Perfect set theorems for closed and analytic sets

Jason Chen

Perfect set theorems are theorems of the following kind:

Theorem Template. *Let X be some kind of set, then one of the following will hold:*

1. *X has a perfect subset (and then possibly something about this perfect subset)*
2. *X is tractable in some way*

This note contains the perfect set theorems for closed and analytic ($\underset{\sim}{\Sigma}_1^1$) sets in the Baire space.

1 Perfect set theorem for closed sets

Definition 1.1. A closed set $X \subseteq \omega^\omega$ is the body $[T]$ of a tree $T \subseteq \omega^{<\omega}$.

The proof of the perfect set theorem demonstrates a common technique that has many different variants for later use. The technique involves eliminating isolated elements, which will remind the analysts of the Cantor-Bendixson derivative. Let us first define an increasing sequence of isolated nodes in a tree.

Definition 1.2. Let T be a tree $\subseteq \omega^{<\omega}$. We say a node s is *isolated in T* iff it is an element of T and doesn't split below (i.e., there are no extensions $\sigma_0, \sigma_1 \succ s$ in T that are incompatible).

Definition 1.3. Given a tree $T \subseteq \omega^{<\omega}$, define an increasing sequence of sets T_α as follows:

$$T_0 = \emptyset$$
$$T_\lambda = \bigcup_{\beta<\lambda} T_\beta \text{ if } \lambda \text{ is a limit ordinal}$$
$$T_{\alpha+1} = T_\alpha \cup \{s \in T \mid s \text{ is isolated in } T \setminus T_\alpha\}$$

Lemma 1.4. *If $T \subseteq \omega^{<\omega}$ is a tree, then there is some countable ordinal δ when $T_\delta = T_{\delta+1}$.*

Proof. This is because, if $T_0 \subsetneq T_1 \subsetneq T_2 \subseteq ... \subsetneq T_\alpha \subsetneq T_{\alpha+1} \subsetneq ...$ lasts $\geq \omega_1$ many steps, we would have collected an uncounable subset of T, which has at most countably many elements. □

So for any tree T, there will be a least ordinal countable ordinal δ for which $T_\delta = T_{\delta+1}$.

Theorem 1.5 (The case where $T \neq T_\delta$). *For any T and δ as above, if $T \neq T_\delta$, then there is an injection from 2^ω into $[T]$.*

Proof. Almost by definition, $T \setminus T_\delta$ is the set of nodes in T that will always split. So we can map $2^{<\omega}$ into T by mapping splits to splits recursively. More concretely, set $f(\langle\rangle) = \langle\rangle$, and if $f(s) \in T \setminus T_\delta$ is defined, then by assumption there are at least two incompatible extensions $f(s)^\frown m, f(s)^\frown n$ below it. Pick the least such pair $m < n$ and map $s^\frown 0$ to $f(s)^\frown m$ and $s^\frown 1$ to $f(s)^\frown n$. (Visually, we are "stretching" the infinite binary tree to "fit" its splits to those on $T \setminus T_\delta$). Finally for $x \in 2^\omega$ set $F(x) = \bigcup_{n\in\omega} f(x \upharpoonright n)$. □

On the other hand, if $T = T_\delta$, then this provides a fertile ground for an effective analysis of T. First, notice that this would imply that each node in T becomes isolated at some point, and then gets collected at the next stage. This is just another way of saying every node of T is isolated in some $T \setminus T_\alpha$ for $\alpha < \delta$, and then it gets in $T_{\alpha+1}$.

Definition 1.6. For each node $s \in T$, call the unique ordinal α where this happens its isolation rank (i.e., least α such that $s \in T_{\alpha+1} \setminus T_\alpha$), written $\rho(s)$.

A few things to notice:

- Nodes with smaller isolation ranks get picked up by T_α earlier in the process.
- The empty sequence always has the maximum isolation rank, because it always gets picked up last.
- If $s \prec t$, then $\rho(s) \geq \rho(t)$.

Observation 1.7. If $x \in [T]$, then the isolation ranks of $x \upharpoonright 0, x \upharpoonright 1, x \upharpoonright 2, ...$ is a non-increasing sequence of ordinals. So this sequence of ordinals must be eventually constant. Since it won't cause any confusion, we also call this eventual ordinal constant the isolation rank ρ_x of x.

Theorem 1.8 (The case where $T = T_\delta$). *For a tree T with $T = T_\delta$ (recall the notations above), if $x \in [T]$ has isolation rank $\rho_x < \delta$, as witnessed by $x \upharpoonright m$, then x is definable from $T, \rho_x, x \upharpoonright m$*

Proof. To say x has isolation rank ρ_x (witnessed by $x \upharpoonright m$) is to say that, in $T \setminus T_{\rho_x}$, the only extensions to $x \upharpoonright m$ are $x \upharpoonright (m+1), x \upharpoonright (m+2), x \upharpoonright (m+3), ...$ But then it is now easy to define x: set $x(i) = j$ iff $\begin{cases} (x \upharpoonright m)(i) = j & i < m \\ (\exists q \in T \setminus T_{\rho_x})(x \upharpoonright m \prec q \wedge q(i) = j) & i \geq m \end{cases}$

In other words, to "compute" $x(i)$ from T and ρ_x, we only need to build T_{ρ_x} and remove it from T, and then just brute-force search through finite sequences of natural numbers in that tree to see if any extends $x \upharpoonright m$. By the assumptions that $x \in [T]$ and $x \upharpoonright m$ is isolated in $T \setminus T_{\rho_x}$, for each length $k \geq m$ a unique extension exists. Moreover, all the extensions cohere (meaning their union is a real number, in this case x).

□

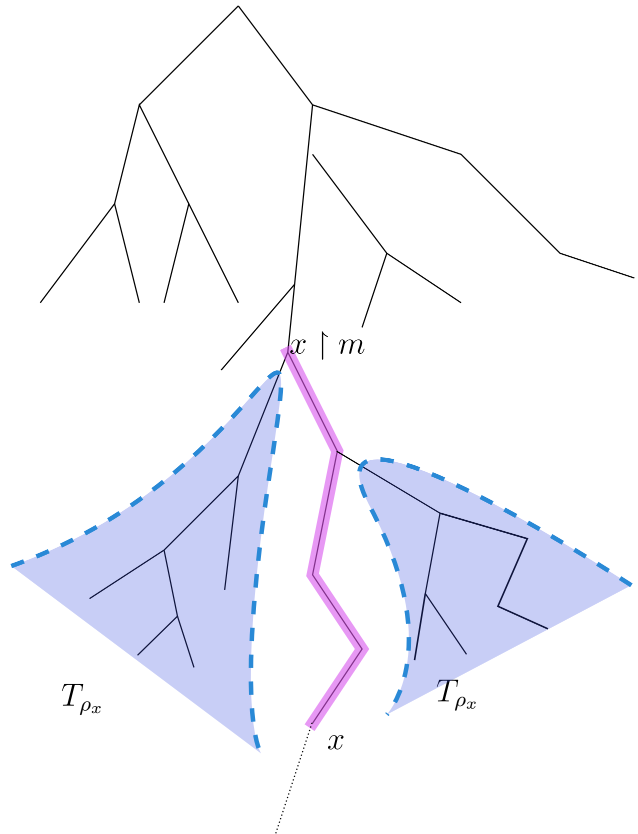

Figure 1: Once T_{ρ_x} is removed from T, $x \upharpoonright m$ becomes isolated and hence x becomes definable by brute-force search through remaining nodes that extend it

Observation 1.9. The definition of x above is absolute to $L[T]$.

Proof. This is because the transfinite recursion constructing T_α's from T, the definition of isolation ranks, and the definition of x (which is arithmetic) only involve Δ_0 formulas with parameters in $L[T]$. $\square$

Corollary 1.10 (Perfect set theorem for closed sets)**.** *Let $X \subseteq \omega^\omega$ be the body of some tree T, then either X has a perfect subset, or $X \in L_{\omega_1}[T]$ (hence countable).*

Proof. We've proved almost everything, except that $X \in L_{\omega_1}[T]$. This is via absoluteness considerations. The construction process T_α and the arithemtical definition of x are both absolute between V and $L[T]$. And $\delta < \omega_1$. We've shown that $X \subseteq L_{\delta+\omega}[T]$, and so X can be defined as the paths through T in that level, which we know are all of them. $\square$

Remark. If you really think about it, all that there is to the proof is already captured in the lyrics to Lemon Tree by Fool's Garden. Isolation is good for you; if all that you can see is the lemon tree then you can define the lemon tree, etc etc.

2 Perfect set theorem for analytic sets

Definition 2.1. An analytic set $X \subseteq \omega^\omega$ is the projection $p[T]$ of the body of a tree $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$. Recall: $p[T] := \{x \mid (\exists y)(x,y) \in [T]\}$

Example 2.2. One of the first analytic sets is Luzin's set A of sequences containing a progressively divisible subsequence:

$$A(x) \Leftrightarrow \exists n_0 < n_1 < n_2 < ... x(n_i) \text{ divides } x(n_{i+1})$$

It is the projection of the tree of attempts searching for such a sequence:

$$\begin{aligned} T := \{(s,t) \mid \ &\text{the } t(0), t(1), ..., t(n)\text{th places of } s \text{ are not a counterexample} \\ &\text{to progressive divisibility for any } n < \text{length}(t)\} \end{aligned}$$

Furthermore, this set is $\utilde{\Sigma}^1_1$-complete, meaning that every analytic set can be obtained as the continuous pre-image of this set.

Remark. Hopefully, the above example brings to mind the set of directed graphs containing a clique or a Hamiltonian path. These are of course classic examples of NP-complete sets. Indeed, in many aspects we have good reasons to think of the NP sets as finitary analogues of the analytic sets and vice versa. As a matter of fact, the sets

$$A := \{x \in \omega^\omega \mid x \text{ codes a countable graph with an infinite clique}\}$$
$$B := \{x \in \omega^\omega \mid x \text{ codes a countable graph with a Hamiltonian path}\}$$

are both $\utilde{\Sigma}^1_1$-complete.

The perfect set theorem for analytic sets is proved using a similar technique to closed sets. First we define the notion of isolation.[1]

Definition 2.3. Given a tree $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$ and (u,v), we say (u,v) is *isolated in* T iff $(u,v) \in T$ and it has no extensions in T that are incompatible in the first coordinate. In other words, every extension of (u,v) in T will be compatible in the first coordinate.

Given a tree $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$, construct again a sequence of sets T_α as follows:

$$T_0 = \emptyset$$
$$T_\lambda = \bigcup_{\beta<\lambda} T_\beta \text{ if } \lambda \text{ is a limit ordinal}$$
$$T_{\alpha+1} = T_\alpha \cup \{(u,v) \in T \mid (u,v) \text{ is isolated in } T \setminus T_\alpha\}$$

Lemma 2.4 and Theorem 2.5 are proved in the same manner as before.

[1] There's some slight clash of notation here. I've defined isolation and isolation rank for closed sets in the previous section. Technically the same definition generalizes to the closed set $[T] \subseteq \omega \times \omega$. But what is terminological consistency in the face of a convenient proof? Hence I've chosen to redefine what isolation means for this specific proof.

Lemma 2.4. *There is a least countable ordinal δ for which $T_\delta = T_{\delta+1}$.*

Theorem 2.5 (The case where $T \neq T_\delta$)**.** *For any T and δ as above, if $T \neq T_\delta$, then there is an injection from 2^ω into $p[T]$.*

Now, if $T = T_\delta$, then every node in T becomes isolated at some point. That is, for every $(u, v) \in T$, there is a least ordinal α such that (u, v) is isolated in $T \smallsetminus T_\alpha$ and then $(u, v) \in T_{\alpha+1}$. Let us call this ordinal α the isolation rank of (u, v), written $\rho(u, v)$. Again, $(\emptyset, \emptyset)$ has the highest isolation rank, and if $(u, v) \prec (u', v')$, then $\rho(u, v) \geq \rho(u', v')$. That is, as you descend down a path, the isolation ranks can't go up (it can only go down or remain the same).

Recall if $x \in p[T]$, then there is some y such that $(x, y) \in [T]$. That is, there is some y such that $(x \restriction n, y \restriction n) \in T$ for all $n \in \omega$.

Observation. So, if $x \in p[T]$, as witnessed by y, then the sequence

$$(\rho(\emptyset, \emptyset), \rho(x \restriction 1, y \restriction 1), \rho(x \restriction 2, y \restriction 2), ...)$$

is a non-increasing sequence of ordinals. Hence it must be eventually constant.

Theorem 2.6 (The case where $T = T_\delta$)**.** *For any T and δ as above, such that $T = T_\delta$, if $x \in p[T]$ as witnessed by y, and $(\rho(\emptyset, \emptyset), \rho(x \restriction 1, y \restriction 1), \rho(x \restriction 2, y \restriction 2), ...)$ is eventually constant starting at $\rho(x \restriction m, y \restriction m)$, then x is definable from $T, x \restriction m, y \restriction m, \rho(x \restriction m, y \restriction m)$.*

Proof. Again, the assumptions imply that $(x \restriction m, y \restriction m)$ is isolated in $T \smallsetminus T_{\rho(x\restriction m, y\restriction m)}$. This means that every extension of the pair $(x \restriction m, y \restriction m)$ in $T \smallsetminus T_{\rho(x\restriction m,y\restriction m)}$ must be compatible in the first coordinate. In other words, every extension of $(x \restriction m, y \restriction m)$in $T \smallsetminus T_{\rho(x\restriction m,y\restriction m)}$ will have the form $(x \restriction k, q)$ for some $k \geq m$ and $q \in \omega^{<\omega}$.

But it is now easy to define x: set $x(i) = j$ iff

$$\begin{cases} (x \restriction m)(i) = j & i < m \\ (\exists (p, q) \in T \smallsetminus T_{\rho(x\restriction m, y\restriction m)})((x \restriction m, y \restriction m) \prec (p, q) \wedge p(i) = j) & i \geq m \end{cases}$$

In other words, to "compute" $x(i)$ from $T, x \restriction m, y \restriction m, \rho(x \restriction m, y \restriction m)$, we only need to build $T_{\rho(x\restriction m,y\restriction m)}$ and remove it from T, and then just brute-force search through pairs of finite sequences of natural numbers to see if any extends $(x \restriction m, y \restriction m)$. Since $(x \restriction m, y \restriction m)$ is isolated and for each length $k \geq m$, extensions of length k exist and all cohere in the first coordinate (recall the isolation rank of $(x \restriction m, y \restriction m)$ is that eventual ordinal constant), this definition defines $x \in p[T]$. $\square$

Corollary 2.7 (Perfect set theorem for analytic sets)**.** *Let $X \subseteq \omega^\omega$ be the projection $p[T]$ the body of some tree T, then either X has a perfect subset, or $X \in L_{\omega_1}[T]$ (hence countable).*

Proof. Again, this is via absoluteness considerations by noticing that the construction process T_α and the arithemtical definition of x are both absolute between V and $L[T]$. $\square$

3 Some more examples of perfect set theorems

Theorem 3.1 (Harrison). *If A is Σ^1_1, then either X has a perfect subset or there exists a computable ordinal α such that every element of A is computable by $\emptyset^{(\alpha)}$.*

Theorem 3.2 (Guaspari, Sacks, Kechris). *If X is Π^1_1, then either X has a perfect subset, or $X \subseteq C_1 := \{x \in \omega^\omega \mid x \in L_{\omega_1^x}\}$*

The set C_1 is called the largest countable thin set.

Theorem 3.3. *If A is $\Sigma^1_2(a)$, then either A has a perfect subset, or $A \in L[a]$*

Under the assumption that $C_2(a) := \omega^\omega \cap L[a]$ is countable, $C_2(a)$ is called the largest countable $\Sigma^1_2(a)$ set. People have been somewhat obsessed with these sets since the 80s.

The above theorem is a consequence of the more general result.

Theorem 3.4 (Mansfield, Solovay). *If X is the projection $p[T]$ of the body of $T \subseteq \omega^{<\omega} \times Y^{<\omega}$ for some set Y, then either X has a perfect subset (moreover the tree corresponding to that perfect set is in $L[T]$), or $X \in L[T]$.*