Minicourse: Lecture 4
Applying Topology to Spaces of Countable Structures

Russell Miller

Queens College & CUNY Graduate Center

DDC Program, Part I: Virtual Semester

Mathematical Sciences Research Institute
Berkeley, CA (remotely)
Autumn 2020

Plan of the Minicourse

Week 1: Specific example: subrings of $\mathbb{Q}$.
Online discussion: Thursday, Sept. 24, 11:00 PDT.

Week 2: Computability and continuity.
Online discussion: Thursday, Oct. 1, 11:00 PDT.

Week 3: Classifications of spaces of structures.
Online discussion: Thursday, Oct. 8, 11:00 PDT.

Week 4: The space of algebraic fields.
Online discussion: Thursday, Oct. 15, 11:00 PDT.

Week 5: Other related questions.
Online discussion: Thursday, Oct. 22, 11:00 PDT.

(Also watch Caleb Springer's MSRI Junior Seminar: Oct. 20, 09:00.)

Algebraic field extensions of $\mathbb{Q}$

Now we consider $\mathfrak{Alg}_{\mathbb{Q}}$, the set of all algebraic field extensions of $\mathbb{Q}$. We often say "subfields of $\overline{\mathbb{Q}}$" synonymously, where $\overline{\mathbb{Q}}$ is the algebraic closure of $\mathbb{Q}$. However, the meaning is that this is a set of isomorphism types. Many distinct subfields of $\overline{\mathbb{Q}}$ are isomorphic, and in $\mathfrak{Alg}_{\mathbb{Q}}$ such subfields are identified.

The process of creating a topology is the same as in Week 3: consider all presentations of such fields on the domain $\mathbb{N}$, as a subspace of Cantor space $2^{\mathbb{N}}$, and mod out by isomorphism. However, we must decide what signature to use.

Indexing of algebraic fields

The isomorphism problem for $\mathfrak{Alg}_{\mathbb{Q}}$ is Π^0_2 :

Lemma

For all algebraic fields K_0 and K_1 of characteristic 0,

$$K_0 \cong K_1 \iff (\forall f \in \mathbb{Q}[X]) \, [f \text{ has a root in } K_0 \iff f \text{ has a root in } K_1].$$

This suggests an indexing for these fields. Fix a computable list $f_0, f_1, \ldots$ of all monic irreducible polynomials in $\mathbb{Q}[X]$, and define

$$I_K = \{n \in \mathbb{N} : f_n \text{ has a root in } K\} \in 2^{\mathbb{N}}.$$

By the Lemma, $K_0 \cong K_1 \iff I_{K_0} = I_{K_1}$.

Indexing of algebraic fields

The isomorphism problem for $\mathfrak{Alg}_{\mathbb{Q}}$ is Π^0_2 :

Lemma

For all algebraic fields K_0 and K_1 of characteristic 0,

$$K_0 \cong K_1 \iff (\forall f \in \mathbb{Q}[X])\,[f \text{ has a root in } K_0 \iff f \text{ has a root in } K_1].$$

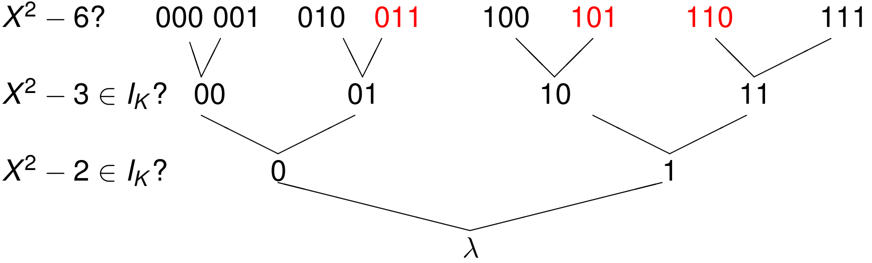
This suggests an indexing for these fields. Fix a computable list $f_0, f_1, \ldots$ of all monic irreducible polynomials in $\mathbb{Q}[X]$, and define

$$I_K = \{n \in \mathbb{N} : f_n \text{ has a root in } K\} \in 2^{\mathbb{N}}.$$

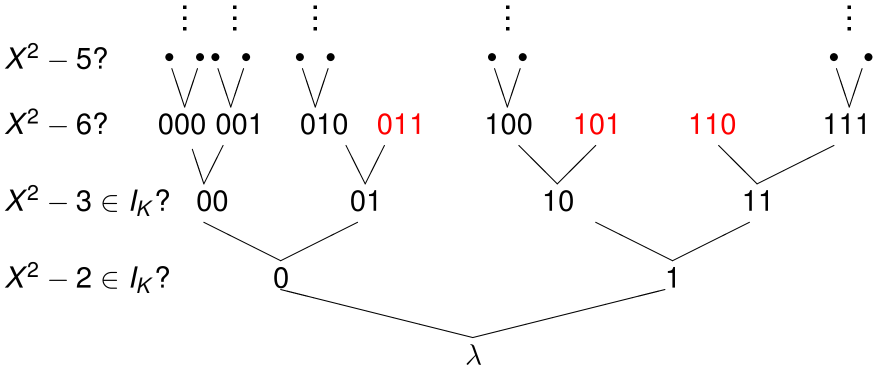
By the Lemma, $K_0 \cong K_1 \iff I_{K_0} = I_{K_1}$.

But beware! Not all $I \in 2^{\mathbb{N}}$ are indices of algebraic fields this way. For example, I might indicate that $X^4 - 2$ has a root, but that $X^2 - 2$ does not. Clearly no field K has such an I as its index I_K.

Picture of $\{I_K : K \subseteq \overline{\mathbb{Q}}\} \subset 2^{\mathbb{N}}$

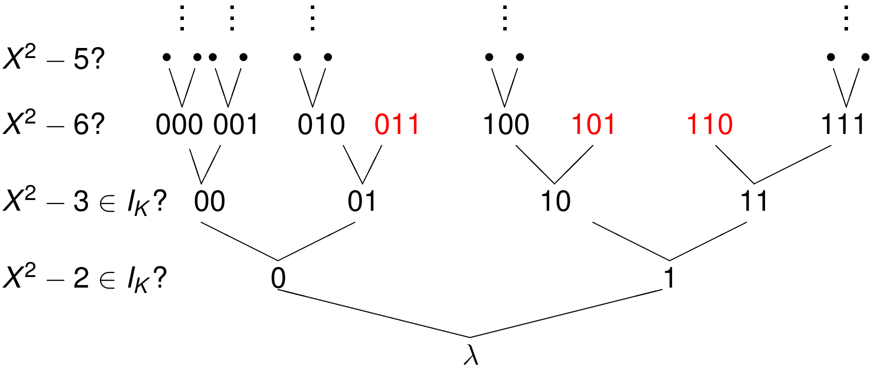


Picture of $\{I_K : K \subseteq \overline{\mathbb{Q}}\} \subset 2^{\mathbb{N}}$



With the red nodes eliminated, there will be no terminal nodes and no isolated paths, and the paths through this tree will be precisely the indices I_K. So they form a subspace homeomorphic to Cantor space.

Picture of $\{I_K : K \subseteq \overline{\mathbb{Q}}\} \subset 2^{\mathbb{N}}$



With the red nodes eliminated, there will be no terminal nodes and no isolated paths, and the paths through this tree will be precisely the indices I_K. So they form a subspace homeomorphic to Cantor space.

Question: Is it decidable which nodes are red?

Working towards I_K

We face two questions.

1. Is $\{I_K \in 2^{\mathbb{N}} : K \in \mathfrak{Alg}_{\mathbb{Q}}\}$?, with the subspace topology, computably homeomorphic to Cantor space?
2. Can we compute I_K from a presentation of K, and vice versa?

For (2), the function $K \mapsto I_K$ is not continuous, just as $\mathbb{Z}[W^{-1}] \mapsto W$ was discontinuous on $\mathfrak{R}_{\mathbb{Q}}$ without $\mathbb{I}$ in the signature. It becomes continuous, with continuous inverse, if we adjoin d-ary *root predicates* R_d to the signature, for all $d > 1$:

$$R_d(a_0, \ldots, a_{d-1}) \iff (\exists x)\, x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0 = 0.$$

With these R_d in the atomic diagram, one can recognize when f_n has no root in K, as well as when it has a root.

Presenting K using I_K

This is not as simple as one thinks. As an example: Suppose I_K says that K contains elements x, y with $x^8 = 2 = y^{12}$. Then $\left(\frac{x^2}{y^3}\right)^4 = 1$, but this does not specify whether $x^2 = \pm y^3$ or $x^2 = \pm i y^3$: either is possible, and the resulting fields $\mathbb{Q}(x, y)$ do not embed into each other.

Presenting K using I_K

This is not as simple as one thinks. As an example: Suppose I_K says that K contains elements x, y with $x^8 = 2 = y^{12}$. Then $\left(\frac{x^2}{y^3}\right)^4 = 1$, but this does not specify whether $x^2 = \pm y^3$ or $x^2 = \pm i y^3$: either is possible, and the resulting fields $\mathbb{Q}(x, y)$ do not embed into each other.

The solution is to find a primitive generator for each of the two possibilities, then determine the minimal polynomial over $\mathbb{Q}$ of each generator, and check I_K to see which of the two minimal polynomials has a root in K.

This requires certain tools.....

Tools we need: Kronecker's Theorem

The *splitting set* S_L of a countable field L is the set of reducible polynomials in $L[X]$. S_L computes which $f \in L[X]$ have roots in L.

Kronecker's Theorem (1882)

- $S_{\mathbb{Q}}$ is decidable.
- If t is transcendental over L within a larger field E, then $S_{L(t)} \leq_T S_L \oplus \Delta(E)$, uniformly.
- If x is algebraic over L within E, then $S_{L(x)} \leq_T S_L \oplus \Delta(E)$, uniformly in the minimal polynomial of x over L.

The algorithms for transcendental and algebraic elements are distinct.

Viz. H.M. Edwards, *Galois Theory* (Springer GTM 101, 1984) §§ 55-60.

This allows you to decide, e.g., whether a given $f(Y)$ has a root in a given number field F – or whether $f(Y)$ would acquire a root when F is extended to $F(x) = F[X]/(g(X))$.

Computable homeomorphism

Kronecker's Theorem shows that the set of red nodes is decidable:

- A node $\sigma 0$ is red iff the field built up to node σ already has a root of the relevant polynomial $f_{|\sigma|}$. (E.g., $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ already contains the square roots of 6.)
- A node $\sigma 1$ is red iff adjoining a root of the relevant polynomial $f_{|\sigma|}$ would generate a root of some earlier f_m with $\sigma(m) = 0$. E.g., for the node 011, adjoining $\sqrt{6}$ to $\mathbb{Q}(\sqrt{3})$ would result in a field containing $\sqrt{2}$, which was ruled out by the "0" in 011.

So the homeomorphism between $2^{\mathbb{N}}$ and the set of indices I_K is computable (in both directions). Each index I_K corresponds to a unique $J_K \in 2^{\mathbb{N}}$, and every $J \in 2^{\mathbb{N}}$ is J_K for some $K \in \mathfrak{Alg}_{\mathbb{Q}}$.

Tools we need: Primitive Element Theorem

Effective Theorem of the Primitive Element

For every finite algebraic field extension E/K, there is a single $x \in E$ such that $E = K(x)$. Moreover, x may be found effectively, uniformly in $\Delta(K)$ and in generators $x_1, \ldots, x_k$ for E over K. (So may its minimal polynomial over K, provided that S_K is decidable.)

To describe the generators, we give polynomials $g_i \in K[X_1, \ldots, X_i]$ such that each $g_i(x_1, \ldots, x_{i-1}, X_i)$ is the minimal polynomial of x_i over $K(x_1, \ldots, x_{i-1})$.

Finding x can be a blind search, since we know it exists. For a more efficient algorithm: Fried & Jarden, *Field Arithmetic* (Springer, 1986).

Homework problem!

We discussed the space $\mathfrak{FBT}$ of (infinite) finite-branching trees in Week 3. With branching predicates in the signature, it is homeomorphic to Baire space $\mathbb{N}^{\mathbb{N}}$.

$\mathfrak{FBT}$ is very similar to $\mathfrak{Alg}_{\mathbb{Q}}$. Once again, the isomorphism problem is Π^0_2: two trees are isomorphic iff every finite subtree of each tree embeds into the other tree.

For $T \in \mathfrak{FBT}$, let $I_T = \{n \in \mathbb{N} : S_n \hookrightarrow T\}$ be the set of finite trees that embed into T. So $T \cong T' \iff I_T = I_{T'}$. With branching predicates in the signature, I_T is computable from T, and isomorphism becomes Π^0_1.

Homework problem!

We discussed the space $\mathfrak{FBT}$ of (infinite) finite-branching trees in Week 3. With branching predicates in the signature, it is homeomorphic to Baire space $\mathbb{N}^{\mathbb{N}}$.

$\mathfrak{FBT}$ is very similar to $\mathfrak{Alg}_{\mathbb{Q}}$. Once again, the isomorphism problem is Π^0_2: two trees are isomorphic iff every finite subtree of each tree embeds into the other tree.

For $T \in \mathfrak{FBT}$, let $I_T = \{n \in \mathbb{N} : S_n \hookrightarrow T\}$ be the set of finite trees that embed into T. So $T \cong T' \iff I_T = I_{T'}$. With branching predicates in the signature, I_T is computable from T, and isomorphism becomes Π^0_1.

So why is $\mathfrak{FBT}$, with branching, not homeomorphic to Cantor space? Where does the argument for $\mathfrak{Alg}_{\mathbb{Q}}$ break down on $\mathfrak{FBT}$?

We will discuss this in the discussion section on October 15!

Basic open sets

$\mathfrak{Alg}_{\mathbb{Q}}$ is more difficult to present than $\mathfrak{R}_{\mathbb{Q}}$ was. Another version of the homeomorphism onto $2^{\mathbb{N}}$ appears in:
Miller, Isomorphism and classification for countable structures, *Computability* **8** (2019) 2, 99–117, DOI 10.3233/COM-180095.

But the simplest version is probably:
For each number field F and each $h \in \mathbb{Q}[X]$ with no roots in F, let

$$\mathcal{U}_{F,h} = \{\text{algebraic fields } K \supseteq \mathbb{Q} : F \hookrightarrow K \ \& \ h \text{ has no roots in } K\}.$$

So F is the "positive" information (essentially finite, since F is a number field) about K, and h is the "negative" information.
(These are the basic open sets used in current joint work with Eisenträger, Springer, and Westrick.)

Indices and presentations

For an isomorphism type K in $\mathfrak{Alg}_{\mathbb{Q}}$, we have defined the index

$$I_K = \{n \in \mathbb{N} : f_n \text{ has a root in } K\}.$$

The set of all indices maps homeomorphically onto Cantor space, and J_K is the image of I_K there, with $I_K \equiv_T J_K$ uniformly.
A *presentation* of K is a field L, isomorphic to K, whose domain is $\mathbb{N}$.
The *atomic diagram* of L is basically the addition and multiplication tables for L, coded as a subset of $\mathbb{N}$. For each presentation L of K, we have the following sets, all $\Delta(L)$-computably enumerable:

$$S_L = \{h \in L[X] : h \text{ factors in } L[X]\} \quad \text{(the splitting set of } L).$$

$$R_L = \{h \in L[X] : h \text{ has a root in } L\} \quad \text{(the \textit{root set of} L)}.$$

$$HTP(L) = \{h \in L[X_1, X_2, \ldots] : h = 0 \text{ has a solution in } L\}$$

The relationships among these, relative to $\Delta(L)$, for each L of type K:

$$I_K \equiv_T J_K \equiv_T R_L \equiv_T S_L \leq_T HTP(L) \leq_T (\Delta(L))'.$$

An application

With $\mathfrak{Alg}_{\mathbb{Q}}$ homeomorphic to Cantor space, we can now use the notions of Baire category to investigate the prevalence of various properties of algebraic fields. A sample result is the following, describing the difficulty of computing the root set R_F of a field from a presentation F of the field (in the signature $(+,\cdot)$). It is well known that R_F is always c.e. in $\Delta(F)$, but can fail to be computable from $\Delta(F)$.

Theorem (M., 2020 CiE Proceedings, LNCS 12098)

These two classes of algebraic fields are both co-meager in $\mathfrak{Alg}_{\mathbb{Q}}$.

- $\{K \in \mathfrak{Alg}_{\mathbb{Q}}$: some presentation L of K has $R_L \not\leq_T \Delta(L)\}$.
  (Direct proof here is joint with Eisenträger-Springer-Westrick.)
- $\{K \in \mathfrak{Alg}_{\mathbb{Q}}$: every presentation L of K has $(R_L)' \leq_T (\Delta(L))'\}$.
  (That is, R_L is low relative to every presentation L.)

To prove the second item, we relativize to $\Delta(L)$, proving $(J_K)' \leq_T (\Delta(L))'$ and then invoking $J_K \oplus \Delta(L) \equiv_T R_L \oplus \Delta(L)$.

Procedure to show $(J_K)' \leq_T (\Delta(L))'$

We need to decide whether a given oracle Turing program Φ, when run with oracle J_K, halts on a given input e. To help us decide, we have our own oracle $(\Delta(L))'$.

Procedure to show $(J_K)' \leq_T (\Delta(L))'$

We need to decide whether a given oracle Turing program Φ, when run with oracle J_K, halts on a given input e. To help us decide, we have our own oracle $(\Delta(L))'$.

With $(\Delta(L))'$, we can compute J_K, so we can run $\Phi^{J_K}(e)$. If we ever see it halt, we have our answer.

Procedure to show $(J_K)' \leq_T (\Delta(L))'$

We need to decide whether a given oracle Turing program Φ, when run with oracle J_K, halts on a given input e. To help us decide, we have our own oracle $(\Delta(L))'$.

With $(\Delta(L))'$, we can compute J_K, so we can run $\Phi^{J_K}(e)$. If we ever see it halt, we have our answer.

For each initial segment σ of J_K, one by one, we ask $(\Delta(L))'$ whether

$$(\forall \tau \supseteq \sigma)\ [\Phi^{\tau}(e) \text{ does not halt within } |\tau| \text{ steps}].$$

If we ever find such a σ, then we know that $\Phi^{J_K}(e)$ never halts. (If it halted, some initial segment τ of J_K would contradict the above.)

This is our decision procedure. It does not always give an answer, but we claim that, for some comeager set of indices J_K, it answers correctly for all programs Φ and inputs e, on all presentations L of K.

The procedure works on a comeager set

All answers given by our procedure are correct, so if it messes up, let Φ and e be a program and input for which it never gives an answer. Then $\Phi^{J_K}(e)$ never halts, but for every initial segment σ of J_K, there is a $\tau \supseteq \sigma$ that would make it halt.

Now consider $\mathcal{U}_\sigma = \{E \in \mathfrak{Alg}_{\mathbb{Q}} : \sigma \subset J_E\}$ (for any σ). If there exists some $K \in \mathcal{U}_\sigma$ with a presentation L on which our procedure never halts for this Φ and e, then there is some $\tau \supseteq \sigma$ that would make the procedure halt (on this Φ and e). This means that, for every $E \in \mathcal{U}_\tau$, the procedure gives the correct answer on this Φ and e and on every presentation L of E. (Arbitrarily much of $\Delta(L)$ might be required to compute $E{\restriction}|\tau|$ from $\Delta(L) \oplus R_L$, depending on the presentation L.) So

$$\{E : \exists \text{ a presentation } L \text{ of } E \text{ s.t. the procedure never halts on } \Phi \text{ \& } e\}$$

is a nowhere dense set: it is not dense within $\mathcal{U}_\sigma$, because it contains no element of $\mathcal{U}_\tau$.

But the countable union of these nowhere dense sets, across all Φ and e, is meager and contains all J_K for which our procedure messes up.

Presentations L with $R_L \not\leq_T \Delta(L)$

If every presentation of K computes I_K, then they all enumerate $\overline{I_K}$. By a theorem of Knight from 1986, $\overline{I_K} \leq_e \Sigma_1$-Th$(K)$, which in turn is $\leq_e I_K$. But we claim that each e-reduction succeeds in reducing $\overline{I_K}$ to I_K only for a nowhere dense set of fields K.

Suppose a single e-reduction gives $\overline{I_K} \leq_e I_K$ for certain fields $K \in \mathcal{U}_{F,h}$. Since F is a number field, we can fix a prime $p > \max([F : \mathbb{Q}], \deg(h))$. Then h has no root in $F(\sqrt[p]{2})$, and F contains no p-th root of 2, so we consider $\mathcal{U}_{F,h\cdot(Y^p-2)}$. Assume the enumeration reduction works for some K here. Now $Y^p - 2$ lies in $\overline{I_K}$, so let $K_0 \subseteq K$ be a number field extending F, with enough elements that the enumeration reduction on I_{K_0} says that $Y^p - 2 \in \overline{I_{K_0}}$.

Presentations L with $R_L \not\leq_T \Delta(L)$, continued

Now h has no root in K_0 since $K_0 \subseteq K$; and with $K_0(\sqrt[p]{2})$ minimal over K_0 of prime degree $p > \deg(h)$, h can have no root there either. So $\mathcal{U}_{K_0(\sqrt[p]{2}),h}$ is a basic open set within $\mathcal{U}_{F,h}$.

But every $E \in \mathcal{U}_{K_0(\sqrt[p]{2}),h}$ will have $K_0 \subseteq E$, so running the e-reduction on a presentation of E will say that $\sqrt[p]{2} \notin E$, which is wrong. Thus the e-reduction fails on an entire open subset of $\mathcal{U}_{F,h}$, and so the set where it succeeds is not a dense subset of $\mathcal{U}_{F,h}$. Since F and h were arbitrary, this enumeration reduction succeeds only on a nowhere dense set.

Thus, if every presentation of K computes I_K, then K lies in the union of these countably many nowhere dense sets (one set for each e-reduction). So co-meager-many K have a presentation L for which $\Delta(L)$ does not compute I_K. But if $\Delta(L)$ computed R_L, then it would compute I_K. So we are done.