Turing Computability of (Non-)Linear Optimization*

Martin Ziegler
Mathematics and Computer Science
Universität Paderborn
33095 GERMANY
ziegler@uni-paderborn.de

Vasco Brattka
Theoretical Computer Science I
FernUniversität Hagen
58084 GERMANY
vasco.brattka@fernuni-hagen.de

Abstract

We consider the classical LINEAR OPTIMIZATION problem, but in the Turing rather than the REAL-RAM model. Asking for mere computability of a function's maximum over some closed domain, we show that the common presumptions 'full-dimensional' and 'bounded' in fact cannot be omitted: The sound framework of Recursive Analysis enables us to rigorously prove this folkloristic observation! On the other hand, convexity of this domain may be weakened to connectedness, and even *non*-linear functions turn out to be effectively optimizable.

1 Motivation

The gap between (theoretical) algorithm design and (practical) implementation reveals a particular challenge in Computational Geometry: Provably correct algorithms keep, upon implementation, reporting not only inaccurate but sometimes even entirely invalid results, especially upon input of degenerate configurations [9] — for obvious reasons: Algorithms in Computational Geometry are most generally developed in the REAL-RAM model [1], capable of operating on real numbers *exactly*. Actual digital computers however can process in each step only a finite amount of information [2].

We therefore find it necessary to consider a different model of real number computation. In fact, Alan Turing himself introduced 'his' machine in order to study computability aspects over $\mathbb{R}$ [14] and thereby initiated the nowadays well-established field of RECURSIVE ANALYSIS [8, 12, 15]. Roughly speaking, a Turing machine is said to compute the real number r if it can output rational approximations of arbitrarily prescribable precision.

*Partially supported by DFG Grants Me872/7-3 and Br1807/4-1

It is straight forward to adopt this model to Computational Geometry. In fact [6] already did so in investigating Turing computability of the extreme points of a convex body, a continuous generalization of classical convex hull problems. [11] considers geometric predicates on convex bodies such as MEMBERSHIP and LINEINTERSECT in exactly the same sense. Also Grötschel, Lovász, and Schrijver in their enlightening book [7] developed polynomial-time Turing-algorithms for CONVEX OPTIMIZATION up to prescribable error $\varepsilon > 0$: although not mentioned explicitly, their WEAK OPTIMIZATION PROBLEM is exactly in the spirit of Recursive Analysis.

The present work generalizes a fundamental result from [7]: Dropping the requirement of *polynomial* time complexity, we present Turing algorithms for the optimization of arbitrary continuous objective functions. Even the convexity condition may be relaxed: It is shown that one can computably optimize over compact, connected, and everywhere full-dimensional domains, so-called 'bodies'. To this end, we establish new computability results for operators on bodies such as union, intersection, and function pre-image. The latter considerations resemble similar ones in [5] which, however, were performed in the domain-theoretic setting.

2 Introduction

Recursive Analysis combines Classical Analysis with discrete Recursion Theory: Results of the former such as INTERMEDIATE VALUE or FIXED POINT Theorems claiming *existence* of certain objects are investigated with respect to the *computability* of these objects [13, 10, 16, 18]. A real number $r \in \mathbb{R}$ for example is computable if there exists some Turing machine which outputs two sequences p_n and q_n of integers such that $|r - p_n/q_n| < 2^{-n}$.

Obviously, every rational number is computable in this sense; and so are some irrationals such as π, e, $\sqrt{2}$. On the other hand not *every* of the uncountable many real numbers can be computed by only countable many Turing machines: any non-recursive subset N of $\mathbb{N}$ gives rise to a non-computable real $\sum_{n\in N} 2^{-n}$.

As usual, Turing machines become able to compute non-numeric objects (e.g., graphs) by *encoding* them into binary strings and operating on these 'names': Following the general framework of TYPE-2 THEORY OF EFFECTIVITY [15], let X be an arbitrary set not exceeding continuum cardinality $|X| \leq \mathfrak{c}$ and $\alpha :\subseteq \Sigma^{\mathbb{N}} \to X$ a fixed partial surjective mapping, a so-called representation. Machine M is said to α-compute $x \in X$ if it outputs an α-name for x, that is, an infinite string $\bar{\sigma} \in \Sigma^{\mathbb{N}}$ such that $\alpha(\bar{\sigma}) = x$. For infinite outputs to make sense after finite time, the output tape is required to be *one-way*.

In the above example $X = \mathbb{R}$, let $\rho(\bar{\sigma})$ be defined for all strings of the form $\bar{\sigma} = \big(\mathrm{bin}(p_1), \mathrm{bin}(q_1), \mathrm{bin}(p_2), \mathrm{bin}(q_2), \ldots\big)$ where $r := \lim_{n\to\infty}(p_n/q_n)$ exists and $|r - p_n/q_n| < 2^{-n}$; here $\mathrm{bin}(i)$ denotes usual binary encoding of integer i. This $\rho : \bar{\sigma} \mapsto r$ is the standard representation of real numbers.

Different representations of the same set X may induce different notions of computability. For instance define $\rho_<(\bar{\sigma})$ to be the real number r iff $\bar{\sigma}$ is (an encoding of) a sequence of rationals converging to r *from below*. It is easy to see that any ρ-computable real is $\rho_<$-computable as well, but not vice versa. If however some $r \in \mathbb{R}$ is both $\rho_<$-computable and $\rho_>$-computable — the latter defined analogously by convergence from above — then r must be ρ-computable, too.

The preceding considerations dealt with computability of single objects x. This was reflected by the fact that the corresponding machine M received no input but had to output a name σ for x 'from scratch'. Let us now fix two sets X and Y with corresponding representations α and β. A function $f : X \to Y$ is (α,β)-computable if some Turing machine, upon input of any α-name $\bar{\sigma}$ for $x \in X$, outputs a β-name $\bar{\tau}$ for $y = f(x)$.

The MAIN THEOREM OF RECURSIVE ANALYSIS now states that any (ρ,ρ)-computable real function must necessarily be continuous! More generally: Let $\Sigma^{\mathbb{N}}$ be equipped with the Cantor Topology and $\alpha :\subseteq \Sigma^{\mathbb{N}} \to X$, $\beta :\subseteq \Sigma^{\mathbb{N}} \to Y$ admissible representations, (i.e., 'well-behaved' w.r.t. the final topologies on X and Y). If $f : X \to Y$ is (α,β)-computable, then f is continuous: see Theorem 3.2.11 in [15].

The issue of non-/uniform computability is best illustrated by comparing the following two statements:

Heredity: x α-computable $\Rightarrow$ $f(x)$ β-computable.

Effectivity: $f : x \mapsto f(x)$ is (α,β)-computable.

Call a representation α of X reducible to α' of X ($\alpha \preceq \alpha'$), if some Turing machine converts α-names into α'-names for *the same* object $x \in X$. Equivalent, i.e., pairwise reducible representations lead to identical notions of uniform computability. It holds, e.g.,

$$\rho \preceq \rho_<,\ \rho \preceq \rho_>, \quad \rho_< \not\preceq \rho,\ \rho_> \not\preceq \rho, \quad \rho \equiv (\rho_< + \rho_>)$$

We mention two ways of constructing new representations from given ones, details can be found in [15]:

- Let $\alpha_1, \ldots, \alpha_n$ be representations of $X_1, \ldots, X_n$, respectively. There exists a canonical representation, denoted $\prod_{i=1}^n \alpha_i$, of $\prod_{i=1}^n X_i$.
- Let α and β be admissible representations of X and Y, respectively. There exists a canonical representation of $C(X,Y)$, the space of continuous functions $f : X \to Y$, which will be denoted by $[\alpha \to \beta]$.

3 Computability of Sets

For subsets of $\mathbb{R}^n$, several notions of computability are treated in literature. For instance, $\emptyset \neq A \subseteq \mathbb{R}^n$ is Turing-located [6] iff the Euclidean distance function

$$d_A : \mathbb{R}^n \to \mathbb{R}, \qquad x \mapsto \inf_{a\in A} \sqrt{\sum\nolimits_{i=1}^n (x_i - a_i)^2}$$

is (ρ^n,ρ)-computable. This d_A can be regarded as continuous analogue to the characteristic function of A which, except for trivial cases $A = \emptyset$ and $A = \mathbb{R}^n$, is never computable nor continuous.

It is easy to see that d_A characterizes the set A uniquely, *provided* A is closed. In other words, we have a representation (from now on called ψ^n) of the hyperspace $\mathfrak{A}^n$ of closed non-empty subsets of $\mathbb{R}^n$: $\bar{\sigma}$ is a ψ^n-name for A iff $\bar{\sigma}$ is a $[\rho^n \to \rho]$-name for d_A. On the other hand, no representation of *all* subsets can exist as there are too many of them: $|2^{\mathbb{R}}| > \mathfrak{c}$.

By definition, $A \in \mathfrak{A}^n$ is Turing-located iff it is ψ^n-computable. Two weaker notions, $\psi^n_<$ and $\psi^n_>$, ask for $[\rho^n \to \rho_>]$- and $[\rho^n \to \rho_<]$-computability of d_A, respectively. Notice the reversed index: $\psi^n_> = [\rho^n \to \rho_<]$.

For compact, i.e. bounded closed, sets it turns out that many problems become computable only if, in addition to the ψ^n-information on A, some *explicit* upper bound on its diameter is supplied. Let

$$\mathfrak{K}^n := \big\{(A,r) : A \in \mathfrak{A}^n, r \in \mathbb{N}, A \subset [-r,r]^n\big\} \subset \mathfrak{A}^n \times \mathbb{N}$$

denote the hyperspace of all non-empty compact subsets and $\kappa^n := (\psi^n \times \mathrm{bin})$ its canonical representation; similarly $\kappa^n_<$ and $\kappa^n_>$. We tacitly assume that $\mathfrak{K}^n$ be embedded into $\mathfrak{A}^n$ the obvious way. *Non*-uniform it holds that compact $A \subset \mathbb{R}^n$ is ψ^n-computable iff it is κ^n-computable. Also, obviously $\kappa^n \preceq \psi^n\big|^{\mathfrak{K}^n}$; but $\psi^n\big|^{\mathfrak{K}^n} \not\preceq \kappa^n$.

It is well-known that κ^n (but not ψ^n) permits a computable version of the famous HEINE-BOREL Theorem, see [15]. Even more, the maximum of arbitrary continuous functions over compact sets can effectively be computed:

Lemma 1 (Corollary 6.2.5 in [15]) *Maximum of continuous functions over a compact set* $C(\mathbb{R}^n,\mathbb{R}) \times \mathfrak{K}^n \ni (f,K) \mapsto \max f[K] \in \mathbb{R}$ *is* $\big([\rho^n \to \rho] \times \kappa^n, \rho\big)$*-computable.*

On the other hand, compact sets are quite often described not by κ^n-names but, for example a convex body, as solutions to a system of linear inequalities:

$$L(A,b) = \big\{x \in \mathbb{R}^n : A\cdot x \le b\big\}, \quad A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m$$

The very important question is thus:
Can $(\rho^{m\times n} \times \rho^m)$*-names for* (A,b) *be converted to* κ^n*-names for* $L(A,b)$ *in a computable way?*
In general, they can't. Even more, LINEAR OPTIMIZATION with degenerate constraint hyperplane configurations is in general not computable! This folkloristic observation, usually avoided by *general position* presumptions, now follows rigorously in the sound framework of Recursive Analysis from its Main Theorem and the following

Example 2 (Discontinuous LP Optimum) *In dimension 2, fix* $f : (x,y) \mapsto x$ *and* $0 \le \varepsilon < 1$*; consider LP:*

$$\begin{array}{rcl} y & \ge & \varepsilon x \\ y & \le & 0 \\ x & \ge & -1 \\ x & \le & 1 \end{array} \qquad \textit{i.e.} \qquad A = \begin{pmatrix} \varepsilon & -1 \\ 0 & 1 \\ -1 & 0 \\ 1 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

$\varepsilon > 0$: $L(A,b) = \mathrm{convHull}\big\{(0,0),(-1,0),(-1,-\varepsilon)\big\}$, $f[L(A,b)] = [-1,0]$, $\max f[L(A,b)] = 0$

$\varepsilon = 0$: $L(A,b) = [-1,1]\times\{0\}$, $\max f[L(A,b)] = 1$ □

Another problem arises from the fact that the hyperspace of compact subsets of $\mathbb{R}^n$ does not contain half-spaces nor is it effectively closed under intersection or pre-image [3]:

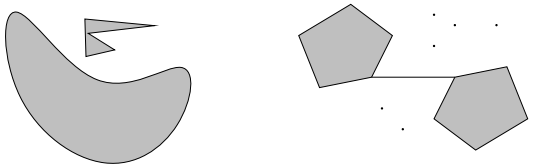
- Binary union $\cup : \mathfrak{A}^n \times \mathfrak{A}^n \to \mathfrak{A}^n$, $(A,B) \mapsto A \cup B$ is $(\psi^n\times\psi^n, \psi^n)$- and $(\kappa^n\times\kappa^n, \kappa^n)$-computable;
- whereas intersection $(A,B) \mapsto A \cap B$ is not even $(\kappa^n \times \kappa^n, \psi^n_<)$-continuous.
- Function pre-image $A \mapsto f^{-1}[A]$ is not $(\kappa^m, \psi^n_<)$-computable, even for computable $f : \mathbb{R}^n \to \mathbb{R}^m$.

4 Computable Solids

It is important to observe that the two-dimensional Example 2 heavily relies on the set $L(A,b)$ to become one-dimensional for $\varepsilon = 0$. In fact, we track down the (non-) computability of LINEAR OPTIMIZATION to an issue of *full-dimensionality* rather than non-degeneracy. More precisely, our new Theorem 4 shows that operations union, intersection, and pre-image become effective when restricting to everywhere full-dimensional sets.

Call set A everywhere full-dimensional (or regular) if it coincides with the closure of its topological interior $\overline{A^\circ}$. This notion generalizes the full-dimensionality from convex sets, cf. Lemma 11 of the full version. A solid is a regular non-empty proper subset of $\mathbb{R}^n$. The below figure exemplifies, in dimension 2,

A) A SOLID AND B) A NON-SOLID SET.



Let $\mathfrak{X}^n$ denote the hyperspace of all solids in $\mathbb{R}^n$, then

$$A \in \mathfrak{X}^n \underset{\text{d)+h)}}{\overset{\text{L. 9}}{\iff}} \emptyset \neq A = \overline{\mathbb{R}^n \setminus B} \quad \text{for } \emptyset \neq B := \overline{\mathbb{R}^n \setminus A}$$

Encode $A \in \mathfrak{X}$ simultaneously by $\psi^n_>$-names for both A and B, i.e., effective approximations from below of *two* distances: to the set and to its closed complement. Let η^n denote the corresponding representation.

On the other hand by virtue of Lemma 10d):

$$A \in \mathfrak{X}^n \iff A = \overline{\mathbb{R}^n \setminus B} \quad \text{for } \textit{some } B \in \mathfrak{A}^n$$

One may thus extend the domain of η^n and encode $A \in \mathfrak{X}^n$ by a $\psi^n_>$-name for A and a $\psi^n_>$-name for *any* B with $A = \overline{\mathbb{R}^n \setminus B}$. Let ξ^n denote this joint representation. We made the important observation that both

approaches are only logically but not computationally equivalent:

Lemma 3 $\eta^n \preceq \xi^n$, *but* $\xi^n \not\preceq \eta^n$. □

It turns out that ξ^n serves better for generating κ^n-names of $L(A,b)$: Many so-far discontinuous operators on closed sets become ξ^n-computable after restricting them to solids. For intersection and preimage, precautions have to be taken to avoid non-solid results:

Theorem 4 *a) Union is* $(\xi^n \times \xi^n, \xi^n)$*-computable.*

b) Intersection of solids set is $(\xi^n \times \xi^n, \xi^n)$*-computable whenever well-defined, i.e., on* $\{(A,B) : A,B \in \mathfrak{X}^n, A \cap B \in \mathfrak{X}^n\} \subset \mathfrak{X}^n \times \mathfrak{X}^n$

c) Pre-image under continuous open *mappings* $f \in C(\mathbb{R}^n, \mathbb{R}^m)$ *is* $\big([\rho^n \to \rho^m] \times \xi^m, \xi^n\big)$*-computable.*

d) $\xi^n \preceq \psi^n$*; even restricted to solids,* $\psi^n|^{\mathfrak{X}^n} \not\preceq \xi^n$.

e) Restricted to non-empty, compact and connected sets, $\psi^n\big|^{\{A \in \mathfrak{K}^n \text{ connected}\}} \preceq \kappa^n$ □

Combining parts d) and e) and Lemma 3 with Lemma 1, it immediately follows that *any* computable function can effectively be maximized over a body (compact, connected, non-empty, everywhere full-dimensional) $B \subset \mathbb{R}^n$, given by either a ξ^n-name, a η^n-name, a ψ^n-name, or a κ^n-name!

Let us now show how such a name for full-dimensional $L(A,b)$ can be obtained from a $(\rho^{m\times n} \times \rho^m)$-name of (A,b): For $\lambda_0 \in \mathbb{R}$, $0 \neq \lambda \in \mathbb{R}^n$, $H_{\lambda_0,\lambda}$ denotes the closed halfspace $\{x \in \mathbb{R}^n : \sum_{i=1}^n \lambda_i x_i \geq \lambda_0\}$.

Theorem 5 *a) The mapping* $\mathbb{R} \times (\mathbb{R}^n \setminus \{0\}) \ni (\lambda_0, \lambda) \mapsto H_{\lambda_0,\lambda} \in \mathfrak{X}^n$ *is* $(\rho \times \rho^n, \xi^n)$*-computable.*

b) Upon $(\rho^{m\times n} \times \rho^m)$*-input of* $(A,b) \in \mathbb{R}^{m\times n} \times \mathbb{R}^m$, $L(A,b)$ *is* ξ^n*-computable,* provided A *contains no zero-rows and* $L(A,b)$ *is full-dimensional.* □

Let us remark that, apart from full-dimensionality (Example 2), boundedness is an essential prerequisite, too, in order to ensure computability of LINEAR OPTIMIZATION:

Example 6 *In dimension 2, fix* $f : (x,y) \mapsto x$. *The LP*

$$x \leq \varepsilon y \qquad \textit{i.e.} \qquad A = (1, -\varepsilon), \quad b = (0)$$

defines, for each $0 \leq \varepsilon < 1$*, a full-dimensional domain* $L(A,b)$. *But* $\max f[L(A,b)]$ *is discontinuous in* ε*:* $\max f\big[L(A,b)\big] = \infty$ *for* $\varepsilon > 0$, $= 0$ *for* $\varepsilon = 0$. □

5 Conclusion

We presented Turing-computability results for BODIES, i.e., non-empty bounded connected regular closed subsets of $\mathbb{R}^n$. As objects cannot be physically distinguished anyway if they differ only on a nowhere dense set, the full-dimensionality requirement is no restriction in practice. Theorem 5 generalizes previous considerations [4] on computability of linear equalities to *in*equalities. Proof techniques however are entirely different: Topology rather than Linear Algebra, see the full version.

References

[1] M. de Berg, M. van Kreveld, M. Overmars, O. Schwarzkopf: "Computational Geometry", Springer (1997).

[2] V. Brattka, P. Hertling: "Feasible Real Random Access Machines", pp.490-526 in *Journal of Complexity* **14** (1998).

[3] V. Brattka, K. Weihrauch: "Computability on Subsets of Euclidean Space I: Closed and Compact Subsets", pp.65-93 in *Theoretical Computer Science* **219** (1999).

[4] V. Brattka, M. Ziegler: "Computability of Linear Equations", submitted; further information available from `http://www.phys.upb.de/~ziegler/basis.html`

[5] Abbas Edalat, A. Lieutier: "Foundation of a Computable Solid Modelling", to appear in *Theoretical Computer Science* (2002).

[6] Xiaolin Ge, Anil Nerode: "On extreme points of convex compact Turing located sets", pp.114-128 in *Logical Foundations of Computer Science*, Springer LNCS **813** (1994).

[7] M. Grötschel, L. Lovász, A. Schrijver: "Geometric Algorithms and Combinatorial Optimization", Springer (1988).

[8] A. Grzegorczyk: "On the definitions of computable real continuous functions", *Fundamenta Mathematicae* **44** (1957) 61-77.

[9] P. Hertling, K. Weihrauch: "Levels of Degeneracy and Exact Lower Complexity Bounds for Geometric Algorithms", pp.237-242 in *Proceedings of the 6th CCCG* (1994).

[10] Ker-I Ko: "Complexity Theory of Real Functions", *Progress in Theoretical Computer Science*, Birkhäuser Boston (1991).

[11] M. Kummer, M. Schäfer: "Computability of Convex Sets", pp.550-561 in *Proceedings of the STACS*, Lecture Notes in Computer Science **900**, Springer (1995).

[12] Daniel Lacombe: "Les ensembles récursivement ouverts ou fermés, et leurs applications à l'Analyse récursive", pp.28-31 in *Compt. Rend. Acad. des Sci. Paris*, **246** (1958).

[13] M.B. Pour-El and J.I. Richards: "Computability in Analysis and Physics", Springer (1989).

[14] Alan M. Turing: "On computable numbers, with an application to the Entscheidungsproblem", pp.230-265 in *Proc. of the London Math. Soc.* **42(2)** (1936).

[15] Klaus Weihrauch: "Computable Analysis", Springer (2000).

[16] Kam-Chau Wong: "Computability of Minimizers and Seperating Hyperplanes", pp.564-568 in *Math. Logic Quaterly* **42** (1996).

[17] M. Ziegler, V. Brattka: "Computing the Dimension of Linear Subspaces", pp.450-458 in *SOFSEM'2000: Theory and Practice of Informatics*, Springer LNCS **1963** (2000).

[18] M. Ziegler, V. Brattka: "A Computable Spectral Theorem", pp.378-388 in *Computability and Complexity in Analysis*, Lecture Notes in Computer Science **2064**, Springer (2001).