Ekeland's variational principle in weak and strong systems of arithmetic

David Fernández-Duque[1] · **Paul Shafer[2]** · **Keita Yokoyama[3]**

Abstract

We analyze Ekeland's variational principle in the context of reverse mathematics. We find that that the full variational principle is equivalent to $\Pi^1_1\text{-CA}_0$, a strong theory of second-order arithmetic, while natural restrictions (e.g. to compact spaces or to continuous functions) yield statements equivalent to weak König's lemma (WKL_0) and to arithmetical comprehension (ACA_0). We also find that the localized version of Ekeland's variational principle is equivalent to $\Pi^1_1\text{-CA}_0$, even when restricted to continuous functions. This is a rare example of a statement about continuous functions having great logical strength.

✉ Paul Shafer
  p.e.shafer@leeds.ac.uk
  http://www1.maths.leeds.ac.uk/~matpsh/

  David Fernández-Duque
  david.fernandezduque@ugent.be
  https://users.ugent.be/~dfernnde/

  Keita Yokoyama
  y-keita@jaist.ac.jp
  http://www.jaist.ac.jp/~y-keita/

[1] Department of Mathematics, Ghent University, Krijgslaan 281 S22, 9000 Ghent, Belgium

[2] School of Mathematics, University of Leeds, Leeds LS2 9JT, UK

[3] School of Information Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi 923-1292, Ishikawa, Japan

1 Introduction

The field of *reverse mathematics*, introduced by Friedman [10], aims to identify the minimal foundational assumptions required to prove specific results from several mathematical fields, including mathematical analysis. There are many advantages to determining these minimal assumptions, among them aiding in extracting computational content from theorems whose original proofs were non-constructive. It is also desirable from a methodological perspective to avoid strong set-theoretic assumptions when possible, and indeed advances in reverse mathematics have shown us that a large portion of known mathematics can be carried out within a relatively small fragment of second-order arithmetic [14]. Second-order arithmetic extends the language of Peano arithmetic by adding variables and quantifiers ranging over sets of natural numbers (see Sect. 2); this is sufficient to formalize many familiar concepts from analysis, including real numbers (and, more generally, points in a complete separable metric space, also called a *Polish space*), continuous functions, and open and closed sets (see Sect. 3). One can then ask which axioms are needed to prove (say) the Stone–Weierstrass theorem in this framework.

Although there are many exceptions, a surprisingly large portion of the theorems analyzed are equivalent to one of a handful of systems of second-order arithmetic. These axiomatic systems are always assumed to extend the 'base theory' of reverse mathematics, *recursive comprehension* (RCA_0), and the ones we consider are *weak König's lemma* (WKL_0), *arithmetical comprehension* (ACA_0), and Π_1^1-*comprehension* (Π_1^1-CA_0), listed in strictly increasing order of strength. Each of RCA_0, ACA_0 and Π_1^1-CA_0 include axioms asserting that sets of the form $\{n \in \mathbb{N} : \varphi(n)\}$ exist, where in the case of RCA_0, φ must express a computable predicate; in the case of ACA_0, φ may contain arbitrary quantifiers over natural numbers (but not over sets of natural numbers); and in the case of Π_1^1-CA_0, φ is of the form $\forall X \psi(n, X)$, where X is a second-order variable and $\psi(n, X)$ contains no additional second-order quantifiers. Meanwhile, the theory WKL_0 asserts that any infinite binary tree has an infinite branch.

It is known, for example, that the Baire category theorem is provable in RCA_0, that the Heine–Borel theorem is equivalent to WKL_0, and that the Stone–Weierstrass theorem is equivalent to ACA_0. As we will see throughout the text, there are many more examples of theorems equivalent to these theories. On the other hand, statements equivalent to Π_1^1-CA_0 are much more difficult to come by. One example is the Cantor–Bendixson theorem, stating that any closed subset of a Polish space can be written as the union of a countable set and a perfect set.

From a logical perspective, there is a vast gap in strength between the systems ACA_0 and Π_1^1-CA_0. ACA_0 may be thought of as the second-order analog of the familiar first-order system of Peano arithmetic (PA). Indeed, ACA_0 is conservative over PA, meaning that for every first-order statement φ (i.e., for every φ that may contain quantifiers over natural numbers but not over sets of natural numbers), φ is provable in ACA_0 if and only if it is provable in PA. Both ACA_0 and PA have proof-theoretic ordinal ε_0. In computational terms, ACA_0 is characterized by the existence of Turing jumps. That is, over RCA_0, ACA_0 is equivalent to the statement "for every set X, the Turing jump of X is also a set." Π_1^1-CA_0, on the other hand, is a squarely impredicative system whose proof-theoretic ordinal is far above Γ_0. In fact, its proof-theoretic ordinal is far above even

the Bachmann–Howard ordinal. In computational terms, Π_1^1-CA_0 is characterized by the existence of hyperjumps. That is, over RCA_0, Π_1^1-CA_0 is equivalent to the statement "for every set X, the hyperjump of X is also a set." See [14] for analyses of the Heine–Borel, Stone–Weierstrass, and Cantor–Bendixson theorems and for characterizations of ACA_0 and Π_1^1-CA_0 in terms of jumps and hyperjumps, and see [13] for an ordinal analysis of Π_1^1-CA_0.

Our goal in this article is to study Ekeland's variational principle [5] in the context of reverse mathematics. This principle states that under certain conditions, lower semi-continuous functions on complete metric spaces always attain 'approximate minima,' which we call *critical points*. In his original paper [5] and the survey [6], Ekeland provides many applications of his variational principle, centered around optimization problems concerning minimal surfaces, partial differential equations, geodesics, the geometry of Banach spaces, control theory, and more. Ekeland's variational principle has been studied extensively, leading to many variants and extensions (see e.g. [1,11]). The variational principle can also be used to give an easy proof of Caristi's fixed point theorem [3], whose logical strength we analyze in forthcoming work [7–9].

A priori, an analysis of Ekeland's variational principle in reverse mathematics is a natural and interesting project. Ekeland's variational principle is a well-known and important result, and understanding its computational content could lead to algorithms for approximating critical points, or at least to determining that no such algorithm exists. From a technical perspective, lower semi-continuous functions have not yet received much attention in reverse mathematics, and developing a theory of these functions in second-order arithmetic is an interesting endeavor on its own (see Sect. 4).

However, *a posteriori* the analysis of Ekeland's variational principle is even more interesting and quite surprising: as we will see, natural restrictions of the result (e.g. to compact spaces, to continuous f) yield statements equivalent over RCA_0 to each of WKL_0, ACA_0, and Π_1^1-CA_0, including what is, to the best of our knowledge, the first statement about continuous functions stemming from analysis that is equivalent to Π_1^1-CA_0.

Before diving into formal systems, let us discuss the variational principle in more detail and sketch a proof. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, where we stipulate that $\sup \varnothing = -\infty$ and $\inf \varnothing = \infty$. If $\mathcal{X}$ is a metric space, recall that a function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is *lower semi-continuous* if for every $x \in \mathcal{X}$ and every $\lambda < f(x)$, there is a $\delta > 0$ such that whenever $d(x, y) < \delta$, it follows that $f(y) > \lambda$. The notion of an *upper semi-continuous* function is defined dually. Clearly, a function that is both upper and lower semi-continuous is continuous. Additionally, recall the following well-known characterization of semi-continuity.

Lemma 1.1 *Let $\mathcal{X}$ be a metric space. A function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is*

(a) *lower semi-continuous if and only if $\{(x, y) \in \mathcal{X} \times \mathbb{R} : f(x) \leq y\}$ is closed;*
(b) *upper semi-continuous if and only if $\{(x, y) \in \mathcal{X} \times \mathbb{R} : f(x) \geq y\}$ is closed.*

If $f : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is continuous but $\mathcal{X}$ is not compact, then f may not attain its infimum, and since every continuous function is also lower semi-continuous, the same holds of lower semi-continuous functions. Nevertheless, Ekeland's variational principle states that f has points that are in a sense approximate local minima, and we

call these points *ε-critical points*. For the sake of brevity, we will often refer to lower semi-continuous functions $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ as *potentials*.

Definition 1.2 Let $\mathcal{X}$ be a metric space, let $\varepsilon > 0$, and let $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be a potential. A point $x_* \in \mathcal{X}$ is an *ε-critical point of f* if

$$(\forall y \in \mathcal{X})[(\varepsilon d(x_*, y) \leq f(x_*) - f(y)) \to y = x_*].$$

Ekeland's variational principle then reads as follows.

Theorem 1.3 (Ekeland [5]) *If $\mathcal{X}$ is a complete metric space and $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ is any potential, then for every $\varepsilon > 0$, f has an ε-critical point x_*.*

Moreover, for any $x_0 \in \mathcal{X}$, we can choose x_* so that

$$\varepsilon d(x_0, x_*) \leq f(x_0) - f(x_*). \tag{1}$$

In the literature, it is often assumed that $f(x_0) \leq \varepsilon + \inf(f)$, in which case (1) is replaced by $d(x_0, x_*) \leq 1$. We refer to Theorem 1.3 without (1) as the *free variational principle* (FVP) and with (1) as the *localized variational principle* (LVP); note that Ekeland instead calls these the *weak* and *strong* principles, respectively.

Let us sketch a proof similar to that presented by Du [4], which is based on one given by Brézis and Browder [2] in a more general order-theoretic setting. Suppose that $\mathcal{X}$ is a complete metric space, let $x_0 \in \mathcal{X}$, and let $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be any potential. For $\varepsilon > 0$, define a partial order $\preccurlyeq_\varepsilon$ on $\mathcal{X}$ given by $y \preccurlyeq_\varepsilon x$ if and only if $\varepsilon d(x, y) \leq f(x) - f(y)$. For $x \in \mathcal{X}$, define $S(x) = \{y \in \mathcal{X} : y \preccurlyeq_\varepsilon x\}$. Then we construct a sequence $(x_n : n < \omega)$ by letting $x_{n+1} \in S(x_n)$ be such that $f(x_{n+1}) < \inf f[S(x_n)] + 2^{-n}$. It is not too hard to check that $(x_n : n < \omega)$ is Cauchy and hence converges to some $x_* \in \mathcal{X}$, and, using the fact that f is lower semi-continuous, that $f(x_*) \leq f(x_n)$ for all $n < \omega$. From this it readily follows that x_* is ε-critical and satisfies (1); we leave the details to the reader.

As it turns out, there are some issues when attempting to formalize this argument in second-order arithmetic. The sets $S(x_n)$ are closed, and finding infima of lower semi-continuous functions on closed sets requires $\Pi^1_1\text{-CA}_0$ in general. It is easier to formalize the argument when $S(x_n)$ is open *and* f is continuous, and we can achieve this by replacing $S(x_n)$ by a suitable open set. However, this modification typically causes (1) to fail. In fact, we obtain that the FVP for continuous functions is equivalent to ACA_0 (Sect. 7), while the LVP is equivalent to $\Pi^1_1\text{-CA}_0$, even when f is assumed to be continuous (Sect. 10). This last result is particularly interesting because it gives an example of a statement about continuous functions that is equivalent to $\Pi^1_1\text{-CA}_0$; typically the mathematics of continuous functions can be carried out in ACA_0 and below.

By tweaking other parameters in the statement, for example taking $\mathcal{X}$ to be compact or even to be a specific space such as the unit interval, we obtain several variants of the variational principle, each equivalent to one of WKL_0, ACA_0 or $\Pi^1_1\text{-CA}_0$. The equivalence between these theories and variational principles is established over RCA_0, and hinges on known equivalences, e.g. the equivalence between ACA_0 and the fact that every increasing sequence of rationals bounded above has a supremum. The main reversals are found in Sect. 9. A full summary of our main results is found in Sect. 11.

2 Subsystems of second-order arithmetic

In this section we review the subsystems of second-order arithmetic we will be working with, loosely following the presentation of Simpson [14]. Let us first settle conventions regarding syntax. The language of second-order arithmetic consists of first-order variables intended to range over natural numbers; second-order variables intended to range over sets of natural numbers; constant symbols 0 and 1; 2-ary function symbols $+$ and $\times$; and 2-ary relation symbols $=$, $<$, and $\in$. Terms and formulas are built from variables, constant symbols, function symbols, relation symbols, propositional connectives ($\neg$, $\wedge$, $\vee$, etc.), and the quantifiers $\forall$ and $\exists$. The equality relation symbol is restricted to first-order objects. Equality for second-order objects is defined via $\in$, and we write $X = Y$ as an abbreviation for $\forall n(n \in X \leftrightarrow n \in Y)$. Similarly, we write $X \subseteq Y$ as an abbreviation for $\forall n(n \in X \rightarrow n \in Y)$.

We use Δ_0^0 to denote the class of all formulas, possibly with parameters, where no second-order quantifiers appear and all first-order quantifiers are *bounded;* that is, of the form $\forall x(x < t \rightarrow \varphi)$ or $\exists x(x < t \wedge \varphi)$. We simultaneously define $\Sigma_0^0 = \Pi_0^0 = \Delta_0^0$, and we recursively define Σ_{n+1}^0 to be the set of all formulas of the form $\exists x \varphi$ with $\varphi \in \Pi_n^0$ and similarly define Π_{n+1}^0 to be the set of all formulas of the form $\forall x \varphi$ with $\varphi \in \Sigma_n^0$. We denote by Π_ω^0 the union of all Π_n^0; these are the *arithmetical formulas.* The classes Σ_n^1 and Π_n^1 are defined analogously by counting alternations of second-order quantifiers and setting $\Sigma_0^1 = \Pi_0^1 = \Delta_0^1 = \Pi_\omega^0$.

We use Robinson arithmetic Q as our background theory, which essentially consists of the axioms of PA without induction (see [12, Definition I.1.1]). When added to the basic axioms of Q, the following schemes axiomatize many of the theories we consider. Below, Γ denotes a class of formulas.

- $\Gamma\text{-CA}$: $\exists X \forall x \left(x \in X \leftrightarrow \varphi(x) \right)$, where $\varphi \in \Gamma$ and X is not free in φ;
- $\Delta_1^0\text{-CA}$: $\forall x \left(\pi(x) \leftrightarrow \sigma(x) \right) \rightarrow \exists X \forall x \left(x \in X \leftrightarrow \sigma(x) \right)$, where $\sigma \in \Sigma_1^0$, $\pi \in \Pi_1^0$, and X is not free in σ;
- $\mathrm{I}\Gamma$: $\varphi(0) \wedge \forall x \left(\varphi(x) \rightarrow \varphi(x+1) \right) \rightarrow \forall x \, \varphi(x)$, where $\varphi \in \Gamma$.

Recursive comprehension, the base theory of second-order arithmetic, is defined as

$$\mathsf{RCA}_0 : \ Q + \mathrm{I}\Sigma_1^0 + \Delta_1^0\text{-CA}.$$

We also define the theory of arithmetical comprehension ACA_0: $\mathsf{RCA}_0 + \Sigma_1^0\text{-CA}$ and the much stronger theory $\Pi_1^1\text{-CA}_0$: $\mathsf{ACA}_0 + \Pi_1^1\text{-CA}$. This gives us three of the four important theories that we consider; the fourth, WKL_0, requires formalizing trees in second-order arithmetic.

RCA_0 suffices to define a bijective pairing function $\langle \cdot, \cdot \rangle \colon \mathbb{N}^2 \to \mathbb{N}$ that is increasing in both coordinates, such as $\langle x, y \rangle = (x+y)(x+y+1)/2$. We also use $\langle \cdot, \cdot \rangle$ to denote a pairing function on sets, which may also be defined in a standard way. With this, a binary relation is a pair of sets $\langle A, R \rangle$ with $A, R \subseteq \mathbb{N}$, where the elements of R are of the form $\langle n, m \rangle$ with $n, m \in A$. When a binary relation F is meant to represent a function, we write $y = F(x)$ instead of $\langle x, y \rangle \in F$.

RCA$_0$ also suffices to implement the typical codings of finite sequences of natural numbers as natural numbers (see [14, Section II.2], for example). The set of all finite sequences of natural numbers is denoted $\mathbb{N}^{<\mathbb{N}}$. We write $\sigma \sqsubseteq \tau$ if σ is an initial segment of τ, $\sigma \sqsubset \tau$ if σ is a proper initial segment of τ, and define $\downarrow \sigma = \{\tau \in \mathbb{N}^{<\mathbb{N}} : \tau \sqsubseteq \sigma\}$. If $x \colon \mathbb{N} \to \mathbb{N}$ and $n \in \mathbb{N}$, we write $x \restriction n$ for the finite sequence $(x(i))_{i<n}$. We extend the use of $\sqsubset$ by writing $\sigma \sqsubset x$ whenever $\sigma = x \restriction n$ for some n.

Definition 2.1 A *tree* is a set $T \subseteq \mathbb{N}^{<\mathbb{N}}$ such that $\forall \sigma (\sigma \in T \to \downarrow \sigma \subseteq T)$. We say that T is a *binary tree* if $T \subseteq \{0, 1\}^{<\mathbb{N}}$, that is, if all entries appearing in elements of T are either 0 or 1. When $T \subseteq \mathbb{N}^{<\mathbb{N}}$ is a tree, we say that an infinite sequence x is a *path through T* if $\forall n(x \restriction n \in T)$. The collection of all paths through T is denoted $[T]$.

Say that a set $X \subseteq \mathbb{N}$ is *finite* if it is bounded above, and say that X is infinite otherwise. With this, we define the axiom WKL to be the natural formalization of the following theorem.

Theorem 2.2 (König) *Every infinite binary tree has an infinite path.*

For reference, we list the theories we have defined in increasing order of strength:

$$\begin{aligned}
\text{RCA}_0 &: \quad \text{Q} + \text{I}\Sigma^0_1 + \Delta^0_1\text{-CA}; \\
\text{WKL}_0 &: \quad \text{RCA}_0 + \text{WKL}; \\
\text{ACA}_0 &: \quad \text{RCA}_0 + \Sigma^0_1\text{-CA}; \\
\Pi^1_1\text{-CA}_0 &: \text{ACA}_0 + \Pi^1_1\text{-CA}.
\end{aligned}$$

The standard reference for second-order arithmetic is Simpson's [14], and we refer the reader to it for a complete treatment of all the material mentioned above.

3 Metric spaces in second-order arithmetic

Part of the appeal of second-order arithmetic as a foundational system for mathematics is that it suffices to develop a large part of mathematical analysis, particularly when dealing with complete separable metric spaces. However, this requires some coding machinery. In this section, we review this machinery and establish notation that will be used throughout.

First, we assume that $\mathbb{Q}$ is represented in some standard way using e.g. pairs of natural numbers and that $\mathbb{R}$ is represented by rapidly converging Cauchy sequences of rationals as in [14, Section II.4]. We use the notation $\mathbb{Q}_{>0} = \{q \in \mathbb{Q} : q > 0\}$, and we define $\mathbb{Q}_{\geq 0}$, $\mathbb{R}_{\geq 0}$, etc. analogously.

Definition 3.1 (RCA$_0$; see [14, Definition II.5.1]) A (code for a) complete separable metric space $\mathcal{X} = \widehat{X}$ is a non-empty set $X \subseteq \mathbb{N}$ together with a sequence of real numbers $d \colon X \times X \to \mathbb{R}_{\geq 0}$ such that $d(a, a) = 0$, $d(a, b) = d(b, a) \geq 0$, and $d(a, b) + d(b, c) \geq d(a, c)$ for all $a, b, c \in X$. A *point* of $\widehat{X}$ is a sequence $x = (x_i)_{i \in \mathbb{N}}$ of elements of X such that for all $i \leq j$, $d(x_i, x_j) \leq 2^{-i}$. We write $x \in \widehat{X}$ to mean that x is a point of $\widehat{X}$. We identify $a \in X$ with the sequence $(a)_{i \in \mathbb{N}}$ and consider X as a dense subset of $\widehat{X}$. We set $d(x, y) = \lim_{n \to \infty} d(x_n, y_n)$, and write $x =_{\mathcal{X}} y$ if $d(x, y) = 0$ (subscripts will be omitted if there is no confusion).

We use either notation $\mathcal{X}$ or $\widehat{X}$ to denote complete separable metric spaces. The symbol $\widehat{X}$ emphasizes that the space is coded by the dense set X (and the metric on X). The following spaces appear throughout.

(1) The real line, $\mathbb{R}$, with dense set the rational numbers, $\mathbb{Q}$, equipped with the usual metric. Closed subintervals of $\mathbb{R}$ may be represented similarly, where the dense set for $[a, b]$ is $\mathbb{Q} \cap [a, b]$.
(2) The Baire space, with dense set the set of sequences $x \in \mathbb{N}^{\mathbb{N}}$ that are eventually zero and

$$d(x, y) = \max \left(\{0\} \cup \{2^{-n} : x(n) \neq y(n)\} \right).$$

(3) The Cantor space, which is $\{0, 1\}^{\mathbb{N}}$ (also denoted $2^{\mathbb{N}}$) seen as a subspace of the Baire space.
(4) $\mathcal{C}([a, b])$ with $a < b$ rational numbers. The dense set is given by piecewise linear continuous functions $f : [a, b] \to \mathbb{R}$ with rational breakpoints, each represented by finitely many pairs $\langle x, f(x) \rangle \in \mathbb{Q} \times \mathbb{Q}$. The metric is $d(f, g) = \max_{x \in [a,b]} |f(x) - g(x)|.$[1]

In the cases of the Baire space and the Cantor space, a sequence that is eventually zero may be represented via an appropriate finite initial segment and hence as a natural number. If σ is a finite sequence, then where convenient we identify σ with $\sigma^\frown 0^{\mathbb{N}}$, which is the infinite sequence beginning with σ and is then identically zero. We remark also that the official dense set for $\mathcal{C}([a, b])$ in [14] is the set of polynomial functions $f : [a, b] \to \mathbb{Q}$ with rational coefficients, but piecewise linear functions are more convenient for us. The two presentations are equivalent over RCA_0 (see [14, Example II.10.3 and Lemma IV.2.4]).

Definition 3.2 (RCA_0; see [14, Definition II.5.6]) Let $\widehat{X}$ be a complete separable metric space. A *(code for a) rational open ball* $B_r(a)$ is an ordered pair $\langle a, r \rangle$, with $a \in X$ and $r \in \mathbb{Q}_{>0}$. We define $B_r(a) \Subset B_q(b)$ if $d(a, b) + r < q$.

A *(code for an) open set* $\mathcal{U}$ in $\widehat{X}$ is a set $U \subseteq \mathbb{N} \times X \times \mathbb{Q}_{>0}$, where a point $x \in \widehat{X}$ is said to belong to $\mathcal{U}$ (abbreviated $x \in \mathcal{U}$) if it satisfies the Σ_1^0 condition $\exists n \exists a \exists r (d(x, a) < r \wedge (n, a, r) \in U)$. A *(code for a) closed set* $\mathcal{C}$ is the same as a code for its complement $\widehat{X} \setminus \mathcal{C}$, except we define $x \in \mathcal{C}$ if $x \notin \widehat{X} \setminus \mathcal{C}$.

The intuition behind the above definition is that a code for an open set is an enumeration of rational open balls. If $U \subseteq \mathbb{N} \times X \times \mathbb{Q}_{>0}$ codes an open set, then the rational open ball $B_r(a)$ appears in the enumeration if there is an n with $(n, a, r) \in U$.

We remark that for $\mathcal{X} \subseteq \mathbb{R}$, we may represent basic open sets in the form $(a - r, a + r)$ rather than $B_r(a)$, and we will often prefer this representation. On occasion we will regard closed balls (or, more precisely, closures of balls) as metric spaces in their own right, in which case if $B_r(a)$ is a ball in $\widehat{X}$, then $\overline{B_r(a)}$ is the subspace with dense set $B_r(a) \cap X$.

We may also reason about compactness within second-order arithmetic.

---

[1] We emphasize that $\mathcal{C}([a, b])$ is the space of uniformly continuous functions $f : [a, b] \to \mathbb{R}$ that have moduli of uniform continuity in the sense of Simpson [14, Definition IV.2.1]. The statement "every continuous function $f : [a, b] \to \mathbb{R}$ has a modulus of uniform continuity" is equivalent to WKL_0 over RCA_0.

Definition 3.3 (RCA$_0$; see [14, Definition III.2.3]) A complete separable metric space $\mathcal{X}$ is *compact* if there is a sequence of finite sequences $((x_{i,j} : j < n_i) : i \in \mathbb{N})$ of points in $\mathcal{X}$ such that

$$(\forall z \in \mathcal{X})(\forall i \in \mathbb{N})(\exists j < n_i)(d(z, x_{i,j}) < 2^{-i}).$$

We say that $\mathcal{X}$ is *Heine–Borel compact* if for every sequence $(B_{r_k}(a_k) : k \in \mathbb{N})$ of rational open balls such that $(\forall x \in \mathcal{X})(\exists k \in \mathbb{N})(x \in B_{r_k}(a_k))$, there is an $N \in \mathbb{N}$ such that $(\forall x \in \mathcal{X})(\exists k < N)(x \in B_{r_k}(a_k))$.

What we call 'compact' might be more properly called 'uniformly (or effectively) totally bounded,' but when working with complete separable metric spaces in second-order arithmetic, the convention is to use 'compact' for this notion, and we do not wish to deviate. WKL$_0$ is required to show that every compact (in the above sense) complete separable metric space is Heine–Borel compact.

Theorem 3.4 (See [14, Theorem IV.1.2 and Theorem IV.1.5]) *The following are equivalent over* RCA$_0$.

(1) WKL$_0$.
(2) *Every compact complete separable metric space is Heine–Borel compact.*
(3) *The unit interval* $[0, 1]$ *is Heine–Borel compact.*

Sequential compactness is an even stronger notion. ACA$_0$ is required to show that every compact space is sequentially compact.

Theorem 3.5 (See [14, Theorem III.2.2 and Theorem III.2.7]) *The following are equivalent over* RCA$_0$.

(1) ACA$_0$.
(2) *Every infinite sequence of points in a compact complete separable metric space has a convergent subsequence.*
(3) *Every infinite sequence of points in* $[0, 1]$ *has a convergent subsequence.*
(4) *Every increasing (or decreasing) sequence of points in* $[0, 1]$ *converges.*
(5) *Every increasing (or decreasing) sequence of rational points in* $[0, 1]$ *converges.*

Note that the equivalence with (5) is not stated explicitly in [14], but it follows from the proof of Theorem III.2.2 given there. We conclude this section by defining continuous functions between complete separable metric spaces. The idea is to code a continuous function $f : \mathcal{X} \to \mathcal{Y}$ by an enumeration of (codes for) pairs of open balls $\langle B_r(a), B_q(b) \rangle$, where $B_r(a) \subseteq \mathcal{X}$ and $B_q(b) \subseteq \mathcal{Y}$. If the pair $\langle B_r(a), B_q(b) \rangle$ appears in the enumeration, then this means that f maps $B_r(a)$ into the closure of $B_q(b)$.

Definition 3.6 (RCA$_0$; [14, Definition II.6.1]) Let $\mathcal{X} = \widehat{X}$ and $\mathcal{Y} = \widehat{Y}$ be complete separable metric spaces. A *continuous partial function* $f : \mathcal{X} \to \mathcal{Y}$ is coded by a set $\Phi \subseteq \mathbb{N} \times X \times \mathbb{Q}_{>0} \times Y \times \mathbb{Q}_{>0}$ that satisfies the properties below. Let us write $B_r(a) \overset{\Phi}{\to} B_s(b)$ for $\exists n \left((n, a, r, b, s) \in \Phi \right)$. Then, for all $a, a' \in X$, all $b, b' \in Y$, and all $r, r', s, s' \in \mathbb{Q}_{>0}$, Φ must satisfy:

(CF1) if $B_r(a) \overset{\Phi}{\to} B_s(b)$ and $B_r(a) \overset{\Phi}{\to} B_{s'}(b')$, then $d(b, b') \leq s + s'$;

(CF2) if $B_r(a) \overset{\Phi}{\to} B_s(b)$ and $B_{r'}(a') \Subset B_r(a)$, then $B_{r'}(a') \overset{\Phi}{\to} B_s(b)$;

(CF3) if $B_r(a) \overset{\Phi}{\to} B_s(b)$ and $B_s(b) \Subset B_{s'}(b')$, then $B_r(a) \overset{\Phi}{\to} B_{s'}(b')$.

A point $x \in \mathcal{X}$ is in the domain of the function f coded by Φ if, for every $\varepsilon > 0$, there are $B_r(a) \overset{\Phi}{\to} B_s(b)$ such that $d(x, a) < r$ and $s < \varepsilon$. If $x \in \text{dom}(f)$, we define the value $f(x)$ to be the unique point $y \in \mathcal{Y}$ such that $d(y, b) \leq s$ for all $B_r(a) \overset{\Phi}{\to} B_s(b)$ with $d(x, a) < r$. A *continuous function* is a continuous partial function $f : \mathcal{X} \to \mathcal{Y}$ with $\text{dom}(f) = \mathcal{X}$. In case $\mathcal{Y} = \mathbb{R}$, we often write $B_r(a) \overset{\Phi}{\to} (u, v)$ for $B_r(a) \overset{\Phi}{\to} B_s(b)$ with $u = b - s$ and $v = b + s$.

To reason about the values of coded continuous functions, it often helps to think in the following way. Suppose that Φ codes a continuous function $f : \mathcal{X} \to \mathcal{Y}$. If $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ are such that for every $\varepsilon > 0$ there are $B_r(a) \overset{\Phi}{\to} B_s(b)$ with $x \in B_r(a)$, $y \in B_s(b)$, and $s < \varepsilon$, then $f(x) = y$.

The following lemmas are useful for constructing codes of open sets, closed sets, and continuous functions.

Lemma 3.7 [14, Lemma II.5.7] *For a given Σ_1^0 formula $\varphi(x)$, the following is provable within* RCA$_0$. *Let $\mathcal{X} = \widehat{X}$ be a complete separable metric space. If $x =_{\mathcal{X}} y$ implies $\varphi(x) \leftrightarrow \varphi(y)$, then there exists (a code for) an open set $\mathcal{U} \subseteq \mathcal{X}$ such that $\varphi(x) \leftrightarrow x \in \mathcal{U}$.*

This lemma guarantees that a Σ_1^0-definable subset of a complete separable metric space is an open set, and thus a Π_1^0-definable subset is a closed set. Note that this lemma holds uniformly. In other words, if $\varphi(x, i)$ is a Σ_1^0 formula that defines a subset of $\mathcal{X}$ for each $i \in \mathbb{N}$, then there exists a sequence of codes for open sets $(\mathcal{U}_i)_{i \in \mathbb{N}}$ such that $\varphi(x, i) \leftrightarrow x \in \mathcal{U}_i$.

Intuitively, the following lemma states that if a function is uniformly continuous in an effective way and the values of the function can be computed on a dense set (e.g., if $f(x)$ is provided by elementary functions, by power series, etc.), then there is a code for the function. This property also holds restricted to any open set.

Lemma 3.8 (RCA$_0$) *Let $\mathcal{X} = \widehat{X}$ and $\mathcal{Y} = \widehat{Y}$ be complete separable metric spaces, and let $\mathcal{U} \subseteq \mathcal{X}$ be an open set. Assume that $(\langle a_i, y_i \rangle : i \in \mathbb{N})$ is a sequence of points in $X \times Y$ and that $h : \mathbb{N} \to \mathbb{N}$ is a function such that*

- *$(a_i)_{i \in \mathbb{N}}$ enumerates all points in $\mathcal{U} \cap X$,*
- *$d_{\mathcal{X}}(a_i, a_j) < 2^{-h(n)}$ implies $d_{\mathcal{Y}}(y_i, y_j) < 2^{-n}$ for all $i, j, n \in \mathbb{N}$.*

Then, there exists (a code for) a continuous function $f : \mathcal{U} \to \mathcal{Y}$ such that $f(a_i) = y_i$ for all $i \in \mathbb{N}$. (In fact, h is a modulus of uniform continuity for f; see [14, Definition IV.2.1].)

Proof Let U be the code for $\mathcal{U}$. Define a code Φ for a continuous partial function $f : \mathcal{X} \to \mathbb{R}$ so that $B_r(a) \overset{\Phi}{\to} B_s(b)$ if and only if the quadruple $(a, r, b, s) \in X \times \mathbb{Q}_{>0} \times Y \times \mathbb{Q}_{>0}$ satisfies

- $B_r(a) \not\in B_{r'}(a')$ for some $(n, a', r') \in U$, and
- $B_r(a) \not\in B_{2^{-h(n)-1}}(a_i)$ and $B_{2^{-n}}(y_i) \not\in B_s(b)$ for some $n, i \in \mathbb{N}$.

The code Φ exists because the above conditions can be described by a Σ_1^0 formula. It is easy to see that Φ satisfies (CF2) and (CF3) of Definition 3.6. We show that Φ also satisfies (CF1). Assume that $B_r(a) \overset{\Phi}{\to} B_s(b)$ and $B_r(a) \overset{\Phi}{\to} B_{s'}(b')$, where in each case the second condition is satisfied by n, a_i, y_i and m, a_j, y_j, respectively. Without loss of generality, we may assume that $n \leq m$ and that h is non-decreasing. Since $a \in B_{2^{-h(n)-1}}(a_i) \cap B_{2^{-h(m)-1}}(a_j)$, we have $d_{\mathcal{X}}(a_i, a_j) < 2^{-h(n)}$. Then $d_{\mathcal{Y}}(y_i, y_j) < 2^{-n}$, and thus $y_j \in B_{2^{-n}}(y_i) \subseteq B_s(b)$. Hence $y_j \in B_s(b) \cap B_{s'}(b')$, which implies that $d_{\mathcal{Y}}(b, b') \leq s + s'$.

Finally, we check that $\mathcal{U} = \mathrm{dom}(f)$ and $f(a_i) = y_i$. If $x \notin \mathcal{U}$, then $x \notin \mathrm{dom}(f)$ by the first condition, thus we have $\mathrm{dom}(f) \subseteq \mathcal{U}$. For the converse, assume that $x \in \mathcal{U}$ and $\varepsilon > 0$. Take $(n', a', r') \in U$ so that $x \in B_{r'}(a')$ and $n \in \mathbb{N}$ so that $2^{-n+1} < \varepsilon$, and then take $r \in \mathbb{Q}_{>0}$ so that $B_{2r}(x) \not\in B_{r'}(a')$ and $r < 2^{-h(n)-1}$. Pick a point $a = a_i \in \mathcal{U} \cap X$ so that $d_{\mathcal{X}}(x, a) < r$ and a point $b \in Y$ so that $d_{\mathcal{Y}}(y_i, b) < 2^{-n}$, and put $s = 2^{-n+1}$. Then we have $B_r(a) \overset{\Phi}{\to} B_s(b)$ and $x \in B_r(a)$. Since ε was arbitrary and $s < \varepsilon$, this means that $x \in \mathrm{dom}(f)$. If $x = a_j$ for some $j \in \mathbb{N}$ in the above, then i can be always chosen to be equal to j, and thus one can find $B_r(a) \overset{\Phi}{\to} B_s(b)$ with $a_j \in B_r(a)$, $y_j \in B_s(b)$ and $s < \varepsilon$ for all $\varepsilon > 0$. This means that $f(a_j) = y_j$ for all $j \in \mathbb{N}$. □

We also remark that continuous partial functions can be patched together, provided they are mutually compatible.

Lemma 3.9 (RCA$_0$; [15, Lemma 3.31]) *Let $\mathcal{X} = \widehat{X}$ and $\mathcal{Y} = \widehat{Y}$ be complete separable metric spaces. Assume that $(\mathcal{U}_i, f_i)_{i \in \mathbb{N}}$ is a sequence of pairs of (codes for) open sets $\mathcal{U}_i \subseteq \mathcal{X}$ and continuous functions $f_i : \mathcal{U}_i \to \mathcal{Y}$ such that $f_i(x) = f_j(x)$ for any $i, j \in \mathbb{N}$ and $x \in \mathcal{U}_i \cap \mathcal{U}_j$. Then, there exists (a code for) a continuous function $\bar{f} : \bigcup_{i \in \mathbb{N}} \mathcal{U}_i \to \mathcal{Y}$ such that $\bar{f}(x) = f_i(x)$ for all $i \in \mathbb{N}$ and $x \in \mathcal{U}_i$.*

Notice that the above lemmas also hold uniformly. For Lemma 3.8, if we are given a sequence $(\langle a_{i,j}, y_{i,j}\rangle_{i \in \mathbb{N}}, \mathcal{U}_j, h_j : j \in \mathbb{N})$, where $\langle a_{i,j}, y_{i,j}\rangle_{i \in \mathbb{N}}, \mathcal{U}_j$ and h_j satisfy the conditions of Lemma 3.8 for all $j \in \mathbb{N}$, then we may obtain a sequence of continuous functions $(f_j : \mathcal{U}_j \to \mathcal{Y})_{j \in \mathbb{N}}$.

Finally, we consider products of finitely many metric spaces $(\mathcal{X}_i : i < n)$. In [14, Example II.5.4], product spaces are defined via the Euclidean metric, but for our purposes it is more convenient to use the 'max' metric instead.

Definition 3.10 (RCA$_0$; see [14, Example II.5.4]) Let $(\mathcal{X}_i : i < n)$ be complete separable metric spaces coded by dense sets $(X_i : i < n)$ and metrics $(d_i : i < n)$. The *product space* $\mathcal{X} = \prod_{i<n} \mathcal{X}_i$ is the complete separable metric space coded by the dense set $X = \prod_{i<n} X_i$ and the metric $d : X \times X \to \mathbb{R}$ where

$$d(\vec{a}, \vec{b}) = \max_{i<n} d_i(a_i, b_i).$$

Products of metric spaces will be very useful to us. For example, we may use them to define Banach spaces in the context of second-order arithmetic: these are simply metric spaces $\mathcal{X}$ equipped with a designated element $\mathbf{0} \in \mathcal{X}$ and total, continuous functions $+ \colon \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ and $\cdot \colon \mathbb{R} \times \mathcal{X} \to \mathcal{X}$ such that if we define $\|x\| = d(x, \mathbf{0})$, then $(\mathcal{X}, \mathbf{0}, +, \cdot, \|\cdot\|)$ satisfies the standard definition of a normed vector space. On occasion we will make free use of the fact that some of the spaces we consider, such as $\mathcal{C}([0, 1])$, come equipped with a standard Banach space structure.

If $\mathcal{X} = \prod_{i<n} \mathcal{X}_i$ is the product of the complete separable metric spaces $(\mathcal{X}_i : i < n)$, then, officially, the basic open sets in $\mathcal{X}$ are those of the form $B_r(\vec{a})$, where $\vec{a} \in \prod_{i<n} X_i$ and $r \in \mathbb{Q}_{>0}$. However, when working with $\mathcal{X}$, it is often more convenient to think of the basic open sets as being the sets of the form $\prod_{i<n} B_{r_i}(a_i)$, where each $B_{r_i}(a_i)$ is a basic open set in $\mathcal{X}_i$. The following lemma says that in RCA_0 we can uniformly translate between the two styles of basic open sets when coding open sets. Therefore we may always use the style of open set that is most convenient.

Lemma 3.11 (RCA_0) *Let $(\mathcal{X}_i : i < n)$ be complete separable metric spaces coded by dense sets $(X_i : i < n)$ and metrics $(d_i : i < n)$. Let $X = \prod_{i<n} X_i$ and d code $\mathcal{X} = \prod_{i<n} \mathcal{X}_i$ as in Definition 3.10.*

(i) *There is a function $f \colon \prod_{i<n}(X_i \times \mathbb{Q}_{>0}) \times \mathbb{N} \to X \times \mathbb{Q}_{>0}$ such that, for every*

$$(a_0, r_0, \ldots, a_{n-1}, r_{n-1}) \in \prod_{i<n}(X_i \times \mathbb{Q}_{>0}),$$

$f(a_0, r_0, \ldots, a_{n-1}, r_{n-1}, \cdot)$ enumerates an official code for the unofficial basic open set

$$\prod_{i<n} B_{r_i}(a_i) \subseteq \mathcal{X}.$$

(ii) *For every $\langle \vec{a}, r \rangle \in X \times \mathbb{Q}_{>0}$, $B_r(\vec{a}) = \prod_{i<n} B_r(a_i)$. Thus the function $g \colon X \times \mathbb{Q}_{>0} \to \prod_{i<n}(X_i \times \mathbb{Q}_{>0})$ given by $g(\vec{a}, r) = (a_0, r, \ldots, a_{n-1}, r)$ translates the code $\langle \vec{a}, r \rangle$ for the official basic open set $B_r(\vec{a})$ to the code $(a_0, r, \ldots, a_{n-1}, r)$ for the equivalent unofficial basic open set $\prod_{i<n} B_r(a_i)$.*

Proof (i) The function f exists by Lemma 3.7 because the unofficial basic open set $\prod_{i<n} B_{r_i}(a_i)$ is Σ^0_1 uniformly in $(a_0, r_0, \ldots, a_{n-1}, r_{n-1})$.

(ii) Let $\vec{x} \in \mathcal{X}$. Then $\vec{x} \in B_r(\vec{a})$ if and only if $d(\vec{a}, \vec{x}) < r$ if and only if $\max_{i<n} d_i(a_i, x_i) < r$, if and only if $(\forall i < n)(d_i(a_i, x_i) < r)$, if and only if $\vec{x} \in \prod_{i<n} B_r(a_i)$. Thus $B_r(\vec{a}) = \prod_{i<n} B_r(a_i)$. $\qquad\square$

4 Semi-continuous functions in second-order arithmetic

Although continuous functions have been extensively studied in the context of second-order arithmetic, lower semi-continuous functions have received less attention. Fortunately, they admit a natural representation in the spirit of Definition 3.6.

Definition 4.1 (RCA$_0$) Let $\mathcal{X} = \widehat{X}$ be a complete separable metric space. A *code for a lower semi-continuous function* $f : \mathcal{X} \to \overline{\mathbb{R}}$ is a set $\Psi \subseteq \mathbb{N} \times X \times \mathbb{Q}_{>0} \times \mathbb{Q}$ that satisfies the properties below. Let us write $B_r(a) \xrightarrow{\Psi} q$ for $\exists n((n, a, r, q) \in \Psi)$. Then, for all $a, a' \in X$, all $q, q' \in \mathbb{Q}$, and all $r, r' \in \mathbb{Q}_{>0}$, Ψ must satisfy:

(LSC1) if $B_r(a) \xrightarrow{\Psi} q$ and $B_{r'}(a') \Subset B_r(a)$, then $B_{r'}(a') \xrightarrow{\Psi} q$, and

(LSC2) if $B_r(a) \xrightarrow{\Psi} q$ and $q' < q$, then $B_r(a) \xrightarrow{\Psi} q'$.

For $x \in \mathcal{X}$ and $y \in \overline{\mathbb{R}}$ we define $f(x) = y$ if

$$y = \sup \left\{ q \in \mathbb{Q} : (\exists \langle a, r \rangle \in X \times \mathbb{Q}_{>0})(B_r(a) \xrightarrow{\Psi} q \wedge d(x, a) < r) \right\}.$$

The relation f thus defined is a *lower semi-continuous function* if $(\forall x \in \mathcal{X})(\exists y \in \overline{\mathbb{R}})$ $(f(x) = y)$. The *support* of f, denoted supp(f), is the collection $\{x \in \mathcal{X} : f(x) \in \mathbb{R}\}$. We do not assume that supp(f) exists as any kind of coded set in RCA$_0$. An assertion of the form $x \in \text{supp}(f)$ is to be taken as an abbreviation for $(\exists y \in \mathbb{R})(f(x) = y)$.

If $b \in \mathbb{Q}$ is such that for every $a \in X$ and $r \in \mathbb{Q}_{>0}$ there is $m \in \mathbb{N}$ such that $(m, a, r, b) \in \Psi$, we say that Ψ is a code of a lower semi-continuous function $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq b}$. If $b = 0$, we call the function coded by Ψ a *potential*.

Suppose that $\mathcal{X}$ is a complete separable metric space. As with Definition 3.6, the idea behind Definition 4.1 is that Ψ enumerates pairs $\langle B_r(a), q \rangle$ with the property that if f is the function being coded by Ψ, then f maps $B_r(a) \cap \text{dom}(f)$ into $[q, \infty]$. One may define *upper semi-continuous partial functions* from $\mathcal{X}$ to $\overline{\mathbb{R}}$ by appropriately dualizing Definition 4.1, in which case we write $B_r(a) \xrightarrow{\Psi} q$ for the dual of $B_r(a) \xrightarrow{\Psi} q$. We remark that according to our definition, any code for a lower semi-continuous function defines a function $f : \mathcal{X} \to \overline{\mathbb{R}}$, as, recalling that $\sup \varnothing = -\infty$, we see that every subset of $\mathbb{R}$ has a supremum in $\overline{\mathbb{R}}$. However, this fact is not provable in RCA$_0$. For this reason, lower semi-continuous functions are defined with the explicit assumption that f is defined everywhere. See Remark 4.4 for further discussion.

Next we show that RCA$_0$ proves a version of Lemma 1.1, which we take as evidence that Definition 4.1 is a reasonable definition of semi-continuity for use in RCA$_0$. Indeed, we prove an RCA$_0$ version of the fact that a function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous if and only if $\{\langle x, y \rangle \in \mathcal{X} \times \mathbb{R} : f(x) \leq y\}$ is closed.

Proposition 4.2 (RCA$_0$) *Let $\mathcal{X}$ be a complete separable metric space.*

(i) *If $f : \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous, then $\{\langle x, y \rangle \in \mathcal{X} \times \mathbb{R} : f(x) \leq y\}$ is closed.*

(ii) *If $\mathcal{C} \subseteq \mathcal{X} \times \mathbb{R}$ is a closed set such that*

- *for every $x \in \mathcal{X}$ and $y, z \in \mathbb{R}$, if $\langle x, y \rangle \in \mathcal{C}$ and $z \geq y$, then $\langle x, z \rangle \in \mathcal{C}$, and*
- *for every $x \in \mathcal{X}$, $\inf\{y \in \mathbb{R} : \langle x, y \rangle \in \mathcal{C}\}$ exists*

then there is a lower semi-continuous function $f : \mathcal{X} \to \overline{\mathbb{R}}$ such that $\mathcal{C} = \{\langle x, y \rangle \in \mathcal{X} \times \mathbb{R} : f(x) \leq y\}$.

Proof (i) Let f be coded by Φ. Then $f(x) > y$ if and only if there exists $(a, r, q) \in X \times \mathbb{Q}_{>0} \times \mathbb{Q}$ such that $B_r(a) \overset{\Phi}{\to} q \wedge d(x, a) < r \wedge y < q$. Thus, $\{\langle x, y \rangle \in \mathcal{X} \times \mathbb{R} : f(x) > y\}$ is Σ^0_1-definable, hence it is open by Lemma 3.7.

(ii) Let $\mathcal{C}$ be a closed set as in the statement of (ii), and let $\mathcal{U}$ be the complement of $\mathcal{C}$. Enumerate a code Φ for a lower semi-continuous function f by enumerating $B_r(a) \overset{\Phi}{\to} q$ whenever $\mathcal{U}$ enumerates an open set $B_s(b) \times (u, v)$ with $B_r(a) \Subset B_s(b)$ and $q < v$. We show that, for every $x \in \mathcal{X}$, $f(x)$ exists and equals $\inf\{z \in \mathbb{R} : \langle x, z \rangle \in \mathcal{C}\}$ (which exists by the assumptions on $\mathcal{C}$). So let $x \in \mathcal{X}$, and let $y = \inf\{z \in \mathbb{R} : \langle x, z \rangle \in \mathcal{C}\}$. We need to show that

$$y = \sup\{q \in \mathbb{Q} : (\exists \langle a, r \rangle \in X \times \mathbb{Q}_{>0})(B_r(a) \overset{\Phi}{\to} q \wedge d(x, a) < r)\}.$$

Suppose that $q \in \mathbb{Q}$ is such that there is a ball $B_r(a) \subseteq \mathcal{X}$ with $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\to} q$. By the definition of Φ, there is a $B_s(b) \times (u, v) \subseteq \mathcal{U}$ where $B_r(a) \Subset B_s(b)$ and $q < v$. From this we can conclude that if $z \in \mathbb{R}$ is such that $\langle x, z \rangle \in \mathcal{C}$, then $q < z$. If not, then by the first assumption on $\mathcal{C}$, there would be a z such that $\langle x, z \rangle \in \mathcal{C}$ and $z \in (u, v)$. This implies that $\langle x, z \rangle \in B_s(b) \times (u, v) \subseteq \mathcal{U}$, which contradicts that $\mathcal{U}$ is the complement of $\mathcal{C}$. Thus $q \leq y$ because $y = \inf\{z \in \mathbb{R} : \langle x, z \rangle \in \mathcal{C}\}$.

We have shown that y is an upper bound on the $q \in \mathbb{Q}$ for which there is a $B_r(a) \subseteq \mathcal{X}$ with $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\to} q$. We need to show that y is least. So consider a $z < y$. Then $\langle x, z \rangle \notin \mathcal{C}$ because $y = \inf\{z \in \mathbb{R} : \langle x, z \rangle \in \mathcal{C}\}$. Then there is a $B_s(a) \times (u, v)$ enumerated into $\mathcal{U}$ that contains $\langle x, z \rangle$. Let $r < s$ be such that $x \in B_r(a) \Subset B_s(a)$, and let $q \in \mathbb{Q} \cap (z, v)$. Then $B_r(a) \overset{\Phi}{\to} q$ by the definition of Φ. Thus $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\to} q$, but $z < q$. So z is not an upper bound on the $q \in \mathbb{Q}$ for which there is a $B_r(a) \subseteq \mathcal{X}$ with $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\to} q$. Thus y is indeed the required supremum.

We have shown that, for all $x \in \mathcal{X}$, $f(x) = \inf\{z \in \mathbb{R} : \langle x, z \rangle \in \mathcal{C}\}$. It is now easy to see that, for $\langle x, y \rangle \in \mathcal{X} \times \mathbb{R}$, $f(x) \leq y$ if and only if $\inf\{z \in \mathbb{R} : \langle x, z \rangle \in \mathcal{C}\} \leq y$ if and only if $\langle x, y \rangle \in \mathcal{C}$. □

The analogous properties of upper semi-continuous functions can be proved by dualizing Proposition 4.2.

The next proposition is an RCA$_0$ version of the fact that a function $f : \mathcal{X} \to \mathbb{R}$ is continuous if and only if it is upper semi-continuous and lower semi-continuous.

Proposition 4.3 (RCA$_0$) *Let $\mathcal{X}$ be a complete separable metric space.*

(i) *If $f : \mathcal{X} \to \mathbb{R}$ is continuous, then there are a lower semi-continuous $g : \mathcal{X} \to \mathbb{R}$ and an upper semi-continuous $h : \mathcal{X} \to \mathbb{R}$ such that $(\forall x \in \mathcal{X})(f(x) = g(x) = h(x))$.*

(ii) *If $g : \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous, $h : \mathcal{X} \to \overline{\mathbb{R}}$ is upper semi-continuous, and $(\forall x \in \mathcal{X})(g(x) = h(x))$, then there is a continuous partial $f : \mathcal{X} \to \mathbb{R}$ with $\mathrm{dom}(f) = \mathrm{supp}(g) = \mathrm{supp}(h)$ such that $(\forall x \in \mathrm{dom}(f))(f(x) = g(x) = h(x))$.*

Proof (i) Let Φ be a code for a continuous $f : \mathcal{X} \to \mathbb{R}$. Define a code Γ for a lower semi-continuous $g : \mathcal{X} \to \overline{\mathbb{R}}$ by enumerating $B_r(a) \overset{\Gamma}{\to} q$ whenever $B_s(b) \overset{\Phi}{\to} (u, v)$

for a $B_s(b)$ with $B_r(a) \not\Subset B_s(b)$ and a (u, v) with $q < u$. It is then easy to check that $g(x)$ is defined and equal to $f(x)$ for all $x \in \mathcal{X}$. The upper semi-continuous partial function $h \colon \mathcal{X} \to \overline{\mathbb{R}}$ such that $(\forall x \in \mathrm{dom}(f))(f(x) = h(x))$ is defined dually.

(ii) Let Γ be a code for a lower semi-continuous $g \colon \mathcal{X} \to \overline{\mathbb{R}}$ and let Ψ be a code for an upper semi-continuous $h \colon \mathcal{X} \to \overline{\mathbb{R}}$ such that $(\forall x \in \mathcal{X})(g(x) = h(x))$. Define a code Φ for a continuous partial function $f \colon \mathcal{X} \to \mathbb{R}$ by enumerating $B_r(a) \overset{\Phi}{\to} (u, v)$ whenever there are $B_s(b), B_t(c) \subseteq \mathcal{X}$ and $p, q \in \mathbb{Q}$ such that $B_r(a) \not\Subset B_s(b)$, $B_r(a) \not\Subset B_t(c)$, $[p, q] \subseteq (u, v)$, $B_s(b) \overset{\Gamma}{\to} p$, and $B_t(c) \overset{\Psi}{\to} q$. It is easy to see that Φ satisfies (CF2) and (CF3) of Definition 3.6. To see that Φ satisfies (CF1) of Definition 3.6, first observe that if $B_r(a) \overset{\Phi}{\to} (u, v)$, then $g(a) > u$ and $h(a) < v$. However, $g(a) = h(a)$, so their common value must be finite and in the interval (u, v). Therefore, if $B_r(a) \overset{\Phi}{\to} (u, v)$ and $B_r(a) \overset{\Phi}{\to} (u', v')$, then (u, v) and (u', v') both contain $g(a)$, and therefore (u, v) and (u', v') intersect.

We show that $\mathrm{dom}(f) = \mathrm{supp}(g) = \mathrm{supp}(h)$ and that $(\forall x \in \mathrm{dom}(f))(f(x) = g(x) = h(x))$. First, suppose that $x \notin \mathrm{supp}(g)$. Then $g(x) = \pm\infty$. Suppose that $g(x) = -\infty$. The only way this can happen is if there is no $B_s(b) \overset{\Gamma}{\to} p$ with $x \in B_s(b)$. But if there is no $B_s(b) \overset{\Gamma}{\to} p$ with $x \in B_s(b)$, then there is also no $B_r(a) \overset{\Phi}{\to} (u, v)$ with $x \in B_r(a)$. Thus $x \notin \mathrm{dom}(f)$. If instead $g(x) = \infty$, then also $h(x) = \infty$. In this case, there can be no $B_t(c) \overset{\Psi}{\to} q$ with $x \in B_t(c)$, which similarly implies that $x \notin \mathrm{dom}(f)$.

Now suppose that $x \in \mathrm{supp}(g)$, and let $\varepsilon \in \mathbb{Q}_{>0}$. Let $(u, v) \subseteq \mathbb{R}$ be such that $g(x) \in (u, v)$ and $v - u < \varepsilon$. By the definition of $g(x)$, there are a $B_s(b) \subseteq \mathcal{X}$ and a $p \in (u, v)$ such that $x \in B_s(b)$ and $B_s(b) \overset{\Gamma}{\to} p$. Likewise, as $h(x) = g(x) \in (u, v)$, there are a $B_t(c) \subseteq \mathcal{X}$ and a $q \in (u, v)$ such that $x \in B_t(c)$ and $B_t(c) \overset{\Psi}{\to} q$. Notice then that $p \leq g(x) \leq q$. Let $B_r(a) \subseteq \mathcal{X}$ be such that $x \in B_r(a)$, $B_r(a) \not\Subset B_s(b)$, and $B_r(a) \not\Subset B_t(c)$. Then $B_s(b), B_t(c), p$, and q witness that $B_r(a) \overset{\Phi}{\to} (u, v)$. Thus for every $\varepsilon \in \mathbb{Q}_{>0}$, there are a $B_r(a) \subseteq \mathcal{X}$ with $x \in B_r(a)$ and a $(u, v) \subseteq \mathbb{R}$ with $v - u < \varepsilon$ and $g(x) \in (u, v)$. This means that $x \in \mathrm{dom}(f)$ and $f(x) = g(x) = h(x)$.
□

Remark 4.4 Our definition of the values of lower semi-continuous functions by suprema of rationals is similar to the definition of Borel measures within RCA_0 in [14, Section X.1]. As with Borel measures, one could understand the values of lower semi-continuous functions in a comparative way instead of requiring that the defining suprema exist. That is, if f is lower semi-continuous, then it is still possible to make sense of inequalities like $f(x) \leq r$ in RCA_0 even if the supremum defining $f(x)$ does not exist. Indeed, in ω-models, the values of lower semi-continuous functions are defined as (relative) left-c.e. reals. With this perspective, Proposition 4.2 is still true without the existence of infima. The proof of Proposition 4.3 shows that a code Ψ for a partial continuous function $g \colon \mathcal{X} \to \mathbb{R}$ can be considered as a pair of codes (Ψ_-, Ψ_+) for lower and upper semi-continuous functions by putting $B_r(a) \overset{\Psi_-}{\to} u$ and $B_r(a) \overset{\Psi_+}{\to} v$ if $B_r(a) \overset{\Psi}{\to} (u, v)$, and then $x \in \mathrm{dom}(g)$ if and only if the values at x

defined by Ψ_- and Ψ_+ coincide. If the values coincide in a comparative sense, then it exists as a real number within RCA_0. Similar modifications may be available for other theorems presented here.

5 Honestly-coded potentials

Recall that we refer to lower semi-continuous functions $f \colon \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ as *potentials*. According to Definition 4.1, for $\mathcal{X}$ a complete separable metric space, Φ a code for a lower semi-continuous function $f \colon \mathcal{X} \to \overline{\mathbb{R}}$, $B_r(a) \subseteq \mathcal{X}$, and $q \in \mathbb{Q}$, we know that $f(x) \geq q$ for all $x \in B_r(a)$ if it is the case that $B_r(a) \overset{\Phi}{\rightharpoonup} q$. We call a code Φ for f *honest* if it contains all of the information of this sort. That is, if the 'if' is an 'if and only if'.

Definition 5.1 (RCA_0) Let $\mathcal{X}$ be a complete separable metric space and let $f \colon \mathcal{X} \to \overline{\mathbb{R}}$ be lower semi-continuous. A code Φ for f is called *honest* if, for every $B_r(a) \subseteq \mathcal{X}$ and every $q \in \mathbb{Q}$, $B_r(a) \overset{\Phi}{\rightharpoonup} q$ if and only if $(\forall x \in B_r(a))(f(x) \geq q)$. If f has an honest code, then we say that f is *honestly-coded*.

Notice that if $f \colon \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous and $\Phi \subseteq \mathbb{N} \times X \times \mathbb{Q}_{>0} \times \mathbb{Q}$ is any set such that, for every $B_r(a) \subseteq \mathcal{X}$ and every $q \in \mathbb{Q}$, $B_r(a) \overset{\Phi}{\rightharpoonup} q$ if and only if $(\forall x \in B_r(a))(f(x) \geq q)$, then Φ is automatically a code (and hence an honest code) for f as in Definition 4.1.

Every lower semi-continuous function admits an honest code, although such a code cannot always be constructed in a weak theory.

Lemma 5.2 *Let $\mathcal{X}$ be a complete separable metric space.*

(i) (WKL_0) *If $\mathcal{X}$ is compact and $f \colon \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous, then there is a code Φ for f such that, for every $B_r(a) \subseteq \mathcal{X}$ and every $q \in \mathbb{Q}$, $B_r(a) \overset{\Phi}{\rightharpoonup} q$ if and only if $(\forall x \in \overline{B_r(a)})(f(x) > q)$.*

(ii) (ACA_0) *If $\mathcal{X}$ is compact and $f \colon \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous, then f has an honest code.*

(iii) (ACA_0) *If $f \colon \mathcal{X} \to \mathbb{R}$ is continuous, then there is an honestly-coded lower semi-continuous $g \colon \mathcal{X} \to \overline{\mathbb{R}}$ such that $(\forall x \in \mathcal{X})(g(x) = f(x))$.*

(iv) ($\Pi_1^1\text{-}\mathsf{CA}_0$) *If $f \colon \mathcal{X} \to \overline{\mathbb{R}}$ is lower semi-continuous, then f has an honest code.*

Proof (i) Work in WKL_0. For $\langle a, r \rangle \in X \times \mathbb{Q}$ and a sequence $(\langle b_i, s_i \rangle : i < n)$ of elements of $X \times \mathbb{Q}$, the assertion that $\overline{B_r(a)} \subseteq \bigcup_{i<n} B_{s_i}(b_i)$ is Σ_1^0 uniformly in $\langle a, r \rangle$ and $(\langle b_i, s_i \rangle : i < n)$, essentially by a uniform version of Simpson [14, Theorem IV.1.7].

Let Ψ be a code for f. Define a new code Φ by enumerating $B_r(a) \overset{\Phi}{\rightharpoonup} q$ whenever there are balls $(B_{s_i}(b_i) : i < n)$ and rationals $\langle t_i : i < n \rangle$ such that $t_i > q$ and $B_{s_i}(b_i) \overset{\Psi}{\rightharpoonup} t_i$ for each $i < n$ and $\overline{B_r(a)} \subseteq \bigcup_{i<n} B_{s_i}(b_i)$.

Suppose that $B_r(a) \overset{\Phi}{\rightharpoonup} q$. Then $\overline{B_r(a)}$ is covered by balls $(B_{s_i}(b_i) : i < n)$ where $f(x) > q$ for all $x \in B_{s_i}(b_i)$ and all $i < n$. Thus $f(x) > q$ for all $x \in \overline{B_r(a)}$.

Conversely, suppose that $f(x) > q$ for all $x \in \overline{B_r(a)}$. Then any enumeration $(B_{s_i}(b_i) : i \in \mathbb{N})$ of every ball $B_{s_i}(b_i)$ for which there is a $t_i \in \mathbb{Q}$ with $t_i > q$ and $B_{s_i}(b_i) \overset{\Psi}{\to} t_i$ is an open cover of $\overline{B_r(a)}$ and hence (essentially by Simpson [14, Theorem IV.1.6]) has a finite subcover. Thus there are a sequence $(B_{s_i}(b_i) : i < n)$ of balls and a sequence $\langle t_i : i < n \rangle$ of rationals witnessing that $B_r(a) \overset{\Phi}{\to} q$.

We have shown that, for every $B_r(a) \subseteq \mathcal{X}$ and every $q \in \mathbb{Q}$, $B_r(a) \overset{\Phi}{\to} q$ if and only if $(\forall x \in \overline{B_r(a)})(f(x) > q)$. Using this fact, it is straightforward to verify that Φ is also a code for f in the sense of Definition 4.1.

(ii) Work in ACA_0. Let Φ be a code for f as in (i). Then, define a code Γ by setting $B_r(a) \overset{\Gamma}{\to} q$ whenever it is the case that $B_s(b) \overset{\Phi}{\to} t$ for every $t < q$ and every $B_s(b) \subseteq \mathcal{X}$ with $B_s(b) \not\subseteq B_r(a)$.

Suppose that $x \in B_r(a)$ and $B_r(a) \overset{\Gamma}{\to} q$. Let $B_s(b)$ be such that $x \in B_s(b)$ and $B_s(b) \not\subseteq B_r(a)$. Then $B_s(b) \overset{\Phi}{\to} t$ for every $t < q$, which means that $f(x) > t$ for every $t < q$. Hence $f(x) \geq q$. Conversely, suppose that $f(x) \geq q$ for all $x \in B_r(a)$. If $B_s(b) \subseteq \mathcal{X}$ and $t \in \mathbb{Q}$ satisfy $t < q$ and $B_s(b) \not\subseteq B_r(a)$, then $f(x) \geq q > t$ for all $x \in \overline{B_s(b)}$ because $\overline{B_s(b)} \subseteq B_r(a)$. Thus $B_s(b) \overset{\Phi}{\to} t$. Hence $B_s(b) \overset{\Phi}{\to} t$ whenever $t < q$ and $B_s(b) \not\subseteq B_r(a)$, so $B_r(a) \overset{\Gamma}{\to} q$.

We have shown that $B_r(a) \overset{\Gamma}{\to} q$ if and only if $(\forall x \in B_r(a))(f(x) \geq q)$. So Γ is an honest code for f.

(iii) In ACA_0, we can define a code Φ by $B_r(a) \overset{\Phi}{\to} q$ if and only if $f(b) \geq q$ for all $b \in B_r(a) \cap X$. One readily checks that Φ is an honest code for a lower semi-continuous g that is equal to f.

(iv) In $\Pi_1^1\text{-}\mathsf{CA}_0$, we can directly define an honest code Φ for f by setting $B_r(a) \overset{\Phi}{\to} q$ if and only if $(\forall x \in B_r(a))(f(x) \geq q)$.   □

6 Continuous envelopes

Honestly-coded lower semi-continuous functions facilitate the calculation of infima, which helps us approximate lower semi-continuous functions by continuous ones. To be precise, lower semi-continuous functions bounded below can be written as the increasing limit of continuous functions using the following construction.

Definition 6.1 Given a potential f on a complete separable metric space $\mathcal{X}$ and an $\alpha \in \mathbb{R}_{>0}$, define the *lower α-envelope of f* by

$$f_\alpha(x) = \inf_{y \in \mathcal{X}} (f(y) + \alpha d(x, y)).$$

Although not needed for our purposes, it is instructive to observe that if $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ is a potential, then f_n converges pointwise to f as $n \to \infty$. What we do need to prove is that continuous envelopes exist, and this can typically not be done within

a weak theory. The construction of envelopes hinges on the following more general lemma.

Lemma 6.2 (ACA$_0$) *Let $\mathcal{X}$ and $\mathcal{Y}$ be complete separable metric spaces, let $h: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_{\geq 0}$ be uniformly continuous, and let $f: \mathcal{Y} \to \overline{\mathbb{R}}_{\geq 0}$ be lower semi-continuous and honestly-coded with non-empty support. Then there is a uniformly continuous function $g: \mathcal{X} \to \mathbb{R}_{\geq 0}$ such that*

$$(\forall x \in \mathcal{X})[g(x) = \inf_{y \in \mathcal{Y}} (h(x, y) + f(y))].$$

Proof Let Φ be a code for h, and let Ψ be an honest code for f. We first show that if $x \in \mathcal{X}$, then $\inf_{y \in \mathcal{Y}}(h(x, y) + f(y))$ indeed exists, which is a consequence of the following Claim. □

Claim Let $x \in \mathcal{X}$ and $q \in \mathbb{Q}$. Then there is a $y \in \mathcal{Y}$ such that $h(x, y) + f(y) < q$ if and only if there are $B_r(\langle a, b \rangle) \subseteq \mathcal{X} \times \mathcal{Y}$, $(u, v) \subseteq \mathbb{R}$, and $p \in \mathbb{Q}$ such that $x \in B_r(a)$, $B_r(\langle a, b \rangle) \overset{\Phi}{\to} (u, v)$, $\neg(B_r(b) \overset{\Psi}{\to} p)$, and $v + p < q$.

Proof of Claim Be aware that $B_r(\langle a, b \rangle) = B_r(a) \times B_r(b)$ by Definition 3.10. Suppose that $y \in \mathcal{Y}$ is such that $h(x, y) + f(y) < q$. Let $u, v, p \in \mathbb{Q}$ be such that $h(x, y) \in (u, v)$, $p > f(y)$, and $v + p < q$. Then there must be a $B_r(\langle a, b \rangle) \subseteq \mathcal{X} \times \mathcal{Y}$ such that $\langle x, y \rangle \in B_r(\langle a, b \rangle)$ and $B_r(\langle a, b \rangle) \overset{\Phi}{\to} (u, v)$. Furthermore, since Ψ is a code for f, it must be that $\neg(B_r(b) \overset{\Psi}{\to} p)$ because $y \in B_r(b)$ but $f(y) < p$. Thus $B_r(\langle a, b \rangle)$, (u, v), and p are as required.

Conversely, suppose that there are such $B_r(\langle a, b \rangle)$, (u, v), and p. Because Ψ is honest and $\neg(B_r(b) \overset{\Psi}{\to} p)$, there must be a $y \in B_r(b)$ such that $f(y) < p$. Also, $h(x, y) \leq v$ because $\langle x, y \rangle \in B_r(a) \times B_r(b) = B_r(\langle a, b \rangle)$ and $B_r(\langle a, b \rangle) \overset{\Phi}{\to} (u, v)$. Hence $h(x, y) + f(y) < v + p < q$. □

Thus, given $x \in \mathcal{X}$, let Q be the set of all $q \in \mathbb{Q}$ for which there are $\langle a, b, r \rangle \in X \times Y \times \mathbb{Q}$ and $u, v, p \in \mathbb{Q}$ such that $B_r(\langle a, b \rangle)$, (u, v), and p witness that there is a $y \in \mathcal{Y}$ such that $h(x, y) + f(y) < q$ as in the Claim. The set Q is bounded below by 0 and is non-empty since $\text{supp}(f) \neq \varnothing$, so $\inf Q$ exists (essentially by Simpson [14, Theorem III.2.2]), and, by the Claim, $\inf Q = \inf_{y \in \mathcal{Y}}(h(x, y) + f(y))$. Henceforth, for each $x \in \mathcal{X}$, let α_x denote $\inf_{y \in \mathcal{Y}}(h(x, y) + f(y))$.

We can make the above argument uniformly for all $a \in X$ by letting $A \subseteq X \times \mathbb{Q}$ be the set of all $\langle a, q \rangle$ for which there are $r \in \mathbb{Q}$, $b \in Y$, and $u, v, p \in \mathbb{Q}$ such that $B_r(\langle a, b \rangle)$, (u, v), and p witness that there is a $y \in \mathcal{Y}$ such that $h(a, y) + f(y) < q$. Then from A we can define the sequence $(\alpha_a : a \in X)$ because, given a sequence of sets of rationals all bounded from below, ACA$_0$ suffices to produce the corresponding sequence of infima (that is ACA$_0$ proves item 4 of Simpson [14, Theorem III.2.2] uniformly).

Define a code Γ for a continuous partial function $g: \mathcal{X} \to \mathbb{R}$ by defining $B_r(a) \overset{\Gamma}{\to} (u, v)$ if and only if $(\forall c \in B_r(a) \cap X)[\alpha_c \in (u, v)]$. It is easy to see that Γ satisfies the requirements of Definition 3.6. We need to show that g is indeed uniformly continuous on all of $\mathcal{X}$ and that $(\forall x \in \mathcal{X})(g(x) = \alpha_x)$.

Claim Let $\varepsilon, \delta \in \mathbb{Q}_{>0}$ be such that

$$(\forall \langle x_0, y_0 \rangle, \langle x_1, y_1 \rangle \in \mathcal{X} \times \mathcal{Y})$$
$$(d_{\mathcal{X} \times \mathcal{Y}}(\langle x_0, y_0 \rangle, \langle x_1, y_1 \rangle) < \delta \rightarrow |h(x_0, y_0) - h(x_1, y_1)| \leq \varepsilon).$$

Let $x_0, x_1 \in \mathcal{X}$. Then $d_{\mathcal{X}}(x_0, x_1) < \delta \rightarrow |\alpha_{x_0} - \alpha_{x_1}| \leq \varepsilon$.

Proof of Claim We show that $(\forall \eta \in \mathbb{Q}_{>0})(\alpha_{x_1} \leq \alpha_{x_0} + \varepsilon + \eta)$, which implies that $\alpha_{x_1} \leq \alpha_{x_0} + \varepsilon$. By a symmetric argument, we also have that $\alpha_{x_0} \leq \alpha_{x_1} + \varepsilon$, which gives the desired $|\alpha_{x_0} - \alpha_{x_1}| \leq \varepsilon$.

Thus let $\eta \in \mathbb{Q}_{>0}$. Let $y \in \mathcal{Y}$ be such that $h(x_0, y) + f(y) \leq \alpha_{x_0} + \eta$. Then

$$\alpha_{x_1} \leq h(x_1, y) + f(y) = (h(x_1, y) - h(x_0, y)) + (h(x_0, y) + f(y)) \leq \varepsilon + \alpha_{x_0} + \eta,$$

where the last inequality is by the choice of y and the fact that $d_{\mathcal{X} \times \mathcal{Y}}(\langle x_0, y \rangle, \langle x_1, y \rangle) = d_{\mathcal{X}}(x_0, x_1) < \delta$. $\qquad \square$

Given $x \in \mathcal{X}$ and $\varepsilon \in \mathbb{Q}_{>0}$, let $\delta \in \mathbb{Q}_{>0}$ be such that

$$(\forall \langle x_0, y_0 \rangle, \langle x_1, y_1 \rangle \in \mathcal{X} \times \mathcal{Y})$$
$$(d_{\mathcal{X} \times \mathcal{Y}}(\langle x_0, y_0 \rangle, \langle x_1, y_1 \rangle) < \delta \rightarrow |h(x_0, y_0) - h(x_1, y_1)| \leq \varepsilon/4),$$

and let $a \in X$ be such that $x \in B_\delta(a)$. Let $(u, v) \subseteq \mathbb{R}$ be such that $[\alpha_a - \varepsilon/4, \alpha_a + \varepsilon/4] \subseteq (u, v)$ and $v - u < \varepsilon$. The Claim tells us that if $c \in B_\delta(a) \cap X$, then $\alpha_c \in [\alpha_a - \varepsilon/4, \alpha_a + \varepsilon/4] \subseteq (u, v)$, which means that $B_\delta(a) \xrightarrow{\Gamma} (u, v)$. Likewise, the Claim implies that $\alpha_x \in (u, v)$. Thus for every $x \in \mathcal{X}$ and every $\varepsilon \in \mathbb{Q}_{>0}$, there are a ball $B_\delta(a) \subseteq \mathcal{X}$ and a $(u, v) \subseteq \mathbb{R}$ such that $x \in B_\delta(a)$, $\alpha_x \in (u, v)$, $v - u < \varepsilon$, and $B_\delta(a) \xrightarrow{\Gamma} (u, v)$. Thus $g(x)$ is defined and equals α_x for all $x \in \mathcal{X}$. Furthermore, the Claim clearly implies that g is uniformly continuous.

Lemma 6.3 (ACA$_0$) *Let $\mathcal{X}$ be a complete separable metric space, let $f \colon \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be lower semi-continuous and honestly coded with non-empty support, and let $\alpha \in \mathbb{R}_{\geq 0}$. Then there is a uniformly continuous function $f_\alpha \colon \mathcal{X} \to \mathbb{R}_{\geq 0}$ such that, for all $x \in \mathcal{X}$,*

$$f_\alpha(x) = \inf_{y \in \mathcal{X}} (f(y) + \alpha d(x, y)).$$

Proof This follows from Lemma 6.2 because the function $h \colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$ given by $h(x, y) = \alpha d_{\mathcal{X}}(x, y)$ is uniformly continuous and bounded below by 0. $\qquad \square$

7 Critical points of continuous functions

In this section we formalize Ekeland's variational principle and show that it has natural restrictions provable in weak systems. Recall that a *potential* on a complete separable metric space $\mathcal{X}$ is a lower semi-continuous function $f \colon \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$. Recall also that

ε-*critical points* were defined in Definition 1.2 and that this definition may readily be formalized in RCA$_0$ using coded lower semi-continuous functions. With this, we are ready to define our formalized version (or, version*s*) of Ekeland's variational principle.

Definition 7.1 Given definable classes $\mathfrak{X}$ of coded complete separable metric spaces and $\mathfrak{F}$ of coded potentials, the *(formalized) free variational principle (FVP) for $\mathcal{X} \in \mathfrak{X}$ and $f \in \mathfrak{F}$* is the statement that, if $\mathcal{X} \in \mathfrak{X}$ and $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ in $\mathfrak{F}$ is such that supp$(f) \neq \varnothing$, then for every $\varepsilon > 0$ there is an $x_* \in$ supp(f) such that

$$\forall x \in \mathrm{supp}(f)\Big(\big(\varepsilon d(x_*, x) \leq f(x_*) - f(x)\big) \to x = x_*\Big).$$

The *localized variational principle* (LVP) is defined similarly, except that $x_0 \in$ supp(f) is given and x_* is chosen so that $\varepsilon d(x_0, x_*) \leq f(x_0) - f(x_*)$.

When not mentioned, we assume that $\mathfrak{X}$ is the class of all coded complete separable metric spaces and $\mathfrak{F}$ is the class of all coded potentials. Given $c > 0$, to indicate that $f(x) \leq c$ for all $x \in$ dom(f), we will say that f is *c-bounded,* and if we want to fix the value of ε to c, we call the statement the *free/localized variational principle for $\varepsilon = c$.* The free/localized variational principles for $\varepsilon = c$ and f c-bounded will be denoted the *c-FVP* and the *c-LVP*, respectively.

Critical points are sometimes also called *pseudo-minima,* as they may serve as approximate minima of functions that do not actually attain their minimum value.

Lemma 7.2 (RCA$_0$) *If $\mathcal{X}$ is any complete separable metric space and $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ is lower semi-continuous with non-empty support, then for any $x_* \in \mathcal{X}$, f attains its minimum at x_* if and only if x_* is an ε-critical point of f for all $\varepsilon > 0$.*

Proof It is straightforward to see that if $x \neq x_*$ and $\varepsilon d(x, x_*) \leq f(x_*) - f(x)$ then $f(x) < f(x_*)$, hence if x_* is not ε-critical it cannot be a minimum. Conversely, if f does not attain its minimum at x_*, then we may choose $x \in \mathcal{X}$ with $f(x) < f(x_*)$. By choosing ε small enough, we obtain $\varepsilon d(x, x_*) \leq f(x_*) - f(x)$. $\qquad\square$

The class $\mathfrak{F}$ will contain either continuous or lower semi-continuous functions. Important special cases for us are the 1-FVP and the 1-LVP, which on occasion are equivalent to the FVP and LVP, respectively.

Lemma 7.3 (RCA$_0$) *Let $\mathfrak{X}$ be a class of complete separable metric spaces and $\mathfrak{F}$ be either the class of continuous or of lower semi-continuous potentials.*

(1) *The FVP and the LVP with $\mathcal{X} \in \mathfrak{X}$ and $f \in \mathfrak{F}$ are equivalent, respectively, to the FVP and the LVP with $\varepsilon = 1$, $\mathcal{X} \in \mathfrak{X}$, and $f \in \mathfrak{F}$.*
(2) *If $\mathfrak{X}$ is either the class of all complete separable metric spaces or of all compact metric spaces then the FVP and the LVP with $\mathcal{X} \in \mathfrak{X}$ and $f \in \mathfrak{F}$ bounded are equivalent, respectively, to the 1-FVP and the 1-LVP with $\mathcal{X} \in \mathfrak{X}$ and $f \in \mathfrak{F}$.*
(3) *If $\mathcal{X}$ is (a) a Banach space, (b) the Baire space, (c) the Cantor space, or (d) a closed ball in $\mathbb{R}^n$ (with any norm), then the FVP and the LVP for $\mathcal{X}$ and bounded, lower semi-continuous potentials are equivalent, respectively, to the 1-FVP and the 1-LVP for $\mathcal{X}$ and lower semi-continuous potentials.*

Proof (1) Replace f by $f' = f/\varepsilon$ and note that any 1-critical point for f' is ε-critical for f.

(2) First replace f by $f' = f/b$, so that f' is bounded by 1 and any ε/b-critical point for f' is an ε-critical point for f. As we can no longer freely scale f', we instead scale the metric and consider $d' = \varepsilon d/b$. Then the reader can verify that any 1-critical point for f' with respect to d' is an ε-critical point for f with respect to d.

(3) The general idea is that in all of these cases, the proof of (2) can be simulated (in a possibly *ad-hoc* way) without modifying the space $\mathcal{X}$. For the sake of illustration we sketch the proof in the case where $\mathcal{X}$ is a closed ball in $\mathbb{R}^n$—we give an informal argument and let the reader verify that it can be carried out in RCA$_0$. The FVP and LVP cases are similar, so we focus on the FVP.

Assume that the 1-FVP holds for $\mathcal{X}$. Without loss of generality we may assume that $\mathcal{X} = \overline{B_\rho(0)}$ for some $\rho > 0$. Let f be a bounded lower semi-continuous function on $\mathcal{X}$ with non-empty domain, and let $\varepsilon > 0$. Without loss of generality we may assume that f is bounded by $1/2$ (adjusting ε if needed).

Let $A \subseteq B_\rho(0) \cap \mathbb{Q}^n$ be finite and such that every point in $B_\rho(0)$ is within distance $1/4$ of some point in A. Let $\delta > 0$ be small enough so that $\delta < \min\{\varepsilon\rho, 1/4\}$ and $d(a, a') > 2\delta$ whenever $a, a' \in A$ are distinct. Define $g\colon \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ by

$$g(x) = \begin{cases} f\left(\frac{\rho}{\delta}(x-a)\right) & \text{if } x \in \overline{B_\delta(a)} \text{ for some } a \in A \\ 1 & \text{otherwise.} \end{cases}$$

By the 1-FVP, let x_* be a 1-critical point of g. We claim that $x_* \in \overline{B_\delta(a)}$ for some $a \in A$. If not, then $g(x_*) = 1$ by the definition of g. Let $a \in A$ minimize $d(x_*, a)$, and note that $d(x_*, a) < 1/4$ by our choice of A. Let $y_0 \in \text{supp}(f)$. Then $x_0 := \frac{\delta}{\rho}y_0 + a \in \overline{B_\delta(a)}$ and $g(x_0) = f(y_0) \leq 1/2$, so that

$$g(x_*) - g(x_0) \geq 1 - 1/2 = 1/4 + 1/4 \geq d(x_0, a) + d(a, x_*) \geq d(x_0, x_*).$$

This contradicts that x_* is a 1-critical point of g. Finally, we claim that $y_* := \frac{\rho}{\delta}(x_* - a)$ is an ε-critical point of f. Let $y \in \text{supp}(f)$ with $y \neq x_*$ be arbitrary, and let $x := \frac{\delta}{\rho}y + a \in B_\delta(a)$. Then since x_* is a 1-critical point of g,

$$g(x_*) - g(x) < d(x_*, x),$$

which by the definition of g becomes

$$f(y_*) - f(y) < d\left(\frac{\delta}{\rho}y_* + a, \frac{\delta}{\rho}y + a\right) = \frac{\delta}{\rho}d(y_*, y) < \varepsilon d(y_*, y),$$

as needed.

The proof that the 1-LVP implies the bounded LVP is similar, except that it is more convenient to replace A by the singleton $\{0\}$. The proofs of the other items are analogous, with the main difference being the choice of g. Let us assume that f is bounded by 1. For (3a) we set $g(x) = f\left(\frac{1}{\varepsilon}x\right)$. For items (3b) and (3c) we use the

observation that $d(\sigma^\frown x, \sigma^\frown y) = 2^{-|\sigma|} d(x, y)$ for every σ, x, and y. Let n be large enough so that $2^{-n} < \varepsilon$. For the FVP define $g(\sigma^\frown y) = f(y)$, where $|\sigma| = n$. That is, given x, $g(x)$ chops off $x{\restriction}n$ and applies f to the remaining sequence y. For the LVP, we define $g((0^n)^\frown y) = f(y)$ and $g(x) = 1$ if x does not begin with n zeroes. We leave to the reader to check that in each of this cases, any 1-critical point for g induces an ε-critical point for f, satisfying the localization condition when appropriate. □

With Lemma 7.3 in mind, we often take $\varepsilon = 1$ or, in the bounded case, f bounded by 1, and we write *critical point* instead of 1-*critical point*. Note however that if we fix $\mathcal{X}$ beforehand, the FVP or LVP for $\mathcal{X}$ with bounded f are not necessarily equivalent to the 1-FVP or the 1-LVP for $\mathcal{X}$, respectively, since we may not be able to perform the required scaling transformations without modifying the space. Claim (3) gives a few examples where such transformations may be performed, and while the list is by no means meant to be exhaustive, we conjecture that a space $\mathcal{X}$ can be found so that the 1-FVP and the FVP for $\mathcal{X}$ are not provably equivalent in RCA_0.

Note also that in the cases (3a)–(3c), if the function f is continuous, then so is g. Thus in these cases, the claim also holds for the respective variational principles restricted to continuous functions. Our construction for (3d) does not preserve continuity, but it should also be possible to define some $\tilde{g}$ which is continuous if f was (similar to the 2-envelope of g). Note however that some care would be required to construct $\tilde{g}$ in RCA_0. In any case, the current formulation of the lemma suffices for our purposes.

Let us now begin our analysis by showing that WKL_0 suffices to construct critical points of continuous functions on compact spaces.

Proposition 7.4 (WKL_0) *The* FVP *holds for compact $\mathcal{X}$ and continuous f; in fact, in this setting f attains its minimum.*

Proof WKL_0 proves that every continuous function from a compact complete separable metric space to $\mathbb{R}$ attains a minimum value (see [14, Theorem IV.2.2] for the version of this fact with 'maximum' in place of 'minimum'). Thus let $x_* \in \mathcal{X}$ be a point at which f attains its minimum value. Then x_* is a critical point of f by Lemma 7.2. □

When working over ACA_0, we may drop the assumption that $\mathcal{X}$ is compact. Our proof is based on those in [2,4].

Theorem 7.5 (ACA_0) *The* FVP *holds for arbitrary $\mathcal{X}$ and continuous f.*

Proof Let $\mathcal{X} = \widehat{X}$ be a complete separable metric space with metric d, and in view of Lemma 7.3.1, assume that $\varepsilon = 1$. First we use ACA_0 to collect some information in order to implement the proof. Define $S \subseteq X \times X \times \mathbb{Q}_{>0}$ by

$$S = \{(a, b, q) \in X \times X \times \mathbb{Q}_{>0} : d(a, b) < f(a) - f(b) + q\},$$

and for $a \in X$ and $q \in \mathbb{Q}_{>0}$, write $S(a, q)$ for $\{b \in X : (a, b, q) \in S\}$. Define a sequence of reals $(u_{a,q} : a \in X \wedge q \in \mathbb{Q}_{>0})$ by $u_{a,q} = \inf_{b \in S(a,q)} f(b)$ for each $a \in X$ and $q \in \mathbb{Q}_{>0}$, which always exists because f is bounded from below and $S(a, q)$ is non-empty (as it contains a). Next, define a sequence of reals $(r_a : a \in X)$ by

$r_a = \sup_{q \in \mathbb{Q}_{>0}} u_{a,q}$ for each $a \in X$, which always exists because $(\forall q \in \mathbb{Q}_{>0})(u_{a,q} \leq f(a))$. That the sequences $(u_{a,q} : a \in X \wedge q \in \mathbb{Q}_{>0})$ and $(r_a : a \in X)$ can be defined in ACA_0 follows from the appropriately uniformized version of the $(1) \Rightarrow (4)$ direction of Simpson [14, Theorem III.2.2] (and the analogous implication with 'greatest lower bound' in place of 'least upper bound' in (4)). Observe that $(\forall a \in X)(\forall q_0, q_1 \in \mathbb{Q}_{>0})(q_0 \leq q_1 \rightarrow u_{a,q_1} \leq u_{a,q_0})$, which means that $(\forall a \in X)(r_a = \lim_{q \to 0^+} u_{a,q})$. Finally, define a function $R \colon X \times \mathbb{Q}_{>0} \to \mathbb{Q}$ describing the rates of convergence of the sequences $(u_{a,q} : q \in \mathbb{Q}_{>0})$ for each $a \in X$ by letting $R(a, p) = 2^{-n}$, where n is least such that $(\forall q \in \mathbb{Q}_{>0})(q \leq 2^{-n} \rightarrow |r_a - u_{a,q}| < p)$.

Now we define a sequence $(a_n : n \in \mathbb{N})$ of points in X converging to a critical point of f and a helper sequence $(q_n : n \in \mathbb{N})$ of points in $\mathbb{Q}_{>0}$. Let a_0 be the first point (that is, the point with the least code) in X, and let $q_0 = 2^{-2} R(a_0, 1)$. Given $(a_i : i \leq n)$ and $(q_i : i \leq n)$, let a_{n+1} be the first point in $S(a_n, q_n)$ such that $f(a_{n+1}) < u_{a_n,q_n} + 2^{-(n+1)}$, and let $q_{n+1} = 2^{-(n+3)} \prod_{i \leq n+1} R(a_i, 2^{-i})$. The important property of $(q_n : n \in \mathbb{N})$ is that, for every $n \in \mathbb{N}$,

$$\sum_{k \geq n} q_k < \min\{2^{-n}, R(a_n, 2^{-n})\},$$

which holds because $(\forall n)(\forall k \geq n)[q_k \leq 2^{-(k+2)} R(a_n, 2^{-n})]$.

As the sequence $(f(a_n) : n \in \mathbb{N})$ is bounded from below, let $\ell = \liminf_{n \in \mathbb{N}} f(a_n)$ (which exists by Simpson [14, Lemma III.2.1] with 'lim inf' in place of 'lim sup'). Observe that $\liminf_{n \in \mathbb{N}} u_{a_n,q_n} = \ell$ as well because $\forall n (u_{a_n,q_n} \leq f(a_{n+1}) < u_{a_n,q_n} + 2^{-(n+1)})$. Let $i_0 < i_1 < i_2 < \cdots$ be an increasing sequence of indices such that $\lim_{n \in \mathbb{N}} f(a_{i_n}) = \ell$. We show that the corresponding sequence $(a_{i_n} : n \in \mathbb{N})$ is Cauchy. Given $M \in \mathbb{N}$, let $N > M$ be large enough so that $\forall n, m(N < n < m \rightarrow |f(a_{i_n}) - f(a_{i_m})| \leq 2^{-(M+1)})$. Then, for every $n, m \in \mathbb{N}$ with $N < n < m$, we have that

$$\begin{aligned} d(a_{i_n}, a_{i_m}) &\leq \sum_{k=i_n}^{i_m - 1} d(a_k, a_{k+1}) \\ &\leq \sum_{k=i_n}^{i_m - 1} [f(a_k) - f(a_{k+1}) + q_k] \\ &\leq f(a_{i_n}) - f(a_{i_m}) + \sum_{k \geq i_n} q_k \\ &\leq 2^{-(M+1)} + 2^{-(M+1)} \leq 2^{-M}. \end{aligned} \tag{2}$$

In the above expression, the first inequality is the triangle inequality, the second inequality is because a_{k+1} is chosen from $S(a_k, q_k)$ for every $k \in \mathbb{N}$, the third inequality is obtained by canceling the appropriate terms in the telescoping sum, and the fourth inequality is by the choice of N and by recalling that $\sum_{k \geq i_n} q_k < 2^{-i_n}$ (and that $i_n \geq n > N > M$). Thus $(a_{i_n} : n \in \mathbb{N})$ is Cauchy. By the $(1) \Rightarrow (3)$ direction of Simpson [14, Theorem III.2.2], $(a_{i_n} : n \in \mathbb{N})$ converges to some $x_* \in \mathcal{X}$.

We show that x_* is a critical point of f. By the continuity of f, $f(x_*) = \lim_{n\in\mathbb{N}} f(a_{i_n}) = \ell$. We show that if $y \in \mathcal{X}$ is such that $d(x_*, y) \leq f(x_*) - f(y)$, then $f(y) = \ell$ as well. Thus consider a $y \in \mathcal{X}$ such that $d(x_*, y) \leq f(x_*) - f(y)$. Clearly $f(y) \leq \ell$ because otherwise we would have the contradiction $d(x_*, y) < 0$. So we show that $f(y) \geq \ell$. To do this, fix $n \in \mathbb{N}$. We see that $d(a_{i_n}, x_*) \leq f(a_{i_n}) - f(x_*) + \sum_{k\geq i_n} q_k$ by considering the inequality (2) as $m \to \infty$. Therefore,

$$d(a_{i_n}, y) \leq d(a_{i_n}, x_*) + d(x_*, y)$$
$$\leq f(a_{i_n}) - f(x_*) + \sum_{k\geq i_n} q_k + f(x_*) - f(y)$$
$$= f(a_{i_n}) - f(y) + \sum_{k\geq i_n} q_k.$$

Thus $d(a_{i_n}, y) < f(a_{i_n}) - f(y) + R(a_{i_n}, 2^{-i_n})$ because $\sum_{k\geq i_n} q_k < R(a_{i_n}, 2^{-i_n})$. It follows that $f(y) \geq u_{a_{i_n}, R(a_{i_n}, 2^{-i_n})}$ because y is in the open set

$$\mathcal{U} = \{x \in \mathcal{X} : d(a_{i_n}, x) < f(a_{i_n}) - f(x) + R(a_{i_n}, 2^{-i_n})\}$$

and $S(a_{i_n}, R(a_{i_n}, 2^{-i_n}))$ is dense in $\mathcal{U}$. Observe also that $q_{i_n} \leq R(a_{i_n}, 2^{-i_n})$, which implies that $u_{a_{i_n}, q_{i_n}} - u_{a_{i_n}, R(a_{i_n}, 2^{-i_n})} < 2^{-i_n}$. Therefore $f(y) > u_{a_{i_n}, q_{i_n}} - 2^{-i_n}$. So we have shown that $\forall n (f(y) > u_{a_{i_n}, q_{i_n}} - 2^{-i_n})$. As $(u_{a_{i_n}, q_{i_n}} : n \in \mathbb{N})$ is a subsequence of $(u_{a_n, q_n} : n \in \mathbb{N})$, it follows that $\liminf_{n\in\mathbb{N}} u_{a_{i_n}, q_{i_n}} \geq \liminf_{n\in\mathbb{N}} u_{a_n, q_n} = \ell$. Thus we conclude that $f(y) \geq \ell$, as desired. This completes the proof that if $d(x_*, y) \leq f(x_*) - f(y)$, then $f(y) = f(x_*) = \ell$. So if $y \in \mathcal{X}$ satisfies $d(x_*, y) \leq f(x_*) - f(y)$, then $y = x_*$. Hence x_* is a critical point of f. □

8 Critical points of arbitrary potentials

Now let us consider the case where f is a possibly discontinuous potential. As it turns out, this general case can be reduced to the continuous case by using envelopes. To be precise, in order to find an α-critical point of a lower semi-continuous potential f, it suffices to find an α-critical point of the β-envelope f_β for any $\beta > \alpha$.

Lemma 8.1 (RCA$_0$) *Let $\mathcal{X}$ be a complete separable metric space, $0 < \alpha < \beta$, $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be a potential with* $\mathrm{supp}(f) \neq \varnothing$, *and suppose that there is a continuous function $f_\beta : \mathcal{X} \to \mathbb{R}_{\geq 0}$ such that, for every $x \in \mathcal{X}$,*

$$f_\beta(x) = \inf_{y\in\mathcal{X}} (f(y) + \beta d(x, y)).$$

Then, for any $x_* \in \mathcal{X}$:

(1) *If x_* is an α-critical point of f_β, then $f(x_*) = f_\beta(x_*)$ and x_* is an α-critical point of f.*

(2) *If f_β attains its minimum at x_*, then f attains its minimum at x_*.*

Proof In view of Lemma 7.2, the second item is a consequence of the first: indeed, if f_β attains its minimum at $x_* \in \mathcal{X}$, then for every $\alpha < \beta$ we have that x_* is α-critical for f_β, so that by the first item it is α-critical for f. Since α is arbitrary, f attains its minimum at x_*.

Thus we focus on the first item. Suppose that x_* is an α-critical point of f_β; we begin by showing that $f(x_*) = f_\beta(x_*)$. Clearly $f_\beta(x_*) \leq f(x_*)$, so suppose for a contradiction that $f_\beta(x_*) < f(x_*)$. Let $\varepsilon = 1$ if $f(x_*) = \infty$ and $\varepsilon = \frac{f(x_*)-f_\beta(x_*)}{2} > 0$ otherwise, and let $\delta \in \mathbb{Q}_{>0}$ be such that $f(y) > f_\beta(x_*) + \varepsilon$ whenever $d(x_*, y) < \delta$. Such a δ can be obtained by letting Φ be the code for f; letting $B_r(a) \subseteq \mathcal{X}$ and $q \in \mathbb{Q}_{>0}$ be such that $B_r(a) \overset{\Phi}{\rightharpoonup} q$, $x_* \in B_r(a)$, and $q > f_\beta(x_*) + \varepsilon$; and then taking $\delta < r - d(a, x_*)$. Now let $y \in \mathcal{X}$ be such that

$$f_\beta(x_*) + \min\{(\beta - \alpha)\delta, \varepsilon\} > f(y) + \beta d(x_*, y).$$

Note that $y \neq x_*$, for otherwise $f_\beta(x_*) > f(x_*) - \min\{(\beta - \alpha)\delta, \varepsilon\} \geq f_\beta(x_*) + \varepsilon$, a contradiction. Moreover, we cannot have that $d(x_*, y) < \delta$, for otherwise

$$
\begin{aligned}
f_\beta(x_*) &> f(y) + \beta d(x_*, y) - \min\{(\beta - \alpha)\delta, \varepsilon\} \\
&> (f_\beta(x_*) + \varepsilon) + \beta d(x_*, y) - \min\{(\beta - \alpha)\delta, \varepsilon\} \\
&\geq f_\beta(x_*) + \beta d(x_*, y),
\end{aligned}
$$

where the first inequality is by the choice of y and the second inequality is by the choice of δ and the assumption $d(x_*, y) < \delta$. Thus it must be that $d(x_*, y) \geq \delta$. Therefore,

$$
\begin{aligned}
\alpha d(x_*, y) &< f_\beta(x_*) - f(y) - (\beta - \alpha)d(x_*, y) + \min\{(\beta - \alpha)\delta, \varepsilon\} \\
&\leq f_\beta(x_*) - f_\beta(y) - (\beta - \alpha)\delta + (\beta - \alpha)\delta = f_\beta(x_*) - f_\beta(y),
\end{aligned}
$$

where the first inequality is by the choice of y and the second inequality is because $f_\beta(y) \leq f(y)$, $\delta \leq d(x_*, y)$, and $\min\{(\beta - \alpha)\delta, \varepsilon\} \leq (\beta - \alpha)\delta$. This means that x_* is not an α-critical point of f_β, which is a contradiction.

We have established that $f(x_*) = f_\beta(x_*)$. We now use this to show that x_* is an α-critical point of f. Assume for a contradiction that this is not the case. Then there is a $y \in \mathcal{X}$ such that $\alpha d(x_*, y) \leq f(x_*) - f(y)$ but $y \neq x_*$. Then,

$$\alpha d(x_*, y) \leq f(x_*) - f(y) \leq f_\beta(x_*) - f_\beta(y)$$

because $f_\beta(x_*) = f(x_*)$ and $f_\beta(y) \leq f(y)$. This contradicts that x_* is an α-critical point of f_β. Thus x_* is indeed an α-critical point of f. $\square$

Theorem 8.2 (i) (ACA$_0$) *The* FVP *holds for arbitrary $\mathcal{X}$ and any honestly-coded potential f.*

(ii) (ACA$_0$) *The* FVP *holds for compact* $\mathcal{X}$ *and any potential* f; *in fact, such an* f *attains its minimum.*

(iii) (Π_1^1-CA$_0$) *The* FVP *holds for arbitrary* $\mathcal{X}$ *and any potential* f.

Proof In view of Lemma 7.2, we may assume that $\varepsilon = 1$; note that an honest code remains honest after scaling f.

(i) As f is assumed to be honestly-coded, the envelope f_2 exists and is continuous by Lemma 6.3. The function f_2 has a critical point by Theorem 7.5, and this critical point is also a critical point of f by Lemma 8.1.

(ii) As $\mathcal{X}$ is compact, f can be honestly coded by Lemma 5.2 item (ii). Thus f has a critical point by item (i) of this theorem. In fact, f_2 is defined and attains its minimum by Proposition 7.4, so that f also attains its minimum.

(iii) Working in Π_1^1-CA$_0$, we can assume that f is honestly-coded by Lemma 5.2 item (iv). Thus f has a critical point by item (i). □

Below we will show that the points in Theorem 8.2 are optimal.

9 Reversals of the variational principle

In this section we show that many of the results of Sect. 7 and Sect. 8 reverse. Let us begin with the weakest version of the FVP we have considered. Recall from Proposition 7.4 that WKL$_0$ suffices to prove that every continuous potential over a compact space has a critical point. We now show that WKL$_0$ is also necessary to prove this version of the FVP.

Proposition 9.1 *The* FVP *for continuous, bounded* f *on the Cantor space or on* $[0, 1]$ *implies* WKL$_0$ *over* RCA$_0$.

Proof We work in RCA$_0$ and prove the contrapositive. If WKL$_0$ fails, then there is an infinite binary tree $T \subseteq 2^{<\mathbb{N}}$ that has no infinite path. Let T° be the set of leaves of T: $T^\circ = \{\sigma \in T : \sigma^\frown 0, \sigma^\frown 1 \notin T\}$. Since T is infinite and has no infinite path, T° is also infinite. For each $\sigma \in T^\circ$, define

$$A_\sigma = \{i < |\sigma| - 1 : \neg(\exists \tau \in T)(|\tau| = |\sigma| + 1) \text{ and } \tau \sqsupseteq (\sigma{\restriction}i)^\frown(1 - \sigma(i+1))\}.$$

For each $\sigma \in T$, define $\tilde{\sigma} \in 2^{<\mathbb{N}}$ with $|\tilde{\sigma}| = 2|\sigma|$ as $\tilde{\sigma}(2i) = 0$ and $\tilde{\sigma}(2i+1) = \sigma(i)$ if $i < |\sigma|$. Put $\tilde{T} = \{\tilde{\sigma} : \sigma \in T\}$, $\tilde{T}^\circ = \{\tilde{\sigma} : \sigma \in T^\circ\}$, and put

$$S = \{\tau \in 2^{<\mathbb{N}} : (\forall \sigma \in T)(\tau \not\sqsubseteq \tilde{\sigma}) \text{ and } (\exists \sigma \in T)(\tau{\restriction}(|\tau| - 1) \sqsubset \tilde{\sigma})\}.$$

Here, S is the set of all binary strings τ which move away from $\tilde{T}$ before reaching a member of $\tilde{T}^\circ$. S can be defined in RCA$_0$ because each τ need only be checked against strings $\tilde{\sigma}$ of length at most $|\tau|$. The elements of $\tilde{T}^\circ \cup S$ are pairwise incomparable, and for every $x \in 2^{\mathbb{N}}$ there is a $\sigma \in \tilde{T}^\circ \cup S$ such that $x \sqsupseteq \sigma$. Indeed, each x must either reach a leaf of $\tilde{T}$ or move away from $\tilde{T}$ before that because $\tilde{T}$ has no infinite path.

Now, define a continuous function $f : 2^{\mathbb{N}} \to [0, 3]$ as follows:

$$f(x) = \begin{cases} 2 - \sum_{i \in A_\sigma} 2^{-2i} & \text{if } x \sqsupseteq \tilde{\sigma} \text{ for some } \sigma \in T^\circ, \\ 3 & \text{if } x \sqsupseteq \tau \text{ for some } \tau \in S. \end{cases}$$

One may easily obtain a code for f by Lemmas 3.8 and 3.9. We show that f has no critical point, which gives the desired contradiction. If $x \sqsupseteq \tau$ for some $\tau \in S$, then take any $\sigma \in T^\circ$ and $y \sqsupseteq \tilde{\sigma}$, and observe that $f(x) - f(y) \geq 3 - 2 \geq d(x, y)$. Thus x is not a critical point. Assume instead that $x \sqsupseteq \tilde{\sigma}$ for some $\sigma \in T^\circ$. Let i_0 be the greatest $i < |\sigma| - 1$ such that $i \notin A_\sigma$, which exists because T is infinite. Then there is a $\sigma' \in T^\circ$ such that $\sigma' \sqsupseteq (\sigma \restriction i_0) ^\frown (1 - \sigma(i_0 + 1))$ and $|\sigma'| > |\sigma|$. By the maximality of i_0, any $\tau \in T$ which extends $\sigma \restriction (i_0 + 1)$ is shorter than σ'. Thus, we have that $i_0 \in A_{\sigma'}$ and that $j \in A_\sigma$ implies that $j \in A_{\sigma'}$ for every $j < i_0$ since $\sigma \restriction i_0 = \sigma' \restriction i_0$. Take $y \sqsupseteq \tilde{\sigma}'$. Then $d(x, y) \leq 2^{-2i_0 - 1}$, and therefore

$$f(x) - f(y) = - \sum_{i \in A_\sigma} 2^{-2i} + \sum_{i \in A_{\sigma'}} 2^{-2i} \geq 2^{-2i_0} - \sum_{i \in A_\sigma, i > i_0} 2^{-2i}$$

$$\geq 2^{-2i_0 - 1} \geq d(x, y).$$

Thus x is not a critical point.

We can simulate the above construction on the unit interval in order to obtain a continuous function on $[0, 1]$ that has no critical point. For a given $\sigma \in 2^{<\mathbb{N}}$, let $l_\sigma = 0.\sigma = \sum_{i < |\sigma|} 2^{-\sigma(i)}, r_\sigma = l_\sigma + 2^{-|\sigma|}$, and $I_\sigma = [l_\sigma, r_\sigma] \subseteq [0, 1]$. Now, for each $\sigma \in T^\circ$, let $g_{\tilde{\sigma}} : I_{\tilde{\sigma}} \to [0, 3]$ be a piecewise linear function such that $g_{\tilde{\sigma}}(l_{\tilde{\sigma}}) = g_{\tilde{\sigma}}(r_{\tilde{\sigma}}) = 3$ and $g_{\tilde{\sigma}}((l_{\tilde{\sigma}} + r_{\tilde{\sigma}})/2) = 2 - \sum_{i \in A_\sigma} 2^{-2i}$, and, for each $\tau \in S$, define $g_\tau : I_\tau \to [0, 3]$ as $g_\tau(x) = 3$. Put $g = \bigcup_{\tau \in \tilde{T}^\circ \cup S} g_\tau$. Then, g is a continuous function on $[0, 1]$ by Lemmas 3.8 and 3.9: for this, observe that for each $\tau \in \tilde{T}^\circ \cup S$, g_τ can be extended to an open subset of $[0, 1]$, and thus g can be decomposed into piecewise linear functions on open subsets of $[0, 1]$. Indeed, if $l_\tau, r_\tau \notin \{0, 1\}$, then one can effectively find $\sigma_0, \sigma_1 \in \tilde{T}^\circ \cup S$ such that $r_{\sigma_0} = l_\tau$ and $r_\tau = l_{\sigma_1}$. Then g is still piecewise linear on an interval $(l_{\sigma_0}, r_{\sigma_1})$. One can check that g has no critical point, as we have seen above. If $x \in I_\tau$ for some $\tau \in S$, then take any $\sigma \in T^\circ$, and observe that $y = (l_{\tilde{\sigma}} + r_{\tilde{\sigma}})/2$ witnesses that x is not a critical point. If $x \in I_{\tilde{\sigma}}$ for some $\sigma \in T^\circ$ then take $\sigma' \in T^\circ$ as in the Cantor space case. Then $y = (l_{\tilde{\sigma}'} + r_{\tilde{\sigma}'})/2$ witnesses that x is not a critical point.   □

Next we consider versions of the variational principle provable in ACA_0. Recall from Theorems 7.5 and 8.2 that ACA_0 is able to prove the FVP when either f is continuous or $\mathcal{X}$ is compact. Let us see that each of these cases implies ACA_0, beginning with the former.

Proposition 9.2 *The* FVP *for continuous, bounded f on the Baire space implies* ACA_0 *over* RCA_0.

Proof As is well-known, ACA_0 is equivalent to the statement "for every injection $h : \mathbb{N} \to \mathbb{N}$, the range of h exists as a set" (see [14, Lemma III.1.3]).

Let $h\colon \mathbb{N} \to \mathbb{N}$ be an injection. For the purposes of this proof, we view the natural numbers as coding the finite sets, with 0 coding $\varnothing$. For each $n \in \mathbb{N}$ and finite set D, let $v_n(D)$ denote the number of $a \in D$ with $h(a) < n$: $v_n(D) = |\{a \in D : h(a) < n\}|$. The fact that h is an injection implies that $v_n(D) \leq n$.

Define a continuous function $g\colon \mathbb{N}^{\mathbb{N}} \to \mathbb{R}_{\geq 0}$ by

$$g(x) = \sum_{n \in \mathbb{N}} 2^{-2n-1+v_n(x(2^{n+1}))} \leq 1.$$

We may use Lemma 3.8 to show that g can indeed be coded as a continuous function in RCA$_0$, as it is straightforward to produce the sequence $(\langle \sigma {^\frown} 0^{\mathbb{N}}, g(\sigma {^\frown} 0^{\mathbb{N}})\rangle : \sigma \in \mathbb{N}^{<\mathbb{N}})$ and, for all $n \in \mathbb{N}$ and $\sigma, \tau \in \mathbb{N}^{<\mathbb{N}}$, to check that

$$d(\sigma {^\frown} 0^{\mathbb{N}}, \tau {^\frown} 0^{\mathbb{N}}) < 2^{-2^{n+1}} \Rightarrow |g(\sigma {^\frown} 0^{\mathbb{N}}) - g(\tau {^\frown} 0^{\mathbb{N}})| < 2^{-n}.$$

To see the above implication, observe that if $d(x, y) < 2^{-2^{n+1}}$, then $x{\restriction}(2^{n+1} + 1) = y{\restriction}(2^{n+1} + 1)$. This means that the terms $2^{-2k-1+v_k(x(2^{k+1}))}$ and $2^{-2k-1+v_k(y(2^{k+1}))}$ agree for $k \leq n$ and therefore that $|g(x) - g(y)| < \sum_{k=n+1}^{\infty} 2^{-k-1} = 2^{-n-1} < 2^{-n}$.

We see that $g(x) \leq 1$ for all $x \in \mathbb{N}^{\mathbb{N}}$, thus the function $f\colon \mathbb{N}^{\mathbb{N}} \to \mathbb{R}_{\geq 0}$ given by $f(x) = 1 - g(x)$ is a continuous potential on the Baire space that is bounded by 1.

By the FVP on the Baire space, let x_* be a critical point of f. For each n, let D_n denote the set coded by $x_*(2^{n+1})$. We claim that D_n must contain every a for which $h(a) < n$. Suppose not, and let a be such that $h(a) < n$ but $a \notin D_n$. Let $y \in \mathbb{N}^{\mathbb{N}}$ be such that $y(2^{n+1})$ is a code for $D_n \cup \{a\}$ and $y(m) = x_*(m)$ for all $m \neq 2^{n+1}$. Then

$$d(x_*, y) = 2^{-2^{n+1}} \leq 2^{-2n-1} \leq 2^{-2n-1+v_n(D_n)} = f(x_*) - f(y),$$

contradicting that x_* is a critical point. Thus D_n contains every a for which $h(a) < n$. We may then extract the range of h from x_* by taking $\mathrm{ran}(h) = \{n : (\exists a \in D_{n+1})(h(a) = n)\}$. $\qquad\square$

Proposition 9.3 *The FVP for honestly-coded f on $[0, 1]$ implies* ACA$_0$ *over* RCA$_0$.

Proof We work in RCA$_0$ and prove the contrapositive. By Theorem 3.5, if ACA$_0$ fails, then there is a strictly increasing sequence $\vec{c} = (c_n : n \in \mathbb{N})$ of rationals in $[0, 1]$ with no supremum.

To define f, enumerate a code Φ by enumerating $(u, v) \overset{\Phi}{\to} q$ if $q \leq u$ or if $q \leq 2$ and there is an n such that $v < c_n$. One readily checks that Φ is a code for the potential $f\colon [0, 1] \to \mathbb{R}_{\geq 0}$ given by

$$f(x) = \begin{cases} 2 & \text{if } \exists n(x < c_n) \\ x & \text{otherwise.} \end{cases}$$

To ensure that Φ is honest, additionally enumerate $(u, v) \overset{\Phi}{\to} q$ whenever there is an n such that $q \leq c_n < v$. To see that the resulting code is honest, consider an open interval

(u, v), and suppose that $(\forall x \in (u, v))(f(x) \geq q)$. Note that f is bounded above by 2, so $q \leq 2$. If there is an n such that $v < c_n$, then $(u, v) \xrightarrow{\Phi} q$. If $c_n < u$ for all n, then $f(x) = x$ on (u, v). Thus $q \leq u$, so $(u, v) \xrightarrow{\Phi} q$. Finally, suppose that there is an n such that $c_n \in (u, v)$, but there is no n such that $v < c_n$. As v is not the supremum of $(c_n : n \in \mathbb{N})$ (because we assumed that there is no such supremum), there is a $y \in (u, v)$ such that $\forall n(c_n < y)$. Then $f(y) = y$, which implies that $q \leq y < v$. By a similar argument, it cannot be that $\forall n(c_n < q)$ because if this were true, then there would be a $z \in (u, v)$ with $z < q$ such that $\forall n(c_n < z)$. As also $f(z) = z$, this contradicts the assumption that $(\forall x \in (u, v))(f(x) \geq q)$. Thus it must be that $q \leq c_n < v$ for some n, which implies that $(u, v) \xrightarrow{\Phi} q$. Hence Φ is honest.

We claim that f has no critical point. Assume towards a contradiction that c_* is a critical point of f. We show that c_* is the supremum of $\vec{c}$. Indeed, it is readily checked that if $c_* < c_n$ for some n, then

$$f(1) = 1 \leq 2 - d(c_*, 1) = f(c_*) - d(c_*, 1),$$

and hence c_* cannot be a critical point. It follows that $c_* \geq c_n$ for all n. Now, if $c' < c_*$ were also an upper bound of $\vec{c}$, then we would have $f(c') = c'$, so that

$$f(c') = c' = c_* - (c_* - c') = f(c_*) - d(c', c_*),$$

and thus c_* cannot be a critical point. Hence c_* is the supremum of $\vec{c}$, contradicting our initial assumption. ☐

Finally, we show that the unrestricted FVP proves Π^1_1-CA$_0$ by appealing to the following characterization of Π^1_1-CA$_0$.

Lemma 9.4 [14, Lemma VI.1.1] *The following are equivalent over* RCA$_0$:

(1) Π^1_1-*CA*$_0$
(2) *for any sequence* $(T_i)_{i \in \mathbb{N}}$ *of subtrees of* $\mathbb{N}^{<\mathbb{N}}$, *there is a set* X *such that for all* $i \in \mathbb{N}$, $i \in X$ *if and only if* T_i *has an infinite path.*

Proposition 9.5 *The* FVP *on the Baire space implies* Π^1_1-*CA*$_0$ *over* RCA$_0$.

Proof By Proposition 9.2, the FVP on the Baire space implies ACA$_0$ over RCA$_0$, so we may work over ACA$_0$.

Recall that we assume that the pairing function $\langle \cdot, \cdot \rangle \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is increasing in both coordinates. For the purposes of this proof, if $x \in \mathbb{N}^{\mathbb{N}}$ and $i \in \mathbb{N}$, then $(x)_i \in \mathbb{N}^{\mathbb{N}}$ is the function defined by $(x)_i(n) = x(\langle i, n \rangle)$. Similarly, if $\sigma \in \mathbb{N}^{<\mathbb{N}}$ and $i \in \mathbb{N}$, then $(\sigma)_i \in \mathbb{N}^{<\mathbb{N}}$ is the longest sequence such that $(\forall n < |(\sigma)_i|)[\langle i, n \rangle \in \text{dom } \sigma \wedge (\sigma)_i(n) = \sigma(\langle i, n \rangle)]$.

Let $(T_i)_{i \in \mathbb{N}}$ be a sequence of subtrees of $\mathbb{N}^{<\mathbb{N}}$. We first define a code for the lower semi-continuous potential $f \colon \mathbb{N}^{\mathbb{N}} \to [0, 1]$ given by

$$f(x) = \sum_{i=0}^{\infty} \{2^{-i} : (x)_i \notin [T_i]\}.$$

To do this, define $B_r(\sigma) \overset{\Phi}{\dashrightarrow} q$ if there is a $\tau \in \mathbb{N}^{<\mathbb{N}}$ such that $B_r(\sigma) \Subset B_{2^{-|\tau|}}(\tau)$ and $q \leq \sum_{i=0}^{\infty}\{2^{-i} : (\tau)_i \notin T_i\}$. For a given $x \in \mathbb{N}^{\mathbb{N}}$, one readily checks that $v = \sum_{i=0}^{\infty}\{2^{-i} : (x)_i \notin [T_i]\}$ (which ACA_0 proves exists) is indeed the supremum of

$$\{q \in \mathbb{Q} : (\exists\langle \sigma, r\rangle \in \mathbb{N}^{<\mathbb{N}} \times \mathbb{Q}_{>0})(B_r(\sigma) \overset{\Phi}{\dashrightarrow} q \wedge d(x, \sigma) < r)\}.$$

This shows that Φ correctly codes the desired potential f and that f is provably total in ACA_0.

By the FVP, let x_* be a critical point of f. Let $X = \{i : (x_*)_i \in [T_i]\}$. We show that $(\forall i \geq 0)(i \in X \leftrightarrow T_i \text{ has a path})$. Clearly, if $i \in X$, then T_i has a path. Suppose for a contradiction that there is a j such that T_j has a path, but $j \notin X$. Let $h \in [T_j]$, and let $y \in \mathbb{N}^{\mathbb{N}}$ be such that for all $i, n \in \mathbb{N}$,

$$y(\langle i, n\rangle) = \begin{cases} h(n) & \text{if } i = j \\ x_*(\langle i, n\rangle) & \text{otherwise,} \end{cases}$$

so that $(y)_j = h$ and $\forall i[i \neq j \rightarrow (y)_i = (x_*)_i]$. Then

$$d(x_*, y) \leq 2^{-\langle j, 0\rangle} \leq 2^{-j} = f(x_*) - f(y).$$

This contradicts that x_* is a critical point of f, which completes the proof. $\qquad\square$

10 The localized variational principle

In this section we compare the strength of the free and localized variational principles in different contexts. First, we show that, in contrast to the FVP, the LVP is not affected by the boundedness of f. We also show that the two principles are equivalent when f is not assumed to be continuous. For the latter, it is clear that the LVP implies the FVP, so we focus on the other direction. The following two lemmas aid the proof of the above facts.

Lemma 10.1 (RCA_0) *Let $\mathcal{X}$ be a complete separable metric space, let $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be lower semi-continuous, and let $\mathcal{C} \subseteq \mathcal{X}$ be closed. Then there is a lower semi-continuous function $g : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ such that*

$$g(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C} \\ f(x) & \text{if } x \notin \mathcal{C}. \end{cases}$$

Proof Let Ψ be a code for f, and let U be a code for the complement of $\mathcal{C}$. Enumerate a code Φ for g by enumerating $B_r(a) \overset{\Phi}{\dashrightarrow} q$ if either

- $q \leq 0$ or
- $B_r(a) \overset{\Psi}{\dashrightarrow} q$ and there is an open ball $B_s(b)$ enumerated in U with $B_r(a) \Subset B_s(b)$.

If $x \in \mathcal{C}$ then there are balls $B_r(a)$ with $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\rightharpoonup} 0$, but there are no balls $B_r(a)$ and $q > 0$ such that $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\rightharpoonup} q$. Hence $g(x) = 0$. If $x \notin \mathcal{C}$, given $q \in \mathbb{Q}_{>0}$, it is clear from the definition of Φ that if there is a ball $B_r(a)$ with $x \in B_r(a)$ and $B_r(a) \overset{\Phi}{\rightharpoonup} q$ then also $B_r(a) \overset{\Psi}{\rightharpoonup} q$. Conversely, suppose that there is a ball $B_r(a)$ with $x \in B_r(a)$ and $B_r(a) \overset{\Psi}{\rightharpoonup} q$. Since $x \notin \mathcal{C}$, $x \in B_s(b)$ for some ball $B_s(b)$ enumerated in U, and hence there are a', r' such that $x \in B_{r'}(a') \Subset B_r(a)$, $B_s(b)$. But then by condition (LSC1) of Definition 4.1, $B_{r'}(a') \overset{\Psi}{\rightharpoonup} q$, so that also $B_{r'}(a') \overset{\Phi}{\rightharpoonup} q$. Since q was arbitrary we conclude that $g(x) = f(x)$ given that both are the supremum of such q in $\overline{\mathbb{R}}$. □

Lemma 10.2 (RCA$_0$) *Let $\mathcal{X}$ be a complete separable metric space, and let $f, g: \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be lower semi-continuous. Then $(f + g)$, $\max\{f, g\}$, $\min\{f, g\}: \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ (defined in the usual way) are lower semi-continuous.*

Proof Let Ψ_0 be a code for f and let Ψ_1 be a code for g. Enumerate a code for $f + g$ by enumerating $B_r(a) \overset{\Phi}{\rightharpoonup} q$ if there are q_0, q_1 such that $B_r(a) \overset{\Psi_0}{\rightharpoonup} q_0$, $B_r(a) \overset{\Psi_1}{\rightharpoonup} q_1$, and $q \leq q_0 + q_1$. Enumerate a code Φ for $\max\{f, g\}$ by enumerating $B_r(a) \overset{\Phi}{\rightharpoonup} q$ if either $B_r(a) \overset{\Psi_0}{\rightharpoonup} q$ or $B_r(a) \overset{\Psi_1}{\rightharpoonup} q$. Enumerate a code Φ for $\min\{f, g\}$ by enumerating $B_r(a) \overset{\Phi}{\rightharpoonup} q$ if both $B_r(a) \overset{\Psi_0}{\rightharpoonup} q$ and $B_r(a) \overset{\Psi_1}{\rightharpoonup} q$. □

Lemma 10.3 (RCA$_0$) *For every complete separable metric space $\mathcal{X}$, the LVP for bounded potentials on $\mathcal{X}$ implies the LVP on $\mathcal{X}$, and the LVP for bounded continuous potentials on $\mathcal{X}$ implies the LVP for continuous potentials on $\mathcal{X}$.*

Proof Given a complete separable metric space $\mathcal{X}$, a potential $f: \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$, and an $x_0 \in \mathrm{supp}(f)$, define $\tilde{f}: \mathcal{X} \to \mathbb{R}_{\geq 0}$ by setting $\tilde{f}(x) = \min\{f(x), f(x_0)\}$. Then $\tilde{f}$ is bounded, is a potential by Lemma 10.2, and is continuous if f is continuous. Suppose that $\varepsilon > 0$ and that $x_* \in \mathcal{X}$ is an ε-critical point of $\tilde{f}$ with $\varepsilon d(x_0, x_*) \leq \tilde{f}(x_0) - \tilde{f}(x_*)$. Then either $\tilde{f}(x_*) = f(x_*)$ or $\tilde{f}(x_*) = f(x_0)$. However, if $\tilde{f}(x_*) = f(x_0)$, then $\varepsilon d(x_0, x_*) \leq \tilde{f}(x_0) - \tilde{f}(x_*) = f(x_0) - f(x_0) = 0$. Thus $d(x_0, x_*) = 0$ and $x_* = x_0$, in which case again $\tilde{f}(x_*) = f(x_*)$. Thus in either case $\tilde{f}(x_*) = f(x_*)$. Therefore $\varepsilon d(x_0, x_*) \leq \tilde{f}(x_0) - \tilde{f}(x_*) = f(x_0) - f(x_*)$. Furthermore, if $x \in \mathcal{X}$ and $\varepsilon d(x_*, x) \leq f(x_*) - f(x)$, then also $\varepsilon d(x_*, x) \leq \tilde{f}(x_*) - \tilde{f}(x)$, so $x = x_*$ because x_* is an ε-critical point of $\tilde{f}$. Thus x_* is also an ε-critical point of f with $\varepsilon d(x_0, x_*) \leq f(x_0) - f(x_*)$. □

Lemma 10.4 (RCA$_0$) *For every complete separable metric space $\mathcal{X}$, the FVP on $\mathcal{X}$ implies the LVP on $\mathcal{X}$.*

Proof Assume the FVP. Let $f: \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be any potential, $\varepsilon > 0$, $x_0 \in \mathrm{supp}(f)$, and let $\mathcal{C}$ be the closed set

$$\mathcal{C} = \{x \in \mathcal{X} : \varepsilon d(x, x_0) \leq f(x_0) - f(x)\}.$$

To see that $\mathcal{C}$ is closed, rewrite $\mathcal{C}$ as $\mathcal{C} = \{x \in \mathcal{X} : f(x) + \varepsilon d(x, x_0) \leq f(x_0)\}$, and observe that the function $g(x) = f(x) + \varepsilon d(x, x_0)$ is lower semi-continuous by

Proposition 4.3 and Lemma 10.2. As in the proof of Proposition 4.2, $g(x) \leq f(x_0)$ is a Π_1^0-definable property of x, so $\mathcal{C}$ is closed.

By Lemmas 10.1 and 10.2, the function $\tilde{f} \colon \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ given by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \mathcal{C} \\ \max\{f(x), \varepsilon d(x, x_0) + f(x_0)\} & \text{if } x \notin \mathcal{C} \end{cases}$$

is lower semi-continuous. Let x_* be an ε-critical point of $\tilde{f}$. It must be that $x_* \in \mathcal{C}$, for otherwise we would have that $x_* \neq x_0$ (as $x_0 \in \mathcal{C}$), but $\tilde{f}(x_*) \geq \varepsilon d(x_*, x_0) + f(x_0) = \varepsilon d(x_*, x_0) + \tilde{f}(x_0)$, meaning that x_0 witnesses that x_* is not an ε-critical point of $\tilde{f}$.

We have that $f(x_*) \leq f(x_0) - \varepsilon d(x_*, x_0)$ because $x_* \in \mathcal{C}$. Moreover, if $x \in \mathcal{X}$ and $f(x) \leq f(x_*) - \varepsilon d(x_*, x)$, then

$$f(x) \leq f(x_*) - \varepsilon d(x_*, x) \leq f(x_0) - \varepsilon d(x_*, x_0) - \varepsilon d(x_*, x) \leq f(x_0) - \varepsilon d(x, x_0)$$

(where the last inequality is by the triangle inequality), so $x \in \mathcal{C}$ and therefore $\tilde{f}(x) = f(x)$. Thus for an $x \in \mathcal{X}$ with $f(x) \leq f(x_*) - \varepsilon d(x_*, x)$,

$$\tilde{f}(x) = f(x) \leq f(x_*) - \varepsilon d(x_*, x) = \tilde{f}(x_*) - \varepsilon d(x_*, x),$$

and therefore $x = x_*$ because x_* is an ε-critical point of $\tilde{f}$. So x_* is also an ε-critical point of f. We conclude that x_* is an ε-critical point of f and that $f(x_*) \leq f(x_0) - \varepsilon d(x_*, x_0)$. $\qquad\square$

Thus the FVP and the LVP for arbitrary f are equivalent. In contrast, the LVP for continuous functions suffices to prove the full FVP. To make this precise, we introduce the notion of *pseudo-fibrations*. Below, by a *closed isometry* we mean a continuous function $f \colon \mathcal{X} \to \mathcal{Y}$ such that $d_{\mathcal{X}}(x_0, x_1) = d_{\mathcal{Y}}(f(x_0), f(x_1))$ for all $x_0, x_1 \in \mathcal{X}$ and such that $f[\mathcal{X}]$ is closed. We check that if f is a closed isometry and $\mathcal{C} \subseteq \mathcal{X}$ is closed, then RCA_0 proves that $f[\mathcal{C}]$ is closed as well.

Lemma 10.5 (RCA_0) *Let $\mathcal{X}$ be a complete separable metric space, let $\mathcal{C} \subseteq \mathcal{X}$ be closed, and let $f \colon \mathcal{X} \to \mathcal{Y}$ be a closed isometry. Then $f[\mathcal{C}] \subseteq \mathcal{Y}$ is closed.*

Proof Let U be a code for the complement of $\mathcal{C} \subseteq \mathcal{X}$, and let V_0 be a code for the complement of $f[\mathcal{X}] \subseteq \mathcal{Y}$. Enumerate a code V for an open $\mathcal{V} \subseteq \mathcal{Y}$ by enumerating the ball $B_s(b)$ in V if either

- $B_s(b)$ is enumerated in V_0 or
- there is a ball $B_r(a)$ enumerated in U such that $d_{\mathcal{Y}}(f(a), b) + s < r$.

We show that a $y \in \mathcal{Y}$ is in $\mathcal{V}$ if and only if there is no $x \in \mathcal{C}$ such that $f(x) = y$. This shows that $\mathcal{V}$ is the complement of $f[\mathcal{C}]$ and hence that $f[\mathcal{C}]$ is closed.

For $y \in \mathcal{Y}$, we know that there is no $x \in \mathcal{X}$ such that $y = f(x)$ if and only if there is a ball $B_s(b)$ containing y that is enumerated in V_0. Thus it suffices to assume that there is an $x \in \mathcal{X}$ such that $y = f(x)$ and show that $y \in \mathcal{V}$ if and only if $x \notin \mathcal{C}$. Note that x is unique because f is an isometry.

First suppose that $x \notin C$. Then $x \in B_r(a)$ for some ball $B_r(a)$ enumerated in U. We have that $d_{\mathcal{Y}}(f(a), y) = d_{\mathcal{X}}(a, x) < r$ because $y = f(x)$ and f is an isometry. Thus by choosing b sufficiently close to y, we see that there is a ball $B_s(b)$ containing y with $d_{\mathcal{Y}}(f(a), b) + s < r$. Thus $y \in \mathcal{V}$.

Conversely, suppose that $y \in \mathcal{V}$. Then there are balls $B_s(b) \subseteq \mathcal{Y}$ and $B_r(a) \subseteq \mathcal{X}$ such that $y \in B_s(b)$, $B_r(a)$ is enumerated in U, and $d_{\mathcal{Y}}(f(a), b) + s < r$. Therefore $d_{\mathcal{Y}}(f(a), f(x)) < r$ because $f(x) = y \in B_s(b)$. Thus $d_{\mathcal{X}}(a, x) < r$ because f is an isometry. Therefore $x \in B_r(a)$, so $x \notin C$. $\qquad\square$

Definition 10.6 Let $\mathcal{X}, \mathcal{Y}$ be complete separable metric spaces. Say that a space $\mathcal{Z}$ is an *$\mathcal{X}$-pseudofibration* of $\mathcal{Y}$ if there are a closed isometry $\iota : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and a continuous function $\pi : \mathcal{Z} \to \mathcal{Y}$ such that

(1) for all $z, z' \in \mathcal{Z}$, $d_{\mathcal{Y}}(\pi(z), \pi(z')) \leq d_{\mathcal{Z}}(z, z')$ and
(2) for all $\langle x, y \rangle \in \mathcal{X} \times \mathcal{Y}$, $\pi\iota(x, y) = y$.

Of course the typical example of an $\mathcal{X}$-pseudofibration of $\mathcal{Y}$ is $\mathcal{X} \times \mathcal{Y}$, but later we will see that there are others. Pseudofibrations are useful due to the following.

Lemma 10.7 (RCA$_0$)

(1) *If $\mathcal{X}$ and $\mathcal{Z}$ are complete separable metric spaces such that $\mathcal{Z}$ is an $\mathcal{X}$-pseudofibration of $\mathbb{R}_{\geq 0}$, then the LVP for continuous potentials on $\mathcal{Z}$ implies the FVP on $\mathcal{X}$.*
(2) *If $\mathcal{X}$ and $\mathcal{Z}$ are complete separable metric spaces such that $\mathcal{Z}$ is an $\mathcal{X}$-pseudofibration of $[0, 1]$, then the LVP for continuous potentials on $\mathcal{Z}$ implies the 1-FVP on $\mathcal{X}$.*

Proof We give a uniform proof for the two claims. Note that by Lemma 7.3.1 we can also assume for the conclusion of the first claim that $\varepsilon = 1$. Let $\mathcal{X}$ be a complete separable metric space, let $\mathcal{Y}$ be either $\mathbb{R}_{\geq 0}$ for the first claim or $[0, 1]$ for the second, and let $\mathcal{Z}$ be an $\mathcal{X}$-pseudofibration of $\mathcal{Y}$ with associated functions $\iota : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and $\pi : \mathcal{Z} \to \mathcal{Y}$. For the first claim, let $f : \mathcal{X} \to \overline{\mathbb{R}}_{\geq 0}$ be lower semi-continuous; and for the second claim, let $f : \mathcal{X} \to [0, 1]$ be lower semi-continuous. By Proposition 4.2, the set

$$\Delta = \{\langle x, y \rangle \in \mathcal{X} \times \mathcal{Y} : f(x) \leq y\}$$

is closed, hence so is $\Gamma = \iota[\Delta] \subseteq \mathcal{Z}$ by Lemma 10.5 because ι is a closed isometry. By Urysohn's lemma, which is provable in RCA$_0$ by Simpson [14, Lemma II.7.3], there is a continuous function $g : \mathcal{Z} \to [0, 1]$ such that $g(z) = 0$ if and only if $z \in \Gamma$. Define a continuous function $\tilde{f} : \mathcal{Z} \to \mathbb{R}_{\geq 0}$ by

$$\tilde{f}(z) = \pi(z) + g(z).$$

Every $z \in \Gamma$ is of the form $z = \iota(x, y)$ for some $\langle x, y \rangle \in \mathcal{X} \times \mathcal{Y}$. For such a z, we then have that

$$\tilde{f}(z) = \pi(z) + g(z) = \pi(z) = \pi\iota(x, y) = y,$$

where the second equality is because $z \in \Gamma$ (and therefore $g(z) = 0$), and the last equality is by Definition 10.6.2.

Now, fix any $x_0 \in \text{supp}(f)$, let $y_0 = f(x_0)$, and let $z_0 = \iota(x_0, y_0)$. By the LVP for continuous potentials on $\mathcal{Z}$, let $z_* \in \mathcal{Z}$ be a critical point of $\tilde{f}$ that satisfies $\tilde{f}(z_*) \leq \tilde{f}(z_0) - d_{\mathcal{Z}}(z_0, z_*)$. We first claim that $z_* \in \Gamma$. To see this, observe that $z_0 \in \Gamma$ by definition and therefore $\tilde{f}(z_0) = \pi(z_0) = y_0$. We then have that

$$
\begin{aligned}
\pi(z_*) + g(z_*) = \tilde{f}(z_*) &\leq \tilde{f}(z_0) - d_{\mathcal{Z}}(z_0, z_*) \\
&= \pi(z_0) - d_{\mathcal{Z}}(z_0, z_*) \\
&\leq \pi(z_0) - d_{\mathcal{Y}}(\pi(z_0), \pi(z_*)) \quad \text{by Definition 10.6.1} \\
&= \pi(z_0) - |\pi(z_0) - \pi(z_*)| \\
&\leq \pi(z_*).
\end{aligned}
$$

Therefore we must have that $g(z_*) = 0$, which means that $z_* \in \Gamma$.

As $z_* \in \Gamma$, it must be that $z_* = \iota(x_*, y_*)$ for some $\langle x_*, y_* \rangle \in \Delta$. Thus $x_* \in \mathcal{X}$ and $y_* \geq f(x_*)$. In fact, it must be that $y_* = f(x_*)$. To see this, suppose for a contradiction that $y_* > f(x_*)$, and let $z = \iota(x_*, f(x_*))$. We then have that

$$
\tilde{f}(z_*) - \tilde{f}(z) = y_* - f(x_*) = d_{\mathcal{Z}}(z, z_*),
$$

where the first equality is because z_* and z are in Γ, and the second equality is because ι is an isometry. Thus z witnesses that z_* is not a critical point of $\tilde{f}$, which is a contradiction. Therefore $y_* = f(x_*)$, so $z_* = \iota(x_*, f(x_*))$.

We now show that x_* is a critical point of the original f, which completes the proof. Let $x \in \mathcal{X}$, and suppose that $f(x_*) \geq f(x) + d_{\mathcal{X}}(x, x_*)$. Let $z = \iota(x, f(x))$. Then

$$
d_{\mathcal{Z}}(z_*, z) = d_{\mathcal{X} \times \mathcal{Y}}(\langle x_*, f(x_*)\rangle, \langle x, f(x)\rangle) = f(x_*) - f(x),
$$

where the first equality is because ι is an isometry, and the second equality is because $f(x_*) - f(x) \geq d_{\mathcal{X}}(x, x_*)$ and we use the max norm on $\mathcal{X} \times \mathcal{Y}$. As z_* and z are in Γ, we have that

$$
\tilde{f}(z_*) - \tilde{f}(z) = f(x_*) - f(x) = d_{\mathcal{Z}}(z_*, z).
$$

Therefore $z = z_*$ because z_* is a critical point of $\tilde{f}$, and since ι is injective, $x = x_*$. We have shown that if $x \in \mathcal{X}$ satisfies $f(x_*) \geq f(x) + d_{\mathcal{X}}(x, x_*)$, then $x = x_*$. Thus x_* is a critical point of f. $\qquad\square$

Of course it follows from Lemma 10.7 that the LVP for continuous functions on $\mathcal{X} \times \mathbb{R}_{\geq 0}$ implies the FVP on $\mathcal{X}$. However, the notion of pseudofibrations will allow us to replace $\mathcal{X} \times \mathbb{R}_{\geq 0}$ by a more familiar space. In what follows, we will mainly deal with $\mathcal{C}([0, 1])$. For given $h \in \mathcal{C}([0, 1])$ and a closed interval $I \subseteq [0, 1]$, we write $\|h\|_I = \sup_{t \in I} |h(t)|$. By the definition (coding) of $\mathcal{C}([0, 1])$, statements of the form $\|h\|_I \leq r$ are always Π_1^0.

Lemma 10.8 (RCA$_0$) *Let $\mathcal{X}$ be a complete separable metric space, and let $\mathcal{Y} = [a, b] \cap \mathbb{R}$, where $a < b$ are elements of $\overline{\mathbb{R}}$. If $\mathcal{X}$ embeds by a closed isometry into $\mathcal{C}([0, 1])$, then $\mathcal{C}([0, 1])$ is an $\mathcal{X}$-pseudofibration of $\mathcal{Y}$.*

Proof Let $\mathcal{Z} = \mathcal{C}([0, 1])$ and let $f \colon \mathcal{X} \to \mathcal{Z}$ be a closed isometry. Let us write f_x instead of $f(x)$. Define $\iota \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ by

$$\big(\iota(x, y)\big)(t) = \begin{cases} 2t f_x(0) + (1 - 2t)y & \text{if } t < 1/2, \\ f_x(2t - 1) & \text{if } t \geq 1/2, \end{cases}$$

and define $\pi \colon \mathcal{Z} \to \mathcal{Y}$ by

$$\pi(g) = \begin{cases} a & \text{if } g(0) < a, \\ b & \text{if } g(0) > b, \\ g(0) & \text{otherwise.} \end{cases}$$

Intuitively, $\iota(x, y)$ represents y by its value on 0 and represents $f(x)$ by its values on $[1/2, 1]$. It is easy to check that a code for π can be constructed using Lemma 3.8. We may also use Lemma 3.8 to construct a code for ι. To do this, let X and Y be the dense sets of $\mathcal{X}$ and $\mathcal{Y}$, respectively. It is straightforward to produce the sequence $(\langle\langle a, q\rangle, \iota(a, q)\rangle : \langle a, q\rangle \in X \times Y)$ and observe that

$$d_{\mathcal{X} \times \mathcal{Y}}(\langle a_0, q_0\rangle, \langle a_1, q_1\rangle) = \max(d_{\mathcal{X}}(a_0, a_1), d_{\mathcal{Y}}(q_0, q_1)) = d_{\mathcal{Z}}(\iota(a_0, q_0), \iota(a_1, q_1)). \tag{3}$$

Thus the hypotheses of Lemma 3.8 are satisfied, so there exists a code for ι. Equation (3) implies that ι is an isometry. To see that ι is closed, consider an $h \in \mathcal{C}([0, 1])$, put $h_\flat(t) = (2t)h(1/2) - (1 - 2t)h(0)$, and put $h_\sharp(t) = h(t/2 + 1/2)$. Then $h \in \iota[\mathcal{X} \times \mathcal{Y}]$ if and only if $\|h - h_\flat\|_{[0, 1/2]} = 0$ (i.e., h is linear on $[0, 1/2]$), $h(0) \in \mathcal{Y}$, and $h_\sharp \in f[\mathcal{X}]$, which is a Π_1^0 condition because $f[\mathcal{X}]$ is closed by assumption. Thus, $\iota[\mathcal{X} \times \mathcal{Y}]$ is closed by Lemma 3.7. It is then easy to check that ι and π make $\mathcal{Z}$ an $\mathcal{X}$-pseudofibration of $\mathcal{Y}$. □

In order to apply Lemma 10.8, we need to show that many of the spaces we are interested in isometrically embed into $\mathcal{C}([0, 1])$. Let us begin with the unit interval.

Lemma 10.9 (RCA$_0$) *There exists a closed isometry $f \colon [0, 1] \to \mathcal{C}([0, 1])$.*

Proof Let $\mathcal{X} = [0, 1]$ and $\mathcal{Y} = \mathcal{C}([0, 1])$. For $r \in [0, 1]$, write $\mathcal{I}_r$ for the constant function with value r. Define $f \colon \mathcal{X} \to \mathcal{Y}$ by $f(x) = \mathcal{I}_x$. It is obvious that f is an isometry and thus a code for f exists by arguing in the style of the proof of Lemma 10.8 and appealing to Lemma 3.8. We may observe that $h \in f[\mathcal{X}]$ if and only if $\|h - \mathcal{I}_{h(0)}\|_{[0, 1]} = 0$, which is a Π_1^0 condition, thus $f[\mathcal{X}]$ is closed by Lemma 3.7. □

Next, let us see that there is also a closed isometry from the Baire space into $\mathcal{C}([0, 1])$.

Definition 10.10 If $I = [a, b]$ is an interval with $0 \le a < b \le 1$, define $\hat{I} \colon [0, 1] \to \mathbb{R}$ to be the piecewise linear function with $\hat{I}(0) = \hat{I}(a) = \hat{I}(b) = \hat{I}(1) = 0$, $\hat{I}\left(\frac{a+b}{2}\right) = 1$, and $\hat{I}$ linear elsewhere. For $n \in \mathbb{N}$, define $I_n = [1 - 2^{-n}, 1 - 2^{-(n+1)}]$ and for $m, n \in \mathbb{N}$ define

$$J_n^m = \left[1 - 2^{-(n+1)} - 2^{-(n+m+1)}, \, 1 - 2^{-(n+1)} - 2^{-(n+m+2)} \right].$$

Let $\mathcal{X}$ be the Baire space and $\mathcal{Y}$ be $\mathcal{C}([0, 1])$. Then, define $F \colon \mathcal{X} \to \mathcal{Y}$ by

$$F(x) = \sum_{n \in \mathbb{N}} 2^{-n} \hat{J}_n^{x(n)}.$$

Lemma 10.11 (RCA_0) *The function F of Definition 10.10 is a closed isometry from the Baire space to $\mathcal{C}([0, 1])$.*

Proof A code for F exists by checking that F is an isometry, arguing as in the proof of Lemma 10.8, and appealing to Lemma 3.8. For $h \in \mathcal{C}([0, 1])$, $h \in F[\mathbb{N}^{\mathbb{N}}]$ if and only if for every $n \in \mathbb{N}$, $\|h\|_{I_n} = 2^{-n}$, and for every $n, m \in \mathbb{N}$, either $\|h\|_{J_n^m} = 0$ or $\|h - 2^{-n} \hat{J}_n^m\|_{I_n} = 0$, which is a Π_1^0 statement. It is easy to check the "only if" direction, so we check the "if" direction. For a given $h \in \mathcal{C}([0, 1])$ which satisfies the assumption, define $h_{\mathbb{N}} \colon \mathbb{N} \to \mathbb{N}$ by letting $h_{\mathbb{N}}(n)$ be the unique $m \in \mathbb{N}$ such that $\|h\|_{J_n^m} \ne 0$ (which can be found effectively). Then $F(h_{\mathbb{N}}) = h$. Thus the image of F is closed by Lemma 3.7. $\qquad\square$

Proposition 10.12 (RCA_0)

(1) *The LVP for bounded, continuous potentials on $[0, 1] \times [0, 1]$ with the max metric implies ACA_0.*
(2) *The LVP for bounded, continuous potentials on $\mathcal{C}([0, 1])$ implies Π_1^1-CA_0.*

Proof Note that in view of Lemma 10.3 the bounded LVP implies the full LVP, so we may remove 'bounded' from both items. For the first item, we note by Proposition 9.3 that the FVP on $[0, 1]$ implies ACA_0. Moreover, by Lemma 7.3.3d, since $[0, 1]$ is a closed ball in $\mathbb{R}^1$, we may replace the FVP by the 1-FVP. Since $[0, 1] \times [0, 1]$ is clearly a $[0, 1]$-pseudofibration of $[0, 1]$, by Lemma 10.7 we have that the LVP for bounded, continuous potentials on $[0, 1] \times [0, 1]$ implies ACA_0 as well.

For the second, let $\mathcal{X}$ be the Baire space and $\mathcal{Z}$ be $\mathcal{C}([0, 1])$. By Lemma 7.3.3b and Proposition 9.5, the 1-FVP on $\mathcal{X}$ implies Π_1^1-CA_0. By Lemma 10.11, $\mathcal{X}$ embeds by a closed isometry into $\mathcal{Z}$, so that by Lemma 10.8, $\mathcal{Z}$ is an $\mathcal{X}$-pseudofibration of $[0, 1]$. Hence, reasoning as above, the LVP for bounded, continuous functions on $\mathcal{Z}$ implies the 1-FVP on $\mathcal{X}$, from which we obtain Π_1^1-CA_0. $\qquad\square$

Remark 10.13 Let $\mathcal{B}$ be the closed unit ball of $\mathcal{C}([0, 1])$. It is easy to see that the isometries used in Lemmas 10.9 and 10.11 map into $\mathcal{B}$, and the proof of Lemma 10.8 can be modified to show that $\mathcal{B}$ is an $\mathcal{X}$-pseudofibration of $[0, 1]$. With this we may replace $\mathcal{C}([0, 1])$ by $\mathcal{B}$ in Proposition 10.12.2.

11 Conclusion

We have considered formalizations of Ekeland's variational principle with various natural restrictions and have shown that they are equivalent to well-known theories of second-order arithmetic. In this final section we synthesize these results to produce the definitive versions of our main results.

There are three main weakenings of the LVP we considered: the free, rather than localized, statement; the case where $\mathcal{X}$ is compact; and the case where f is continuous. When the LVP is weakened in the three ways simultaneously, we obtain a statement equivalent to WKL_0.

Theorem 11.1 *Over* RCA_0*, the following are equivalent:*

(a) WKL_0*;*
(b) *the* FVP *for continuous* f *and compact* $\mathcal{X}$*;*
(c) *the* FVP *for continuous* f *on the Cantor space or on* [0, 1]*.*

Proof That (a) implies the other items is Proposition 7.4, while (c) implies (a) by Proposition 9.1. □

If we choose two, but not three, of the weakenings, we instead obtain statements equivalent to ACA_0. Moreover, if we impose only one of the three weakenings, we typically obtain statements that are not provable in ACA_0. The only exception to this is the case where $\mathcal{X}$ is compact.

Theorem 11.2 *Over* RCA_0*, the following are equivalent:*

(a) ACA_0*;*
(b) *the* FVP *for continuous* f*;*
(c) *the* FVP *for compact* $\mathcal{X}$*;*
(d) *the* LVP *for compact* $\mathcal{X}$*;*
(e) *the* LVP *for continuous* f *and compact* $\mathcal{X}$*.*

Moreover in item (b) *we may take* $\mathcal{X}$ *to be the Baire space, in items* (c) *and* (d) *we may take* $\mathcal{X} = [0, 1]$*, and in* (d) *and* (e) *we may take* $\mathcal{X} = [0, 1] \times [0, 1]$ *with the max metric.*

Proof The equivalence between (a) and (b) is given by Theorem 7.5 and Proposition 9.2, while the equivalence between (a) and (c) is given by Theorem 8.2 and Proposition 9.3. By Lemma 10.4, (c) implies (d). Clearly, (d) implies (e), which by Proposition 10.12 implies (c). □

Other than compactness, if we impose at most one weakening to the LVP we obtain statements equivalent to $\Pi_1^1\text{-}\mathsf{CA}_0$.

Theorem 11.3 *The following are equivalent:*

(a) $\Pi_1^1\text{-}CA_0$*;*
(b) *the* FVP*;*
(c) *the* LVP *for continuous* f*;*

(d) *the* LVP.

Moreover, in (b) and (d) we may take $\mathcal{X}$ to be the Baire space, and in any of (b)–(d) we may take $\mathcal{X}$ to be (the closed unit ball of) $\mathcal{C}([0, 1])$.

Proof The equivalence between (a) and (b) is given by Theorem 8.2 and Proposition 9.5. That (b) implies (d) is Lemma 10.4. Clearly, (d) implies (c), and the latter implies (a) by Proposition 10.12 (with Remark 10.13).   □

We have also considered variants of the variational principle where f is bounded, but this restriction has not affected the proof-theoretic strength of the theorem; indeed, in each of these theorems we may additionally assume that f is bounded.

References

1. Bosch, C., García, A., García, C.L.: An extension of Ekeland's variational principle to locally complete spaces. J. Math. Anal. Appl. **328**(1), 106–108 (2007)
2. Brézis, H., Browder, F.E.: A general principle on ordered sets in nonlinear functional analysis. Adv. Math. **21**(3), 355–364 (1976)
3. Caristi, J.: Fixed point theorems for mappings satisfying inwardness conditions. Trans. Am. Math. Soc. **215**, 241–251 (1976)
4. Du, W.-S.: A simple proof of Caristi's fixed point theorem without using Zorn's lemma and transfinite induction. Thai J. Math. **14**(2), 259–264 (2016)
5. Ekeland, I.: On the variational principle. J. Math. Anal. Appl. **47**(2), 324–353 (1974)
6. Ekeland, I.: Nonconvex minimization problems. Bull. Am. Math. Soc. **1**(3), 443–474 (1979)
7. Fernández-Duque, D., Shafer, P., Towsner, H., Yokoyama, K.: Abstracts of the 2017 European summer meeting of the association for symbolic logic: Caristi's fixed point theorem and strong systems of arithmetic. Bull. Symb. Log. **24**(2), 243–244 (2018)
8. Fernández-Duque, D., Shafer, P., Towsner, H., Yokoyama, K.: Abstracts of the 2018 North American annual meeting of the association for symbolic logic: reverse mathematics, Ekeland's variational principle, and Caristi's fixed point theorem. Bull. Symb. Log. **24**(4), 499 (2018)
9. Fernández-Duque, D., Shafer, P., Towsner, H., Yokoyama, K.: Symposium, eleventh panhellenic logic (PLS 11, 2017: the reverse mathematics of Caristi's fixed point theorem. Bull. Symb. Log. **24**(3), 384–385 (2018)
10. Friedman, H.: Some systems of second order arithmetic and their use. Proceedings of the International Congress of Mathematicians (Vancouver, BC, 1974), vol. 1, pp. 235–242 (1975)
11. Georgiev, P.G.: The strong Ekeland variational principle, the strong drop theorem and applications. J. Math. Anal. Appl. **131**(1), 1–21 (1988)

12. Hájek, P., Pudlák, P.: Metamathematics of First Order Arithmetic. Springer, Berlin, Heidelberg, New York (1993)
13. Pohlers, W.: Subsystems of set theory and second order number theory. Handb. Proof Theory **137**, 209–335 (1998)
14. Simpson, S.G.: Subsystems of Second Order Arithmetic, 2nd edn. Perspectives in Logic, Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie (2009)
15. Yokoyama, K.: Standard and non-standard analysis in second order arithmetic, Tohoku Mathematical Publications, vol. 34. Mathematical Institute, Tohoku University, Sendai (2009). Thesis, Tohoku University, Sendai (2007)