MATHEMATICS SURROUNDING APPLICATIONS OF RIEMANN'S EXISTENCE THEOREM

MICHAEL E. ZIEVE

ABSTRACT. In these notes we collect in one place the various background material involved in applications of Riemann's existence theorem to questions about polynomials or curves. This will eventually include material on Riemann surfaces, algebraic topology, Galois theory, group theory, ramification, elliptic curves, rigidity, and other topics.

1. INTRODUCTION

We begin with some algebraic consequences of Riemann's existence theorem, which give group-theoretic descriptions of the possible ramification configurations for Galois extensions of $\mathbb{C}(x)$.

Theorem 1.1. *Pick any $S \subset \mathbb{C} \cup \{\infty\}$. The following properties of a finite group G are equivalent:*

(1) *There is a Galois extension $L/\mathbb{C}(x)$ such that $G = \mathrm{Gal}(L/\mathbb{C}(x))$ and $L/\mathbb{C}(x)$ is unramified over each point outside S; and*
(2) *G can be generated by $\#S - 1$ elements.*

One immediate consequence is a solution to the inverse Galois problem over $\mathbb{C}(x)$: every finite group is the Galois group of an extension of $\mathbb{C}(x)$. Moreover, Theorem 1.1 shows which finite groups occur as Galois groups of extensions of $\mathbb{C}(x)$ whose branch locus is contained in a specified set. The following more precise version allows one to prescribe the ramification type as well as the branch points. (For background material on branch points and inertia groups, see Section 6.)

Theorem 1.2. *Pick any distinct $y_1, \ldots, y_\ell \in \mathbb{C} \cup \{\infty\}$, any finite group G, and any $\sigma_1, \ldots, \sigma_\ell \in G$. The following are equivalent:*

(1) *There is a Galois extension $L/\mathbb{C}(x)$ such that $G = \mathrm{Gal}(L/\mathbb{C}(x))$ and $L/\mathbb{C}(x)$ is unramified over each point outside $\{y_1, \ldots, y_\ell\}$, and moreover for each i there is a point of L lying over y_i whose inertia group in $L/\mathbb{C}(x)$ is generated by σ_i; and*

Date: February 5, 2011.

(2) *there are G-conjugates $\widehat{\sigma}_i$ of σ_i such that $\widehat{\sigma}_1\widehat{\sigma}_2\dots\widehat{\sigma}_\ell = 1$ and G is generated by the $\widehat{\sigma}_i$'s.*

Although the above results are algebraic statements, there is no known algebraic proof of either implication in either result. Instead, every known proof relies on algebraic topology and results on Riemann surfaces. Unfortunately, every known proof that (2) implies (1) is non-constructive: it shows that L exists without showing how to write down such a field L. Still, these results are extremely useful, since knowledge of the possibilities for the ramification data is sufficient for many applications. Sample applications include descriptions of

- indecomposable $f, g \in \mathbb{C}[x]$ such that $f(x) - g(y)$ is reducible
- the curves of a given genus with a given automorphism group.

These descriptions include determinations of the relevant Galois groups, counts of the number of examples with each Galois group, and ramification-theoretic descriptions of the relevant polynomials and curves; different methods are required to actually write down the polynomials and curves explicitly.

These results can be refined to count the number of isomorphism classes of extensions $L/\mathbb{C}(x)$ as above, in terms of group-theoretic data. Moreover, one can generalize the results by replacing the base field $\mathbb{C}(x)$ by any function field (i.e., any finite extension of $\mathbb{C}(t)$).

2. Lüroth's theorem

The following result is known as Lüroth's theorem:

Theorem 2.1. *Let K and L be fields such that $K \subsetneq L \subseteq K(x)$, where x is transcendental over K. Then $L = K(y)$ for some $y \in K(x)$.*

This result follows at once from basic algebraic geometry. For, there is an equivalence of categories between smooth projective curves up to birational equivalence and one-dimensional algebraic function fields up to isomorphism. Via this equivalence, the result translates to the assertion that the image of a rational map from the projective line over K to another curve is again birational to the projective line over K, which is true since such an image is a genus-zero curve (by Riemann–Hurwitz) with a K-rational point (the image of any K-rational point on the projective line).

In Section 2.1 we will give an elementary proof of Lüroth's theorem, which only relies on Gauss's lemma about irreducibility of polynomials in $(K(x))[y]$. This proof provides further information about how to find y in practice; as an application, we will show how to use this extra

information to compute the subfield of $K(x)$ fixed by any finite group of K-automorphisms. However, before reading this proof, the reader may wish to read Section 3, in which we explain the consequences of Lüroth's theorem in our setting. In Section 2.2 we will discuss generalizations and the history of the result.

2.1. **Elementary proof of Lüroth's theorem.** We now give another proof of Lüroth's theorem.

First, note the following 'almost proof' in case K has characteristic zero. Namely, suppose $K \subsetneq L \subseteq K(x)$, and pick any $u \in K(x) \setminus K$. Then $L/K(u)$ is a finite separable extension, so the Primitive Element Theorem implies that $L = (K(u))(v) = K(u, v)$. This shows that L is generated over K by two elements; it is much more difficult to show that a single element will suffice.

The proof below relies only on Gauss's lemma. In our context, this lemma says:

Lemma 2.2. *Let R be a unique factorization domain.*

(1) *If $f \in R[x] \setminus R$ is irreducible in $R[x]$, then f is irreducible in* $\mathrm{Frac}(R)[x]$.
(2) *For $g, h \in R[x]$, we have* $\mathrm{C}(g) \cdot \mathrm{C}(h) = \mathrm{C}(g \cdot h)$.

Here $\mathrm{Frac}(R)$ denotes the field of fractions of R, and $\mathrm{C}(g)$ is the greatest common divisor of the coefficients of g. Since we will make further use of some of the ideas involved in the proof of Gauss's lemma, we recall that proof here.

Proof. For $g, h \in R[x]$, write $g = \mathrm{C}(g) \cdot g_0$ and $h = \mathrm{C}(h) \cdot h_0$ where $g_0, h_0 \in R[x]$ have content 1. If $g_0 h_0$ does not have content 1, let π be an irreducible element of R which divides every coefficient of $g_0 h_0$. Since g_0 and h_0 have content 1, they each have a term whose coefficient is not divisible by π. Write αx^i and βx^j for the highest-degree terms of g_0 and h_0 which are not divisible by π; then π does not occur in the prime factorization of either α or β, so π does not divide $\alpha\beta$. Hence the coefficient of x^{i+j} in $g_0 h_0$ is not divisible by π, a contradiction which proves the second assertion.

Let $f \in R[x]$ be the product of two nonconstant polynomials $g, h \in \mathrm{Frac}(R)[x]$. Writing d_g for the least common multiple of the denominators of the coefficients of g, and defining d_h analogously, it follows that $\widehat{g} := d_g g$ and $\widehat{h} := d_h h$ are nonconstant polynomials in $R[x]$ such that $\widehat{g} \cdot \widehat{h} = d_g d_h f$. Then $\mathrm{C}(\widehat{g}) \cdot \mathrm{C}(\widehat{h}) = d_g d_h \, \mathrm{C}(f)$, so $\tilde{g} := \widehat{g}/\,\mathrm{C}(g)$ and $\tilde{h} := \widehat{h}/\,\mathrm{C}(h)$ are nonconstant polynomials in $R[x]$ such that $\mathrm{C}(f)\tilde{g}\tilde{h} = f$, whence f is irreducible. $\square$

We will apply Gauss's lemma in case $R = K[y]$, where K is a field. We begin with a warmup application. Here, for $u \in K(x)$, we write $\deg(u)$ for the bigger of the degrees of the numerator and denominator of u.

Lemma 2.3. *If K is a field and $u(x) \in K(x) \setminus K$, then $[K(x) : K(u(x))] = \deg(u)$.*

Proof. Write $u(x) = a(x)/b(x)$ where $a, b \in K[x]$ are coprime. Then x is a root of the polynomial $f := a(T) - u(x)b(T) \in K(u(x))[T]$. Since This polynomial has degree one in $u(x)$, so coprimality of $a(T)$ and $b(T)$ implies that f is irreducible in $(K[T])[u(x)]$, which we rewrite as $(K[u(x)])[T]$. Then Gauss's lemma implies that f is irreducible in $(K(u(x)))[T]$, so f is a constant multiple of the minimal polynomial of x over $K(u(x))$, which proves the lemma. □

Proof of Theorem 2.1. Since $L \neq K$, some nonconstant $u \in K(x)$ lies in L, so that $[K(x) : L] \leq [K(x) : K(u)] = \deg(u)$. Thus x is algebraic over L; let $f(t) := a_0 t^n + a_1 t^{n-1} + \cdots + a_n \in L[t]$ be the minimal polynomial for x over L, so that $a_i \in L$ and $a_0 = 1$. Each a_i is in L and hence in $K(x)$. We will show that $L = K(a_j)$ for each j such that $a_j \notin K$ (there always exists such an a_j, since x is transcendental over K).

Let d be the least common multiple of the denominators of the various a_i's, and write $\widehat{f}(t) := d \cdot f(t)$. Then $\widehat{f}(t)$ lies in $K[x][t]$, and is not divisible by any nonconstant polynomial in $K[x]$ (since f is monic). Let m be the x-degree of $\widehat{f}$, so that $m > 0$ and m is the largest degree of any of the polynomials da_i. Say $m = \deg(da_i)$, and write $a_i = b(x)/c(x)$ where $b, c \in K[x]$ are coprime. Thus $m \geq \max(\deg(b(x)), \deg(c(x)))$. Now, $b(t) - a_i c(t)$ is a polynomial in $L[t]$ having root $t = x$, so it is divisible by the minimal polynomial of x over L, namely $f(t)$; say $f(t) \cdot q(t) = b(t) - a_i c(t)$. Multiply by $c(x)$ to get $c(x) \cdot f(t) \cdot q(t) = c(x)b(t) - b(x)c(t)$. Then the right side is not divisible by any nonconstant polynomial in $K[x]$, so by Gauss's lemma we can write $c(x) \cdot f(t) \cdot q(t) = \widehat{f}(t) \cdot \widehat{q}(t)$ with $\widehat{q}(t) \in K[x][t]$. Since $c(x)b(t) - b(x)c(t)$ has x-degree at most $m = \deg_x \widehat{f}(t)$, we must have $\widehat{q}(t) \in K[t]$. But by symmetry, $c(x)b(t) - b(x)c(t)$ has t-degree m and is not divisible by any nonconstant polynomial in $K[t]$, so $\widehat{q} \in K^*$ and thus $\widehat{f}$ is an element of K^* times $c(x)b(t) - b(x)c(t)$; comparing degrees in t shows that $n = m$. Thus $[K(x) : K(a_i)] \leq m = n = [K(x) : L] \leq [K(x) : K(a_i)]$, so $L = K(a_i)$. Finally, if $a_j \notin K$ then $[K(x) : L] \leq [K(x) : K(a_j)] = \deg(a_j) \leq \deg(da_j) \leq \deg(da_i) = m = [K(x) : L]$, so $L = K(a_j)$ as desired. □

We now use the main assertion in this proof to compute the subfield of $K(x)$ fixed by a finite group G of K-automorphisms of $K(x)$. Calling this subfield L, the above proof shows that any nonconstant coefficient of the minimal polynomial of x over L will generate L over K. The roots of this minimal polynomial are precisely the images of x under $\mathrm{Gal}(K(x)/L)$, i.e., the elements $g(x)$ with $g \in G$. Hence this minimal polynomial is $\prod_{g\in G}(T - g(x))$, and its coefficients are (up to multiplication by ± 1) the elementary symmetric polynomials in the values $g(x)$ with $g \in G$. Thus, for any specific G, we simply compute these elementary symmetric polynomials until we find one whose value isn't in K, and then that value y will satisfy $L = K(y)$.

Remark 2.4. The possible groups G are known. If K has characteristic zero, then Klein showed that G is either cyclic, dihedral, A_4, S_4, or A_5. If K has characteristic $p > 0$, then Dickson showed that the only other possibilities for G are $\mathrm{PGL}(2, p^n)$, $\mathrm{PSL}(2, p^n)$, and subgroups of the group of upper-triangular matrices in $PGL(2, p^n)$ (each such group is the semidirect product of an elementary abelian p-group by a cyclic group of order coprime to p). Further, for any K, one knows explicitly all subgroups of $\mathrm{Aut}_K(K(x))$ isomorphic to any of the above groups.

2.2. **Generalizations and historical remarks.** Lüroth proved Lüroth's theorem in case $K = \mathbb{C}$ in 1876. It was first proved for general fields K by Steinitz in 1910, by the above argument.

It is true more generally that if $K \subseteq L \subseteq M$ and M is finitely generated over K, then also L is finitely generated over K. For, it is shown in most introductory algebra textbooks that a maximal K-algebraically independent subset S of L can be extended to a maximal K-algebraically independent subset T of M, and that the minimal number of generators of M/K is at least as large as $\#T$ (here $\#T$ is the *transcendence degree* of M/K). Then L is algebraic over $K(S)$, and $[L : K(S)] = [L.K(T) : K(T)] \leq [M : K(T)] < \infty$.

One can also ask more generally about minimal numbers of generators of finitely-generated extensions. For instance, suppose $K \subsetneq L \subseteq K(x_1, \ldots, x_n)$ where the x_i are algebraically independent over K. If L/K has transcendence degree 1, then $L = K(\alpha)$. This was proved for $K = \mathbb{C}$ by Gordan in 1887, and for arbitrary K by Igusa in 1951. If $\mathbb{C} \subsetneq L \subseteq \mathbb{C}(x_1, \ldots, x_n)$ where $L/\mathbb{C}$ has transcendence degree 2, then $L = \mathbb{C}(\alpha, \beta)$. This was proved by Castelnuovo in 1894. All known proofs are difficult. The result is not true in general for other types of fields K, such as $\mathbb{Q}$ or $\mathbb{R}$. Finally, there are fields L with $\mathbb{C} \subsetneq L \subsetneq \mathbb{C}(x_1, x_2, x_3)$ such that $L/\mathbb{C}$ has transcendence degree 3 but cannot be generated by three elements.

3. Monodromy groups

In this section we apply Lüroth's theorem to build a dictionary between decompositions of a polynomial $f(x)$ and fields between $\mathbb{C}(x)$ and $\mathbb{C}(f)$, which in turn correspond to groups between two associated Galois groups. The relevant Galois group is *not* the Galois group of the polynomial $f(x)$ itself; instead it is defined as follows:

Definition 3.1. The *monodromy group* of $f(x) \in \mathbb{C}[x]$ is the Galois group of $f(x) - t$ over the field $\mathbb{C}(t)$.

Note that the Galois group of $f(x)$ gives information about the roots of $f(x)$, whereas the monodromy group gives information about the inverse mapping to f. The monodromy group is a group of permutations of the set of roots of $f(X) - t$ in an algebraic closure of $\mathbb{C}(t)$; the subgroup fixing x is a one-point stabilizer of the monodromy group.

For $f(x) \in \mathbb{C}(x) \setminus \mathbb{C}$, Lüroth's theorem implies that any field L satisfying $\mathbb{C}(f(x)) \subseteq L \subseteq \mathbb{C}(x)$ must have the form $L = \mathbb{C}(h(x))$ for some $h(x) \in \mathbb{C}(x) \setminus \mathbb{C}$. Since $f(x) \in \mathbb{C}(h(x))$, it follows that $f = g \circ h$ for some $g \in \mathbb{C}(x)$. Convesely, if $f = g \circ h$ then $\mathbb{C}(h(x))$ lies between $\mathbb{C}(f(x))$ and $\mathbb{C}(x)$. This is not quite a bijection between intermediate fields and decompositions, since there can be more than one choice of h corresponding to a single field L. This issue is resolved in the following result, which is an immediate consequence of Lemma 2.3:

Lemma 3.2. *For $h_1, h_2 \in \mathbb{C}(x) \setminus \mathbb{C}$, the fields $\mathbb{C}(h_1(x))$ and $\mathbb{C}(h_2(x))$ are identical if and only if $h_1 = \mu \circ h_2$ for some degree-one $\mu \in \mathbb{C}(x)$.*

This result motivates the following definition.

Definition 3.3. For $f \in \mathbb{C}(x)$, a *decomposition* of f is an expression $f = g \circ h$ with $g, h \in \mathbb{C}(x)$. The decomposition is *nontrivial* if both g and h have degree at least 2. Finally, two decompositions $f = g \circ h$ and $f = g_1 \circ h_1$ are *equivalent* if there is a degree-one $\mu \in \mathbb{C}(x)$ such that $g \circ \mu = g_1$ and $h = \mu \circ h_1$.

Theorem 3.4. *Pick $f \in \mathbb{C}(x) \setminus \mathbb{C}$, let G be the monodromy group of f, and let Ω be the splitting field of $f(x) - t$ over $\mathbb{C}(t)$. The maps $\rho\colon (g \circ h) \mapsto \mathbb{C}(h)$ and $\phi\colon L \mapsto \mathrm{Gal}(\Omega/L)$ define bijections between*

1. *the set of equivalence classes of decompositions of f,*
2. *the set of fields lying between $\mathbb{C}(f(x))$ and $\mathbb{C}(x)$, and*
3. *the set of subgroups of G.*

Moreover, if $f = g \circ h$ then $\deg(h) = [\mathbb{C}(x) : \mathbb{C}(h)] = [G : \mathrm{Gal}(\Omega/\mathbb{C}(h))]$.

We now show that any decomposition of a polynomial (as a composition of rational functions) is equivalent to a decomposition involving only polynomials.

Lemma 3.5. *If $g, h \in \mathbb{C}(x) \setminus \mathbb{C}$ satisfy $g \circ h \in \mathbb{C}[x]$, then there exists a degree-one $\mu \in \mathbb{C}(x)$ for which both $g \circ \mu^{-1}$ and $\mu \circ h$ lie in $\mathbb{C}[x]$. (Here μ^{-1} is the functional inverse of μ, hence is itself a degree-one rational function.)*

Proof. Note that polynomials are precisely the rational functions under which the unique preimage of infinity is infinity. Write $c := h(\infty)$. Then ∞ is the unique h-preimage of c, and c is the unique g-preimage of ∞. Thus, the result holds whenever μ satisfies $\mu(c) = \infty$. If $c = \infty$ we can use $\mu = x$; otherwise, we can use $\mu = 1/(x - c)$. □

4. Translating properties of polynomials into properties of monodromy groups

Indecomposability of a polynomial translates to primitivity of the monodromy group. Irreducibility of $(f(x) - f(y))/(x - y)$ translates to double transitivity of the monodromy group. More generally, the degrees of the factors of $f(x) - f(y)$ are the *subdegrees* of the monodromy group, namely, the lengths of the orbits of a one-point stabilizer.

5. Permutation groups

Equivalence between three notions of primitive groups. Normal subgroups of primitive groups are transitive. Any minimal normal subgroup of a finite group is a power of a simple group. A primitive group has at most two minimal normal subgroups. Description of the centralizer of the subgroup of a primitive group generated by its minimal normal subgroups. O'Nan–Scott theorem. Classification of doubly transitive groups. Primitive groups with an n-cycle which aren't doubly transitive are contained in $\mathrm{AGL}(1, n)$ (and only exist when n is prime). Classification of primitive groups with an n-cycle. Classification of monodromy groups of indecomposable polynomials.

6. Decomposition and inertia groups, and the genus formula

Define places, decomposition groups, and inertia groups. Then describe the construction of inertia groups via completions.

Pick $f \in \mathbb{C}[x] \setminus \mathbb{C}$, let Ω be the splitting field of $f(x) - t$ over $\mathbb{C}(t)$, and let G be the monodromy group of f. Then, for any place P of Ω which lies over the infinite place of $\mathbb{C}(t)$, the inertia group of P in $\Omega/\mathbb{C}(t)$ is cyclic and transitive (and hence is generated by a cycle of length $\deg(f)$). This is a huge constraint on G.

Pick any $t_0 \in \mathbb{C}$, and write $f(x) - t_0 = \prod_{i=1}^{k} p_i(x)^{e_i}$ where the p_i are pairwise coprime degree-one polynomials in $\mathbb{C}[x]$, and the e_i are positive integers. Then, for any place P of Ω which lies over the place $t = t_0$, the inertia group of P in $\Omega/\mathbb{C}(t)$ is cyclic, and its orbits have lengths $e_1, e_2, \ldots, e_k$; thus, each generator of this inertia group is the product of disjoint cycles of lengths $e_1, e_2, \ldots, e_k$.

Riemann–Hurwitz.

7. More Galois theory

Fried's description of the factors of $f(x) - g(y)$. Etc.

8. Topological vesion of Riemann's existence theorem

9. Analytic version of Riemann's existence theorem

10. Algebraic version of Riemann's existence theorem

11. Rigidity and fields of definition

12. Elliptic curves

13. Heights and absolute values

Moriwaki's height function on $\mathbb{Q}(x)$ which extends the standard height on $\mathbb{Q}$, and which also satisfies the analogue of Northcott's theorem (only finitely many elements have height less than any prescribed bound). Connection with absolute values and local height functions.

References

[1] J. D. Dixon and B. Mortimer, Permutation Groups, Springer-Verlag, New York, 1996.

[2] M. Kuga, *Galois' dream: group theory and differential equations*, Birkhäuser Boston, Inc., Boston, MA, 1993.

[3] R. Lidl, G. L. Mullen and G. Turnwald, Dickson Polynomials, John Wiley & Sons, Inc., New York, 1993.

[4] P. Müller, *Primitive monodromy groups of polynomials*, in: Recent Developments in the Inverse Galois Problem 385–401, Amer. Math. Soc., Providence, RI, 1995.

[5] H. Stichtenoth, Algebraic Function Fields and Codes (second edition), Springer-Verlag, Berlin, 2009.

[6] H. Völklein, Groups as Galois Groups: an introduction, Cambridge University Press, New York, 1996.

[7] H. Wielandt, Finite Permutation Groups, Academic Press, New York-London, 1964.

E-mail address: zieve@umich.edu
URL: www.math.lsa.umich.edu/∼zieve/