

FIELD THEORY

PETE L. CLARK

CONTENTS

About these notes	2
0.1. Some conventions and remarks for the reader	2
1. Introduction to Fields	3
2. Some examples of fields	4
2.1. Examples familiar from undergraduate mathematics	4
2.2. Fields of fractions	5
2.3. Fields of functions	8
2.4. Completion	9
3. Field Extensions	12
3.1. Introduction	12
3.2. Some impossible constructions	16
3.3. Subfields of algebraic numbers	17
4. Algebraically closed fields	18
5. Splitting fields	21
5.1. "Uniqueness" of the algebraic closure	21
5.2. Splitting fields	23
5.3. Stable extensions; normal extensions	24
6. Separable algebraic extensions	27
6.1. Separable polynomials	27
6.2. Separable algebraic field extensions	30
6.3. Purely inseparable extensions	31
6.4. Structural results on algebraic extensions	33
6.5. The primitive element theorem	34
7. Galois Extensions	34
7.1. Introduction	35
7.2. Finite Galois extensions	37
7.3. An abstract Galois correspondence	39
7.4. The finite Galois correspondence	41
7.5. Galois groups of infinite algebraic extensions	43
8. Structure of Transcendental Extensions	43
8.1. Applications to algebraically closed fields	45
8.2. An axiomatic approach to independence	46

ABOUT THESE NOTES

The purpose of these notes is to give a treatment of the theory of fields. Some aspects of field theory are popular in algebra courses at the undergraduate or graduate levels, especially the theory of finite field extensions and Galois theory. However, a student of algebra (and many other branches of mathematics which use algebra in a nontrivial way, e.g. algebraic topology or complex manifold theory) inevitably finds that there is more to field theory than one learns in one's standard "survey" algebra courses.¹ For me, when teaching graduate courses in algebra and arithmetic/algebraic geometry, I often find myself "reminding" students of field theoretic facts that they have evidently not seen before, or not seen in the form I wish to use them. I also wish to fill in certain gaps in my own field-theoretic knowledge; especially, I have long wished to gain a deeper understanding of positive characteristic algebraic geometry, and has become clear that the place to begin study of the "pathologies"² of positive characteristic is the study of finitely generated field extensions in positive characteristic.

These notes are meant to be comprehensible to students who have taken a basic graduate course in algebra. In theory one could get away with less – the exposition is for the most part self-contained. As far as algebraic prerequisites one hardly needs to know any group or ring theory *per se*; mostly one just needs to know what these things mean. But I would imagine that the so-called mathematical maturity necessary to read these notes without undue effort is not possessed by the average undergraduate student.

0.1. Some conventions and remarks for the reader.

By convention, all of our rings are associative and have a multiplicative unity, called 1. Again by convention, a homomorphism of rings necessarily carries 1 to 1.

These notes contain many exercises, including some which ask for proofs of stated results. In general I am not at all opposed to the idea of a text giving complete details for all of its arguments.³ However, it is in the nature of this particular subject that there are many more results than proof techniques, to the extent that giving complete proofs of all results would create a lengthy repetitiveness that may discourage the reader to read the proofs that we do give.

As a rule, exercises that ask for proofs of stated results are meant to require no new ideas beyond what was (even recently) exposed in the text. A reader who feels otherwise should contact me: there may well be an unintended gap in the exposition. On the other hand, if exercises are given at all, it certainly spruces things up to have some more challenging and interesting exercises. I have also not hesitated to give exercises which can in principle be solved using the material up to that point but become much easier after later techniques are learned.

At some point I fell victim to the disease of not liking the look of a paragraph of

¹Obviously this phenomenon is not unique to field theory...

²The term was used by Mumford, but with evident affection. A more accurate word would be "idiosyncrasy".

³In fact I agree with Robert Ash that the prevailing negative reputation of such texts is undeserved: the royal road to a particular destination may or may not exist, but it seems perverse to claim that it *ought not* to exist.

text in which only a few words appear on the last line. In order to remedy this in the exercises I have sometimes omitted the words “Show that”. I hope the intended meaning remains clear.

1. INTRODUCTION TO FIELDS

A **field** is a commutative ring in which each nonzero element has a multiplicative inverse. Equivalently, a field is a commutative ring R in which the only ideals are (0) and R itself.

So if F is a field, S is a ring, and $\varphi : F \rightarrow S$ is a homomorphism of rings, then since the kernel of φ is an ideal of F , φ is either injective (if its kernel is 0) or identically the zero map (if its kernel is F). Moreover, the latter case implies that $1_S = \varphi(1_F) = 0$, which happens iff S is the zero ring. So any homomorphism from a field into a nonzero ring – in particular into any field or integral domain – is injective. Thus if $\varphi : F \rightarrow K$ is a homomorphism between fields, we may equally well speak of the **field embedding** φ .

Variations on the definition: In older terminology, a field could be non-commutative, i.e., any ring in which each nonzero element has a two-sided multiplicative inverse. We now call such things “division rings” or “division algebras.” One also sometimes encounters non-associative division algebras, e.g. Cayley’s octonions.

The two branches of mathematics in which general fields play a principal role are field theory (of course) and linear algebra. Most of linear algebra could be developed over a general division algebra rather than over a general field. In fact for the most part the theory is so similar that it is not really necessary to consider division algebras from the outset: one can just check, if necessary, that a certain result which is true for vector spaces over a field is also true for left modules over a division algebra. On the other hand, when one studies things like roots of polynomials and lattices of finite degree extensions, one immediately finds that non-commutative division algebras behave in quite different and apparently more complicated ways.

Example 1.1: There are exactly two complex numbers z such that $z^2 = -1$: $z = i$ and $z = -i$. In general, any nonzero polynomial $P(t)$ with coefficients in a field can have no more solutions than its degree. But in Hamilton’s quaternion algebra \mathbb{H} there are clearly at least three solutions: $i^2 = j^2 = k^2 = -1$, and in fact there are uncountably many: a quaternion squares to -1 iff it is of the form $xi + yj + zk$ with $x^2 + y^2 + z^2 = 1$.

Example 1.2: Let K/\mathbb{Q} be a quartic field (i.e., a field extension of \mathbb{Q} which has dimension 4 as a \mathbb{Q} -vector space). Then there are at most three intermediate subfields $\mathbb{Q} \subsetneq F \subsetneq K$. (More precisely there is either zero, one or three such fields, and the first case happens “most of the time.”) However, any noncommutative division algebra B/\mathbb{Q} of degree 4 as a \mathbb{Q} -vector space has infinitely many nonisomorphic quadratic subfields.⁴

It is certainly true that the study of division algebras is closely related to field

⁴This is certainly not obvious, but it is far enough away from our interests that we shall say nothing more about it here.

theory – via Brauer groups and Galois cohomology – so that one can put one’s understanding of a field F and its finite extensions to excellent use in studying noncommutative division algebras over F . In fact, notwithstanding the above two examples, the finite dimensional, central division algebra over a field F are significantly easier to understand than finite dimensional extension fields of F : e.g. we understand quaternion algebras over \mathbb{Q} far better than quartic number fields.

2. SOME EXAMPLES OF FIELDS

2.1. Examples familiar from undergraduate mathematics.

Example 2.1.1: First of all there is the field of real numbers \mathbb{R} . One also encounters the complex numbers $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\}$ and the rational numbers $\mathbb{Q} = \{\frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\}\}$.

Example 2.1.2: For a prime p , the ring $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ of integers modulo p is a field. In fact it is enough to show that it is an integral domain, since any finite integral domain must be a field: if a is a nonzero element of a finite integral domain, there must exist $0 < i < j$ such that $a^i = a^j$, and then by cancellation we get $1 = a^{j-i} = a^{j-i-1}a$. To check that \mathbb{F}_p is an integral domain, suppose that x, y are nonzero elements in \mathbb{F}_p such that $0 = xy$. Equivalently, we have integers x, y not divisible by p but such that $p \mid xy$. This contradicts the uniqueness of factorization of integers into primes, i.e., the “Fundamental Theorem of Arithmetic.”

Nonexample 2.1.3: The ring of integers $\mathbb{Z}/n\mathbb{Z}$ is not a field unless n is prime: if $n = n_1 \cdot n_2$ with $n_1, n_2 > 1$, then $(n_1 \pmod{n}) \cdot (n_2 \pmod{n}) = 0 \pmod{n}$ exhibits zero divisors.

Example 2.1.4: Suppose there were a field \mathbb{F}_4 with four elements. One of the elements would be 0, another would be 1, and that leaves two further elements x and y . The nonzero elements of any field form a group under multiplication, so in this case the group would have order 3 and therefore be cyclic. In particular x has order 3, hence so does x^2 , so x^2 is equal to neither 0, 1 or x , and therefore $x^2 = y = x^{-1}$ and $y^2 = x = y^{-1}$. Also $x + y$ cannot equal x or y ; if $x + y = 0$, then $x = -y = y$ since $-1 = 1$ in \mathbb{F}_4 . Therefore we must have $x + y = 1$, i.e., $y = x - 1 = x + 1 = x^2$. We have thus uniquely worked out the addition and multiplication table for our putative field of order four, and one can check directly that all the field axioms are satisfied: there is, indeed, a field of order four. Moreover there is a unique such field up to isomorphism, and moreover the mapping $x \rightarrow y = x^2$ gives the unique nontrivial isomorphism from \mathbb{F} to itself.

Nonexample 2.1.5: Suppose \mathbb{F} is a field of order 6. The additive group of \mathbb{F} is then commutative of order 6 and hence cyclic, generated say by x . But then $2x \neq 0$, $3x \neq 0$ but $2x \cdot 3x = 6x = 0$, so \mathbb{F} is not an integral domain. Similar arguments hold whenever the order of \mathbb{F} is divisible by two distinct primes p, q . Therefore the order of a finite field must be the power of a prime p^f , and the characteristic of \mathbb{F} , i.e., the least $n \in \mathbb{Z}^+$ such that $n \cdot 1 = 0$ must be p . In other words, the additive group of \mathbb{F} must be an elementary abelian p -group: $(\mathbb{F}, +) \cong (\mathbb{Z}/p\mathbb{Z})^a$.

We will see later that for every prime power p^a there is a finite field \mathbb{F} of order p^a , that any two finite fields of order p^a are isomorphic, and that the automorphism group of a finite field of order p^a is cyclic of order a , generated by the “Frobenius map” $x \mapsto x^p$.

2.2. Fields of fractions. If R is an integral domain, then it is well-known that one can define a field F whose elements are viewed as fractions $\frac{a}{b}$ with $a, b \in R, b \neq 0$. Formally speaking one considers ordered pairs $(a, b) \in R^2, b \neq 0$ and introduces the equivalence relation $(a, b) \sim (c, d) \iff ad = bc$, i.e., exactly the same construction that one uses to define the rational numbers in terms of the integers. The field F is called the field of fractions, or quotient field, of the integral domain R .

Remark: The reader may well recognize this as a special case of localization in the study of ring theory, but the concept is (even) more elementary and transparent when restricted to fields. This is a common feature of field theory: treating it as a special case of the theory of commutative rings is, in many instances, overkill.

Exercise 2.2.1 (Functoriality of the field of fractions): Let $\varphi : R \rightarrow S$ be an injective homomorphism of integral domains. Show that φ extends uniquely to a homomorphism from the fraction field $F(R)$ of R to the fraction field $F(S)$ of S .

Exercise 2.2.2 (Universal property of the field of fractions): Let R be an integral domain with fraction field F and let K be a field. Show that for any injective homomorphism $\varphi : R \rightarrow K$, there exists a unique extension to a homomorphism $F \rightarrow K$.

Exercise 2.2.3: Show that for any integral domain R , the cardinality of R is equal to the cardinality of its field of fractions $F(R)$. Note that this generalizes the fact that any finite integral domain is a field.

Thus any method which produces a supply of integral domains will also produce a supply of fields (of course distinct integral domains may have isomorphic fraction fields, a trivial example being \mathbb{Z} and \mathbb{Q} itself; there are in fact uncountably many isomorphism classes of integral domains with fraction field \mathbb{Q}).

Proposition 1. *If R is an integral domain, then the univariate polynomial ring $R[t]$ is also an integral domain. Moreover, if F is the fraction field of R , then the fraction field of $R[t]$ is $F(t)$, the field of all quotients of polynomials with F -coefficients.*

Exercise 2.2.4: Prove Proposition 1.

Example 2.2.5: Applying the Proposition with $R = F$ a field, we get a field $F(t)$ of rational functions in F . E.g., the field $\mathbb{C}(t)$ is the field of meromorphic functions on the Riemann sphere (see next section). Moreover, for any field F , $F[t]$ is a domain, so $F[t_1, t_2] := F[t_1][t_2]$ is also an integral domain. The fraction field is easily seen to be $F(t_1, t_2)$, i.e., the fraction field of $F[t_1, \dots, t_n]$ is $F(t_1, \dots, t_n)$ the field of rational functions in n indeterminates.

Although successive applications of Proposition 1 will yield polynomial rings in

only finitely many indeterminates, nothing stops us from considering larger polynomial rings: let $\mathbb{T} = \{t_i\}$ be any set of indeterminates, and R any commutative ring. One can consider the polynomial ring $R[\mathbb{T}]$, defined as the union (or, if you like, direct limit) of polynomial rings $R[S]$ where $S \subset \mathbb{T}$ is a finite subset. In other words, we consider the ring of polynomials in an arbitrary infinite set S of indeterminates, but any given polynomial involves only finitely many indeterminates. One can again show that if R is an integral domain, so is $R[\mathbb{T}]$. The corresponding fraction field $R(\mathbb{T})$ is the field of all quotients of polynomials in all these indeterminates.

Exercise 2.2.6: Let F be a field and \mathbb{T} a nonempty set of indeterminates. Show that the cardinality of the rational function field $F(\mathbb{T})$ is $\max(\aleph_0, \#F, \#\mathbb{T})$.

Another way of manufacturing integral domains is to start with a commutative ring R and take the quotient by a prime ideal \mathfrak{p} . Then we can get a field by (if necessary, i.e., if \mathfrak{p} is not maximal) taking the field of fractions of R/\mathfrak{p} . For example with $R = \mathbb{Z}$ we get the finite fields \mathbb{F}_p .

Example 2.2.7: Let $R = F[T]$ and \mathfrak{p} a prime ideal. Then, since R is a PID, $\mathfrak{p} = (P(t))$, where $P(t)$ is an irreducible polynomial. Moreover, assuming $P(t) \neq 0$, \mathfrak{p} is maximal, so without having to take quotients we get a field

$$F_P = F[t]/(P(t)),$$

whose dimension as an F -algebra is the degree of P . This is perhaps the most important single example in all of field theory.

Let R be an integral domain. We say that R is finitely generated (over \mathbb{Z}) if there exists $n \in \mathbb{Z}^+$ and elements $\alpha_1, \dots, \alpha_r \in R$ such that the least subring of R containing all the α_i 's is R itself. Another way of saying this is that the natural map

$$\mathbb{Z}[T_1, \dots, T_n] \rightarrow R, T_i \mapsto \alpha_i$$

is surjective. In other words, an integral domain is finitely generated iff it is, for some n , the quotient of the ring $\mathbb{Z}[T_1, \dots, T_n]$ by some prime ideal \mathfrak{p} .

Proposition 2. *For a field F , the following are equivalent:*

- There exist $\alpha_1, \dots, \alpha_n \in F$ so that the only subfield of F containing all the α_i 's is F itself.*
- F is the fraction field of $\mathbb{Z}[x_1, \dots, z_n]/\mathfrak{p}$ for some prime ideal \mathfrak{p} .*

Exercise 2.2.8: Prove Proposition 2.

A field satisfying the equivalent conditions of Proposition 2 is said to be **finitely generated**. Note that applying part b) and Exercise XX we see that any finitely generated field is finite or countably infinite; in particular the fields \mathbb{R} , \mathbb{C} are not finitely generated. Conversely, a countable field need not be finitely generated: if \mathbb{T} is a countably infinite set of indeterminates, then by Exercise X.X the field $\mathbb{Q}(\mathbb{T})$ is countable; moreover it is both plausible and true that $\mathbb{Q}(\mathbb{T})$ is not finitely generated, but we lack the tools to prove this at the moment: we will return to this later on in the context of the concept of **transcendence degree**.

One can also speak of finite generation in a relative sense:

Proposition 3. For a subfield $f \subset F$, the following are equivalent:

- There exist elements $\alpha_1, \dots, \alpha_n \in F$ such that the only subfield of F containing f and the α_i 's is F itself.
- F is isomorphic to the fraction field of $f[x_1, \dots, x_n]/\mathfrak{p}$ for some prime ideal \mathfrak{p} .

Exercise 2.2.9: Prove Proposition 3.

If f is a subfield of F and $\alpha_1, \dots, \alpha_n \in F$, we write $f(\alpha_1, \dots, \alpha_n)$ for the smallest subfield of F containing f and the α_i 's. The notation is sensible because this field can be described concretely as the set of all rational expressions $\frac{P(\alpha_1, \dots, \alpha_n)}{Q(\alpha_1, \dots, \alpha_n)}$ for $P, Q \in k[t_1, \dots, t_n]$. (In particular there is a unique such smallest subfield.)

So for instance one can speak of fields which are finitely generated over the complex numbers \mathbb{C} , and such fields are especially important in algebraic geometry.

To relate the absolute and relative notions of finite generation, we need to introduce the concept of a **prime subfield**. Let F be a field. Like any integral domain, it is either the case that $n \cdot 1 \neq 0$ for all $n \in \mathbb{Z}^+$, or the least $n \in \mathbb{Z}^+$ such that $n \cdot 1 = 0$ is a prime number p . In the first case we say that F has **characteristic zero**; in the second case we say F has **positive characteristic p** .

Proposition 4. Let F be a field.

- If F has characteristic 0, there is a unique homomorphism $\iota : \mathbb{Q} \rightarrow F$.
- If F has characteristic p , there is a unique homomorphism $\iota : \mathbb{F}_p \rightarrow F$.

Proof. For any ring R , there exists a unique ring homomorphism $\iota : \mathbb{Z} \rightarrow R$, which takes the integer n to n times the multiplicative identity in R . For $R = F$ a field, the map ι is an injection iff F has characteristic 0. So if F has characteristic 0, ι is injective, and by Exercise 2.3.2 it extends uniquely to a homomorphism $\iota : \mathbb{Q} \rightarrow F$. Any homomorphism from \mathbb{Q} to F would have to restrict to the canonical injection on \mathbb{Z} and is therefore equal to ι . If F has characteristic $p > 0$, then ι factors through to give a map $\iota : \mathbb{F}_p \rightarrow F$. That this homomorphism is unique can be seen in any number of ways: we leave it to the reader to find an explanation that she finds simple and convincing. \square

It follows that \mathbb{Q} (resp. \mathbb{F}_p) is the unique minimal subfield of any field F of characteristic 0 (resp. $p > 0$). We refer to \mathbb{Q} (resp. \mathbb{F}_p) as the **prime subfield** of F . Note that since there are no nontrivial automorphisms of either \mathbb{Q} or \mathbb{F}_p (this follows by applying the proposition with $F = \mathbb{Q}$ or $F = \mathbb{F}_p$), the prime subfield sits inside F in an especially canonical way.

Exercise 2.2.10: Let F be a field and f its prime subfield. Show that the concepts “ F is finitely generated” and “ F is finitely generated over f ” coincide.

Exercise 2.2.11: For any field F , there exists a set of indeterminates \mathbb{T} and a prime ideal \mathfrak{p} of $\mathbb{Z}[\mathbb{T}]$ such that F is isomorphic to the fraction field of $\mathbb{Z}[\mathbb{T}]/\mathfrak{p}$.

If F is infinitely generated (i.e., not finitely generated over its prime subfield) then the set \mathbb{T} in Exercise 2.3.11 will of course have to be infinite. In such a case this “presentation” of F is not, in truth, so useful: e.g., with certain limited exceptions

(to be discussed!) this is not a very insightful way of viewing the complex field \mathbb{C} .

Exercise 2.2.12: Let R be a commutative ring, $\iota : R \rightarrow S$ an injective ring homomorphism, and $\alpha \in S$. Show that there is a unique minimal subring of S containing R and α , namely the set of all polynomials $P(\alpha), P \in R[t]$. This subring is accordingly denoted $R[\alpha]$.

2.3. Fields of functions.

Let U be a domain – i.e., a nonempty connected open subset – of the complex plane. In complex analysis one studies the set $\text{Hol}(U)$ of all functions holomorphic (a.k.a. analytic) on all of U and also the larger set $\text{Mer}(U)$ of all meromorphic functions on U , i.e., functions which are holomorphic on the complement of a discrete set $X = \{x_i\}$ and such that for each x_i there exists a positive integer n_i such that $z^{n_i} f$ is holomorphic at x_i . Under the usual pointwise addition and multiplication of functions, $\text{Hol}(U)$ is a ring (a subring of the ring of all continuous \mathbb{C} -valued functions on U). Similarly, one can view $\text{Mer}(U)$ as a ring in a natural way – note that strictly speaking it is not a ring of functions on U because an element of $\text{Mer}(U) \setminus \text{Hol}(U)$ will not be defined on all of U .

Theorem 5. *Let U be a domain in the complex plane.*

- a) $\text{Hol}(U)$ is an integral domain.
- b) $\text{Mer}(U)$ is a field.
- c) $\text{Mer}(U)$ is the field of fractions of $\text{Hol}(U)$.

Proof. a) A consequence of the principle of analytic continuation is that the zero set of a not-identically-zero holomorphic function is discrete in U . For $0 \neq f, g \in \text{Hol}(U)$, the zero set of fg is the union of the zero sets of f and g so is again discrete and thus certainly a proper subset of U .

b) Because $0 \neq f \in \text{Hol}(U)$ has a discrete zero set $\{x_i\}$ and for each x_i , there exists a positive integer n_i such that $\frac{f}{z^{n_i}}$ extends to a continuous nonzero function at x_i , it follows that $\frac{1}{f_i}$ is meromorphic.

c) This lies deeper: it is a consequence of Weierstrass' factorization theory, in particular of the fact that for any discrete subset $X = \{x_i\}$ of U and any sequence of positive integers $\{n_i\}$ there exists a holomorphic function on U with zero set X and order of vanishing n_i at x_i . \square

Exercise 2.3.1: Show that $\text{Mer}(\mathbb{C})$ is not finitely generated over \mathbb{C} .

More generally, if M is a connected complex manifold, there is a ring $\text{Hol}(M)$ of “global” holomorphic functions on M and a field $\text{Mer}(M)$ of meromorphic functions. It need not be the case that $\text{Mer}(M)$ is the fraction field of $\text{Hol}(M)$.

Example 2.3.2: Take $M = \mathbb{C} \cup \{\infty\}$ to be the Riemann sphere. Then the only holomorphic functions on M are the constant functions, whereas $\text{Mer}(M) = \mathbb{C}(z)$, the rational functions in z .

In various branches of geometry one meets many such “fields of functions”: a very general example, for the highly trained reader, is that if X is an integral (reduced and irreducible) scheme, then the ring of all functions regular at the generic point

η is a field. If X itself is a scheme over a field k , then this field is written $k(X)$ and called the **field of rational functions** on X . For example, the field of rational functions on the complex projective line \mathbb{P}^1/\mathbb{C} is the rational function field $\mathbb{C}(t)$. This is essentially the same example as the Riemann sphere above, but couched in more algebraic language.

We wish to bring to the reader's attention is that one must restrict to functions of a rather special kind in order to get a field of functions. Using the ideas of the previous subsection, it seems fruitful to first consider *rings* R of functions on a topological space X . Then we want R to be an integral domain in order to speak of fraction field $F(R)$ of "meromorphic functions" on X .

Suppose that X is a topological space and we consider the ring $R = R_c(X, \mathbb{C})$ of all continuous functions $f : X \rightarrow \mathbb{C}$. A moment's thought indicates that for the "reasonable" topological spaces one considers in geometry, R will not be an integral domain. The question comes down to: do there exist functions $f_1, f_2 : X \rightarrow \mathbb{C}$ neither of which is zero on all of X but such that the product $f_1 \cdot f_2$ is identically zero?

Here are some easy observations. First, if X is not connected, the answer is certainly yes: write $X = Y_1 \cup Y_2$ where Y_i are disjoint open sets. Take f_1 to be the characteristic function of Y_1 and $f_2 = 1 - f_1$ to be the characteristic function of Y_2 .

In fact R is not a domain even if X is the Euclidean plane: let D_1, D_2 be two disjoint closed disks, say with centers z_i and radii equal to 1. Certainly there exist continuous functions $f_i : X \rightarrow \mathbb{C}$ such that $f_i(z_i) = 1$ and $f_i(z) = 0$ if z lies outside of D_i . Indeed it is well-known that f_i may be chosen to be infinitely differentiable, and the argument generalizes to all manifolds and indeed to paracompact Hausdorff spaces (the key point being the existence of suitable partitions of unity).

On the other hand, suppose the space X is **irreducible**: that is, if Y_1, Y_2 are two proper closed subsets of X then $Y_1 \cup Y_2 \neq X$. Then, applying this to $Y_i = f_i^{-1}(0)$, we get that the zero set of $f_1 f_2$ is $Y_1 \cup Y_2 \neq X$, so $C_c(X, \mathbb{C})$ is a domain, and one can take its fraction field, which consists of functions which are defined on some dense open subset of X . If you have never studied algebraic geometry than you will doubtless be thinking, "What kind of crazy topological space would be irreducible?" However, the Zariski topology on a smooth, connected algebraic variety over (say) the complex field \mathbb{C} is irreducible. The function fields one gets in this manner are precisely the fields which are finitely generated over \mathbb{C} .

2.4. Completion.

None of the constructions of fields we have discussed so far give rise to either \mathbb{R} or \mathbb{C} in a reasonable way. Since these fields are uncountable, a purely algebraic discussion of their structure is rather complicated: the right way to think about these fields is via a mixture of algebra and topology, e.g. one thinks of \mathbb{R} as the completion of the field of rational numbers with respect to the standard absolute value.

An **absolute value** on a field K is a real-valued function $x \rightarrow \|x\|$ satisfying:

(AV1) $\|x\| \geq 0$ for all $x \in K$, with equality iff $x = 0$.

(AV2) $\|xy\| = \|x\|\|y\|$ for all $x, y \in K$.

(AV3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in K$.

It is immediate that an absolute value gives rise to a metric on K , via $d(x, y) = \|x - y\|$. We can therefore complete the metric space to get a metric space \hat{K} with a canonically embedded, dense copy of K . The key point is that \hat{K} also has a canonical field structure.

DESCRIBE THIS, USING equivalence classes of Cauchy sequences.

Example 2.4.1: The completion of \mathbb{Q} with the standard Archimedean absolute value $\|\frac{p}{q}\| = |\frac{p}{q}|$ is the real field \mathbb{R} .

Remark 2.4.2: It is sometimes suggested that there is a circularity in this construction, in that the definition of completion refers to a metric and the definition of a metric refers to the real numbers. (In particular, Bourbaki's *General Topology* refrains from making any reference to real numbers or metric spaces for many hundreds of pages until the reals can be rigorously constructed.) But one should not worry about this. On the one hand, from our present point of view we can consider the reals as being already constructed and then it is a true, non-tautologous statement that the metric completion of the rationals is the reals. But moreover, a careful look at the construction in terms of equivalence classes of Cauchy sequences shows that one absolutely can construct the real numbers in this way, just by being careful to avoid referring to the real numbers in the course of the completion process. In other words, the real numbers can be defined as the quotient of the ring of Cauchy sequences of rational numbers (where the definition of Cauchy sequence uses only the metric as defined on rational numbers) by the maximal ideal of sequences converging to zero. After one constructs the real numbers in this way, one notes that the \mathbb{Q} -valued metric on \mathbb{Q} extends to an \mathbb{R} -valued metric on \mathbb{R} : no problem.

Example 2.4.3: If k is any field, then defining $\|0\| = 0$ and $\|x\| = 1$ for all $x \neq 0$ gives an absolute value on k . The induced metric is the discrete metric and therefore k is, in a trivial way, complete and locally compact. This absolute value (and any other absolute value inducing the discrete topology) is called **trivial**; such absolute values are usually either explicitly or implicitly excluded from consideration.

Example 2.4.4: For a prime number p , we define an absolute value $x \mapsto \|x\|_p$ on \mathbb{Q} , as follows: $\|\frac{a}{b}\|_p = p^{\text{ord}_p(b) - \text{ord}_p(a)}$, where for an integer a , $\text{ord}_p(a)$ denotes the largest power of p dividing a . (To get the degenerate cases to work out correctly, we set $\text{ord}_p(0) = \infty$ and $p^{-\infty} = 0$.) The induced metric on \mathbb{Q} is called the p -adic metric: in this metric, a number is close to zero if, after cancelling common factors, its numerator is divisible by a high power of p . Since the induced topology has no isolated points, the completeness of the metric would contradict the Baire category theorem, hence the completion is an uncountable field, called \mathbb{Q}_p , the field of p -adic

numbers.

Example 2.4.5: Let k be any field and $K = k(t)$. Any element $r(t) \in K$ can be written as $t^a \frac{P(t)}{Q(t)}$ where $P(0)Q(0) \neq 0$ for a uniquely determined integer a . Define $\|r(t)\|_\infty := e^{-a}$. (There is no particular reason to use the number $e = 2.718\dots$; and real number greater than 1 would serve as well.)

Exercise 2.4.6: Show that $\|\cdot\|_\infty$ gives an absolute value on $K(t)$.

An element $r(t) \in K(t)$ is close to 0 iff it is divisible by a high power of t .

Exercise 2.4.7: Show that the completion of $K(t)$ with respect to $\|\cdot\|_\infty$ is isomorphic to the **Laurent series field** $K((t))$, whose elements are formal power series $\sum_{n=n_0}^{\infty} a_n t^n$ with $n_0 \in \mathbb{Z}$, $a_n \in f$. (Hint: It is enough to show that the norm $\|\cdot\|_\infty$ extends to all of $K((t))$ and that $K(t)$ is dense in $K((t))$ in the induced topology.)

Exercise 2.4.8: Show that the fields \mathbb{Q}_p are locally compact in their natural topology. Show that $K((t))$ is locally compact iff K is finite.

Remark 2.4.9: If k is a field complete with respect to an absolute value $|\cdot|$ and V is a finite-dimensional vector space over k , then viewing $V \cong k^{\dim V}$ gives V the canonical structure of a topological space – i.e., we can endow it with the product topology, and this topology is independent of the choice of basis. In particular, if k is locally compact, so is V . Moreover it has the canonical structure of a uniform space, and if k is complete then so is V . In particular, if $k \hookrightarrow l$ is a field embedding such that l is finite-dimensional as a k -vector space, then l is a complete uniform space and is locally compact iff k is. This implies that any finite extension of the fields \mathbb{R} , \mathbb{Q}_p or $\mathbb{F}_p((t))$ have a canonical locally compact topology.

Theorem 6. (*Classification of locally compact valued fields*) Let $\|\cdot\|$ be a nontrivial valuation on a field K . TFAE:

- (i) The metric topology on K is locally compact.
- (ii) Either $(K, \|\cdot\|) = \mathbb{R}$ or \mathbb{C} ; or the induced metric is complete and non-Archimedean and the residue field is finite.
- (iii) K is a finite extension of \mathbb{R} , of \mathbb{Q}_p or of $\mathbb{F}_p((t))$.

Proof. Give reference! □

There are more elaborate ways to construct complete fields. For instance, suppose R is an integral domain and \mathfrak{p} is a prime ideal of R . Then in commutative algebra one learns how to complete R with respect to \mathfrak{p} , getting a homomorphism $R \rightarrow \hat{R}$ in which \hat{R} is an integral domain, the image $\mathfrak{p}\hat{R}$ is the unique maximal ideal of \hat{R} , and \hat{R} is complete with respect to a canonical uniform structure. We can then take the fraction field to get a complete field \hat{K} . Let us just mention one simple example to give the flavor: let f be a field and $R = f[x_1, \dots, x_n]$ and $\mathfrak{p} = (x_1, \dots, x_n)$. Then the completion is $\hat{R} = f[[x_1, \dots, x_n]]$, the ring of formal power series in the

indeterminates x_1, \dots, x_n , and its quotient field is $f((x_1, \dots, x_n))$, the field of formal Laurent series in these indeterminates, i.e., the set of all formal sums $\sum_I a_I x^I$ where $I = (i_1, \dots, i_n) \in \mathbb{Z}^n$ is a multi-index, $a_I \in k$, $x^I = x^{i_1} \cdots x^{i_n}$, and the set of indices I in which at least one i_j is negative and $a_I \neq 0$ is finite.

Such fields arise in algebraic and analytic geometry: $\mathbb{C}((x_1, \dots, x_n))$ is the field of germs of meromorphic functions at a nonsingular point P on an n -dimensional analytic or algebraic variety.

Exercise 2.4.10: Show that the field $k((x_1, x_2))$ is properly contained in $k((x_1))(x_2)$.

3. FIELD EXTENSIONS

3.1. Introduction.

Let K be a field. If $\iota : K \rightarrow L$ is a homomorphism of fields, one says that L is an **extension field** of K . As a matter of psychology, it often seems more convenient to think of L as “lying above K ” rather than K as embedding into L . We often write L/K instead of $\iota : K \rightarrow L$, notwithstanding the fact that the latter notation hides important information, namely the map ι .⁵

Much of field theory is devoted to an understanding of the various extension fields of a given field K . Since any field K has extensions of all sufficiently large cardinalities – $K(\mathbb{T})$ for any large enough set \mathbb{T} – one obviously cannot literally hope to understand all field extensions of K . However there are two important classes (sets!) of field extensions that one can at least hope to understand: the first is the class of all finitely generated field extensions of K , and the second is the class of all algebraic field extensions of K .

If L/K is a field extension, then L is a K -algebra and in particular a vector space over K . Therefore it has a well-determined (but possibly infinite) dimension, denoted by $[L : K]$. One says that the extension L/K is **finite** if $[L : K] < \infty$, i.e., if L is a finite-dimensional K -vector space. For instance, one has $[\mathbb{C} : \mathbb{R}] = 2 < \infty$, so \mathbb{C}/\mathbb{R} is a finite field extension.

Warning 3.1.1: The term “finite field extension” is ambiguous: it could presumably also refer to an extension of fields L/K in which L and K are both finite fields. In practice, one should expect the term to have the former meaning – i.e., the finiteness refers to the degree of the extension, and not to either field – but be prepared to seek clarification if necessary.

As an immediate application we can rederive the fact that the order of a finite field is necessarily a prime power. Namely, let \mathbb{F} be a finite field, and let \mathbb{F}_p be its prime subfield. Since \mathbb{F} is finite, it is certainly finite-dimensional over \mathbb{F}_p (any infinite dimensional vector space over any field is infinite), say of dimension d . Then \mathbb{F} as an \mathbb{F}_p -vector space is isomorphic to \mathbb{F}_p^d , so its cardinality is p^d .

⁵Beware: the notation L/K has nothing to do with cosets or quotients!

Theorem 7. (*Degree multiplicativity in towers*) Let $F \subset K \subset M$ be field extensions. Then $[M : F] = [M : K][K : F]$.

Proof. Let $\{b_i\}_{i \in I}$ be an F -basis for K and $\{a_j\}_{j \in J}$ be a K -basis for M . We claim that $\{a_i b_j\}_{(i,j) \in I \times J}$ is an F -basis for M . This suffices, since then $[K : F] = \#I$, $[M : K] = \#J$, $[M : F] = \#(I \times J) = \#I \times \#J$.

Let c be an element of M . Then there exist $\alpha_j \in K$, all but finitely many of which are zero, such that $c = \sum_{j \in J} \alpha_j a_j$. Similarly, for each $j \in J$, there exist $\beta_{i,j} \in F$, all but finitely many of which are zero, such that $\alpha_j = \sum_{i \in I} \beta_{i,j} b_i$, and thus

$$c = \sum_{j \in J} \alpha_j a_j = \sum_{(i,j) \in I \times J} \beta_{i,j} a_i b_j,$$

so that $\{a_i b_j\}$ spans M as an F -vector space. Now suppose the set $\{a_i b_j\}$ were linearly dependent. By definition, this means that there is some finite subset $S \subset I \times J$ such that $\{a_i b_j\}_{(i,j) \in S}$ is linearly dependent, and thus there exist $\beta_{i,j} \in F$, not all zero, such that

$$\sum_{(i,j) \in S} (\beta_{i,j} b_j) a_i = 0.$$

Since the a_i 's are K -linearly independent elements of M , we have that for all i , $\sum \beta_{i,j} b_j = 0$, and then similarly, since the b_j 's are linearly independent elements of K we have $\beta_{i,j} = 0$ for all j . \square

Remark 3.1.2: In general the degree $[L : K]$ of a field extension is a cardinal number, and the statement of Theorem 7 is to be interpreted as an identity of (possibly infinite) cardinals. On the other hand, when M/K and K/F are finite, the argument shows that M/F is finite and the result reduces to the usual product of positive integers. Moreover the finite case is the one that is most useful.

Let L/K be an extension of fields and $\alpha \in L$. We say that α is **algebraic** over K if there exists some polynomial $P(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0 \in K[t]$ such that $P(\alpha) = 0$. If α is not algebraic over K it is said to be **transcendental** over K . A complex number which is algebraic over \mathbb{Q} is called an **algebraic number**.

Examples 3.1.3: i is algebraic over \mathbb{R} since it satisfies the equation $i^2 + 1 = 0$. It is also algebraic over \mathbb{Q} for the same reason. Indeed for any $a \in \mathbb{Q}$, $a^{\frac{1}{n}}$ is algebraic over \mathbb{Q} . This is almost tautological, since by $a^{\frac{1}{n}}$, one generally means any complex number α such that $\alpha^n = a$, so α satisfies $t^n - a = 0$.

The following exercise gives less trivial examples.

Exercise 3.1.4: Let $\frac{a}{b}$ be any rational number. Show that $\cos(\frac{a}{b}\pi)$ and $\sin(\frac{a}{b}\pi)$ are algebraic.

Exercise 3.1.5:

- Show that the set of all algebraic numbers is countably infinite.
- More generally, let K be any infinite field and L/K be any field extension. Show that the cardinality of the set of elements of L which are algebraic over K is equal to the cardinality of K .

So “most” real or complex numbers are transcendental. This was observed by Cantor and stands as a famous early application of the dichotomy between countable and uncountable infinite sets. Earlier Liouville had constructed particular transcendental numbers, like $\sum_{n=1}^{\infty} 10^{-n!}$: an easy application of the Mean Value Theorem shows that a number which “too well approximated” by rational numbers cannot be algebraic. It is of course a different matter entirely to decide whether a particular, not obviously algebraic, number which is given to you is transcendental. Let us say only that both e and π were shown to be transcendental in the 19th century; that there have been some interesting results in transcendence theory in the 20th century (in particular e^{π} and $2^{\sqrt{2}}$ are transcendental); and that to this day the transcendence of a reasonable looking constant – e.g. $\zeta(3) = \sum_{n=1}^{\infty} \frac{1}{n^3}$ – is much beyond our reach.

The problem of determining whether particular numbers are transcendental, although certainly of interest, has little to do with modern field theory. (It is rather regarded as a branch of number theory, a part of the theory of Diophantine approximation.)

We now come back to the theory of field extensions by asking the reader to complete the following exercise:

Exercise 3.1.6 (Universal property of polynomial rings): Let $\iota : R \rightarrow S$ be a homomorphism of commutative rings, and let $\alpha_1, \dots, \alpha_n$ be elements of S . There is a unique R -algebra homomorphism $\Phi : R[t_1, \dots, t_n] \rightarrow S$ which takes $t_i \mapsto \alpha_i$.

Now let L/K be a field extension and $\alpha \in L$. By Exercise 3.1.6 there is a unique K -algebra homomorphism $\Phi : K[t] \rightarrow L$, $t \mapsto \alpha$. Let I be the kernel of Φ . Since $K[t]/I$ embeds in L , it is an integral domain, so I is a prime ideal. Since $K[t]$ is a principal ideal domain, there are only two choices:

Case 1: $I = 0$, i.e., Φ embeds $K[t]$ into L . This means precisely that α satisfies no polynomial relations with K -coefficients, so occurs iff α is transcendental over K .

Case 2: $I = (P(t))$ is generated by a single irreducible polynomial $P(t)$. Since the units of $K[t]$ are precisely the nonzero elements of K , it follows that there is a unique monic polynomial $P(t)$ (i.e., with leading coefficient 1) that generates I . We call this the **minimal polynomial** of α . Evidently for $Q \in K[t]$ we have $Q(\alpha) = 0 \iff P(t) \mid Q(t)$. In particular $P(\alpha) = 0$, so that α is algebraic, and moreover Φ induces an isomorphism $K[t]/(P(t)) \rightarrow L$. If P has degree d , then we say α is algebraic of degree d ; moreover, a K -basis for the left-hand side is $1, t, \dots, t^{d-1}$, so $[L : K] = d = \deg(P)$.

Let us summarize:

Theorem 8. *Let L/K be a field extension and $\alpha \in L$.*

a) TFAE:

(i) α is algebraic of degree d over K .

(ii) The K -vector space $K[\alpha]$ is finite, of degree d .

- (iii) The K -vector space $K(\alpha)$ is finite, of degree d .
 b) If α is algebraic of degree d , then $K[\alpha] = K(\alpha) \cong K[t]/(P(t))$, where $P(t) \in K[t]$ is the unique monic polynomial of degree d such that $P(\alpha) = 0$.
 c) If α is transcendental over K , then $K[t] \cong K[\alpha] \subsetneq K(\alpha) \cong K(t)$.

It follows that the set of all rational expressions $\frac{P(\pi)}{Q(\pi)}$ with $P, Q \in \mathbb{Q}[t]$ is isomorphic to the rational function field $\mathbb{Q}(t)$! In other words, there is no genuinely algebraic distinction to be made between “fields of numbers” and “fields of functions.”

A field extension L/K is **algebraic** if every $\alpha \in L$ is algebraic over K .

Corollary 9. *A finite extension L/K of fields is necessarily an algebraic extension.*

Proof. Indeed, the contrapositive form of this statement, that if there exists $\alpha \in L$ which is transcendental over K , then $[L : K]$ is infinite, follows from the relation $[L : K] = [L : K(\alpha)][K(\alpha) : K]$ and Theorem 8c). Specifically, $K(\alpha) \cong K(t)$, the field of rational functions, which contains (properly!) the infinite dimensional K -vector space $K[t]$ so is itself infinite dimensional. \square

The converse does not hold: many (in some sense, most) fields admit infinite algebraic extensions. A detailed analysis of algebraic field extensions is still ahead of us, but it is easy to see that the extension $\mathbb{Q}[\bigcup_{n \geq 2} 2^{\frac{1}{n}}]$ is an infinite algebraic extension, since it contains subextensions of arbitrarily large finite degree.

Exercise 3.1.7 (Direct limits) Let (I, \leq) be a directed set: recall that this means that I is partially ordered under \leq and for any $i, j \in I$ there exists $k \in I$ with $i \leq k$ and $j \leq k$. A **directed system of sets** is a family of sets $\{X_i\}_{i \in I}$ together with field homomorphisms $\iota(i, j) : X_i \rightarrow X_j$ for all $i \leq j$ satisfying the natural compatibility condition: for all $i \leq j \leq k$, $\iota(i, k) = \iota(j, k) \circ \iota(i, j)$. By definition, the **direct limit** $\lim_I X$ is the quotient of the disjoint union $\bigsqcup_{i \in I} X_i$ by the equivalence relation $(x, X_i) \sim (\iota(i, j)x, X_j)$ for all $i \leq j$.

a) Show that there are natural maps $\iota_i : X_i \rightarrow \lim_I X_i$. State and prove a universal mapping property for the direct limit.

b) Suppose that the maps $\iota(i, j)$ are all injective. Show that the maps $\iota_i : X_i \rightarrow \lim_I X_i$ are all injective. Explain why in this case $\lim_I X_i$ is often informally referred to as the “union” of the X_i ’s.

c) In any concrete category \mathcal{C} – i.e., a category whose objects are sets, for which the set of all morphisms from an object A to an object B is a subset of the set of all functions from A to B , and for which composition and identity of morphisms coincide with the usual notions of functions – one has the notion of a directed system $\{A_i\}$ of objects in \mathcal{C} , i.e., we have sets A_i indexed by the directed set (I, \leq) and for all $i \leq j$, the function $\iota(i, j) : A_i \rightarrow A_j$ is a morphism in \mathcal{C} . Give a definition of the direct limit $\lim_I A_i$ in this more general context. Show that the direct limit exists in the following categories: monoids, groups, commutative groups, rings, commutative rings, fields.

d) Give an example of a concrete category in which directed limits do not necessarily exist.⁶

e) Show that a field extension L/K is algebraic iff it is the direct limit of its finite subextensions.

⁶Suggestion: impose some finiteness condition on one of the above categories.

3.2. Some impossible constructions.

The results we have derived so far do not look very deep (or perhaps even very interesting) to modern eyes, but they were recognized in the 19th century to imply a negative solution to several of the longest standing open problems in mathematics. Namely, the Greeks were interested in **constructibility** of quantities using a compass and a straightedge. Without getting overly bogged down in the details, here are the basic ideas: one starts out with two distinct points in the plane, which we may as well view as being a unit distance apart. We have at our disposal an unmarked straightedge, so that given any two points we may construct the line passing through them, and a compass, such that given any previously constructed point P_1 and any previously constructed pair of points P_2, P_3 , we may draw a circle whose center is P_1 and whose radius is the distance between P_2 and P_3 (of course $P_1 = P_2$ or $P_1 = P_3$ is allowed). Let us say that a positive real number α is **constructible** if we can after a finite sequence of steps construct points P, P' whose distance between them is α (more precisely, α times the unit distance we started with), and let us agree that a negative number α is constructible iff $|\alpha|$ is constructible. If you have not actually spent time during your adulthood considering ruler and compass constructions, you might be surprised that the supply of constructible numbers is in some sense rather large:

- Exercise 3.2.2: a) Show that the constructible numbers form a subfield of \mathbb{R} .
 b) Show that if $\alpha > 0$ is constructible, then so is $\sqrt{\alpha}$.

However, there are some severe restrictions on constructible numbers: a constructible number is built up in a series of steps: $\alpha_1 = 1, \alpha_2, \dots, \alpha_n = \alpha$ corresponding to a tower of fields $F_1 = \mathbb{Q}, F_2 = F_1(\alpha_2), \dots, F_n = F_{n-1}(\alpha_n)$. To get from F_i to $F_{i+1} = F_i(\alpha_i)$, we are either intersecting two lines – which corresponds to solving a linear equation with coefficients in F_{i-1} , so $F_i = F_{i-1}$ – or intersecting a line defined over F_{n-1} with a circle whose coefficients lie in F_{i-1} which yields solutions in either F_{i-1} or a quadratic extension of F_{i-1} – or we are intersecting two circles with equations defined over F_{i-1} , which leads to solutions over at worst a quadratic extension of a quadratic extension of F_{i-1} . (Note quadratic, not quartic: any two distinct circles intersect in at most two points, and thus the common intersection can also be expressed as the intersection of a line and a circle.)

Thus any constructible number α lies in a field which is at the top of a tower of quadratic field extensions, and therefore $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ divides a power of 2 and is therefore itself a power of 2. The impossibility of three classically sought after constructions follows easily.

First we cannot “double the cube”, i.e., given a cube with sides of our unit length, we cannot construct a cube whose volume is twice that of the given cube, because the length of a side would be $\sqrt[3]{2}$, and $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$. Similarly we can construct angles that we cannot trisect; in particular, we can construct an angle of 60 degrees (i.e., we can construct $\cos 60^\circ = \frac{1}{2}$ and $\sin 60^\circ = \frac{\sqrt{3}}{2}$, but we cannot construct $\cos 20^\circ$ since it satisfies an irreducible cubic polynomial over \mathbb{Q} (exercise). Finally, still less could we hope to “square the circle”, i.e., construct a square whose

area is that of a unit circle, for that would involve constructing a side length of $\sqrt{\pi}$ and π is not even algebraic!

3.3. Subfields of algebraic numbers. Let L/K be an arbitrary extension of fields. Consider the set $\text{Cl}_L(K)$ of all elements of L which are algebraic over K . For example, when $K = \mathbb{Q}$, $L = \mathbb{C}$ we are examining the set of all algebraic numbers, which is certainly a proper subset of \mathbb{C} .

Proposition 10. *The set $\text{Cl}_L(K)$ is a subfield of K .*

We often refer to $\text{Cl}_L(K)$ as the **algebraic closure of K in L** .

I can see no reason not to prove this result in a more general context, that of integral extensions of integral domains. The generalized proof is not much harder and will be extremely useful for any student of algebra. So: let R be an integral domain and S a domain which extends R , i.e., there is an injective homomorphism $R \rightarrow S$. We say that $\alpha \in S$ is **integral over R** if α satisfies a monic polynomial with R -coefficients:

$$\exists a_{n-1}, \dots, a_0 \in R \mid \alpha^n + a_{n-1}\alpha^{n-1} + \dots + a_1\alpha + a_0 = 0.$$

We say that the extension S/R is **integral** if every element of S is integral over R .

Note that if R and S are fields, $\alpha \in S$ is integral over R is by definition precisely the same as being algebraic over R . The next result in fact revisits the basic finiteness property of algebraic elements in this slightly generalized context.

Lemma 11. *For $R \subset S$ an extension of domains and $\alpha \in R$, TFAE:*

- a) α is integral over R .
- b) The subring $R[\alpha]$ is finitely generated as an R -module.

Proposition 12. *Let $R \subset S$ be an integral extension. If R is a field, so is S .*

Proof: Let L be the fraction field of S . If $0 \neq \alpha \in S$ is integral over R , then by Lemma 11, $R[\alpha]$ is a finite-dimensional R -submodule of L , so it is a subfield, i.e., is equal to $R(\alpha)$. So $R(\alpha) = R[\alpha] \subset S$, meaning that S contains α^{-1} .

Theorem 13. *Let $R \subset T$ be integral domains and $\alpha \in T$. If there exists an intermediate ring $R \subset S \subset T$ such that $\alpha \in S$ and S is finitely generated as an R -module, then α is integral over R .*

Remark 3.3.1: If R were a Noetherian ring, we would have a very easy proof: S is finitely generated over R , hence the submodule $R[\alpha]$ is finitely generated over R , so α is integral over R . But with more care, one can prove the argument without the Noetherian hypothesis (and without any knowledge of modules over Noetherian rings). In particular fields are Noetherian rings – and a subspace of a finite dimensional vector space is finite dimensional! – so that the reader who is only interested in the case of field extensions can skip the following.

Proof. Let s_1, \dots, s_n be a finite set of generators for S over R , and express each of the elements $s_i\alpha$ in terms of these generators:

$$s_i\alpha = \sum_{j=1}^n r_{ij}s_j, \quad r_{ij} \in R.$$

Let M be the $n \times n$ matrix $\alpha I_n - (r_{ij})$; then recall from linear algebra that

$$MM^* = \det(M) \cdot I_n,$$

where M^* is the “adjugate” matrix (of cofactors). If $s = (s_1, \dots, s_n)$ (the row vector), then the above equation implies $0 = sM = sMM^* = s \det(M) \cdot I_n$. The latter matrix equation amounts to $s_i \det(M) = 0$ for all i . Since we can express 1 as an R -linear combination of the s_i 's, this gives $\det(M) = 0$, which means that α is a root of the monic polynomial $\det(T \cdot I_n - (a_{ij}))$. \square

Corollary 14. *Let S/R be an extension of integral domains. Then $\text{Cl}_S(R)$, the set of elements of S which are integral over R is a subring of S , called the **integral closure** of R in S .*

Proof. Let $\alpha, \beta \in S$ be integral over R . Since -1 is integral over R , it suffices to show that $\alpha + \beta$ and $\alpha\beta$ are integral over R . Consider the extension $R[\alpha] \subset R[\alpha, \beta] = R[\alpha][\beta]$. Since β is integral over R , it is integral over $R[\alpha]$, hence $R[\alpha][\beta]$ is finitely generated over $R[\alpha]$. Similarly $R[\alpha]$ is finitely generated over R , and it is easy to show that this means that $R[\alpha, \beta]$ is finitely generated as an R -module. Therefore $\alpha + \beta$ and $\alpha\beta$ lie in the intermediate ring $R[\alpha, \beta]$ which is finitely generated as an R -module, so by Theorem 13 $\alpha + \beta$ and $\alpha\beta$ are integral over R . \square

Applying Corollary 14 with R and S equal to fields gives Proposition 10.

4. ALGEBRAICALLY CLOSED FIELDS

A field K is **algebraically closed** if it admits no proper algebraic extension.

Exercise 4.1: Show that the following properties of a field K are equivalent:

- (i) K admits no proper algebraic extension.
- (ii) K admits no proper finite extension.
- (iii) Every irreducible polynomial $P(t) \in K[t]$ has degree 1.
- (iv) For any nonconstant $P(t) \in K[t]$, there exists $\alpha \in K$ such that $P(\alpha) = 0$.

Thus any of these properties could be taken as the definition of algebraic closure.

Theorem 15. *(Fundamental Theorem of Algebra)*

The complex field \mathbb{C} is algebraically closed.

Because the existence of a nonconstant complex polynomial without a complex root leads to absurdities in many areas of mathematics, there are many different proofs, e.g. using degree theory or complex analysis. It is often held that “fundamental theorem of algebra” is a misnomer, in that the result concerns a structure – the complex numbers – whose definition is in part analytic/topological. We do not dispute this. Nevertheless the true algebraist hankers for an algebraic proof, and indeed this is possible. We may, in fact, view Theorem 15 as a special case of the following result.

Theorem 16. *(Artin-Schreier) Suppose K is a field with the following properties:*

- (i) *There do not exist $n \in \mathbb{Z}^+$ and $x_1, \dots, x_n \in K$ such that $-1 = x_1^2 + \dots + x_n^2$.*
- (ii) *Every polynomial $P \in K[t]$ of odd degree has a root in K .*
- (iii) *For any $x \in K^\times$, exactly one of x and $-x$ is a square in K .*

Then $K[\sqrt{-1}] = K[t]/(x^2 + 1)$ is algebraically closed.

Theorem 16 \implies Theorem 15: Take $K = \mathbb{R}$. Hypothesis (i) holds in \mathbb{R} (as in every orderable field) and (ii) and (iii) are easy consequences of the Intermediate Value Theorem. Thus we deduce that $\mathbb{C} = \mathbb{R}[\sqrt{-1}]$ is algebraically closed.

We must defer the proof until later: it requires Galois theory (and also a bit of group theory).

Proposition 17. *Let L/K be a field extension, with L algebraically closed. Then $\overline{K} = \text{Cl}_K(L)$ is algebraically closed.*

Proof. Suppose not: then there exists a proper finite extension M of \overline{K} . Obviously any finite extension is finitely generated: there exist $\alpha_1, \dots, \alpha_n \in M$ such that $M = \overline{K}[\alpha_1, \dots, \alpha_n]$. Each α_i satisfies a polynomial with coefficients a_{ij} in \overline{K} , i.e., which are algebraic over K . Therefore α_i is finite over the extension $K[\{a_{ij}\}]$ which is itself finite over K , so α_i is finite over K . Therefore M is finite over K , so every element of M is algebraic over K , i.e., $M = \text{Cl}_K(L)$. \square

Corollary 18. *The field $\overline{\mathbb{Q}}$ of all algebraic numbers is algebraically closed.*

This follows immediately, since $\overline{\mathbb{Q}} = \text{Cl}_{\mathbb{Q}}(\mathbb{C})$.

If K is a field, then a field \overline{K} which is on the one hand algebraically closed and on the other hand an algebraic extension of K is called an **algebraic closure** of K . The content of Corollary 18 is that to find an algebraic closure of a field K it suffices to embed it into some algebraically closed field.

Following S. Lang,⁷ we will show that such an embedding can always be found.

First:

Lemma 19. *Let F be any field, and $P_1, \dots, P_n \in F[t]$ any finite set of nonconstant polynomials with coefficients in F . Then there exists a finite field extension K/F such that each polynomial $P_i(t)$ has a root in K .*

Proof. It suffices to show the Lemma with $n = 1$. Indeed, assuming that this is true, we can find a finite field extension F_1/F such that P_1 has a root in F_1 , then a finite field extension F_2/F_1 in which P_2 has a root, and so forth. At the end we get a finite tower $F \subset F_1 \subset \dots \subset F_n$ of finite field extensions, so F_n/F is a finite field extension in which all P_i 's have roots.

The $n = 1$ case is a triumphant example of the beautiful tautologies that exist in abstract algebra. Namely, pick an irreducible factor Q of P_1 and consider the ring $R = F[t]/(Q)$. Since Q is irreducible, (Q) is a maximal ideal, and hence R/F is a field extension, of finite degree $\deg(Q)$. Moreover in R the class of t modulo Q is a root of Q , since $Q(t) \equiv 0 \pmod{(Q(t))}$! \square

Theorem 20. *Every field K can be embedded in an algebraically closed field L . Thus every field has at least one algebraic closure, namely $\text{Cl}_L(K)$.*

Proof. We first claim that it suffices to find a field K_1 such that every nonconstant polynomial with coefficients in K has roots in K_1 . If so, then applying the claim

⁷Lang himself was following E. Artin.

to K_1 we get an extension field K_2 in which each polynomial with coefficients in K_1 has a root in K_2 , and so forth: we generate a sequence of field extensions

$$K \subset K_1 \subset \dots \subset K_n \subset \dots$$

Then the union $L = \bigcup_n K_n$ is itself a field, and any nonconstant polynomial $P \in L[t]$, having only finitely many coefficients, actually lies in $K_n[t]$ for sufficiently large n , and thus has a root in K_{n+1} and therefore also in L .

To prove the claim, we construct a large polynomial ring $R = K[\mathbb{T}]$ over K ; more precisely, the indeterminates t_f will correspond bijectively to the nonconstant polynomials $f \in K[t]$. Consider the ideal I of R generated by all polynomials of the form $f(t_f)$. First we must check that I is a proper ideal: if not, there exists a finite subset $\{f_1, \dots, f_n\}$ and elements $g_1, \dots, g_n \in R$ such that

$$g_1 f_1(t_{f_1}) + \dots + g_n f_n(t_{f_n}) = 1.$$

Let F/K be a finite extension such that each polynomial $f_i(t)$ has a root $\alpha_i \in F$. If we then consider the equation in the ring $F[\mathbb{T}]/(t_{f_i} - \alpha_i)$ we get $0 = 1$, a contradiction. Thus we may choose a maximal ideal $\mathfrak{m} \supset I$, and consider the field $K_1 := R/\mathfrak{m}$. In this field the image of t_f is a root of f , so that indeed K_1 is an extension of K in which each nonconstant polynomial in K has a root. \square

Remark 4.2: Note that we used the fact that every proper ideal in a commutative ring is contained in a maximal ideal. This fact is on the one hand quite easy to prove: one considers the set of all proper ideals J containing a given proper ideal I : the union of every chain of proper ideals containing I is a proper (since 1 is not in any of the ideals in the chain, it is not in the union) ideal containing I , so by Zorn's Lemma we are entitled to a maximal such ideal. However, by using Zorn's Lemma we are invoking the Axiom of Choice. It can be shown that this invocation is absolutely necessary: the fact that every field can be embedded in an algebraically closed field implies the Axiom of Choice.

Later we will study algebraic closures of a field F in more detail. In particular we will see that they are essentially unique in that any two algebraic closures of F are isomorphic as F -algebras. However, in general they are far from being *canonically* isomorphic: the ambiguity in the choice of automorphism corresponds precisely to the group of F -algebra automorphisms of an algebraic closure \bar{F} . In most cases, this automorphism group is extremely large and complicated.

Exercise 4.3: Let K be a field and $P \in K[t]$ be a polynomial of degree $n \geq 1$. Let \bar{K} be an algebraic closure of K , so that over \bar{K} , P factors into n linear factors $\prod_{i=1}^n (t - x_i)$. One says that P is **separable** if $x_i \neq x_j$ for all $i \neq j$.⁸

a) Show that P is separable iff $\gcd(P, P') = 1$. (Here P' is the formal derivative of P , i.e., the unique linear map from $K[t]$ to itself such that $(t^n)' = nt^{n-1}$.)

b) Let K be a field and $n \in \mathbb{Z}^+$. If K has positive characteristic p , assume that $\gcd(n, p) = 1$. Let $a \in K$ be arbitrary. Show that the polynomial $t^n - a$ is separable. Deduce that there is a finite field extension L/K containing n elements z_1, \dots, z_n such that $z_i^n = 1$ for all i .

c) Show that no finite field is algebraically closed.

d) Show that if K is a field and \bar{K} is an algebraic closure, then $\#\bar{K} = \max(\aleph_0, \#K)$.

⁸Much more on this later, of course.

e) Show that there are algebraically closed fields of all infinite cardinalities.

Exercise 4.4: Let $\iota : \overline{\mathbb{Q}} \rightarrow \overline{\mathbb{Q}}$ be a field embedding. Show that ι is an isomorphism (i.e., is surjective). (Suggestion: viewing $\overline{\mathbb{Q}}$ as the direct limit of its finite subextensions may be helpful here.)

5. SPLITTING FIELDS

Our next order of business is to develop the theory of splitting fields. On the one hand, this is naturally viewed as a generalization of the notion of algebraic closure. On the other hand, our previous construction of an algebraically closed field containing any given field gives the splitting fields a place to live, and thus makes the basic theory rather straightforward to establish.

5.1. “Uniqueness” of the algebraic closure.

We have seen two fundamental facts: first, that any field K can be embedded in an algebraically closed field L ; and second, that the set $\text{Cl}_L(K)$ of all elements of L which are algebraic over K is an algebraically closed, algebraic field extension of K .

Notice, however, that we are not yet entitled to call $\text{Cl}_L(K)$ “the algebraic closure” of K , the reason being that the construction depended upon a choice of algebraically closed field L/K . Since any field is contained in algebraically closed fields of all sufficiently large cardinalities, we have a proper class’s worth of ambiguity.

To be a little more concrete, consider the case of the rational field \mathbb{Q} . We have met two different algebraically closed, algebraic field extensions of \mathbb{Q} . The first is the field of algebraic numbers, i.e., the algebraic closure of \mathbb{Q} in the algebraically closed field \mathbb{C} . The second is the algebraic closure of \mathbb{Q} in the algebraically closed field L constructed in the proof of Theorem 20. A moment’s thought shows that L is a countable field, so that certainly L is not isomorphic to the complex numbers.

In some naive sense we want to ask: are the fields $\text{Cl}_L(\mathbb{Q})$ and $\text{Cl}_{\mathbb{C}}(\mathbb{Q})$ the same? But first we must decide what we mean by “the same”! If we mean “equal” then depending upon one’s perspective the answer is either negative or the question is meaningless. To get a flavor of the pitfalls of asking whether two objects which are not given to you as elements of a common set are equal, consider: is the underlying set of the complex field \mathbb{C} “equal” to the set of ordered pairs $(a, b) \in \mathbb{R}^2$? Or is it “equal” to the set of cosets of $t^2 + 1$ in the polynomial ring $\mathbb{R}[t]$? (Or is it “equal” to a certain subring of the 2×2 real matrices? Or . . .) It’s an essentially meaningless question.⁹ What we can say of course is first that these fields are isomorphic, and second that we have a particular isomorphism in mind: e.g. $(a, b) \mapsto a + bt + (\text{mod } t^2 + 1)$. In other words, all our usual descriptions of the complex field are the same in that they are *canonically* isomorphic.

It turns out that two algebraic closures K_1, K_2 of a field F satisfy a condition

⁹Better put: it’s a difficult philosophical problem whether the question is meaningful, but it is a problem with no mathematical content.

which is a bit better than being isomorphic as fields but distinctly worse than being canonically isomorphic. Namely:

Theorem 21. (*“Uniqueness” of the algebraic closure*) *Let F be a field and K_1, K_2 be algebraically closed, algebraic extensions of F . Then there exists an F -algebra isomorphism $\Phi : K_1 \rightarrow K_2$, i.e., an isomorphism of fields such that $\Phi(x) = x$ for all $x \in F$. In general, Zorn’s Lemma is needed to construct Φ .*

Proof. Consider the partially ordered set whose elements are triples (L_1, L_2, φ) , where: $F \subset L_i \subset K_i$ are intermediate fields, and $\varphi : L_1 \rightarrow L_2$ is an F -algebra isomorphism. We say that $(L_1, L_2, \varphi) \leq (L'_1, L'_2, \varphi')$ if $L_i \subset L'_i$ and φ' restricted to L_1 is φ . It is immediate that any chain in this poset has an upper bound: just take the union of the elements of the chain. Therefore by Zorn’s Lemma we are entitled to a maximal such element $\varphi : L_1 \rightarrow L_2$. It remains to show that $L_i = K_i$. In fact it is enough to show that $L_1 = K_1$, because then $L_2 = \varphi(K_1)$ is an algebraically closed subfield of K_2 containing F , and the only such field is K_2 itself. So suppose that there exists $\alpha \in K_1 \setminus L_1$, and let $P \in F[t]$ be the minimal polynomial of α . Because K_2 is algebraically closed, there exists some element $\alpha' \in K_2$ such that $P(\alpha') = 0$. We claim that there is a unique isomorphism $\psi : L_1[\alpha] \rightarrow L_2[\alpha']$ determined by $\psi(x) = \varphi(x)$ for $x \in L_1$ and $\psi(\alpha) = \alpha'$. Indeed, there is a unique map $L_1[t] \rightarrow L_2[\alpha']$ taking $x \in L_1 \mapsto \varphi(x)$, $t \mapsto \alpha'$, whose kernel is the set of all $P = \sum_i a_i t^i \in L_1[t]$ such that

$$0 = \psi\left(\sum_i a_i t^i\right) = \sum_i \varphi(a_i)(\alpha')^i = \varphi(Q)(\alpha'),$$

where $\varphi(Q) \in L_2[t]$ is $\sum_i \varphi(a_i) t^i$. But since the minimal polynomial of α' over K is $P(t)$ – the minimal polynomial of α , this holds iff $P(t) \mid \varphi(Q)$ iff $P(t) \mid Q$. Thus the kernel of ψ is precisely the ideal generated by $P(t)$, so ψ descends to give an isomorphism

$$\psi : L_1[t]/(P(t)) \rightarrow L_2[\alpha'].$$

Since $L_1[t]/(P(t)) \cong L_1[\alpha]$ (as an L_1 -algebra, hence certainly as an F -algebra), we have constructed an isomorphism $L_1[\alpha] \rightarrow L_2[\alpha']$, contradicting the maximality of (L_1, L_2, φ) . \square

Remark 5.1.1: Certainly the use of Zorn’s Lemma in the proof looks “serious.” This is of course not the same as proving that Zorn’s Lemma is needed. The latter is an assertion of axiomatic set theory: there is a model of the Zermelo-Frankel axioms without the Axiom of Choice in which some field admits nonisomorphic algebraic closures. We are certainly not going to prove that here: see e.g. XXX.

But morally the lack of canonicity in the isomorphism between algebraic closures is expressing the fact that there are, in general, many different isomorphisms to choose: the set of F -algebra isomorphisms from one algebraic closure to another is in bijection with the group of F -algebra automorphisms of a given algebraic closure \overline{F} . As we shall see later, this automorphism group is typically a large and interesting object: in fact it is either trivial (which in characteristic 0 can only happen if F is already algebraically closed), or is the group of order two (e.g. $F = \mathbb{R}$), or it is an uncountably infinite group!

The proof of Theorem 21 easily extends to give the following generalization:

Corollary 22. *Let K_1/F and K_2/F be two algebraic field extensions. If $\varphi : K_1 \rightarrow K_2$ is any F -algebra embedding and $\overline{K_i}$ is any algebraic closure of K_i , then φ extends to an isomorphism $\overline{K_1} \rightarrow \overline{K_2}$. (Again, this requires an appeal to Zorn's Lemma.)*

Exercise 5.1.2: Prove Corollary 22.

Corollary 23. *Let K/F be an algebraic field extension and L/F a field extension with L algebraically closed (but possibly transcendental over F). Then there is an F -algebra embedding $K \hookrightarrow L$.*

Exercise 5.1.3: Prove Corollary 23.

Thus every algebraically closed field is big enough to contain at least one isomorphic copy of every algebraic extension of one of its subfields. However, in general there may be multiple copies, and that is a crucial complication in the theory. For example, let $K = \mathbb{Q}[\sqrt[3]{2}] \subset \mathbb{C}$. As a field (or equivalently, as a \mathbb{Q} -algebra) K is isomorphic to $\mathbb{Q}[t]/(t^3 - 2)$, but the irreducible polynomial $t^3 - 2$ has two other complex roots: $\zeta_3 \sqrt[3]{2}$ and $\zeta_3^2 \sqrt[3]{2}$, where $\zeta_3 = e^{2\pi i/3}$ is a primitive cube root of unity. Clearly these other two roots, being nonreal complex numbers, do not lie in the real field K . This gives rise to two further embeddings from the “abstract cubic field” $\mathbb{Q}[t]/(t^3 - 2)$ into the complex numbers: namely we can take $t \mapsto \zeta_3 \sqrt[3]{2}$ and get the subfield $\mathbb{Q}[\zeta_3 \sqrt[3]{2}]$ or $t \mapsto \zeta_3^2 \sqrt[3]{2}$ and get the subfield $\mathbb{Q}[\zeta_3^2 \sqrt[3]{2}]$. Of course, these three subfields lie in $\overline{\mathbb{Q}}$, the field of algebraic numbers. On the other hand, the subfield generated by any two of these fields is $K = \mathbb{Q}[\zeta_3, \sqrt[3]{2}]$, a field extension of \mathbb{Q} of degree 6. This field K has nicer properties: any subfield $L \subset \overline{\mathbb{Q}}$ which is not equal to K is not isomorphic to K . By Corollary 23 every isomorphism $L \rightarrow K$ extends to a field endomorphism σ of $\overline{\mathbb{Q}}$, which by Exercise 4.4 is necessarily an automorphism. Hence the claim is equivalent to the assertion that the subfield K is **stable** under all automorphisms $\sigma : \overline{\mathbb{Q}} \rightarrow \overline{\mathbb{Q}}$: i.e., $\sigma(K) = K$.

Compare with the following concept in group theory: a subgroup H of a group G is called **characteristic** if for every automorphism $\sigma : G \rightarrow G$ we have $\sigma(H) = H$. The idea is that such a subgroup H can be “characterized” by properties of G which are invariant under automorphism. Examples: the minimal subgroup of G such that G/H is abelian (a.k.a. the commutator subgroup); the intersection of all maximal subgroups of G (the Frattini subgroup); any subgroup whose order is different from any other subgroup (e.g. A_n in S_n).

5.2. Splitting fields.

What we showed in the previous section was that an algebraic closure \overline{F} of F is characterized – up to F -algebra isomorphism – as being an algebraic field extension of F in which every nonconstant polynomial $P(t) \in F[t]$ has a root. (A priori an algebraic closure \overline{F} must have the stronger property that every nonconstant polynomial with coefficients in \overline{F} has a root, but we saw above that, since algebraic extensions of algebraic extensions are themselves algebraic, it is enough to require that every nonconstant polynomial with coefficients in F has a root.)

It is equally well the case that \overline{F} is characterized (up to...) as the unique algebraic extension of F in which each nonconstant polynomial $P \in F[t]$ **splits** into

linear factors: indeed, if $F(t) = (t - \alpha_1)G(t)$ for $\alpha_1 \in \overline{F}$, then $G(t)$ is a polynomial with coefficients in \overline{F} and as above therefore itself has a root, and so forth. This allows us to generalize the notion of algebraic closure, as follows:

Definition: Let F be a field and $S = \{P_i(t)\}$ be a set of nonconstant polynomials with F -coefficients. A **splitting field** for (F, S) is a field extension K/F satisfying the following properties:

- (SF1) Every $P_i \in S$ splits into linear factors in K .
- (SF2) No proper subextension of K satisfies (SF1), i.e., if $F \subset K' \subset K$ and every $P_i \in S$ splits in K' , then $K' = K$.

Exercise 5.2.1: Suppose K/F is a splitting field for (F, S) , and K' is an F -algebra isomorphic to K . Show that K' is also a splitting field for (F, S) .

Theorem 24. (*Existence and “uniqueness” of splitting fields*) Let F be a field and $S \subset F[t]$ a set of nonconstant polynomials.

- a) Any algebraic closure \overline{F} contains a unique splitting field for S , namely the subfield of \overline{F} obtained by adjoining to F all roots α_{ij} of all polynomials $P_i \in S$.
- b) Splitting fields are unique up to F -algebra isomorphism.

Proof. It is no problem to see that the recipe of part a) does indeed construct a splitting field for F and S : clearly every polynomial in S splits in $F(\alpha_{ij})$ and conversely any subfield of \overline{F} in which all the polynomials in F split must contain all the α_{ij} 's. One way to see the uniqueness up to isomorphism is to reduce to the case of uniqueness up to isomorphism of algebraic closures. Namely, let K_1, K_2 be two splitting fields for F and S . It is easy to see that (SF2) implies that K_i/F is algebraic, so let \overline{K}_i be an algebraic closure of K_i . Since K_i is algebraic over F , \overline{K}_i is equally well an algebraic closure of F , so by Theorem XX there exists an F -algebra isomorphism $\Phi : \overline{K}_1 \rightarrow \overline{K}_2$. Then $\Phi(K_1)$ is a subfield of \overline{K}_2 which is a splitting field for F and S , and we just saw that each algebraic closure contains a unique splitting field, so $\Phi(K_1) = K_2$ and $\Phi : K_1 \rightarrow K_2$ is an F -algebra isomorphism. \square

Corollary 25. Let F be a field and K/F a splitting field (for some subset $S \subset F[t]$). Let L/K be any extension and $\sigma : L \rightarrow L$ be any automorphism. Then $\sigma(K) = K$.

Proof. Let \overline{L} be an algebraic closure of L . By Corollary XX, we may extend σ to an automorphism $\sigma : \overline{L} \rightarrow \overline{L}$. Let $\overline{K} = \text{Cl}_{\overline{L}}(K)$, the unique algebraic closure of K contained in \overline{L} . Since K/F is algebraic, \overline{K} is also an algebraic closure of F . Since $\sigma(\overline{K})$ is also an algebraic closure of K contained in \overline{L} , by the aforementioned uniqueness we have $\sigma(\overline{K}) = \overline{K}$. Now by Exercise X.X, $\sigma(K)$ is a subfield of \overline{K} which is a splitting field for (F, S) , so by Theorem 24 we conclude $\sigma(K) = K$. \square

Exercise 5.2.2: Show that the field $K = \mathbb{Q}[\zeta_3, \sqrt[3]{2}]$ discussed in §5.1 is the splitting field of $F, \{t^3 - 2\}$. Conclude that if $L \subset \mathbb{C}$ is such that $L \neq K$, then L is not isomorphic to K .

5.3. Stable extensions; normal extensions.

Thus we have shown that any splitting field has strong stability properties under automorphisms. In general, if M/K is a field extension, a subextension L (i.e.,

$K \subset L \subset M$) is said to be M/K -**stable** if for every K -algebra automorphism σ of M we have $\sigma(K) \subset K$. Note that this condition applied to σ^{-1} gives $\sigma^{-1}(K) \subset K$ and then $\sigma(\sigma^{-1}(K)) = K \subset \sigma(K)$, so we may equally well define stability by $\sigma(K) = K$ for all F -automorphisms σ of M . A subfield L of M is called a **stable subfield** if it is stable under all field automorphisms of M (i.e., it is M/f -stable, where f is the prime subfield). An F -algebra K is **universally stable** if whenever $F \subset K \subset M$ we have that K is M/F -stable.

Examples 5.3.1: The Corollary shows that if K/F is a splitting field then it is universally stable. Similarly, a splitting field K/\mathbb{Q} is a stable subfield of \mathbb{C} . As it happens, the converse is true:

Theorem 26. *Let F be a field and M/F an extension, with M algebraically closed. Then a subextension L is M/F -stable iff L/F is a splitting field: in particular L must be algebraic. In particular, $L \subset M$ is a stable subfield iff L is a splitting field over the prime subfield.*

At the moment we will concentrate on the case of algebraic field extensions K/F . Thus we must show:

Theorem 27. *Let K/F be algebraic and \overline{K} an algebraic closure of K . TFAE:*

- a) *For every F -embedding $\sigma : K \rightarrow \overline{K}$ we have $\sigma(K) = K$.*
- b) *K/F is the splitting field of a subset $S \subset F[t]$.*
- c) *If $P \in F[t]$ is irreducible and has a root in K , it splits completely in K .*

Proof. c) \implies b): Let S be the set of irreducible polynomials which have at least one root in K . Then by the hypothesis of c), K is the splitting field for (F, S) .

b) \implies a) is Corollary 25. a) \implies c): Suppose $P(t)$ is irreducible over F and there exist $\alpha, \alpha' \in \overline{K}$, $P(\alpha) = P(\alpha') = 0$, and such that α lies in K but α' does not. Let $F' = F(\alpha)$. Thus F' is an F -subalgebra of K , and as an abstract F -algebra is isomorphic to $F[t]/(P(t))$. There is therefore a unique F -algebra embedding $\sigma : F' \hookrightarrow \overline{K}$ that maps $\alpha \mapsto \alpha'$. By Corollary XX we may extend σ_2 to an isomorphism $\overline{K} \rightarrow \overline{K}$ and then restrict it to K to get an embedding $\sigma : K \hookrightarrow \overline{K}$. Since $\alpha' \in \sigma(F') \subset \sigma(K)$, $\sigma(K) \neq K$. \square

Definition: An arbitrary field extension K/F is called **normal** if for every irreducible polynomial $P \in F[t]$, if P has a root in K , it splits completely in K . The term “normal” is often applied only to algebraic field extensions; by Theorem 27, an algebraic field extension is normal iff it is universally stable iff it is a splitting field of some set of polynomials.

Trivial Examples 5.3.2: An extension K/F is normal if: $K = F$ (obviously); $K = \overline{K}$ is algebraically closed (every polynomial in F splits in K); if $\text{Cl}_K(F) = F$ (no irreducible polynomial $P \in F[t]$ has a root in K).

Exercise 5.3.3: The extension \mathbb{R}/\mathbb{Q} is not normal.

Example 5.3.4: Any quadratic extension K/F is normal. Indeed, necessarily $K = F[t]/(P(t))$ for some quadratic polynomial P , and $P(t)$, being a quadratic polynomial in K which has at least one root, clearly must split. And obviously L is minimal with this property. Note that this works even in characteristic 2, when the

roots need not be distinct: e.g. if $K = \mathbb{F}_2(x)$ the polynomial $t^2 - x$ splits over K as $(t - \sqrt{x})^2$. In this case there is only one F -embedding of K into \overline{K} : we must take $\alpha \mapsto \alpha$.

Example 5.3.5: Let $n > 2$ be an integer. The extension $K = \mathbb{Q}[\sqrt[n]{2}]/\mathbb{Q}$ is a non-normal degree extension. Indeed, let $\zeta_n = e^{2\pi i/n}$; then the other roots of $t^n - 2$ in \mathbb{C} are $\zeta_n^k \cdot \sqrt[n]{2}$ with $0 \leq k < n$, which are not even real numbers unless $k = 0$ or $k = \frac{n}{2}$. So $t^n - 2$ does not split over K . In this case, any extension of K which is normal over \mathbb{Q} must contain all the roots of $t^n - 2$, hence must contain $\sqrt[n]{2}$ and ζ_n . Therefore the smallest normal extension is the splitting field of $t^n - 2$, which is $M = \mathbb{Q}[\sqrt[n]{2}, \zeta_n]$.

Example 5.3.6: If F has characteristic $p > 0$ and $P(t) = t^{p^a} - a$ is irreducible over F , then $F[t]/(P(t))$ is normal, because it is the splitting field for $P(t)$.

Exercise 5.3.7: If K/F is an extension, then K/F is normal iff $\text{Cl}_K(F)/F$ is normal.

Thus normality is really only an interesting concept for algebraic extensions, and it is often defined only in this restricted context. (However, it seemed interesting to place it in the larger context of stable extensions.)

Exercise 5.3.8: Show that a direct limit of normal extensions is normal.

Exercise 5.3.9: Show that, “unfortunately”, if K/F is normal and L/K is normal, then L/F need not be normal. (Normality does not satisfy the **tower property**.)

Exercise 5.3.10: a) Let L/F be an extension and K_1, K_2 be subextensions. Show that $K_1 \cap K_2$, the set of all elements common to both K_1 and K_2 , is again an extension field of F .

b) As above, but with any collection of intermediate field extensions $\{K_i\}_{i \in I}$.

Proposition 28. *Let L/F be an extension, and $\{K_i\}_{i \in I}/F$ be a collection of subextensions. If each K_i/F is normal, then so is the intersection $K = \bigcap_i K_i$.*

Proof. Using Exercise 5.3.7, we may replace K_i by $\text{Cl}_{K_i}(F)$ and L by $\text{Cl}_L(F)$ and thus assume that all field extensions are algebraic. Let $P \in F[t]$ be an irreducible polynomial. If P has a root in K , P has a root in each K_i , hence K_i contains a splitting field for (F, P) . Splitting fields are unique inside any given algebraic extension, so this means that each K_i contains the common splitting field for (F, P) , hence K contains it, so P splits in K . \square

Let K/F be any field extension. As above, \overline{K}/F is certainly normal. Since the intersection of any family of normal subextensions of \overline{K} is normal, it follows that there is a unique smallest subextension L , $F \subset K \subset L \subset \overline{K}$, such that L/F is normal. If we define a **normal closure** of an extension K/F to be an extension L/K which is normal over F and such that no proper subextension is normal over F , then we just constructed a normal closure, by intersecting all normal subextensions inside an algebraic closure of K . This shows that any normal closure of K/F is algebraic over K , and by the usual tricks with uniqueness up to F -isomorphism of algebraic closure one can show that the normal closure of an extension is also unique up to F -isomorphism.

Proposition 29. *Let K/F be finite of degree n . Then the degree of the normal closure M of K/F (inside any algebraic closure \overline{K}) is at most $n!$.*

Proof. In case $K = F[t]/(P(t))$ for an irreducible polynomial P of degree n , the argument is easy: the normal closure of K is just the splitting field for P . The polynomial P has at least one root in K , so it can be factored as a linear polynomial times $Q(t)$ a polynomial of degree $n - 1$. If $Q(t)$ already splits into distinct linear factors, we're done: otherwise there is an irreducible factor of degree at most $n - 1$, and by applying the same argument we make a field extension of degree at most $n - 1$ and get another root. The result follows by induction.

In general, although we have not yet ventured an example, it is possible for K/F to be finite yet not to be monogenic as a ring over F . But in any case there exists $\alpha \in K \setminus F$; let P be the minimal polynomial for α over F ; then $F \subset F[t]/(P(t)) = K' \subset K$. The polynomial P has a root in K so in order to form the normal closure we must ensure that it splits completely: by the above this can be done by a field extension L'/K' of degree at most $d! = [K' : F]!$. Then the field extension $L'K/F$ has degree at most $n \cdot (d - 1)!$ (FIX ME!!) \square

6. SEPARABLE ALGEBRAIC EXTENSIONS

Let K/F be an algebraic field extension. We have already explored one desirable property for K/F to have: normality. Normality can be expressed in terms of stability under F -homomorphisms into any extension field, and also in terms of irreducible polynomials: every irreducible polynomial in $F[t]$ with a root in $K[t]$ must split. There is another desirable property of an algebraic extension L/K called **separability**. In some sense it is dual to normality, but this is hard to believe at first because there is a large class of fields F for which all algebraic extensions K/F are separable, including all fields of characteristic 0. (For that matter, there are fields for which every algebraic extension is normal, like \mathbb{R} and \mathbb{F}_p .) Like normality, separability can also be expressed in terms of polynomials and also in terms of embedding conditions. We begin with a study of polynomials.

6.1. Separable polynomials.

A nonconstant polynomial $P \in F[t]$ is **separable** if over an algebraic closure \overline{F} , $P(t)$ splits into *distinct* linear factors. Equivalently, if P has degree n , then there are n distinct elements $\alpha_1, \dots, \alpha_n \in \overline{F}$ such that $P(\alpha_i) = 0$ for all i . Note that both of these conditions are easily seen to be independent of the chosen algebraic closure.

Exercise: let F be a field and K/F be any extension. Show that a polynomial $P \in F[t]$ is separable as a polynomial over F iff it is separable when viewed as a polynomial over K .

Exercise X.X: Let F be a field of characteristic $p > 0$ and $\alpha \in F^\times$. Show that the polynomial $t^p - \alpha$ is irreducible iff $\alpha \neq \beta^p$ for any $\beta \in F$.

Over any field F it is no trouble to come up with a polynomial that is not separable: t^2 . What is of more interest is whether there is an inseparable irreducible polynomial in $F[t]$. Note that some authors define a polynomial to be separable if all its irreducible factors are separable and others only discuss in/separability for

irreducible polynomials. Although these conventions certainly “work” as well, I find the current definition to be more convenient and more thematic. First, Exercise XX shows that with this definition, separability is faithfully preserved by base extension. Since the way one will check whether an irreducible polynomial is separable is by considering it over the algebraic closure, where of course it is a product of separable (linear!) polynomials, our definition seems simpler. Moreover, in the theory of algebras one does meet reducible polynomials: for any nonconstant $P \in F[x]$, we may consider the finite-dimensional F -algebra $A_P = F[x]/(P(x))$. Then our definition makes it true that P is separable iff A_P is a **separable algebra**, i.e., an algebra which is semisimple and remains semisimple after arbitrary base change.

In general it is far from obvious whether the field extension obtained by adjoining a root of an irreducible polynomial is normal. Fortunately, it is much easier to determine whether a polynomial, especially an irreducible polynomial, is separable.

Proposition 30. (*Derivative criterion*) Let $P \in F[t]$ be a nonconstant polynomial.

- a) P is separable iff $\gcd(P, P') = 1$, where P' is the formal derivative of P .
- b) If P is irreducible, it is separable iff $P'(t) \neq 0$.
- c) An irreducible polynomial is always separable in characteristic 0. In characteristic $p > 0$, an irreducible polynomial is inseparable iff there exists $Q \in F[t]$ such that $P = Q(t^p)$.

Proof: a) Using Exercise X.X, we may check the separability of P over an algebraic closure \bar{F} and thus assume that $P(t) = c \prod_{i=1}^n (t - \alpha_i)$. A simple calculus exercise shows that there exist $i \neq j$ such that $\alpha_i = \alpha_j \iff P'(\alpha_i) = 0$, i.e., iff P and $P'(t)$ have some common linear factor $(t - \alpha_i)$.

b) If $P(t)$ is irreducible, then the ideal $(P(t))$ is maximal in $F[t]$. Therefore $\gcd(P, P') = 1 \iff F[t] = (P, P')$ iff P' is not a multiple of $P(t)$. But since $\deg(P') < \deg(P)$ the only way that P could divide P' is if $P' = 0$. But if P is nonconstant, how could P' be zero? If

$$P(t) = \sum_{i=0}^n a_i t^i, \quad a_n \neq 0$$

then

$$P'(t) = \sum_{i=1}^n i a_i t^{i-1}.$$

The coefficient of t^{n-1} is $n a_n$; since $a_n \neq 0$, this coefficient – and hence P' itself – is nonzero unless F has positive characteristic $p \mid n$. In that case, $P'(t) = 0$ iff $a_i = 0$ for all i which are not divisible by p , i.e., if every nonzero monomial term is of the form $a_i t^{p^i}$, in which case $P = Q(t^p)$ where $Q = \sum_i a_i t^i$.

Lemma 31. Let F be a field of characteristic $p > 0$, $a \in \mathbb{Z}^+$, and $\alpha \in F^\times$. TFAE:

- (i) There exists $\beta \in F$ such that $\beta^p = \alpha$.
- (ii) The polynomial $P(t) = t^{p^a} - \alpha$ is reducible over F .

Proof: Because the polynomial $t^p - \alpha$ is inseparable, it has a unique root in an algebraic closure \bar{F} , namely an element β such that $\beta^p = \alpha$. We must show that the reducibility of $P(t)$ is equivalent to this β lying in F . Moreover, let γ be an element of \bar{F} such that $\gamma^{p^a} = \alpha$. Then $P(t) = (t - \gamma)^{p^a}$, so that the element γ is unique; moreover, since $(\gamma^{p^{a-1}})^p = \alpha$ and α has a unique p th root in \bar{F} , we must

have $\gamma^{p^{a-1}} = \beta$. That (i) implies (ii) is now easy: if $\beta \in F$, then we may write $P(t) = (t - \beta)^p$ so P is irreducible over F .

Conversely, assume that $P(t)$ is reducible over F , i.e., there exist $0 < i < p^r$ such that $(t - \gamma)^i \in F[t]$. The coefficient of t^{i-1} in this polynomial is $-i\gamma$, so if i is prime to p this implies that $\gamma \in F$, hence $\beta \in F$ which gives (i). So we may therefore assume that $i = p^b \cdot j$ where $1 \leq b \leq a-1$ and $\gcd(p, j) = 1$. Then $(t^{p^b} - \gamma^{p^b})^j \in F[t]$, and arguing as before we get that $\gamma^{p^b} \in F$, and therefore $\beta = (\gamma^{p^b})^{p^{a-b}} \in F$.

A field F is **perfect** if every irreducible polynomial over F is separable. It follows immediately from Prop XXc) that every field of characteristic 0 is perfect. In other words, the entire discussion of separability is nonvacuous only in positive characteristic, so for the remainder of this section we assume that all fields are of positive characteristic. Unless otherwise specified, p shall always denote a prime number which is the characteristic of the field(s) in question.

If F has characteristic $p > 0$, we consider the Frobenius homomorphism

$$f: F \rightarrow F, x \mapsto x^p.$$

Let $F^p = f(F)$ be the image, a subfield of F .

Proposition 32. a) A field of characteristic $p > 0$ is perfect iff the Frobenius homomorphism is surjective: $F^p = F$.

b) Therefore finite fields and algebraically closed fields are perfect.

Proof: Assume $F^p = F$, and let $P(t) = \sum_i a_i t^{pi}$ be an irreducible inseparable polynomial. We can then write $a_i = b_i^p$ and then

$$P(t) = \sum_i (b_i)^p (t^i)^p = \left(\sum_i b_i t^i \right)^p = Q(t)^p,$$

hence $P(t)$ is not irreducible after all. Therefore F is perfect. Inversely, if the Frobenius homomorphism is not surjective, then there exists some $\alpha \in F$ which is not a p th power, and then by Lemma XX the inseparable polynomial $t^p - \alpha$ is irreducible, so F is not perfect. This gives part a). As for part a), like any field homomorphism, the Frobenius map is injective, and an injective map from a finite set to itself is necessarily surjective. If F is algebraically closed, then for any $\alpha \in F$ the polynomial $t^p - \alpha$ has a root in F , i.e., $\alpha \in F^p$.

For any positive integer a , we may consider the \mathfrak{p}^a , the map which takes $\alpha \mapsto \alpha^{p^a}$, which can also be described as the a th power of the Frobenius map. We write $F^{p^a} = \mathfrak{p}^a(F)$. If F is not perfect then we get an infinite descending chain of proper subfields

$$F \supsetneq F^p \supsetneq F^{p^2} \supsetneq \dots$$

Indeed, if $\alpha \in F \setminus F^p$, then $\alpha^{p^{a-1}} \in F^{p^{a-1}} \setminus F^{p^a}$. This gives another proof that an imperfect field is infinite.

Exercise X.X.X: Let F be a field of characteristic p , with an algebraic closure \overline{F} . Define $F^{1/p} = \{\beta \in \overline{F} \mid \beta^p \in F\}$.

a) Show that $F^{1/p}$ is a subextension of \overline{F}/F .

b) Similarly define a tower of subextensions

$$F \subset F^{1/p} \subset F^{1/p^2} \subset \dots F^{1/p^a} \subset \dots \subset \overline{F},$$

and show that if F is imperfect, all these inclusions are strict.

c) Define $F^{1/p^\infty} = \bigcup_{a=1}^{\infty} F^{1/p^a}$. Show that F^{1/p^∞} is perfect and is the intersection of all perfect subextensions of \overline{F} . It is called the **perfect closure** of F .

Purely inseparable polynomials: Say that a polynomial $P(t) \in F[t]$ is **purely inseparable** if there exists exactly one $\alpha \in \overline{F}$ such that $P(\alpha) = 0$. As above, there are certainly purely inseparable polynomials over $F - (t - \alpha)^n$ for any $\alpha \in F$ and $n \in \mathbb{Z}^+$ – and what is of interest is the purely inseparable irreducible polynomials, which by the discussion thus far clearly can only exist in characteristic $p > 0$.

Proposition 33. *Let F be a field of characteristic $p > 0$. The irreducible, purely inseparable monic polynomials $P(t) \in F[t]$ are precisely those of the form $t^{p^a} - \alpha$ for some $a \in \mathbb{Z}^+$ and some $\alpha \in F \setminus F^p$.*

Proof: By Lemma XX, any polynomial of the form $t^{p^a} - \alpha$ for $\alpha \in F \setminus F^p$ is irreducible. Conversely, let $P(t)$ be a purely inseparable polynomial. By XXXXX, there exists a polynomial $P_2(t)$ such that $P(t) = P_2(t^p)$. Since P is irreducible, so is P_2 . If there exist distinct $\alpha, \beta \in \overline{F}$ such that $P_2(\alpha) = P_2(\beta)$ then there are unique and distinct elements $\alpha^{\frac{1}{p}}, \beta^{\frac{1}{p}}$ in \overline{F} such that $P(\alpha^{\frac{1}{p}}) = P(\beta^{\frac{1}{p}}) = 0$, contradicting the pure inseparability of α . Therefore P_2 must itself be irreducible purely inseparable, and an evident inductive argument finishes the proof.

Exercise XX: Show that the polynomial $t^6 - x$ over the field $\mathbb{F}_3[x]$ is irreducible and inseparable but not purely inseparable.

6.2. Separable algebraic field extensions.

Let F be a field and $P(t)$ an irreducible, inseparable polynomial over F of degree $d > 1$. Consider the finite field extension $K = F[t]/(P(t))$ of F . It exhibits some strange behavior. First, the only F -algebra embedding $\sigma : K \rightarrow \overline{K}$ is the inclusion map. Indeed, such embeddings correspond bijectively to the assignments of $t \in K$ to a root α of P in \overline{K} , and by assumption there are less than d such elements. It follows that the group $\text{Aut}(K/F)$ of F -algebra automorphisms of F has cardinality smaller than d .

Let us define a finite extension K/F to be **separable** if the set of F -algebra homomorphisms $K \rightarrow \overline{K}$ has cardinality exactly $[K : F]$. We will call an algebraic extension K/F separable if each of its finite subextensions separable.

For a finite extension K/F , we define the **separable degree** $[K : F]_s$ to be the cardinality of the set of F -algebra embeddings $\sigma : K \rightarrow \overline{F}$. By definition K/F is separable iff $[K : F]_s = [K : F]$.

Theorem 34. *The separable degree is multiplicative in towers: if $L/K/F$ is a tower of finite field extensions, then $[L : F]_s = [L : K]_s [K : F]_s$.*

Proof: ...

Since the degree is also multiplicative in towers it follows that:

Corollary 35. *Let $L/K/F$ be a tower of finite field extensions. Then L/F is separable iff $[L : K] = [L : K]_s$ and $[K : F] = [K : F]_s$.*

Theorem 36. *Let L/F be an algebraic field extension. TFAE:*

- (i) L/F is separable.
- (ii) Every irreducible polynomial $P \in F[t]$ which has a root in L is separable.
- (iii) L is obtained by adjoining to F a set of roots of separable polynomials.

Proof: It is easy to see that if these properties are equivalent for all finite subextensions of L/F then they are equivalent for L/F . Therefore we may assume that L/F is finite. Assume first that L/F is separable. If P is irreducible and has a root in L then $K := K[t]/(P)$ is a subextension of the finite separable extension L/K and by Corollary XX is therefore separable, which as above implies that P is separable. That (ii) implies (iii) is immediate; so assume that $L = F[\alpha_1, \dots, \alpha_n]$ where the minimal polynomial P_i of α_i is separable. (Obviously we can take the set of generators to be finite, since L/F is finite.) Then we get that $F[\alpha_1]/F$ is separable. The polynomial P_2 may become reducible over $F[\alpha_1]$; let P'_2 be an irreducible factor such that $P'_2(\alpha_2) = 0$. But no problem: any divisor of a separable polynomial is still separable, so $F[\alpha_1, \alpha_2] = F[\alpha_1]/(P'_2)$ is therefore separable over $F[\alpha_1]$. By Corollary XX this implies that $F[\alpha_1, \alpha_2]/F$ is separable. Proceeding in this manner we get that $L = F[\alpha_1, \dots, \alpha_n]/F$ is separable.

Corollary 37. *Any subextension of a separable field extension is separable.*

Proof: Apply condition (ii) of Theorem XX.

Corollary 38. *Let $\{K_i\}$ be a family of separable extensions of F inside a common algebraic closure \bar{F} . Then the compositum $K = \prod_i K_i$ of the K_i 's is separable.*

Exercise X.X.X: Apply condition (iii) of Theorem XX.

It follows that for any algebraic field extension K/F , there exists a largest separable subextension $\text{SepCl}_K(F)$, called the **separable closure** of F in K .

6.3. Purely inseparable extensions.

Theorem 39. *For an algebraic field extension K/F , TFAE:*

- (i) There is only one F -algebra embedding $K \hookrightarrow \bar{K}$.
- (ii) Every irreducible polynomial $P \in F[t]$ with a root in K is purely inseparable.
- (iii) K is obtained by adjoining to F roots of purely inseparable polynomials.
- (iv) The separable closure of K in F – i.e., the largest separable subextension – is F .

Proof: ...

A field extension satisfying the equivalent conditions of Theorem XX is said to be **purely inseparable**.

Corollary 40. *Let K/F be a purely inseparable extension of finite degree. Then $[K : F]$ is a power of p .*

Proof: One may reduce to the case of a simple extension $K = F[\alpha]$, and then α is purely inseparable over F so has minimal polynomial of the form $t^{p^a} - \alpha$ for some $a \in \mathbb{Z}^+$.

Corollary 41. *A purely inseparable extension is normal.*

Proof: This follows immediately from condition (i) of the Theorem.

Corollary 42. *If $L/K/F$ is a tower of fields, then L/F is purely inseparable iff L/K and K/F are both purely inseparable. In particular, a subextension of a purely inseparable extension is purely inseparable.*

Proof: Suppose L/F is purely inseparable, so that L is obtained by adjoining to F roots of purely inseparable polynomials. Then L is certainly obtained by adjoining to K roots of purely inseparable polynomials, so L/K is purely inseparable. If an irreducible polynomial $P \in F[t]$ has a root in K , then it certainly has a root in L , so by condition (ii) of Theorem XX it is purely inseparable. Conversely, suppose L/K and K/F are purely inseparable and let $\alpha \in L$. Then since α is the root of a purely inseparable polynomial with coefficients in K , there exists $a \in \mathbb{N}$ such that $\alpha^{p^a} \in K$. Similarly, there exists b such that $\alpha^{p^{a+b}} = (\alpha^{p^a})^{p^b} \in F$, so α is purely inseparable over F .

Corollary 43. *Let $\{K_i\}$ be a family of purely inseparable field extensions of F inside a common algebraic closure \overline{F} . Then the compositum $K = \bigwedge_i K_i$ is purely inseparable over F .*

Proof: This follows immediately from condition (iii) of the Theorem.

Therefore, for any algebraic field extension K/F , we may define the **purely inseparable closure** of F in K to be the largest subextension of K which is purely inseparable over F .

Exercise X.X.X: Show that the purely inseparable closure of F in an algebraic closure \overline{F} is the perfect closure F^{1/p^∞} .

The flavor of these results is that many formal properties are common to both separable and purely inseparable extensions. The exceptions to this rule are the following: first, purely inseparable extensions are always normal, whereas this is most certainly not the case for separable extensions. A more subtle difference is expressed in Theorem XX: if K/F is **not** purely inseparable, then it must have a nontrivial separable subextension. However, if K/F is **not** separable, that does not mean that it has a nontrivial purely inseparable subextension.

Example (Morandi, p. 48): Let k be a field of characteristic 2, $F = k(x, y)$ (rational function field), u a root in \overline{F} of the separable irreducible quadratic polynomial $t^2 + t + x$, $S = F(u)$ and $K = S(\sqrt{uy})$. Clearly K/S is purely inseparable and S/F is separable. But there is no nontrivial purely inseparable subextension of K/F . Equivalently, we will show that if $a \in K$, $a^2 \in F$, then already $a \in F$. An F -basis for K is $1, u, \sqrt{uy}, u\sqrt{uy}$. If $a^2 \in F$, write

$$a = \alpha + \beta u + \gamma \sqrt{uy} + \delta u\sqrt{uy}, \quad \alpha, \beta, \gamma, \delta \in F.$$

Since $a^2 \in F$, the coefficient of $u = 0$, i.e.,

$$\beta^2 + (\gamma + \delta)^2 y + \delta^2 xy = 0.$$

If $\delta = 0$ then $\beta^2 + \gamma^2 y = 0$, so $\gamma = 0$ since y is not a square in F . But then $\beta = 0$ and $a \in F$. If $\delta \neq 0$, then

$$x = \frac{\beta^2 + (\gamma + \delta)^2 y}{\delta^2 y} = \left(\frac{\gamma}{\delta} + 1\right)^2 + \left(\frac{\beta}{\delta}\right)^2 y,$$

so that $x \in F^2(y)$, which is not the case. So $\delta = 0$ and $a \in F$.

6.4. Structural results on algebraic extensions.

Proposition 44. *Suppose an algebraic extension K/F is both separable and purely inseparable. Then $K = F$.*

Exercise: Prove Proposition XX.

Proposition 45. (*Déviissage for algebraic field extensions*) *For any algebraic field extension K/F , the extension $K/\text{SepCl}_K(F)$ is purely inseparable.*

Exercise: Prove Proposition XX.

In general this result is not valid the other way around: an algebraic field extension K/F need not be separable over its purely inseparable closure. Indeed, in Example XX of the previous section the purely inseparable closure F_i was F and K/F was not separable. The following two results give more information on when K is separable over F_i .

Theorem 46. *For an algebraic extension K/F , let F^s and F^i be, respectively, the separable and purely inseparable closures of F in K . TFAE: (i) $K = F_s F_i$. (ii) K is separable over F_i .*

Proof: If (i) holds, then K is obtained by adjoining to F_i roots of separable polynomials with coefficients in F , hence by polynomials with coefficients in F_s . Conversely, if K/F_i is separable then $K/F_i F_s$ is separable; and since K/F_s is inseparable, $K/F_i F_s$ is inseparable. By Proposition XX, we conclude $K = F_i F_s$.

Corollary 47. *The equivalent conditions of Theorem XX hold when K/F is normal. In particular they hold for \overline{F}/F , giving $\overline{F} = F^{\text{sep}} F^{1/p^\infty}$.*

Proof: Let $\alpha \in K \setminus F^i$. Then α is not purely inseparable over F , i.e., the minimal polynomial P of α has at least one other distinct root, say β , in an algebraic closure. But since K/F is normal, $\beta \in F$. By the Extension Theorem, there exists an F -algebra automorphism s of L such that $s(\alpha) = \beta$. This shows that the set of elements in F which are fixed by every automorphism of L/F is precisely F^i . Let Q be the minimal polynomial of α over F^i , and let $\alpha_1, \dots, \alpha_r$ be the distinct roots of Q in \overline{F} . Since the group G of automorphisms of K/F^i acts on the α_i 's by permutations, the separable polynomial $R(t) = \prod_{i=1}^r (t - \alpha_i)$ is invariant under G , i.e., it lies in $F^i[x]$. This shows that K/F_i is obtained by adjoining roots of separable polynomials and is therefore separable. The second sentence of the Corollary follows immediately from the first.

Corollary 48. *For a finite extension K/F , $[K : F]_s = [\text{SepCl}_K(F) : F]$. In particular $[K : F]_s \mid [K : F]$.*

Proof: We have $[K : F]_s = [K : \text{SepCl}_K(F) : F]_s [\text{SepCl}_K(F) : F]_s$. But the separable degree of a purely inseparable extension is 1, so the conclusion follows.

For a finite extension K/F one may therefore define the **inseparable degree** $[K : F]_i$ of a finite extension to be $[K : F]/[K : F]_s = [K : \text{SepCl}_K(F)]$. (Note that from the above discussion we may not in general call it the “purely inseparable degree.”)

Remark: If we like, we can define the separable degree of an infinite algebraic extension L/K to be the cardinal number $[\text{SepCl}_K(F) : F]$ and the inseparable degree to be the cardinality $[K : \text{SepCl}_K(F)]$, as e.g. is done in Lang’s *Algebra*. But I do not know any particular application of this.

A field is **separably closed** if it does not admit any proper separable field extension.

Proposition 49. *The separable closure of F in any algebraic closure \overline{F} is separably closed.*

Exercise X.X.X: Prove Proposition XX.

One often writes F^{sep} for a separable closure of F . Like the algebraic and normal closures, this extension is unique up to non-canonical F -algebra isomorphism, so it is common to speak of “the separable closure” but we shall endeavor not to do so here.

Corollary 50. *Let K/F be a normal algebraic extension. Then the separable closure F^s of F in K is also normal.*

Proof: For any embedding σ of K into \overline{F} , the image $\sigma(F^s)$ lies in K (by normality of K) and is evidently also a separable subextension of K/F . Therefore we must have $\sigma(F^s) = F^s$.

Corollary 51. *A field F is perfect iff its separable closure is algebraically closed.*

Proof: If F is perfect there is no distinction between algebraic and separable extensions, so the result is clear. Inversely, suppose that F is not perfect, so there exists $\alpha \in F \setminus F^p$ and a corresponding purely inseparable field extension $F[\alpha^{1/p}]/F$ defined by the irreducible inseparable polynomial $P = t^p - \alpha$. According to Theorem XX only a separable irreducible polynomial can acquire a root in a separable field extension, so the polynomial P remains irreducible over the separable closure of F .

6.5. The primitive element theorem.

Both the following result and its corollary go by the name “Primitive Element Theorem.”

Theorem 52. *Let K/F be a finite field extension. TFAE:*

- (i) *K/F is simple: there exists $\alpha \in K$ such that $K = F[\alpha]$.*
- (ii) *The set of subextensions L of K/F is finite.*

Moreover, these conditions hold when K/F is separable.

7. GALOIS EXTENSIONS

7.1. Introduction.

For any field extension K/F we define $\text{Aut}(K/F)$ to be the group of F -algebra automorphisms of K , i.e., the set of all field isomorphisms $\sigma : K \rightarrow K$ such that $\sigma(x) = x$ for all $x \in F$. This is a group under composition.

Let G be a subgroup of $\text{Aut}(K/F)$, i.e., a group of F -algebra automorphisms of K . We define the **fixed field**

$$K^G = \{x \in K \mid \sigma(x) = x \forall \sigma \in G\}.$$

Note that the notation comes from representation theory: if R is a commutative ring, M an R -module and G is a group, then one has the notion of an R -linear representation of G on M , i.e., a homomorphism from G to the group of R -module automorphisms of M . In such a situation one can “take invariants”, i.e., consider the subset of M on which G acts trivially: this is denoted M^G . The present definition is an instance of this with $R = F$, $M = K$.

It is immediate to check that K^G is a subextension of K/F . (In fact in the more general setting detailed above, one checks that M^G is an R -submodule of M .)

Definition: We say that K/F is **Galois** if $K^{\text{Aut}(K/F)} = F$. Equivalently, for any element $x \in K \setminus F$, there exists $\sigma \in \text{Aut}(K/F)$ such that $\sigma(x) \neq x$.

This “top down” definition of a Galois extension is the generalization to arbitrary extensions of a definition of E. Artin for finite extensions. It has the merit of making it easy to exhibit a large class of Galois extensions: if K is any field and G is any group of automorphisms of K , then K/K^G is, tautologically, a Galois extension.

Example: Let G be the 2-element subgroup of the complex numbers generated by complex conjugation. Then $\mathbb{C}^G = \mathbb{R}$, so \mathbb{C}/\mathbb{R} is a Galois extension.

Example: Let L/K be a separable quadratic extension, so that $L = K[t]/(P(t))$, where $P(t)$ is a separable polynomial. Then $P(t)$ splits over L into $(t - \alpha)(t - \bar{\alpha})$, so that the automorphism group of L/K has order 2, the nontrivial element being the unique K -automorphism σ of L which sends $\alpha \mapsto \bar{\alpha}$. Since $L^{\text{Aut}(L/K)}$ is a subextension of the degree 2 extension L/K , it could only be L or K , and since $\sigma(\alpha) = \bar{\alpha} \neq \alpha$, we conclude that the fixed field is K and the extension is Galois. In contrast the automorphism group of an inseparable quadratic extension is trivial, so this extension is not Galois.

Example: Let $K = \mathbb{Q}[t]/(t^3 - 2) = \mathbb{Q}[\sqrt[3]{2}]$. Since K contains exactly one of the three roots of $t^3 - 2$ in $\overline{\mathbb{Q}}$, $\text{Aut}(K/\mathbb{Q})$ is the trivial group and K/\mathbb{Q} is not Galois. On the other hand, the automorphism group of the normal closure $M = \mathbb{Q}[\zeta_3, \sqrt[3]{2}]$ of K/\mathbb{Q} has order 6: since everything is separable, there are three embeddings of $\mathbb{Q}[\sqrt[3]{2}]$ into M , and each of these extends in two ways to an automorphism of M . Any automorphism s of M is determined by an $i \in \{0, 1, 2\}$ and $j \in \{0, 1\}$ such that

$$s : \sqrt[3]{2} \mapsto \zeta_3^i \sqrt[3]{2}, \quad \zeta_3 \mapsto (\zeta_3)^{(-1)^j}.$$

Since there are six possibilities and six automorphisms, all of these maps must indeed give automorphisms. In particular, there is an order 3 automorphism σ which takes $\sqrt[3]{2} \mapsto \zeta_3 \sqrt[3]{2}$ and fixes ζ_3 and an order 2 automorphism τ which fixes $\sqrt[3]{2}$ and maps $\zeta_3 \mapsto \zeta_3^{-1}$. One checks that $\tau\sigma\tau = \tau\sigma\tau^{-1} = \sigma^{-1}$, i.e., $\text{Aut}(L/\mathbb{Q}) \cong S_3$, the symmetric group on three elements. Indeed, these three elements can be viewed as the three roots of $t^3 - 2$ in M . Finally, the subgroup fixed by $\{1, \sigma\}$ is precisely K , whereas the generator $\sqrt[3]{2}$ of K/\mathbb{Q} is not fixed by σ , so that we conclude that $M^{\text{Aut}(M/\mathbb{Q})} = \mathbb{Q}$ and M/\mathbb{Q} is Galois.

These examples already suggest that a finite extension K/F is Galois iff it is normal and separable, and in this case $\#\text{Aut}(K/F) = [K : F]$. We will show in the next section that these conditions are all equivalent.

Example: The extension $\overline{\mathbb{Q}}/\mathbb{Q}$ is Galois. We cannot show this by some sort of direct computation of $G_{\mathbb{Q}} := \text{Aut}(\overline{\mathbb{Q}}/\mathbb{Q})$: this group is uncountably infinite and has a very complicated structure. Indeed, as an algebraic number theorist I am more or less honorbound to inform you that the group $G_{\mathbb{Q}}$ is the single most interesting group in all of mathematics! We will see that the Galois theory of infinite algebraic extensions cannot be developed in exactly the same way as in the finite case, but is, in theory, easily understood by a reduction to the finite case.

Example: The extension \mathbb{C}/\mathbb{Q} is Galois, as is $\mathbb{C}/\overline{\mathbb{Q}}$. In particular the automorphism group of the complex field is (much) larger than just $\{1, c\}$. In fact we will show that if F has characteristic zero and K is algebraically closed, then K/F is Galois. These results are not part of ‘‘Galois theory’’ as it is usually understood, but rather are facts about automorphism groups of transcendental extensions.

Example: For any field F , $\text{Aut}(F(t)/F)$ is the group of linear fractional transformations: the group $GL_2(F)$ of 2×2 matrices $[[ab]][[cd]]$ with $ad \neq bc$ acts by automorphisms on $F(t)$, via $t \mapsto \frac{at+b}{ct+d}$. Scalar matrices – those with $b = d = 0$, $a = c$ – act trivially, so the action factor through to the quotient $PGL_2(F)$ of $GL_2(F)$ by the subgroup F^\times of scalar matrices. It is a standard fact (more in the vein of algebraic geometry than pure field theory) that this is the entire automorphism group of $F(t)$.

Proposition 53. *The extension $F(t)/F$ is Galois iff F is infinite.*

Proof: We will need to use a fact from the next section: if G is a finite group of automorphisms acting on a field K , then $[K : K^G] = \#G < \infty$. Therefore if F is finite, $F(t)^{\text{Aut}(F(t)/F)}$ has finite index in $F(t)$, so is certainly not equal to F . Conversely assume F is infinite. . .

Remark Aside: I am not aware of a simple necessary and sufficient condition for an extension K/F which is finitely generated, but of infinite degree, to be Galois. When K/F is regular of transcendence degree 1 (two terms which we have not yet defined), one can give such a criterion in terms of the Jacobian $J(C)$ of the corresponding algebraic curve C/F , namely K/F is Galois iff $\dim J(C) = 0$ or ($\dim J(C) = 1$ and $J(C)(F)$ is infinite). In particular no such field of *genus*

$g \geq 2$ is Galois. One can give some examples of Galois extensions of higher transcendence degree – e.g. the proof of Proposition XX easily adapts to show that $F(t_1, \dots, t_n)/F$ is Galois if F is infinite – but the general problem seems to be a quite subtle one in birational arithmetic geometry.

7.2. Finite Galois extensions.

Proposition 54. (*Linear independence of characters*) *Let M be a monoid and K a field. The set $X(M, K)$ of all monoid homomorphisms $M \rightarrow K^\times$ is linearly independent as a subset of the K -vector space K^M of all functions from M to K .*

Proof: If the set $X(M, K)$ were not linearly independent, then (by definition) there would exist finitely many elements $\chi_1, \dots, \chi_N \in X(M, K)$ and $a_1, \dots, a_N \in K$, not all zero, such that for all $x \in M$

$$(1) \quad \alpha_1 \chi_1(x) + \dots + \alpha_N \chi_N(x) = 0.$$

Choose $m \in M$ such that $\chi_1(m) \neq \chi_2(m)$ (after all χ_1 and χ_2 are distinct functions). Applying the dependence relation with mx in place of x , we get

$$(2) \quad \alpha_1 \chi_1(m) \chi_1(x) + \alpha_2 \chi_2(m) \chi_2(x) + \dots + \alpha_N \chi_N(m) \chi_N(x) = 0.$$

Multiplying (2) by $\chi_1(m)^{-1}$ and subtracting this from (1), we get

$$(3) \quad \alpha_2 \left(\frac{\chi_2(m)}{\chi_1(m)} - 1 \right) \chi_2(x) + \dots + \alpha_N \left(\frac{\chi_N(m)}{\chi_1(m)} - 1 \right) \chi_N(x) = 0.$$

Since $\chi_1(m) \neq \chi_2(m)$, the coefficient of $\chi_2(x)$ is nonzero, and this produces a linear dependence relation in $N - 1$ characters. Proceeding in this way we will get a linear dependence relation in 1 character, which is absurd as the only one element subset of K^M which is linearly independent is the identically zero function, whereas characters take on only nonzero values. This contradiction completes the proof.

Theorem 55. *If K/F is a finite field extension, $\text{Aut}(K/F)$ is a finite group of cardinality at most $[K : F]$.*

Proof: First recall that the set of F -algebra embeddings σ of K into an algebraic closure \bar{F} is finite, so in particular the subset of such with $\sigma(K) = K$ is finite. This holds because $K = F(\alpha_1, \dots, \alpha_n)$, and an embedding σ is determined by sending each α_i to one of the at most $d_i = [F[\alpha_i] : F]$ roots of the minimal polynomial of α_i over F in \bar{F} . Therefore the set of such embeddings has cardinality at most $d_1 \cdots d_n$. Note that when $K = F[\alpha]$ is simple this is exactly the bound we want, so that e.g. if K/F is separable we are already done.

Now for the general case. Let $\text{Aut}(K/F) = \{\sigma_1, \dots, \sigma_N\}$ and suppose, for a contradiction, that $N > m = [K : F]$. Let $\alpha_1, \dots, \alpha_m$ be an F -basis for K , and consider the $N \times m$ matrix A whose (i, j) entry is $\sigma_i(\alpha_j)$. This matrix has rank at most $m < N$, so that its rows are K -linearly dependent: there exist $c_1, \dots, c_N \in K$, not all 0, such that for all $1 \leq j \leq m$ we have

$$\sum_i c_i \sigma_i(\alpha_j) = 0.$$

For each $x \in K^\times$, there exist a_1, \dots, a_m in F such that $x = \sum_j a_j \alpha_j$. Then

$$\sum_i c_i \sigma_i(x) = \sum_i c_i \sigma_i \left(\sum_j a_j \alpha_j \right) = \sum_i c_i \left(a_j \sum_j \sigma_j(\alpha_j) \right)$$

$$= \sum_j a_j \left(\sum_i c_i \sigma_i(\alpha_j) \right) = 0.$$

But taking $M = K^\times$ all the automorphisms σ_i give characters $M \rightarrow K^\times$ hence are K -linearly independent. Therefore in the last equation we must have $c_i = 0$ for all i , a contradiction.

Proposition 56. *Let K be a field and G a finite group of automorphisms of K . Then $[K : K^G] = \#G$.*

Proof: Let $N = \#G$. Since $G \subset \text{Aut}(K/K^G)$, Theorem XX gives $\#G = N \leq [K : K^G]$. Suppose we had $N < [K : K^G]$. Then there exist $N + 1$ K^G -linearly independent elements $\alpha_1, \dots, \alpha_{N+1}$ of K . Let $G = \{\sigma_1, \dots, \sigma_N\}$, and as before consider the matrix A whose (i, j) entry is $\sigma_i(\alpha_j)$. This time rank considerations tell us that the columns are linearly dependent over K . After relabelling the α_i 's if necessary, let n be the largest number such that the first $n - 1$ columns are K -linearly independent (so $n \leq N$). Then there exist $c_i \in K$, not all zero, such that $\sum_{i=1}^n c_i \sigma_j(\alpha_i) = 0$ for all j , and by minimality we have all the c_i 's are nonzero. ...

Theorem 57. (*Omnibus theorem for finite Galois extensions*) *Let K/F be a finite extension. TFAE:*

- (i) $K^{\text{Aut}(K/F)} = F$ (“ K/F is Galois.”)
- (ii) $\# \text{Aut}(K/F) = [K : F]$.
- (iii) K/F is normal and separable.
- (iv) K/F is the splitting field of a separable polynomial.

Proof: Let $G = \text{Aut}(K/F)$. (i) implies (ii) by Proposition XX. (ii) implies (i): we have $F \subset K^G \subset K$, and $[K : K^G] = \#G = [K : F]$, so $K^G = F$.

(iii) implies (iv): if K/F is separable then by the Primitive Element Theorem $K = F[t]/(P(t))$ for some irreducible, separable polynomial P . Since it is normal, P splits in K and therefore K/F is the splitting field of the separable polynomial P . (iv) implies (iii) is essentially the same: since K/F is a splitting field, it is normal; since it is obtained by adjoining roots of separable polynomials, it is separable.

(iv) \iff (ii): We know that the number of embeddings of K into \overline{F} is equal to the separable degree of K/F and that this equals $[K : F]$ iff K/F is separable; moreover, every F -algebra embedding $s : K \rightarrow \overline{F}$ has $s(K) = K$ – i.e., gives an automorphism of K iff K/F is normal.

Corollary 58. *A finite extension is a subextension of a finite Galois extension iff it is separable. Any algebraic closure \overline{F} of K contains a unique minimal extension M of K such that M/F is Galois, namely the normal closure of K/F in \overline{F} .*

Proof: Since Galois extensions are separable and subextensions of separable extensions are separable, for K/F to be contained in a finite Galois extension it is clearly necessary for it to be separable. If so, then the normal closure M of K/F , being a compositum of the separable extensions $s(K)$ as s ranges over the finite set of distinct F -algebra embeddings of K into \overline{F} is separable and normal, hence Galois. M/K is even the minimal extension of K which is normal over F , so certainly it is the minimal such Galois extension.

Remark: In view of Corollary XX, it is reasonable to call the normal closure of a finite separable field extension the **Galois closure**.

7.3. An abstract Galois correspondence.

Let X be a set and G a group of automorphisms of X , i.e., a subgroup of the group $\text{Sym}(X)$ of all bijections $s : X \rightarrow X$. Let $\Lambda(X)$ be the collection of all subsets of X and $\Lambda(G)$ be the collection of all subgroups of G . Both $\Lambda(X)$ and $\Lambda(G)$ are partially ordered sets under inclusion.

For a subset $Y \subset X$, we define

$$G_Y = \{g \in G \mid gy = y \forall y \in Y\},$$

which is a subgroup of G . Dually, for a subgroup H of G , we define

$$X^H = \{x \in X \mid gx = x \forall g \in H\},$$

which is a subset of X . (We could define in the same way X^S for any subset $S \subset G$, but one checks immediately that if H is the subgroup generated by S , $X^S = X^H$, so this extra generality leads nowhere.) To be very formal about it, we have thus defined a map

$$\Phi : \Lambda(X) \rightarrow \Lambda(G), \quad Y \mapsto G_Y$$

and a map

$$\Psi : \Lambda(G) \rightarrow \Lambda(X), \quad H \mapsto X^H.$$

Let us explore what can be said about these two maps in this extreme level of generality. Statements that we do not prove are exercises in unwinding the definitions and left to the reader. (We do recommend that the reader perform these exercises!)

First, both Φ and Ψ are **anti**-homomorphisms of the partially ordered sets, i.e., if $Y_1 \subset Y_2$, then $\Phi(Y_2) \subset \Phi(Y_1)$, and similarly if $H_1 \subset H_2$ then $\Psi(H_2) \subset \Psi(H_1)$. This implies that $\Psi \circ \Phi : \Lambda(X) \rightarrow \Lambda(X)$ and $\Phi \circ \Psi : \Lambda(G) \rightarrow \Lambda(G)$ are homomorphisms of partially ordered sets:

$$Y_1 \subset Y_2 \implies X^{G_{Y_1}} \subset X^{G_{Y_2}},$$

$$H_1 \subset H_2 \implies G_{X^{H_1}} \subset G_{X^{H_2}}.$$

Moreover, for all $Y \subset X$ and $H \subset G$ we have

$$(GC) \quad Y \subset X^H \iff H \subset G_Y.$$

Indeed, both containments assert precisely that every element of H acts trivially on every element of Y . If $H = G_Y$ we certainly have the second containment, therefore by (GC) we have

$$(4) \quad Y \subset X^{G_Y}.$$

Dually with $Y = X^H$ we certainly have the first containment hence (GC) gives

$$(5) \quad H \subset G_{X^H}.$$

Proposition 59. *Let H be a subgroup of G , Y a subset of X and $\sigma \in G$. We have:*

$$a) \quad \sigma G_Y \sigma^{-1} = G_{\sigma Y}.$$

$$b) \quad \sigma X^H = X^{\sigma H \sigma^{-1}}.$$

Proof: We have $g \in G_{\sigma Y} \iff \forall y \in Y, g\sigma y = \sigma y \iff \forall y \in Y, \sigma^{-1}g\sigma y = y \iff \sigma^{-1}g\sigma \in G_Y \iff g \in \sigma G_Y \sigma^{-1}$. Similarly, $y \in \sigma X^H \iff \sigma^{-1}y \in X^H \iff \forall h \in H, h\sigma^{-1}y = \sigma^{-1}y \iff \forall h \in H, (\sigma h \sigma^{-1})y = y \iff y \in \sigma H \sigma^{-1}$.

Let us now introduce the following simplified (and symmetric) notation: for $Y \subset X$, we write Y' for G_Y ; for $H \subset G$, we write H' for X^H . Equations (4) and (5) now read as $Y \subset Y''$ and $H \subset H''$. Let us call a subset Y of X (resp. a subgroup H of G) **closed** if $Y'' = Y$ (resp. if $H'' = H$).

Proposition 60. *For any $Y \in \Lambda(X)$ and $H \in \Lambda(G)$, we have $Y' = Y'''$ and $H' = H'''$. Hence Y' is a closed subgroup of G and X' is a closed subset of X .*

Proof: By (4) we have $Y' \subset (Y')''$ and $Y \subset Y''$. Applying a prime to the latter containment reverses it and hence gives $Y' \supset (Y'')'$. Therefore $Y' = Y'''$. The argument for H is identical.

Remark: This shows that the operators $'$ on the posets $\Lambda(X)$ and $\Lambda(G)$ are what are called **closure operators**. In general, if (S, \leq) is a partially ordered set, then a map $c : S \rightarrow S$ is a closure operator if for all $s \in S$, $s \leq c(s)$, $s \leq t \implies c(s) \leq c(t)$ and $c(c(s)) = c(s)$ for all $s \in S$.

Corollary 61. *Let $\Lambda_c(X)$ be the closed subsets of X and $\Lambda_c(G)$ be the closed subgroups of G . Let Φ_c be Φ restricted to $\Lambda_c(X)$ and Ψ_c be Ψ restricted to $\Lambda_c(G)$. Then*

$$\Phi_c : \Lambda_c(X) \rightarrow \Lambda_c(G), \quad \Psi_c : \Lambda_c(G) \rightarrow \Lambda_c(X)$$

give mutually inverse anti-automorphisms of posets.

In fact the proof is immediate from the previous result; again, it is a good exercise for the reader to chase through the definitions and notation to see this.

Corollary 62. *a) A closed subgroup H of G is normal iff its corresponding closed subset $Y = H' = H'''$ is stable under all automorphisms of G : for all $\sigma \in G$, $\sigma Y = Y$.*

b) A closed subset Y of X is stable under all automorphisms of G iff the corresponding closed subgroup $H = Y' = Y'''$ is normal in G .

Again, this follows immediately from Proposition XX.

Exercise X.X: Show that if H is a normal subgroup of G , so is its closure H'' .

Example: Suppose $\#X > 1$; let $x \in X$ and take $Y = X \setminus x$. Then Y is not a closed subset of X , since any group of automorphisms of X which fixes every element of Y must also fix x .

Example: If $X = \{1, 2\}$ and $G = S_2$ is the full symmetry group of X , then the closed subsets are \emptyset and X ; the corresponding closed subgroups are G and the trivial subgroup e . In particular all subgroups are closed. If $X = \{1, 2, 3\}$ and G is the full symmetry group S_3 . The closed subsets are \emptyset , $\{1\}$, $\{2\}$, $\{3\}$ and X . The corresponding closed subgroups are S_3 , $\langle(23)\rangle$, $\langle(13)\rangle$, $\langle(12)\rangle$ and the trivial subgroup e . In particular the (unique) subgroup $H = \langle(123)\rangle$ of order 3 is not closed: $H' = \emptyset$ and $H'' = S_3$. If X is any set, the only subsets invariant under $G = \text{Sym}(X)$ are \emptyset and X itself, so $\text{Sym}(S)$ does not have any nontrivial, proper

closed normal subgroups. On the other hand, if $\#X \geq 3$ then $\text{Sym}(S)$ always has a nontrivial, proper normal subgroup (i.e., it is not a simple group): if S is finite, so $\text{Sym}(S) \cong S_n$ take the alternating group A_n (the only possible choice if $n \geq 5$); if S is infinite, take the the subgroup H of elements $g \in \text{Sym}(S)$ such that $X \setminus X^{(g)}$ is finite. Then $\#H = \#S$ while $\#\text{Sym}(S) = 2^{\#\text{Sym}(S)}$.

Key example: Let K/F be a field extension, $X = K$ and $G = \text{Aut}(K/F)$. Then every closed subset of X is a subextension K^H of K/F . Corollary XX shows that there is a bijective correspondence between the closed subextensions of K/F and the closed subgroups of $\text{Aut}(K/F)$. Of course the key word in the previous sentence is “closed”: if e.g. $\text{Aut}(K/F)$ is the trivial group (e.g. $K = \mathbb{R}$) then the statement is completely vacuous. In the next section we will show that if K/F is a finite Galois extension, the best possible behavior occurs.

Exercise X.X: Let Λ and Λ' be two partially ordered sets. A **Galois connection** between Λ and Λ' is a pair of order-reversing maps $\Phi : \Lambda \rightarrow \Lambda'$, $\Psi : \Lambda' \rightarrow \Lambda$ satisfying the analogue of identity (GC)¹⁰ above: for $x \in \Lambda$, $y \in \Lambda'$, $\Phi(x) \leq y \iff x \leq \Psi(y)$.

a) Check that the entire discussion (except for the bit about conjugation and normality) goes through in this level of generality: we get closure operators on Λ and Λ' such that Φ and Ψ give mutually inverse anti-automorphisms on the subsets of closed elements: $\Phi : \Lambda_c \xrightarrow{\sim} \Lambda'_c$, $\Psi : \Lambda'_c \xrightarrow{\sim} \Lambda_c$.

b) Look for Galois connection. in your everyday (mathematical) life, paying special attention to the closure process. For example, consider the polynomial ring $R = k[t_1, \dots, t_n]$ over an algebraically closed field k . Let Λ be the set of ideals I of R , and let Λ' be the set of **algebraic** subsets of affine n -space \mathbb{A}^n over k : that is, the subsets of k^n of the form $\bigcap_{i \in I} P_i^{-1}(0)$, where $\{P_i\}_{i \in I}$ is a set of elements of R . Define $\Phi : \Lambda \rightarrow \Lambda'$ by $I \mapsto V(I)$, the set of points of (a_1, \dots, a_n) such that $P(a_1, \dots, a_n) = 0$ for all $P \in I$. Define $\Psi : \Lambda' \rightarrow \Lambda$ by $S \mapsto I(S)$, the ideal of all elements of R which vanish at every $(x_1, \dots, x_n) \in S$. It is no problem to see that this gives a Galois connection. What are the closed ideals? What are the closed algebraic subsets?

7.4. The finite Galois correspondence.

Let K/F be a finite Galois extension, so that by the general nonsense of the previous section, we get a bijective correspondence between closed subextensions L of K and closed subgroups of $G = \text{Aut}(K/F)$.

Theorem 63. (*Fundamental theorem of Galois theory*) *If K/F is finite Galois, then every subgroup H of $G = \text{Aut}(K/F)$ is closed, i.e., of the form $H = \text{Gal}(K/L)$ for a unique subextension L/K . Conversely, every subextension L is closed, i.e., of the form K^H for a unique subgroup H of G . Therefore the maps $L \mapsto \text{Gal}(K/L)$ and $H \mapsto K^H$ give mutually inverse inclusion-reversing bijections between the set of subextensions of L/K and the set of subgroups of G . Moreover, a subextension L is Galois over F iff the corresponding subgroup $\text{Gal}(K/L)$ is normal in G , and in this case $\text{Aut}(L/F)$ is canonically isomorphic to the quotient $\text{Aut}(K/F)/\text{Aut}(K/L)$.*

¹⁰In particular “GC” stands for **G**alois **C**onnection.

Proof: Let L be a subextension of K/F . It is clear (either from the previous section, or just because it's a trivial statement) that $L \subset K^{\text{Aut}(K/L)}$. But by XXX we know that $[K : K^{\text{Aut}(K/L)}] = \#\text{Aut}(K/L)$. Since K/F is Galois, so is K/L , hence $\#\text{Aut}(K/L) = [K : L]$. Therefore we must have $K^{\text{Aut}(K/L)} = L$. Moreover, if H is a subgroup of G , we again clearly have $H \subset G_{KH}$; but we also have $[G : G_{KH}] = \frac{\#G}{\#\text{Aut}(K/K^H)} = [G : H]$, so $H = G_{KH}$. This shows the Galois correspondence is perfect. Now applying Corollary XX we get that $H = \text{Gal}(K/L)$ is normal in G iff L is stable under all F -algebra automorphisms σ of K . Since K/F is itself normal, this holds iff L is stable under all F -algebra embeddings into an algebraic closure \bar{F} , i.e., iff L/F is normal. Finally, suppose that L/F is normal. Then every F -automorphism of K restricts to an F -automorphism of L , giving a natural map $\text{Aut}(K/F) \rightarrow \text{Aut}(L/F)$ which is easily checked to be a homomorphism of groups. The map is surjective by the Extension Theorem XX. Its kernel is the subgroup of F -algebra automorphisms of K which fix every element of L , i.e., $\text{Aut}(K/L)$.

This theorem is probably the single most important result in field theory. It reduces the study of the lattice of subextensions of a finite Galois extension K/F to the corresponding lattice of subgroups of the finite group $\text{Aut}(K/F)$, which is much easier to study, e.g. is *a priori* finite. Indeed, if K/F is any finite separable extension, then one may – and should! – apply the Galois correspondence to the Galois closure M/F . In particular this provides a much more conceptual proof of the fact that a finite separable extension has only finitely many subextensions – we may embed it in a finite Galois extension, which has that property by the Fundamental Theorem of Galois Theory.

When K/F is Galois, we write $\text{Gal}(K/F)$ for $\text{Aut}(K/F)$ and speak of $\text{Gal}(K/F)$ as the **Galois group** of K/F . We note that some authors (e.g. Shifrin, Kaplansky) use the notation $\text{Gal}(K/F)$ for the automorphism group of an arbitrary field extension, but from the perspective of infinite Galois theory (coming up!) and modern number theory this seems dangerously misleading. Namely, it would then be tempting to call any automorphism group of a finite extension “a Galois group” and this is most certainly at odds with contemporary terminology. Indeed, perhaps the single outstanding problem in field theory is to decide whether, for any finite group G , there is a Galois extension K/\mathbb{Q} such that $\text{Gal}(K/\mathbb{Q}) \cong G$. However, the corresponding statement that any finite group is the automorphism group of some finite extension K/\mathbb{Q} – possibly with $[K : \mathbb{Q}] > \#G$ – is a much weaker one, and indeed is a 1978 theorem of E. Fried and J. Kollar.

Composita of Galois extensions: let F be a field and K_1, K_2 two Galois extensions of F . After choosing an algebraic closure \bar{F} of F , since K_1 and K_2 are splitting fields, there is a unique F -algebra embedding of K_i into \bar{F} . Since composita of normal (resp. separable) extensions are normal (resp. separable), the compositum $K = K_1 \vee K_2$ is a finite Galois extension. What is the relationship of $\text{Gal}(K/F)$ to $\text{Gal}(K_1/F)$ and $\text{Gal}(K_2/F)$? As above we get surjective restriction maps $\iota_i : \text{Gal}(K/F) \rightarrow \text{Gal}(K_i/F)$, and hence a diagonal map $\iota = (\iota_1, \iota_2) : \text{Gal}(K/F) \rightarrow \text{Gal}(K_1/F) \times \text{Gal}(K_2/F)$. This composite homomorphism ι need not (of course?) be surjective: e.g. it will not be if $K_1 = K_2$ are

nontrivial extensions of F . Rather ι is always injective: since K is generated as a field by K_1 and K_2 , a pair of automorphisms σ_i of K_i can extend in at most one way to an automorphism of K . Therefore $\text{Gal}(K/F)$ can naturally be viewed as a subgroup of the product $\text{Gal}(K_1/F) \times \text{Gal}(K_2/F)$.

This is in fact rather useful: let \mathcal{C} be any class of finite groups which is closed under formation of direct products and passage to subgroups, and suppose that K_i/F are two \mathcal{C} -Galois extensions, i.e., finite Galois extensions whose Galois groups lie in \mathcal{C} . Then the compositum $K_1 \vee K_2$ is a \mathcal{C} -Galois extension. E.g. we may profitably take \mathcal{C} to be the class of all finite abelian groups, or the class of all finite solvable groups. When we turn to infinite Galois theory we will see that we are allowed to take infinite composita as well, and this observation will show that any field admits a maximal \mathcal{C} -Galois extension.

Exercise: Let $K_1, K_2/F$ be two finite Galois extensions, and $K = K_1K_2$ their compositum. Let H be the image of the map $\iota : \text{Gal}(K/F) \rightarrow \text{Gal}(K_1/F) \times \text{Gal}(K_2/F)$. Show that H is normal in $\text{Gal}(K_1/F) \times \text{Gal}(K_2/F)$, and that the quotient $(\text{Gal}(K_1/F) \times \text{Gal}(K_2/F))/H \cong \text{Gal}(K_1 \cap K_2/F)$. In particular, ι is an isomorphism iff $K_1 \cap K_2 = F$.

7.5. Galois groups of infinite algebraic extensions.

Theorem 64. *For an algebraic field extension K/F , TFAE:*

- (i) $K^{\text{Aut}(K/F)} = F$. (“ K/F is Galois.”)
- (ii) K is normal and separable.
- (iii) K is the splitting field of a set (possibly infinite) of separable polynomials.

Proof: The equivalence of (ii) and (iii) follows from our characterization of normal and separable algebraic extensions.

(i) \implies (ii): (Morandi, p. 40something)

(ii) \implies (i): Let $\alpha \in K \setminus F$. Then the minimal polynomial P for α over F splits in K and has at least one other distinct root β . There is a unique F -algebra embedding $\sigma : F[\alpha] \rightarrow K$ that sends α to β ; as usual, we can extend σ to an automorphism of \overline{F} and then the restriction of σ to K is an automorphism of K (since K is normal) for which $\sigma(\alpha) \neq \alpha$. Therefore $K^{\text{Aut}(K/F)} = F$.

Let us now revisit the abstract setting of section XX in the somewhat less trivial present framework: $X = K = F^{\text{sep}}$, $G = \text{Aut}(K/F)$. Then the maps $L \mapsto \text{Gal}(K/L)$ and $H \mapsto K^H$ give a bijective correspondence between **closed** subextensions L of K/F and **closed** subgroups H of G . The key fact is the following

Lemma 65. *Every subextension L of K/F is closed, i.e. $K^{\text{Gal}(K/L)} = L$.*

Proof:

8. STRUCTURE OF TRANSCENDENTAL EXTENSIONS

Let K/F be an extension. A finite subset $S = \{x_1, \dots, x_n\} \subset K$ is **algebraically independent** over F if for the only polynomial $P(t_1, \dots, t_n) \in F[t_1, \dots, t_n]$ such that $P(x_1, \dots, x_n) = 0$ is $P = 0$. An arbitrary subset $S \subset K$ is said to be algebraically independent if all of its finite subsets are algebraically independent. (To be absolutely precise, we must impose some ordering on the elements of S in order

to substitute them in as values of an n -variable polynomial, but the definition is obviously independent of the chosen ordering.) We say that K/F is **purely transcendental** if it is of the form $F(S)$ for some algebraically independent subset S of K .

Proposition 66. *Let K/F be an extension and $S = \{x_i\}$ be an ordered set of elements of K . TFAE:*

- (i) *The natural map $\Phi : F[\{t_i\}] \rightarrow K$ given by $t_i \mapsto x_i$ is an injection.*
- (ii) *The map Φ extends uniquely to an isomorphism $F(\{t_i\}) \rightarrow F(S)$.*
- (iii) *S is algebraically independent over F .*

A subset S of K/F is a **transcendence basis** if it is algebraically independent and $K/F(S)$ is algebraic. In other words, a transcendence basis for K/F effects a decomposition of K/F into a tower $K/F(S)/F$ of a purely transcendental extension followed by an algebraic extension.

Example: The empty set is – perhaps by definition – always algebraically independent. If K/F is algebraic, then the only algebraically independent subset is the empty set, which is a transcendence basis.

Lemma 67. *Let K/F be an extension, $S \subset K$ an algebraically independent set, and $x \in K$. Then $S \cup \{x\}$ is algebraically independent iff x is transcendental over $F(S)$.*

Proof: If S is an algebraically independent subset and $x \in K$ is transcendental over $F(S)$, then suppose for a contradiction that $S \cup \{x\}$ were dependent: i.e., there exists finite ordered subset $S_n = (x_1, \dots, x_n)$ of S and a nonzero polynomial $P \in F[t_1, \dots, t_n, t_{n+1}]$ such that $P(x_1, \dots, x_n, x) = 0$. But the transcendence of x over $F(S)$ implies that the polynomial $P(x_1, \dots, x_n, t_{n+1})$ is identically zero, so that the polynomial $Q(t_1, \dots, t_n) := P(t_1, \dots, t_n, 0)$ is not identically zero and $Q(x_1, \dots, x_n) = 0$, contradicting the independence of (x_1, \dots, x_n) . The other direction is even easier.

Corollary 68. *a) An algebraically independent subset S of K is a transcendence basis iff it is maximal, i.e., not properly contained in any other algebraically independent set.*

b) Every algebraically independent subset S of K is contained in a transcendence basis.

Proof: Part a) follows immediately from Lemma XX: a maximal algebraically independent set S is precisely one for which $K/F(S)$ is algebraic, i.e., a transcendence basis. Moreover the union of a chain of algebraically independent sets is algebraically independent, so part b) follows from part a) by Zorn's Lemma.

Note that applying Corollary XX to $S = \emptyset$, we deduce that every field extension K/F admits a transcendence basis.

Exercise X.X.X: Let $\{x_i\}_{i \in S}$ be a transcendence basis for the (nonalgebraic) field extension K/F . Let $n_\bullet : S \rightarrow \mathbb{Z}^+$ be any function. Show that $\{x_i^{n_i}\}$ is also a transcendence basis.

Definition: The **transcendence degree** of a field extension K/F is the minimum cardinality of a transcendence basis. We defer the obvious question – can there exist two transcendence bases of different cardinalities? – until §X.X.

The transcendence degree of an extension is related to $\#K$ and $\#F$ as follows:

Proposition 69. *Let K/F be a transcendental field extension, with transcendence degree κ . Then*

$$\#K = \max(\#F, \kappa, \aleph_0).$$

Proof: Since K/F is transcendental, K is infinite. Moreover, κ and $\#F$ are cardinalities of subsets of K , so clearly $\#K \geq \max(\#F, \kappa, \aleph_0)$. Conversely, let S be a transcendence basis; then $F(S)$ has cardinality $\max(\#, \kappa)$ and $K/F(S)$ is algebraic and $F(S)$ is infinite, so $\#K = \#F(S)$.

8.1. Applications to algebraically closed fields.

Theorem 70. *(Automorphism extension theorem) Let K be an extension of F , with K algebraically closed. Then every automorphism of F can be extended to at least one automorphism of K .*

Proof: Let $\{x_i\}_{i \in S}$ be a transcendence basis for K/F . There is a unique automorphism of $F(S)$ which extends ι and maps each x_i to itself. Since K is the algebraic closure of $F(S)$, by Corollary XX we can further extend to an automorphism of K .

For any field K , let \mathbb{F} be its prime subfield. An **absolute transcendence basis** for K is a transcendence basis for K/F .

Corollary 71. *a) Two algebraically closed fields K_1 and K_2 are isomorphic iff they have the same characteristic and the same absolute transcendence degree.*

b) Suppose K_1, K_2 are two algebraically closed fields of the same characteristic and $\#K_1 = \#K_2$ is uncountable. Then $K_1 \cong K_2$.

Proof: Evidently any pair of isomorphic fields $K_1 \cong K_2$ have the same characteristic and absolute transcendence degree. If K_1 is algebraically closed with prime subfield \mathbb{F} and transcendence degree κ , then for a set S of indeterminates of cardinality κ , then K_1 is isomorphic to the algebraic closure of $\mathbb{F}(S)$, which shows that the characteristic and the absolute transcendence degree determine the isomorphism class of an algebraically closed field. Proposition ?? implies that the absolute transcendence degree of any uncountable field is equal to its cardinality, and part b) then follows immediately from part a).

Remark: The fact that any two algebraically closed fields of given cardinality and, say, continuum cardinality, are isomorphic has important applications in model theory: it shows that the first order theory of algebraically closed fields of a given characteristic is **complete**.

Theorem 72. *Let K/F be an extension of fields, of transcendence degree κ . TFAE:*

(i) For any extension field K' of F with transcendence degree $\kappa' \leq \kappa$, there exists an F -algebra embedding $K' \hookrightarrow K$.

(ii) K is algebraically closed.

Exercise X.X.X: Prove Theorem X.X.

Theorem 73. *Let K be an algebraically closed field. The group $\text{Aut}(K)$ of all automorphisms of K has cardinality $2^{\#K}$.*

Proof: Step 0: Note that $2^{\#K}$ is also the cardinality of the set of all functions from K to K , so is the largest conceivable value of $\#\text{Aut}(K)$.

Step 1: We must check the result for $\overline{\mathbb{F}_p}$ and $\overline{\mathbb{Q}}$. In the former case we have identified the automorphism group as $\hat{\mathbb{Z}}$, which indeed has cardinality $c = 2^{\aleph_0} = 2^{\#\overline{\mathbb{F}_p}}$. In the latter case we can by no means “identify” $\text{Aut}(\overline{\mathbb{Q}})$, but to see that it has continuum cardinality it suffices, by the automorphism extension theorem, to exhibit a simpler Galois extension K/\mathbb{Q} which has continuum cardinality. Indeed one can take K to be quadratic closure of \mathbb{Q} , i.e., the compositum of all quadratic field extensions of \mathbb{Q} . The automorphism group here is $(\mathbb{Z}/2\mathbb{Z})^{\aleph_0} = c$.

Step 2: By the automorphism extension theorem, the cardinality of the automorphism group of any algebraically closed field is at least that of the continuum, which by Step 0 gives the answer for all countable fields, i.e., for all fields of countable absolute transcendence degree.

Step 3: Otherwise K is uncountable so there exists an absolute transcendence basis S with $\#S = \#K$. Now the natural action of $\text{Sym}(S)$ on S gives rise to an injection $\text{Sym}(S) \hookrightarrow \text{Aut}(\mathbb{F}(S))$, i.e., by permutation of indeterminates. By the automorphism extension theorem, this shows that $\#\text{Aut}(K) \geq \#\text{Sym}(S) = 2^{\#S}$.

Corollary 74. *Suppose K/F is an extension with K algebraically closed. Then $K^{\text{Aut}(K/F)}$ is the purely inseparable closure of F in K . In particular, $K^{\text{Aut}(K/F)} = F$ iff F is perfect.*

Proof: If x lies in the purely inseparable closure of F in K , then for some $e \in \mathbb{Z}^+$, $x^{p^e} \in F$. Since x has no Galois conjugates, we must have $\sigma(x) = x$ for every $\sigma \in \text{Aut}(K/F)$. Let \overline{F} be the algebraic closure of F in K . By the usual Galois theory we have $\overline{F}^{\text{Aut}(\overline{F}/F)}$ is the purely inseparable closure of F in \overline{F} , and by the automorphism extension theorem we conclude that $K^{\text{Aut}(K/F)} \cap \overline{F}$ is the purely inseparable closure of F in K . If $x \in K$ is transcendental over F , then by Theorem X.X.X there exists an ordered transcendence basis $S = (x, \{x_\alpha\})$ containing x . By Exercise X.X.X, $S' = (x^2, \{x_\alpha\})$ is also a transcendence basis hence there exists an automorphism $F(S) \rightarrow F(S')$ sending $x \mapsto x^2$, which, as usual, extends to an F -algebra automorphism σ of K with $\sigma(x) = x^2 \neq x$.

Another fact which is true about automorphism groups of algebraically closed field extensions K/F is that any bijection φ between algebraically independent subsets I and I' of K extends to an F -automorphism of F . For this it is necessary and sufficient that φ extend to a bijection on transcendence bases $S \supset I$, $S' \supset I'$. A moment's thought shows that this holds provided that all transcendence bases of K/F have the same cardinality and need not hold otherwise. This brings us to the next section.

8.2. An axiomatic approach to independence.

We wish to prove the following result.

Theorem 75. *Let K/F be a field extension. Then any two transcendence bases for K/F have the same cardinality, so that the transcendence degree of K/F is the cardinality of any transcendence basis.*

Of course this is strikingly similar to the situation in ordinary linear algebra. We could therefore go back to our linear algebra texts, consult the proof of the cardinality independence of bases in vector spaces, and attempt to mimic it in the present context. This approach will succeed. Of course in order to do this we will have to find some sort of precise analogy between linear independence and algebraic independence. In mathematics, once we determine that situations A and B are analogous (to the extent that certain proofs can be carried over from one context to the other), do we just dutifully copy down the similar proofs and keep the analogy in the back of our mind in case we need it later? Depending on taste, this is a reasonable approach to take, perhaps more reasonable for the mind which is able to quickly remember what it once knew. As for myself, I would at the same time worry that it would take me some time and energy to recreate the analogy if I hadn't written it down, and I would also be curious whether A and B might be common instances of a more general construction that it might be interesting or useful to know explicitly. So we shall follow the second course here, with apologies to those with different tastes.

Let us begin by placing alongside the analogies between linear independence of a subset S of an F -vector space V and algebraic independence of a subset S of an F -algebra K .

In both contexts we have a set, say X , and a collection of subsets S of X that we are calling **independent**, subject to:

- (LI1) The empty set is independent.
- (LI2) A set is independent iff all its finite subsets are independent.
- (LI3) Any subset of an independent set is independent.

Notice that it follows from (LI2) and (LI3) that the union $S = \bigcup_i S_i$ of any chain of independent subsets is independent: if not, there would exist a finite dependent subset S' of S , but S' would have to be a subset of some S_i , contradicting the independence of S_i . Combining this with (LI1) and applying Zorn's Lemma, we get

- (A) Maximal independent sets exist, and every independent set is contained in some maximal independent set.

Now could it be that (LI1) through (LI3) also imply the following desirable property?

- (B) All maximal independent sets have the same cardinality.

A moment's thought reveals that this is not the case. Suppose we have a set X which is partitioned into disjoint subsets:

$$X = \coprod_i X_i.$$

Call a subset $S \subset X$ independent iff it is contained in X_i for some i , it is evident that (LI1) through (LI3) are satisfied and the maximal independent sets are simply the X_i 's, which we are evidently not entitled to conclude have the same cardinality.

So we need another axiom. Consider the following:

(LI4) If S_1 and S_2 are independent subsets of X with $\#S_1 < \#S_2$, then there exists $x \in X \setminus S_1$ such that $S_1 \cup \{x\}$ is independent.

A set X equipped with a family of subsets $\{S_i\}$ satisfying axioms (LI1) through (LI4) is called an **independence space**.

In an independence space, if S_1 and S_2 are independent sets with $\#S_1 < \#S_2$, then S_1 is non-maximal. Therefore a maximal independent set has cardinality at least as large as any other independent set, so by symmetry all maximal independent sets have the same cardinality: independence spaces satisfy (B). Conversely, (LI1) through (LI3) and (B) clearly imply (LI4).

In this new language, Theorem 75 takes the form

Theorem 76. *If K/F is a field extension, then the collection of algebraically independent subsets of K is an independence space.*

Unfortunately it is not so obvious how to show that the collection of algebraically independent subsets of K satisfies (LI4). So let us try a different approach, in terms of something called spanning sets. We notice that to each subset S of a vector space its linear span \overline{S} gives an abstract closure operator: namely we have

- (CL1) $S \subset \overline{S}$
- (CL2) $S \subset S' \implies \overline{S} \subset \overline{S'}$
- (CL3) $\overline{\overline{S}} = \overline{S}$.

But the linear span satisfies two other properties, the first of which is not surprising in view of what has come before:

(SO4) if $x \in \overline{S}$, there exists a finite subset $S' \subset S$ such that $x \in \overline{S'}$.

Famously, linear span also satisfies the following **Exchange Lemma**:¹¹

(SO5) If $y \in \overline{S \cup x}$ and y is not in \overline{S} , then $x \in \overline{S \cup y}$.

(Proof: If $y \in \overline{S \cup x}$, there exist $s_1, \dots, s_n \in S$ and scalars a_1, \dots, a_n, a such that $y = a_1 s_1 + \dots + a_n s_n + ax$. If y is not in the span of S , then $a \neq 0$, so

¹¹This is an absolutely prototypical example of a *lemma*: the exchange lemma is the essential kernel of content in the theory of linearly independence, and yet it is itself not very memorable or appealing, so is doomed to be overshadowed by the figurehead theorems that it easily implies.

$$x = y - \frac{-a_1}{a} s_1 + \dots + \frac{-a_n}{a} s_n \in \overline{S \cup y}.)$$

Now, suppose K/F is a field extension and S is a subset of K . We will define \overline{S} to be the algebraic closure of $F(S)$ in K . It is immediate that this “algebraic closure” operator satisfies (SO1) through (SO4). Let us check that it also satisfies (SO5): suppose $y \in \overline{S \cup x}$ and y is not in the algebraic closure of S . Then there exists a finite subset x_1, \dots, x_n of S such that y is algebraic over $F(x_1, \dots, x_n, x)$: i.e., there exists a polynomial $f(t_1, \dots, t_n, t_{n+1}, t_{n+2})$ with F -coefficients such that $f(x_1, \dots, x_n, x, t_{n+2}) \neq 0$ and $f(x_1, \dots, x_n, x, y) = 0$. Writing

$$f(x_1, \dots, x_n, t_{n+1}, t_{n+1}) = \sum_{i=0}^g A_i(x_1, \dots, x_n, t_{n+2}) t_{n+1}^i,$$

observe that not all the polynomials $A_i(x_1, \dots, x_n, t_{n+2})$ can be zero. Since y is not algebraic over $F(S)$, it follows that not all of the elements $A(x_1, \dots, x_n, y)$ are zero, and therefore $f(x_1, \dots, x_n, t_{n+1}, t_{n+1}, y) \neq 0$. Since $f(x_1, \dots, x_n, x, y) = 0$, it follows that x is algebraic over $F(S, y)$ as asserted.

Suppose again that X is any set equipped with a **spanning operator** $S \mapsto \overline{S}$, i.e., an operator satisfying the three closure axioms (CL1) through (CL3) and also (CL4) and (CL5). A subset S of X is a **spanning set** if $\overline{S} = X$. A subset S of X is **independent** if for all $s \in S$, s is not in $\overline{S \setminus s}$. A **basis** is an independent spanning set.

Note that it is immediate to show that the independent sets for a spanning operator satisfy (LI1) through (LI3). In particular, we have (A), that bases exist and any independent set is contained in a basis. Again it is not obvious that (LI4) is satisfied. Rather we will show (B) directly – which is what we really want anyway – and by the above remarks that implies (LI4).

In the following results X is always a set equipped with a spanning operator $S \mapsto \overline{S}$.

Proposition 77. *For a subset $S \subset X$, TFAE:*

- (i) S is a minimal spanning set of X .
- (ii) S is a maximal independent set of X .
- (iii) S is a basis.

Proof: (This is the usual thing.) (i) \implies (iii): Suppose S is minimal spanning but not dependent; then by definition there exists $s \in S$ such that $x \in \overline{S \setminus s}$, so that $\overline{S \setminus s}$, being a closed set containing S , also contains the closure of S , i.e., X , and we found a smaller spanning set. (iii) \implies (ii): if S is a basis and $S \cup \{x\}$ is independent then x does not lie in \overline{S} which is absurd since S is a spanning set. (ii) \implies (i) is similar: if S were a maximal independent set but not a spanning set, then there exists $x \in X \setminus \overline{S}$ and then $S \cup \{x\}$ is independent.

Theorem 78. *Let S be an independent subset of X and T a spanning set. There exists a subset $T' \subset T$ such that $S \cup T'$ is a basis and $S \cap T' = \emptyset$.*

Proof: Let \mathcal{I} be the collection of all subsets T' of T such that $S \cap T' = \emptyset$ and $S \cup T'$ is independent. Observe that $\emptyset \in \mathcal{I}$, so \mathcal{I} is not itself empty. As usual, \mathcal{I} is closed under unions of increasing chains so by Zorn’s Lemma has a maximal

element T' . Let $x \in T$, and suppose that x is not in $\overline{S \cup T'}$. Then $T'' := T' \cup \{x\}$ is a strictly larger subset of T such that $S \cup T''$ is still independent, contradicting the maximality of T' . Therefore

$$X = \overline{T} \supset \overline{\overline{S \cup T'}} = \overline{S \cup T'},$$

so $S \cup T'$ is a basis.

Corollary 79. *If X admits a finite spanning set, it admits a finite basis.*

Proof: Apply Theorem 78 with $S = \emptyset$.

Theorem 80. *Any two bases B, B' of X have the same cardinality.*

Proof: Case 1: Suppose $B = \{x_1, \dots, x_n\}$ is a finite basis, and let B' be any other basis. Let $m = \#B \cap B'$. If $m = n$ then $B \subset B'$ and by Proposition 77 distinct bases are at least incomparable, so $B = B'$. So suppose (WLOG) that $B \cap B' = \{x_1, \dots, x_m\}$ with $m < n$. The set $B \setminus x_{m+1}$ cannot be a spanning set, whereas B' is, so there exists $y \in \overline{B' \setminus x_{m+1}}$. The set $B_1 := (B \setminus x_{m+1}) \cup y$ is independent. By the Exchange Lemma (SO5), $x_{m+1} \in \overline{B_1}$. Hence $B \subset \overline{B_1}$, and since B is a spanning set, so is B_1 . Thus B_1 is a basis. Notice that B_1 has n elements and also $\{x_1, \dots, x_m, y\} \subset B_1 \cap B'$, so that we have replaced B by another basis of the same cardinality and sharing at least one more element with B' . Repeating this procedure will produce a finite sequence of bases B_2, B_3 , each of cardinality n , such that the last basis B_k is contained in, and thus equal to, B' .

Case 2: We may now suppose that B and B' are both infinite. For every $x \in X$, we claim the existence of a subset E_x with the property that $x \in \overline{E_x}$ and for any subset E of B such that $x \in \overline{E}$, $E_x \subset E$. Assuming the claim for the moment, we complete the proof. Consider the subset $S = \bigcup_{x \in B'} E_x$ of B . Since each E_x is finite, $\#S \leq \#B'$. On the other hand, for all $x \in B'$, $x \in \overline{E_x} \subset \overline{S}$, so $B' \subset \overline{S}$ and therefore $\overline{S} \supset \overline{B'} = X$. Therefore S is a spanning subset of the basis B , so $S = B$ and thus $\#B \leq \#B'$. By reversing the roles of B and B' in the argument we conclude $\#B = \#B'$.

It remains to prove the claim on the existence of E_x . In turn we claim that if E' and E'' are two subsets of B such that $x \in \overline{E'} \cap \overline{E''}$ and x is not in the span of any proper subset of E' , then $E' \subset E''$; this certainly suffices. Assuming to the contrary that there exists $y \in E' \setminus E''$. Then x is not in the span of $E' \setminus y$ and is in the span of $(E' \setminus y) \cup y$, so by (SO5) y is in the span of $(E' \setminus y) \cup x$. Since x is in the span of E'' , we get that y is in the span of $(E' \setminus y) \cup E''$. But this contradicts the fact that the $(E' \setminus y) \cup E'' \cup \{y\}$, being a subset of B , is independent.

Remark: A set X endowed with a spanning operator as above is often called a **finitary matroid**. (The word “finitary” refers to (SO4).) Combinatoricists are especially interested in finite matroids, which includes the class of finite-dimensional vector spaces over finite fields but not that of independent subsets of a field extension (except in the trivial case of an algebraic field extension).

We saw above how to go from a finitary matroid to an independence space, namely by decreeing a subset $S \subset X$ to be dependent if there exists $x \in S$ such that $x \in \overline{S \setminus x}$. Conversely, to every independence space we can associate a finitary matroid: define the span \overline{Y} of a subset Y to be the set of $x \in X$ such that $S \cup x$

is dependent. This complete equivalence between concepts of linear independence and spanning seems a bit unexpected, even in the context of vector spaces.

For finite matroids, combinatorialists know at least half a dozen other equivalent axiomatic systems: e.g. in terms of graphs, circuits, “flat” subspaces and projective geometry. As above, demonstrating the equivalence of any two of these systems is not as easy as one might expect. This phenomenon of multiple nonobviously equivalent axiomatizations has been referred to, especially by G. Rota, as **cryptomorphism**. Of course every twenty-first century student of mathematics has encountered cryptomorphism (although it seems that the multiplicity is especially large for finite matroids!). In several essays, Rota saw cryptomorphism as a warning not to take any particular axiomatization of a theory or structure too seriously. This seems fair, but since the different axiomatizations can lead to different and possibly easier proofs, perhaps it should also be viewed as an instance of the inherent richness of mathematical concepts.