_i \cap W_j)$ so since V_i is irreducible we must have $V_i \cap W_j = V_i$ for some j, meaning $V_i \subseteq W_j$. By symmetry we must also have W_j contained in some $V_{i'}$, but then $V_i \subseteq V_{i'}$ which forces $i = i'$ and then equality holds every, so $V_i = W_j$.
- Thus, each V is equal to some W. In the same way, we see each W is equal to some V, so we are done.

- Exercise: If k is finite, show that the irreducible affine algebraic sets in $\mathbb{A}^n(k)$ are $\emptyset$ and single points.
- Exercise: If k is infinite, show that the irreducible affine algebraic sets in $\mathbb{A}^2(k)$ are $\emptyset$, $\mathbb{A}^2(k)$, single points, and curves of the form $V(f)$ for a monic irreducible polynomial $f \in k[x, y]$. [Hint: Show that if $f, g \in k[x, y]$ are relatively prime, then (f, g) contains a nonzero polynomial in $k[x]$ and a nonzero polynomial in $k[y]$.]
- Although it may appear that I and V should behave like inverses, they are not quite.
  - For example, even in $\mathbb{A}^1(k)$, we have $V(x^2) = \{0\}$ so that $I(V(x^2)) = (x)$. The point here is that $I = (x^2)$ is not a radical ideal, and in this case, $I(V(I)) = \text{rad}(I)$.
  - However, even if I is radical, it is not always true that $I(V(I)) = \text{rad}(I)$: for example, in $\mathbb{A}^1(\mathbb{R})$ we have $V(1 + x^2) = \emptyset$ so that $I(V(1 + x^2)) = \mathbb{R}[x]$.
  - Indeed, there is no subset S of $\mathbb{A}^1(\mathbb{R})$ with $I(S) = (1+x^2)$ since the only set S with $I(S) \supseteq (1+x^2)$ is the empty set. The issue here is that $\mathbb{R}$ is not algebraically closed: if instead we work in $\mathbb{C}$, then $S = \{i, -i\}$ does have $I(S) = (1 + x^2)$.
- When the field k is algebraically closed, we do in fact solve all of the issues described above; this is the main content of Hilbert's Nullstellensatz, which we will discuss next time.

0.16 (Nov 4) Affine Varieties, Curves, and Projective Space

- We first show a weak version of the Nullstellensatz:
- Theorem (Hilbert's Nullstellensatz, weak version): If k is an algebraically closed field and I is a proper ideal of $k[x_1, \ldots, x_n]$, then $V(I) \neq 0$.
  - Proof: Since $I \subseteq J$ implies $V(I) \supseteq V(J)$, it suffices to show the result when I is a maximal ideal (since any non-maximal ideal is contained in a maximal ideal, and so its vanishing locus is at least as big as that of the maximal ideal).
  - If I is maximal, then $k[x_1, \ldots, x_n]/I$ is a field extension of k: we claim it is equal to k. Assuming this, then for each i we would have $\overline{x_i} = a_i$ for some $a_i \in k$ in the quotient ring, meaning that $x_i - a_i \in I$ and thus I contains $(x_1 - a_1, \ldots, x_n - a_n)$.
  - But then since $(x_1 - a_1, \ldots, x_n - a_n)$ is actually maximal (the quotient ring is isomorphic to the field k), we must have $I = (x_1 - a_1, \ldots, x_n - a_n)$, in which case $V(x_1 - a_1, \ldots, x_n - a_n) = \{(a_1, \ldots, a_n)\}$ is nonempty as claimed.
  - To establish the claimed statement, we first show the following result, which is known as Zariski's lemma:
  - Lemma: A field L that is finitely generated over k as a ring is finitely generated over k as a module.
    - Proof (Lemma): Induct on the number n of generators. For the base case $n = 1$, if $L = k[\alpha]$ as a ring then α must be algebraic over k since $k[\alpha]$ is not a field if α is transcendental (it does not contain $1/\alpha$). If α has minimal polynomial m of degree d, then L is spanned by $\{1, \alpha, \ldots, \alpha^{d-1}\}$ as a vector space over k, so L is finitely generated over k as a module.
    - For the inductive step, suppose the result holds for any extension generated by $n - 1$ elements, let $L = k[\alpha_1, \ldots, \alpha_n]$ as a ring and take $E = k[\alpha_1]$. Then L is finitely generated over E as a ring hence as a module by the inductive hypothesis, so we need only show that E is finitely generated over k as a module. If α_1 is algebraic over k the result follows by the above, so assume α_1 is transcendental.
    - Since L is finitely generated as a module over E, each α_i is integral over E, which is to say, it satisfies some equation $\alpha_i^{d_i} + c_{d_{i-1},i}\alpha_i^{d_i-1} + \cdots + c_{0,i} = 0$ for $c_\star \in E$. If r is the lcm of all of the denominators appearing in all of the elements of the rational functions $c_\star \in E = k(\alpha_1)$ then scaling each equation by r^{d_i} shows that $r\alpha_i$ is integral over E for each α_i.

    - * Then for any $\beta \in L = k[\alpha_1, \dots, \alpha_n]$ we see that there exists some N such that $r^N \beta$ is integral over E. But this claim is false for $\beta = \frac{1}{r+1}$: since r is transcendental over k, all of the elements $\frac{r^N}{r+1}$ are transcendental over k, and thus cannot be integral. This is the desired contradiction, so the result follows.
  - By Zariski's lemma, since $L = k[x_1, \dots, x_n]/I$ is a field extension of k that is finitely generated as a ring (since I is finitely generated) it is necessarily finitely generated as a module. This means it has finite degree over k, but since k is algebraically closed, we must have $L = k$ as required.

- Using the so-called "Rabinowitsch trick" we may bootstrap this statement into the full Nullstellensatz:

- Theorem (Hilbert's Nullstellensatz, strong version): If k is an algebraically closed field and I is any ideal of $k[x_1, \dots, x_n]$, then $I(V(I)) = \text{rad}(I)$.
  - Explicitly, if the polynomial g vanishes whenever $f_1, \dots, f_r$ vanish, then there exists some N and some $c_i \in k[x_1, \dots, x_n]$ such that $g^N = c_1 f_1 + \cdots + c_r f_r$.
  - Proof: It is easy to see that $\text{rad}(I) \subseteq I(V(I))$: if $f^n \in I$, then f^n and hence f vanishes on $V(I)$.
  - Now suppose $g \in I(V(f_1, \dots, f_r))$ and define the ideal $J = (f_1, \dots, f_r, x_{n+1}g - 1)$ of $k[x_1, \dots, x_n, x_{n+1}]$.
  - By hypothesis, $V(J) = \emptyset$ since all of $f_1, \dots, f_r$ vanishing implies that $x_{n+1}g - 1 = -1$.
  - Thus, the weak Nullstellensatz, J cannot be a proper ideal, so $1 \in J$. This means $1 = \sum_{i=1}^r A_i(x_1, \dots, x_{n+1}) f_i + B(x_1, \dots, x_{n+1})(x_{n+1}g - 1)$ for some polynomials A_i, B.
  - Plug in $x_{n+1} = 1/y$ and then clear denominators in y to obtain an equation of the form $y^N = \sum_{i=1}^r C_i(x_1, \dots, x_n, y) f_i + D(x_1, \dots, x_n, y)(g - y)$.
  - Now evaluate both sides at $y = g$: this yields $g^N = \sum_{i=1}^r C_i(x_1, \dots, x_n, g(x_1, \dots, g_n)) f_i$, which is precisely of the desired form.

- Per the Nullstellensatz and its various implications, we see that I and V give nice bijections between various sets in $\mathbb{A}^n(k)$ and ideals of $k[x_1, \dots, x_n]$.
  - By the full Nullstellensatz, since $I(V(I)) = \text{rad}(I)$, we obtain a correspondence between radical ideals and affine algebraic sets.
  - Furthermore, by the weak Nullstellensatz, if I is a proper ideal then $V(I)$ must contain some point $(a_1, \dots, a_n)$, whence I is contained in $I(\{(a_1, \dots, a_n)\}) = (x_1 - a_1, \dots, x_n - a_n)$. But since the quotient of $k[x_1, \dots, x_n]$ by $(x_1 - a_1, \dots, x_n - a_n)$ is isomorphic to k via the evaluation map $p \mapsto p(a_1, \dots, a_n)$, the latter ideal is maximal. Thus, the maximal ideals of $k[x_1, \dots, x_n]$ correspond precisely with points $(a_1, \dots, a_n)$.
  - Also, by the full Nullstellensatz, if I is a prime ideal, then $I(V(I)) = \text{rad}(I) = I$ since prime ideals are radical, and so by property (8) earlier, we see that $V(I)$ is irreducible. Thus, the prime ideals of $k[x_1, \dots, x_n]$ correspond with irreducible affine algebraic sets.
  - To summarize, we have the following correspondences:

    [Affine algebraic sets] $\underset{V}{\overset{I}{\rightleftarrows}}$ [Radical Ideals]

    [Irreducible affine algebraic sets] $\underset{V}{\overset{I}{\rightleftarrows}}$ [Prime Ideals]

    [Points of $\mathbb{A}^n(k)$] $\underset{V}{\overset{I}{\rightleftarrows}}$ [Maximal Ideals]

- Now we can bring function fields into the discussion.

- Definition: If k is algebraically closed, an irreducible affine algebraic set in $\mathbb{A}^n(k)$ is called an affine variety. The coordinate ring of an affine variety V is the ring $\Gamma(V) = k[x_1, \dots, x_n]/I(V)$, and its associated field of rational functions (or function field) $k(V)$ is the field of fractions of $\Gamma(V)$.
  - Exercise: Let $\mathcal{F}(V, k)$ be the ring of k-valued functions on V. We say $f \in \mathcal{F}(V, k)$ is a polynomial function if there exists $g \in k[x_1, \dots, x_n]$ such that $f(P) = g(P)$ for all $P \in V$. Show that $\Gamma(V)$ is the set of equivalence classes of polynomial functions under the relation $g_1 \sim g_2$ if $g_1(P) = g_2(P)$ for all $P \in V$.

- By the exercise above, the coordinate ring of V can be thought of as the collection of distinct polynomial functions on V, and thus the field of rational functions is, quite explicitly, the collection of rational functions on V.

- Rational functions can have poles, which are points $P \in V$ where the function is not defined.

- Definition: If V is an affine variety, we say $f \in k(V)$ is defined at a point P if $f = a/b$ for some $a, b \in \Gamma(V)$ and $b(P) \neq 0$. If f is defined at P, its value $f(P)$ is the ratio $a(P)/b(P) \in k$. The local ring of V at P, denoted $\mathcal{O}_P(V)$, is the set of rational functions $f \in k(V)$ that are defined at P.
  - As we have essentially discussed already in the context of the function field $F(t)$ in one variable, the local ring $\mathcal{O}_P(V)$ is in fact a local ring, with maximal ideal $m_P(V)$ given by the elements $f \in \mathcal{O}_P(V)$ that vanish at P.
  - The points P for which f is not defined are the poles of f, since they are necessarily zeroes of its denominator.
  - Exercise: Show that $\Gamma(V) = \cap_{P \in V} \mathcal{O}_P(V)$: in other words, that a function with no poles is a polynomial. (Note of course that k is algebraically closed!)

- We will emphasize here that there may be numerous ways to write $f = a/b$ as a quotient of polynomials, and it may be necessary to work with different "equivalent" formulas in order to verify that f is defined at a particular point P.

- Example: Consider the affine variety $V = V(y^2 - x^2 + 1)$ in $\mathbb{A}^2(k)$ for $k = \mathbb{C}$ and the rational function $f = \dfrac{x-1}{y} \in k(V)$.
  - It is clear from the expression $f = \dfrac{x-1}{y}$ that f is defined at all points $P = (x, y) \in V$ where $y \neq 0$.
  - However, because $\Gamma(V) = k[x, y]/(y^2 - x^2 + 1)$, we see that $y^2 = x^2 - 1$ in $\Gamma(V)$, by factoring and rearranging we see that $\dfrac{x-1}{y} = \dfrac{y}{x+1}$ inside $k(V)$. Therefore, f is also equal to $\dfrac{y}{x+1}$, and this latter expression shows that f is also defined at the point $(1, 0)$.
  - On the other hand, there is no way to rewrite $f = \dfrac{x-1}{y}$ in such a way that it is defined at $(-1, 0)$: if $\dfrac{x-1}{y} = \dfrac{p}{q}$ then $(x-1)q = yp$ but then evaluating both sides at $P = (-1, 0)$ produces $-2q(P) = 0$, which is a contradiction.
  - Remark: More generally, the same argument shows that if the expression for $f(P)$ is of the form $a/0$ for $a \neq 0$, then f is not defined at P. (If, of course, we obtain an expression $0/0$, then f could possibly be defined at P.)

- Definition:If V is an affine variety with function field $k(V)$, its dimension is defined to be the transcendence degree of $k(V)$ over k. An affine curve is an affine variety of dimension 1.
  - Examples:$V(y - x)$ and $V(y^2 - x^2 + 1)$ are affine curves in $\mathbb{A}^2(k)$.
  - If we think of $V = V(I)$ as being cut out from $\mathbb{A}^n(k)$ by the generators of I, then the dimension (as defined above) agrees with the intuitive topological sense of the dimension of $V(I)$ as a (hyper)surface, when $k = \mathbb{C}$.

- We outline some facts about affine plane curves (i.e., affine curves in $\mathbb{A}^2(k)$):
  1. Via the correspondence $C \mapsto V(f)$, an affine plane curve C is the same as a nonconstant monic irreducible polynomial $f \in k[x, y]$. We define the degree of C to be the degree of the corresponding polynomial f.
     - Proof: As noted in an exercise earlier, the irreducible affine sets in $\mathbb{A}^2(k)$ are $\emptyset$ (dimension 0), single points (dimension 0), $\mathbb{A}^2(k)$ (dimension 2), and the sets of the form $V(f)$ where f is a monic irreducible polynomial (these are the only sets of dimension 1, so they are the only curves).

2. If P is a point of the affine curve $C = V(f)$, we say P is a singular point if $f_x(P) = f_y(P) = 0$, and otherwise we say P is a nonsingular point (or smooth point or simple point). We say that C itself is smooth if all points of C are smooth points.

    - The main idea here is that a point P is singular if and only if C does not have a well-defined tangent line at P.
    - To find the tangent line(s) to a curve at a point P, we simply expand the defining polynomial f as a local Taylor series centered at $P = (x_0, y_0)$, i.e., as $f = a_{0,0} + a_{1,0}(x - x_0) + a_{0,1}(y - y_0) + a_{2,0}(x - x_0)^2 + a_{1,1}(x - x_0)(y - y_0) + a_{0,2}(y - y_0)^2 + \cdots$. Then the tangent lines are obtained by factoring the lowest-degree homogeneous component appearing in the factorization.
    - In particular, since $a_{0,0} = f(P) = 0$, $a_{1,0} = f_x(P)$, and $a_{0,1} = f_y(P)$ by the usual Taylor expansion, we see that there is a unique tangent line precisely when the linear term does not vanish (i.e., P has multiplicity 1), which is to say, precisely when $f_x(P)$ and $f_y(P)$ are not both zero.
    - Example: The point $(0, 0)$ lies on the variety $V(x + x^3 - 2y - y^5)$. Writing the curve locally near $(0, 0)$ yields $f = (x - 2y) + x^3 - y^5$, and the lowest-degree homogeneous component is $x - 2y$. Here, the curve has a unique tangent line at $(0, 0)$ given by $x - 2y = 0$ (which one may check explicitly using calculus).
    - Example: The variety $V(y^2 - x^2 - x^3)$ has a singular point at $(0, 0)$. Writing the curve locally near $(0, 0)$ yields $f = -x^2 + y^2 - x^3$, and the lowest-degree homogeneous component is $(-x^2 + y^2) = (-x + y)(-x - y)$. Here, the curve has two different tangent lines, $y = x$ and $y = -x$, which is made very clear by actually graphing $y^2 = x^3 + x^2$ (the curve crosses itself at $(0, 0)$).
    - Example: The variety $V(y^2 - x^3)$ has a singular point at $(0, 0)$. Writing the curve locally near $(0, 0)$ yields $f = y^2 - x^3$, and the lowest-degree homogeneous component is y^2. Here, the curve has a double tangent line $y = 0$, which can be seen by graphing $y^2 = x^3$ (the curve has a cusp at $(0, 0)$).
    - The degree of the lowest term with a nonzero coefficient in the local expansion of f at P is called the multiplicity of P. One may show that for sufficiently large n, the multiplicity of C at P is equal to $\dim_k(m_P^{n+1}/m_P^n)$, where m_P is the maximal ideal of the local ring $\mathcal{O}_P$ at P.

3. A point P is on an affine curve C is smooth if and only if the local ring $\mathcal{O}_P(C)$ is a discrete valuation ring.

    - Proof: We will show more specifically that if $L : \{ax + by + c = 0\}$ is any line through P not tangent to C at P, then the image of L in $\mathcal{O}_P(C)$ is a local uniformizer.
    - For this, apply a linear change of variables to move P to $(0, 0)$, to make y the tangent direction, and x the line through P not tangent to C at P. Then (regardless of the behavior at P), the maximal ideal $m_P(V)$ of the local ring $\mathcal{O}_P(V)$ is generated by x and y.
    - Furthermore, following the linear change of variables, the local expansion of f is $f = y+$[terms of degree $\geq$ 2], which is of the form $yg(x, y) + x^2h(x, y)$ for some polynomials $g, h \in k[x, y]$ where $g(0, 0) = 1$. In the coordinate ring $\Gamma(C) = k[x, y]/(f)$, we have $\overline{yg} = -\overline{x^2h}$, and so $y = -x^2\dfrac{h}{g}$ in the function field $k(V)$. Since $g(0, 0) = 1$ is not zero, this shows y is defined at P and in fact evaluates to zero at P, so $y \in m_P$ and y is a multiple of x. Thus, x generates m_P, which establishes the claimed result.
    - For the converse, we invoke the fact noted above that for sufficiently large n, the multiplicity of C at P is equal to $\dim_k(m_P^{n+1}/m_P^n)$, where m_P is the maximal ideal of the local ring $\mathcal{O}_P$ at P. Here, since the multiplicity is equal to 1 by hypothesis, the valuation of y must actually equal 1.

4. If $C_1 = V(f)$ and $C_2 = V(g)$ are two distinct affine plane curves sharing no common component, then their intersection $C_1 \cap C_2 = V(f, g)$ is finite. We may associate a divisor to this intersection $C_1 \cap C_2$ as $\sum_{P \in C_1 \cap C_2} n_P P$, where n_P is the intersection number of $C_1 \cap C_2$ at P given by $n_P = \dim_k \mathcal{O}_P(\mathbb{A}^2)/(f, g)$.

    - For polynomials in one variable, the ideal (f, g) is principal and generated by the gcd of f and g. (One may check that the intersection number at a point P, under the definition above, is the power of $t - P$ that divides their gcd.)
    - For polynomials in two variables (f, g) will no longer be principal, but it still carries the natural sense of being a "common divisor". Thus, we can think (roughly) of the divisor $\sum_{P \in C_1 \cap C_2} n_P P$ as describing the precise way in which the curves intersect.
    - It is not particularly obvious that this value $\dim_k \mathcal{O}_P(\mathbb{A}^2)/(f, g)$ is really the right definition. It is not hard to see that the value is invariant under linear changes of coordinates, and that the intersection

number is 1 whenever P is a simple point of C_1 and C_2 where C_1 and C_2 meet transversally (i.e., their tangent lines at P are different). It is also additive when we take unions of curves.

- We will not really use this particular formulation of divisors; it is merely some motivation for how divisors arise in a fairly natural way in the context of curves.

- With all of this in hand, we can see that if C is a smooth affine curve over an algebraically closed field k, then the points of C all correspond to primes of the associated function field $k(C)$, since by hypothesis the local rings are all DVRs.
  - However, even for $\mathbb{A}^1(k)$, there is one prime missing, namely, the prime at ∞, which does not arise as the local ring corresponding to any point of C.
- To fix this issue, we instead need to work instead with projective varieties, which will neatly solve this issue of "missing primes".

0.17 (Nov 8) Homework #2 Discussion

0.18 (Nov 15) Projective Space and Projective Curves

- First, the basics of projective space:
  - Definition: For a field k, we define projective n-space $\mathbb{P}^n(k) = \{[x_0 : x_1 : \cdots : x_n] : x_i \in k \text{ not all zero}\}/\sim$, where $P \sim Q$ if $P = \lambda Q$ for some nonzero $\lambda \in k$. Equivalently, $\mathbb{P}^n(k)$ is the set of lines through the origin in $\mathbb{A}^{n+1}(k)$.
  - We use the notation $[x_0 : x_1 : \cdots : x_n]$ to evoke the idea of considering only the ratios between the coordinates, since (for example) in $\mathbb{P}^1(k)$ the points $[1 : 1]$ and $[2 : 2]$ are the same. The coordinates x_i of a point $P \in \mathbb{P}^n(k)$ are not well-defined, but since the equivalence is only up to scaling by a nonzero constant, the statement "$x_i = 0$" is still well-defined, as are the ratios x_i/x_j.
  - For the set $U_i = \{[x_0 : x_1 : \cdots : x_n] : x_i = 1\}$, we can see that U_i looks exactly like $\mathbb{A}^n(k)$ (if we just delete the coordinate $x_i = 1$), and $\mathbb{P}^n(k) = \cup_{i=0}^n U_i$.
  - The complement of the set U_i is the hyperplane $x_i = 0$, and it looks exactly like $\mathbb{P}^{n-1}$ (if we just delete the coordinate $x_i = 0$).
  - Thus, somewhat informally, we have $\mathbb{P}^n(k) = \mathbb{A}^n(k) \cup \mathbb{P}^{n-1}(k)$, where we can think of $\mathbb{A}^n(k)$ as being the points with $x_n = 1$ and $\mathbb{P}^{n-1}(k)$ as being the points with $x_n = 0$.
  - Example: We have $\mathbb{P}^1(k) = \{[x : 1] : x \in k\} \cup \{[1, 0]\}$, which looks like $\mathbb{A}^1$ along with a point at ∞.
- We cannot sensibly plug a projective point into an arbitrary polynomial, since the result is not well-defined even up to scaling[4]. However, for our purposes we only need to describe vanishing sets, which (at least) have a chance of being better behaved.
  - For example, for $f(x, y) = x^2 - y^2$, it is reasonable to say that the projective point $[1 : 1]$ should be in the vanishing set for f: not only do we have $f(1, 1) = 0$, but in fact for any point $[t : t]$ equivalent to $[1 : 1]$, we have $f(t, t) = 0$ as well.
  - On the other hand, for $g(x, y) = x - y^2$, it is less reasonable to say that $[1 : 1]$ should be in the vanishing set for f: although $f(1, 1) = 0$, in general $f(t, t) = t - t^2$ need not be zero for other values of t.
  - One option would be to say that $P \in \mathbb{P}^n(k)$ is in the vanishing set of $f \in k[x_0, \ldots, x_n]$ if $f(P) = 0$ for *all* choices of coordinates for P.
  - Exercise: Suppose k is an infinite field, $P \in \mathbb{A}^{n+1}\backslash\{0\}$, and $f \in k[x_0, \ldots, x_n]$. If we write $f = f_0 + f_1 + \cdots + f_d$ for homogeneous[5] polynomials f_i of degree i, show that $f(\lambda P) = 0$ for all $\lambda \in k^\times$ if and only if $f_i(P) = 0$ for all i. [Hint: Use linear algebra and the fact that Vandermonde determinants are nonvanishing.]

[4] For example, if $f(x, y) = x^2 + y^2$, we could try to define f on the projective point $[1 : 2]$ by plugging in $x = 1$ and $y = 2$, thus yielding the value 5, but this clashes with attempting to define f on $[2 : 4]$ by plugging in $x = 2$ and $y = 4$ to obtain 20, since $[1 : 2] = [2 : 4]$ as points in $\mathbb{P}^1(k)$.

[5] Recall that a polynomial is homogeneous of degree d if all of its monomial terms have total degree d. For example, $x^2y - 3x^3 + xyz$ is homogeneous of degree 3.

- Per the exercise above, we see that when k is an infinite field, requiring $f(P) = 0$ for all choices of coordinates for P is equivalent to requiring that all of the homogeneous components of f vanish.
- For consistency with finite fields (which have nonzero polynomials that vanish everywhere, causing issues with the argument above), we instead define the vanishing of a polynomial f on a projective point P in terms of homogeneous components.

- Definition: If $f \in k[x_0, \dots, x_n]$ is a polynomial with $f = f_0 + f_1 + \cdots + f_d$ for homogeneous polynomials f_i of degree i, we say that f vanishes at $P \in \mathbb{P}^n(k)$, and write $f(P) = 0$, if $f_i(P) = 0$ for each i.
  - Note that $f_i(\lambda P) = \lambda^i f(P)$ so the vanishing condition on f_i does not depend on which equivalent coordinates are used for P.
  - Example: The polynomial $f(x, y) = x^2 - y^2$ vanishes at the projective point $[1 : 1]$ since its only nonzero homogeneous component $x^2 - y^2$ vanishes at P, but the polynomial $g(x, y) = x - y^2$ does not since its homogeneous components are x and $-y^2$ and these do not vanish at $[1 : 1]$.
  - The main theme is that when we want to work with polynomials in projective space, we want to consider only homogeneous polynomials.

- Now that we have given a reasonable definition of vanishing for projective points, we can define the projective versions of the operators V and I:

- Definition: If S is any set of polynomials in $k[x_0, \dots, x_n]$, we define the vanishing locus $V(S) = \{P \in \mathbb{P}^n(k) : f(P) = 0 \text{ for all } f \in S\}$. Conversely, if X is any set of points in $\mathbb{P}^n(k)$, we define the ideal of functions vanishing on X as $I(X) = \{f \in k[x_0, \dots, x_n] : f(P) = 0 \text{ for all } P \in X\}$.
  - Exercise: Identify $V(x_0)$, $V(x_0^2)$, $V(x_1 - x_0)$, $V(x_1 - x_0^2)$, $V(x_1^2 - x_0^2)$, $V(x_0, x_1)$, $V(x_0, x_1, x_2)$, and $V(x_0 x_1 - x_2^2)$ in $\mathbb{P}^2(k)$.
  - Exercise: Show that all of the basic properties of the affine operators I and V also hold for the projective I and V (suitably modified):
    1. If I is the ideal generated by $T \subseteq k[x_0, \dots, x_n]$, then $V(T) = V(I)$.
    2. $V(0) = \mathbb{P}^n(k)$, $V(1) = \emptyset$, and $V(\{a_i x_j - a_j x_i\}_{0 \le i,j \le n}) = \{[a_0 : a_1 : \cdots : a_n]\}$.
    3. $I(\emptyset) = k[x_0, \dots, x_n]$, $I(\mathbb{P}^n) = 0$ when k is infinite, and $I(\{[a_0 : a_1 : \cdots : a_n]\}) = (\{a_i x_j - a_j x_i\}_{0 \le i,j \le n})$.
    4. $V(\cup_i I_i) = \cap_i V(I_i)$ and $V(IJ) = V(I) \cup V(J)$.
    5. For ideals I and J, if $I \subseteq J$ then $V(I) \supseteq V(J)$, and for sets X and Y, if $X \subseteq Y$ then $I(X) \supseteq I(Y)$.
    6. For any subset S of $k[x_0, \dots, x_n]$, $S \subseteq I(V(S))$ and $V(S) = V(I(V(S)))$.
    7. For any subset X of $\mathbb{P}^n(k)$, $X \subseteq V(I(X))$ and $I(X) = I(V(I(X)))$. Furthermore, $I(X)$ is a radical ideal.

- Owing to our definition of vanishing in terms of homogeneous components, the ideals of sets in $\mathbb{P}^n(k)$ have an additional property:

- Definition: An ideal I of $k[x_0, \dots, x_n]$ is homogeneous if, for any $f \in I$ with homogeneous decomposition $f = f_0 + f_1 + \cdots + f_d$, it is true that each component $f_i \in I$.
  - It is easy to see that $I(X)$ is homogeneous, since for any $f = f_0 + f_1 + \cdots + f_d \in I(X)$, by definition of vanishing we see that for any $P \in X$ we have $f_i(P) = 0$ and so $f_i \in I(X)$.
  - Exercise: Show that an ideal I of $k[x_0, \dots, x_n]$ is homogeneous if and only if I is generated by finitely many homogeneous polynomials.

- We also have a projective version of the Nullstellensatz, which is essentially the same as the affine version except that we must account for the fact that the vanishing locus of the ideal $(x_0, x_1, \dots, x_n)$ in $\mathbb{P}^n$ is empty since $[0 : 0 : \cdots : 0]$ is not a point of $\mathbb{P}^n$:

- Theorem (Projective Nullstellensatz): Let k be an algebraically closed field and I be a homogeneous ideal of $k[x_0, \dots, x_n]$. Then the following hold:

1. (Weak) $V(I) = \emptyset$ if and only if I contains all monomials of sufficiently large degree, if and only if $\text{rad}(I)$ contains $(x_0, \dots, x_n)$.
2. (Strong) If $V(I) \neq \emptyset$, then $I(V(I)) = \text{rad}(I)$.
   - The proofs are similar to those of the affine Nullstellensatz, and are left as exercises.
   - Owing to the fact that its vanishing locus is trivial, and thus can essentially be ignored when doing computations, the ideal $(x_0, x_1, \dots, x_n)$ in $k[x_0, \dots, x_n]$ is called the irrelevant ideal.

- Next, we define algebraic sets, varieties, and coordinate rings in $\mathbb{P}^n$. The ideas proceed essentially the same way:
- Definition: A projective algebraic set is a set in $\mathbb{P}^n(k)$ of the form $V(I)$ for some ideal I of $k[x_0, \dots, x_n]$. A projective algebraic set V is reducible if it can be written as $V = V_1 \cup V_2$ where $V_1, V_2 \neq V$, and it is irreducible otherwise. A projective variety is an irreducible projective algebraic set.
  - By essentially the same arguments as in the affine case, V is irreducible if and only if $I(V)$ is a prime ideal of $k[x_0, \dots, x_n]$, and any projective algebraic set can be written uniquely as a union of irreducible components $V_1 \cup V_2 \cup \cdots \cup V_n$ such that $V_i \not\subseteq V_j$ for any $i \neq j$.
- Definition: If V is a projective variety, then its (homogeneous) coordinate ring is the integral domain $\Gamma(V) = k[x_0, \dots, x_n]/I(V)$.
  - As before, we may decompose the polynomials $f \in \Gamma(V)$ as $f = f_0 + f_1 + \cdots + f_d$ where f_i is homogeneous of degree i.
  - Since $I(V)$ is prime, the coordinate ring is an integral domain, so its fraction field is well defined. Unlike in the affine case, however, the elements of this fraction field do not generally determine functions on V, because a ratio of polynomials need not be a function on V.
  - The first obvious issue is that for a ratio $\dfrac{f}{g} = \dfrac{f_0 + f_1 + \cdots + f_d}{g_0 + g_1 + \cdots + g_d}$, the various homogeneous terms in the numerator and denominator will not transform the same way if we choose a different representative for the projective point $P \in V$ at which we are attempting to evaluate f/g. (For example: what is the value of $\dfrac{x + y^2}{x + y}$ at the projective point $[1 : 1]$?)
  - To handle this issue, we must only have a single homogeneous component in the numerator and denominator. But even here, in order for the ratio to be well-defined, the degrees of the numerator and denominator must be equal.
  - When we restrict to rational functions of this form, however, we do obtain well-defined functions on projective points: if f, g are both homogeneous of degree d, then $\dfrac{f(\lambda P)}{g(\lambda P)} = \dfrac{\lambda^d f(P)}{\lambda^d g(P)} = \dfrac{f(P)}{g(P)}$, so the ratio f/g is well defined regardless of the representative of P we use.
- Definition: If V is a projective variety, its function field $k(V)$ is the set of elements z in the fraction field of $\Gamma(V)$ such that z can be written in the form $z = \dfrac{f}{g}$ for some homogeneous polynomials $f, g \in k[x_0, \dots, x_n]$ of the same degree. We say z is defined at a point $P \in V$ if $z = \dfrac{f}{g}$ for some g with $g(P) \neq 0$. The local ring of V at P is $\mathcal{O}_P(V) = \{z \in k(V) : z \text{ is defined at } P\}$ with maximal ideal $m_P(V) = \{z \in \mathcal{O}_P(V) : z(P) = 0\}$.
  - As in the affine case, we may require different expressions $z = f/g$ at different points P.
- Example: Consider the affine variety $V = V(Y^2 + Z^2 - X^2)$ in $\mathbb{P}^2(\mathbb{C})$ and the rational function $f = \dfrac{X - Z}{Y} \in k(V)$.
  - It is clear from the expression $f = \dfrac{X - Z}{Y}$ that f is defined at all points $P = [X : Y : Z] \in V$ where $Y \neq 0$, which is to say, at all points of the form $[X : 1 : Z]$ after rescaling. The only points of V with $Y = 0$ are those with $X^2 = Z^2$, which gives two points: $[1 : 0 : 1]$ and $[1 : 0 : -1]$.

- However, because $\Gamma(V) = k[x,y]/(Y^2 + Z^2 - X^2)$, we see that $Y^2 = X^2 - Z^2$ in $\Gamma(V)$, by factoring and rearranging we see that $\frac{X-Z}{Y} = \frac{Y}{X+Z}$ inside $k(V)$. Therefore, f is also equal to $\frac{Y}{X+Z}$, and this latter expression shows that f is also defined at the point $[1:0:1]$ (and in fact it vanishes there).
- On the other hand, there is no way to rewrite $f = \frac{X-Z}{Y}$ in such a way that it is defined at $[1:0:-1]$: if $\frac{X-Z}{Y} = \frac{p}{q}$ then $(X-Z)q = Yp$ but then evaluating both sides (as polynomials in X, Y, Z) at $X = 1$, $Y = 0$, $Z = -1$ produces $-2q(1,0,-1) = 0$, which is a contradiction since this means $q(P) = 0$.
- Remark: Note that this is just the projective version of the example we did earlier for the affine variety $V = V(y^2 + 1 - x^2)$ in $\mathbb{A}^2$.

- As clearly indicated by the similarity of the calculations in the example above and the nearly-identical affine example from earlier, there is quite a lot of interplay between projective and affine spaces.
  - One such correspondence is obtained by viewing $\mathbb{P}^n$ as the lines through the origin in $\mathbb{A}^{n+1}$, so for any set S in $\mathbb{P}^n$ we may write down the set of its corresponding points in $\mathbb{A}^{n+1}$ by converting the point $[x_0 : x_1 : \cdots : x_n]$ to the point $(x_0, x_1, \ldots, x_n)$.
  - Explicitly, if $S \subseteq \mathbb{P}^n$, the cone $C(S)$ of S in $\mathbb{A}^{n+1}$ is the set $\{(x_0, x_1, \ldots, x_n) : [x_0 : x_1 : \cdots : x_n] \in S\} \cup \{(0, 0, \ldots, 0)\}$.
  - Exercise: When V is a nonempty projective algebraic variety, show that $I_{\text{affine}}(C(V)) = I_{\text{projective}}(V)$, and when I is a homogeneous ideal with $V_{\text{projective}}(I) \neq \emptyset$, show that $C(V_{\text{projective}}(I)) = V_{\text{affine}}(I)$.

- Although the cone of a variety shares the same underlying ideal, and thus has the same coordinate ring and function field, its dimension is different.
  - We would like instead to think of $\mathbb{P}^n$ as being $\mathbb{A}^n$ plus a hyperplane at ∞, and so an affine variety in $\mathbb{A}^n$ should give rise to one that looks essentially the same in $\mathbb{P}^n$, except for having some additional points in the hyperplane at ∞.
  - The main idea, as exemplified by comparing the example above to its affine version, is that of homogenization and dehomogenization.

- Definition: If $F \in k[x_0, x_1, \ldots, x_n]$ is a polynomial, its dehomogenization with respect to x_0 is $F_* = F(1, x_1, \ldots, x_n)$. Inversely, if $f \in k[x_1, \ldots, x_n]$ is a polynomial, its homogenization with respect to x_0 is $f^* = x_0^{\deg(f)} f(x_1/x_0, x_2/x_0, \ldots, x_n/x_0)$.
  - More explicitly, if $f \in k[x_1, \ldots, x_n]$ has homogeneous decomposition $f = f_0 + f_1 + \cdots + f_d$, then $f^* = x_0^d f_0 + x_0^{d-1} f_1 + \cdots + f_d$.
  - Example: The homogenizations of $x_1^2 + x_2$, $4 + x_1 x_3 - 3x_4^5$, and 1 are $x_1^2 + x_0 x_2$, $4x_0^5 - x_0^3 x_1 x_3 - 3x_4^5$, and 1 respectively.
  - Example: The dehomogenizations of $x_0^2 + 3x_0 x_1 + x_1 x_2$, $x_0^3 + 4x_0 x_2^2 + x_3^3$, and x_0^2 are $1 + 3x_1 + x_1 x_2$, $1 + 4x_2^2 + x_3^3$, and 1 respectively.
  - The main idea is that dehomogenizing removes the variable x_0 by setting it equal to 1 (thereby usually creating a non-homogeneous polynomial in the remaining variables $x_1, \ldots, x_n$) while homogenizing takes a non-homogeneous polynomial in $x_1, \ldots, x_n$ and makes it homogeneous in $x_0, x_1, \ldots, x_n$ by using the extra variable x_0 to make all of the terms have the same degree.
  - Homogenization and dehomogenization are essentially inverses of one another, aside from occasionally losing powers of x_0.
  - Exercise: Show that $(FG)_* = F_* G_*$, $(fg)^* = f^* g^*$, $(f^*)_* = f$, $(F_*)^* = F/x_0^{v_{x_0}(f)}$, $(F+G)_* = F_* + G_*$, and $x_0^{\deg(f)+\deg(g)-\deg(f+g)}(f+g)^* = x_0^{\deg(g)} f^* + x_0^{\deg(f)} g^*$.

- Our interest here is that homogenizing an affine equation creates a projective one, and dehomogenizing a projective equation yields an affine one, thereby giving a correspondence between affine varieties and projective varieties.

- Motivating Example: Homogenizing the affine equation $x_1 + x_2 = 1$ yields the projective equation $x_1 + x_2 = x_0$. An affine point (x_1, x_2) satisfying $x_1 + x_2 = 1$ then yields a projective point $[1 : x_1 : x_2] = [y_0 : y_1 : y_2]$ satisfying $y_1 + y_2 = y_0$. If we compare the affine points to the projective ones, we see that the projective variety consists of the points $[1 : x_1 : x_2]$, which all correspond to affine points, along with one additional point $[0 : 1 : -1]$ which we think of as the point at ∞ on this line.
- Motivating Example: Dehomogenizing the projective equation $x_2^2 x_0 = x_1^3 + x_1 x_0^2$ yields the affine equation $x_2^2 = x_1^3 + x_1$. A projective point $[x_0 : x_1 : x_2]$ satisfying $x_2^2 x_0 = x_1^3 + x_1 x_0^2$ then yields an affine point $(y_1, y_2) = (x_1/x_0, x_2/x_0)$ satisfying $y_2^2 = y_1^3 + y_1$, as long as $x_0 \neq 0$. When we dehomogenize, the projective points $[x_0 : x_1 : x_2]$ with $x_0 = 0$ "disappear" from the affine curve (note here that there is only one such point, namely $[0 : 1 : 0]$).

- In order to make this precise, we can extend homogenization to ideals and then to affine algebraic sets:
- Definition: If $I = (f_1, \dots, f_k)$ is an ideal of $k[x_1, \dots, x_n]$, the homogenization of I is the ideal $I^* = (f_1^*, \dots, f_k^*)$ generated by the homogenizations of the generators of I. Conversely, if J is an ideal of $k[x_0, x_1, \dots, x_n]$, the dehomogenization of J is the ideal $J_* = \{g_* : g \in J\}$ of dehomogenizations of the elements of J (and is generated by the dehomogenizations of the generators of J).
  - Per the definition, we see immediately that I^* is a homogeneous ideal.
  - Note also that the homogenization of I is not simply the set of homogenizations of elements of I (the latter is not generally an ideal, since it is not closed under scaling by x_0), but rather the ideal generated by these homogenizations; this is why the two definitions appear slightly different. The dehomogenization of J is, however, just the set of dehomogenizations of elements of J.
  - Example: For $I = (x_1^2, x_1 + x_2^2)$, we have $I^* = (x_1^2, x_0 x_1 + x_2^2)$.
  - Example: For $J = (x_0^2 x_1, x_1^2 + x_2^2)$, we have $J_* = (x_1, x_1^2 + x_2^2) = (x_1)$.
- Definition: If V is an affine algebraic set, then for $I = I_{\text{affine}}(V)$ we define the homogenization of V to be the projective algebraic set $V^* = V_{\text{projective}}(I^*)$. Conversely, if W is a projective algebraic set, then for $J = I_{\text{projective}}(W)$ we define the dehomogenization of W to be the affine algebraic set $W_* = V_{\text{affine}}(J_*)$.
  - Example: If $W = V(x_0^2 x_1 + x_0 x_2^2)$ in $\mathbb{P}^2$, then $J = I(W) = (x_0^2 x_1 + x_0 x_2^2)$ so $J_* = (x_1 + x_2^2)$ and thus $W_* = V(x_1 + x_2^2)$ in $\mathbb{A}^2$.
  - Example: If $V = V(x_1 + x_2^2)$ in $\mathbb{A}^2$, then $I = I(V) = (x_1 + x_2^2)$ so $I^* = (x_0 x_1 + x_2^2)$ and then $V^* = V(x_0 x_1 + x_2^2)$ in $\mathbb{P}^2$.
  - In the pair of examples above, we can see that $(W_*)^* = V(x_0 x_1 + x_2^2)$ is not equal to W (the defining polynomial is now missing a factor of x_0, resulting in a loss of most points of the form $[0 : x_1 : x_2]$ from W), whereas $(V^*)_*$ is equal to V.
- These examples typify the general behavior of homogenizing and dehomogenizing: up to some minor issues regarding losing powers of x_0, these operations are essentially inverses, and thus allow us to go back and forth between $\mathbb{P}^n$ and $\mathbb{A}^n$.
- Proposition (Homogenization of Affine Sets): Let k be an algebraically closed field, let $H_\infty = V(x_0)$ denote the hyperplane at ∞ inside $\mathbb{P}^n$, and let $U_0 = \mathbb{P}^n \backslash H_\infty$ be its complement. Also let V, V_1, V_2 be affine algebraic sets and W, W_1, W_2 be projective algebraic sets.
  1. For any V we have $(V^*)_* = V$.
  2. For $\varphi_0 : \mathbb{A}^n \to \mathbb{P}^n$ with $\varphi_0(x_1, \dots, x_n) = [1 : x_1 : \cdots : x_n]$, we have $V^* \cap U_0 = \varphi_0(V)$.
  3. If $V_1 \subseteq V_2 \subseteq \mathbb{A}^n$ then $V_1^* \subseteq V_2^* \subseteq \mathbb{P}^n$, and if $W_1 \subseteq W_2 \subseteq \mathbb{P}^n$ then $(W_1)_* \subseteq (W_2)_* \subseteq \mathbb{A}^n$.
  4. If V is irreducible in $\mathbb{A}^n$ then V^* is irreducible in $\mathbb{P}^n$.
  5. If $V \subseteq \mathbb{A}^n$ then V^* is the smallest projective algebraic set in $\mathbb{P}^n$ that contains $\varphi_0(V)$.
  6. If $V = \cup_i V_i$ is an irreducible decomposition in $\mathbb{A}^n$, then $V^* = \cup_i V_i^*$ is an irreducible decomposition in $\mathbb{P}^n$.
  7. If V is a nonempty proper algebraic subset of $\mathbb{A}^n$, then no component of V^* lies in or contains H_∞.

8. If W is a nonempty proper algebraic subset of $\mathbb{P}^n$ and no component of W lies in or contains H_∞, then W_* is a proper algebraic subset of $\mathbb{A}^n$ and $(W_*)^* = W$. More generally, for any W we have $W = (W_*)^* \cup (H_\infty \cap W)$.
   - Proofs: (1)-(3) are immediate from properties of (de)homogenization and V and I.
   - (4) follows by observing that if I is prime then so is I^*.
   - For (5), suppose that $W \subseteq \mathbb{P}^n$ contains $\varphi_0(V)$. Then for any $f \in I(W)$, we must have $f_* \in I(V)$, so $f = x_0^{v_{x_0}(f)}(f_*)^* \in I(V)$ as well. Therefore, $I(W) \subseteq I(V)^*$ so W contains V^*.
   - For (6), note that (4) shows that each V_i^* is irreducible and (3) shows that none of them contain another. Also, $\cup_i V_i^*$ is also the smallest projective algebraic set that contains $\cup_i \varphi_0(V_i) = \varphi_0(V)$ so it equals V^* by (5).
   - For (7), we may assume by (6) that V is irreducible. Then V^* is not a subset of H_∞ because (2) tells us that $V^* \cap U_0 = \varphi_0(V) \neq \emptyset$. For the other part, by the Nullstellensatz, since V is a proper subset of $\mathbb{A}^n$ we have $I(V) \neq 0$: then for any nonzero $f \in I(V)$ we see that f^* is not a multiple of x_0. But if V^* contains H_∞, then $I(V)^* \subseteq I(H_\infty) = (x_0)$, which is a contradiction.
   - For (8), assume again by (6) that W is irreducible. Since $\varphi_0(W_*) \subseteq W$, it is enough to show that $W \subseteq (W_*)^*$, which follows by (3) from showing $I(W_*) \subseteq I(W)$. So suppose $f \in I(W_*)$: then $f^n \in I(W)_*$by the Nullstellensatz, which means $x_0^k(f^n)^* \in I(W)$ for some power k. But since $I(W)$ is prime and $x_0 \notin I(W)$ because W is not contained in $V(x_0) = H_\infty$, this means $(f^n)^* = (f^*)^n \in I(W)$, so again by primality this means $f^* \in I(W)$ as required. The second part follows by analyzing the various possible cases for $W \cap H_\infty$.

- Per (1), (4), and (5) above, if we identify V with its image $\varphi_0(V)$ in $\mathbb{P}^n$, we can view V^* as being the projective closure of the affine variety V.
   - In particular, if we ignore varieties contained entirely within the hyperplane H_∞, then (7) and (8) tell us that we have a natural bijection between nonempty affine and projective varieties.
   - Based on the simplicity of the relationship between V and V^*, it is natural to expect that the function fields of V and V^* should be the same, which is in fact the case:

- Proposition (Equivalence of Function Fields): If V is an affine variety with projective closure V^*, then the function fields $k(V)$ and $k(V^*)$ are isomorphic. Furthermore, if P is any point on V with corresponding point P^* on V^*, then the isomorphism of $k(V)$ and $k(V^*)$ also yields an isomorphism of $\mathcal{O}_P(V)$ with $\mathcal{O}_{P^*}(V^*)$ and of $m_P(V)$ with $m_{P^*}(V^*)$.
   - Proof: Suppose that $\overline{f} \in \Gamma(V^*)$ is a homogeneous polynomial in $k[x_0, \dots, x_n]/I(V^*)$. By dehomogenizing, we get a residue class $\overline{f}_* \in k[x_1, \dots, x_n]/I(V)$; note that this residue class is well defined by (1) from the proposition above.
   - We may extend this map on coordinate rings to one on the fraction fields $k(V^*)$ to $k(V)$ by taking $\psi(f/g) = f_*/g_*$ for any homogeneous polynomials f and g of the same degree. (Note that this map is simply dehomogenization; the point is that it is still well defined up to equivalence.)
   - On the other hand, we also have a natural inverse map from $k(V)$ to $k(V^*)$ by homogenizing: explicitly, we may take $\tau(p/q) = x_0^{\deg(q)-\deg(p)} p^*/q^*$, which is a quotient of homogeneous polynomials of the same degree and is therefore an element of $k(V^*)$.
   - Since ψ is clearly a ring homomorphism, we see that $k(V^*)$ is isomorphic to $k(V)$. It is also easy to see that this isomorphism restricts to an isomorphism on the corresponding local rings, since $g(P) \neq 0$ if and only if $g_*(P^*) \neq 0$, and $f(P) = 0$ if and only if $f_*(P^*) = 0$.

0.19 (Nov 18) Rational Maps and Extensions of Function Fields

- We can also homogenize and dehomogenize with respect to other variables (e.g., x_1, x_2, ... , x_n); there is no particular reason to use x_0 specifically, other than convenience.
   - In particular, since the intersection of the hyperplanes $V(x_0)$, $V(x_1)$, ... , $V(x_n)$ in $\mathbb{P}^n$ is empty, their complements cover $\mathbb{P}^n$. Each of these complements corresponds to a copy of $\mathbb{A}^n$ obtained by dehomogenizing with respect to the corresponding variable.

- Thus, if W is any projective variety, we may analyze any point of W "affinely", inside one of the possible dehomogenizations of W.
- Since our interest is in plane curves, we will use uppercase letters (X,Y,Z) for the variables of projective equations and lowercase letters (x,y,z) for the variables of affine equations, and indicate which variable is being homogenized or dehomogenized.
- For example, we may dehomogenize the projective equation $Y^2Z = X^3 + 3XZ^2$ with respect to X to obtain the affine equation $y^2z = 1+3z^2$, or with respect to Y to obtain the affine equation $z = x^3+3xz^2$, or with respect to Z to obtain the affine equation $y^2 = x^3 + 3x$.
- By dehomogenizing, we see that the projective point $[1 : 2 : 1]$ on $Y^2Z = X^3 + XZ^2$ corresponds to the affine point $(y, z) = (2, 1)$ on $y^2z = 1 + z^2$, the affine point $(x, z) = (1/2, 1/2)$ on $z = x^3 + xz^2$, and the affine point $(x, y) = (1, 2)$ on $y^2 = x^3 + 3x$.
- Likewise, the projective point $[0 : 1 : 0]$ on $Y^2Z = X^3 + XZ^2$, which disappears on the two affine curves $y^2z = 1 + z^2$ and $y^2 = x^3 + x$ since the dehomogenized variable is zero, still corresponds to a point $(x, z) = (0, 0)$ on $z = x^3 + xz^2$.
- We think of the projective point $[0 : 1 : 0]$ as being a "point at infinity" on the affine curve $y^2 = x^3 + x$; if we want to study its local ring, we can simply work instead with the other dehomogenization $z = x^3+xz^2$, where it corresponds to $(0, 0)$.

- We can see, therefore, that working with projective curves nearly addresses the issue of "missing" points at infinity, which (in the language of function fields) corresponds to primes at infinity.
  - Furthermore, by working with appropriate dehomogenizations, we may import all of our terminology and results about affine plane curves from earlier without much change: e.g., projective plane curves are of the form $V(f)$ for a homogeneous irreducible polynomial $f \in k[X, Y, Z]$, a point of V is nonsingular when f_x, f_y, f_z are not all zero, and so forth.
  - Exercise: Show that if $f \in k[X_1, \dots, X_n]$ is homogeneous of degree d, then $X_1 f_{X_1} + \cdots + X_n f_{X_n} = df$. (This is a famous result of Euler.) Deduce that for a homogenous $f \in k[X, Y, Z]$, if two of f_X, f_Y, f_Z are zero then the third is as well.
  - Let us now be more precise about the correspondence between smooth projective curves and function fields over an algebraically closed field k.
  - If we have a smooth projective curve C, then its associated function field $k(C)$ is a function field that has transcendence degree 1 over its constant field k (in our language). Each point P on the curve has an associated local ring $\mathcal{O}_P(C)$ that corresponds to a prime of the function field; each of these local rings is a DVR since C is smooth.
  - If we have a function field K/k of transcendence degree 1, to construct the associated curve, first choose any prime P of K (necessarily of degree 1 since k is algebraically closed). Then the associated DVR $\mathcal{O}_P$ is finitely generated as a ring over k, so since it is a domain, it is isomorphic to $k[x_1, \dots, x_n]/I$ for some prime ideal I. If we take $C = V(I)$, we obtain an affine curve whose function field is K. With some additional work, one can eventually show that the projective closure of this affine curve (whose function field is also K, as we showed above) is smooth.

- So far, we have mostly been assuming that the constant field k is algebraically closed. In particular, since we want to focus on function fields over $\mathbb{F}_q$, we need to remove this assumption.
  - Explicitly, suppose V is a variety over k and E is a subfield of k. We would naively like to define the set of E-points of V as $V \cap \mathbb{A}^n(E)$ if V is affine, and as $V \cap \mathbb{P}^n(E)$ if V is projective.
  - We may make this more precise using Galois actions: specifically, assuming that $k = \overline{E}$, then the Galois group of k/E acts naturally on the k-points of V.

- Definition: Let E be a field with algebraic closure k, and let $G = \text{Gal}(E/k)$. If V is a variety over k, we define the E-points of V to be the set of points of V over k that are fixed by G.
  - Explicitly, P is an E-point of V if and only if $\sigma(P) = P$ for all $\sigma \in \text{Gal}(E/k)$.
  - The set of E-points of V is precisely $V \cap \mathbb{A}^n(E)$ if V is affine, and is $V \cap \mathbb{P}^n(E)$ if V is projective, since the given condition is equivalent to saying that all of the coordinates of the point lie in E.

- Example: For $E = \mathbb{F}_5$ and $V = V(y^2 - x^2 - 1)$ in $\mathbb{A}^2$, the set of E-points of V is $(x, y) = (0, 1)$, $(0, 4)$, $(2, 0)$, and $(3, 0)$.
- Example: For $E = \mathbb{F}_3$ and $V = V(Y^2Z^2 - XZ^3 - X^4)$ in $\mathbb{P}^2$, the set of E-points of V is $[X : Y : Z] = [0 : 0 : 1]$, $[0 : 1 : 0]$, $[1 : 0 : 2]$.
- We can also define the elements of the coordinate ring and function field of V over E, namely, as the elements of $\Gamma(V)$ and $k(V)$ fixed by E, respectively.

- Definition: If E is a field with algebraic closure k, we say that a variety V is defined over E if $I(V)$ can be generated by polynomials with coefficients in E.
  - We will think of all varieties as implicitly being defined over an algebraically closed field, even if it they are actually defined over a subfield.
  - Thus, we may meaningfully speak of the points of V on arbitrary algebraic extensions of E.

- We now discuss maps between varieties. The most natural starting point is to consider maps defined by polynomials:

- Definition: If V is an affine variety in $\mathbb{A}^n(k)$ and W is an affine variety in $\mathbb{A}^m(k)$, a map $\varphi : V \to W$ is called a polynomial map from V to W if there exist polynomials $T_1, \dots, T_m \in k[x_1, \dots, x_n]$ such that $\varphi(a_1, \dots, a_n) = (T_1(a_1, \dots, a_n), T_2(a_1, \dots, a_n), \dots, T_m(a_1, \dots, a_n))$.
  - Example: The map $\varphi : \mathbb{A}^1 \to \mathbb{A}^1$ with $\varphi(a) = a^2 + a$ is a polynomial map, as is the map $\varphi : \mathbb{A}^1 \to \mathbb{A}^3$ with $\varphi(a) = (a, a^2, a^3)$.
  - Example: The map $\varphi : V(x^2 + y^2 - 1) \to \mathbb{A}^1$ with $\varphi(x, y) = x$ is a polynomial map.
  - Example: The map $\varphi : \mathbb{A}^2 \to V(x^2 + y^2 - z^2)$ with $\varphi(a, b) = (2ab, a^2 - b^2, a^2 + b^2)$ is a polynomial map. Note that this map is well-defined because $(2ab)^2 + (a^2 - b^2)^2 - (a^2 + b^2)^2$ is indeed zero for all $(a, b) \in \mathbb{A}^2$, so $(2ab, a^2 - b^2, a^2 + b^2) \in V(x^2 + y^2 - z^2)$.
  - Example: The map $\varphi : V(y - x^2) \to V(z - xy)$ with $\varphi(x, y) = (x, y, x^3)$ is a polynomial map. Note that this map is well-defined because for all $(x, y) \in V(y - x^2)$ we have $y = x^2$, and then $(x, y, x^3) \in V(z - xy)$.

- Polynomial maps are equivalent to homomorphisms of coordinate rings:

- Proposition (Polynomial Maps and Coordinate Rings): If V and W are affine varieties, then any polynomial map $\varphi : V \to W$ induces a homomorphism $\tilde{\varphi} : \Gamma(W) \to \Gamma(V)$ on coordinate rings via "plugging in": $\tilde{\varphi}(f) = f \circ \varphi$. Conversely, any homomorphism $\tilde{\varphi} : \Gamma(W) \to \Gamma(V)$ is induced by a unique polynomial map $\varphi : V \to W$ with $\tilde{\varphi}(f) = f \circ \varphi$.
  - Proof: First suppose $\varphi : V \to W$ is a polynomial map. For any $f \in k[x_1, \dots, x_n]$, define $\psi(f) = f \circ \varphi$. Clearly, ψ is a ring homomorphism (since it is just polynomial evaluation). Furthermore, this map ψ descends to a well-defined map $\tilde{\varphi} : \Gamma(W) \to \Gamma(V)$: this follows by noting that if $f \in \Gamma(W)$ is the $I(W)$-residue of a polynomial $G(x_1, \dots, x_n)$, then $\tilde{\varphi}(f) = f \circ \varphi$ is the $I(V)$-residue of the polynomial $G(T_1, \dots, T_m)$.
  - For the converse, we can simply reconstruct the map φ from its action on each variable x_i. Explicitly, suppose that $\tilde{\varphi} : \Gamma(W) \to \Gamma(V)$ is a homomorphism. Then $\tilde{\varphi}$ maps $x_i + I(W)$ to some polynomial $T_i + I(V)$ for each $1 \le i \le m$. Then the map $\varphi(a_1, \dots, a_n) = (T_1(a_1, \dots, a_n), \dots, T_m(a_1, \dots, a_n))$ is a polynomial map from $\mathbb{A}^n$ to $\mathbb{A}^m$, and it induces a map $\hat{\varphi} : \Gamma(\mathbb{A}^m) \to \Gamma(\mathbb{A}^n)$. From the information given we know that $\hat{\varphi}(I(W)) \subseteq I(V)$, so $\varphi(V) \subseteq W$. Thus, $\varphi|_V$ is a polynomial map from V to W, and $\tilde{\varphi}(f) = f \circ \varphi$ as required.

- Definition: If V and W are affine varieties, a polynomial map $\varphi : V \to W$ is an isomorphism if it possesses an inverse polynomial map $\psi : W \to V$ (i.e., with $\varphi \circ \psi = \mathrm{id}_W$ and $\psi \circ \varphi = \mathrm{id}_V$).
  - By the above, we see that V and W are isomorphic if and only if their coordinate rings are isomorphic as k-algebras (i.e., if their coordinate rings are isomorphic as rings where the isomorphism also fixes k).
  - Example: The map $\varphi : V(x - y) \to V(x - 2y)$ with $\varphi(x, y) = (2x, y)$ is an isomorphism with inverse $\psi(x, y) = (x/2, y)$.

- Exercise: Show that the isomorphisms $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ are the invertible affine linear transformations, of the form $\varphi(x) = Ax + b$ where A is an invertible $n \times n$ matrix and b is any vector of constants. (Hint: First show that the degree of each coordinate in φ and ψ must be 1.)

- We would like to write down a similar definition for projective varieties, which we can do at the cost of a bit of added complexity.
  - The most immediate issue is that we need to insist that all of the polynomials T_i be homogeneous of the same degree, in order to ensure that "plugging in" to a polynomial map is well defined.
  - However, this is not the only obstruction; difficulties also arise in the event that all of the polynomials T_i vanish simultaneously, since then the resulting value does not yield a well-defined point in $\mathbb{P}^1$.

- Definition: If V and W are projective varieties, a rational map from V to W is a map of the form $\varphi = [\varphi_0 : \varphi_1 : \cdots : \varphi_m]$ where the $\varphi_i \in k[x_0, \dots, x_n]$ are homogeneous polynomials of the same degree, and such that for all $f \in I(W)$, we have $f \circ \varphi = f(\varphi_0(x_0, \dots, x_n), \dots, \varphi_m(x_0, \dots, x_n)) \in I(V)$.
  - If φ is a rational map, then for $P \in V$ we can evaluate $\varphi(P) = [\varphi_0(P) : \varphi_1(P) : \cdots : \varphi_m(P)] \in W$ as long as not all of the values $\varphi_i(P)$ are zero. We can see that this value $\varphi(P)$ is well defined because the φ_i are homogeneous of the same degree, and $\varphi(P) \in W$ precisely because $f \circ \varphi \in I(V)$ for any $f \in I(W)$.
  - To illustrate, consider the map $\varphi : V(X^2 + Y^2 - Z^2) \to \mathbb{P}^1$ given by $\varphi[X : Y : Z] = [X + Z : Y]$. On its face, this would appear to be a perfectly well-defined function, since for any equivalent representative $[\lambda X : \lambda Y : \lambda Z]$ we have $\varphi[\lambda X : \lambda Y : \lambda Z] = [\lambda X + \lambda Z : \lambda Y] = [X + Z : Y] = \varphi[X : Y : Z]$.
  - However, for the point $P = [1 : 0 : -1]$ in $V(X^2 + Y^2 - Z^2)$, the definition states $\varphi(P) = [0 : 0]$, which is not a point of $\mathbb{P}^1$.
  - Notice, though, that if we work inside $\Gamma(V)$, we see that $[X + Z : Y] = [(X + Z)(X - Z) : Y(X - Z)] = [-Y^2 : Y(X - Z)] = [-Y : X - Z]$ and this latter expression is defined at $[1 : 0 : -1]$ since it evaluates to $[0 : 2]$.
  - We would like to extend our interpretation of the value of $\varphi(P)$ in a way that allows us to make these kinds of manipulations.

- Definition: If $\varphi : V \to W$ is a rational map, we say that $\varphi = [\varphi_0 : \cdots : \varphi_m]$ is defined at P if there exist homogeneous polynomials $\psi_0, \dots, \psi_n$ of the same degree such that $\varphi_i \psi_j \equiv \varphi_j \psi_i \pmod{I(V)}$ for all pairs (i, j), and where $\psi_i(P) \neq 0$ for some i, and we write $\varphi(P) = [\psi_0(P) : \cdots : \psi_m(P)]$.
  - The idea here is that, inside $\Gamma(V)$, we view the homogeneous coordinates $[\varphi_0 : \cdots : \varphi_m]$ and $[\psi_0 : \cdots : \psi_m]$ as being projectively equivalent.
  - We call these "rational maps" because if we work affinely, they arise from rational functions.

- Definition: If V and W are varieties, a morphism from V to W is a rational map that is defined at all points of V. An isomorphism is a morphism possessing an inverse morphism.
  - If $\varphi : V \to W$ is a morphism, then φ induces an injective homomorphism on function fields $\tilde{\varphi} : k(W) \to k(V)$ via composition: $\tilde{\varphi}(f) = f \circ \varphi$.
  - As in the affine case for polynomial maps, the converse is true as well: any injective k-algebra homomorphism on function fields $\tilde{\varphi} : k(W) \to k(V)$ (i.e., a ring homomorphism fixing k) yields a morphism $\varphi : V \to W$.
  - Example: The map $\varphi : V(Y^2 Z - X^3 - XZ^2) \to \mathbb{P}^1$ given by $\varphi[X : Y : Z] = [Y : Z]$ is a morphism. (Note that there are no points of $V(Y^2 Z - X^3 - XZ^2)$ where φ is undefined, since if $Y = Z = 0$ then X would also be zero.)
  - Example: The map $\varphi : V(X^2 + Y^2 - Z^2) \to \mathbb{P}^1$ given by $\varphi[X : Y : Z] = [X + Z : Y]$ is a morphism, since it is defined at all points of $V(X^2 + Y^2 - Z^2)$ as shown earlier.
  - Example: The map $\psi : \mathbb{P}^1 \to V(X^2 + Y^2 - Z^2)$ given by $\psi[S : T] = [S^2 - T^2 : 2ST : S^2 + T^2]$ is a morphism. In fact, it is the inverse of the previous morphism, since we have $(\varphi \circ \psi)[S : T] = \varphi[S^2 - T^2 : 2ST : S^2 + T^2] = [2S^2 : 2ST] = [S : T]$ and $(\psi \circ \varphi)[X : Y : Z] = \psi[X + Z : Y] = [(X + Z)^2 - Y^2 : 2Y(X + Z) : (X + Z)^2 + Y^2] = [2X(X + Z) : 2Y(X + Z) : 2Z(X + Z)] = [X : Y : Z]$.

- Example: The map $\psi : V(Y^2Z - X^3 - XZ^2) \to V(Y^2Z - X^3 - XZ^2)$ given by $\psi[X : Y : Z] = [X : -Y : Z]$ is a morphism. In fact, it is an isomorphism, since it is its own inverse.
- Example: If k has characteristic q and V is defined over $\mathbb{F}_q$, the map $\varphi : V \to V$ given by $\varphi[X_0 : X_1 : \cdots : X_n] = [X_0^q : X_1^q : \cdots : X_n^q]$ is a morphism called the Frobenius morphism.
- Example: The map $\varphi : \mathbb{P}^1 \to \mathbb{P}^2$ given by $\varphi[X : Y] = [X^2 : XY : Y^2]$ is a morphism giving an embedding of $\mathbb{P}^1$ into $\mathbb{P}^2$ (it is an example of the general family of d-uple embeddings). The image of φ is the variety $V(XZ - Y^2)$.
- Example: The map $\psi : V(Y^2Z - X^3 - Z^3) \to V(Y^2Z - X^3 - Z^3)$ given by $\psi[X : Y : Z] = [2XY(Y^2 - 9Z^2) : Y^4 + 18Y^2Z^2 - 27Z^4 : 8Y^3Z]$ is a morphism. (Actually checking that it is well-defined is rather unpleasant, but it does work out!) This particular morphism arises as "multiplication by 2" on the elliptic curve $V(Y^2Z - X^3 - Z^3)$.

- Restricting now to the case of projective curves, we have the following facts:

1. If C_1 is a smooth projective curve, then any rational map $\varphi : C_1 \to C_2$ is automatically a morphism.
   - The idea here is that if P is any point on C_1, then since C_1 is smooth at P (meaning that the local ring $\mathcal{O}_P(V)$ is a DVR), we may choose a local uniformizer t at P.
   - Then we can rescale the components of $\varphi = [\varphi_0 : \varphi_1 : \cdots : \varphi_m]$ by an appropriate power of t in order to make the minimum valuation among the φ_i equal to zero, at which point we see that φ is defined at P.
2. If $\varphi : C_1 \to C_2$ is a nonconstant morphism of projective curves, then φ is surjective, and $k(C_1)$ is a finite-degree extension of $\tilde{\varphi}(k(C_2))$.
   - The first statement follows from the result that the image of a morphism of a projective variety is itself a projective variety (this is usually phrased as saying that projective varieties are complete). Thus, the image $\varphi(C_1)$ is a subvariety of C_2: if its dimension is 1 then since C_2 is irreducible this means $\varphi(C_1) = C_2$, and otherwise if its dimension is 0 then $\varphi(C_1)$ would be a single point and φ would be constant, which we assumed it was not.
   - The fact that $k(C_1)$ is an extension of $k(C_2)$ follows from the fact that φ is surjective, and the fact that the extension has finite degree follows because both $k(C_1)$ and $k(C_2)$ have transcendence degree 1 over k.
3. If $\iota : k(C_2) \to k(C_1)$ is an injection fixing k, then there is a unique nonconstant morphism $\varphi : C_1 \to C_2$ such that $\tilde{\varphi} = \iota$.

- By putting (3) together with (2), and then letting Galois groups act on both sides (so as to remove the requirement that k be algebraically closed) we obtain our claimed equivalence of categories from much earlier:

1. (Objects) Algebraic function fields K/k of transcendence degree 1 where $K \cap \overline{k} = k$
   (Morphisms) Field injections fixing 1 (up to isomorphism)
2. (Objects) Smooth projective curves defined over k
   (Morphisms) Non-constant rational maps defined over k (up to isomorphism)

- We will now study morphisms of curves in the context of function fields: from the equivalence above, this is the same as studying field extensions of algebraic function fields K/k_2 over E/k_1.
  - The main advantage here is that in addition to exploiting number-theoretic ideas about field extensions and ramification, we can also exploit geometric ideas about morphisms between curves.
  - We may view function field extensions as consisting of two parts: constant field extensions (where we simply extend scalars in the constant field, by going from $E/\mathbb{F}_q$ to $E\mathbb{F}_{q^n}/\mathbb{F}_{q^n}$ for some n) and field extensions fixing the constant field (going from $E/\mathbb{F}_q$ to $K/\mathbb{F}_q$ where K is a finite-degree extension of E).
  - On the curves side, constant field extensions are vaguely trivial: they correspond simply to viewing the curve as being defined over $\mathbb{F}_{q^n}$ rather than over $\mathbb{F}_q$. (Very usefully, we can also think of constant field extensions in terms of the Frobenius morphism, as we will discuss in a moment.)
  - Field extensions fixing the constant field correspond to morphisms from one curve to another; for this reason these extensions are called geometric extensions (since they arise "geometrically" rather than from merely changing the field of definition).

0.20 (Nov 22) Primes in Extensions

- With all of this in mind, if K is a function field over $\mathbb{F}_q$, we can give another very useful interpretation of the zeta function $\zeta_K(s)$ in terms of counting points.
  - Explicitly, if X is the smooth projective curve corresponding to $K/\mathbb{F}_q$, then for $P \in X(\overline{\mathbb{F}_q})$ we define the degree of P to be the degree of the residue field $\mathcal{O}_P/m_P$ over $\mathbb{F}_q$.
  - The connection with the degree of a divisor of K is as follows: a divisor $D = \sum_P n_P P$ is defined over $\mathbb{F}_{q^n}$ precisely when it is fixed by the nth power of the Frobenius map.
  - For an automorphism σ we have $\sigma(D) = \sum_P n_P \sigma(P) = \sum_P n_{\sigma^{-1}(P)} P$, and so we see $\sigma(D) = D$ precisely when $n_{\sigma^{-1}(P)} = n_P$ for all points P. By repeatedly applying σ, we see that this is equivalent to saying that all of the Galois conjugates of P have the same coefficient n_P.
  - Thus, for example, a point of X defined over $\mathbb{F}_{q^2}$ has a single nontrivial Galois conjugate $\sigma(P)$, and the corresponding prime divisor over $\mathbb{F}_q$ is $P + \sigma(P)$, has degree 2.
  - In the same way, a point of X defined over $\mathbb{F}_{q^n}$ has a total of n Galois conjugates (including itself), and so the corresponding prime divisor over $\mathbb{F}_q$ is $P + \sigma(P) + \cdots + \sigma^{n-1}(P)$, which has degree n.

- Definition: If X is a smooth projective curve over $\overline{\mathbb{F}_q}$, the zeta function of X is defined as $\zeta_X(s) = \sum_{P \in X} \dfrac{1}{N(P)^s} = \prod_{P \in X}(1 - N(P)^{-s})^{-1}$ where $N(P) = q^{\deg(P)}$ as usual.
  - It is not hard to see, per the discussion above, that $\zeta_X(s) = \zeta_K(s)$ where K is the function field of X.

- We can give another formula for $\zeta_X(s)$ in terms of the cardinalities $N_n = \#X(\mathbb{F}_{q^n})$, the number of $\mathbb{F}_{q^n}$-points of X.
  - Explicitly, since $X(\overline{\mathbb{F}_q}) = \bigcup_{n \geq 1} X(\mathbb{F}_{q^n})$, and $X(\mathbb{F}_{q^a}) \subseteq X(\mathbb{F}_{q^b})$ whenever $a|b$, we can see that with $u = q^{-s}$ as usual, we have

$$\log \zeta_X(s) = \sum_{P \in X} -\log(1 - u^{\deg P}) = \sum_{n=1}^{\infty} \sum_{P \in X} \frac{u^{n \deg P}}{n} = \sum_{n=1}^{\infty} \frac{\#X(\mathbb{F}_{q^n})}{n} u^n$$

    where the last equality follows because a point P shows up a total of a times in the sum for $n = k$ whenever P is defined over $\mathbb{F}_{q^{k/a}}$.
  - Example: For $C = \mathbb{P}^1$, we have $\#C(\mathbb{F}_{q^n}) = q^n + 1$. Then $\log \zeta_C(s) = \sum_{n=1}^{\infty} \dfrac{q^n + 1}{n} u^n = -\log(1-u) - \log(1 - qu) = \dfrac{1}{(1 - q^{-s})(1 - q^{1-s})}$, which agrees with our usual zeta function for the rational function field $\mathbb{F}_q(t)/\mathbb{F}_q$.
  - Notice also that if φ is the q-power Frobenius map, then the fixed points of φ^n are the points of $X(\mathbb{F}_{q^n})$.
  - By using the Weil conjectures we can make this quite explicit: from $\zeta_X(u) = \dfrac{\prod_{i=1}^{2g}(1 - \pi_i u)}{(1-u)(1-qu)}$ we have $\log \zeta_X(u) = \sum_{i=1}^{2g} \log(1 - \pi_i u) - \log(1-u) - \log(1-qu) = \sum_{n=1}^{\infty} [\sum_{i=1}^{2g} \pi_i^n - 1 - q^n] u^n$, and so we get the formula $\#C(\mathbb{F}_{q^n}) = q^n + 1 - \sum_{i=1}^{2g} \pi_i^n$.
  - By the Riemann hypothesis, we obtain the inequality $|\#C(\mathbb{F}_{q^n}) - q^n - 1| \leq 2g q^{n/2}$, which tells us that for large n, the number of points on $C(\mathbb{F}_{q^n})$ is $q^n + O(q^{n/2})$. (This is really just a rephrasing of our results for the general prime number theorem for function fields.)

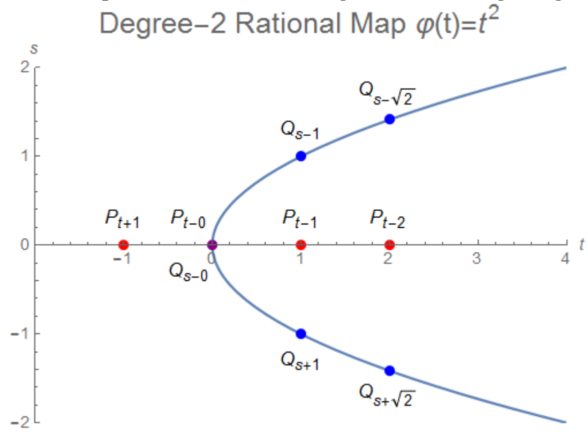
- Now we will discuss geometric extensions of function fields, which are finite-degree extensions of the form L/K where L and K have the same constant field $\mathbb{F}_q$.
  - If $\tilde{P}$ is a prime of L with associated valuation ring $\mathcal{O}_{\tilde{P}}$ in L, then it is not hard to verify that the intersection $\mathcal{O}_{\tilde{P}} \cap K$ is a discrete valuation ring of K and that its maximal ideal is $\tilde{P} \cap K$; explicitly, this follows by noting that the restriction of the valuation $v_{\tilde{P}}$ to K is still a discrete valuation, possibly after rescaling to ensure that there is an element of valuation 1 (the precise details may be found in Theorem 15.26 of Dummit/Foote).

- Definition: If L/K is a geometric extension of function fields and $\tilde{P}$ is a prime of L with associated valuation ring $\mathcal{O}_{\tilde{P}}$ in L and P is a prime of K, we say $\tilde{P}$ lies over P and write $\tilde{P}|P$ if $\mathcal{O}_P = K \cap \mathcal{O}_{\tilde{P}}$ and $P = K \cap \tilde{P}$.
  - This is the natural analogue of the corresponding situation for prime ideals of the ring of integers of a number field: if L/K is an extension of number fields with corresponding rings of integers $\mathcal{O}_L$ and $\mathcal{O}_K$, then each prime ideal $\tilde{P}$ of $\mathcal{O}_L$ lies over a unique prime ideal P of $\mathcal{O}_K$ with $P = \mathcal{O}_K \cap \tilde{P}$.
  - Furthermore, if we extend P to be an ideal of $\mathcal{O}_L$ (i.e., by considering the ideal of $\mathcal{O}_L$ generated by P), since $\mathcal{O}_L$ is a Dedekind domain we see that $\tilde{P}$ contains P and thus $\tilde{P}$ divides P (as an ideal); this is the source of the notation $\tilde{P}|P$.
  - For example, in the ring of integers $\mathbb{Z}[\sqrt{7}]$ of the quadratic field $\mathbb{Q}(\sqrt{7})/\mathbb{Q}$, the prime ideal $\tilde{P} = (3, \sqrt{7}-1)$ of $\mathbb{Z}[\sqrt{7}]$ lies above the prime ideal $P = (3)$ of $\mathbb{Z}$. Quite visibly we have $(3) \subseteq (3, \sqrt{7}-1)$ as ideals of $\mathbb{Z}[\sqrt{7}]$, and indeed one may calculate that $(3) = (3, \sqrt{7}-1) \cdot (3, \sqrt{7}+1)$, so that the ideal $(3, \sqrt{7}-1)$ divides the ideal (3).
- There are two important quantities associated to $\tilde{P}|P$:
- Definition: If $\tilde{P}$ lies over P, the relative degree $f = f(\tilde{P}|P)$ is the vector space dimension $\dim_{\mathcal{O}_P/P}(\mathcal{O}_{\tilde{P}}/P)$, and the ramification index $e = e(\tilde{P}|P)$ is the integer with $P\mathcal{O}_{\tilde{P}} = (\tilde{P})^e$.
  - The relative degree and ramification index are the function-field analogues of the corresponding quantities in number fields.
  - The relative degree is well defined because $\mathcal{O}_{\tilde{P}}/P$ is a finite-degree extension of $\mathcal{O}_P/P$, since the degree is at most the degree of L/K by the simple observation that any basis of $\mathcal{O}_{\tilde{P}}/P$ over $\mathcal{O}_P/P$ remains linearly independent in L/K (we will prove a stronger version of this statement later).
  - Furthermore, the ramification index is well defined because $P\mathcal{O}_{\tilde{P}}$ is an ideal of the DVR $\mathcal{O}_{\tilde{P}}$ hence is some power of its maximal ideal $\tilde{P}$. As a consequence, we see that $v_P(a) = e \cdot v_{\tilde{P}}(a)$ for all $a \in K$.
  - Exercise: Show that the relative degree and ramification index compose in towers; explicitly, that if $\tilde{\tilde{P}}|\tilde{P}$ and $\tilde{P}|P$ in a tower of extensions $M/L/K$, then $e(\tilde{\tilde{P}}|P) = e(\tilde{\tilde{P}}|\tilde{P}) \cdot e(\tilde{P}|P)$ and $f(\tilde{\tilde{P}}|P) = f(\tilde{\tilde{P}}|\tilde{P}) \cdot f(\tilde{P}|P)$.
- Example: Consider the quadratic extension of rational function fields given by $\mathbb{R}(t)/\mathbb{R}(t^2)$ with constant field $\mathbb{R}$.
  - When we view this field extension L/K as a morphism of curves $\varphi : C_1 \to C_2$, we have the map $\varphi : \mathbb{P}^1(\mathbb{R}) \to \mathbb{P}^1(\mathbb{R})$ with $\varphi[x_0 : x_1] = [x_0^2 : x_1^2]$, which when we dehomogenize in x_0 we see is the affine map $\varphi : \mathbb{A}^1(\mathbb{R}) \to \mathbb{A}^1(\mathbb{R})$ with $\varphi(t) = t^2$ (i.e., the squaring map).
  - For concreteness, if we view $\mathbb{R}(t^2)$ as being the rational function field $\mathbb{R}(s)$, the corresponding map of function fields from $R(s)$ to $R(t)$ takes $s \mapsto t^2$.
  - Here are some degree-1 primes of K, with corresponding primes lying above them in L, and their values of e and f:

Prime of K	Prime of L	$e(Q_i\|P_i)$	$f(Q_i\|P_i)$
P_{t-1}, prime at $t-1$ $(= s^2-1)$	Q_{s-1}, prime at $s-1$ and Q_{s+1}, prime at $s+1$	1	1
P_{t-2}, prime at $t-2$ $(= s^2-2)$	$Q_{s-\sqrt{2}}$, prime at $s-\sqrt{2}$ and $Q_{s+\sqrt{2}}$, prime at $s+\sqrt{2}$	1	1
P_{t-0}, prime at t $(= s^2)$	Q_{s-0}, the prime at s	2	1
P_{t+1}, prime at $t+1$ $(= s^2+1)$	Q_{s^2+1}, the prime at s^2+1	1	2

  - Since degree-1 primes of K correspond to points of $C_2 = \mathbb{P}^1(\mathbb{R})$, the primes lying above a given point $P \in C_2$ in L are simply the points in the preimage $\varphi^{-1}(P)$. Thus, for example, the primes of L lying above the point 1 (i.e., the prime P_{t-1} associated to $t-1$) are simply the primes associated to $\varphi^{-1}(1) = \{1, -1\}$.
  - Over an algebraically closed field, the above calculation yields all of the primes of L. However, over a non-algebraically-closed field, we must instead glue together Galois conjugates of preimages, which yield primes of degree grater than 1.
  - Here, for example, over the complex numbers the prime P_{-1} would have two primes Q_i and Q_{-i} lying over it, but over $\mathbb{R}$ these two primes become glued together and yield instead the degree-2 prime Q_{s^2+1}.

- Geometrically, we can view the map on the curves using the following diagram:



- The diagram above illustrates the various primes identified above using the morphism $\varphi : C_1 \to C_2$ (we view C_1 as the y-axis and C_2 as the x-axis, with the morphism between them following the graph of the parabola). The prime above P_{t+1} is "missing" from the diagram since it corresponds to a degree-2 prime, which is not a point on the curve.

- Exercise: For the cubic extension $\mathbb{R}(t)/\mathbb{R}(t^3)$, determine all primes lying over P_{t^3-0}, P_{t^3-1}, and P_{t^3+8} and determine their relative degrees and ramification indices.

- The relative degree and ramification index have a nice relationship with the overall extension degree $[L : K]$:

- Theorem (Relative Degree and Ramification Index): Let L/K be a geometric extension of function fields with $[L : K] = n$. Then the following hold:

  1. If $\tilde{P}$ lies over P, then $e(\tilde{P}|P) \cdot f(\tilde{P}|P) \leq n$.
     - Proof: Choose a generator π for $\tilde{P}$ and let $\{c_1, \dots, c_f\}$ be a linearly independent set in $\mathcal{O}_{\tilde{P}}/\tilde{P}$. We will show that the elements of the form $\pi^{i-1}c_j \in L$ for $1 \leq i \leq e(\tilde{P}|P)$ and $1 \leq j \leq f(\tilde{P}|P)$ are linearly independent over K, which immediately yields $e(\tilde{P}|P) \cdot f(\tilde{P}|P) \leq [L : K] = n$.
     - So suppose we had a linear dependence $\sum_{i,j} a_{i,j}\pi^{i-1}c_j = 0$ for $a_{i,j} \in K$. By clearing denominators and rescaling by a power of a uniformizer, without loss of generality we may assume that all of the coefficients $a_{i,j} \in \mathcal{O}_P$ and that some $a_{i,j} \notin P$.
     - Group the sum as $[a_{1,1}c_1 + \cdots + a_{1,f}c_f] + \pi[a_{2,1}c_1 + \cdots + a_{2,f}c_f] + \cdots + \pi^{e-1}[a_{e,1}c_1 + \cdots + a_{e,f}c_f] = 0$ and let $x_j = a_{j,1}c_1 + \cdots + a_{j,f}c_f$.
     - For each x_j, if all coefficients $a_{i,j}$ are in P then π^e divides x_j (by the definition of the ramification index), and otherwise, x_j is a unit in $\mathcal{O}_{\tilde{P}}$ since it is not zero modulo $\tilde{P}$.
     - Then since some $a_{i,j} \notin P$, at least one x_j is not divisible by π^e, and we can then calculate $v_{\tilde{P}}[x_1 + \pi x_2 + \cdots + \pi^{e-1}x_f] = \min_j[x_j \neq 0 \pmod{\tilde{P}}] < e$, which is a contradiction because the sum $x_1 + \pi x_2 + \cdots + \pi^{e-1}x_f$ is zero.
  2. If L/K is separable, let R be the integral closure of $\mathcal{O}_P$ in L. Then R is a Dedekind domain and the ideal factorization of PR is $PR = (\tilde{P}_1)^{e_1} \cdots (\tilde{P}_g)^{e_k}$ where $\tilde{P}_1, \dots, \tilde{P}_g$ are the primes of L lying over P with ramification indices $e_i = e_i(\tilde{P}_i|P)$.
     - This is a special case of a general result known as Dedekind's factorization theorem (and the same factorization procedure works to enumerate all of the ideals lying a particular prime P of K when K is a number field).
     - More generally, if A is a Dedekind domain with fraction field K, $B = A[\alpha]$, f is the minimal polynomial of α over K, and P is a prime ideal of A, then if $f(x) \equiv \prod g_i(x)^{e_i} \bmod P$ for distinct monic irreducibles $g_i(x)$, then the prime ideal factorization of PB inside B is $PB = \prod(P, g_i(\alpha))^{e_i}$.
     - We omit the full proof.
  3. (The efg Theorem) If L/K is any extension of function fields $\tilde{P}_1, \dots, \tilde{P}_g$ are the primes of L lying over P, then $\sum_{i=1}^{g} e(\tilde{P}_i|P) \cdot f(\tilde{P}_i|P) = n$.

    - The result holds in general, however we will only give the proof of the separable case (the inseparable case requires a bit more development).
    - Proof (Separable Case): Let R be the integral closure of $\mathcal{O}_P$ in L. By (2), we have $PR = (\tilde{P}_1)^{e_1} \cdots (\tilde{P}_g)^{e_k}$ where $\tilde{P}_1, \ldots, \tilde{P}_g$ are the primes of L lying over P with ramification indices $e_i = e_i(\tilde{P}_i|P)$.
    - It is also not hard to verify that R is a free $\mathcal{O}_P$-module of rank n. (Exercise: check this.)
    - Then applying the Chinese remainder theorem yields $R/PR \cong (R/\tilde{P}_1^{e_1}) \oplus \cdots \oplus (R/\tilde{P}_g^{e_g})$, so since $R/\tilde{P}^e \cong \mathcal{O}_{\tilde{P}}/\tilde{P}^e$ we need only compute the dimensions of each of the terms.
    - Exercise: If $\mathcal{O}$ is a discrete valuation ring with maximal ideal m, show that $\dim_{\mathcal{O}/m}(m^n/m^{n+1}) = 1$ and deduce that $\dim_{\mathcal{O}/m}(\mathcal{O}/m^n) = n$.
    - From the exercise above, we see that $\dim_{\mathcal{O}_{\tilde{P}}/\tilde{P}}[R/\tilde{P}^e] = e$. Since by definition we also have $\dim_{\mathcal{O}_P/P}(\mathcal{O}_{\tilde{P}}/\tilde{P}) = f$, by the usual properties of vector space dimensions and field degrees we deduce that $\dim_{\mathcal{O}_P/P}[R/\tilde{P}^e] = ef$.
    - Finally, taking the dimension of R/PR as an $\mathcal{O}_P/P$-vector space yields $n = \dim_{\mathcal{O}_P/P}[R/PR] = \sum_{i=1}^{g} \dim_{\mathcal{O}_P/P}[R/\tilde{P}_i^{e_i}] = \sum_{i=1}^{g} e(\tilde{P}_i|P) \cdot f(\tilde{P}_i|P)$, as claimed.

4. If L/K is Galois, then the Galois group $\mathrm{Gal}(L/K)$ acts transitively on the primes $\tilde{P}_i$ lying above P. These primes $\tilde{P}_i$ for $1 \le i \le g$ all have the same ramification index e and relative degree f, and $efg = n$.
    - Proof: To see that the Galois group acts transitively, suppose otherwise, so that $\sigma(\tilde{P}_i) \neq \tilde{P}_j$ for any $\sigma \in \mathrm{Gal}(L/K)$. Then by the Chinese remainder theorem, there exists $t \in \tilde{P}_i$ such that $\sigma(t) \notin \tilde{P}_j$ for any $\sigma \in \mathrm{Gal}(L/K)$.
    - But then $\prod_{\sigma \in \mathrm{Gal}(L/K)} \sigma(t)$ is fixed by the entire Galois group, so it is an element of K and also in $\tilde{P}_i$ hence is an element of P hence also of $\tilde{P}_j$. But $\tilde{P}_j$ is a prime ideal containing a product of elements $\prod_{\sigma \in \mathrm{Gal}(L/K)} \sigma(t)$ not contained it, which is a contradiction.
    - Therefore, the Galois group acts transitively on the $\tilde{P}_i$. For the second part we simply observe that conjugate primes have the same e and f, and so since all of the primes are conjugate, they all have the same e and f; then (3) immediately yields $efg = n$.

- From the theorem above, we see that the ramification index e, relative degree f, and number of primes g are always between 1 and the extension degree $n = [L : K]$. The "maximal cases" have special names:
- Definition: Suppose L/K is an extension of function fields with $[L : K] = n$ and P is a prime of K with $\tilde{P}_1, \ldots, \tilde{P}_g$ lying above it.
    1. We say P is totally ramified if $e = n$ and $f = 1$ (so that $g = 1$). In this case, there is also a unique prime $\tilde{P}$ lying above P and $P\mathcal{O}_{\tilde{P}} = (\tilde{P})^e$.
    2. More generally, we say P is ramified if there is some $\tilde{P}$ lying above P such that $e(\tilde{P}|P) > 1$; otherwise, if $e(\tilde{P}|P) = 1$ for all $\tilde{P}$ lying above P, and L/K is separable, we say P is unramified.
    3. We say P splits completely if all e_i and f_i are equal to 1 (so that $g = n$). In this case, there is the maximal possible number of primes $\tilde{P}_1, \ldots, \tilde{P}_n$ that lie above P.
    4. We say P is totally inert if $e = 1$ and $f = n$ (so that $g = 1$). In this case, there is a unique prime $\tilde{P}$ lying above P and $P\mathcal{O}_{\tilde{P}} = \tilde{P}$. (The idea is that P "stays prime" when we extend to L.)

    - Example: For the quadratic extension of rational function fields given by $\mathbb{R}(t)/\mathbb{R}(t^2)$ with constant field $\mathbb{R}$ from earlier, the primes P_{t-1} and P_{t-2} split completely, the prime P_{t-0} is totally ramified, and the prime P_{t+1} is inert.
    - These terms are also used for number fields, where we can give some additional examples to illustrate the reasons for focusing on these cases.
    - Example: For the quadratic extension $\mathbb{Q}(i)/\mathbb{Q}$ with ring of integers $\mathbb{Z}[i]$, by Dedekind's factorization theorem, the primes above (5) are $(2 + i)$ and $(2 - i)$ so the prime (5) splits completely in $\mathbb{Q}(i)$. There is a unique prime above (3), namely (3) itself; the prime (3) is totally inert in $\mathbb{Q}(i)$. There is a unique prime above (2), namely $(1 + i)$; since $(2) = (1 + i)^2$, the prime (2) is totally ramified.

- Example: For the quadratic extension $\mathbb{Q}(\sqrt{7})/\mathbb{Q}$ with ring of integers $\mathbb{Z}[\sqrt{7}]$, the primes above (3) are $(3, 1+\sqrt{7})$ and $(3, 1-\sqrt{7})$; the prime (3) splits completely in $\mathbb{Q}(\sqrt{7})$. There is a unique prime above (5), namely (5); the prime (5) is totally inert in $\mathbb{Q}(\sqrt{7})$. There is a unique prime above (7), namely $(\sqrt{7})$; since $(7) = (\sqrt{7})^2$, the prime (7) is totally ramified.
- Exercise: In the quadratic extension $\mathbb{Q}(\sqrt{11})/\mathbb{Q}$ with ring of integers $\mathbb{Z}[\sqrt{11}]$, use Dedekind's factorization theorem to determine whether the primes (2), (3), (5), (7), and (11) are split, inert, or ramified.
- Example: Consider the cubic extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ where $\alpha^3 - \alpha - 1 = 0$; one may show that the ring of integers of $\mathbb{Q}(\alpha)$ is $\mathbb{Z}[\alpha]$. By Dedekind's factorization theorem, since $\alpha^3 - \alpha - 1$ is irreducible modulo 2, the factorization of (2) is simply (2), so (2) is totally inert. Since $\alpha^3-\alpha-1 = (\alpha-2)(\alpha^2+2\alpha+3)$ modulo 5, we see that the factorization of (5) is $(5) = (5, \alpha-2)(5, \alpha^2+2\alpha+3)$, so (5) is split but not totally split. Since $\alpha^3 - \alpha - 1 = (\alpha - 10)^2(\alpha - 3)$ modulo 23, we see that the factorization of (23) is $(23) = (23, \alpha - 10)^2(23, \alpha-3)$, so (23) is ramified but not totally ramified. Also, since $\alpha^3-\alpha-1 = (\alpha-4)(\alpha-13)(\alpha-42)$ modulo 59, we see that the factorization of (59) is $(59) = (59, \alpha - 4)(59, \alpha - 13)(59, \alpha - 42)$ so (59) is totally split.

- One key result from algebraic number theory is that the number of ramified primes in an extension $K/\mathbb{Q}$ is always finite, and in fact, every ramified prime divides the discriminant of the extension d_K.
  - This result can be extended to relative extensions by defining the discriminant ideal $d_{L/K} = \det[\mathrm{tr}(x_i x_j)]$ where $\{x_1, \dots, x_d\}$ is a maximal $\mathcal{O}_K$-linearly independent subset of $\mathcal{O}_L$.
  - One may show that $d_{L/K}$ is well defined and that it is an ideal of the ring of integers $\mathcal{O}_K$: the main result is then that a prime ideal P of K is ramified if and only if P divides the discriminant $d_{L/K}$.
- Our next goal is to give a similar construction in the function field case.

0.21 (Nov 29) Ramification and the Riemann-Hurwitz Theorem

- We now define the different and discriminant for extensions of function fields. Some preliminary terminology:
- Definition: Let L/K be an extension of function fields. If P is a prime of K, we define the conorm of P to be the divisor $\mathrm{Con}_{L/K}(P) = \sum_{\tilde{P}|P} e(\tilde{P}|P)\tilde{P}$ of L, and then extend the definition to all divisors D of K by linearity. Also, if $\tilde{P}$ is a prime of L, we define the norm of P to be the divisor $N_{L/K}(\tilde{P}) = f(P'|P)P$ of K, and then extend the definition to all divisors $\tilde{D}$ of L by linearity.
  - From the efg-theorem, we see immediately that $N_{L/K} \circ \mathrm{Con}_{L/K}$ is multiplication by $n = [L : K]$.
  - Exercise: Show that the norm and conorm of divisors compose in towers.
  - Exercise: Show that for any $a \in K^\times$, it is true that $\mathrm{Con}_{L/K}[\mathrm{div}_K a] = \mathrm{div}_L(a)$.
  - With some effort, one may show that the norm and conorm map principal divisors to principal divisors: indeed, for any $a \in K^\times$ one has $\mathrm{Con}_{L/K}[\mathrm{div}_K a] = \mathrm{div}_L(a)$, and for any $\tilde{a} \in L^\times$ one also has $N_{L/K}[\mathrm{div}_L \tilde{a}] = \mathrm{div}_K[N_{L/K}\tilde{a}]$, where the norm on the right-hand side is the field norm[6] of the element $\tilde{a}$.
  - Exercise: When L/K is separable, show that for any $\tilde{a} \in L$, it is true that $N_{L/K}[\mathrm{div}_L \tilde{a}] = \mathrm{div}_K[N_{L/K}\tilde{a}]$.
- The norm and conorm behave well with respect to degrees:
- Proposition (Degrees of Norm and Conorm): If L/E is a function field extension of K/F (i.e., L has constant field E and K has constant field F), then for any divisor D of K and $\tilde{D}$ of L we have $\deg_K N_{L/K}(\tilde{D}) = [E : F] \deg_L(\tilde{D})$ and $\deg_L \mathrm{Con}_{L/K}(D) = \dfrac{[L : K]}{[E : F]} \deg_K(D)$.
  - Proof: By linearity it suffices to check the results for prime divisors:

[6] If L/K is any finite-degree extension and $a \in L$, the trace $\mathrm{tr}_{L/K}(a)$ and norm $N_{L/K}(a)$ are defined to be the trace and determinant (respectively) of the "multiplication-by-a" map on L considered as a vector space over K. When L/K is Galois with Galois group G, the trace is the sum $\sum_{\sigma \in G} \sigma(a)$ of all of the Galois conjugates of a, while the norm is the product $\prod_{\sigma \in G} \sigma(a)$ of all of the Galois conjugates of a. In the non-Galois case, one can instead take the σ_i to be the $[L : K]$ possible embeddings of L into a fixed algebraic closure of K.

$\deg_K N_{L/K}(\tilde{P}) = f(\tilde{P}|P)\deg_K(P) = [\mathcal{O}_{\tilde{P}}/\tilde{P} : \mathcal{O}_P/P]\cdot[\mathcal{O}_P/P : F] = [\mathcal{O}_{\tilde{P}}/\tilde{P} : E]\cdot[E : F] = \deg_L(\tilde{P})\cdot[E : F]$

$\deg_L \mathrm{Con}_{L/K}(P) = \deg_L[\sum_{\tilde{P}|P} e(\tilde{P}|P)\tilde{P}] = \sum_{\tilde{P}|P} e(\tilde{P}|P)\deg_L(\tilde{P}) = \sum_{\tilde{P}|P} e(\tilde{P}|P)\dfrac{f(\tilde{P}|P)\deg_K(P)}{[E:F]} = \dfrac{[L:K]}{[E:F]}\deg_K(P).$

- Definition: Let L/K be a separable extension of function fields, $A = \mathcal{O}_P$ be any discrete valuation ring of K, and B be the integral closure of K in L, which one may show is $B = \bigcap_{\tilde{P}|P}\mathcal{O}_{\tilde{P}}$.
  1. The discriminant ideal $\partial_{B/A}$ is the ideal of A given by $\det[\mathrm{tr}(x_i x_j)]$, where $\{x_1, \dots, x_d\}$ is any free basis for B as an A-module.
  2. The codifferent $C_{B/A}$ is the set of elements $C_{B/A} = \{x \in L : \mathrm{tr}_{L/K}(xb) \in A \text{ for all } b \in B\}$; it is the maximal B-submodule of L consisting of those elements whose trace lands in A.
  3. The codifferent can be shown to be an invertible fractional ideal of B. Its inverse $D_{B/A}$ is called the different of B/A.
- In our situation, we can make the calculations more explicit:
- Proposition (Different and Discriminant): Let L/K be a separable extension of function fields and $\tilde{P}$ be a prime of L lying over P in K, and let $A = \mathcal{O}_P$ and $B = \bigcap_{\tilde{P}|P}\mathcal{O}_{\tilde{P}}$ be the integral closure of $\mathcal{O}_P$ in L.
  1. For any basis $\{z_1, \dots, z_n\}$ of L/K, there exists a dual basis $\{z_1^*, \dots, z_n^*\}$ such that $\mathrm{tr}_{L/K}(z_i z_j^*) = \delta_{i,j}$.
     - Proof: The dual space of L/K has the same dimension as L/K, so since there is a natural pairing between them (namely, the evaluation map), we see that $\hat{L}$ is a 1-dimensional vector space over L. Then since the trace map is nonzero, all of the maps of the form $\lambda_i(z_j) = \delta_{i,j}$ are multiples of the trace map. By selecting appropriate multiples we obtain a dual basis.
  2. If $\{z_1, \dots, z_n\}$ is an integral basis of $\mathcal{O}_{\tilde{P}}$ over $\mathcal{O}_P$, then $C_{B/A} = \sum_{i=1}^n \mathcal{O}_P \cdot z_i^*$ where $\{z_1^*, \dots, z_n^*\}$ is the dual basis of $\{z_1, \dots, z_n\}$.
     - Proof: Suppose $z \in C_{B/A}$. Then $z = \sum_{i=1}^n x_i z_i^*$ for some $x_i \in K$ since $\{z_1^*, \dots, z_n^*\}$ is a basis of L/K.
     - Since $z \in C_{B/A}$ we see that $\mathrm{tr}_{L/K}(zz_j) \in \mathcal{O}_P$ for all j. But $\mathrm{tr}_{L/K}(zz_j) = x_j$ by the definition of the dual basis, and so all of the $x_j \in \mathcal{O}_P$.
     - On the other hand, by the same calculation, any $z = \sum_{i=1}^n x_i z_i^*$ with all $x_i \in \mathcal{O}_P$ is an element of $C_{B/A}$, since $\mathrm{tr}_{L/K}(zz_j) \in \mathcal{O}_P$ so since $\{z_1, \dots, z_n\}$ is an integral basis of $\mathcal{O}_{\tilde{P}}$ this means $\mathrm{tr}_{L/K}(z\tilde{z}) \in \mathcal{O}_P$ for any $\tilde{z} \in \mathcal{O}_{\tilde{P}}$.
     - Exercise: If $\mathcal{O}_{\tilde{P}} = \mathcal{O}_P[\alpha]$ has integral basis $\{1, \alpha, \dots, \alpha^{n-1}\}$ where α has minimal polynomial $f(t) = (t-\alpha)(c_{n-1}t^{n-1} + \cdots + c_0)$ over K, show that the dual basis is $\{\dfrac{c_0}{f'(\alpha)}, \dfrac{c_1}{f'(\alpha)}, \dots, \dfrac{c_{n-1}}{f'(\alpha)}\}$. [Hint: Use the identity $\sum_{i=1}^n \dfrac{r_i^k}{f'(r_i)}\dfrac{f(t)}{t-r_i} = t^k$ where the r_i are the roots of f.]
  3. There exists $t \in L$ such that $C_{B/A} = t\cdot B$, and for this t it is true that $v_{\tilde{P}}(t) \le 0$ for all $\tilde{P}|P$. Furthermore, $C_{B/A} = t'\cdot B$ for another $t' \in L$ if and only if $v_{\tilde{P}}(t) = v_{\tilde{P}}(t')$ for all $\tilde{P}|P$.
     - Proof: Take $\{z_1, \dots, z_n\}$ to be an integral basis of $\mathcal{O}_{\tilde{P}}$ over $\mathcal{O}_P$ and pick an element $x \in K$ such that $v_P(x) \ge -v_{\tilde{P}}(z_i^*)$ for all $\tilde{P}|P$ and all i.
     - Then $v_{\tilde{P}}(xz_i^*) = e(\tilde{P}|P)v_P(x) + v_{\tilde{P}}(z_i^*) \ge 0$ and so $x \cdot C_{B/A} \subseteq B$ since B is the set of elements $u \in L$ with $v_{\tilde{P}}(u) \ge 0$ for all $\tilde{P}|P$.
     - It is easy to see that $x \cdot C_{B/A}$ is in fact an ideal of B, so $x \cdot C_{B/A} = yB$ for some $y \in B$. If we take $t = y/x$ then $C_{B/A} = t \cdot B$ and $v_{\tilde{P}}(t) \le 0$ for all $\tilde{P}|P$ as claimed.
     - The second part follows by chasing valuations (we omit the details).
     - Exercise: Show that if $\mathcal{O}_{\tilde{P}} = \mathcal{O}_P[\alpha]$ has integral basis $\{1, \alpha, \dots, \alpha^{n-1}\}$ where α has minimal polynomial $f(t)$ over K, then we can take $t = f'(\alpha)$ here. [This connection between the different and the derivative $f'(\alpha)$ is why it has the name "different".]
  4. For all but finitely many primes P of K, it is true that $C_{B/A} = B$.

- Proof: Take $\{z_1, \dots, z_n\}$ to be any basis of L/K, with corresponding dual basis $\{z_1^*, \dots, z_k^*\}$. Then the minimal polynomials for $\{z_1, \dots, z_n\}$ and $\{z_1^*, \dots, z_k^*\}$ have only finitely many coefficients in total and hence these coefficients have finitely many poles in total.
- For any prime P not among these finitely many poles, all of the z_i and z_i^* are in $B = \bigcap_{\tilde{P}|P} \mathcal{O}_{\tilde{P}}$, and so we have $\sum_{i=1}^n \mathcal{O}_P \cdot z_i \subseteq \bigcap_{\tilde{P}|P} \mathcal{O}_{\tilde{P}} \subseteq \sum_{i=1}^n \mathcal{O}_{\tilde{P}} \cdot z_i^* \subseteq \bigcap_{\tilde{P}|P} \mathcal{O}_{\tilde{P}} = \sum \mathcal{O}_P \cdot z_i$, so equality must hold everywhere. Thus we see that $B = \bigcap_{\tilde{P}|P} \mathcal{O}_{\tilde{P}} = \sum_{i=1}^n \mathcal{O}_P \cdot z_i$ meaning that $\{z_1, \dots, z_n\}$ is an integral basis of B. Then by (2) this means $B = C_{B/A}$.

5. The norm of the different is the discriminant: explicitly, $N_{L/K} D_{B/A} = \partial_{B/A}$.
   - This calculation is fairly lengthy so we omit it, but the main idea is just to write everything out in terms of the basis given in (2) for the codifferent.
6. With notation as in (3), if we define the different exponent $d(\tilde{P}|P)$ to be $d(\tilde{P}|P) = -v_{\tilde{P}}(t)$ where $C_{B/A} = t \cdot B$, then we have $d(\tilde{P}|P) \geq e(\tilde{P}|P) - 1$ with equality if and only if $e(\tilde{P}|P)$ is not divisible by the characteristic of K (in this situation we say $\tilde{P}$ is tamely ramified).
   - This is the function-field analogue of a classical theorem of Dedekind that was generalized by Hilbert.
   - Here is a brief sketch of the argument: first, by taking completions, we may convert the problem to one in positive characteristic. Second, we may replace K with its maximal unramified extension, in which case $\tilde{P}|P$ is now totally ramified.
   - In this case, if we choose a uniformizer π for $\tilde{P}$, we have $B = A[\pi]$ so an integral basis is given by $\{1, \pi, \dots, \pi^{e-1}\}$.
   - Furthermore, the minimal polynomial of π over K has degree $e = e(\tilde{P}|P)$ and will be Eisenstein-irreducible at P hence of the form $f(t) = t^e + a_{n-1}t^{e-1} + \cdots + a_0$ where the coefficients $a_i \in P$, and the different $D_{B/A}$ is generated by the derivative $f'(\pi)$, as follows by the exercises given above with (3) and (4).
   - But $f'(\pi) = e\pi^{e-1} + \cdots + a_1$ and since the first term is divisible by π^{e-1} while the others are divisible by π^e (since all of the $a_i \in P$) we see $v_\pi(D_{B/A}) = v_\pi(f'(\pi)) \geq e - 1$, with equality if and only if e is not divisible by the characteristic of K.
   - Since $v_\pi(D_{B/A}) = d(\tilde{P}|P)$ we obtain the claimed result.
7. For any prime P of K, there is a prime $\tilde{P}|P$ that ramifies if and only if $\partial_{B/A} \subseteq P$, which is (essentially) to say, precisely when P divides the discriminant $\partial_{B/A}$. Therefore, there are only finitely many ramified primes of K.
   - Proof: By (6), if P is ramified, meaning that $e(\tilde{P}|P) > 1$, then $d(\tilde{P}|P) > 1$ also, but then (5) forces $\partial_{B/A} \subseteq P$.
   - The part about having only finitely many ramified primes follows immediately from (4).

- Since there are only finitely many primes $\tilde{P}$ for which $d(\tilde{P}|P) > 0$, we may construct a divisor associated to the different:
- Definition: Let L/K be a separable extension of function fields. We define the different divisor to be the sum $\text{Diff}(L/K) = \sum_P \sum_{\tilde{P}|P} d(\tilde{P}|P)\tilde{P}$.
- Our main result now is the Riemann-Hurwitz theorem:
- Theorem (Riemann-Hurwitz): Suppose L/E is a separable function-field extension of K/F. Then $2g_L - 2 = \dfrac{[L:K]}{[E:F]}(2g_K - 2) + \deg[\text{Diff}(L/K)]$.
  - The Riemann-Hurwitz theorem is really a topological result, and we can give some geometric motivation for where it comes from in the situation of Riemann surfaces where $E = F = \mathbb{C}$ (in which case the extension L/K is geometric).
  - Here, the function field extension L/K corresponds to a morphism $\varphi : C_L \to C_K$ of smooth projective complex curves of degree $d = [L : K]$, which we view as surfaces over $\mathbb{R}$.
  - Then the morphism φ (which is surjective, as we mentioned previously) represents a d-sheeted covering of C_K by C_L, where each unramified point of C_K has exactly d preimages in C_L.

- If φ were unramified everywhere, then (e.g., by considering a triangulation of C_K) we see that the Euler characteristic $2-2g_L$ of C_L would be d times the Euler characteristic $2-2g_K$ of K: this is precisely the statement of Riemann-Hurwitz above.
- At ramified points of φ, the sheets of the covering collide, which introduces an error term into the calculation: this is the correction that is accounted for by $\deg[\mathrm{Diff}(L/K)]$.
- More precisely, for a tamely ramified point $\tilde{P}|P$ of ramification index e, a total of e sheets in the covering collide at $\tilde{P}$, which results in a net error of $e-1$ in the calculation of the Euler characteristic. But $e-1$ is precisely the value $d(\tilde{P}|P)$ that appears as the coefficient of P in the different $\mathrm{Diff}(L/K)$, so the different correctly accounts for the error. (For wildly ramified points the ideas are similar but more involved.)
- The precise proof for function fields requires the following two facts about Weil differentials, which we have not actually defined. (However, the corresponding statements for Riemann surfaces are fairly natural.)
  1. For any differential ω of K/F, there exists a unique differential $\tilde{\omega}$ of L/E such that $\mathrm{tr}_{L/K}(\tilde{\omega}(\alpha)) = \omega(\mathrm{tr}_{E/F}\alpha)$ for all α in the adeles of K/F. Furthermore, $\mathrm{div}_L(\tilde{\omega}) = \mathrm{Con}_{L/K}(\mathrm{div}_K\omega) + \mathrm{Diff}(L/K)$.
  2. For any differential ω of K, the degree $\deg(\mathrm{div}\,\omega) = 2g-2$ where g is the genus of K. (Note that this is something we have previously encountered in the Riemann-Roch theorem, since the divisor of any differential lies in the canonical class C.)
- Proof: Using the first fact above, we have $\mathrm{div}_L(\tilde{\omega}) = \mathrm{Con}_{L/K}(\mathrm{div}_K\omega) + \mathrm{Diff}(L/K)$.
- Now take degrees of both sides: this yields $\deg(\mathrm{div}_L\tilde{\omega}) = \deg[\mathrm{Con}_{L/K}(\mathrm{div}_K\omega)] + \deg \mathrm{Diff}(L/K)$.
- Applying the second fact and the properties of the conorm then gives the claimed result: $2g_L - 2 = \dfrac{[L:K]}{[E:F]}(2g_K-2) + \deg[\mathrm{Diff}(L/K)]$.

- The Riemann-Hurwitz theorem has many useful consequences, which we will now investigate.
  - For example, if L/K is separable, we immediately see that $2g_L - 2 \geq 2g_K - 2$ and so $g_L \geq g_K$. (We will remark that this inequality is *not* necessarily true when L/K is inseparable.)
- Another less trivial consequence is the following theorem of Lüroth:
- Theorem (Lüroth): If $L = F(x)$ and K is a subfield of L properly containing F, then there exists $u \in K$ such that $K = F(u)$.
  - In other words, every transcendental subextension of $F(x)$ is also purely transcendental over F.
  - Proof: Since K properly contains F, K has transcendence degree 1 over F, so $[L:K]$ is finite.
  - Let M be the maximal separable extension of K inside L. If $\mathrm{char}(L) = 0$ then $K = M$, and otherwise L/M is a purely inseparable extension of finite degree.
  - From the general theory of inseparable extensions in characteristic p, we must have $[L:M] = p^n$ and then $x^{p^n} \in M$. But since $\mathrm{div}_-(x^{p^n}) = p^n P_\infty$ we see that $[L:F(x^{p^n})] = p^n$ and thus $F(x^{p^n}) = M$.
  - Therefore, we can reduce to the case where L/K is separable. Then since $g_L = 0$, by Riemann-Hurwitz we must also have $g_K = 0$.
  - Since L has a prime of degree 1 (namely, $\mathrm{div}_+(x)$), the prime below it in K must also have degree 1. Thus, from our analysis earlier using Riemann-Roch of genus-0 extensions with a prime of degree 1, we deduce $K = F(u)$ for some u, as claimed.

0.22 (Dec 2) The abc Conjecture and the S-Unit Equation

- In the first lecture, we gave the proof of Fermat's last theorem for polynomials. We now embark on a proof of the abc conjecture for function fields, which makes precise the general idea that if $a+b=c$, then at least one of a,b,c must have a "large" prime divisor.
  - Recall that the Mason-Stothers theorem says if $a,b,c \in F[t]$ are nonconstant and relatively prime with $a+b=c$, and not all of a',b',c' are zero, then $\max(\deg a, \deg b, \deg c) \leq \deg \mathrm{rad}(abc) - 1$.

- Thus, there cannot be too many high powers appearing in the factorization of abc, since the degree of $\text{rad}(abc)$ is at least $1 + \max(\deg a, \deg b, \deg c)$.

- Here is a precise formulation of this phenomenon for integers:
- Conjecture (abc Conjecture): For any $\epsilon > 0$, there exists a constant $C_\epsilon > 0$ such that for any relatively prime nonzero integers a, b, c with $a + b = c$, it is true that $\max(|a|, |b|, |c|) \le C_\epsilon \cdot \text{rad}(abc)^{1+\epsilon}$.
  - Another way of phrasing this result, if we rearrange to make a, b, c positive and define the "quality" of a triple (a, b, c) to be $q = (\log c)/(\log \text{rad}\, abc)$, then there are only finitely many triples of quality exceeding $1 + \epsilon$ for any $\epsilon > 0$.
  - Equivalently, the abc conjecture says that the limit supremum of the quality over all triples (a, b, c) is equal to 1.
  - For illustration, here are some triples of high quality:
    * $2 + 3^{10} \cdot 109 = 23^5$ (quality 1.6299)
    * $11^2 + 3^2 5^6 7^3 = 2^{21} \cdot 23$ (quality 1.6260)
    * $19 \cdot 1307 + 7 \cdot 29^2 \cdot 31^8 = 2^8 \cdot 3^{22} \cdot 5^4$ (quality 1.6235)
    * $1 + 2 \cdot 3^7 = 5^4 \cdot 7$ (quality 1.5808)
  - We will note that the $1 + \epsilon$ is necessary (i.e., the result cannot be sharpened to say that there are only finitely many triples of quality greater than 1).
  - Specifically, if we take $a = 3^{2^n} - 1$, $b = 1$, and $c = 3^{2^n}$, then $\max(a, b, c) = c = 3^{2^n}$ and $\text{rad}(abc) \le 3 \cdot \dfrac{3^{2^n} - 1}{2^n} \cdot 2$ since $2^n | 3^{2^n} - 1$. Therefore, for $n \ge 2$ we see that the quality is at least $\log(3^{2^n})/\log(3^{2^n} \cdot 3/2^n) > 1$.
  - Mochizuki announced a proof of the abc conjecture in 2012. As of 2021, the correctness of the claimed proof is still disputed, and it is not clear whether the conjecture has actually been resolved or not. Some well-known critics include Peter Scholze (recent Fields medalist) and Jacob Stix.

- Let us now describe and prove the abc conjecture for function fields, which will necessarily involve some reformulation of the central idea.
  - Recall that for a rational number $\alpha = p/q$ in lowest terms, its height is $\text{ht}(\alpha) = \max(\log |p|, |q|)$.
  - We can reformulate the abc conjecture in terms of heights by rewriting the statement homogeneously and taking logarithms:
  - Conjecture (Reformulated abc Conjecture): For rational numbers u $(= a/c)$ and v $(= b/c)$ with $u+v = 1$, we have $\max(\text{ht}\, u, \text{ht}\, v) \le C_\epsilon + (1 + \epsilon) \sum_{p|abc} \log p$.
  - The analogue of height for rational function fields for $u = a/c \in F(t)$ would be $\max(\deg a, \deg c)$, which as we have seen is equal to the extension degree $[K : F(u)]$.
  - This formulation generalizes nicely to function fields:

- Definition: If K/F is a function field, the degree (or height) of an element $u \in K \backslash F$ is the extension degree $\deg(u) = [K : F(u)]$.
  - In the context of projective curves, we can think of u as being a rational map from the curve C whose function field is K to the projective line $\mathbb{P}^1(F)$.
  - Then the degree $\deg(u)$ represents the degree of this rational map (as a function), and unramified points in $\mathbb{P}^1$ have exactly $\deg(u)$ preimages on C.
  - The analogue of the radical for function fields is the support: the set of primes where a divisor (or element) has positive valuation.

- With all of this in hand, we can give the analogue of the abc conjecture for function fields:
- Theorem (Function-Field abc Conjecture): Let K/F be a function field where F is perfect. For any $u, v \in K^\times$ with $u+v = 1$, we have $\deg_S u = \deg_S v \le 2g_K - 2 + \sum_{P \in \text{Supp}(A+B+C)} \deg_K P$, where $A = \text{div}_+ u$, $B = \text{div}_+ v$, $C = \text{div}_- u = \text{div}_- v$, and where $\deg_S x$ represents the degree of the maximal separable extension $[M : F(x)]$.

- The result holds in general, but we will just do the case where K/F is separable (the inseparable case follows directly from this one when applied to the maximal separable extension M of F inside K).
- The main idea now is to ignore primes $\tilde{P}$ not lying over P_0, P_1, or P_∞, which are the positive parts of the divisors of u, $1-u=v$, and $1/u$ respectively.
- Proof (Separable Case): Let $n = \deg u = [K : F(u)]$. By Riemann-Hurwitz and the fact that $d(\tilde{P}|P) \geq e(\tilde{P}|P)-1$, we have $2g_K - 2 \geq -2n + \sum_{\tilde{P}|P}[e(\tilde{P}|P)-1]\deg_K P \geq -2n + \sum_{\tilde{P}|P, P\in\text{Supp}(A+B+C)}[e(\tilde{P}|P)-1]\deg_K P$.
- We then have $A = \text{Con}_{K/F(u)}P_0$, $B = \text{Con}_{K/F(u)}P_1$, and $C = \text{Con}_{K/F(u)}P_\infty$, and thus

$$\begin{aligned}\sum_{P\in\text{Supp}(A)}[e(\tilde{P}|P_0)-1]\deg_K P &= \deg_K[\text{Con}_{K/F(u)}P_0] - \sum_{P\in\text{Supp}(A)}\deg_K P \\ &= [K : F(u)]\cdot \deg_{F(u)} P_0 - \sum_{P\in\text{Supp}(A)}\deg_K P \\ &= n - \sum_{P\in\text{Supp}(A)}\deg_K P.\end{aligned}$$

- Similarly,

$$\begin{aligned}\sum_{P\in\text{Supp}(B)}[e(\tilde{P}|P_1)-1]\deg_K P &= n - \sum_{P\in\text{Supp}(B)}\deg_K P \\ \sum_{P\in\text{Supp}(C)}[e(\tilde{P}|P_\infty)-1]\deg_K P &= n - \sum_{P\in\text{Supp}(C)}\deg_K P.\end{aligned}$$

- Plugging these into the inequality from the beginning then yields

$$2g_K - 2 \geq -2n + \sum_{\tilde{P}|P, P\in\text{Supp}(A+B+C)}[e(\tilde{P}|P)-1]\deg_K P = -2n + 3n - \sum_{P\in\text{Supp}(A+B+C)}\deg_K P$$

  and finally rearranging yields the desired result $n \leq 2g_K - 2 + \sum_{P\in\text{Supp}(A+B+C)}\deg_K P$.

- Using function-field *abc*, we can study Fermat's equation in arbitrary function fields:
- Theorem (Fermat in Positive Characteristic): Let K/F be a function field with F perfect and let n be an integer not divisible by the characteristic of F. Then the equation $x^n + y^n = 1$ has no nonconstant solutions in K if $g_K = 0$ and $n \geq 3$, or if $g_K \geq 1$ and $n > 6g_K - 3$.
  - Proof: Suppose that $u^n + v^n = 1$. Then by function-field *abc*, we see that $\max(\deg_S(u^n), \deg_S(v^n)) \leq 2g_K - 2 + \sum_{P\in\text{Supp}(A+B+C)}\deg_K P$ where $A = \text{div}_+(u)$, $B = \text{div}_+(v)$, and $C = \text{div}_-(u) = \text{div}_-(v)$.
  - If M is the maximal separable extension of $F(u)$ in K, then since $p \nmid n$, $F(u)/F(u^n)$ is separable, so $\deg_S(u^n) = n\deg_S(u)$ and $\deg_S(u^n) = n\deg_S(u)$.
  - From the same argument as in the proof of the *abc* conjecture, we have $\sum_{P\in\text{Supp}(A)}\deg_K P \leq \deg_S u$, $\sum_{P\in\text{Supp}(B)}\deg_K P \leq \deg_S v$, and $\sum_{P\in\text{Supp}(C)}\deg_K P \leq \min[\deg_S u, \deg_S v]$.
  - Thus, we have $n\cdot\sum_{P\in\text{Supp}(A)}\deg_K P \leq n\deg_S(u) = \deg_S(u^n) \leq 2g_K - 2 + \sum_{P\in\text{Supp}(A+B+C)}\deg_K P$ and similarly for $\text{Supp}(B)$ and $\text{Supp}(C)$.
  - Adding these three inequalities yields $n\cdot\Sigma \leq 6g_K - 6 + 3\Sigma$ where $\Sigma = \sum_{P\in\text{Supp}(A+B+C)}\deg_K P$.
  - If $g_K = 0$ and $n \geq 3$, then we get a contradiction from $(n-3)\Sigma \leq -6$ since the left-hand side is nonnegative.
  - If $g_K \geq 1$ and $n > 6g_K - 3$, then we have $(6g_K - 6)\Sigma < (n-3)\Sigma \leq 6g_K - 6$ which is a contradiction since it forces $\Sigma \leq 0$, which would mean u, v are constants.
- Another useful application of the same ideas is to studying S-units.
  - In a number field K, if $S = \{P_1, \ldots, P_k\}$ is a finite set of primes of K, we say $\alpha \in K$ is an S-unit if the principal fractional ideal (α) is a product of powers of primes in S (positive or negative powers are allowed, since we are working with fractional ideals).

- There is a generalization of Dirichlet's unit theorem that says the rank of the multiplicative group of S-units is equal to $r + (\#S)$ where r is the rank of the unit group (which, for completeness, is equal to $r_1 + r_2 - 1$ where r_1 is the number of real embeddings of K and r_2 is the number of conjugate pairs of complex embeddings of K).

- The natural analogue in the function-field case is to consider the support of the divisor of an element:
- Definition: Let K/F be a function field and $S = \{P_1, \dots, P_k\}$ be a finite set of primes. An element $u \in K^\times$ is called an S-unit if the support of $\operatorname{div}(u)$ is contained in S, which is to say, when $\operatorname{div}(u) = \sum_{P\in S} c_P P$.
  - In other words, u is an S-unit if the primes occurring with nonzero coefficients in $\operatorname{div}(u)$ are elements of S.
  - The divisor map gives an obvious homomorphism from the multiplicative group of S-units to the divisor group of K; the kernel of this map is $F^\times$.
  - Furthermore, since the degree of any principal divisor is 0, the image of the S-units in the divisor group D_K is free of rank $(\#S) - 1$.
- Our object of study here is the S-unit equation $u + v = 1$ where u, v are S-units.
  - In characteristic p, if $u + v = 1$ then taking p^nth powers shows that $u^{p^n} + v^{p^n} = 1$ also, so there will be infinitely many solutions if there is one solution.
  - We would like to avoid this sort of triviality, so we will instead focus on separable solutions, where we say $u \in K$ is separable if $K/F(u)$ is separable.
- Theorem (S-Unit Equation): Let K/F be a function field with F perfect and let S be a finite set of primes of K. Then there exist finitely many pairs (u, v) of separable non-constant S-units (u, v) of K such that $u+v = 1$, and any other solution has the form (u^{p^n}, v^{p^n}) where (u, v) is a separable nonconstant solution and p is the characteristic of F.
  - Proof: Suppose that $u + v = 1$ and u, v are nonconstant separable S-units.
  - By the *abc* theorem, with $A = \operatorname{div}_+ u$, $B = \operatorname{div}_+ v$, $C = \operatorname{div}_- u = \operatorname{div}_- v$ as usual, we have $\deg u \le 2g_K - 2 + \sum_{P\in \operatorname{Supp}(A+B+C)} \deg_K P \le 2g_K - 2 + \sum_{P\in S} \deg_K P$.
  - Furthermore, if $A = \sum_{P\in S} a_P P$ then $\deg u = \sum_{P\in S} a_P \deg_K P$, and so $\sum_{P\in S} a_P \deg_K P \le 2g_K - 2 + \sum_{P\in S} \deg_K P$.
  - Note that the upper bound here is constant, and so the a_P are uniformly bounded above. Furthermore, since $\operatorname{Supp}(A) \subseteq S$, there are only finitely many possibilities for A.
  - The same argument shows that there are only finitely many possibilities for B and for C, and thus there are only finitely many possibilities for $\operatorname{div}(u) = A - C$ and for $\operatorname{div}(v) = B - C$.
  - If there were infinitely many solutions to the S-unit equation, then by the pigeonhole principle two pairs would have the same divisor.
  - However, since $\operatorname{div}(u) = \operatorname{div}(u')$ if and only if $u/u' \in F$ is constant, this would yield two solutions $u + v = 1$ and $\alpha u + \beta v = 1$ where $(\alpha, \beta) \ne (1, 1) \in F^2$. But by basic linear algebra this yields $u, v \in F$, a contradiction.
  - For the inseparable part, suppose $u + v = 1$ where u is not separable.
  - For M the maximal separable extension of $F(u)$ in K, by basic facts about inseparable extensions, we know $[K : M] = p^m$ and $u, v = 1 - u$ are both pth powers (in fact, they are p^mth powers).
  - So if $u = u_1^p$ and $v = v_1^p$ we have $(u_1 + v_1)^p = u_1^p + v_1^p = u + v = 1$ so $u_1 + v_1 = 1$. If u_1 is separable, we are done; otherwise, iterate this procedure.
  - This process must eventually terminate since if P is a prime with $\operatorname{ord}_P(u) > 0$, and $u = (u_k)^{p^k}$ then $\operatorname{ord}_P(u)$ is a multiple of p^k, which can only hold for $p^k \le \operatorname{ord}_P(u)$.
  - Thus, we see that any inseparable solution must be of the form (u^{p^n}, v^{p^n}) where (u, v) is a separable solution, as claimed.

- As a corollary, we can deduce that there are only finitely many non-constant separable solutions to Fermat's equation over a finite field:
- Corollary (Fermat Over Finite Fields): Suppose $K/\mathbb{F}_q$ is a function field and $n > 3$ is not divisible by $p = \text{char}(\mathbb{F}_q)$. Then there are finitely many non-constant separable solutions (u, v) to the Fermat equation $u^n + v^n = 1$.
  - Proof: Suppose that $u^n + v^n = 1$ and u, v are nonconstant (note that $\mathbb{F}_q$ is perfect so u, v are automatically separable).
  - We showed earlier in our study of Fermat's equation in function fields that for $A = \text{div}_+ u$, $B = \text{div}_+ v$, $C = \text{div}_- u = \text{div}_- v$ then $(n-3) \cdot \sum_{P \in \text{Supp}(A+B+C)} \deg_K P \leq 6g_K - 6$.
  - Thus, for $n > 3$ we see that $\deg_K P$ is bounded above for $P \in \text{Supp}(A + B + C)$.
  - Since there are only finitely many primes over $\mathbb{F}_q$ of bounded degree (as we showed during our lengthy discussion of the zeta function) this means that u^n and v^n are S-units where $S = \{P : \deg_K P \leq \dfrac{6g_K - 6}{n-3}\}$.
  - Thus, by the S-unit theorem above, there are only finitely many separable solutions for (u^n, v^n) and hence finitely many possibilities for (u, v), as claimed.

0.23 (Dec 6) Proof of the Riemann Hypothesis

- We now embark on our last item, a proof of the Riemann hypothesis for function fields.
  - Let $F = \mathbb{F}_q$ be a finite field and write F_n for the degree-n extension $\mathbb{F}_{q^n}$ of $\mathbb{F}$.
  - We have a correspondence between smooth projective curves and function fields: if C is a nonsingular curve over $\overline{F}$ then its corresponding function field is $\overline{F}(C)$ with constant field $\overline{F}$.
  - Then the Galois group $\text{Gal}(\overline{F}/F)$ acts on C via the coordinate action on points, and also on $\overline{F}(C)$ and the primes of $\overline{F}(C)$.
  - The map sending $\alpha \in C$ to the pair $(\mathcal{O}_\alpha, P_\alpha)$ is onto but not one-to-one: we have $P_\alpha = P_{\alpha'}$ if and only if $\alpha' = \sigma\alpha$ for some $\sigma \in \text{Gal}(\overline{F}/F)$. Equivalently, this says that if P is a prime of degree $d = \deg_K P$ of $\overline{F}(C)$, then P corresponds to a total of d points on C.
  - Recall also that we have the Frobenius map $\varphi : \mathbb{P}^n \to \mathbb{P}^n$ defined via $\varphi[x_0 : \cdots : x_n] = [x_0^q : \cdots : x_n^q]$. Also note that the Frobenius map is a topological generator of $\text{Gal}(\overline{K}/K) \cong \text{Gal}(\overline{F}/F)$; we denote the corresponding element of the Galois group by π. (We think of π as acting on the coefficients of functions $f \in \overline{K}$ by taking the qth power, while we think of φ as acting on coordinates of points by taking the qth power.)
  - If C is defined over F, then φ is a morphism from C to C, and furthermore, $C(F_n) = C(\overline{F})^{\varphi^n}$, the points of C fixed by φ^n.
  - In particular, for $K = C(F)$ we have $K = C(\overline{F})^{\varphi}$, and a prime P of K corresponds to a rational point of C (i.e., P has degree 1) if and only if $\pi^* P = P$.
- We briefly outline the main approach before delving into the various components of the argument:
  - Suppose C has genus g and N_1 is the number of degree-1 primes of K (equivalently, the number of points of $C(F)$).
  - We will first show (1) that if $q > (g+1)^4$ and q is an even power of the characteristic p, then $N_1(K) \leq q + 1 + (2g+1)\sqrt{q}$. The main idea of this argument is to use Riemann-Roch to construct an element with high-order zeroes at all rational points but very few poles.
  - Then we will show (2) a lower bound of similar magnitude (i.e., of the form $q + 1 + C\sqrt{q}$), and so together these results tells us (3) that $N_1(K) = q + O(\sqrt{q})$.
  - Then by using the other components of the Weil conjectures and the properties of the zeta function, we will deduce (4) that the zeta function's inverse zeroes all have magnitude $\sqrt{q}$, which is the Riemann hypothesis for C.

- Recall also the statement of Riemann-Roch: for any divisor A of K, we have $\ell(A) = \deg(A) - g + 1 + \ell(\mathcal{C} - A)$ where $\mathcal{C}$ is the canonical class of A.

- Proposition (RH1): Suppose that $C(F) \neq \emptyset$ and let x be any F-rational point of C with corresponding prime P_x. Define the spaces $R_m = L(mP_x) = \{f \in K : \operatorname{div} f + mP_x \geq 0\}$ for $m \geq 0$. We then have the following properties of R_m:

  1. $\dim R_{m+1} \leq \dim R_m + 1$.
     - Proof: If f, g both have a pole of order $m+1$ at P_x, then for some constant $c \in F$ (given by the ratio of the leading terms of the corresponding power series), $f - cg$ has a pole of order at most m at P_x, hence is in R_m.
  2. $\dim R_m \leq m + 1$.
     - Proof: Immediate from (1) and the fact that $\dim(R_0) = 1$.
  3. $\dim R_m \geq m - g + 1$ with equality when $m > 2g - 2$.
     - Proof: Immediate from Riemann-Roch.
  4. If $f \in R_m$, then $f \circ \varphi$ is a qth power and $\operatorname{div}(f \circ \varphi) = q \cdot \pi^{-1}(\operatorname{div} f)$.
     - Proof: Suppose $f = \dfrac{F}{G}$ for some homogeneous polynomials F, G at the point α.
     - Then $(f \circ \varphi)(\alpha) = \dfrac{F(\varphi(\alpha))}{G(\varphi(\alpha))} = \dfrac{\pi^{-1}(F(\alpha)^q)}{\pi^{-1}(G(\alpha)^q)} = \pi^{-1}[f(\alpha)^q]$ which means $f \circ \varphi = \pi^{-1} f^q$. The given result then follows by taking divisors.
  5. $R_m \circ \varphi \subseteq R_{mq}$, where $R_m \circ \varphi$ consists of the functions $f \circ \varphi$ for $f \in R_m$.
     - Proof: If $f \in R_k$ then $\pi^{-1} f$ is also, since $\operatorname{ord}_{P_x}(\pi^{-1} f) = \operatorname{ord}_{\pi P_x}(f) = \operatorname{ord}_{P_x}(f)$, so $f \circ \varphi = \pi^{-1} f^q \in R_{mq}$.
  6. $\dim(R_m)^{p^e} = \dim R_m$ for any $e \geq 0$.
     - Proof: The map $f \mapsto f^{p^e}$ gives a quasilinear isomorphism of the vector spaces R_m and $(R_m)^{p^e}$ (i.e., an isomorphism up to being the p^eth-power map on the scalars).
  7. $\dim(R_m \circ \varphi) = \dim R_m$.
     - Proof: By definition, composition with φ is an onto map from R_m to $R_m \circ \varphi$, so we just need to check it is one-to-one.
     - But $f_1 \circ \varphi = f_2 \circ \varphi$ is equivalent to $\pi^{-1}(f_1^q) = \pi^{-1}(f_2^q)$ whence $f_1^q = f_2^q$ whence $f_1 = f_2$, as required.

- Now we consider other ways to construct elements of the spaces R_m.
  - If A is a subset of R_m and B is a subset of R_n, then the (ring) product AB, consisting of the collection of finite sums of products $a_1 b_1 + \cdots + a_k b_k$ for $a_i \in A$ and $b_i \in B$, is a subset of R_{m+n}.
  - We then have a natural map from the tensor product $A \otimes_{\overline{F}} B$ to the product AB by mapping the simple tensor $a \otimes b$ to ab and then extending linearly.
  - By parts (1)-(2) of Proposition RH1, since $\dim(R_{a+1}/R_a) \leq 1$ for each a, we can choose a basis of R_m of the form $\{f_1, f_2, \ldots, f_t\}$ such that $v_x(f_i) < v_x(f_{i+1})$ for each i.
  - Now apply the above to $A = R_l^{p^e}$, a subset of R_{lp^e} whose dimension equals $\dim R_l$ by part (6), and $B = R_m \circ \varphi$, a subset of R_{mq} by part (5) whose dimension equals $\dim R_m$ by part (7).
  - The simple tensors of $R_l^{p^e} \otimes_{\overline{F}} (R_m \circ \varphi)$ are of the form $g^{p^e} \otimes (f \circ \varphi)$ for $g \in R_l$. Thus, by expressing everything in terms of the given basis elements $\{f_1 \circ \varphi, \ldots, f_t \circ \varphi\}$ of $R_m \circ \varphi$, we see that every tensor is of the form $\sum_{i=1}^{t} g_i^{p^e} \otimes (f_i \circ \varphi)$ for some $g_i \in R_l$.
  - We claim that all of these tensors are linearly independent, which will allow us to compute the dimension of the tensor product.

- Proposition (RH2): If $lp^e < q$, then the tensor product $R_l^{p^e} \otimes_{\overline{F}} (R_m \circ \varphi)$ is isomorphic to the product $R_l^{p^e}(R_m \circ \varphi)$, and thus has dimension $(\dim R_l)(\dim R_m)$.

- Proof: As above, take the basis $\{f_1, f_2, \dots, f_t\}$ of R_m such that $v_x(f_i) < v_x(f_{i+1})$ for each i, and consider the natural surjection from $R_l^{p^e} \otimes_{\overline{F}} (R_m \circ \varphi)$ to $R_l^{p^e}(R_m \circ \varphi)$.
- We wish to show that this map is also injective, so suppose that $\sum_{i=1}^t g_i^{p^e} \otimes (f_i \circ \varphi)$ is in the kernel of this homomorphism for some $g_i \in R_l$. If all of the g_i are zero we are done, so suppose otherwise and let $g_r \neq 0$ with r minimal.
- Then $g_r^{p^e}(f_r \circ \varphi) = -\sum_{i=r+1}^t g_i^{p^e}(f_i \circ \varphi)$. Now compare orders at P_x on both sides: we have $\text{ord}_x[g_r^{p^e}(f_r \circ \varphi)] = p^e \text{ord}_x(g_r) + q\text{ord}_x(f_r)$ and $\text{ord}_x[-\sum_{i=r+1}^t g_i^{p^e}(f_i \circ \varphi)] \geq \min_{i>r}[p^e \text{ord}_x(g_i) + q\text{ord}_x(f_i)] \geq -lp^e + q\text{ord}_x(f_{r+1})$ since each $g_i \in R_l$ and the f_i are arranged in order of increasing valuation at x.
- Thus we see $p^e \text{ord}_x(g_r) \geq -lp^e + q\,\text{ord}_x(f_{r+1}/f_r) \geq -lp^e + q > 0$, and so g_r is zero at x.
- But $g_r \in R_l$ so the only place it could possibly have a pole is at x, but it doesn't by the above. Hence g_r has no poles and is zero at x, so it equals zero. This is the desired contradiction, and so in fact all of the g_i are zero.
- Thus, the map is in fact an isomorphism is claimed. Then the dimension calculation follows immediately from (6)-(7) of Proposition RH1.

- Now we can prove the first direction in the main result:
- Proposition (RH3): Suppose that $q > (g+1)^4$ and q is an even power of the characteristic p. Then $N_1(K) \leq q + 1 + (2g+1)\sqrt{q}$.
  - The idea here is to use the previous results to construct a function with high-order zeroes at all but one of the rational points (i.e., the elements of K) but with control over the pole order, using judicious selections of the various parameters that we can choose (e, l, m, etc.).
  - Proof: If $N_1(K) = 0$ the result is trivial so suppose otherwise and take x to be a rational point, as earlier.
  - First, take $lp^e < q$. Consider the homomorphism $\delta : R_l^{p^e}(R_m \circ \varphi) \to R_l^{p^e} R_m$ with $\delta(\sum_{i=1}^t g_i^{p^e}(f_i \circ \varphi)) \to \sum_{i=1}^t g_i^{p^e} f_i$; note that this homomorphism is well defined because we showed that the tensor product is isomorphic to the ring product in Proposition RH2.
  - From Proposition RH2, the dimension of the domain is equal to $(\dim R_l)(\dim R_m)$, which is at least $(l-g+1)(l-m+1)$ by Proposition RH1 part (3) whenever $l, m \geq g$.
  - Furthermore, the image of δ is contained in R_{lp^e+m} which (again by RH1 part (3)) has dimension $lp^e + m - g + 1$ since $lp^e + m > 2g - 2$ because we are already assuming $l, m \geq g$.
  - So, by the nullity-rank theorem, as long as the dimension $(l-g+1)(l-m+1)$ of the domain is larger than the dimension $lp^e + m - g + 1$ of the image, there exists a nonzero element of the kernel of δ.
  - Making this assumption, if $f \in \ker(\delta)$, then for any $\alpha \neq x$ in $C(F)$, we have $f(\alpha) = \sum_{i=1}^t g_i(\alpha)^{p^e} f_i(\varphi(\alpha)) = \sum_{i=1}^t g_i(\alpha)^{p^e} f_i(\alpha) = 0$ since φ fixes α. Thus, f vanishes everywhere except at α.
  - Furthermore, because $p^e < q$, by RH1 part (4) we see that f is a p^eth power, so in fact f vanishes to order at least p^e everywhere except α.
  - So, computing degrees yields the bound $p^e(N_1(K) - 1) \leq \deg_{\overline{K}}(\text{div}_+ f) = \deg_{\overline{K}}(\text{div}_- f) \leq lp^e + mq$, which is equivalent to $N_1(K) \leq 1 + l + mq/p^e$.
  - Now we just have to select l, m, and e to achieve the desired bound while also satisfying the assumptions we have made: that $lp^e < q$, that $l, m \geq g$, and that $(l-g+1)(l-m+1) > lp^e + m - g + 1$.
  - Exercise: Verify that if $q = p^{2b}$ and $q > (g+1)^4$, then taking $e = b$, $m = p^b + 2g$, and $l = \lfloor \frac{g}{g+1} p^b \rfloor$ satisfies all of the required relations and yields the bound $N_1(K) \leq 1 + p^b + (p^b + 2g)q = 1 + q + (2g+1)\sqrt{q}$.

- Now we establish a lower bound; the main idea here is to exploit Galois theory.
  - Suppose that L/K is a finite (geometric) Galois extension of K with Galois group $G = \text{Gal}(L/K) \cong \text{Gal}(\overline{L}/\overline{K})$.
  - The Frobenius element π in $\text{Gal}(\overline{L}/L)$ restricts to the Frobenius element of $\text{Gal}(\overline{K}/K)$, which by mild abuse of notation we will also call π. Furthermore, π commutes with G, since $\overline{L} = L\overline{K}$.

- Let T be the set of primes of $\overline{K}$ having K-degree 1, and let $\overline{T}$ be the set of primes of $\overline{L}$ lying above the primes in T.
- If $P \in T$ then $\pi P = P$ (indeed, these are precisely the primes in T) and π permutes transitively the primes above P in $\overline{T}$, so if $\tilde{P}|P$ then there exists $\sigma \in G$ such that $\pi\tilde{P} = \sigma\tilde{P}$.
- Furthermore, we observe that if $\tilde{P}|P$ is unramified, then σ is unique: this follows because $efg = \#G$ and $f = 1$ because $\overline{F}$ is algebraically closed, hence $\#G = g$ (so since the action is transitive, there must be exactly one group element for each prime above P).
- Thus, we get a map $\Omega : \overline{T}^{\text{unram}} \to G$. For each $\sigma \in G$, let $\tilde{T}(\sigma)$ denote the primes in $\overline{T}^{\text{unram}}$ with $\pi\tilde{P} = \sigma\tilde{P}$.
- Since the $\tilde{T}(\sigma)$ partition the unramified primes in $\overline{T}$ and since all of the fibers of Ω have the same cardinality, we see that $\#\overline{T} = \#\overline{T}^{\text{unram}} + \#\overline{T}^{\text{ram}} = (\#G) \cdot N_1(K) + O(1)$, since the number of ramified primes is bounded and is independent of q.
- Equivalently, if we define $N(\sigma, \overline{L}/\overline{K}) = \#\tilde{T}(\sigma)$ for each $\sigma \in G$, this says $\sum_{\sigma \in G} N_1(\sigma, \overline{L}/\overline{K}) = \#\overline{T}^{\text{unram}} = \#\overline{T} + O(1) = (\#G) \cdot N_1(K) + O(1)$.
- In the language of curves, we are really just counting points on the curve C_1 (function field $\overline{L}$) that covers C_2 (function field $\overline{K}$) and comparing the actions of π and G on the preimages of the points in $C_2(F)$.

- Proposition (RH4): Let $\tilde{g} = g_{\overline{L}} = g_L$ and $\sigma \in G$. Suppose that $q > (\tilde{g}+1)^4$ and q is an even power of the characteristic p. Then $N_1(\sigma, \overline{L}/\overline{K}) \le q + 1 + (2\tilde{g}+1)\sqrt{q}$.

  - Proof: This is essentially the same argument as for RH1-RH3 but we instead start with a prime $\tilde{P}_x$ of $\overline{L}$ and define $R_m = L(m\tilde{P}_x)$. All of the same properties of the spaces R_m hold from RH1.
  - Then for RH2, we instead use the map $\delta_\sigma : R_l^{p^e}(R_m \circ \varphi) \to R_l^{p^e}(R_m \circ \sigma)$ defined by $\delta_\sigma(\sum_{i=1}^t g_i^{p^e}(f_i \circ \varphi)) = \sum_{i=1}^t g_i^{p^e}(f_i \circ \sigma)$.
  - The argument for RH2-RH3 then goes through in essentially the same way, except for noting that if $f \in R_m$ then $f \circ \sigma \in L(m\sigma^{-1}\tilde{P}_x)$, meaning that the image of δ_σ is contained in $L(lp^e\tilde{P}_x + m\sigma^{-1}\tilde{P}_x)$, but this still has dimension $lp^e + m - g + 1$ since its degree is not changed. The resulting functions have poles at $\tilde{P}_x$ and also at $\sigma^{-1}\tilde{P}_x$, but the remainder of the argument is not much altered.

- Proposition (RH5): Suppose that $q > (\tilde{g}+1)^4$ and q is an even power of the characteristic p. For all $\sigma \in G$, it is true that $q + 1 + (\#G) \cdot (N_1(K) - q - 1) + O(\sqrt{q}) \le N_1(\sigma, \overline{L}/\overline{K})$.

  - Proof: By RH4, we know that $q + 1 + (2\tilde{g}+1)\sqrt{q} - N_1(\sigma, \overline{L}/\overline{K}) \ge 0$ for each σ.
  - Summing the left-hand side over all σ yields $[(q+1) + (2\tilde{g}+1)\sqrt{q}]\#G - \sum_{\sigma\in G} N_1(\sigma, \overline{L}/\overline{K})$.
  - Now using the fact that $\sum_{\sigma \in G} N_1(\sigma, \overline{L}/\overline{K}) = (\#G) \cdot N_1(K) + O(1)$, and the fact that each term is at most the entire sum (since all terms are nonnegative), yields $q + 1 + (2\tilde{g}+1)\sqrt{q} - N_1(\sigma, \overline{L}/\overline{K}) \le [(q+1) + (2\tilde{g}+1)\sqrt{q}]\#G - (\#G) \cdot N_1(K) + O(1)$.
  - Rearranging yields $q + 1 + (\#G) \cdot (N_1(K) - q - 1) + (2\tilde{g}+1)\sqrt{q} - (2\tilde{g}+1)\sqrt{q}\,\#G + O(1) \le N_1(\sigma, \overline{L}/\overline{K})$, and this implies the required estimate $q + 1 + (\#G) \cdot (N_1(K) - q - 1) + O(\sqrt{q}) \le N_1(\sigma, \overline{L}/\overline{K})$.

- We can now use this result to extract the desired lower bound on $N_1(K)$.

  - If there exists $\alpha \in K$ such that $K/F(\alpha)$ is Galois, then applying RH5 to $(\overline{K}, \overline{F}(\alpha))$ yields $q + O(\sqrt{q}) \le N_1(\sigma, \overline{L}/\overline{K})$ for each $\sigma \in G$.
  - For each σ, we have $N_1(\overline{F}(x)) = q+1$ since $\overline{F}(x)$ is a rational function field, so summing over all σ yields $(\#G) \cdot (q + O(\sqrt{q})) \le \#G \cdot N_1(K)$ so we get $q + O(\sqrt{q}) \le N_1(K)$ as required.
  - Unfortunately, there may not exist such an $\alpha \in K$. However, we can at least find an α such that $K/F(\alpha)$ is separable, and if we then take L to be the Galois closure of $K/F(\alpha)$, then by definition L is Galois over K and the argument above will apply to L.
  - By exploiting the Galois action appropriately, we can then push our calculations back down into K. It may also happen that L could include a constant field extension, but this will turn out to be of no importance since we will see that constant field extensions don't affect things.

- Proposition (RH6): Let K/F have genus g and assume $q = p^{2b}$. Suppose there exists $\alpha \in K$ such that the Galois closure L of $K/F(\alpha)$ has constant field F. If $q > (g+1)^4$ then $N_1(K) = q + O(\sqrt{q})$.
  - Proof: It suffices to show the lower bound, since RH3 supplies the upper bound.
  - Let $G = \text{Gal}(\overline{L}/\overline{F}(\alpha))$ and $H = \text{Gal}(\overline{L}/\overline{K})$. Suppose $\tilde{\tilde{P}}$ is a prime of $\overline{L}$ lying over a degree-1 prime P of $\overline{F}(x)$, and let $\tilde{P}$ be the prime of $\overline{K}$ below $\tilde{\tilde{P}}$.
  - Suppose that $\pi\tilde{\tilde{P}} = \tau\tilde{\tilde{P}}$ for some $\tau \in H$. Then $\pi\tilde{P} = \tau\tilde{P} = \tilde{P}$ (note $\tau\tilde{P} = \tilde{P}$ since H is the Galois group of $\overline{L}/\overline{K}$ hence fixes the prime $\tilde{P}$ of $\overline{K}$), but then $\pi\tilde{P} = \tilde{P}$ implies that $\tilde{P}$ has degree 1 in $\overline{K}$.
  - This means that $N_1(\tau, \overline{L}/\overline{F}(\alpha)) = N_1(\tau, \overline{L}/\overline{K})$.
  - Now, for each $\sigma \in G$, by RH5 applied to $\overline{L}/\overline{F}(\alpha)$, we have $q + O(\sqrt{q}) \le N_1(\sigma, \overline{L}/\overline{F}(\alpha))$.
  - Summing over all elements of H: this yields $(\#H)q+O(\sqrt{q}) \le \sum_{\tau\in H} N_1(\tau, \overline{L}/\overline{F}(\alpha)) = \sum_{\tau\in H} N_1(\tau, \overline{L}/\overline{K}) = (\#H)N_1(K) + O(1)$, and so dividing by $\#H$ yields the desired lower bound $q + O(\sqrt{q}) \le N_1(K)$.
- At last, we can finish the proof of the Riemann hypothesis.
  - We will use a few facts about constant field extensions, the most critical of which is that if $K_n = K\mathbb{F}_{q^n}$ then $L_{K_n}(u) = \prod_{i=1}^{2g}(1 - \pi_i^n u)$.
  - To see this, we observe that if P is any prime of K, then P splits into $d = \gcd(n, \deg_K P)$ primes in K_n each of degree $(\deg_K P)/d$ with $f = n/d$.
  - This follows by noting that $\deg_K P = [\mathcal{O}_P/P : F]$ and that $\mathcal{O}_{\tilde{P}}/\tilde{P} = (\mathcal{O}_P/P)F_n$, so $\deg_{K_n} \tilde{P} = [\mathcal{O}_{\tilde{P}}/\tilde{P} : F_n] = [(\mathcal{O}_P/P)F_n : F_n] = \text{lcm}(\deg_K P, n)$ using the fact that $F_m F_n = F_{\text{lcm}(m,n)}$ has degree $\text{lcm}(m,n)/n = m/\gcd(m,n)$ over F_n. Then $f(\tilde{P}|P) = \dim_{\mathcal{O}_P/P}(\mathcal{O}_{\tilde{P}}/P) = n/d$.
  - Then $L_{K_n}(u) = \prod_{i=1}^{2g}(1 - \pi_i^n u)$ follows by plugging all of this into the zeta function and using the fact that $N(\tilde{P}) = N(P)^{f(\tilde{P}|P)}$.
  - We also note that constant field extensions are unramified everywhere (which follows from a calculation involving the different) and that K_n has the same genus g as K (which follows from Riemann-Hurwitz).
- Theorem (Riemann Hypothesis for Function Fields): Let K be a function field over $\mathbb{F}_q$. Then $\zeta_K(s) = \dfrac{L_K(u)}{(1-u)(1-qu)}$ where $L_K(u) = \prod_{i=1}^{2g}(1 - \pi_i u)$ for elements π_i with $|\pi_i| = q^{1/2}$.
  - Proof: First we observe that it is enough to show the Riemann hypothesis for some constant field extension of K.
  - This follows because for $K_n = K\mathbb{F}_{q^n}$ we know $L_{K_n}(u) = \prod_{i=1}^{2g}(1-\pi_i^n u)$. Thus, if the Riemann hypothesis holds for K_n, then $|\pi_i^n| = (q^n)^{1/2}$ whence $|\pi_i| = q^{1/2}$, so the Riemann hypothesis holds for K.
  - So now we may take a constant field extension K_N of K, which allows us to select q so that $q > (g+1)^4$ and $q = p^{2b}$, and then (possibly by taking another constant field extension) we can assume there exists $\alpha \in K_N$ such that $K_n/F_N(\alpha)$ is separable.
  - Then by our results, we know that $N_1(K_m) = q^m + O(q^{m/2}) = N_m(K)$.
  - Since we also know that $N_m(K) = q^m + 1 - \sum_{i=1}^{2g} \pi_i^m$ from the log-derivative of the zeta function, we see that $q^n + O(q^{m/2}) = N_1(K_m) = N_m(K) = q^m + 1 - \sum_{i=1}^{2g} \pi_i^m$, so $\sum_{i=1}^{2g} \pi_i^m = O(q^{m/2})$.
  - This holds for any multiple n of m as well, which implies that as a power series, $\sum_{i=1}^{2g} \pi_i^n u^n$ has radius of convergence $\ge q^{-1/2}$. This implies $|\pi_i| \le q^{1/2}$, but since by the reflection identity we also have $\pi_i \pi_{2g-i} = q$, we must have equality.
  - This establishes the Riemann hypothesis for K, so we are done.

0.24 (Dec 9) Homework #3 Discussion

Well, you're at the end of my handout. Hope it was helpful.