Computable Randomness and Differentiability in $\mathbb{R}^n$

Aleksander Galicki
Department of Computer Science
University of Auckland

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy
2018

This thesis studies connections between computable randomness in $\mathbb{R}^n$ and various properties related to differentiability. This research was inspired by results in two recent papers: one by Brattka, Miller and Nies and another one by Freer, Kjos-Hanssen, Nies and Stephan. In those papers it was shown that computable randomness on the unit interval can be characterized by differentiability properties of computable Lipschitz functions and computable monotone function. We generalize to $\mathbb{R}^n$ most of those results. Moreover, we show several new results of this kind both on the real line and on $\mathbb{R}^n$. In the process, we prove effective versions of several notable classical results such as: Rademacher's theorem, Aleksandrov's theorem, Sard's theorem for monotone Lipschitz functions and Brenier's theorem. In most cases we prove converse results as well.

Acknowledgements

There is one person without whom writing this thesis would not have been nearly as enjoyable. Thank you, Andre.

Contents

1 Introduction **1**
- 1.1 What this thesis is about . . . 1
- 1.2 Computable analysis and algorithmic randomness . . . 1
- 1.3 Randomness and "a.e. theorems" . . . 3
- 1.4 Randomness and differentiability of real-valued functions on $\mathbb{R}^n$ . . . 5
- 1.5 Non-differentiability sets . . . 7
- 1.6 Computable randomness and differentiability . . . 7
- 1.7 Summary of contributions . . . 9
- 1.8 Structure of the thesis . . . 10

2 Preliminaries **11**
- 2.1 Computable analysis . . . 11
  - 2.1.1 Computable elements of $\mathbb{R}^n$ . . . 11
  - 2.1.2 Computable functions on $\mathbb{R}^n$ . . . 12
- 2.2 Computable metric spaces . . . 13
- 2.3 Computable measures . . . 14
- 2.4 Dyadic rationals and dyadic cubes in $\mathbb{R}^n$ . . . 15
- 2.5 Computable randomness . . . 16
  - 2.5.1 Computable randomness in $\mathbb{R}^n$ . . . 18
- 2.6 Differentiability of real-valued functions on $\mathbb{R}^n$ . . . 21
  - 2.6.1 Functions on the real line . . . 21
  - 2.6.2 Functions on $\mathbb{R}^n$, partial and directional derivatives . . . 22
  - 2.6.3 Differentiability of partial functions . . . 23
  - 2.6.4 Derivatives and antiderivatives . . . 23
  - 2.6.5 Non-differentiability sets of functions . . . 24
  - 2.6.6 Approximate continuity, essential values and Lebesgue points . 24
- 2.7 Lipschitz, convex and monotone functions . . . 25
  - 2.7.1 Properties of continuous convex functions on $\mathbb{R}$ . . . 27
  - 2.7.2 Monotone functions . . . 27
  - 2.7.3 Subdifferentials of convex functions . . . 29
  - 2.7.4 Minty's correspondence . . . 31

2.7.5 Other properties of convex functions on $\mathbb{R}^n$ 32
2.7.6 Convex conjugate functions 32
2.7.7 Monge-Ampère measures 33
2.7.8 Symmetric derivative of a measure 33

3 Computable randomness and differentiability on the unit interval 35
3.1 Porosity, differentiability and betting 35
3.1.1 Martingales, measures and monotone functions 35
3.1.2 Porosity points 36
3.1.3 Computable monotone functions are differentiable at computably random reals 38
3.2 Zahorski's construction on the real line 39
3.2.1 Classical argument 40
3.2.2 Effective version 40
3.3 Convex functions, their derivatives and probability measures 41
3.3.1 Non-differentiability sets of computable convex functions . . . 42
3.3.2 Derivatives of computable convex functions 45
3.3.3 Differentiability of monotone weakly computable functions . . 49
3.3.4 New characterisations of computable randomness 50
3.3.5 Atoms and differentiability of computable probability measures on the real line 51

4 Computable randomness and differentiability in $\mathbb{R}^n$ 54
4.1 Effective Rademacher theorem 56
4.1.1 Existence of partial derivatives 57
4.1.2 Existence of directional derivatives 58
4.1.3 Linearity of directional derivatives 59
4.1.4 The proof 60
4.2 Computable monotone Lipschitz functions preserve non-randomness . 61
4.3 An effective version of Sard's theorem for monotone Lipschitz functions 66
4.4 Differentiability of monotone functions on $\mathbb{R}^n$ 68
4.5 Convex functions on $\mathbb{R}^n$ 68
4.6 Almost everywhere computable monotone functions on $\mathbb{R}^n$ 69

5 Effective Brenier theorem 72
5.1 Optimal transport 73
5.1.1 Kantorovich's optimal transportation problem 73
5.1.2 Monge's optimal transportation problem 74
5.1.3 Optimal transportation theorem for quadratic cost 74
5.1.4 Kantorovich duality 76
5.2 Proof of the Effective Brenier theorem 77

5.3 Subspaces of Lipschitz functions 80
5.4 Optimal cost is computable 86

6 Controlling non-differentiability in $\mathbb{R}^n$ 89
6.1 Reinterpreting the effective Zahorski construction 90
6.2 Effective Zahorski construction on $\mathbb{R}^n$ 93
6.2.1 The Monge-Ampère equation and regularity of optimal transport 94
6.2.2 Exhibiting a non-differentiable computable measure 95
6.3 Consequences of the main theorem 103
6.3.1 Monotone and convex functions 103
6.3.2 Probability measures on $[0, 1]^n$ 104
6.3.3 Critical values of computable monotone Lipschitz functions . . 105
6.3.4 The Monge-Ampère equation 105

Notation index 108

Bibliography 111

Chapter 1

Introduction

1.1 What this thesis is about

The subject of this thesis lies at the interface of computable analysis and algorithmic randomness. Computable randomness is a randomness notion defined in terms of effective betting strategies. One of the main results in the paper of Brattka, Miller and Nies [BMN16] shows that computably random points on the unit interval can be characterized as points where all computable monotone real-valued functions of one variable are differentiable. An analogous result for computable Lipschitz functions has been proven in the paper of Freer, Kjos-Hanssen, Nies and Stephan [FKHNS14].

Those results suggest that there are close connections between computable randomness and differentiability of certain classes of objects in Euclidean spaces. In order to advance our understanding of this particular phenomenon, we pursue two directions of research. Firstly, we seek to generalize, where possible, known results to higher dimensions (that is to $\mathbb{R}^n$). Secondly, we study how computable randomness and differentiability interact for other classes of objects that are closely related to monotone and Lipschitz functions.

1.2 Computable analysis and algorithmic randomness

The theory of *algorithmic randomness* (see Nies [Nie09], Downey and Hirschfeldt [DH10]) is the area of mathematics that formalizes and studies the intuitive notion of randomness. Intuitively, an element of some space is random if it does not exhibit any exceptional properties. This approach can be formalized, utilizing measure theory and computability, by identifying exceptional properties with effective null sets. Different types of effective null sets correspond to different notions of algorithmic randomness. This use of computability theory to rigorously specify which properties are exceptional, explains the word "algorithmic" in "algorithmic randomness". Below we provide some informal examples of known randomness notions.

> **Computable randomness**
> An infinite binary sequence is said to be *computably random*, if no effective betting strategy can make unbounded profits while betting on this sequence. This notion of randomness can be generalized to $\mathbb{R}^n$ by identifying elements of $\mathbb{R}^n$ with their binary expansions.

> **Weak 2-randomness**
> An element of a computable measure space (X, μ) is said to be weakly 2-random if it does not belong to any effective G_δ null set $V \subseteq X$.

> **Martin-Löf randomness**
> Let (X, μ) be a computable measure space. A $V \subseteq X$ is a Martin-Löf test if there is a computable sequence $(G_i)_{i\in\mathbb{N}}$ of effectively open subsets such that $V = \cap G_i$ and $\mu(G_j) \leq 2^{-j}$ for all j.
> $x \in X$ is said to be Martin-Löf random if it does not belong to any Martin-Löf test.

Computable analysis (see Weihrauch [Wei00], Pour-El and Richards [PER89] and Ker-I Ko [Ko91]), building on tools and foundations of computability theory, studies effective versions of notions and results from analysis. In this area various notions of computability of mathematical objects from analysis (for example, real numbers, real-valued functions on $\mathbb{R}^n$, measures, closed subsets of $\mathbb{R}^n$) are rigorously defined and studied.

It is known that for functions from integers to integers, all sensible known notions of computability are equivalent. The situation is different for most classes of objects studied in analysis, such as real-valued functions on $\mathbb{R}^n$, real numbers, etc. In those cases, there are multiple known notions of computability which are pairwise incompatible.

We customarily use the term "effective" to denote a known and rigorously defined notion of computability without specifying which one. As such, this word is reserved to informal discussions, while all mathematical results either explicitly specify which notions of computability are used, or this information is clear from the context. The main notion of computability for real-valued functions on $\mathbb{R}^n$ is that of Grzegorczyk-Lacombe (we will define this notion rigorously later). Every time we write about "computable functions" on $\mathbb{R}^n$, we mean this particular notion of computability.

A typical result in computable analysis assumes effectiveness of some or all of the objects in the premises and then asserts (some level of) effectiveness of objects in the conclusion. Consider the following two examples.

> **Theorem (A)**
> If $f : \mathbb{R} \to \mathbb{R}$ is a computable Lipschitz function, then its derivative exists at every computably random z.

> **Theorem (B)**
> If $f : \mathbb{R} \to \mathbb{R}$ is computable and C^2, then its derivative is a computable function.

Results such as Theorem (A) are effective versions of known theorems from analysis. Not all results are of this kind: Theorem (B) is an example of a statement for which its non-effective version is trivial.

Since algorithmic randomness studies effective null sets and measure theory plays a prominent role in analysis, there are numerous connections between algorithmic randomness and computable analysis. Theorem (A) is an example of a result at the interface of those two areas.

1.3 Randomness and "a.e. theorems"

Randomness notions are examples of mathematical properties that hold almost everywhere (that is, on a set of full measure with respect to some fixed measure). Such properties appear naturally in various contexts and play important roles in many established areas of mathematics, such as analysis, ergodic theory, integration theory, geometric measure theory and others. Consider statements of the following form:

(AEP) given an object e of some class C, the property $P(e, x)$ holds for almost all $x \in X$ with respect to some measure μ on X.

A statement of this kind can be interpreted as saying that elements of the class C exhibit a high degree of regularity with respect to the property P: given $e \in C$, its *set of irregularity*, $I_P(e) = \{x : \neg P(e, x)\}$, is negligible.

Many known mathematical results appear in this form. We call those *a.e. theorems.* Notable examples of a.e. theorems include:

> **Rademacher's theorem (Rademacher, 1919)**
> If $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz, then it is differentiable at almost all $x \in \mathbb{R}^n$.

Birkhoff's Ergodic Theorem (Birkhoff, 1931)

Let (X, μ) be a probability space, and let $T : X \to X$ be ergodic. Let $f : X \to \mathbb{R}$ be $L^1(X)$. Then for almost all X,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(T^j(x)) = \int f \, \mathrm{d}\mu.$$

Lebesgue differentiation theorem (Lebesgue, 1910)

If $f \in L^1_{loc}(\mathbb{R}^n)$, then for almost all $x \in \mathbb{R}^n$,

$$f(x) = \lim_{r\to 0} \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} f \, \mathrm{d}\lambda.$$

A natural question to ask is how known randomness notions are related to a.e. theorems from various areas of mathematics. One meaningful way of approaching this question is to consider effective variants of (AEP) statements. Effectivization restricts the class C to some countable subset of its effective members and ensures the complement of the union of all sets of irregularities

$$R_P = \bigcap_{e \in C \text{ is effective}} \{ \, x \mid P(e, x) \}$$

is a set of full measure. This means that every effective variant of an a.e. theorem can be seen as a definition of a randomness notion. The question then is to compare such notions to known randomness notions. Recent research suggests, somewhat surprisingly, that more often than not effective versions of a.e. theorems correspond to only a small number of well known algorithmic randomness notions. Below we describe two examples related to two of the above-mentioned a.e. theorems.

By results of Bienvenu, Day, Hoyrup, Mezhirov and Shen [BDH+12], and, independently, by Franklin, Greenberg, Miller and Ng [FGMN12], a particular effective version of Birkhoff's ergodic theorem characterizes Martin-Löf randomness. Specifically, they have shown the following result:

Theorem 1.3.1. *Let (X, μ) be a computable probability space, and let $T : X \to X$ be a computable ergodic map. Then $x \in X$ is Martin-Löf random if and only if for all effectively closed subsets $C \subseteq X$,*

$$\lim_{n\to\infty} \frac{\#\{i < n \ : \ T^i(x) \in C\}}{n} = \mu(C).$$

Pathak, Rojas and Simpson [PRS14] matched an effective form of the Lebesgue differentiation theorem to Schnorr randomness via the following result:

Theorem 1.3.2. *For all $x \in [0,1]^n$ the following are pairwise equivalent.*

1. *x is Schnorr random.*
2. *$\lim_{Q \to x} \frac{\int_Q f(x)\,\mathrm{d}x}{\lambda(Q)}$ exists for all L^1-computable functions $f \in L^1([0,1]^n)$, where the limit is taken over all cubes Q containing x as the diameter of Q tends to 0.*

As it was mentioned before, often there are multiple incompatible notions of computablity known for a given class of mathematical objects. This means that one a.e. theorem often can be effectivized in a number of ways and different effectivizations could characterize different randomness notions. For example, Pathak, Rojas and Simpson [PRS14] characterized Schnorr randomness via another effective version of Birkhoff's ergodic theorem. Similarly, at the end of this thesis, we will prove an effective version of the Lebesgue differentiation theorem characterizing computable randomness.
A typical result of this kind consists of two parts:

- The *forward direction* part of the form "if z is random, then f is differentiable at z". It is an effective version of a given a.e. theorem.
- The corresponding effective version of the *converse direction*. This usually involves exhibiting an effective object with a prescribed set of irregularities.

Such results are interesting because they improve our understanding both of randomness notions and analytical properties used to characterize them. We will return to this point later.

1.4 Randomness and differentiability of real-valued functions on $\mathbb{R}^n$

Analysis in general and differentiability of functions on $\mathbb{R}^n$ in particular, is a rich source of a.e. theorems. Many classes of well-behaved functions are known to be a.e. differentiable, for example:

- Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, by Rademacher's theorem;
- real-valued functions belonging to $W^{1,p}(\mathbb{R}^n)$, by Calderon's theorem from 1951;
- monotone functions on $\mathbb{R}^n$, via result by Mignot [Mig76];
- cone-monotone real-valued functions on $\mathbb{R}^n$, by the result of Chabrillac and Crouzeix [CC87].

This opens up the possibility of studying differentiability of effective functions through the lens of algorithmic randomness. Moreover, it makes it possible to characterize algorithmic randomness notions (on $\mathbb{R}^n$) in terms of differentiability properties of effective functions. It is possible to show that non-differentiability points of computable real-valued functions on $\mathbb{R}^n$ form an effective G_δ set. This means that differentiability of computable functions can be used to characterize randomness notions weaker than or equal to weak-2-randomness. As it happens, most of major randomness notions fall into this category. A lot of research in this areas has been published, especially for computable real-valued functions on the unit interval:

- Brattka, Miller and Nies [BMN16] characterized weak-2-randomness, Martin-Löf randomness, computable randomness and Schnorr randomness in terms of differentiability of various classes of a.e. differentiable computable real-valued functions on the unit interval. They have studied the following classes of functions: a.e. differentiable functions, functions of bounded variation, absolutely continuous functions and monotone functions.

- Freer et al. [FKHNS14] have studied differentiability of computable Lipschitz functions. They have characterized computable randomness and Schnorr randomness using two different subclasses of computable Lipschitz real-valued functions on the unit interval.

- Nies in [Nie14] has shown that p-randomness can be characterized in terms of differentiability of polynomial time computable real-valued monotone functions on the unit interval.

Apart from the notion of differentiability of functions, there are other, closely related notions, which appear frequently in a.e. theorems. Some of them have been studied in the context of algorithmic randomness as well:

- Miyabe [Miy13] has characterized weak randomness in terms of Lebesgue points of a.e. computable real-valued function on X, where X is some computable metric space.

- Pathak, Rojas and Simpson [PRS14] and, independently, Rute [Rut13], have characterized Schnorr randomness in terms of Lebesgue points of L^1-computable real-valued functions on $\mathbb{R}^n$.

- Bienvenu, Hölzl, Miller and Nies [BHMN14], using an effective version of the Denjoy-Young-Saks theorem, characterized computable randomness in terms of Dini derivatives of computable real-valued functions on the unit interval.

1.5 Non-differentiability sets

Two-directional results mentioned in the previous section can be seen as randomness-focused characterizations of *non-differentiability sets* of effective functions. By a non-differentiability set of f we mean the set of points where f is not differentiable. Characterizations of non-differentiability sets of functions on $\mathbb{R}^n$ have been studied in analysis as well. Below we mention several results in this area that are particularly relevant in the context of this thesis.

Several important results are contained in the seminal paper of Zahorski [Zah46], who characterized non-differentiability sets of various classes of real-valued functions on the real line. In particular, he fully characterized non-differentiability sets of continuous functions, Lipschitz functions and monotone functions.

> *Remark* 1.5.1 (Zahorski's construction). The main step in his characterization of non-differentiability sets of continuous functions was the construction, for any G_δ set A of measure zero, of a monotone Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ such that its non-differentiability set is equal to A.

Zajicek [Zaj79] characterized non-differentiability sets of convex functions on $\mathbb{R}^n$.

The problem of characterizing non-differentiability sets of Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ has attracted a lot of attention. This interest is related to the quest of formulating and proving a converse to Rademacher's theorem. It turned out to be a highly non-trivial problem which took two decades of efforts to finally solve it. For more details, please see the paper by Alberti, Csörnyei and Preiss [APC11], and the paper by Preiss and Speight [PS15]. The vital part of this solution is the construction, for a given G_δ set $A \subseteq \mathbb{R}^n$ of measure zero, of a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^n$ which is non-differentiable at all elements of A. The paper with this construction, announced close to a decade ago, has not been published to this day.

1.6 Computable randomness and differentiability

The focus of this thesis is on a particular subset of results mentioned in Section 1.4 — those related to computable randomness and differentiability of computable monotone and computable Lipschitz functions. Our starting point is the following pair of theorems:

> **Theorem 1.6.1** ([BMN16]).
> *A real number $z \in [0, 1]$ is computably random $\iff$*
> *every computable monotone $f : [0, 1] \to \mathbb{R}$ is differentiable at z.*

Theorem 1.6.2 ([FKHNS14]).
A real number $z \in [0,1]$ *is computably random* $\iff$
every computable Lipschitz $f : [0,1] \to \mathbb{R}$ *is differentiable at* z.

Computable randomness is defined in terms of betting strategies on infinite binary strings. This definition is very natural in the space of infinite binary sequences. It can be extended to more general spaces, in particular to $\mathbb{R}^n$, but the definition becomes much less natural. Monotonicity and Lipschitz continuity, on the other hand, are two very natural notions from analysis in $\mathbb{R}^n$. It is curious how differentiability links those natural, but seemingly distant, notions.

Generalizing to $\mathbb{R}^n$ forward directions of those two results would involve proving effective versions of two known theorems from analysis: Rademacher's theorem and Mignot's result from [Mig76]. Generalizing converse directions is a trickier problem. As it was mentioned, the converse direction of Rademacher's theorem has attracted a lot of interest and a non-effective solution has been announced, but the construction of a Lipschitz function with prescribed non-differentiability properties has not been published yet. The converse to Mignot's theorem has not been studied. On the unit interval, both converse directions of Theorem 1.6.2 and Theorem 1.6.1 depended on an effective version of the construction mentioned in Remark 1.5.1. Generalizing this particular construction to $\mathbb{R}^n$ is one of the topics of this thesis.

Several other notions are naturally related to monotonicity and Lipschitz continuity. Among those are convex functions and positive measures. In higher dimensions there are several important a.e. theorems related to differentiability of monotone, Lipschitz and convex functions. As we will see, some of those characterize computable randomness too.

An important source of motivation for studying connections between algorithmic randomness and analysis is that such studies often deepen our understanding of both areas. Let us mention two examples from this thesis:

1. In Section 6.1 we describe a very natural generalization to $\mathbb{R}^n$ of Zahorski's construction. However, finding this generalization is much easier if the starting point is (the divergence set of) an effective betting strategy, rather then just an arbitrary G_δ null-set. In this case, starting from a computable randomness point of view made it easier to find a natural generalization of an important construction from classical analysis.

2. One of intermediate results on our way to effectivize Aleksandrov's theorem is an interesting preservation property: we showed that computable monotone Lipschitz functions on $\mathbb{R}^n$ preserve computable non-randomness.

1.7 Summary of contributions

Most of our contributions are related to characterizations of computable randomness in $\mathbb{R}^n$. We treat two cases, $n = 1$ and $n \geq 1$, separately. For $n = 1$, the starting point was a pair of theorems, Theorem 1.6.2 and Theorem 1.6.1. All our results on the unit interval are demonstrated in Chapter 3. In particular, we have proven four new characterizations of computable randomness on the real line in terms of differentiability:

1. our Theorem 3.3.18 contains a stronger version of Theorem 1.6.1 — for almost everywhere computable monotone functions,
2. the same theorem shows characterization of computable randomness in terms of twice-differentiability of computable convex functions, it is an effective version of Aleksandrov's theorem and its converse on the real line,
3. our Theorem 3.3.25 contains two characterizations of computable randomness in terms of differentiability of computable measures.

Moreover, we have fully characterized non-differentiability sets of real-valued convex functions on the real line in Theorem 3.3.3. This result was then used to characterize sets of atoms of computable probability measures on $\mathbb{R}$ in Proposition 3.3.22. Finally, our Theorem 3.1.9 is a new, relativized version of the forward direction of Theorem 1.6.1. The proof is based on ideas present in [Nie14].

In Chapter 4, we have proven effective version of several known a.e. theorems on $\mathbb{R}^n$. Our Theorem 4.1.1 is an effective version of Rademacher's theorem. It is a generalization of the forward direction of Theorem 1.6.2 to $\mathbb{R}^n$. Our Theorem 4.3.2 is a computable version of Sard's theorem for monotone Lipschitz functions. Theorem 4.4.1 is a generalization to $\mathbb{R}^n$ of the forward direction of Theorem 1.6.1. Theorem 4.6.3 is a stronger result for almost everywhere computable monotone functions on $\mathbb{R}^n$. Both of those results are effective versions of Mignot's theorem. Theorem 4.5.5 is an effective version of Aleksandrov's theorem.

Apart from effective versions of known a.e. theorems, we proved a couple of results of independent interest. Theorem 4.2.8 characterizes computable randomness on $\mathbb{R}^n$ in terms bounded Martin-Löf tests. It is a generalization to $\mathbb{R}^n$ of the known characterization by Merkle, Mihailović and Slaman [MMS06]. We use this characterization to prove Theorem 4.2.9, which shows that every computable monotone Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^n$ preserves computable non-randomness.

In Chapter 5 we prove an effective version of Brenier's theorem. This is a result of independent interest, however we are interested in Brenier's theorem mainly because it allows to generalize Zahorski's construction. This result will be used in Chapter 6.

In Chapter 6 we generalized to $\mathbb{R}^n$ the construction mentioned in Remark 1.5.1. Our construction is very natural and due to this naturalness it can be seen as *the* generalization of the idea of Zahorski. In order to achieve this, firstly we had to reinterpret Zahorski's construction in terms of transport maps. Then we used Brenier's theorem, an important result from the field of optimal transport, to make this reinterpreted idea work in $\mathbb{R}^n$. In order to effectivize this construction, we used our computable version of Brenier's theorem, which is an important result of independent interest. Using this effectivization, we proved Theorem 6.0.1 — which is the converse result both to our effective version of Aleksandrov's theorem and our effective version of Mignot's theorem. In one aspect our generalization of Zahorski's construction is not a perfect one. The same construction on the real line yields a monotone Lipschitz map. However, in higher dimensions, this Lipschitz continuity becomes Hölder continuity. Which means that as it is, this construction can not be used to obtain the converse direction for Rademacher's theorem. Using a combination of Theorem 6.0.1 and results from Chapter 4, we obtained several characterizations of computable randomness in $\mathbb{R}^n$. Apart from the already mentioned characterizations in terms of effective Aleksandrov's theorem and effective Mignot's theorem, we have shown:

1. Theorem 6.3.3, which is a characterization of computable randomness in terms of differentiability of computable absolutely continuous probability measures on $\mathbb{R}^n$. This result can be seen as an effective version of the Lebesgue differentiation theorem for functions that are densities of computable absolutely continuous measures on $[0,1]^n$.

2. We have proven the converse to our effective version of Sard's theorem. This yielded Theorem 6.3.5 which characterizes computable randomness in terms of critical values of computable monotone Lipschitz functions.

3. Finally, we have characterized computable randomness in terms of the Monge-Ampère equation. This is Theorem 6.3.6.

1.8 Structure of the thesis

The structure of the thesis is straightforward. In Chapter 2 we introduce the notation, basic notions and fundamental results used in the rest of the thesis.

Chapter 3 is devoted to results on the unit interval.

In Chapter 4 we prove most of our forward results on $\mathbb{R}^n$. Effective versions of Rademacher's theorem, Aleksandrov's theorem, Mignot's theorem and Sard's theorems are proven there.

In Chapter 5 we prove our computable version of Brenier's theorem.

In Chapter 6 are our converse results. There we show how to generalize Zahorski's construction.

Chapter 2

Preliminaries

The purpose of this chapter is to provide all necessary definitions, define non-standard notation used in this thesis and formulate basic results used later. The notation index is located at the end of the thesis in Section 6.3.4.

2.1 Computable analysis

Since we are interested in specific interactions between effective versions of results from analysis and algorithmic randomness, we need to define rigorously what "effective" in this context means. An area known as computable analysis provides the necessary concepts and tools. As its name suggests, this area studies various effective versions of notions from analysis. In the following subsections we provide the definitions and some of results used throughout this thesis. For more comprehensive introduction, please consult Weihrauch [Wei00], Pour-El and Richards [PER89] and Ko [Ko91].

2.1.1 Computable elements of $\mathbb{R}^n$

Definition 2.1.1. A sequence $(q_i)_{i\in\mathbb{N}}$ of elements of $\mathbb{Q}^n$ is said to be *a Cauchy name for* $x \in \mathbb{R}^n$ if $\lim q_i = x$ and $|x - q_i| \leq 2^{-i}$ for all i. By CN_x we denote the set of all Cauchy names for x.

We say $x \in \mathbb{R}^n$ is *computable* if there is a computable Cauchy name for x.

We say a sequence $(x_i)_{i\in\mathbb{N}}$ of elements of $\mathbb{R}^n$ is computable if there is a computable double sequence (of elements of $\mathbb{Q}^n$) $(q_{i,j})_{i,j\in\mathbb{N}}$ such that for every i, $(q_{i,j})_{j\in\mathbb{N}}$ is a Cauchy name for x_i.

Remark 2.1.2. The above notion of computability won't be affected if all rationals (in the definition) are required to be dyadic.

Notation 2.1.3. If $x \in \mathbb{R}^n$ is computable, we often assume that some of its computable Cauchy names (let's denote it by $(q_i)_{i\in\mathbb{N}}$) is fixed and then by $(x)_i$ we denote q_i, so that $|x-(x)_i| \leq 2^{-i}$ for all i.

The following fact is well-known (see [Wei00]):

Proposition 2.1.4. *There is no computable sequence of elements in $\mathbb{R}^n$ that enumerates all computable elements in $\mathbb{R}^n$.*

Several times we will use the following effective version of Baire's Category Theorem:

Theorem 2.1.5 (cf. Corollary 7 in [Bra01])**.** *Let $A \subseteq \mathbb{R}^n$ be a dense Π^0_2 set. There is a dense computable sequence of elements belonging to A.*

2.1.2 Computable functions on $\mathbb{R}^n$

Various notions of computability for functions from $\mathbb{N}$ to $\mathbb{N}$ are known to be equivalent. The situation is different for real-valued functions on $\mathbb{R}^n$: several of partly equivalent notions have been studied. We will rely on the so called *Grzegorczyk-Lacombe* notion of computability. This approach is known to be equivalent to the one used in [PER89] and the one used in [Ko91]. Our presentation of this approach is based on [Ko91] (see Section 2.5 there), as this formulation is easily relativizable.

Definition 2.1.6. Let $D \subseteq \mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *computable on D* if there is an oracle Turing machine M such that for all $x \in D$, all $\phi \in CN_x$ and all $i \in \mathbb{N}$, $M^\phi(i)$ halts and outputs $q \in \mathbb{Q}^m$ with dyadic components such that

$$|q - f(x)| \leq 2^{-i}.$$

Here we assume some computable enumeration of $\mathbb{Q}^n$ is fixed and we treat Cauchy names as functions from $\mathbb{N}$ to $\mathbb{N}$, so that they can be used as oracles.

We say $f : \mathbb{R}^n \to \mathbb{R}^m$ is *computable* if it is computable on $\mathbb{R}^n$. Similarly, we say $f : [0,1]^n \to \mathbb{R}^m$ is computable if f is computable on $[0,1]^n$.

Notation 2.1.7 (Computed values at a given step)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be computable and let $x \in \mathbb{R}^n$. We often assume that a Turing machine M computing f (in the sense of Definition 2.1.6) is fixed. Likewise, we often assume that one of Cauchy names for x, ϕ, is also fixed. In that case by $(f(x))_t$ we denote the dyadic rational computed by $M^\phi(t)$, so that

$$|f(x) - (f(x))_t| \leq 2^{-t}.$$

In Chapter 3 we will prove a number of one-dimensional results. In order to use them later (particularly in Chapter 4), those results would have to formulated using the notion of relativized computability (defined below).

Definition 2.1.8. Let $D \subseteq 2^\omega \times \mathbb{R}^n$. We say a function $f : 2^\omega \times \mathbb{R}^n \to \mathbb{R}^m$ is computable on D if there is a 2−oracle Turing machine M such that for all $(A, x) \in D$, all $\phi \in CN_x$ and all $i \in N$, $M^{A,\phi}(i)$ halts and outputs $d \in \mathbb{Q}^m$ with dyadic components such that

$$|q - f(A, x)| \leq 2^{-i}.$$

We say a function $f : D \to \mathbb{R}^m$ is computable if there is a function $g : 2^\omega \times \mathbb{R}^n \to \mathbb{R}^m$ that is computable on D and $f = g$ on D.

Theorem 2.1.9 (see Corollary 9.4 in [Bra08])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function. It is computable if and only if its graph, Γf, is a Π^0_1 set.*

2.2 Computable metric spaces

Computability on $\mathbb{R}^n$ can be extended to a large class of metric spaces, called *computable metric spaces*. This concept has been introduced in Weihrauch [Wei93] and studied in Brattka and Presser [BP03] among others.

Definition 2.2.1 (Computable metric spaces)**.** A *computable metric space* is a triple $(X, d, (\alpha_i)_{i\in\mathbb{N}})$, where:

- (X, d) is a complete separable metric space,
- $\{\alpha_i : i \in \mathbb{N}\}$ is dense in X, and
- $d(\alpha_i, \alpha_j)$ is computable uniformly in i, j.

The elements of $(\alpha_i)_{i\in\mathbb{N}}$ are called the *basic points*. Without loss of generality, $i \mapsto \alpha_i$ can be assumed to be injective.
For $x \in X$ and $r > 0$, let $B(x, r)$ denote the metric ball $\{y \in X : d(x, y) < r\}$. We say $B(x, r)$ is a *basic ball*, if x is a basic point and r is a positive rational.

Definition 2.2.2 (Cauchy names, computable points, computable functions)**.** Let $(X, d_X, (\alpha_i)_{i\in\mathbb{N}})$ be a computable metric space. A *Cauchy name* for $x \in X$ is a sequence $(x_i)_{i\in\mathbb{N}}$ of basic points such that $d_X(x_i, x) \leq 2^{-i}$ for all i. Again, we can identify sequences of basic points with functions from $\mathbb{N}$ to $\mathbb{N}$.
Let $(Y, d_Y, (\beta_i)_{i\in\mathbb{N}})$ be a computable metric space. A function $f : X \to Y$ is *computable* if there exists an oracle Turing machine M such that whenever $A_x : \mathbb{N} \to \mathbb{N}$ is a Cauchy name for $x \in X$, then M^{A_x} computes a Cauchy name for $f(x)$. That is, for every $j \in \mathbb{N}$, $M^{A_x}(j)$ halts and outputs $a \in \mathbb{N}$ with

$$d_Y(\beta_a, f(x)) \leq 2^{-j}.$$

Proposition 2.2.3. *Let $(X, d, (\alpha_i)_{i\in\mathbb{N}})$ be a computable metric space. The distance $d : X \times X \to \mathbb{R}$ is a computable function.*

2.3 Computable measures

We follow Gács [Gác05] and Hoyrup and Rojas [HR09] in defining the notion of computability of measures.

Notation 2.3.1. Given some abstract measure space X, by $P(X)$ we denote the set of all probability measures on X.

Since we work mostly with Polish spaces X equipped with their Borel $\sigma-$algebras, $P(X)$ usually denotes the set of Borel probability measures on X.

Given a Polish metric space (X, d), the set $P(X)$ of Borel probability measures over X endowed with the weak topology is a Polish space.

Definition 2.3.2 (Prokhorov metric)**.** The *Prokhorov metric* π on $P(X)$ is defined by:

$$\pi(\mu, \nu) = \inf\{\epsilon > 0 \mid \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ for every Borel set } A\},$$

where $A^\epsilon = \{x \ : \ d(x, A) \leq \epsilon\}$.

Suppose $(X, d, (\alpha_i)_{i\in\mathbb{N}})$ is a computable metric space. Let $(\delta_i)_{i\in\mathbb{N}}$ be an effective enumeration of those elements of $P(X)$ which are concentrated on finite subsets of basic points and assign rational values to them. Then $(P(X), \pi, (\delta_i)_{i\in\mathbb{N}})$ is a computable metric space compatible with the weak topology on $P(X)$. Following Gács [Gác05] and Hoyrup and Rojas [HR09], we define computable measures as computable elements of $(P(X), \pi, (\delta_i)_{i\in\mathbb{N}})$.

Recall that a real number r is said to be *left c.e.* if there is a nondecreasing computable sequence of rationals $(q_i)_{i\in\mathbb{N}}$ such that $r = \lim_i q_i$.

Proposition 2.3.3 (see Theorem 4.2.1 in [HR09])**.** *Let $\mu \in P(X)$. The following are equivalent:*

1. *μ is computable, and*
2. *$\mu(B_i)$ is left c.e. uniformly in i, where $(B_i)_{i\in\mathbb{N}}$ is a computable numbering of basic open balls.*

Remark 2.3.4. The original Theorem 4.2.1 in [HR09] considered finite union of open balls, rather then single balls. However, this is easily seen to be equivalent to our formulation.

The characterization of computable probability measures in Proposition 2.3.3 is somewhat counterintuitive, as $\mu(B_i)$ is required to be left c.e. rather than computable. In general, this can not be strengthened - $\mu(B_i)$ is not necessarily computable uniformly in i. However, in some important cases it is, indeed, computable. For example, if μ is absolutely continuous, $\mu(B_i)$ is computable uniformly in i. Moreover, for a given computable μ, it is always possible to find a computable sequence of open balls $(\hat{B}_i)_{i\in\mathbb{N}}$ so that $\mu(\hat{B}_i)$ is computable uniformly in i and $(\hat{B}_i)_{i\in\mathbb{N}}$ forms a subbasis (of the underlying space X). See the proof of Corollary 5.2.1 in [HR09].

Theorem 2.3.5 ([see Corollary 4.3.2 in [HR09])**.** *Let $\mu \in P(X)$. Let $(f_i)_{i\in\mathbb{N}}$ be a sequence of uniformly computable functions from X to $\mathbb{R}$, i.e. such that the function $(i, x) \mapsto f_i(x)$ is computable. If moreover f_i has a bound M_i computable uniformly in i, then the function $(\mu, i) \mapsto \int f_i \, d\mu$ is computable.*

2.4 Dyadic rationals and dyadic cubes in $\mathbb{R}^n$

Notation 2.4.1. For every $i \in \mathbb{N}^+$, let $\mathbb{D}_i^n$ denote the set of points in $\mathbb{R}^n$ with all coordinates of the form $k2^{-i}$ for some integer k. Let $S \subseteq \mathbb{R}$. By $\mathbb{D}_i^{n,S}$ we denote the set $\mathbb{D}_i^n \cap S^n$. Define $\mathbb{D}_*^n = \cup_i \mathbb{D}_i^n$ and $\mathbb{D}_*^{n,S} = \cup_i \mathbb{D}_i^{n,S}$.

Let $\mathcal{D}^n$ denote the collection of half-open basic dyadic cubes in $\mathbb{R}^n$. That is

$$\mathcal{D}^n = \left\{2^{-k}([m_1, m_1+1) \times \cdots \times [m_n, m_n+1)) : \ k \in \mathbb{Z}, \ m_1, \ldots, m_n \in \mathbb{Z}\right\}.$$

If $D \subset \mathbb{R}^n$ is a cube, by $l(D)$ we denote the side length of D.

For $k \in \mathbb{Z}$, let $\mathcal{D}^n(k)$ denote the collection of basic (half-open) dyadic cubes in $\mathbb{R}^n$ with its side length equal to 2^{-k}. For $x \in \mathbb{R}^n$ and $i \in \mathbb{N}$, define $\mathcal{D}^n(i, x)$ to be the unique element of $\mathcal{D}^n(i)$ containing x.

The following proposition is known as the "1/3−shift trick" in $\mathbb{R}^n$.

Proposition 2.4.2 (cf. Theorem 3.8 in [Tap12])**.** *Consider $\mathbb{R}^n$ equipped with the usual Euclidean metric. For any ball $B = B(x, r) \subset \mathbb{R}^n$, there exists $k \in \mathbb{Z}$, $Q \in \mathcal{D}^n(k)$ and $t \in \{0, 1/3, 2/3\}^n$ such that $B \subset (Q + t)$ and $6r < 2^{-k} \leq 12r$.*

We will also need the following version of Whitney's covering lemma (see Section 1.1 in [Rog04]):

Lemma 2.4.3 (Dyadic covering)**.** *If $A \subseteq \mathbb{R}^n$ is open, then there is a countable collection of (basic) closed dyadic cubes $(Q_i)_{i \in \mathbb{N}}$ with disjoint interiors such that $A = \cup_i Q_i$ and for all j,*

$$1 \leq \frac{d(Q_j, \partial A)}{\sqrt{n} l(Q_j)} \leq 4.$$

Recall that the distance, $d(A, B)$, between two subsets A and B, is defined as

$$d(A, B) = \inf_{x \in A, y \in B} |x - y|.$$

The above lemma says that every open set A is a union of closed basic dyadic cubes Q with $l(Q)$ being proportional to the distance between Q and the boundary of A. In particular, every ball in $\mathbb{R}^n$ of radius r contains a basic dyadic cube whose side length is proportional to r.

Remark 2.4.4. Additional notation related to dyadic cubes in $\mathbb{R}^n$ will be introduced in the next section.

2.5 Computable randomness

Computable randomness is one of the more natural notions of algorithmic randomness. It is usually defined in terms of success sets of effective betting strategies: a sequence is said to be computably random if no computable betting strategy can make an unbounded profit while betting on this sequence. Betting strategies are usually formalized as *martingales* (see Chapter 7 in the book by Nies [Nie09]).

Definition 2.5.1. We say a function $B : 2^{<\omega} \to \mathbb{R}_+ \cup \{0\}$ is a *martingale* if the following condition holds for all $\sigma \in 2^{<\omega}$:

$$2B(\sigma) = B(\sigma 0) + B(\sigma 1).$$

We say B *succeeds* on $Z \in 2^\omega$ if $\limsup_n B(Z\restriction_n) = \infty$.
We say B *diverges* on Z if either B succeeds on Z or

$$\liminf_n B(Z\restriction_n) < \limsup_n B(Z\restriction_n).$$

Definition 2.5.2. A martingale B is called *computable* if $B(\sigma)$ is a computable real number uniformly in σ.
We say $Z \in 2^\omega$ is *computably random* if no computable martingale succeeds on Z.

Remark 2.5.3. It is known that in the above definition "succeeds" can be replaced with "diverges" without changing the notion. That is, Z is computably random iff no computable martingale diverges on Z.

Remark 2.5.4. If $Z \in 2^\omega$ is not computably random, there exists a computable martingale B succeeding on Z. We may assume B has the following two additional properties:

- B only takes positive rational values (see Proposition 7.3.8 in [Nie09]) and
- we may also assume that B has the *savings property*:

  $$B(\sigma\tau) \geq B(\sigma) - 1, \text{ for all } \sigma, \tau \in 2^{<\omega}.$$

  For example, see the proof of Proposition 6.3.8 in [DH10].

A martingale B can be seen as a betting strategy for betting on bits of a string in the ascending order. $B(\sigma)$ can be interpreted as B's capital after betting on all bits of σ. Suppose, after betting on all bits of σ, B bets an amount α, with $0 \leq \alpha \leq B(\sigma)$, that the next bit is 0. If B is right, it wins α and hence $B(\sigma 0) = B(\sigma) + \alpha$, while $B(\sigma 1) = B(\sigma) - \alpha$.

Both the following definition and the theorem are due to Miyabe and Rute [MR13].

Definition 2.5.5. A total computable function $M : 2^\omega \times 2^{<\omega} \to \mathbb{R}$ is an *oracle martingale* if $M^Z(\cdot) = M(Z, \cdot)$ is a martingale for every $Z \in 2^\omega$.

We say A is *computably random uniformly relative to* B if there is an oracle martingale M such that $M(B, \cdot)$ succeeds on A.

Rute and Miyabe called oracle martingales *uniform computable martingales*. However, we will use this concept in the context of relativization, hence our preference for a different name.

The following form of van Lambalgen's theorem for computable randomness will play an important role in Chapter 4.

Theorem 2.5.6 (Theorem 1.3 in [MR13])**.** *$A \oplus B$ is computably random if and only if A is computably random uniformly relative to B and B is computably random uniformly relative to A.*

2.5.1 Computable randomness in $\mathbb{R}^n$

Since we work mostly in $\mathbb{R}^n$, we need a more general definition of computable randomness. One way of defining computable randomness in $\mathbb{R}^n$ is to define the concept of binary expansions (for elements in $\mathbb{R}^n$) and to declare those elements as computably random, whose binary expansions are computably random. Since computable randomness is known to be invariant under computable permutations, this definition does not depend on a particular choice of computing binary expansions (with some sensible restrictions). An approach to define computable randomness in more general spaces has been described by Rute [Rut16]. Below we adopt this approach for our purposes.

Definition 2.5.7. We say (X, μ) is a *computable probability space* if X is a computable Polish space and μ is a computable probability measure on X.
A pair $U, V \subseteq X$ is *μ-a.e. decidable pair* if

1. U and V are Σ^0_1 sets,
2. $V \cap U = \emptyset$, and
3. $\mu(U \cup V) = 1$.

A set A is a *μ-a.e. decidable set* if there is a μ-a.e. decidable pair U, V such that $U \subseteq A \subseteq X \setminus V$. The code for the μ-a.e. decidable set A is the pair of codes for the Σ^0_1 sets U and V.
Let $\mathcal{A} = (A_i)_{i\in\mathbb{N}}$ be a computable sequence of a.e. decidable sets. Let $\mathcal{B}$ be the closure of $\mathcal{A}$ under finite Boolean combinations. We say $\mathcal{A}$ is an *(a.e. decidable) generator of* (X, μ) if given a Σ^0_1 set $U \subseteq X$ one can find (effectively from the code of U) a c.e. family $(B_i)_{i\in\mathbb{N}}$ of sets in $\mathcal{B}$ such that $\mu(U \setminus \cup_j B_j) = 0$.

Definition 2.5.8. Let $\mathcal{A} = (A_i)_{i\in\mathbb{N}}$ be an a.e. decidable generator of (X, μ). Recall each A_i is coded by an a.e. decidable pair (A^0_i, A^1_i) where $A^0_i \subseteq A_i \subseteq X \setminus A^1_i$. For $\sigma \in 2^{<\omega}$ of length $s > 1$ define

$$[\sigma]_{\mathcal{A}} = A_0^{\sigma(0)} \cap A_1^{\sigma(1)} \cap \cdots \cap A_{s-1}^{\sigma(s-1)}.$$

When $\sigma = \epsilon$, we let $[\sigma]_{\mathcal{A}} = X$.
When possible, define $x \upharpoonright_{\mathcal{A}} n$ as the unique σ of length n such that $x \in [\sigma]_{\mathcal{A}}$. Also when possible, define the $\mathcal{A}-$name of x as $name_{\mathcal{A}}(x) = \lim_{n\to\infty} x \upharpoonright_{\mathcal{A}} n$. A point without an $\mathcal{A}-$name will be called an *unrepresented point.* Each $[\sigma]_{\mathcal{A}}$ will be called a cell, and the collection of $\{[\sigma]_{\mathcal{A}}\}_{\sigma\in 2^{<\omega}}$ will be called an *(a.e. decidable) cell decomposition of* (X, μ).

Cell decompositions allow one to translate between Cantor space and other spaces. With our interest in $\mathbb{R}^n$ in mind, we will fix one particular cell decomposition (for every n) and we will use it throughout this thesis. Note that $\mathcal{A}_{2^\omega} = ([\sigma])_{\sigma\in 2^{<\omega}}$ is a cell decomposition of $(2^\omega, \lambda)$. We call this *the natural cell decomposition.*

Definition 2.5.9. Define $X = \{Z \in 2^\omega : Z \text{ is co-infinite and infinite}\}$. Let $F : X \to [0,1) \setminus \mathbb{D}^1_*$ be defined by

$$F(Z) = 0.Z = \sum_{i \in Z} 2^{-i-1}.$$

Fix $n \in \mathbb{N}$. For $i \in \mathbb{N}$ with $0 \le i \le n-1$ and $Z \in 2^\omega$, define

$$p^n_i(Z) = \{Z(kn+i) \ : \ k \in \mathbb{N}\}.$$

Define $F_n : \{Z \in 2^\omega : p^n_i(Z) \in X \text{ for all } 0 \le i \le n-1\} \to ([0,1) \setminus \mathbb{D}^1_*)^n$ by

$$F_n(Z) = 0.Z = (F(p^n_0(Z)), \ldots, F(p^n_{n-1}(Z))).$$

It is known that F is a bijection. F^{-1} maps real numbers from $[0,1)$ that are not dyadic rationals to their binary expansions (for example, see 1.8.10-1.8.13 in [Nie09]). Analogously, F_n^{-1} maps elements from $[0,1)^n \setminus \mathbb{D}^n_*$ to their "binary expansions", where the binary expansion of $x = (x_1, \ldots, x_n)$ is the sequence resulting from bit-wise interleaving of binary expansions of $x_1, \ldots, x_n$.

Notation 2.5.10 (Correspondence between clopens in the Cantor space and basic dyadic cubes in $[0,1]^n$)**.** In the terminology of Rute, F is an *a.e. computable isomorphism* between $(2^\omega, \lambda)$ and $([0,1], \lambda_1)$. That is, both F and its inverse are a.e. computable and measure preserving. Similarly, F_n is an a.e. computable isomorphism between $(2^\omega, \lambda)$ and $([0,1]^n, \lambda_n)$. By Proposition 7.7 in [Rut16], $(F_n)^{-1}$ and the natural cell decomposition of $(2^\omega, \lambda)$ induce a cell decomposition on $([0,1]^n, \lambda_n)$. Let us denote this cell decomposition by $\mathcal{A}_n$. With respect to this cell decomposition, $[\sigma]_{\mathcal{A}_n}$ is the set of those $x \in [0,1]^n \setminus \mathbb{D}^n_*$, whose binary expansions $F_n^{-1}(x)$ extend σ. That is, $[\sigma]_{\mathcal{A}_n}$ is always a finite union of open basic dyadic cubes in $[0,1]^n$. From now on, when it is clear from the context that we are working in $[0,1]^n$, we will write $[\sigma]$ instead of $[\sigma]_{\mathcal{A}_n}$. Please note that in this notation $[\sigma]$ always denotes a finite union of *open* dyadic cubes. We will be writing $\overline{[\sigma]}$ to denote the closure of $[\sigma]$.

Definition 2.5.11. We say $z \in [0,1]^n$ is computably random if its binary expansion, that is $Z \in 2^\omega$ with $Z = F_n^{-1}(z)$, is defined and it is computably random.

Let $z \in \mathbb{R}^n$. Let $p \in \mathbb{Z}^n$ be such that $z - p \in [0,1]^n$. We say z is computably random if $z - p$ is computably random.

The above definition seems to depend on the choice of F_n. However, by Theorem 5.7 in [Rut16], every choice of F_n, provided it is an a.e. computable isomorphism, results in the same notion of computable randomness.

Notation 2.5.12. Let $\sigma \in 2^{<\omega}$ and let $p \in \mathbb{Z}^n$. When it won't cause ambiguities, as a notational convenience, we will write $[\sigma]_p$ instead of $[\sigma] + p$.

Definition 2.5.13. A *Martin-Löf test* is a uniformly computable sequence $(U_i)_{i \in \mathbb{N}}$ of Σ_1^0 subsets of $[0,1]^n$ such that $\lambda(U_i) \leq 2^{-i}$ for all i. We say $(U_i)_{i \in \mathbb{N}}$ *covers* $z \in [0,1]^n$ if $z \in \bigcap_i U_i$.
We say a Martin-Löf test $(U_i)_{i \in \mathbb{N}}$ is *bounded* if there is a computable pre-measure $\nu : 2^{<\omega} \to [0, \infty)$ satisfying

$$\lambda(U_i \cap [\sigma]) \leq 2^{-i} \nu(\sigma)$$

for all $i \in \mathbb{N}$ and $\sigma \in 2^{<\omega}$.

We will need the following characterization of computable randomness in the unit cube due to Rute:

Proposition 2.5.14 (cf. Theorem 5.3 in [Rut16])**.** *Let $z \in [0,1]^n \setminus \mathbb{D}_*^n$. The following two are equivalent:*

1. *z is not computably random, and*
2. *there is a bounded Martin-Löf test $(U_i)_{i \in \mathbb{N}}$ that covers z.*

It is worth noting that the above characterization is a straightforward generalization of a result by Merkle, Mihailović and Slaman from [MMS06].

2.6 Differentiability of real-valued functions on $\mathbb{R}^n$

2.6.1 Functions on the real line

Definition 2.6.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We say f is *differentiable* at $a \in \mathbb{R}^n$ if the following limit exists

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

In that case, the value of that limit is denoted by $Df(a)$. We call $Df(a)$ the *derivative* of f at a.

We will be using several closely related notions. Below we define them and introduce the relevant notation that will be used in this thesis.

Definition 2.6.2. Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let $x \in \mathbb{R}$.
Define

$$\overline{D}f(x) = \limsup_{h \to 0} \frac{f(x+h) - f(x)}{h},$$
$$\underline{D}f(x) = \liminf_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

We call $\overline{D}f(x)$ (resp. $\underline{D}f(x)$) *lower derivative* (resp. *upper derivative*) of f at x. Similarly we define *lower and upper dyadic derivatives of f at x*:

$$\overline{D}_2 f(x) = \limsup_{h \to 0, h \in \mathbb{D}^1_*} \frac{f(x+h) - f(x)}{h},$$
$$\underline{D}_2 f(x) = \liminf_{h \to 0, h \in \mathbb{D}^1_*} \frac{f(x+h) - f(x)}{h}.$$

If the following limits exist, their respective values are called *right-sided and left-sided derivatives of f at x:*

$$D_+ f(x) = \lim_{h \to 0, h > 0} \frac{f(x+h) - f(x)}{h},$$
$$D_- f(x) = \lim_{h \to 0, h < 0} \frac{f(x+h) - f(x)}{h}.$$

Collectively, $D_+ f(x)$ and $D_- f(x)$ are known as *one-sided derivatives of f at x.*

2.6.2 Functions on $\mathbb{R}^n$, partial and directional derivatives

Definition 2.6.3. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *differentiable* at $a \in \mathbb{R}^n$ if there is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - T(h)|}{|h|} = 0.$$

In that case, the linear transformation T is unique and it is denoted by $Df(a)$. We call this transformation the *derivative* of f at a.

Definition 2.6.4 (Partial derivatives, gradient)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let $a \in \mathbb{R}^n$. The limit

$$\lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}$$

if it exists, is denoted $D_i f(a)$, and called *the ith partial derivative* of f at a.

When all partial derivatives of f at a exist, define *the gradient of f at a*, $\nabla f(a) \in \mathbb{R}^n$, by $\nabla f(a) = (D_1 f(a), \ldots, D_n f(a))$.

Definition 2.6.5 (Directional derivatives)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, let $a \in \mathbb{R}^n$ and let $v \in S^{n-1}$. The limit

$$\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}$$

if it exists, is denoted $Df(a; v)$, and called *the directional derivative* of f at a, in the direction of v.
The limit

$$\lim_{t \to +0} \frac{f(a + tv) - f(a)}{t}$$

if it exists, is denoted $D_+ f(a; v)$, and called *the one-sided directional derivative* of f at a, in the direction of v.

2.6.3 Differentiability of partial functions

In order to reason rigorously about twice differentiability of convex functions, we need to be able to define the concept of differentiability for functions that are defined almost everywhere. A minor modification of Definition 2.6.3 suffices:

Definition 2.6.6. Let $A \subseteq \mathbb{R}^n$ be of full measure. A function $f : A \to \mathbb{R}^m$ is *differentiable* at $a \in A$ if there is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0, a+h \in A} \frac{|f(a + h) - f(a) - T(h)|}{|h|} = 0.$$

In that case, the linear transformation T is unique and it is denoted by $Df(a)$. We call this transformation the *derivative* of f at a.

2.6.4 Derivatives and antiderivatives

Since we are interested in functions that are not necessarily differentiable on the whole domain but are differentiable almost everywhere, we use the terms *derivative* and *antiderivative* in the following sense:

Definition 2.6.7. We say g is a derivative of f if $g(x) = Df(x)$ on all x where $Df(x)$.

We say f is an antiderivative of g if g is a derivative of f.

2.6.5 Non-differentiability sets of functions

Characterizing points of non-differentiability of effective functions is one of the main themes of this thesis. This warrants a separate notation.

Notation 2.6.8. Let $A \subseteq \mathbb{R}^n$. Let $f : A \to \mathbb{R}^m$ be a function. Define

$$N_f = \{z \in A \mid f \text{ is not differentiable at } z\}.$$

2.6.6 Approximate continuity, essential values and Lebesgue points

To control (non)differentiability of functions, we will be using several very closely related notions.

Definition 2.6.9. Let $f : \mathbb{R}^n \to \mathbb{R}^m$. We say $l \in \mathbb{R}^m$ is the *approximate limit of f as $y \to x$*, written

$$\operatorname{ap}\lim_{y \to x} f(y) = l,$$

if for each $\epsilon > 0$,

$$\lim_{r \to 0} \frac{\lambda\left(B(x,r) \cap \{y \ : \ |f(y) - l| \geq \epsilon\}\right)}{\lambda\left(B(x,r)\right)} = 0.$$

We say f is *approximately continuous at x* if

$$\operatorname{ap}\lim_{y \to x} f(y) = f(x).$$

Definition 2.6.10. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally integrable function. Let $x_0 \in \mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. We say f has an *essential value α at x_0* if

$$\lim_{r \to 0} \frac{\int_{B(x,r)} |f(t) - \alpha| \, \mathrm{d}t}{\lambda\left(B(x,r)\right)} = 0.$$

Definition 2.6.11. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally integrable function. We say $x \in \mathbb{R}^n$ is a *Lebesgue points of f* if $f(x)$ is an essential value of f at x.

Remark 2.6.12. It is known that if x is a Lebesgue point of a locally integrable function $f : \mathbb{R}^n \to \mathbb{R}$, then x is a point of approximate continuity. See Section 1.7 in [EG92]. Conversely, suppose $f : \mathbb{R}^n \to \mathbb{R}$ is locally integrable and bounded (that is $|f(x)| \leq M$ for some $M \in \mathbb{R}$). If f has an approximate limit α at x, then f has an essential value α at x (see Remark 6.7 in [Vuo82]).

This means that for locally integrable bounded functions, points of approximate continuity coincide with Lebesgue points.

On the real line the following result is known.

Theorem 2.6.13 (see Theorem 5.5(a) in [Bru78])**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and measurable and fix $a \in \mathbb{R}$. Define*

$$g(x) = \int_a^x f(t)\,\mathrm{d}t. \tag{2.1}$$

Then for any point $x_0 \in \mathbb{R}$ at which f is approximately continuous, g is differentiable at x_0 and $Dg(x_0) = f(x_0)$.

For semi-continuous functions, the converse also holds:

Theorem 2.6.14 (Theorem 5.8 in [Bru78])**.** *Let f be bounded in a neighborhood I of x_0 and lower (or upper) semi-continuous at x_0. Then f is approximately continuous at x_0 if and only if f is the derivative of its integral (that is g defined in* (2.1)*) at x_0.*

2.7 Lipschitz, convex and monotone functions

This thesis is concerned mostly with three classes of functions on $\mathbb{R}^n$: Lipschitz, convex and monotone ones. Here we provide relevant definitions, set-up the required notation and list some of the properties used later. This is by no means a comprehensive introduction, rather a small subset of notions and results needed in this thesis. For more information on Lipschitz functions, please consult lecture notes by Heinonen [Hei05] and the book by Lindenstrauss and Preiss [LPT12]. Regarding convex and monotone functions, we relied on the books by Borwein and Vanderwerff [BV10] and by Phelps [Phe93], and a great survey paper on monotone functions by Alberti and Ambrosio [AA99]. Finally, since the three classes feature prominently in the area known as variational analysis, the following two books are worth mentioning: the one by Rockafellar and Wets [RW97] and the one by Ekeland and Témam [ET99].

Definition 2.7.1 (Lipschitz functions)**.** A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *Lipschitz* if there exists $L \in \mathbb{R}^+$ such that

$$|f(x) - f(y)| \leq L|x - y| \text{ for all } x, y \in \mathbb{R}^n.$$

The least such L is called *the Lipschitz constant* for f. We denote it by $\mathbf{Lip}(f)$. We say f is $K-$Lipschitz if $\mathbf{Lip}(f) \leq K$.

The following result shows that Lipschitz functions (from $\mathbb{R}^n$ to $\mathbb{R}^n$) do not stretch sets too much:

Theorem 2.7.2 (see Section 2.4.1 in [EG92])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz function. Let $A \subset \mathbb{R}^n$ be Borel. Then*

$$\lambda\left(f(A)\right) \leq \left(\mathbf{Lip}(f)\right)^n \lambda\left(A\right).$$

Definition 2.7.3 (Convexity)**.** A set $C \subseteq \mathbb{R}^n$ is said to be *convex* if $tx+(1-t)y \in C$ whenever $x, y \in C$ and $0 \leq t \leq 1$.
Let $C \subseteq \mathbb{R}^n$ be convex. A function $f : C \to [-\infty, +\infty]$ is *convex* if the following condition holds for all $x_0, x_1 \in C$ and all $t \in [0, 1]$:

$$f((1-t)x_0 + tx_1) \leq (1-t)f(x_0) + tf(x_1),$$

whenever the right-hand side is defined.

The set $\{x \in \mathbb{R}^n \mid f(u) < +\infty\}$ is called *the effective domain of f*.

We say that a function $f : C \to [-\infty, +\infty]$ is *proper* if it is nowhere equal to $-\infty$ and is not identically equal to $+\infty$.

It is clear from the above definition, that convex functions are not necessarily continuous. However, a convex function on $\mathbb{R}^n$ is continuous on the interior of its effective domain.

The following concept is useful in the context of convex functions:

Definition 2.7.4. The *epigraph* of a function $f : C \to [-\infty, +\infty]$ is defined by

$$\text{epi}f = \{(x, t) \in C \times \mathbb{R} \mid f(x) \leq t\}.$$

A function is convex if and only if its epigraph is convex.

Up until Chapter 6, we will only deal with continuous convex functions. In Chapter 6 we will have to deal, at least in the proofs, with discontinuous convex functions. Most of them will be proper *lower semi-continuous*:

Definition 2.7.5. A function $f : \mathbb{R}^n \to [-\infty, +\infty]$ is said to be *lower semi-continuous* if it satisfies the following condition:

$$\forall \hat{x} \in \mathbb{R}^n, \quad \lim_{x \to \hat{x}} f(x) \geq f(\hat{x}). \tag{2.2}$$

2.7.1 Properties of continuous convex functions on $\mathbb{R}$

Theorem 2.7.6 (see Theorem 2.1.2 in [BV10])**.** *Let $A \subset \mathbb{R}$ be an open interval and suppose $f : A \to \mathbb{R}$ is convex. Then*

1. *$D_+ f(x)$ and $D_- f(x)$ exist and are finite at each $x \in A$,*
2. *$D_+ f(x)$ and $D_- f(x)$ are non-decreasing functions on A,*
3. *$D_+ f(x) \leq D_- f(y) \leq D_+ f(y)$ for all $x < y$ with $x, y \in A$,*
4. *f is locally Lipschitz on A, in particular if $[a, b] \subseteq A$ and*

$$M = \max\{|D_+ f(a)|, |D_- f(b)|\},$$

then

$$|f(x) - f(y)| \leq M|x - y| \text{ for all } x, y \in [a, b].$$

Sets of points of non-differentiability of real valued convex functions on the real line have a simple characterization: they are precisely the countable subsets of the real line.

Proposition 2.7.7 (see Theorem 2.1.2 and 2.2.15 in [BV10])**.** *$N \subseteq \mathbb{R}$ is countable if and only if there is a convex function $f : \mathbb{R} \to \mathbb{R}$ such that $N_f = N$.*

We will discuss other properties of convex functions later, after introducing monotone functions.

2.7.2 Monotone functions

Monotone operators were first introduced in Minty [Min60] and Zarantonello [Zar60]. This notion can be seen both as a nonlinear generalization of linear endomorphisms with positive semidefinite matrices, and a multidimensional generalization of nondecreasing functions of a real variable. We are mainly interested in the latter. Monotone operators are often defined as set-valued functions from a Banach space to its dual, however we are only interested in (set-valued) functions from $\mathbb{R}^n$ to $\mathbb{R}^n$.

A set-valued function $u : \mathbb{R}^n \to \mathbb{R}^n$ takes every point $x \in \mathbb{R}^n$ to some set $u(x) \subseteq \mathbb{R}^n$. When no ambiguities may arise, we call these set-valued functions simply functions.

Notation 2.7.8. Let set-valued functions u, v, real numbers a, b and a set $B \subseteq \mathbb{R}^n$ be given. For all $x \in \mathbb{R}^n$ we set

$$\begin{aligned}
\text{domain of } u,\ \operatorname{Dm} u &= \{x : u(x) \neq \emptyset\},\\
\text{image of } u,\ \operatorname{Im} u &= \{y : \exists x, y \in u(x)\},\\
\text{graph of } u,\ \Gamma u &= \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in u(x)\},\\
\text{(generalized) inverse of } u,\ [u^{-1}](x) &= \{y : x \in u(y)\},\\
[au + bv](x) &= \{ay + by' : y \in u(x), y' \in v(x)\},\\
u(B) &= \{y : \exists x \in B, y \in u(x)\}.
\end{aligned}$$

Definition 2.7.9 (Monotone functions)**.** We say a (possibly partial) set-valued function $u : \mathbb{R}^n \to \mathbb{R}^n$ is *monotone* if

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0 \text{ for all } x_1, x_2 \in \mathbb{R}^n \text{ and all } y_1 \in u(x_1), y_2 \in u(x_2).$$

Proposition 2.7.10 (Basic properties)**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a monotone (set-valued) function. Then*

1. *T^{-1} is a monotone function,*
2. *αT is monotone for any $\alpha > 0$,*
3. *if $T' : \mathbb{R}^n \to \mathbb{R}^n$ is monotone, then $T + T'$ is monotone.*

The following sub-class of monotone functions is of particular importance:

Definition 2.7.11 (Maximal monotone functions)**.** A set-valued monotone function u is said to be *maximal* if its graph is not properly included in the graph of another monotone function, that is if the following implication holds:

$$\Gamma(u) \subseteq \Gamma(v) \wedge v \text{ is monotone } \implies v = u.$$

Definition 2.7.12 (Maximal monotone extensions)**.** Let $u : \mathbb{R}^n \to \mathbb{R}^n$ and $\overline{u} : \mathbb{R}^n \to \mathbb{R}^n$ be monotone functions. We say $\overline{u}$ *is an extension of* u if $\Gamma(u) \subseteq \Gamma(\overline{u})$. Moreover, if $\overline{u}$ is maximal, we say $\overline{u}$ *is a maximal (monotone) extension of* u.

Remark 2.7.13. Clearly for every monotone function u, there exists a maximal extension. Furthermore, if the domain of u is dense in $\mathbb{R}^n$, then the maximal extension is unique. We will discuss this in more detail in Section 4.6. This means that in the majority of cases discussed in this thesis, we may and we will write about *the* maximal extension.

The usual notion of continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ can be extended to set-valued functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ in such a way that both notions coincide for single-valued functions. We are not particularly interested in this extended notion of continuity (a rigorous exposition can be found in Section 5.B in [RW97]). However, we need to mention a particular relationship between single-valuedness and continuity in the case of maximal monotone functions.

Proposition 2.7.14 (Continuity of maximal monotone functions). *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a maximal monotone function. Then:*

1. *T is single-valued almost everywhere on $int(Dm\ T)$.*
2. *T is continuous at a point $x \in Dm\ T$ if and only if T is single-valued at x, in which case necessarily $x \in int(Dm\ T)$.*

Proposition 2.7.15. *Every continuous single-valued monotone function with domain $\mathbb{R}^n$ is maximal.*

As a consequence, all computable monotone functions are maximal.

2.7.3 Subdifferentials of convex functions

Derivatives of convex functions are closely related to monotone set-valued functions. On the real line, a function is monotone if and only if it coincides (almost everywhere) with a derivative of a convex function. In the case of functions on $\mathbb{R}^n$, the situation is somewhat more nuanced.

The following notion that generalizes the notion of a derivative will be used frequently in this thesis:

Definition 2.7.16. Let Ω be an open subset of $\mathbb{R}^n$ and let $u : \Omega \to (-\infty, +\infty]$ be a proper function. The *subdifferential* of u, is the set-valued function $\partial u : \Omega \to \mathbb{R}^n$ defined by

$$\partial u(x_0) = \bigcap_{x \in \Omega} \{p \; : \; u(x) \geq u(x_0) + \langle p, (x - x_0) \rangle\}$$

for all x_0 where $u(x_0)$ is finite. If $u(x_0) = +\infty$, then $\partial u(x_0)$ is defined to be empty.
Given $A \subset \Omega$, we define $\partial u(A) = \bigcup_{x \in A} \partial u(x)$.

Let $S \subseteq \mathbb{R}^n$ and let $x \in \partial S$. A hyperplane $H \subseteq \mathbb{R}^n$ is said to be *a supporting hyperplane* of S at x, if $x \in H$ and S is entirely contained in one of the two closed half-spaces bounded by H.

When $\partial u(x)$ is not empty, its elements are called *subgradients of* u *at* x. A vector $g \in \mathbb{R}^n$ is a subgradient of u at x if the affine function (of z) $u(x) + \langle g, (z - x) \rangle$ defines a supporting hyperplane to the epigraph of u at $(x, u(x))$.

Since the fundamental results by Rockafellar [Roc66, Roc70] we know that subdifferentials of convex functions and monotone functions are very closely related and this relationship can be summarized in the following two theorems:

Theorem 2.7.17 (Rockafellar [Roc66])**.** *If $f : \mathbb{R}^n \to (-\infty, +\infty]$ is a lower semicontinuous proper convex function, then ∂f is a maximal monotone function.*

Theorem 2.7.18 (Rockafellar [Roc70], also see Theorem 12.25 in [RW97])**.** *A function $T : \mathbb{R}^n \to \mathbb{R}^n$ has the form $T = \partial f$ for some proper, lower semicontinuous, convex function $f : \mathbb{R}^n \to (-\infty, +\infty]$ if and only if T is a maximal cyclically monotone function. Then f is determined by T uniquely up to an additive constant.*

Remark 2.7.19. The class of cyclically monotone functions is a proper subclass of the class of monotone functions. However, we are not interested in the notion of cyclical monotonicity — plain monotonicity will suffice for our purposes.

The following theorem shows how derivatives, gradients and subdifferentials of (proper lower semicontinuous) convex functions are related.

Theorem 2.7.20. *Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a lower semicontinuous proper convex function and let $x \in \mathbb{R}^n$. The following are equivalent:*

1. *f is differentiable at x,*
2. *$\nabla f(x)$ exists and*
3. *$\partial f(x)$ is a singleton.*

Moreover, when any of the above conditions is true, we have

$$\partial f(x) = \{\nabla f(x)\} = \{Df(x)\}.$$

2.7.4 Minty's correspondence

We proceed with recalling a connection between monotone functions and Lipschitz functions discovered by Minty [Min62] and some of its consequences relevant to our work.
Minty showed that the Cayley transformation

$$\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n \text{ defined by } \Phi(x, y) = \frac{1}{\sqrt{2}}(y + x, y - x)$$

transforms the graph of a monotone function into a graph of a 1-Lipschitz function. Note that when $n = 1$ this is a clockwise rotation of $\pi/4$.

Notation 2.7.21. By I_n we denote the identity function on $\mathbb{R}^n$. When it won't lead to a confusion we tend to drop the subscript and write I instead of I_n.

We will rely on the following consequences of Minty's discovery.

Proposition 2.7.22 (cf. Proposition 1.2 in [AA99])**.** *Let $u : \mathbb{R}^n \to \mathbb{R}^n$ be a monotone function. Then*

1. *if u is maximal, $\Gamma(u)$ is closed, and $u(x)$ is a convex, closed (possibly empty) set for every $x \in \mathbb{R}^n$;*
2. *u is maximal if and only if $(u + I)$ is onto, i.e., if and only if the domain of $(u + I)^{-1}$ is $\mathbb{R}^n$;*
3. *$(u + I)$ and $(u + I)^{-1}$ are monotone and $(u + I)^{-1}$ is 1-Lipschitz.*

Proposition 2.7.23 (cf. Theorem 12.65 in [RW97]). *Let $u : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous monotone function. Let $z \in \mathbb{R}^n$ and define $f = (u + I)^{-1}$ and $\hat{z} = u(z) + z$. The following two are equivalent:*

1. *u is differentiable at z, and*

2. *f is differentiable at $\hat{z}$ and $Df(\hat{z})$ is invertible.*

Those two propositions will allow us to translate questions about differentiability of monotone functions into questions about differentiability of Lipschitz functions.

2.7.5 Other properties of convex functions on $\mathbb{R}^n$

Proposition 2.7.24 (See Lemma 2.1.8 in [BV10]). *The convex functions on $\mathbb{R}^n$ form a convex cone closed under taking pointwise suprema: if f_γ is convex for each $\gamma \in \Gamma$ then so is $x \mapsto \sup_{\gamma\in\Gamma} f_\gamma(x)$.*

We will also need the following known fact:

Proposition 2.7.25. *Let $g : \mathbb{R} \to \mathbb{R}$ be a monotone function with $Dm\ g = \mathbb{R}$. Let $C \subseteq \mathbb{R}$ be its set of continuity. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \int_0^x g(t)\,\mathrm{d}t$. Then f is convex, $Df = g_{|C}$ and $N_f = \mathbb{R} \setminus C$.*

Proof. Convexity of f follows from monotonicity of g and Theorem 6.2 in [Bru78]. For every $x \in C$, by Theorem 2.6.13 we have $Df(x) = g(x)$. Thus $Df = g_{|C}$ and hence $N_f \subseteq \mathbb{R}\setminus C$. To complete the proof, we need to show $\mathbb{R}\setminus C \subseteq N_f$. Let $z \in \mathbb{R}\setminus C$. Since g is monotone and discontinuous at z, we have either

$$g(z) < \lim_{x\to_+ z, x\in C} g(x) = D_+f(z)$$

or

$$g(z) > \lim_{x\to_- z, x\in C} g(x) = D_-f(z).$$

Either way, we have $D_+f(z) > D_-f(z)$ and then f is not differentiable at z. □

2.7.6 Convex conjugate functions

Definition 2.7.26. The *Fenchel conjugate* of a function $f : \mathbb{R}^n \to [\infty, +\infty]$ is the function $f^* : \mathbb{R}^n \to [\infty, +\infty]$ defined by

$$f^*(\phi) = \sup_{x\in\mathbb{R}^n} \{\langle\phi, x\rangle - f(x)\}.$$

The conjugate function f^* is always convex and if the domain of f is nonempty, then f^* never takes the value $-\infty$.

If f is proper, lower semicontinuous and convex, then one has $f^{**} = f$ and in this case f and f^* are said to be *conjugate*.

2.7.7 Monge-Ampère measures

Definition 2.7.27. Let $\Omega \subseteq \mathbb{R}^n$ be open and let $u : \Omega \to \mathbb{R}$ be continuous. The class $\mathcal{S} = \{E \subseteq \Omega \ : \ \partial u(E) \text{ is Lebesgue measurable}\}$ is a Borel $\sigma-$algebra. The set function $Mu : \mathcal{S} \to [0, +\infty]$ defined by

$$Mu(A) = \lambda\left(\partial u(A)\right)$$

is a measure, called *the Monge-Ampère measure associated with u.*

"Monge-Ampère" in the name comes from the fact that those measures are related to the so called *Monge-Ampère equation*. We will discuss this further in Section 6.2.1. For more details about Monge-Ampère measures, please see the book by Gutiérrez [Gut01].

2.7.8 Symmetric derivative of a measure

In studying differentiability of measures we will rely on the concept of *symmetrical derivatives*, which we define below together with a few non-standard but related notions.

Definition 2.7.28. Let μ be a Borel measure on $\mathbb{R}^n$. Define the *symmetric derivative* of μ at $x \in \mathbb{R}^n$ to be

$$D_\lambda \mu(x) = \lim_{r \to 0} \frac{\mu(B(x,r))}{\lambda(B(x,r))},$$

at those points x at which this limit exists. When this limit exists, we say μ *is differentiable at* x, otherwise we say μ *is not differentiable at* x.

Let $X \in 2^\omega$ be the binary expansion of $x \in \mathbb{R}^n$. Define the *dyadic derivative* of μ at x to be

$$D_2 \mu(x) = \lim_{i \to \infty} \frac{\mu([X \upharpoonright_{ni}])}{\lambda([X \upharpoonright_{ni}])},$$

at those points x at which this limit exists.
Similarly, we define

$$\underline{D}_2 \mu(x) = \liminf_{i \to \infty} \frac{\mu([X \upharpoonright_{ni}])}{\lambda([X \upharpoonright_{ni}])},$$

$$\overline{D}_2 \mu(x) = \limsup_{i \to \infty} \frac{\mu([X \upharpoonright_{ni}])}{\lambda([X \upharpoonright_{ni}])}.$$

Chapter 3

Computable randomness and differentiability on the unit interval

In this chapter we study characterizations of computable randomness on $\mathbb{R}$ in terms of various differentiability properties.

Our first result, Theorem 3.1.9, is a relativized version of the $\Rightarrow$ direction of Theorem 1.6.1. We will use this result in Chapter 4 to prove our effective version of Rademacher's theorem. Our proof uses the notion of porosity. This is a notion smallness of sets in metric spaces. Our proof is simpler than the one in [BMN16]. To a significant extent this proof uses ideas from a paper by Nies [Nie14]. However, the idea of using porosity in this context is certainly older. For real functions of one variable, porosity appears in sets where different types of derivatives disagree (see [BT84, Tho85] and [BLPT86]).

In Section 3.2 we revisit the construction mentioned in Remark 1.5.1.

In Section 3.3 we study how computable randomness is related to differentiability properties of computable convex functions and computable measures. This section contains the majority of our new results on $\mathbb{R}$.

3.1 Porosity, differentiability and betting

3.1.1 Martingales, measures and monotone functions

There is a correspondence, which will be used repeatedly in this chapter, between martingales, atomless (positive) Borel measures on $[0, 1]$ and non-decreasing continuous functions on the $[0, 1]$. Here we set up the notation only, for formal details and proofs, please consult Section 3.2 in [BMN16].

Notation 3.1.1 (Slopes of functions)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be a function. For every $\sigma \in 2^{<\omega}$, we define

$$S_f(\sigma) = \frac{f(0.\sigma + 2^{-|\sigma|}) - f(0.\sigma)}{2^{-|\sigma|}}.$$

Similarly, for all $a, b \in \mathbb{R}$, let

$$S_f(a, b) = \frac{f(a) - f(b)}{a - b}.$$

Notation 3.1.2. For every martingale M there is *a corresponding measure* μ_M on the unit interval defined by

$$\mu_M([\sigma]) = 2^{-|\sigma|}M(\sigma) \text{ for every } \sigma \in 2^{<\omega}.$$

We say a martingale M is *atomless* if μ_M is atomless.

For the other direction, let μ be an atomless (positive) Borel measure on the unit interval. We define *the corresponding martingale* M_μ by letting

$$M_\mu(\sigma) = 2^{|\sigma|}\mu([\sigma]) \text{ for every } \sigma \in 2^{<\omega}.$$

Likewise, there is a correspondence between atomless martingales and continuous non-decreasing functions on the unit interval. Below we introduce a relativized version.

Notation 3.1.3. Let M be an oracle martingale such that M^A is atomless for every $A \in 2^\omega$. We define $f_M : 2^\omega \times [0, 1] \to \mathbb{R}$ on the unit interval by letting

$$f_M^A(x) = f_M(A, x) = \mu_{M^A}([0, x])$$

for all $A \in 2^\omega, x \in [0, 1]$.
For the other direction, let $f : 2^\omega \times [0, 1] \to \mathbb{R}$ be computable and such that

$$f^A(x) = f(A, x)$$

is non-decreasing for every $A \in 2^\omega$. We define an oracle martingale M_f by letting

$$M_f^A(\sigma) = S_{f^A}(\sigma)$$

for every $\sigma \in 2^{<\omega}$.

3.1.2 Porosity points

The notion of porosity, which originated in works of Denjoy, plays an important role in this chapter.

Let (X, d) be a metric space. A point $x \in X$ is said to be a *porosity point* of $S \subseteq X$ if

$$\mathbf{por}(x, S) = \limsup_{r \to 0} \gamma(x, r, S)/r > 0,$$

where $\gamma(x, r, S)$ is defined for any $r > 0$ as

$$\sup\{r' > 0 : \quad \text{for some } z \in X,\ B(z, r') \subseteq B(x, r) \text{ and } B(z, r') \cap S = \emptyset\}.$$

A set S is said to be *porous* if all its points are porosity points of S. A set is said to be $\sigma-$*porous* if it is a countable union of porous sets.
Informally, if x is a porosity point of $S \subseteq \mathbb{R}^n$, then it is possible to find relatively large balls disjoint from S (called *holes in* S) arbitrarily close to x. The following definitions formalize an effective version of this notion.

Definition 3.1.4. Let C be a subset of 2^ω and let $Z \in C$. Define

$$\mathbf{por}_2(Z, C) = \liminf_{i \to \infty} \{|\sigma| - i \mid \sigma \succ Z\restriction_i \wedge\ [\sigma] \cap C = \emptyset\}.$$

If $\mathbf{por}_2(Z, C) < \infty$, then we say that Z is a *dyadic porosity point of* C. When $\rho = \mathbf{por}_2(Z, C)$ is known, we say Z is a ρ*-porosity point of* C.

Since we are interested in computable betting strategies, we need to restrict our attention to subsets of 2^ω for which finding holes can be done effectively.

Definition 3.1.5. Let $X, A \in 2^\omega$. We say X is an $A-$*porosity point* if there exists a $\Pi^0_1(A)$ set $C \subseteq 2^\omega$ such that:

- X is a dyadic porosity point of C and
- the set $H = \{\sigma \mid [\sigma] \cap C = \emptyset\}$ is truth-table reducible to A.

If $A = \emptyset$, we say X *is a computable porosity point.*

The reason of why truth-table reducibility was needed in the above definition will become clear in the proof the next proposition.

Proposition 3.1.6. *Let $X, A \in 2^\omega$. If X is an $A-$porosity point, then there exists an oracle martingale M such that M^A succeeds on X.*

Proof. Let C and H be as in Definition 3.1.5. Let $\rho = \mathbf{por}_2(X, C)$. M^A bets in the following way. Given $\sigma \in 2^{<\omega}$, if there is $\tau \in H$ with $|\tau| = |\sigma| + \rho$ and $\tau \succ \sigma$ (M^A is able to perform such checks, since H is truth-table reducible to A), then M^A bets against it. Otherwise M^A doesn't bet. (That is, we let $M^A(\tau) = 0$ and $M^A(\hat{\tau}) = \frac{2^\rho}{2^\rho - 1} M^A(\sigma)$ for all $\hat{\tau} \succ \sigma$ with $\hat{\tau} \neq \tau$ and $|\hat{\tau}| = |\sigma| + \rho$.) Since X is a dyadic porosity point of C, M^A succeeds on X. □

3.1.3 Computable monotone functions are differentiable at computably random reals

In this section we will prove the $\Rightarrow$ implication of the following result from [BMN16]: $z \in [0,1]$ is computably random iff every computable monotone $f : [0,1] \to \mathbb{R}$ is differentiable at z. The proof presented here follows ideas from [Nie14], where an analogous result has been shown for polynomial time computable functions.
We require the following two lemmata from [Nie14]:

Lemma 3.1.7 (Lemma 11, [Nie14])**.** *Suppose $f\colon [0,1] \to \mathbb{R}$ is a nondecreasing function. Suppose for a real $z \in [0,1]$, with binary representation $z = 0.Z$, there is a rational p such that*

$$\overline{D}_2 f(z) < p < \overline{D} f(z).$$

Let $\sigma^* \prec Z$ be any string such that $\forall \sigma\, [\sigma^* \preceq \sigma \prec Z \Rightarrow S_f(\sigma) \leq p]$. Then the closed set

$$\mathcal{C} = \overline{[\sigma^*]} \setminus \bigcup_{\sigma \succeq \sigma^*} \{[\sigma] \mid S_f(\sigma) > p\}, \tag{3.1}$$

which contains z, is porous at z.

Lemma 3.1.8 (Lemma 12, [Nie14])**.** *Suppose $f\colon [0,1] \to \mathbb{R}$ is a nondecreasing function. Suppose for a real $z \in [0,1]$, with binary representation $z = 0.Z$, there is a rational q such that*

$$\underline{D} f(z) < q < \underline{D}_2 f(z).$$

Let $\sigma^* \prec Z$ be any string such that $\forall \sigma\, [\sigma^* \preceq \sigma \prec Z \Rightarrow S_f(\sigma) \geq q]$. Then the closed set

$$\mathcal{C} = \overline{[\sigma^*]} \setminus \bigcup_{\sigma \succeq \sigma^*} \{[\sigma] \mid S_f(\sigma) < q\}, \tag{3.2}$$

which contains z, is porous at z.

Our formulation of the lemmata is slightly different from the one in [Nie14]. Both formulations are equivalent, however. The original lemmata were formulated in terms of *pseudo-derivatives*. However, as it is explained in a remark that follows the proof of Fact 10 in [Nie14], "for continuous functions with domain $[0,1]$, the lower and upper pseudo-derivatives of f coincide with the usual lower and upper derivatives." Also, please note that our notation for basic dyadic intervals is slightly different (our $[\sigma]$ denotes an open interval).

Theorem 3.1.9. *Let $f : 2^\omega \times [0,1] \to \mathbb{R}$ be a computable function such that $f^A(x) = f(A,x)$ is monotone for every $A \in 2^\omega$. Let $z \in [0,1]$ and let $A \in 2^\omega$. If f^A is not differentiable at z, then there exists an oracle martingale M such that M^A diverges on Z, the binary expansion of z.*

Proof. Consider the oracle martingale M_f. If M_f^A diverges on Z, we are done. Suppose M_f^A converges on Z. This means that $D_2 f^A(z)$ exists (since $D_2 f^A(z) = \lim_{i\to\infty} M_f^A(Z\upharpoonright_i)$) while $Df^A(z)$, by our assumptions, does not. In this case we will show Z is an $A-$porosity point.

Since $D_2 f^A(z)$ exists but $Df^A(z)$ does not, for some positive $p, q \in \mathbb{Q}$, $\overline{D}_2 f^A(z) < p < q < \overline{D} f^A(z)$ or $\underline{D} f^A(z) < p < q < \underline{D}_2 f^A(z)$.
Suppose $\overline{D}_2 f^A(z) < p < q < \overline{D} f^A(z)$. Choose σ^* and $\mathcal{C}$ as in Lemma 3.1.7 so that

$$\overline{D}_2 f(z) < p < \overline{D} f(z).$$

Pick $s \in \mathbb{N}$ such that $2^{-s} < \frac{q-p}{2}$. Define

$$H^A = \left\{ \sigma \mid \sigma \succeq \sigma^* \wedge \left(S_{f^A}(\sigma) - \frac{p+q}{2} \right)_s \geq 0 \right\}.$$

Observe that H^A is truth-table reducible to A. Moreover, the set $C^A = \overline{[\sigma^*] \setminus \bigcup_{\sigma \in H^A} [\sigma]}$ is contained in $\mathcal{C}$ and hence it is porous at z.
Define $\widetilde{C}^A \subset 2^\omega$ by $\widetilde{C}^A = [\sigma^*] \setminus \bigcup_{\sigma \in H^A} [\sigma]$. Let Z be the binary expansion of z. By Lemma 2.4.3 every interval in $\mathbb{R}$ contains a relatively large basic dyadic interval and hence Z is also a dyadic porosity point of $\widetilde{C}^A$.

The case when $\underline{D} f^A(z) < q < \underline{D}_2 f^A(z)$ can be dealt with in a similar manner, using Lemma 3.1.8.

We have shown that Z is an $A-$porosity point. By Proposition 3.1.6, there exists an oracle martingale M such that M^A succeeds (and hence diverges) on Z. □

3.2 Zahorski's construction on the real line

All known proofs of $\Leftarrow$ directions of Theorem 1.6.1 and Theorem 1.6.2 rely on an effectivization of a particular construction found in the seminal paper by Zahorski [Zah46].

Zahorski characterized non-differentiability sets of continuous real-valued functions on the real line. A crucial part of his argument was constructing a monotone Lipschitz function not differentiable at a given G_δ null set. For the purposes of this thesis we will call this part *the Zahorski construction*, even if Zahorski's argument was more complicated and included other important constructions. A modern version of that part of Zahorski's argument can be seen in [FP09]. It should be noted that while Zahorski was the first to fully characterize non-differentiability sets of continuous real-valued functions on the real line, constructions similar to the one we have in mind, had been known before.

We are particularly interested in how this construction can be used to prove results of the following kind: given a non-random point z, exhibit an effective function $f : \mathbb{R} \to \mathbb{R}$ not differentiable at z.

Below we briefly review the construction itself and its effective version. Reproving known results is not the purpose of this section, hence all proofs in this section are sketches.

3.2.1 Classical argument

Let $A \subset \mathbb{R}$ be a G_δ null-set. For the sake of simplicity, suppose it is a subset of the unit interval. The goal is to construct a monotone Lipschitz function not differentiable precisely at elements of A. The basic idea is to exploit the very close relationship between approximate continuity of a bounded measurable function and differentiability of its integral. Recall Theorem 2.6.13: if $f : \mathbb{R} \to \mathbb{R}$ is bounded, measurable and approximately continuous at $z \in (0, 1)$, then $g(x) = \int_0^x f(t)\,\mathrm{d}t$ is differentiable at z. Furthermore, the following simple fact is known:

Fact 3.2.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a locally integrable non-negative function. Let $z \in (0, 1)$. If the map $g(x) = \int_0^x f(t)\,\mathrm{d}t$ is differentiable at z, then the following limit exists:*

$$\lim_{r \to 0} \frac{1}{\lambda\left((z-r, z+r)\right)} \int_{z-r}^{z+r} f(t)\,\mathrm{d}t. \tag{3.3}$$

Proof. Follows from Theorem 7.24 in [Rud87]. □

With the above facts in mind, consider a bounded non-negative and locally integrable function $f : \mathbb{R} \to \mathbb{R}$ with the following properties:

1. f is approximately continuous at all points of $\mathbb{R} \setminus A$ and
2. for every $x \in A$, the limit in Eq. (3.3) does not exist.

Define $g(x) = \int_0^x f(t)\,\mathrm{d}t$. Then the non-differentiability set of g, N_g, is equal to A. Moreover, $g_{|[0,1]}$ is a monotone Lipschitz function. The main point of this idea is to convert the question about differentiability into a relatively easier question about approximate continuity.
Now let us overview the effectivization of this idea from [BMN16].

3.2.2 Effective version

The starting point of the effective construction is slightly different. Let M be a computable martingale with the saving property. Instead of an arbitrary G_δ null-set, we have an effective null-set A of a particular shape — the set of reals x such that M diverges on the binary expansion of x. We are interested in exhibiting a computable monotone function that is not differentiable at all elements of A. Observe that this is

a somewhat weaker property than in the previously discussed classical construction, as we are interested in a computable monotone function whose non-differentiability set contains A and not necessarily is equal to A.
Since M is computable and has the saving property, μ_M is a computable absolutely continuous (w.r.t. λ) measure on $\mathbb{R}$. Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \mu_M((-\infty, x)) = \int_{-\infty}^{x} D_\lambda \mu_M(t)\,\mathrm{d}t.$$

Then g is a computable monotone function that is not differentiable at any $z \in [0, 1]$ whose binary expansion belongs to A. Moreover, g can be made Lipschitz by assuming M is bounded (for details, please see the proof of Theorem 4.2 in [FKHNS14]).

Remark 3.2.2. In this thesis we say a measure is absolutely continuous if it is absolutely continuous w.r.t. the Lebesgue measure.

An obstacle in higher dimensions

The main idea in the above construction is that approximate discontinuities of $D_\lambda \mu_M$ are translated into non-differentibility points of g, the antiderivative of $D_\lambda \mu_M$. It is easy to see at least one major difficulty in generalizing this approach to higher dimensions: where as on the real line Lebesgue integration provides an easy way to obtain antiderivatives, in higher dimensions this is no longer the case.

We will generalize this construction in Chapter 6. In particular, Section 6.1 discusses how to overcome the mentioned difficulty in higher dimensions.

3.3 Convex functions, their derivatives and probability measures

In this section we mainly study differentiability of computable convex functions of one variable. Monotone functions and probability measures also feature prominently in this section.

Convex functions are very well behaved and play an important role in such areas as optimization, control theory and variational analysis. This class has been studied both in the classical and in the effective context.

Derivatives of computable convex functions of one variable have been studied in Ding-Zhu and Ko [DK89]. In particular, Du and Ko noticed that the derivative is (uniformly) computable on the set of points where it does exist. The second and higher derivatives of computable real functions of one variable were considered for example in Zhong [Zho98].

As described in Section 2.7, convex functions are very closely related to monotone functions. Most of results in this section rely on this fact.

3.3.1 Non-differentiability sets of computable convex functions

Recall from Section 2.7.1 that non-differentiability sets of convex real-valued functions on the real line are exactly the countable subsets of $\mathbb{R}$. When f is computable, it is relatively easy to show (and we will do so in this subsection) that all non-differentiability points are computable. However, not all countable sets of computable real numbers are sets of non-differentiability of computable convex functions.

Definition 3.3.1. We say $A \subset \mathbb{R}$ is a *cnd set* (cnd for convex non-differentiability) if there exists a computable convex function $f : \mathbb{R} \to \mathbb{R}$ such that $N_f = A$.

The next two results provide two useful characterizations of cnd sets.

Proposition 3.3.2. *Let $A \subset \mathbb{R}$ be a set. The following two are equivalent:*

1. *there exists a computable sequence of real numbers $(q_i)_{i\in\mathbb{N}}$ and a c.e. set $W \subseteq \mathbb{N}$ such that*
$$A = \{q_i : i \notin W\}.$$
2. *there exists a sequence of uniformly Π^0_1 sets $(P_i)_{i\in\mathbb{N}}$ and a computable sequence of positive reals $(r_i)_{i\in\mathbb{N}}$ such that $A = \bigcup_i P_i$ and for every i, $\#(P_i) \leq 1$ and $P_i \subseteq [-r_i, r_i]$.*

Proof (1) $\Rightarrow$ (2). Let $(q_i)_{i\in\mathbb{N}}$ and W be as in the statement of the theorem. For every $i \in \mathbb{N}$, let $S_i = \{q_i\}$ when $i \notin W$ and let $S_i = \emptyset$ otherwise. Then $\bigcup_i S_i = A$, $(S_i)_{i\in\mathbb{N}}$ is the required sequence of Π^0_1 classes and $(|q_i|)_{i\in\mathbb{N}}$ is the required sequence of computable reals. □

Proof (1) $\Leftarrow$ (2). The following procedure for a given $k \in \mathbb{N}$, computes a real number p_k that is contained in P_k when P_k is not empty.

Fix $k \in \mathbb{N}$ and consider $S = P_k$. Let $i \in \mathbb{N}$. Enumerate open intervals O_j with dyadic endpoints belonging to the complement of S until $[-r_i, r_i] \setminus \cup_j O_j$ is contained in a closed interval S_i with $\lambda(S_i) \leq 2^{-i}$. We may assume $S_{i+1} \subseteq S_i$ for all i and S_1 is not empty. To compute $p_{k,i}$, an approximation of p_k at stage i, we let $p_{k,i}$ to be the leftmost endpoint of S_i if S_i is not empty and let $p_i = p_{i-1}$ otherwise.

Finally, define $W = \{i : \mathbb{N} | P_i = \emptyset\}$. Then $\bigcup_i P_i = \{p_i | i \notin W\}$. □

Theorem 3.3.3. *$A \subset \mathbb{R}$ is a cnd set iff there exists a computable sequence of real numbers $(q_i)_{i\in\mathbb{N}}$ and a c.e. set $W \subseteq \mathbb{N}$ such that*

$$A = \{q_i : i \notin W\}.$$

Proof ⇐. Let $f : \mathbb{R} \to \mathbb{R}$ be a computable convex function such that $N_f = A$. Recall our notation for slopes. For every $i \in \mathbb{Z}$, define

$$P_i = \bigcap_{h\in\mathbb{Q}} \left\{x : S_f(x, x+h) - S_f(x-h, x) \geq 2^{-i}\right\}.$$

From Theorem 2.7.6 it follows that f is not differentiable at x if and only if left and right derivatives of f at x differ, that is $D_+f(x) - D_-f(x) > 0$. Hence every point of N_f is contained in some P_i. By monotonicity of D_+f and D_-f, for any interval $[a, b]$,

$$\text{if } \#([a,b] \cap P_i) \geq m, \text{ then } D_+f(b) - D_-f(a) \geq 2^{-i}m. \tag{3.4}$$

Thus, every P_i is a discrete subset of N_f.

Since we need to find Π^0_1 sets containing at most one element, we need to "split" P_i sets. This can be done by intersecting P_i sets with sufficiently short intervals. The details follow.

For every $j \in \mathbb{Z}$, let $(D^j_i)_{i\in\mathbb{N}}$ be a computable enumeration of all closed intervals with dyadic endpoints $[a, b]$ such that

$$\left|S_f(b, b+2^{-j}) - S_f(a-2^{-j}, a)\right| < \frac{4}{3} \cdot 2^{-j}.$$

Observe that $\bigcup_j (D^j_i)_{i\in\mathbb{N}}$ covers $\bigcup_j P_j$. Furthermore, via (3.4), for every i, $\#\left(D^j_i \cap P_j\right) \leq 1$.

For all $i \in \mathbb{N}, j \in \mathbb{Z}$, define $P_{i,j} = D^j_i \cap P_j$ so that $\#(P_{i,j}) \leq 1$ and let $p_{i,j}$ be the right endpoint of D^j_i so that $P_{i,j} \subset [-p_{i,j}, p_{i,j}]$. Every $P_{i,j}$ is a Π^0_1 set and $N_f = \cup_{i\in\mathbb{N},j\in\mathbb{Z}} P_{i,j}$.

Now we can apply Proposition 3.3.2 to $(P_{i,j})_{i,j\in\mathbb{N}}$ and $(p_{i,j})_{i,j\in\mathbb{N}}$ to get the required result.

□

Proof ⇒. Let $(q_i)_{i\in\mathbb{N}}$ and $W \subseteq \mathbb{N}$ be as in the statement of the theorem. We will exhibit a computable convex function $F : \mathbb{R} \to \mathbb{R}$ such that $N_F = A$.

For $s \in \mathbb{N}$, let W_s denote an approximation of W at stage s, so that $W_s \subseteq W$ and $W = \cup W_s$.

For every i, let $(q_{i,s})_{s\in\mathbb{N}}$ be a computable Cauchy name for q_i.

For a given i, we define (a sequence of functions from $\mathbb{R}$ to $\mathbb{R}$) $(g_{i,s})_{s\in\mathbb{N}}$ in the following way. Fix $s \in \mathbb{N}$. Let $a_s = q_{i,s} + 2^{-s}$ and $b_s = q_{i,s} + 2^{-s+1}$. If $i \in W_s$, then let $g_{i,s} = g_{i,s-1}$

if $g_{i,s-1}$ has been defined and let $g_{i,s} = 0$ otherwise. If $i \notin W_s$, then define $g_{i,s} : \mathbb{R} \to \mathbb{R}$ as

$$g_{i,s}(x) = \begin{cases} 0, \text{ when } x < a_s, \\ 2^{s-i}(x - a_s), \text{ when } x \in [a_s, b_s), \\ 2^{-i}, \text{ when } x \geq b_s. \end{cases}$$

For all $i \in \mathbb{N}$ and $x \in \mathbb{R}$, define

$$g_i(x) = \lim_{s \to \infty} g_{i,s}(x),$$
$$f_i(x) = \int_0^x g_i(t)\, \mathrm{d}t.$$

Finally, define

$$F_1(x) = \sum_{i \in W} f_i(x),$$
$$F_2(x) = \sum_{i \notin W} f_i(x),$$
$$F(x) = F_1(x) + F_2(x).$$

The following two claims complete the proof.

Claim 3.3.4. *F is a computable convex function.*

Convexity of F follows from Proposition 2.7.24, because F is a countable sum of convex functions.
To show computability of F, let $x \in \mathbb{R}$ and $s \in \mathbb{N}$.
Since $g_i \leq 2^{-i}$, we can find $j \in \mathbb{N}$ such that

$$\int_0^x \sum_{i \geq j} g_i(t)\, \mathrm{d}t \leq 2^{-s-1}.$$

Note that every f_i is a computable (uniformly in i) function. Hence $y = \sum_{i<j} f_i(x)$ can be computed from x, s. Let $(y_i)_{i \in \mathbb{N}}$ be a Cauchy name for y. Then

$$|F(x) - y_{s+1}| \leq 2^{-s-1} + |F(x) - y| = 2^{-s-1} + \left| \sum_{i \geq j} f_i(x) \right| \leq 2^{-s}.$$

It follows that F is a computable function. □

Claim 3.3.5. *$N_F = A$.*

Observe that when $i \in W$, then g_i is a continuous function. Moreover, $G_1 = \sum_{i \in W} g_i$ is also a continuous bounded function. Then

$$F_1(x) = \sum_{i \in W} \int_0^x g_i(t)\,\mathrm{d}t = \int_0^x \sum_{i \in W} g_i(t)\,\mathrm{d}t = \int_0^x G_1(t)\,\mathrm{d}t.$$

By Theorem 2.6.13, F_1 is a C^1 function.

When $i \notin W$, g_i is a lower semi-continuous function discontinuous precisely at q_i. Then $G_2 = \sum_{i \notin W} g_i$ is a non-decreasing function discontinuous precisely at elements of A. Since $F_2(x) = \int_0^x G_2(t)\,\mathrm{d}t$, by Proposition 2.7.25, $N_{F_2} = A$. The claim follows. □

From the proof of the previous proposition, we can extract the following useful lemma:

Lemma 3.3.6. *If $A \subseteq \mathbb{R}$ is a cnd set, then there is a computable convex function $u : \mathbb{R} \to \mathbb{R}$ such that $N_u = A$, and $\sup_{x,y \in \mathbb{R}} |Du(x) - Du(y)| = 1$.*

Remark 3.3.7. While there is no computable convex function non-differentiable at all computable reals, for any computable sequence of real numbers, there exists a computable convex function non-differentiable precisely at elements of the sequence. In particular, there is a computable convex function non-differentiable at all rationals. However, the following proposition shows that every computable convex function is differentiable on a dense computable sequence of reals.

Proposition 3.3.8. *Let $f : \mathbb{R} \to \mathbb{R}$ be a computable convex function. There exists a computable $r \in \mathbb{R}$ such that f is differentiable on $r + \mathbb{Q}$.*

Proof. Let $(q_i)_{i \in \mathbb{N}}$ be a computable sequence of real numbers and let $W \subseteq \mathbb{N}$ be a c.e. set such that $N_f = \{q_i : i \notin W\}$.

Let $(p_i)_{i \in \mathbb{N}}$ be a computable sequence of real numbers that enumerates all real numbers of the form $q_i + q$ for all $i \in \mathbb{N}$ and all $q \in \mathbb{Q}$.

Via Proposition 2.1.4, there exists a computable irrational real number r that is not an element of $(p_i)_{i \in \mathbb{N}}$. Then $\{r + q : q \in \mathbb{Q}\} \cap \{p_i : i \in \mathbb{N}\} = \emptyset$ and hence $\{r + q : q \in \mathbb{Q}\} \cap N_f = \emptyset$. □

3.3.2 Derivatives of computable convex functions

Monotone real-valued functions on the real line correspond to derivatives of convex real-valued functions on $\mathbb{R}$ in the following sense. If $f : \mathbb{R} \to \mathbb{R}$ is convex, both its left and right derivatives are defined everywhere and are non-decreasing (recall Theorem 2.7.6). In this case N_f is precisely the set of discontinuity points of both its left and right derivatives. Conversely, if $g : \mathbb{R} \to \mathbb{R}$ is non-decreasing, then $f(x) = \int_0^x g(t)\,\mathrm{d}t$ is convex and N_f is the set of discontinuity points of g. In this subsection we show an analogous characterization in the effective setting.

Let $f : \mathbb{R} \to \mathbb{R}$ be a computable convex function. We know that Df is monotone and it is defined outside of some cnd set. However, Df can be discontinuous and thus not computable in the sense of Grzegorczyk-Lacombe (not even relative to any oracle).

The following fact (stated in [DK89]) shows $Df(x)$ is computable (from x) where it is defined.

Fact 3.3.9. *Let $f : \mathbb{R} \to \mathbb{R}$ be a computable convex function and define $g(x) = Df(x)$. Then g is computable on its set of continuity.*

Proof. By convexity of f, we get that $S_f(x-t, x) \leq Df(x) \leq S_f(x, x+t)$ for all $t > 0$. Moreover, whenever $Df(x)$ exists, $\lim_{t\to 0} S_f(x, x+t) - S_f(x-t, x) = 0$.

Suppose $Df(x)$ exists. For every $s \in \mathbb{N}$ we can find a rational t_s such that

$$S_f(x, x+t_s) - S_f(x-t_s, x) < 2^{-s-1}.$$

Then $|Df(x) - S_f(x-t_s, x)| < 2^{-s}$. □

Fact 3.3.9 justifies the following definition.

Definition 3.3.10. We say a monotone function $f : \mathbb{R} \to \mathbb{R}$ is *weakly computable* if it is computable on its set of continuity.

Our notion of weak computability is closely related to several other natural notions of effectiveness when restricted to monotone functions. One of them is the notion of *almost everywhere computable functions* (see Subsections 7.1 and 7.2 in [Rut16]). We say a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is almost everywhere computable if it is computable on a subset of full measure.

The other notion is that of computability on $I_{\mathbb{Q}} = [0,1] \cap \mathbb{Q}$ (see Section 7 in the arXiv version of [BMN16]). A partial function $f : [0,1] \to \mathbb{R}$ is said to be *computable on $I_{\mathbb{Q}}$* if its domain contains $I_{\mathbb{Q}}$ and $f(q)$ is computable uniformly in $q \in \mathbb{Q}$.

Proposition 3.3.11. *Let $f : \mathbb{R} \to \mathbb{R}$ be an a.e. computable monotone function. Define $u = (\overline{f} + I)^{-1}$, where $\overline{f}$ is the maximal extension of f and I is the identity function. Then u is a computable 1-Lipschitz function.*

Proof. From Proposition 2.7.22 we know u is a 1−Lipschitz function, let us show that it is computable. Since it is Lipschitz, we are only required to show that it is computable uniformly on rationals. Define $\hat{f} = \overline{f} + I$. Clearly, it is also an a.e.

computable function. It is a known fact (for example, see Proposition 7.2 in [Rut16]) that every a.e. computable function is computable on some dense Π^0_2 subset. Hence there is a dense Π^0_2 set $A \subseteq \mathbb{R}$ such that $\hat{f}$ is computable on A. Via Theorem 2.1.5, let $(x_i)_{i\in\mathbb{N}}$ be a dense computable sequence of elements belonging to A.
Let $y \in \mathbb{Q}$, let $s \in \mathbb{N}$. To compute $(u(y))_s$, find $i, j \in \mathbb{N}$ such that $\hat{f}(x_i) < y$, $\hat{f}(x_j) > y$ and $|f(x_i) - f(x_j)| \leq 2^{-s}$; declare $u(y)$ at stage s to be x_i.

□

Proposition 3.3.12. *Let $f : \mathbb{R} \to \mathbb{R}$ be an a.e. computable monotone function and let $C \subseteq \mathbb{R}$ be its set of continuity. There exists a computable convex function $g : \mathbb{R} \to \mathbb{R}$ such that $Dg = f|_C$.*

Proof. Define $g(x) = \int_0^x f(t)\,\mathrm{d}t$. From Proposition 2.7.25 we know $Dg = f|_C$. Let us show that g is computable.

Firstly, suppose $f(0) = 0$. Define $u = (\overline{f} + I)^{-1}$, where $\overline{f}$ is the maximal extension of f. It is a computable 1-Lipschitz function. Define

$$F(x) = \int_0^x (f(t) + t)\,\mathrm{d}t.$$

Suppose f is computable at $x \in \mathbb{R}$ (and hence $x \in C$). Then we have

$$x(f(x) + x) = xu^{-1}(x) = \int_0^x u^{-1}(t)\,\mathrm{d}t + \int_0^{u^{-1}(x)} u(t)\,\mathrm{d}t =$$
$$\int_0^x (f(t) + t)\,\mathrm{d}t + \int_0^{f(x)+x} u(t)\,\mathrm{d}t = F(x) + \int_0^{f(x)+x} u(t)\,\mathrm{d}t.$$

And hence

$$F(x) = x(f(x) + x) - \int_0^{f(x)+x} u(t)\,\mathrm{d}t.$$

This means that $F(x)$ is computable whenever $f(x)$ is computable. Thus F is a convex function that is a.e. computable. Moreover, by Theorem 2.7.6 it is locally Lipschitz and the Lipschitz constant for a given interval can be effectively approximated from above (since $t \mapsto f(t) + t$ is a.e. computable). Using this and Theorem 2.1.5, we can calculate $F(x)$ from x. Thus F is computable. It follows that

$$g(x) = \int_0^x f(t)\,\mathrm{d}t = F(x) - \int_0^x t\,\mathrm{d}t$$

is a computable convex function such that $Dg = f|_C$.

Now suppose $f(0) \neq 0$. Let u be defined as previously and let $z = u(0)$. Note that z is computable. Define $\hat{f}(x) = f(x + z) + z$ so that $\hat{f}(0) = 0$. Let $\hat{g}(x) = \int_0^x \hat{f}(t)\,\mathrm{d}t$.

It is a computable convex function with $D\hat{g} = \hat{f}|_{\hat{C}}$ where $\hat{C} = C - z$ is the set of continuity of $\hat{f}$. Observe that for all $x \in \mathbb{R}$,

$$\hat{g}(x-z) - z(x-z) = \int_z^x f(t)\,\mathrm{d}t.$$

Then $g(x) = \hat{g}(x-z) - z(x-z)$ is a computable convex function with $Dg = f|_C$. □

Now we can prove equivalence between several notions of effectiveness for monotone functions.

Proposition 3.3.13. *Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function. The following are equivalent:*

1. *f is weakly computable,*
2. *f is computable on a dense computable sequence of reals,*
3. *f is computable on $r + \mathbb{Q}$ for some computable r,*
4. *f is a.e. computable, and*
5. *the graph of its maximal extension is a Π^0_1 class.*

Proof. 1. ⇒ 4. holds since monotone functions are a.e. continuous.

4. ⇒ 3. follows from Proposition 3.3.8 and Proposition 3.3.12.

3. ⇒ 2. is trivial.

2. ⇒ 1. follows from the fact that $f(x)$ can be effectively approximated from above and from below when f is continuous at x.

1. ⇒ 5. Let $\overline{f}$ be the maximal extension of f. By Proposition 3.3.11, $(\overline{f}+I)^{-1}$ is a computable Lipschitz function, hence, by Theorem 2.1.9, its graph is a Π^0_1 set. It follows that the graph of $(\overline{f}+I)$ is a Π^0_1 set as well and hence the graph of $\overline{f}$ is a Π^0_1 class too.

5. ⇒ 1. If the graph of the maximal extension of f, $\overline{f}$, is Π^0_1, then the graph of $g = (\overline{f}+I)^{-1}$ is also Π^0_1. Since g is a Lipschitz function with a Π^0_1 graph, it must be computable (via Theorem 2.1.9). This means that whenever $(\overline{f}+I) = g^{-1}$ is continuous at x, we can compute $(\overline{f}+I)(x)$ (for every s we can find y_0, y_1 with $x \in [g(y_0), g(y_1)]$ and $|y_0 - y_1| \leq 2^{-s}$). This means $(\overline{f}+I)$ is weakly computable and hence f must be weakly computable too. □

With Proposition 3.3.8 and Proposition 3.3.13 in mind, while a monotone weakly computable function $f : \mathbb{R} \to \mathbb{R}$ is not necessarily computable on $I_{\mathbb{Q}}$, there exists a computable real r such that $x \mapsto f(x+r)$ is computable on $I_{\mathbb{Q}}$.

Remark 3.3.14. Fact 3.3.9 and Proposition 3.3.12 show that in a precise sense, monotone weakly computable functions correspond to derivatives of computable convex functions. Both left and right derivatives of a computable convex function are monotone weakly computable functions and every monotone weakly computable function restricted to its set of continuity is a derivative of some computable convex function.

Finally, note that just like every continuous monotone real function of one variable is computable relative to some oracle, every monotone function is a weakly computable function relative to some oracle. This suggests that the class of monotone weakly computable functions is an appropriate class for studying monotone but not necessarily continuous functions in the context of computable analysis.

3.3.3 Differentiability of monotone weakly computable functions

The following proposition is a stronger version of the $\Rightarrow$ direction of Theorem 1.6.1. The original proof of Theorem 1.6.1 could not be easily extended to prove our result, since that proof relied on the fact that the martingale defined using slopes of f is computable. Now that f is no longer computable on the whole domain, this approach does not work. However, using Minty's correspondence (see Section 2.7.4), we can reduce the new problem to the previous one.

Proposition 3.3.15. *Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone weakly computable function. If $z \in \mathbb{R}$ is computably random, then f is differentiable at z.*

Proof. Let $z \in \mathbb{R}$ be computably random. By Proposition 3.3.13, there exists a computable real number r such that $x \mapsto f(x+r)$ is computable (and hence continuous) at dyadic rationals. Since z is computably random iff $z+r$ is computably random, we may assume f is continuous at dyadic rationals. Similarly, we may assume $z \in [0,1]$.

By Proposition 3.3.13 f is a.e. computable and hence by Proposition 3.3.11 $u = (\overline{f} + I)^{-1}$ is a computable 1-Lipschitz function. By Proposition 2.7.23, f is differentiable at z iff u is differentiable at $f(z)+z$ and $Du(f(z)+z) \neq 0$.

Claim 3.3.16. *u is differentiable at $f(z)+z$.*

Proof. It is sufficient to prove that $\hat{z} = f(z) + z$ is computably random. Assume otherwise. Note that $u(\hat{z}) = z$. We may assume $\hat{z}$ is not a dyadic rational, for otherwise $u(\hat{z})$ is computable and hence not computably random.

Let $(V_i)_{i\in\mathbb{N}}$ be a bounded Martin-Löf test with $\hat{z} \in \cap V_i$ and let $\nu : 2^{<\omega} \to \mathbb{R}$ be as in Definition 2.5.13 (so that $\lambda(V_i \cap [\sigma]) \leq 2^{-i}\nu(\sigma)$ for all i, σ). We extend ν to a computable absolutely continuous probability measure on $\mathbb{R}$ which we also denote by

ν. Let $(a_{i,j})_{i,j\in\mathbb{N}}, (b_{i,j})_{i,j\in\mathbb{N}}$ be double computable sequences of dyadic rationals such that $V_i = \bigcup_j (a_{i,j}, b_{i,j})$ for all i.

For every i, define $U_i = \bigcup_j (u(a_{i,j}), u(b_{i,j}))$. Observe that since u is monotone, for all $a, b \in \mathbb{R}$, the interior of $u((a,b))$ is equal to $(u(a), u(b))$. Since u is 1-Lipschitz and computable, $(U_i)_{i\in\mathbb{N}}$ is a Martin-Löf test. Moreover, since $\hat{z}$ is not a dyadic rational, $z = u(\hat{z})$ belongs to $\cap U_i$. Define $\nu_u : 2^{<\omega} \to \mathbb{R}$ by letting $\nu_u(\sigma) = \nu(u^{-1}((\sigma)))$ for all σ. Since u^{-1} is continuous on dyadic rationals, ν_u is computable and well-defined.

$$\lambda\left(U_i \cap [\sigma]\right) = \lambda\left(u(V_i \cap u^{-1}([\sigma]))\right) \leq$$
$$\lambda\left(V_i \cap u^{-1}([\sigma])\right) = \lambda\left(V_i \cap u^{-1}((\sigma))\right) \leq 2^{-i}\nu(u^{-1}((\sigma)) = 2^i \nu_u(\sigma).$$

Hence $(U_i)_{i\in\mathbb{N}}$ is a bounded Martin-Löf test and z is not computably random. □

Claim 3.3.17. *$Du(f(z) + z) \neq 0$.*

Proof. Let Z be the binary expansion of z. Suppose $Du(f(z) + z) = 0$, we will show z is not computably random. Note that $u(f(z) + z) = u(u^{-1}(z)) = z$ and since $Du(u^{-1}(z)) = Du(f(z) + z) = 0$, $\overline{D}u^{-1}(z) = +\infty$ and then we have

$$\limsup_{\sigma \prec Z} S_{u^{-1}}(\sigma) = +\infty.$$

Since u^{-1} is a monotone function, $M(\sigma) = S_{u^{-1}}(\sigma)$ is a martingale that succeeds on Z. Since u^{-1} computable on dyadic rationals, M is a computable martingale. □

□

3.3.4 New characterisations of computable randomness

The following result shows two new characterisations of computable randomness on $\mathbb{R}$: in terms of differentiability of monotone weakly computable functions and in terms of twice differentiability of computable convex functions. The result about monotone weakly computable functions generalizes a known theorem from [BMN16] to discontinuous monotone functions. The other result can be seen as a bi-directional effective version of Aleksandrov's Theorem on the real line. To our knowledge it is the first result that characterises a randomness notion in terms of twice differentiability.

Theorem 3.3.18. *Let $z \in \mathbb{R}$. The following are equivalent:*

(1) z is computably random,

(2) all monotone weakly computable functions from $\mathbb{R}$ to $\mathbb{R}$ are differentiable at z,

(3) all computable convex functions from $\mathbb{R}$ to $\mathbb{R}$ are twice-differentiable at z.

(1) $\iff$ (2). For the (2) $\implies$ (1) implication, suppose z is not computably random. Then there is a computable monotone function $g : \mathbb{R} \to \mathbb{R}$ not differentiable at z. The other direction follows from Proposition 3.3.15.

□

(1) $\iff$ (3). For the (3) $\implies$ (1) implication, suppose z is not computably random. Then there is a computable monotone function $g : \mathbb{R} \to \mathbb{R}$ not differentiable at z. Define $f(x) = \int_0^x g(t)\,\mathrm{d}t$, then f is a computable convex function such that $Df = g$. Hence f is not twice differentiable at z.

The other direction follows from Proposition 3.3.15 and the fact that derivatives of computable convex functions are weakly computable (via Fact 3.3.9). □

3.3.5 Atoms and differentiability of computable probability measures on the real line

Recall Monge-Ampère measures introduced in Section 2.7.7. Below we show that computable measures on the real line are, in some specific sense, Monge-Ampère measures of computable convex functions.

Proposition 3.3.19. *Let $u : \mathbb{R} \to \mathbb{R}$ be a computable convex function. If*

$$\sup_{y,x\in\mathbb{R}} |Du(x) - Du(y)| = 1, \tag{3.5}$$

then the Monge-Ampère measure, Mu, is a computable probability measure on $\mathbb{R}$.

Proof. Since $\sup_{y,x\in\mathbb{R}} |Du(x) - Du(y)| = 1$ and Du is monotone, $Du(\mathbb{R})$ is an interval of measure 1 and thus Mu is a probability measure. To show it is computable, by Proposition 2.3.3, it is sufficient to show that $Mu((q, p))$ is left-c.e. uniformly in $q, p \in \mathbb{Q}$. To this end, fix $q, p \in \mathbb{Q}$ with $q < p$.

Let $f = Du$ and let r be a computable real number such that f is computable on $\mathbb{Q}+r$. Let $(q_i)_{i\in\mathbb{N}}$ and $(p_i)_{i\in\mathbb{N}}$ be two computable sequences of elements in $\mathbb{Q}+r$ such that $q \le q_{i+1} \le q_i < p_i \le p_{i+1} \le p$ for all i. Then $\partial u((q, p)) = \bigcup_i \partial u((q_i, p_i))$. Since Mu is countably additive, by Proposition 1.3.3 in [Bog07],

$$Mu((q, p)) = \lambda(\partial u(q, p)) = \lim_{i\to\infty} \lambda(\partial u((q_i, p_i))) = \lim_{i\to\infty} f(p_i) - f(q_i).$$

Clearly, $\lim_{i\to\infty} f(p_i) - f(q_i)$ is left-c.e. uniformly in q, p. □

Proposition 3.3.20. *Let μ be a computable probability measure on the real line. Then $f(x) = \mu((-\infty, x))$ is a monotone weakly computable function with*

$$\sup_{y,x\in\mathbb{R}} |f(x) - f(y)| = 1.$$

Proof. Observe that when x is not an atom of μ, $1 = \mu(-\infty, x) + \mu(x, +\infty)$ and then both $\mu(-\infty, x)$ and $\mu(x, +\infty)$ are computable from x. The claim follows. □

Notation 3.3.21.
Let μ be a measure. By $\mathcal{A}(\mu)$ we denote the set of atoms of μ.

We have the following important consequence:

Proposition 3.3.22. *$A \subset \mathbb{R}$ is a cnd set if and only if there exists a computable probability measure μ on $\mathbb{R}$ with $\mathcal{A}(\mu) = A$.*

Proof. Suppose $A \subset \mathbb{R}$ is a cnd set. Via Lemma 3.3.6, there exists a computable convex function $u : \mathbb{R} \to \mathbb{R}$ such that Eq. (3.5) holds and $N_u = A$. By Proposition 3.3.19, $\mu = Mu$ is a computable probability measure. Moreover, $\mathcal{A}(\mu) = A$ since $\mathcal{A}(\mu)$ is precisely the set of discontinuities of Du.

For the other direction, suppose μ is a computable probability measure on $\mathbb{R}$. Then, by Proposition 3.3.20, $f(x) = \mu((-\infty, x))$ is a monotone weakly computable function. We know that $\mathcal{A}(\mu)$ is the set of discontinuities of f, which is a cnd set. □

The following notion will play an important role in Chapter 4.

Definition 3.3.23. Let μ be a probability measure on $\mathbb{R}^n$. We say μ *admits a computable dyadic pre-measure* if $\mu([\sigma] + p)$ is computable uniformly in $\sigma \in 2^{<\omega}$ and $p \in \mathbb{Z}^n$.

Proposition 3.3.24. *Let μ be a computable probability measure on $\mathbb{R}$. There exists a computable real number r such that for every $q \in \mathbb{Q}$, the measure $\mu_{r,q}$, defined by $\mu_{r,q}(A) = \mu(A + r + q)$, admits a computable dyadic pre-measure.*

Proof. By Proposition 3.3.22 and Proposition 3.3.8, there is a computable real r such that $\mathbb{Q} - r \cap \mathcal{A}(\mu) = \emptyset$. We may assume r is irrational, for otherwise μ does not have rational atoms and then for every $q \in \mathbb{Q}$, $\mu_{0,q}$ admits a computable dyadic pre-measure. Fix $q \in \mathbb{Q}$. Define a measure $\mu_{r,q}$ by $\mu_{r,q}(A) = \mu(A + r + q)$ so that

$\mathcal{A}(\mu_{r,q}) = \mathcal{A}(\mu) + r + q$. Since $\mathbb{Q} - q = \mathbb{Q}$, we have $\mathbb{Q} - r - q = \mathbb{Q} - r$ and then

$$\emptyset = \mathbb{Q} - r \cap \mathcal{A}(\mu) = \mathbb{Q} - r - q \cap \mathcal{A}(\mu) = \mathbb{Q} \cap \mathcal{A}(\mu) + r + q = \mathbb{Q} \cap \mathcal{A}(\mu_{r,q}).$$

Since $\mu_{r,q}$ is computable and does not have rational atoms, it follows that $\mu_{r,q}([\sigma]+p)$ is computable uniformly in $\sigma \in 2^{<\omega}$ and $p \in \mathbb{Z}$. □

Theorem 3.3.25. *Let $z \in \mathbb{R}$ and let $Z \in 2^\omega$ be its binary expansion. The following are pairwise equivalent:*

1. *z is computably random,*
2. *all computable probability measures are differentiable at z, and*
3. *all computable absolutely continuous probability measures are differentiable at z.*

(1) $\implies$ (2). Suppose μ is a computable probability measure on $\mathbb{R}$ which is not differentiable at z. Define $f(x) = \mu((-\infty, x))$ so that by Proposition 3.3.20 f is a monotone weakly computable function. It is easy to check that f is not differentiable at z. It follows that z is not computably random. □

(2) $\implies$ (3). Trivial. □

(3) $\implies$ (1). Suppose $z \in \mathbb{R}$ is not computably random. Let Z be the binary expansion of z.

Without loss of generality, we may assume $z \in [0,1] \setminus \mathbb{Q}$, for otherwise it would be sufficient to consider a computable shift of z by a suitable irrational number.

Let M be a computable martingale with the saving property such that M succeeds on Z. Then the corresponding probability measure μ_M is computable and absolutely continuous. Furthermore, $\overline{D}_2\mu_M(z) = +\infty$ and thus μ_M is not differentiable at z. □

Chapter 4

Computable randomness and differentiability in $\mathbb{R}^n$

The main purpose of this chapter and Chapter 6 is to generalize to $\mathbb{R}^n$ results from Chapter 3 that relate computable randomness to various differentiability properties of effective functions on $\mathbb{R}^n$. The material is split in two chapters: *forward* results, i.e. results of the form "computable randomness at a given point implies a differentiability property X at this point for a given class of effective functions", are dealt with in this chapter, while the *converse* results are left for Chapter 6. The reason for splitting forward and converse results over two separate chapters is that proving techniques required for the former are very different to those needed for the latter.

Most (forward) results in this chapter generalize those in Chapter 3 and represent effective versions of known theorems from analysis. In particular, we prove generalizations of the forward directions of Theorem 1.6.1, Theorem 1.6.2 and Theorem 3.3.18, and these generalizations correspond to classical results from analysis which we state below.

Theorem 4.0.1 (Rademacher [Rad19], 1919)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz function, then it is differentiable almost everywhere on $\mathbb{R}^n$.*

Theorem 4.0.2 (Mignot [Mig76], 1976)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is a monotone function, then it is differentiable almost everywhere on $\mathbb{R}^n$.*

Theorem 4.0.3 (Aleksandrov [Ale39], 1939)**.** *If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then it is twice-differentiable almost everywhere on $\mathbb{R}^n$.*

It is worth mentioning that results in this chapter generalize some of the results from [BMN16] and [FKHNS14], while our effective version of Rademacher's theorem answers a question formulated in [FKHNS14].

As in Chapter 3, we focus on three particular classes of functions: Lipschitz, monotone and convex. These classes are closely related. As we saw in Chapter 3, computable randomness and differentiability of effective members of those classes are related too. When we say that Lipschitz, monotone and convex functions are closely related, we have in mind two particular mathematical facts. In 1962, Minty [Min62] discovered that there is a 1-1 correspondence between graphs of Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ and graphs of monotone functions on $\mathbb{R}^n$. Several years later, Rockafellar [Roc66, Roc70] proved that subdifferentials of convex functions form a proper subclass of monotone functions. A good exposition of classical results related to this area can be found in Alberti and Ambrosio [AA99] and in the Chapter 12 of (the book of) Rockafellar and Wets [RW97].

Historically, Rademacher's theorem and Aleksandrov's theorem had been proven before the mentioned connections (between Lipschitz, monotone and convex functions on $\mathbb{R}^n$) were discovered. On the other hand, Mignot's proof of his result utilized the correspondence discovered by Minty and Rademacher's theorem.

Our approach is to make full use of the results of Minty and Rockafellar. That is, we will base our result about effective monotone functions on our results about effective Lipschitz functions and then our results about monotone functions will be used to prove our effective form of Aleksandrov's theorem.

In fact, the majority of this chapter is devoted to showing the following three results about interactions between differentiability of computable Lipschitz functions on $\mathbb{R}^n$ and computable randomness:

- We will start by proving an effective version of Rademacher's theorem. It is a natural starting point of our developments, since Lipschitz functions enjoy properties that make it possible to deduce an effective version of Rademacher's theorem from a one-dimensional result proven in Chapter 3 and known preservation properties of computable randomness.

- Next, we turn our attention to the question of whether computable monotone Lipschitz functions preserve the property of not being computably random. We will show that when $z \in \mathbb{R}^n$ is not computably random, $f(z)$ is not computably random either, for every computable monotone Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^n$.

- Finally, we will prove an effective version of Sard's theorem for monotone Lipschitz functions.

Our results about monotone and convex functions will follow from results about Lipschitz functions relatively easily.

4.1 Effective Rademacher theorem

In this section we prove the following effective version of Rademacher's theorem.

Theorem 4.1.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable Lipschitz function and let $z \in [0,1]^n$ be computably random. Then f is differentiable at z.*

As an immediate consequence of Theorem 4.1.1 we get the following corollary:

Corollary 4.1.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a computable Lipschitz function and let $z \in \mathbb{R}^n$ be computably random. Then f is differentiable at z.*

Before proceeding to the proof, we need to recall the following known fact about differentiability of Lipschitz functions on $\mathbb{R}^n$.

Lemma 4.1.3. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz and let $x \in \mathbb{R}^n$. Let A be a dense subset of S^{n-1}. If $v \mapsto Df(x; v)$ is defined and is linear on A, then f is differentiable at x.*

Proof. Let T be such that $Df(x; v) = \langle T, v\rangle$ for all $v \in A$. In what follows we show that T is the derivative of f at x.

For all $v \in A$ and $h > 0$, define

$$D(v, h) = \frac{f(x + hv) - f(x)}{h} - \langle T, v\rangle.$$

We will show that for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|D(v, h)| < \epsilon \text{ whenever } 0 < h < \delta,\ v \in S^{n-1}.$$

Let $L = \mathbf{Lip}(f)$. For any $v, v' \in S^{n-1}$ and $h > 0$ we get

$$|D(v, h) - D(v', h)| \leq (L + |T|)|v - v'|. \tag{4.1}$$

By density of A and compactness of S^{n-1}, we can find $v_1, \ldots, v_p \in A$ so that for every $v \in S^{n-1}$, there exists v_k with $1 \leq k \leq p$ such that

$$|v - v_k| \leq \frac{\epsilon}{2(L + |T|)}. \tag{4.2}$$

By the definition of the directional derivative, we get

$$\lim_{h \to 0^+} D(v, h) = 0 \text{ for all } v \in A. \tag{4.3}$$

From (4.3), we get that there is $\delta > 0$ such that

$$D(v_i, h) < \epsilon/2 \text{ whenever } 0 < h < \delta \text{ and } 1 \leq i \leq p.$$

Finally, for $v \in S^{n-1}$ and $0 < h < \delta$, let v_k be such that (4.2) holds and then we have

$$|D(v,h)| \leq |D(v_k,h)| + |D(v,h) - D(v_k,h)| \leq \epsilon/2 + (L + |T|)|v - v_k| < \epsilon.$$

□

Overview of the proof

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable Lipschitz function. Suppose $z \in \mathbb{R}$ is computably random.

Firstly observe that computable elements of S^{n-1} are dense. Thus, via Lemma 4.1.3, it is sufficient to show that $D_+f\,(z; v)$ exists and is linear on the set of all computable unit vectors $v \in \mathbb{R}^n$. This is demonstrated in three distinct steps:

(1) We show that all partial derivatives of f at z exist. This follows from Theorem 3.1.9.

(2) We show the existence of all one-sided directional derivatives of f at z for computable directions. This follows from the previous step and relies on preservation properties of computable randomness.

(3) Finally, we show that the function $T(v) = D_+f\,(z; v)$ is linear on the set of computable directions. More specifically, we prove that any point x where the directional derivative is not linear and the failure of linearity is witnessed by a computable direction, belongs to a Π^0_1 null set and thus is not weakly random. Since z is computably random, this completes the proof.

In the three following subsections we prove results corresponding to the outlined steps.

4.1.1 Existence of partial derivatives

We will rely on the following result from Chapter 3:

Theorem 3.1.9. *Let $f : 2^\omega \times [0,1] \to \mathbb{R}$ be a computable function such that $f^A(x) = f(A,x)$ is monotone for every $A \in 2^\omega$. Let $z \in [0,1]$ and let $A \in 2^\omega$. If f^A is not differentiable at z, then there exists an oracle martingale M such that M^A diverges on Z, the binary expansion of z.*

Lemma 4.1.4. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable Lipschitz function and let $z \in \mathbb{R}^n$ be computably random. All partial derivatives of f at z exist.*

Proof. Fix $i \leq n$. Suppose $D_i f(z)$ does not exist. We may assume no component of z is a dyadic rational. Let Z be the binary expansion of z.
Let $L \geq \mathbf{Lip}(f)$ be rational and define $M = (L, \ldots, L) \in \mathbb{R}^n$. Let $g : \mathbb{R}^n \to \mathbb{R}$ be the function defined by

$$g(x) = f(x) + \langle M, x \rangle.$$

g is a *cone-monotone* computable function. That is, for every $j \leq n$ and all $x \in \mathbb{R}^n$,

$$g(x + he_j) \geq g(x) \text{ whenever } 0 \leq h.$$

Moreover, the non-differentiability sets of f and g coincide. Hence, $D_i g(z)$ does not exist.
Let $y \in \mathbb{R}^{n-1}$ be an element such that $y_j = z_j$ for all $j < i$ and $y_j = z_{j+1}$ for all $j > i$. Let Y be the binary expansion of y.
Recall the following notation. For $j \in \mathbb{N}$ with $0 \leq j \leq n-1$ and $Z \in 2^\omega$, let

$$p_j^n(Z) = \{Z(kn + j) \ : \ k \in \mathbb{N}\}.$$

For all $X \in 2^\omega$, define $t_i(X)$ to be $x \in \mathbb{R}^n$ such that $x_j = 0.p_j^n(X)$ for all $j < i$, $x_i = 0$ and $x_{j+1} = 0.p_j^n(X)$ for all $j > i$.
Define $\hat{g} : 2^\omega \times [0, 1] \to \mathbb{R}$ by

$$\hat{g}(X, h) = g(t_i(X) + he_i)$$

and let $g_y = \hat{g}(Y, \cdot)$. Then $\hat{g}$ satisfies all relevant assumptions of Theorem 3.1.9 and $Dg_y(z_i)$ exists if and only if $D_i g(z)$ exists. By our assumptions, $Dg_y(z_i)$ does not exist. By Theorem 3.1.9, there exists an oracle martingale M such that M^Y does not converge on Z_i, the binary expansion of z_i. This means Z_i is not computably random uniformly relative to Y. Theorem 2.5.6 implies $Y \oplus Z_i$ is not computably random. Since $Y \oplus Z_i$ is a computable permutation of Z, Z is not computably random either. □

4.1.2 Existence of directional derivatives

Now that we've proven existence of partial derivatives, existence of directional derivatives (for computable directions) follows from preservation properties of computable randomness. The details are below.

For any two distinct computable unit vectors $u, v \in \mathbb{R}^n$, fix (say, via the Gram-Schmidt process) two orthonormal bases B_u, B_v of $\mathbb{R}^n$ with $v \in B_v$ and $u \in B_u$. Let $\Theta_{u \to v} : \mathbb{R}^n \to \mathbb{R}^n$ denote the change of basis map (that takes B_u to B_v) such that $\Theta_{u \to v}(u) = v$. This function is computable, linear and invertible, and thus it does preserve computable randomness (see [Rut16]).

Lemma 4.1.5. *Let $u, v \in \mathbb{R}^n$ be distinct unit vectors, let $x \in \mathbb{R}^n$ and let $\Theta = \Theta_{v \to u}$. Then $Df(x; u)$ exists if and only if $Dg(y; v)$ exists, where $g = f \circ \Theta$ and $y = \Theta^{-1}(x)$.*

Proof.
For any $t \neq 0$ we have

$$\frac{g(y+tv)-g(y)}{t} = \frac{f(\Theta(y+tv))-f(\Theta(y))}{t} = \frac{f(x+tu)-f(x)}{t}.$$

By taking the limits of both sides we get the required result. □

Lemma 4.1.6. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable Lipschitz function and let $z \in \mathbb{R}^n$ be computably random. If $u \in \mathbb{R}^n$ is a computable unit vector, then $Df(z; u)$ exists.*

Proof. Let u be a computable unit vector in $\mathbb{R}^n$ and let $v = e_1$. We apply Lemma 4.1.5 to f, v, u and z, so that $Df(z; u)$ exists whenever $Dg(y; v)$ exists, where g is a computable Lipschitz function and $y \in \mathbb{R}^n$ is the image of z under a computable linear and invertible map. Hence y is computably random (again, we use the result from [Rut16] that computable randomness is preserved by a.e. computable isomorphisms). The required result follows from the fact that $Dg(y; v)$ exists iff $D_1 g(y)$ exists (and we know $D_1 g(y)$ exists). □

4.1.3 Linearity of directional derivatives

In the last step of the proof, we need to show that $Df(z; \cdot)$ is linear on computable elements of S^{n-1}. Suppose this is not the case. We will show that in this case z belongs to a Π^0_1 null set. This will get us the required result, since this would mean z is not weakly random.
For $u \in \mathbb{R}^n$, define

$$\mathcal{K}_u = \{z \mid D_+ f(z; u) \text{ exists}\}.$$

For $q \in \mathbb{Q}^+$ and $u, v \in \mathbb{R}^n$, define $L_{u,v,q}$ to be the set of points where linearity of $Df(z; \cdot)$ fails and the failure is witnessed by u, v and q. More formally, let

$$L_{u,v,q} = \mathcal{K}_u \cap \mathcal{K}_v \cap \mathcal{K}_{u+v} \cap \{z \mid |D_+ f(z; u+v) - D_+ f(z; u) - D_+ f(z; v)| \geq q\}.$$

We also need the following notation for directional slopes:

Notation 4.1.7. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, let $x \in \mathbb{R}^n$, let $v \in S^{n-1}$ and let $h > 0$. Define

$$\delta^v_f(x, h) = \frac{f(x+hv)-f(x)}{h}.$$

Lemma 4.1.8. *Let $q \in \mathbb{Q}^+$. Suppose $v, u \in \mathbb{R}^n$ are computable. If $z \in L_{v,u,q}$, then there is a Π^0_1 null-set containing z.*

Proof. Since $D_+ f(z; v), D_+ f(z; u)$ and $D_+ f(z; v+u)$ exist, there is $p > 0$ such that $\left|\delta^v_f(z,h) + \delta^u_f(z,h) - \delta^{v+u}_f(z,h)\right| \geq q$ for all $h \leq p$. Hence the set of all x such that

$$\forall h \left(h \leq p \implies \left|\delta^v_f(x,h) + \delta^u_f(x,h) - \delta^{v+u}_f(x,h)\right| \geq q\right),$$

where h range over positive rationals, contains z. It is clearly a Π^0_1 set and it is a null set, since its complement contains all points of differentiability of f and by the classical Rademacher's theorem f is a.e. differentiable. □

4.1.4 The proof

Proof of Theorem 4.1.1.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable Lipschitz function and let $z \in [0,1]^n$ be computably random. Via Lemma 4.1.6, all directional derivatives $Df(z; v)$ exists whenever v is computable. From Lemma 4.1.8 we know that $Df(z; v)$ is linear on all computable v. The required result follows from Lemma 4.1.3. □

4.2 Computable monotone Lipschitz functions preserve non-randomness

Before we proceed to study differentiability properties of effective monotone functions, we need to prove one property of computable monotone Lipschitz functions of independent interest. In this section we will show that computable monotone Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ preserve (computable) non-randomness. That is, if $f : \mathbb{R}^n \to \mathbb{R}^n$ is a computable monotone Lipschitz function and $z \in \mathbb{R}^n$ is not computably random, then $f(z)$ is not computably random either.

Lemma 4.2.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a computable Lipschitz function. If $A \subset \mathbb{R}^n$ is a compact Π^0_1 null-set, then $f(A)$ is contained in a compact Π^0_1 null-set.*

Proof. Without loss of generality we may assume f is a 1-Lipschitz function. Otherwise, if f is not 1-Lipschitz, we can consider $g = f/K$ (where $K \geq \mathbf{Lip}(f)$ is some fixed computable real number) which is 1-Lipschitz.

For every i, we can effectively (uniformly in i) find a finite cover of A by the closed balls $\overline{B}(x^i_1, 2^{-i}), \ldots, \overline{B}(x^i_{m_i}, 2^{-i})$, such that $\lim_{i\to\infty} \sum_{j\leq m_i} \lambda\left(\overline{B}(x^i_j, 2^{-i})\right) = 0$.

Define B_i as the union of closed balls: $\overline{B}(f(x^i_1), 2^{-i}), \ldots \overline{B}(f(x^i_{m_i}), 2^{-i})$. It is clear that $\cap_i \overline{B}_i$ is a compact Π^0_1 null-set containing $f(A)$. □

Lemma 4.2.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a computable Lipschitz function such that $f^{-1}(C)$ is connected whenever C is a connected open set. Let $x \in \mathbb{Q}^n$ and $r \in \mathbb{Q}$. Then the interior of $f(B(x, r))$ is a Σ^0_1 set (uniformly in x, r).*

Proof. Again, we may assume f is 1-Lipschitz. Let $A = B(x, r)$. Denote the interior of $f(A)$ by $\tilde{A}$ and the complement of A by $\neg A$. Let $(x_{i,j})_{i,j\in\mathbb{N}}$ be a computable double sequence of elements belonging to ∂A and let $p : \mathbb{N} \to \mathbb{N}$ be a computable function such that $\partial A \subseteq \bigcup_{j\leq p(i)} \overline{B}(x_{i,j}, 2^{-i})$ for every i.
For every $i \in \mathbb{N}$, define

$$f[\partial_i A] = \bigcup_{j\leq p(i)} \overline{B}(f(x_{i,j}), 2^{-i+2}),$$

$$f[A^\circ_i] = \bigcup_{q\in\mathbb{Q}^n \wedge |x-q|<r-2^{-i}} B(f(q), 2^{-i}),$$

$$\tilde{A}_i = f[A^\circ_i] \setminus f[\partial_i A].$$

Observe, that $\tilde{A}_i$ is a Σ^0_1 set uniformly in x, r and i. Define $A_f = \bigcup_i \tilde{A}_i$. Then A_f is a Σ^0_1 set uniformly in x, r. Let us show $A_f = \tilde{A}$.

Claim 4.2.3. *$A_f \subseteq \tilde{A}$.*

Proof. Fix i. We will show $\tilde{A}_i \subseteq \tilde{A}$ by demonstrating $y \in f[A_i^\circ] \setminus \tilde{A} \implies y \in f[\partial_i A]$. Suppose $y \in f[A_i^\circ] \setminus \tilde{A}$. There exists $q \in A$ such that $y \in B = B(f(q), 2^{-i})$. Let $F = f^{-1}(B)$. Since $y \notin \tilde{A}$, $B \not\subseteq \tilde{A}$ and there is a point $\hat{y} \in B$ which is not in the closure of $\tilde{A}$. Thus $f^{-1}(\hat{y}) \cap \neg A \neq \emptyset$. Since F is a connected open set (and hence a path-connected open set) that intersects both A and $\neg A$, F intersects the boundary of A. Pick a point $z \in F \cap \partial A$. There exists j such that $|z - x_{i,j}| \leq 2^{-i}$. Then $|f(z) - f(x_{i,j})| \leq 2^{-i}$ and hence $B \subseteq f[\partial_i A]$. □

Claim 4.2.4. $\tilde{A} \subseteq A_f$.

Proof. Let $\tilde{x} \in \tilde{A}$. Pick $N \in \mathbb{N}$ such that $2^{-N} > d(\tilde{x}, \partial\tilde{A})$. Then $d(f^{-1}(\{\tilde{x}\}), \partial A) > 2^{-N}$. Hence $f^{-1}(\{\tilde{x}\}) \cap \partial_i A = \emptyset$ for all $i > N$. It follows that $\tilde{x} \notin f[\partial_i A]$ and $\tilde{x} \in f[A_i^\circ]$ for $i > N$ and thus $\tilde{x} \in A_f$. □

□

Lemma 4.2.5. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a computable Lipschitz function such that $f^{-1}(A)$ is connected whenever A is a connected open set. Let $(V_i)_{i\in\mathbb{N}}$ be a computable sequence of Σ^0_1 sets with $\lambda(V_i) \leq 2^{-i}$ and $V_{i+1} \subseteq V_i$ for all i. There exists a computable sequence of Σ^0_1 sets $(V_i^f)_{i\in\mathbb{N}}$ such that for all i, $f^{-1}(V_i^f) \subseteq V_i$ and $\lambda\left(V_i^f\right) \leq 2^{-i}$. Moreover, for all z, if $z \in \cap_i V_i$ and $f(z)$ is not weakly random, then $f(z) \in \cap_i V_i^f$.*

Proof. Let $(B_{i,j})_{i,j\in\mathbb{N}}$ be a computable double sequence of basic open balls such that $V_i = \cup_j B_{i,j}$ for all i. For all i, j, define $B^f_{i,j}$ as the interior of $f(B_{i,j})$. By Lemma 4.2.2, $\left(B^f{}_{i,j}\right)_{i,j\in\mathbb{N}}$ is a computable double sequence. Since f is Lipschitz, by Theorem 2.7.2, there exists a constant $k \in \mathbb{N}$ such that $\lambda(f(V_i)) \leq 2^k \lambda(V_i)$ for all i.
For all i, define $V_i^f = \cup_j B^f_{i+k,j}$. Then for all i,

$$f^{-1}\left(V_i^f\right) \subseteq V_{i+k} \subseteq V_i$$

and

$$\lambda\left(V_i^f\right) \leq \lambda(f(V_{i+k})) \leq 2^k \lambda(V_{i+k}) \leq 2^{-i}.$$

Finally, let $z \in \cap_i V_i$ be such that $f(z)$ is not weakly random. Since $f(z)$ is not weakly random, by Lemma 4.2.1, $f(z)$ belongs to the interior of $f(B)$ whenever B is a computable basic open ball containing z. It follows that $f(z) \in \cap_i V_i^f$. □

Remark 4.2.6. The above lemma is applicable to the class of computable Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ for which preimages of connected open sets are connected. In particular, this class includes injective (computable Lipschitz) maps and monotone (computable Lipschitz) maps. To see the latter, let f be a monotone Lipschitz map. The generalized inverse of f, f^{-1}, is a maximal monotone function. By the result of Veselý (Theorem 1 in [Ves92]), f^{-1} maps connected open sets to path-connected sets.

Recall that a probability measure μ on $\mathbb{R}^n$ is said to admit a computable dyadic pre-measure if $\mu([\sigma]_p)$ is computable uniformly in $\sigma \in 2^{<\omega}$ and $p \in \mathbb{Z}^n$. In Chapter 3 we have shown that every computable probability measure on the real line can be "shifted" by a computable number so that the new probability measure admits a computable pre-measure. Below we generalize this result to all computable probability measures on $\mathbb{R}^n$.

Proposition 4.2.7. *Let μ be a computable probability measure on $\mathbb{R}^n$. There exists a computable $r \in \mathbb{R}^n$ such that for every $q \in \mathbb{Q}^n$, the measure $\mu_{r,q}$, defined by $\mu_{r,q}(A) = \mu(A + r + q)$, admits a computable dyadic pre-measure.*

Proof. Fix $i \in \mathbb{N}$ with $1 \leq i \leq n$. For all Borel $A \subseteq \mathbb{R}$, define

$$\mu_i(A) = \mu(\{x \in \mathbb{R}^n \ : \ x_i \in A\}).$$

μ_i is clearly a computable probability measure on the real line. Via Proposition 3.3.24, let r_i be a computable real number such that for every $p \in \mathbb{Q}$, the measure $\mu_{i,r_i,p}$, defined by

$$\mu_{i,r_i,p}(A) = \mu_i(A + r_i + p),$$

admits a computable dyadic pre-measure and does not have rational atoms.

Observe that $P_i = \{x \in \mathbb{R}^n \mid x_i \in (\mathbb{Q} + r_i)\}$ is a null-set with respect to μ, that is $\mu(P_i) = 0$.

Define $r \in \mathbb{R}^n$ by $r = \sum_{1 \leq i \leq n} e_i r_i$. r is computable and it is easy to see that for every $q \in \mathbb{Q}^n$, boundaries of cubes of the form $[\sigma] + r + q$ are contained in $\bigcup_i P_i$ and hence are μ-null-sets. It follows that for every $q \in \mathbb{Q}^n$, $\mu([\sigma] + r + q)$ is computable uniformly in σ. □

Using the above proposition we can prove the following generalization of Proposition 2.5.14.

Theorem 4.2.8. *Let $z \in \mathbb{R}^n$ with no dyadic components. The following are equivalent:*

1. *z is not computably random, and*

2. *there exists a Martin-Löf test $(V_i)_{i\in\mathbb{N}}$ covering z and a computable probability measure ν on $\mathbb{R}^n$ such that*

$$\lambda\left(V_i \cap [\sigma]_p\right) \le 2^{-i}\nu([\sigma]_p),$$

for all i, $\sigma \in 2^{<\omega}$ and $p \in \mathbb{Z}^n$.

Proof (1) $\implies$ (2). We may assume $z \in [0,1]^n$.

This direction follows directly from Proposition 2.5.14 and the fact that every computable pre-measure $\nu : 2^{<\omega} \to \mathbb{R}$ can be extended to a computable probability measure on $\mathbb{R}^n$. □

Proof (2) $\implies$ (1). Let $(V_i)_{i\in\mathbb{N}}$ and ν be as in the statement of the theorem. By Proposition 4.2.7, let $r \in \mathbb{R}$ be a computable real such that ν_r, defined by $\nu_r(A) = \nu(A+r)$, admits a computable dyadic pre-measure. For every i, define $\hat{V}_i = V_i - r$, so that $(\hat{V}_i)_{i\in\mathbb{N}}$ is a Martin-Löf test with $z - r \in \cap_i V_i$. For every $\sigma \in 2^{<\omega}$ and $i \in \mathbb{N}$, we have

$$\lambda\left(\hat{V}_i \cap [\sigma]\right) = \lambda\left(V_i \cap ([\sigma] + r)\right) \le 2^{-i}\nu_r([\sigma]).$$

Since ν_r admits a computable dyadic pre-measure, from Proposition 2.5.14 it follows that $z - r$ is not computably random. Hence z is not computably random either. □

Theorem 4.2.9. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a computable monotone Lipschitz function and suppose $z \in \mathbb{R}^n$ is not computably random. Then $f(z)$ is not computably random either.*

Proof. Before we proceed further, we will make the following two assumptions that won't affect the generality of our proof:

(a1) we may assume that neither z, nor $f(z)$ belongs to a compact Π^0_1 null-set, for otherwise, by Lemma 4.2.1, $f(z)$ would not be weakly random;

(a2) since computable randomness is invariant under invertible computable linear transformations, we may assume $z \in (0,1)^n$ and $f(z) \in (0,1)^n$.

Let $(V_i)_{i\in\mathbb{N}}$ be a bounded Martin-Löf test that covers z and let $\nu : 2^{<\omega} \to \mathbb{R}$ be a computable pre-measure with $\lambda\left(V_i \cap [\sigma]\right) \le 2^{-i}\nu(\sigma)$ for all i, σ. We can extend ν to a computable probability measure on $[0,1]^n$ (with its support contained in $[0,1]^n$) which

we will call $\hat{\nu}$. Define a computable probability measure μ on $\mathbb{R}^n$ by $\mu(A) = \hat{\nu}(f^{-1}(A))$ for all Borel A.
We can apply Lemma 4.2.5 to f, z and $(V_i)_{i\in\mathbb{N}}$. Let $(V_i^f)_{i\in\mathbb{N}}$ be as in the conclusion of Lemma 4.2.5. Since $f(z)$ is not weakly random, $(V_i^f)_{i\in\mathbb{N}}$ is a Martin-Löf test that covers $f(z)$. For every i, $\sigma \in 2^{<\omega}$ and $p \in \mathbb{Z}^n$ we have

$$\lambda\left(V_i^f \cap [\sigma]_p\right) \leq \lambda\left(f(V_i \cap f^{-1}([\sigma]_p))\right) \leq \mathbf{Lip}f \cdot \lambda\left(V_i \cap f^{-1}([\sigma]_p)\right) \leq \mathbf{Lip}f \cdot 2^{-i}\mu([\sigma]_p).$$

Then, by Theorem 4.2.8 applied to $\left(V_{i+k}^f\right)_{i\in\mathbb{N}}$ with a suitably chosen k, $f(z)$ is not computably random. □

4.3 An effective version of Sard's theorem for monotone Lipschitz functions

Let $f : \mathbb{R}^n \to \mathbb{R}^m$. Recall that $z \in \mathbb{R}^n$ is said to be a *critical point of* f if either f is not differentiable at z, or $\det Df(z) = 0$ (which means that $Df(z)$, as a matrix, is singular). If z is a critical point of f, then $f(z)$ is said to be a *critical value of* f.

The main result in this subsection, Theorem 4.3.2, can be seen as an effective version of Sard's theorem for monotone Lipschitz function. Its classical version (for all Lipschitz functions), proven by Mignot ([Mig76], also see Theorem 9.65 in [RW97]), states that for a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^n$, the set of its critical values is a null-set.

When a linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ is not invertible, it is neither surjective nor injective and $T(\mathbb{R}^n)$ is a proper subspace of $\mathbb{R}^n$. When f is a Lipschitz function and $Df(z)$ is singular, f, intuitively, "collapses the mass" around z. The following lemma formalizes this intuition:

Lemma 4.3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz function. Suppose $z \in \mathbb{R}^n$ is such that $Df(z)$ is singular. Then for every $\epsilon > 0$, there exists a basic open ball $O_\epsilon = B(x_\epsilon, r_\epsilon)$ containing z, such that $\lambda(f(O_\epsilon)) \leq \epsilon\lambda(O_\epsilon)$.*

Proof. Fix $\epsilon > 0$ and let $k = \mathbf{Lip}(f)$. Let $V = B(0, 1/2) \subset \mathbb{R}^n$ and define $\gamma = \frac{1}{\lambda(V)}$ (so that for any ball $B \subseteq \mathbb{R}^n$ of diameter d and any cube $C \subseteq \mathbb{R}^n$ of side length d, $\gamma = \frac{\lambda(C)}{\lambda(B)}$).

Define $\epsilon' = \frac{\epsilon}{\gamma k^{n-1} 2^n (\sqrt{n})^n}$. Since f is differentiable at z, there exists $\delta > 0$ such that

$$|f(x) - f(z) - Df(z)(x - z)| \leq \epsilon' |x - z| \tag{4.4}$$

for all $x \in \mathbb{R}^n$ with $|x - z| \leq \delta$. There is an open n-cube C with side length equal to $s = \frac{\delta}{\sqrt{n}}$ such that

- its center, $c \in \mathbb{R}^n$, has rational components, and
- $|z - c| \leq \frac{\delta}{4\sqrt{n}}$ and (4.4) holds for all $x \in C$.

Let L be the mapping defined by $L(x) = f(z) + Df(z)(x - z)$. Since $Df(z)$ is singular, L is not onto and its range is contained in some hyperplane H.

As a consequence of (4.4) we have $|f(x) - L(x)| \leq \epsilon'\delta$ for all $x \in C$. Thus, $f(C) \subseteq L(C) + [-\epsilon'\delta, \epsilon'\delta]^n$. Since L is a k-Lipschitz mapping, the image of C under L lies in the intersection of H with a closed ball with radius $k\delta$ centered at $f(z)$. Then $L(C)$ is contained in a rotated $(n-1)$-dimensional cube of side $2k\delta$. This shows that $f(C)$ lies in a rotated box $\hat{C}$ with

$$\lambda\left(\hat{C}\right) = (2k\delta)^{n-1} 2\epsilon'\delta = 2(2k)^{n-1}\epsilon'(\sqrt{n})^n \left(\frac{\delta}{\sqrt{n}}\right)^n = \gamma^{-1}\epsilon \cdot \lambda(C).$$

Finally, define $O_\epsilon = B(c, \frac{\delta}{2\sqrt{n}})$ so that $O_\epsilon \subset C$, $z \in O_\epsilon$ and $\gamma = \frac{\lambda(C)}{\lambda(O_\epsilon)}$. Then

$$\lambda(f(O_\epsilon)) \leq \lambda(f(C)) \leq \lambda\left(\hat{C}\right) = \gamma^{-1}\epsilon \cdot \lambda(C) = \epsilon \cdot \lambda(O_\epsilon).$$

□

Theorem 4.3.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a computable monotone Lipschitz function and let $z \in \mathbb{R}^n$. If $f(z)$ is computably random, then it is not a critical value of f.*

Proof. As in the proofs of Lemma 4.2.1 and Theorem 4.2.9, we may assume the following:

(a1) z and $f(z)$ do not belong to a compact Π^0_1 null-set;

(a2) both z and $f(z)$ belong to the interior of $[0, 1]^n$;

(a3) f is 1-Lipschitz.

The proof is by contraposition. If f is not differentiable at z, then, by Theorem 4.1.1, z is not computably random. Therefore, by Theorem 4.2.9, $f(z)$ is not computably random too.
Suppose f is differentiable at z and $Df(z)$ is singular. Via Lemma 4.3.1, let $(B_i)_{i\in\mathbb{N}}$ be a sequence of basic open balls in $\mathbb{R}^n$ such that for all i, $z \in B_i$ and $\lambda(f(B_i)) \leq 2^{-i}\lambda(B_i)$.
For all Borel $A \subseteq \mathbb{R}^n$, define

$$\nu(A) = \lambda(f^{-1}(A) \cap (0, 1)^n).$$

Then ν is a computable probability measure on $\mathbb{R}^n$.
For every $i \in \mathbb{N}$, define $V_i \subseteq [0, 1]^n$ as the union of all basic open balls B such that $\lambda(B) \leq 2^{-i}\nu(B)$. Then V_i a Σ^0_1 set uniformly in i. Moreover, for all i, $\lambda(V_i) \leq 2^{-i}$ and $B_i \subseteq V_i$. Hence $(V_i)_{i\in\mathbb{N}}$ is a Martin-Löf test that covers $f(z)$.
Fix i and τ. Then $\lambda(V_i \cap [\tau]) \leq 2^{-i}\nu(\tau)$. Hence, by Theorem 4.2.8, $f(z)$ is not computably random. □

4.4 Differentiability of monotone functions on $\mathbb{R}^n$

The fact that monotone functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ are a.e. differentiable has been proven by Mignot [Mig76]. Mignot used Rademacher's theorem and Minty's correspondence (see Section 2.7.4).

In this section we will show that computable randomness implies differentiability of computable monotone functions on $\mathbb{R}^n$. Our proof follows a path similar to the one taken by Mignot: we use Theorem 4.1.1 and the correspondence observed by Minty. However, we also use the preservation property we showed in Section 4.2.

Theorem 4.4.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a computable monotone function and let $z \in \mathbb{R}^n$ be computably random. Then f is differentiable at z.*

Proof. Define $g = (f + I)^{-1}$, so that g is a computable Lipschitz function. Let $y = f(z) + z$ so that $g(y) = z$. By Theorem 4.2.9, y is computably random and hence g is differentiable at y and by Theorem 4.3.2 $Dg(y)$ is invertible. Hence, by Proposition 2.7.23, f is differentiable at z. □

4.5 Convex functions on $\mathbb{R}^n$

Proposition 4.5.1. *If $u : \mathbb{R}^n \to \mathbb{R}$ is a computable convex function, then ∇u is an a.e. computable monotone function.*

Proof. Fix $i \in \mathbb{N}$ with $1 \leq i \leq n$. It is sufficient to show that $D_i u$ is an a.e. computable function.

For all $x \in \mathbb{R}^n$ define $u^i_x : \mathbb{R} \to \mathbb{R}$ by $u^i_x(h) = u(x + he_i)$. Observe that all u_x are convex. Moreover, whenever $D_i u(x)$ exists, it is equal to $Du^i_x(0)$. When $Du^i_x(0)$ exists, it can be computed from x by an algorithm analogous to the one in the proof of Fact 3.3.9. Since $D_i u(x)$ exists almost everywhere, the required result follows. □

If $u : \mathbb{R}^n \to \mathbb{R}$ is convex, then ∂u is a maximal monotone function that extends ∇u. The following proposition shows that the graph of ∂u is a Π^0_1 set.

Proposition 4.5.2. *If $u : \mathbb{R}^n \to \mathbb{R}$ is a computable convex function, then the graph of ∂u is a Π^0_1 set.*

Proof. Rewriting the definition of the subdifferential of u, we have

$$\Gamma(\partial u) = \{(x, y) \ : \ y \in \partial u(x)\} = \bigcap_{z \in \mathbb{R}^n} \{(x, y) \ : \ u(z) \geq u(x) + \langle y, (z - x)\rangle\}.$$

By continuity of u, we get

$$\Gamma(\partial u) = \bigcap_{q \in \mathbb{Q}^n} \{(x, y) \ : \ u(q) \geq u(x) + \langle y, (q - x) \rangle\}.$$

Clearly, $\Gamma(\partial u)$ is a Π^0_1 set. □

Proposition 4.5.3. *Let $u : \mathbb{R}^n \to \mathbb{R}$ be a computable convex function. If the generalized inverse of ∂u, $(\partial u)^{-1}$, is continuous (that is, single-valued everywhere), then it is a computable function.*

Proof. This is consequence of Proposition 4.5.2 and Corollary 9.4 from [Bra08]: Proposition 4.5.2 implies that $\Gamma((\partial u)^{-1})$ is a Π^0_1 set, while Corollary 9.4 from [Bra08] implies that a continuous function $f : \mathbb{R}^n \to \mathbb{R}^n$ is computable iff its graph is a Π^0_1 set. □

Proposition 4.5.4. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable convex function. If $z \in \mathbb{R}^n$ is computably random, then ∇f is differentiable at z.*

Proof. Let $z \in \mathbb{R}^n$ be computably random.
Define $u(x) = f(x) + \frac{1}{2}\langle x, x \rangle$. It is a computable convex function and its subdifferential, $g = \partial u$, is equal to $\partial f + I$. Then, by Proposition 2.7.22, g^{-1} is Lipschitz and, by Proposition 4.5.3, it is computable. Furthermore, ∇u is differentiable at z iff ∇f is differentiable at z.

By Theorem 4.2.9, $y = g(z)$ is computably random. Theorem 4.1.1 implies that g^{-1} is differentiable at y and, by Theorem 4.3.2, $D(g^{-1})(y)$ is invertible. Hence, by Proposition 2.7.23, g is differentiable at z. □

The following effective version of Aleksandrov's theorem is a trivial corollary to Proposition 4.5.4.

Theorem 4.5.5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a computable convex function. If $z \in \mathbb{R}^n$ is computably random, then f is twice-differentiable at z.*

Proof. Follows from Proposition 4.5.4. □

4.6 Almost everywhere computable monotone functions on $\mathbb{R}^n$

In the previous section we have shown that gradients of computable convex functions are differentiable at computably random elements of $\mathbb{R}^n$. In this section, we will extend this result to a.e. computable monotone functions. Note that the proof of

Proposition 4.5.4 would work for a.e. computable monotone functions as well, if we could show that an analog of Proposition 4.5.3 holds for a.e. computable monotone functions and not just for subdifferentials of computable convex functions. That is, we would have to show that whenever the generalized inverse f^{-1} of an a.e. computable maximal monotone function f is continuous, it is also computable. To prove this, we need to show that graphs of maximal extensions of a.e. computable monotone functions are Π^0_1 sets.

Recall that *the convex hull of* $A \subseteq \mathbb{R}^n$ is the intersection of all the convex sets containing A. This set is always convex and it is also the smallest convex set containing A.

Given a monotone (set-valued, partial) function $f : \mathbb{R}^n \to \mathbb{R}^n$, we know that a maximal extension always exists. In general, this extension is not unique. Moreover, classical proofs of this existence statement usually depend on the Axiom of Choice (for example, see the proof of Proposition 12.6 in [RW97]). However, it is known that a monotone function f such that the interior of the convex hull of its domain is not empty and the closure of its domain is convex, has a maximal extension that is unique within the closure of the convex hull of the domain of f. Liqun Qi [Qi83] attributes this result to Philippe Benilan. A similar result has been re-proven in the paper by Crouzeix, Anaya and Sandoval [CAS07]. Below we state the precise statement of the result of interest.

Theorem 4.6.1 (see Theorem 2.6 in [CAS07], also see [Qi83])**.** *Let $f \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a set-valued partial function monotone on its domain. Let C be the closed convex hull of the domain of f. Suppose the interior of C is not empty. Let $V \subset C$ be open, convex and such that $\overline{V} = \overline{V \cap S}$ for some $S \subseteq Dm\ f$. Define*

$$\Gamma = \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n \mid \langle x^* - y^*, x - y\rangle \geq 0 \textit{ for all } y \textit{ and all } y^* \in f(y)\}.$$

If $\overline{f}$ is a maximal extension of f, then the graph of $\overline{f}$ coincides with Γ on V. That is, for all $x \in V$,

$$\overline{f}(x) = \{y \in \mathbb{R}^n \mid (x, y) \in \Gamma\}.$$

The above theorem gives us the following fact: if $f : \mathbb{R}^n \to \mathbb{R}^n$ is a monotone function and its domain is dense in $\mathbb{R}^n$, then there is a unique maximal extension of f and the graph of this extension is equal to

$$\Gamma = \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n \mid \langle x^* - y^*, x - y\rangle \geq 0 \text{ for all } y \text{ and all } y^* \in f(y)\}.$$

Proposition 4.6.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an a.e. computable monotone function. The graph of its maximal extension is a Π^0_1 subset of $\mathbb{R}^{2n}$.*

Proof. Let $(x_i)_{i\in\mathbb{N}}$ be a computable sequence of elements of $\mathbb{R}^n$ such that it is dense in $\mathbb{R}^n$ and $f(x_i)$ is computable uniformly in i. Existence of such a sequence follows from Theorem 2.1.5.
Define
$$V = \bigcap_i \{(\hat{x}, x) \ : \ \langle \hat{x} - f(x_i), x - x_i \rangle \geq 0\}.$$
V is, clearly, a Π^0_1 subset of $\mathbb{R}^{2n}$. By Theorem 4.6.1, V is the graph of the maximal monotone extension of f. □

Theorem 4.6.3. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an a.e. computable monotone function. If $z \in \mathbb{R}^n$ is computably random, then f is differentiable at z.*

Proof. Let $z \in \mathbb{R}^n$ be computably random.
Define $g : \mathbb{R}^n \to \mathbb{R}^n$ by $g = \overline{f} + I$, where $\overline{f}$ is the maximal extension of f, so that g is an a.e. computable monotone function and g^{-1} is Lipschitz. Since g^{-1} is Lipschitz, it is maximal and hence g is maximal as well. By Proposition 4.6.2 we know that the graph of g is a Π^0_1 set and hence g^{-1}, by Theorem 2.1.9, is a computable Lipschitz function. Furthermore, g is differentiable at z iff f is differentiable at z.

By Theorem 4.2.9, $y = g(z)$ is computably random. It follows that g^{-1} is differentiable at y and, by Theorem 4.3.2, $D(g^{-1})(y)$ is invertible. Hence, by Proposition Proposition 2.7.23, g is differentiable at z. □

Chapter 5

Effective Brenier theorem

The goal of this chapter is to prove an effective version of an important result from optimal transport theory, the so called Brenier's theorem. This result will be used in the next chapter to prove a number of converse results which require exhibiting objects with prescribed non-differentiability properties.

Our interest in optimal transport is mainly because this theory deals with *transport maps*:

Definition 5.0.1. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a map. We say T is a *transport map from* $\mu \in P(\mathbb{R}^n)$ *to* $\nu \in P(\mathbb{R}^n)$, or that T *transports* μ onto ν (in symbols, $\nu = T\#\mu$), if for all measurable A,

$$\nu(A) = \mu(T^{-1}(A)).$$

Our main theorem is the following (the notion of *Brenier maps* mentioned in the theorem will be defined later, in Theorem 5.1.2):

Theorem 5.0.2 (Effective Brenier theorem). *Let μ, ν be absolutely continuous computable probability measures on $\mathbb{R}^n$ such that supp $(\mu) = [0,1]^n$ and the support of ν is bounded. There exists a computable convex function $\phi : [0,1]^n \to \mathbb{R}$ such that $\nabla\phi$ transports μ onto ν. Moreover, $\nabla\phi$ is the restriction of the Brenier map to $[0,1]^n$.*

In the next section we provide a very brief introduction to the theory of optimal transport — just enough to formulate the original (non-effective) version of Brenier's result and later to prove our version.

After the introduction to optimal transport, in Section 5.2, we proceed with the proof of Theorem 5.0.2. The proof itself relies on two intermediate results, whose proofs are given later, in Section 5.3 and Section 5.4. One of those intermediate results is of independent interest: we will show that a $K-$Lipschitz function $f : [0,1]^n \to \mathbb{R}$ is computable iff $\{f\}$ is a Π^0_1 subset of a space of uniformly bounded $K-$Lipschitz functions.

5.1 Optimal transport

The history of optimal transport history started in 1781 when the French mathematician Gaspard Monge formulated the problem of *déblais* (a hole in the ground) and *remblais* (a heap of soil). Given the shapes of déblais and remblais, there are many ways of matching the elements (of soil) from their initial positions in the remblais to their final positions in the déblais. The cost (for example, in terms of energy) of moving an element from remblais to déblais depends on both its initial position and on its final position. Monge's problem is: to find an optimal (with respect to some cost function) transference plan of elements from remblais to déblais.

In the 1940s, the Soviet mathematician Leonid Kantorovich reformulated and extended the original Monge problem. In the more modern formulation, both déblais and remblais are modelled as probability measures μ and ν on some topological spaces X and Y, while the cost of moving is modelled as some function $c\colon X \times Y \to \mathbb{R}$. The *Monge-Kantorovich problem* is to find a probability measure π on $X \times Y$ with marginals μ and ν minimizing the total cost $\int_{X\times Y} c(x, y)\,\mathrm{d}\pi(x, y)$. This is a relaxation of the original Monge problem, as the transference plan is modeled by a probability measure, not a map from X to Y. For several decades the theory of optimal transport has been used in many problems arising in probability theory, economics and statistical mechanics.

Relatively recently the field of optimal transport gained a new popularity, often due to numerous (new) discoveries of connections to other areas of mathematics. According to Villani [Vil03], this new popularity can be traced to a single short note by Yann Brenier [Bre87]. What is now known as Brenier's theorem (a result to which several other mathematicians made significant contributions) states that for a quadratic cost function and under some mild assumptions on μ and ν, an optimal transport map exists, is, in some specific way, unique, and is monotone.

In this section we present the bare minimum needed for this chapter. Since we are only interested in $\mathbb{R}^n$, our presentation is simplified in this regard. For more information, please consult any of the recent books available on this subject. In particular, we recommend two books by Villani [Vil03, Vil09], lecture notes by Ambrosio and Gigli [AG13] and another book by Santambrogio [San15]. Our presentation mostly follows [Vil03].

5.1.1 Kantorovich's optimal transportation problem

Let X and Y be some Borel subsets of $\mathbb{R}^n$. For a given pair of probability measures $\mu \in P(X), \nu \in P(Y)$, an *admissible transference plan* is a probability measure $\pi \in P(X \times Y)$ such that

$$\pi(A \times Y) = \mu(A), \qquad \pi(X \times B) = \nu(B) \tag{5.1}$$

holds for all measurable subsets A of X and B of Y.
By $\Pi(\mu,\nu)$ we denote the set of all admissible transference plans:

$$\Pi(\mu,\nu) = \{\pi \in P(X \times Y) \mid \text{ Eq. (5.1) holds for all measurable } A, B\}.$$

For a given *cost function* $c : X \times Y \to \mathbb{R}$, *Kantorovich's optimal transportation problem* is to find $\pi \in \Pi(\mu,\nu)$ that minimizes

$$I_c[\pi] = \int_{X\times Y} c(x,y)\,\mathrm{d}\pi(x,y). \tag{5.2}$$

The quantity $I_c[\pi]$ is called the *total transportation cost* associated with π. The *total transportation cost* between μ and ν is the value

$$\mathbb{I}_c(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} I_c[\pi].$$

A transference plan $\pi \in \Pi(\mu,\nu)$ is said to be *optimal* when $I_c[\pi] = \mathbb{I}_c(\mu,\nu)$.

5.1.2 Monge's optimal transportation problem

We are interested in transference plans induced by transport maps, that is, plans of the form $\pi_T = (I \times T)\#\mu \in \Pi(\mu,\nu)$ where T is some transport map and I is the identity map on $\mathbb{R}^n$. The total transportation cost associated with a transport map T is

$$I_c[T] = I_c[\pi_T] = \int_{\mathbb{R}^n} c(x, T(x))d\mu(x). \tag{5.3}$$

A transport map T for which the cost is optimal, that is for which $I_c[\pi_T] = \mathbb{I}_c(\mu,\nu)$, is called an *optimal transport map.* The problem of minimizing $I_c[T]$ over the set of all transport maps is known as the *Monge optimal transportation problem.* In general, such a map is not guaranteed to exist. However, under some assumptions on μ, ν and c, it does exist.

5.1.3 Optimal transportation theorem for quadratic cost

Notation 5.1.1. Let μ, ν be Borel measures on $\mathbb{R}^n$. Define $M_2(\mu,\nu)$ by

$$M_2(\mu,\nu) = \int_{\mathbb{R}^n} \frac{|x|^2}{2}\,\mathrm{d}\mu(x) + \int_{\mathbb{R}^n} \frac{|x|^2}{2}\,\mathrm{d}\nu(x).$$

For all $f, g : \mathbb{R}^n \to \mathbb{R}$ with $f \in L^1(\mu)$ and $g \in L^1(\nu)$, define

$$J_{\mu,\nu}(f,g) = \int_{\mathbb{R}^n} f\,\mathrm{d}\mu + \int_{\mathbb{R}^n} g\,\mathrm{d}\nu.$$

When μ and ν are fixed, we omit the subscript and write $J(f,g)$ instead of $J_{\mu,\nu}(f,g)$.

The theorem stated below is the one we need to effectivize. It contains two statements. The first one says that a transference plan is optimal if and only if it is concentrated on the graph of the subdifferential of a convex lower-semicontinuous function ϕ. And then (ϕ, ϕ^*), inevitably, is a minimizer of the Eq. (5.5) problem. From our point of view, it is the second part that is important. Later, it will enable us to show that under some conditions on measures, ϕ may be assumed to be computable.

The second statement is the mentioned earlier Brenier's theorem. It says that when measures are sufficiently "nice", $\nabla\phi$ is actually the unique solution to the Monge transportation problem. It is this statement that supplies the transport map to our generalized Zahorski construction.

Theorem 5.1.2 (cf Theorem 2.12 [Vil03] and Theorem 5.10 [Vil09])**.**
Let μ, ν be probability measures on $\mathbb{R}^n$, with $M_2(\mu,\nu) < +\infty$. We consider the Monge-Kantorovich transportation problem associated with the quadratic cost function $c(x,y) = |x-y|^2$. Then,

1. ***(Knot-Smith optimality criterion)*** *$\pi \in \Pi(\mu,\nu)$ is optimal if and only if there exists a convex lower-semicontinuous function ϕ such that*

$$\mathit{supp}\,(\pi) \subseteq \Gamma(\partial\phi). \tag{5.4}$$

   Furthermore, a pair (ϕ, ϕ^*) of lower-semicontinuous proper conjugate convex functions on $\mathbb{R}^n$ is a minimizer in the problem

$$\inf\{J(\phi,\psi) \mid \forall (x,y) \;\; \langle x, y\rangle \le \phi(x) + \psi(y)\} \tag{5.5}$$

   if and only if ϕ satisfies Eq. (5.4) for some $\pi \in \Pi(\mu,\nu)$.

2. ***(Brenier's theorem)*** *If μ is absolutely continuous, then there is a unique optimal π, which is*

$$\pi = (I \times \nabla\phi)\#\mu,$$

   *where $\nabla\phi$ is the unique (i.e. uniquely determined $\mu-$almost everywhere) gradient of a convex function such that $\nu = \nabla\phi\#\mu$ and thus $\mathit{supp}\,(\pi) \subseteq \Gamma(\partial\phi)$. Moreover $\nabla\phi$ is the unique solution to the Monge transportation problem. We call $\nabla\phi$ the **Brenier map** transporting μ onto ν.*

Remark 5.1.3. Please note that our effective version of the above statement, that is our Theorem 5.0.2, merely states that under certain conditions on both measures, the resulting Brenier map is the gradient of a computable convex function.

5.1.4 Kantorovich duality

A well-known phenomenon in optimization is that some minimization problems, like Kantorovich's optimal transportation problem, admit dual formulations. In the context of optimal transport, the following result is known:

Theorem 5.1.4 (Kantorovich duality)**.**
Let X and Y be Polish spaces. Let $\mu \in P(X)$ and $\nu \in P(Y)$, and let $c : X \times Y \to \mathbb{R}_+ \cup \{+\infty\}$ be a lower-semicontinuous cost function.
Define Φ_c to be the set of all pairs of measurable functions $\phi \in L^1(\mu)$ and $\psi \in L^1(\nu)$ satisfying

$$\phi(x) + \psi(y) \leq c(x, y) \tag{5.6}$$

for $\mu-$almost all $x \in X$ and $\nu-$almost all $y \in Y$.
Then

$$\inf_{\pi \in \Pi(\mu,\nu)} I_c[\pi] = \sup_{(\phi,\psi)\in\Phi_c} J(\phi, \psi). \tag{5.7}$$

Moreover, the infimum on the left-hand side of Eq. (5.7) is attained. Furthermore, it does not change the value of the supremum in the right-hand side of Eq. (5.7) if one restricts the definition of Φ_c to bounded and continuous functions.

The left side of Eq. (5.7) in this context is called *the primal problem* and it corresponds to the usual Kantorovich's optimal transportation problem. The right hand side of Eq. (5.7) is called *the dual problem.*

In order to prove our effective version of Brenier's theorem, we will have to demonstrate that the value of the infimum in Eq. (5.5) under certain conditions is computable. The following known fact, in conjunction with Theorem 5.1.4, will be used to show this.

Lemma 5.1.5. *Let μ and ν be two absolutely continuous probability measures on $\mathbb{R}^n$ with $M_2(\mu, \nu) < +\infty$. Then*

$$\inf \{J(\phi, \psi) \mid \forall (x, y) \ \ \langle x, y\rangle \leq \phi(x) + \psi(y)\} = M_2(\mu, \nu) - \sup_{(\phi,\psi)\in\Phi_c} J(\phi, \psi). \tag{5.8}$$

Moreover, if $(\phi(x), \psi(y))$ is a minimizer for the left-hand side of Eq. (5.8), then $\left(\frac{|x|^2}{2} - \phi(x), \frac{|y|^2}{2} - \psi(y)\right)$ is a maximizer for the dual problem, that is

$$J\left(\frac{|x|^2}{2} - \phi(x), \frac{|y|^2}{2} - \psi(y)\right) = \sup_{(\phi,\psi)\in\Phi_c} J(\phi, \psi).$$

A proof of this fact can be found in Subsection 2.1.2 in [Vil03].

5.2 Proof of the Effective Brenier theorem

Before proving Theorem 5.0.2, we need three additional ingredients: a simple lemma which will be stated and proven here, and two theorems whose proofs will be given in Section 5.3 and Section 5.4. Let us start with explaining those two theorems.
Fix $K \in \mathbb{D}^1_*$ and $n \in \mathbb{N}$ with $K > 0$ and $n \geq 1$. Define

$$L_K = \{f \in C([0,1]^n, \mathbb{R}) : \mathbf{Lip} f \leq K \text{ and } |f| \leq K\}.$$

In Section 5.3 we will show that each L_K is a computable metric space (which will be denoted by $\mathbf{L}_K$). Moreover, our Theorem 5.3.14 states that if $\{f\} \subseteq L_K$ is a Π^0_1 set in $\mathbf{L}_K$, then f is a computable function.

The second result, proven in Section 5.4, is Corollary 5.4.3 which says that $(\mu, \nu) \mapsto \mathbb{I}_p(\mu, \nu)$ (where $\mathbb{I}_p(\mu, \nu)$ is the optimal cost w.r.t. the cost function $c_p(x, y) = d(x, y)^p$) is computable.
Finally, the following known fact is the third bit required to prove our effective version of Brenier's theorem:

Lemma 5.2.1 (Potentials are unique up to an additive constant)**.** *Let μ, ν be absolutely continuous probability measures on $\mathbb{R}^n$.*

Suppose supp $(\mu) = [0,1]^n$ and supp (ν) is bounded. If ϕ, ϕ' are convex functions such that $\nabla\phi = \nabla\phi'$ (on $[0,1]^n$) is the Brenier map (transporting μ onto ν), then ϕ, ϕ' are Lipschitz on $[0,1]^n$ and $\phi - \phi'$ is constant on $[0,1]^n$.

Proof. To see that ϕ is Lipschitz on $[0,1]^n$, observe that all partial derivatives of ϕ are uniformly bounded a.e. on $[0,1]^n$ (since supp (ν) is bounded). Let $A \subseteq [0,1]^n$ be the set of points in $[0,1]^n$ where ϕ is differentiable. Then $\phi_{|A}$ is K-Lipschitz for some K and hence ϕ must be K-Lipschitz (on $[0,1]^n$) as well.
Since $[0,1]^n$ is the closure of a connected open set, this means that the difference $\phi - \phi'$ is constant on $[0,1]^n$ (this is a known fact, for example see Problem 5.10.10 in [Eva98]) and this concludes the proof. □

Below we prove that under certain conditions, the Brenier map is an a.e. computable monotone map.

Theorem 5.0.2 (Effective Brenier theorem). *Let μ, ν be absolutely continuous computable probability measures on $\mathbb{R}^n$ such that supp $(\mu) = [0,1]^n$ and the support of ν is bounded. There exists a computable convex function $\phi : [0,1]^n \to \mathbb{R}$ such that $\nabla\phi$ transports μ onto ν. Moreover, $\nabla\phi$ is the restriction of the Brenier map to $[0,1]^n$.*

Proof. In this proof we will use the following custom notation:

Notation 5.2.2. For every $\psi : [0,1]^n \to \mathbb{R}$, define $\psi_\infty : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$\psi_\infty(x) = \begin{cases} \psi(x), & \text{if } x \in [0,1]^n, \\ +\infty, & \text{otherwise.} \end{cases}$$

For every $\phi : \mathbb{R}^n \to \mathbb{R}$, define $\phi_{[0,1]^n} : [0,1]^n \to \mathbb{R}$ as the restriction of ϕ to $[0,1]^n$. Lastly, define ϕ_∞ as $(\phi_{[0,1]^n})_\infty$.

From Lemma 5.2.1 we know that whenever ϕ is a convex function such that $\nabla\phi$ is the Brenier map transporting μ onto ν, ϕ is $K-$Lipschitz for some integer K. Consider the following set

$$S = \{\phi \in \mathbf{L}_K \ : \ \phi(0) = 0, \quad \phi \text{ is convex and } J(\phi_\infty, \phi_\infty^*) = M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu)\}.$$

Below, in a series of claims, we establish that S is a Π^0_1 singleton that contains the restriction $\phi_{[0,1]^n}$ of a convex function ϕ such that $\nabla\phi$ is the Brenier map transporting μ onto ν. Theorem 5.3.14 implies that $\phi_{[0,1]^n}$ is computable.

Claim 5.2.3. *If (ϕ, ϕ^*) is a pair of lower-semicontinuous proper conjugate convex functions on $\mathbb{R}^n$, then $J(\phi_\infty, \phi_\infty^*) \leq J(\phi, \phi^*)$.*

Since $\operatorname{supp}(\mu) = [0,1]^n$, we have

$$\int_{\mathbb{R}^n} \phi_\infty \, \mathrm{d}\mu = \int_{[0,1]^n} \phi_\infty \, \mathrm{d}\mu = \int_{[0,1]^n} \phi \, \mathrm{d}\mu = \int_{\mathbb{R}^n} \phi \, \mathrm{d}\mu.$$

Furthermore,

$$\int_{\mathbb{R}^n} \phi_\infty^* \, \mathrm{d}\nu = \int_{\mathbb{R}^n} \sup_{x \in \mathbb{R}^n} (\langle x, y\rangle - \phi_\infty(x)) \, \mathrm{d}\nu(y) \leq \int_{\mathbb{R}^n} \sup_{x \in \mathbb{R}^n} (\langle x, y\rangle - \phi(x)) \, \mathrm{d}\nu(y) = \int_{\mathbb{R}^n} \phi^* \, \mathrm{d}\nu.$$

The required claim follows.

Claim 5.2.4. *If (ϕ, ϕ^*) is a pair of lower-semicontinuous proper conjugate convex functions on $\mathbb{R}^n$ which is a minimizer for Eq. (5.5) and $\phi(0) = 0$, then $\phi_{[0,1]^n}$ belongs to S.*

By Lemma 5.1.5 we know that the value of (5.5) (that is, $\inf J$) is equal to

$$M_2(\mu,\nu) - \sup_{(\phi,\psi)\in\Phi_c} J(\phi,\psi)$$

where c is the quadratic cost. By Theorem 5.1.4 we have

$$\sup_{(\phi,\psi)\in\Phi_c} J(\phi,\psi) = \inf_{\pi\in\Pi(\mu,\nu)} I_c[\pi] = \mathbb{I}_2(\mu,\nu)$$

(where $\mathbb{I}_2(\mu,\nu)$ is the optimal cost w.r.t. the quadratic cost) and hence the value of (5.5) is equal to $M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu)$.
Suppose (ϕ,ϕ^*) is a pair of lower-semicontinuous proper conjugate convex functions on $\mathbb{R}^n$ with

$$J(\phi,\phi^*) = M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu).$$

Observe that $(\phi_\infty,\phi_\infty^*)$ is (also) a pair of lower-semicontinuous proper conjugate convex functions on $\mathbb{R}^n$. From Claim 5.2.3, we know that

$$J(\phi_\infty,\phi_\infty^*) \leq M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu),$$

and hence

$$J(\phi_\infty,\phi_\infty^*) = M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu).$$

Thus, $\phi_{[0,1]^n} \in S$.

Claim 5.2.5. $\#S = 1$.

Since Eq. (5.5) admits a minimizing pair of convex functions, S is not empty. Indeed, if (ψ,ψ^*) is a pair of lower-semicontinuous proper conjugate convex functions, then $(\psi - \psi(0), \psi^* + \psi(0))$ is also a pair of lower-semicontinuous proper conjugate convex functions. Moreover $J(\psi,\psi^*) = J(\psi - \psi(0), \psi^* + \psi(0))$. Hence by Claim 5.2.4, S is not empty.
Suppose ϕ and ψ belong to S. Then $(\phi_\infty,\phi_\infty^*)$ and $(\psi_\infty,\psi_\infty^*)$ are minimizers for Eq. (5.5). From Theorem 5.1.2 we know that $\nabla\phi_\infty = \nabla\psi_\infty$ on $[0,1]^n$ and from Lemma 5.2.1 we infer that $\psi_\infty - \phi_\infty$ is constant on $[0,1]^n$. Since $\phi(0) = \psi(0)$, we conclude $\psi = \phi$. The claim follows.

Claim 5.2.6. *S is a Π^0_1 set.*

Let $(x_i)_{i\in\mathbb{N}}$ be a computable dense sequence of elements of $[0,1]^n$. Define $C_0 \subseteq \mathbf{L}_K$ as

$$C_0 = \bigcap_{i,j\in\mathbb{N}, t\in\mathbb{Q}\cap[0,1]} \left\{\phi \ : \ \phi(0)=0,\ \phi(tx_i + (1-t)x_j) \leq t\phi(x_i) + (1-t)\phi(x_j)\right\}.$$

Clearly C_0 is a Π^0_1 subset of $\mathbf{L}_K$. Moreover, by continuity, C_0 is the set of convex functions f belonging to $\mathbf{L}_K$ with $f(0) = 0$.
Define $M \subseteq \mathbf{L}_K$ by

$$M = \left\{\phi \in \mathbf{L}_K \ : \ J(\phi_\infty,\phi_\infty^*) = M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu)\right\}.$$

Let us show M is Π^0_1. By Corollary 5.4.3, $\mathbb{I}_2(\mu,\nu)$ is a computable real number. Since supports of μ and ν are bounded, by Theorem 2.3.5, $M_2(\mu,\nu)$ is a computable real number too. Thus,

$$M_2(\mu,\nu) - \mathbb{I}_2(\mu,\nu)$$

is a computable real number.

Observe that for $y \in \mathbb{R}^n$,

$$\phi_\infty^*(y) = \sup_{x\in\mathbb{R}^n} \{\langle x, y\rangle - \phi_\infty(x)\} = \sup_{x\in[0,1]^n} \{\langle x, y\rangle - \phi_\infty(x)\} = \sup_{x\in[0,1]^n} \{\langle x, y\rangle - \phi(x)\} = \max_{x\in[0,1]^n} \{\langle x, y\rangle - \phi(x)\}.$$

The function $x \mapsto \langle x, y\rangle - \phi(x)$ is a computable (uniformly in $y \in \mathbb{Q}^n \cap [0,1]^n$) real-valued Lipschitz function on $[0,1]^n$. This means ϕ_∞^* is a computable function whenever ϕ is computable. Therefore,

$$J(\phi_\infty, \phi_\infty^*) = \int_{\mathbb{R}^n} \phi_\infty \,\mathrm{d}\mu + \int_{\mathbb{R}^n} \phi_\infty^* \,\mathrm{d}\nu = \int_{[0,1]^n} \phi \,\mathrm{d}\mu + \int_{\mathbb{R}^n} \phi_\infty^* \,\mathrm{d}\nu$$

is computable uniformly in ϕ. Hence M is Π_1^0.
Since $S = C_0 \cap M$, S is a Π_1^0 subset of $\mathbf{L}_K$. □

5.3 Subspaces of Lipschitz functions

The main goal of this subsection is to prove that when a Lipschitz function $f : [0,1]^n \to \mathbb{R}$ is such that $\{f\}$ is a singleton Π_1^0 subset of a suitably defined metric space, then f is a computable function (in the Grzegorczyk-Lacombe sense).
Fix $K \in \mathbb{D}_*^1$ and $n \in \mathbb{N}$ with $K > 0$ and $n \geq 1$. Define

$$L_K = \{f \in C([0,1]^n, \mathbb{R}) : \mathbf{Lip} f \leq K \text{ and } |f| \leq K\}.$$

Since elements of L_K are uniformly bounded and equicontinuous, by Ascoli's theorem, L_K is a compact closed subspace of $C([0,1]^n, \mathbb{R})$ (the space of real valued continuous functions on $[0,1]^n$ endowed with the supremum norm $\|\cdot\|_\infty$).

Definition 5.3.1. We say a computable metric space $(X, d, (\alpha_i)_{i\in\mathbb{N}})$ is *effectively compact* if there is a computable function $\nu : \mathbb{N} \to \mathbb{N}$ such that

$$X = \bigcup_{j\leq\nu(i)} B(\alpha_j, 2^{-i}) \quad \text{for all } i.$$

Let d_∞ be the metric on L_K corresponding to the supremum norm. We will show that (L_K, d_∞) with a suitably chosen dense sequence of elements is an effectively compact computable metric space.

Before we can show that L_K is effectively compact, we need to choose a sequence of its elements that is dense. For this purpose we will select a particular sequence of *piecewise affine* functions.

Definition 5.3.2. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *piecewise affine* if for some $k \in \mathbb{N}$, there exists a finite set of affine functions $f_i(x) = A_i x + b_i$, $i = 1, \dots, k$ such that the inclusion $f(x) \in \{f_i(x), \dots, f_k(x)\}$ holds for all x. The functions f_i are called *selection functions*. The set of pairs (A_i, b_i) is called a collection of *matrix-vector pairs* corresponding to f.

It is known that piecewise affine functions are dense in L_K (it will become clear from Lemma 5.3.4 and Lemma 5.3.5).

Definition 5.3.3. Let $A \subset [0,1]^n$ be a finite set. Let $f : A \to [-K, K]$ be a function. Define the *lower interpolant of* f, $\underline{f}$, by

$$\underline{f}(x) = \max\{\max_{y \in A}(f(y) - K|x - y|), -K\}.$$

Observe that lower interpolants are piecewise affine functions. The following Lemma is a simple modification of Proposition 3.1 from [Bel07].

Lemma 5.3.4. *Let $f \subset [0,1]^n \to [-K, K]$ be a partial function with a finite domain. The following are equivalent:*

1. $\mathbf{Lip}(f) \leq K$*, and*
2. $\underline{f}$ *extends* f *and belongs to* L_K*.*

Proof. For the $1 \implies 2$ implication, let A be the domain of f.

Since $x \mapsto (f(y) - K|x-y|)$ is $K-$Lipschitz for every y, A is finite and $-K \leq \underline{f} \leq K$, we have $\underline{f} \in L_K$.

Let $x \in A$. Since $\mathbf{Lip}(f) = K$, for every $y \in A$, we get $f(x) \geq f(y) - K|x - y|$. Then $\underline{f}(x) = \max\{\max_{y \in A}(f(y) - K|x-y|), -K\} = f(x)$ and hence $\underline{f}$ is an extension of f belonging to L_K.

For the $2 \implies 1$ direction, observe that $\mathbf{Lip}(f) > K$ implies that any extension of f can not be $K-$Lipschitz. □

Lemma 5.3.5. *Let $f \in L_K$ and let $i \in \mathbb{N}$. Let f_i be the restriction of f to $\mathbb{D}_i^{n,[0,1]}$. Then*

$$\|f - \underline{f_i}\|_\infty \leq K\sqrt{n}2^{-i}.$$

Proof. Let $x \in [0,1]^n$. For any $y \in \mathbb{D}_i^{n,[0,1]}$ we have

$$\left|f(x) - \underline{f}_i(x)\right| \leq |f(x) - f(y) - \underline{f_i}(x) + \underline{f_i}(y)| \leq |f(x) - f(y)| + |\underline{f_i}(x) - \underline{f_i}(y)|,$$

and therefore

$$\left|f(x) - \underline{f}_i(x)\right| \leq \min_{y \in \mathbb{D}_i^{n,[0,1]}} \left(|f(x) - f(y)| + |\underline{f_i}(x) - \underline{f_i}(y)|\right) \leq$$

$$2K \min_{y \in \mathbb{D}_i^{n,[0,1]}} |x - y| = 2K \cdot d\left(x,\ \mathbb{D}_i^{n,[0,1]}\right) \leq K\sqrt{n}2^{-i}.$$

□

The above fact justifies the following definition.

Definition 5.3.6. Fix a computable enumeration $(\gamma_i)_{i\in\mathbb{N}}$ of $K-$Lipschitz functions

$$\gamma : \mathbb{D}_*^{n,[0,1]} \to \mathbb{D}_*^{1,[-K,K]}$$

such that the domain of γ is finite. Let $(\hat{\gamma}_i)_{i\in\mathbb{N}}$ be the (computable) sequence of lower interpolants of (functions from) $(\gamma_i)_{i\in\mathbb{N}}$.

Proposition 5.3.7. $\mathbf{L}_K = (L_K, d_\infty, (\hat{\gamma}_i)_{i\in\mathbb{N}})$ *is an effectively compact computable Polish space.*

Proof. Firstly note that $(i, j) \mapsto d_\infty(\hat{\gamma}_i, \hat{\gamma}_j)$ is computable uniformly in i, j. This implies that $\mathbf{L}_K$ is a computable Polish metric space.
For every $i \in \mathbb{N}$, define $r(i) = \#\mathbb{D}_i^{n,[0,1]}$. Let $\left(y_j^i\right)_{i,j}$ be a computable double sequence enumerating elements of $\mathbb{D}_i^{n,[0,1]}$ (that is, y_j^i is the $j-$th element of $\mathbb{D}_i^{n,[0,1]}$).
Fix $s \in \mathbb{N}$. To show the required result, we will demonstrate how to compute (uniformly in s) a finite set $I_s \subset \mathbb{N}$ so that

$$L_K = \bigcup_{j\in I_s} B(\hat{\gamma}_j, 2^{-s}).$$

Let $k \in \mathbb{N}$ be such that $\epsilon = 2^{-k} \leq \frac{2^{-s}}{3K\sqrt{n}+1}$. Define

$$S_k = \left\{x \in [-K, K]^{r(k)}\ :\ \forall (i, j \leq r(k))\ \ |x_i - x_j| \leq K\left|y_i^k - y_j^k\right|\right\}.$$

Since $K > 0$, S_k is a *convex body*, that is, a compact convex set with non-empty interior. It follows that $\mathbb{D}_*^{r(k),[-K,K]}$ is dense in S_k. Since conditions of the form $|x_i - x_j| \leq K\left|y_i^k - y_j^k\right|$ are decidable when x_i, x_j, y_i^k, y_j^k and K are dyadic, we can effectively enumerate elements of $S_k \cap \mathbb{D}_*^{r(k),[-K,K]}$. Clearly, S_k is a Π_1^0 set (uniformly in k). Furthermore, the boundary of S_k is also a Π_1^0 set (uniformly in k), since it can be written as

$$\partial S_k = \left\{x \in [-K, K]^{r(k)}\ :\ \forall (i, j \leq r(k))\ \ |x_i - x_j| = K\left|y_i^k - y_j^k\right|\right\}.$$

Since both S_k and ∂S_k are Π^0_1 sets (uniformly in k) and S_k, as a convex body, is homeomorphic to the unit ball in $\mathbb{R}^{r(k)}$, the distance function $x \mapsto d(x, S_k)$ is a computable (uniformly in k) function (this is a straight consequence of Corollary 2.3 from Miller [Mil02]).
For every $z \in S_k$, define $f_z : \mathbb{D}^{n,[0,1]}_k \to [-K, K]$ by setting $f_z(y^k_i) = z_i$ for all $i \leq r(k)$.

Let $f \in L_K$. Define $z \in \mathbb{R}^{r(k)}$ by letting $z_i = f(y^k_i)$ for all $i \leq r(k)$. Clearly $z \in S_k$ and f_z is the restriction of f to $\mathbb{D}^{n,[0,1]}_k$. By Lemma 5.3.4, $\underline{f_z} \in L_K$. By Lemma 5.3.5 we have $d_\infty(f, \underline{f_z}) \leq K\sqrt{n}\epsilon$. It follows that for every $f \in L_K$, there is $x \in S_k$ such that f_x is the restriction of f to $\mathbb{D}^{n,[0,1]}_k$, $\underline{f_x} \in L_K$ and $d_\infty(f, \underline{f_x}) \leq K\sqrt{n}\epsilon$. Thus

$$L_K = \bigcup_{x \in S_k} B(\underline{f_x}, K\sqrt{n}\epsilon).$$

Suppose there exists a finite subset $C_k \subseteq \mathbb{D}^{r(k),[-K,K]}_*$ with $C_k \subset S_k$ such that

$$S_k \subseteq \bigcup_{x \in C_k} B(x, \epsilon). \tag{5.9}$$

Let $x \in S_k$. Let $y \in C_k$ be such that $|x - y| \leq \epsilon$. Then $d_\infty(f_x, f_y) \leq |x - y| \leq \epsilon$ and hence $d_\infty(\underline{f_x}, \underline{f_y}) \leq \epsilon + 2K\sqrt{n}\epsilon$. Then

$$B(\underline{f_x}, K\sqrt{n}\epsilon) \subseteq B(\underline{f_y}, \epsilon + 3K\sqrt{n}\epsilon),$$

and hence, since $\epsilon 3K\sqrt{n} + \epsilon \leq 2^{-s}$,

$$L_K \subseteq \bigcup_{z \in C_k} B\left(\underline{f_z}, \epsilon + 3K\sqrt{n}\epsilon\right) \subseteq \bigcup_{z \in C_k} B\left(\underline{f_z}, 2^{-s}\right).$$

Since for any $x \in \mathbb{D}^{r(k),[-K,K]}_*$, we can calculate j_x, so that $\underline{f_x} = \hat{\gamma}_{j_x}$, this is sufficient to prove the proposition, provided we can show how to compute C_k. This is done below.
Find $m \in \mathbb{N}$ such that $\sqrt{r(k)}2^{-m} + 4 \cdot 2^{-m} \leq 2^{-s}$. Let

$$D_k = \left\{ x \in \mathbb{D}^{r(k),[-K,K]}_m \mid \left(d(x, S_k) - \frac{1}{2}\sqrt{r(k)}2^{-m} \right)_m \leq 2^{-m} \right\},$$

so that

$$\left\{ x \in \mathbb{D}^{r(k),[-K,K]}_m \mid d(x, S_k) \leq \frac{1}{2}\sqrt{r(k)}2^{-m} \right\} \subseteq D_k \text{ and} \tag{5.10}$$

$$D_k \subseteq \left\{ x \in \mathbb{D}^{r(k),[-K,K]}_m \mid d(x, S_k) \leq \frac{1}{2}\sqrt{r(k)}2^{-m} + 2^{-m+1} \right\}. \tag{5.11}$$

D_k is a finite subset of $\mathbb{D}^{r(k),[-K,K]}_m$ and it is computable uniformly in k.

Since S_k is closed and (5.11) holds, for every $x \in D_k$, there exits $x_S \in S_k$ with

$$d(x, x_S) \leq \sqrt{r(k)}2^{-m} + 2^{-m+1}.$$

Since $\mathbb{D}_*^{r(k),[-K,K]}$ is dense in S_k and $\mathbb{D}_*^{r(k),[-K,K]} \cap S_k$ is computably enumerable, for every $x \in D_k$ we can find $\hat{x} \in \mathbb{D}_*^{r(k),[-K,K]} \cap S_k$ such that

$$\left(d(x, \hat{x}) - \frac{1}{2}\sqrt{r(k)}2^{-m}\right)_m \leq 3 \cdot 2^{-m}.$$

Let $\tilde{D}_k$ be the set of such $\hat{x}$'s. Note that $\tilde{D}_k$ is computable uniformly in k.
Let $y \in S_k$. Via (5.10), there exists $x \in D_k$ such that

$$|y - x| \leq \frac{1}{2}\sqrt{r(k)}2^{-m}.$$

Then there exists $\tilde{x} \in \tilde{D}_k$ with $|x - \tilde{x}| \leq \frac{1}{2}\sqrt{r(k)}2^{-m} + 4 \cdot 2^{-m}$. We have

$$|\tilde{x} - y| \leq \sqrt{r(k)}2^{-m} + 4 \cdot 2^{-m} \leq 2^{-s}.$$

Let $C_k = \tilde{D}_k$. It is easy to see that (5.9) holds. This concludes the proof. □

The following proposition seems to be a generally known fact. However, we could not find a proof of it in the literature, hence we prove it here.

Proposition 5.3.8. *Let $(X, d, (\alpha_i)_{i\in\mathbb{N}})$ be an effectively compact computable metric space. There is a computable surjection $\tilde{F} : 2^\omega \to X$.*

Proof idea

The idea is very straightforward: using effective compactness, we compute a tree of open balls and map the Cantor space onto that tree. The (somewhat technical) details are below.

Proof. Let ν be the computable function from Definition 5.3.1. Without loss of generality, we may assume X is infinite and ν is unbounded, so that for every $j \in \mathbb{N}$, there exists $i \in \mathbb{N}$ such that $j \leq \nu(i)$.
Define

$$P = \prod_{i\in\mathbb{N}} \{1, \ldots, \nu(i+2)\}.$$

P, obviously, is a computable image of the Cantor space. Fix a computable surjection $c : 2^\omega \to P$. Suppose there exists a computable function $F : P \times \mathbb{N} \to \mathbb{N}$ with the following properties:

$$d\left(\alpha_{F(A,i)}, \alpha_{F(A,i+1)}\right) \leq 2^{-i+1} \text{ for all } A \in P, i \in \mathbb{N}, \tag{5.12}$$

$$\left\{\lim_{i\to+\infty}\alpha_{F(A,i)} \mid A\in P\right\} \text{ is dense in } X. \tag{5.13}$$

Define $\tilde{F}: 2^\omega \to X$ by

$$\tilde{F}(A) = \lim_{i\to+\infty}\alpha_{F(c(A),i)}.$$

The (5.12) property and the fact that c is computable ensure that $\tilde{F}$ is a computable function, while (5.13) and surjectivity of c imply that the range of $\tilde{F}$ is dense in X. Since $\tilde{F}(2^\omega)$ is closed in X and dense, it must be equal to X and thus $\tilde{F}$ is a surjection. Let us show that such a function F exists.

Definition of F

Let $A \in P$. Define $F(A,1) = A(1)$. Fix $j \in \mathbb{N}$. Suppose $F(A,j) = m$ and we wish to compute $F(A,j+1)$. Define

$$C(m,j) = \left\{i \le \nu(j+1) \ : \ (d(\alpha_m,\alpha_i))_{j+1} \le 2^{-j} + 2^{-j-1}\right\}.$$

This set is finite, non-empty, and computable uniformly in m, j. Furthermore,

$$d(\alpha_m,\alpha_i) \le 2^{-j+1} \text{ for all } i \in C(m,j) \tag{5.14}$$

and

$$\{i \le \nu(j+1) \ : \ d(\alpha_m,\alpha_i) \le 2^{-j}\} \subseteq C(m,j). \tag{5.15}$$

Define $F(A,j+1)$ to be the $k-$th lowest element in $C(m,j)$, where

$$k = A(j+1) \mod \#(C(m,j)).$$

Clearly F is a computable function. The (5.12) property follows from (5.14). The four claims below establish (5.13).

Claim 5.3.9. *If $F(A,j) = m$, then $F(P,j+1)$ contains $C(m,j)$.*

This is a straightforward consequence of our definitions of F and P.

□

Claim 5.3.10. $[1,\nu(j+1)] \subseteq \bigcup_{m\le\nu(j)} C(m,j)$.

This is a consequence of (5.15). In particular, we have

$$\bigcup_{m\le\nu(j)}\{i \le \nu(j+1) \ : \ d(\alpha_m,\alpha_i) \le 2^{-j}\} \subseteq \bigcup_{m\le\nu(j)} C(m,j).$$

Since for every $i \le \nu(j+1)$, there exists $m \le \nu(j)$ such that $d(\alpha_m,\alpha_i) \le 2^{-j}$, the claim follows.

□

Claim 5.3.11. *For all j, $[1, \nu(j)] \subseteq F(P, j)$.*

The proof is by induction on j. By our construction, $[1, \nu(1)] \subseteq F(P, 1)$.
Let $j \in \mathbb{N}$ and suppose $[1, \nu(j)] \subseteq F(P, j)$. Then, by Claim 5.3.9, $F(P, j+1)$ contains $\bigcup_{m \leq \nu(j)} C(m, j)$. Hence, by Claim 5.3.10, $[1, \nu(j+1)] \subseteq F(P, j+1)$. $\square$

Claim 5.3.12. *For every $i \in \mathbb{N}$, there exists $A_i \in P$ such that $\lim_{j \to \infty} F(A_i, j) = i$.*

Let $i \in \mathbb{N}$. Let A and j be such that $F(A, j) = i$. Such j exists by Claim 5.3.11 combined with the fact that by our assumptions ν is unbounded. Moreover, we may assume $j \leq \nu(i)$. Since for all $\hat{j} \geq j$, $i \in C(i, \hat{j})$, there exists $A_i \in P$ such that $A_i(k) = i$ for all $k \geq j$. This establishes the claim and completes the proof of the proposition. $\square$

Proposition 5.3.13. *If $f \in L_K$ is computable in $\mathbf{L}_K$, then it is a computable (in the sense of Grzegorczyk-Lacombe) function.*

Proof. Firstly observe that $(\hat{\gamma}_i)_{i \in \mathbb{N}}$ is a (computable) sequence of computable (in the sense of Grzegorczyk-Lacombe) functions.
Let $s \in \mathbb{N}$. Since $f \in L_K$ is computable in $\mathbf{L}_K$, we can find $\hat{\gamma}_{j_s}$ so that $\|f - \hat{\gamma}_{j_s}\|_\infty \leq 2^{-s-1}$. Then

$$\left|\left(\hat{\gamma}_{j_s}(x)\right)_{s+1} - f(x)\right| \leq 2^{-s},$$

for all $x \in [0, 1]^n$. $\square$

Theorem 5.3.14. *If $\{f\} \subseteq L_K$ is a Π^0_1 set in $\mathbf{L}_K$, then f is a computable (in the sense of Grzegorczyk-Lacombe) function.*

Proof. From Proposition 5.3.8 we know that there exists a computable surjection $F_K : 2^\omega \to \mathbf{L}_K$. Using this surjection, it is possible to enumerate all basic closed balls that intersect $\{f\}$. This implies that f is a computable point in $\mathbf{L}_K$ and hence, by Proposition 5.3.13, it is a computable function. $\square$

5.4 Optimal cost is computable

Definition 5.4.1 (Wasserstein metrics)**.** Let (X, d) be a Polish metric space. For $p \in \mathbb{N}$ with $p \geq 1$, define a cost function c_p by $c_p(x, y) = d(x, y)^p$. For $\mu, \nu \in P(X)$, define the *Wasserstein metric of order p* by

$$W_p(\mu, \nu) = \mathbb{I}_p(\mu, \nu)^{1/p}$$

where $\mathbb{I}_p$ is the optimal transport cost between μ and ν with respect to c_p.

It is known that when d is bounded, W_p metrizes the weak topology on $P(X)$. Furthermore, by Theorem 4.1.1 in [HR09], W_1 is computable and it is *computably equivalent* to the Prokhorov metric π (see the Definition 2.3.2). That is, given a Cauchy name of μ with respect to π, we can compute a Cauchy name of μ with respect to W_1 and vice versa. Since we are mainly concerned with the quadratic cost, we need to prove an analogous result for $p > 1$.

Recall that $(\delta_i)_{i\in\mathbb{N}}$ is a computable sequence of basic points in $P(X)$ - those elements of $P(X)$ which are concentrated on finite subsets of basic points of X and assign rational values to them.

Proposition 5.4.2 (Wasserstein metrics are computable). *Let $(X, d, (\alpha_i)_{i\in\mathbb{N}})$ be a computable metric space where d is bounded. Then $(P(X), W_p, (\delta_i)_{i\in\mathbb{N}})$ is a computable metric space and W_p computably equivalent to W_1 (and hence to π).*

Proof. Without loss of generality we may assume $p > 1$. Let us show $W_p(\delta_i, \delta_j)$ is computable uniformly in i, j. Let $\mu = \delta_{i_1}$ and $\nu = \delta_{i_2}$ for some $i_1, i_2 \in \mathbb{N}$. Suppose μ is concentrated on $a_1, \ldots, a_n \in X$ and ν is concentrated on $b_1, \ldots, b_m \in X$. For all $i \leq n, j \leq m$, define $z_{i,j} = (a_i, b_j)$ and $p_{i,j} = d(a_i, b_j)^p$. Then $\pi \in \Pi(\mu, \nu)$ if and only if π is concentrated on all $z_{i,j}$ and for all i, j,

$$\sum_{\hat{i}\leq n} \pi(z_{\hat{i},j}) = \nu(b_j), \text{ and } \sum_{\hat{j}\leq m} \pi(z_{i,\hat{j}}) = \mu(a_i).$$

Define $F : \mathbb{R}^{nm} \to \mathbb{R}$ by

$$F(x_{1,1}, \ldots, x_{n,m}) = \sum_{i\leq n, j\leq m} p_{i,j} x_{i,j}.$$

Then we have

$$I_{d^p}[\pi] = F\left(\pi(z_{1,1}), \ldots, \pi(z_{n,m})\right).$$

And then

$$\mathbb{I}_p(\mu, \nu) = \min F(x_{1,1}, \ldots, x_{n,m}),$$
$$\text{subject to}$$
$$\sum_{\hat{i}\leq n} x_{\hat{i},j} = \nu(b_j), \text{ and } \sum_{\hat{j}\leq m} x_{i,\hat{j}} = \mu(a_i) \text{ for all } i, j.$$

This is a linear optimization problem (over variables $x_{1,1}, \ldots, x_{n,m}$) and thus it can be solved uniformly in μ and ν (that is, uniformly in i_1 and i_2). This shows that $(P(X), W_p, (\delta_i)_{i\in\mathbb{N}})$ is a computable metric space.

To show that W_p is computably equivalent to W_1, we will use the following inequalities (see 7.1.2 in [Vil03]):

$$W_1 \leq W_p \leq W_1^{1/p} \text{diam}(X)^{1-1/p}. \tag{5.16}$$

Fix integer D with $D \geq \text{diam}(X)$. Let $(\mu_i)_{i \in \mathbb{N}}$ be a Cauchy name of $\mu \in P(X)$ with respect to π. Since W_1 is computable and $W_1(\mu, \mu_i) \to 0$, given $j \in \mathbb{N}$, we can effectively find $k(j) \in \mathbb{N}$ such that

$$W_1(\mu, \mu_{k(j)}) \leq 2^{-jp} D^{-p}.$$

Then, via (5.16), for every j,

$$W_p(\mu, \mu_{k(j)}) \leq 2^{-j}.$$

Hence $\left(\mu_{k(i)}\right)_{i \in \mathbb{N}}$ is a Cauchy name of μ with respect to W_p.
The other direction is trivial since $W_1 \leq W_p$. □

Corollary 5.4.3 (Optimal cost is computable)**.** *Let $(X, d, (\alpha_i)_{i \in \mathbb{N}})$ be a computable metric space where d is bounded. Then $(\mu, \nu) \mapsto \mathbb{I}_p(\mu, \nu)$ is computable.*

Chapter 6

Controlling non-differentiability in $\mathbb{R}^n$

In Chapter 3 we have shown several characterizations of computable randomness on the real line in terms of differentiability properties of certain classes of functions and measures. Every such result has two directions: the $\Rightarrow$ direction that asserts computable randomness is sufficient for some property to hold and the converse direction stating that computable randomness is necessary. In Chapter 4 we have generalized some of the mentioned $\Rightarrow$ results to $\mathbb{R}^n$. In this chapter we focus on proving the converses.

All the converse results on the unit interval that we have in mind rely on the effective version of the Zahorski construction described briefly in Section 3.2.1. The main insight of this chapter is that the Zahorski construction can be generalized to $\mathbb{R}^n$ using Brenier's theorem. This construction then can be effectivized using the computable version of Brenier's theorem proven in Chapter 5.

The main result of this chapter is the following theorem:

Theorem 6.0.1 (Main theorem)**.**
Let $z \in [0,1]^n$ be not computably random. There exists a computable convex function $u : [0,1]^n \to \mathbb{R}$ and a computable function $g : [0,1]^n \to \mathbb{R}^n$ such that

- $\nabla u = g$ *on* $(0,1)^n$ *and,*
- ∇u *is not differentiable at* z*.*

The above theorem can be seen as the converse result for three theorems from Chapter 4: Theorem 4.5.5, Theorem 4.4.1 and Theorem 4.6.3. Using it, we will generalize to $\mathbb{R}^n$ some of the results from Chapter 3, in particular, Theorem 3.3.18 and Theorem 3.3.25.

The structure of this chapter is following:

- We start by describing how the Zahorski construction can be reinterpreted in terms of transport maps and how the theory of optimal transport is relevant in this context. The whole of Section 6.1 is devoted to this issue.
- Section 6.2 is where all the details needed for the ideas outlined in Section 6.1 are filled in. This section contains the proof of Theorem 6.0.1 and all the intermediate results.
- Finally, in the last section, we use Theorem 6.0.1 to show several new characterizations of computable randomness in $\mathbb{R}^n$.

6.1 Reinterpreting the effective Zahorski construction

Let us recall the effective version of Zahorski's construction from [BMN16]. Suppose $z \in (0,1)$ is not computably random and let $Z \in 2^\omega$ be its binary expansion. Our goal is to exhibit a computable monotone function on the unit interval that is not differentiable at z.
Let M be a computable martingale with the saving property diverging on Z. Since M is atomless, the corresponding measure μ_M is computable and absolutely continuous. Define $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \mu_M((0,x)) = \int_0^x D_\lambda \mu_M(t)\, \mathrm{d}t. \tag{6.1}$$

Then f is a computable monotone function not differentiable at z. As it was explained in the first chapters of this thesis, points of approximate discontinuity of $D_\lambda \mu_M$ become points where f is not differentiable.
We can summarize this approach in the following way:

The Zahorski construction on $\mathbb{R}$ via integration

Obtain an absolutely continuous measure μ not differentiable at $z \in (0,1)$. Then $D_\lambda \mu$ is not approximately continuous at z and $x \mapsto \int_0^x D_\lambda \mu(t)\, \mathrm{d}t$ is a monotone function not differentiable at z.

A key role in this approach is played by the very close relationship between integrals and antiderivatives that is well known for real-valued functions of one variable. The major obstacle in generalizing this idea to higher dimensions is that in $\mathbb{R}^n$ ($n > 1$), this elegant relation is lost. In order to generalize, we would have to reinterpret this construction without relying on integration as a mean of obtaining antiderivatives. Instead of integration, we will rely on the notion of transport maps.

It can be shown that the function f defined in (6.1) is a transport map from μ_M to λ_1. That is, $\lambda_1 = f\#\mu_M$. Intuitively, a transport map f from μ_M to λ_1 turns oscillations of $\frac{\mu_M(B_r(z))}{\lambda(B_r(z))}$ into oscillations of f's slopes at z, thus turning non-differentiability of μ_M at z into non-differentiability of f at z.
In fact, any transport map f from μ_M to λ_1 that maps intervals to intervals, is not differentiable at z. For any such function f, for any $z \in \mathbb{R}$ and $r > 0$, we have

$$\frac{\mu_M(B(z,r))}{\lambda\left(B(z,r)\right)} = \frac{\lambda\left(f(B(z,r))\right)}{\lambda\left(B(z,r)\right)} = \frac{f(z+r) - f(z-r)}{2r}.$$

If $D_\lambda \mu_M(z)$ does not exist, then $\lim_{r\to 0} \frac{f(z+r)-f(z-r)}{2r}$ does not exist either. On the other hand, when f is differentiable at z, then $\lim_{r\to 0} \frac{f(z+r)-f(z-r)}{2r}$ exists. Thus if μ is not differentiable at z, then f is not differentiable at z.

Remark 6.1.1. As we know from Brenier's theorem, f is the unique optimal transport map from μ_M onto λ_1. While Brenier's theorem shows existence and uniqueness of the optimal map, it does not provide an explicit formula for the map. On the other hand, we gave an explicit formula for the transport map f. In fact, in the case of measures on $\mathbb{R}$, optimal maps (with respect to quadratic cost) have a known form:

Proposition 6.1.2. *Let μ and ν be probability measures on $\mathbb{R}$. Suppose μ is atomless. Let F and G be their respective cumulative distribution functions. Then $T = G^{-1} \circ F$ is the (optimal) monotone transport map from μ onto ν.*

For more information about the one-dimensional optimal transport problem, please see Section 2.2 in [Vil03].

The above considerations make it possible to rephrase the Zahorski construction in the following way:

The Zahorski construction on $\mathbb{R}$ via transport maps

Obtain an absolutely continuous measure μ not differentiable at $z \in (0,1)$. Then the monotone transport map from μ to λ_1 is not differentiable at z.

In this reformulation we rely on the fact that differentiability of the source measure μ affects differentiability of the transport map in the way that suits us (see Fact 3.2.1). And this particular phenomenon is known in higher dimensions too. Consider the following theorem:

Theorem 6.1.3 (see Theorem 7.24 in [Rud87]). *If*

(a) V is open in $\mathbb{R}^n$,

(b) $T : V \to \mathbb{R}^n$ is continuous, and

(c) T is differentiable at some point $z \in V$, then

$$\lim_{r\to 0} \frac{\lambda\left(T(B(z,r))\right)}{\lambda\left(B(z,r)\right)} = \left|\det DT(z)\right|. \tag{6.2}$$

This theorem can be seen as a generalization of Fact 3.2.1. In order to see this, suppose T is from $\mathbb{R}$ to $\mathbb{R}$ and of the form considered in Fact 3.2.1. That is $T(x) = \int_0^x f(t)\,\mathrm{d}t$. Then the limit (6.2) from Theorem 6.1.3 can be rewritten to match the limit (3.3) from Fact 3.2.1:

$$\lim_{r\to 0} \frac{\lambda\left(T(B(z,r))\right)}{\lambda\left(B(z,r)\right)} = \lim_{r\to 0} \frac{T(z+r)-T(z-r)}{\lambda\left((z-r,z+r)\right)} = \lim_{r\to 0} \frac{1}{\lambda\left((z-r,z+r)\right)} \int_{z-r}^{z+r} f(t)\,\mathrm{d}t.$$

Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous transport map from μ to λ_n, where μ is some absolutely continuous probability measure on $\mathbb{R}^n$. Then the limit in (6.2), assuming $x \in (0,1)^n$, can be rewritten as

$$\lim_{r\to 0} \frac{\lambda\left(T(B(x,r))\right)}{\lambda\left(B(x,r)\right)} = \lim_{r\to 0} \frac{\mu(B(x,r))}{\lambda\left(B(x,r)\right)} = D_\lambda \mu(x).$$

Hence, if μ is not differentiable at some point x, then T is not differentiable at x either. The remaining two questions are:

1. is there always a monotone transport map $T : \mathbb{R}^n \to \mathbb{R}^n$ from μ onto λ_n and

2. under what conditions can T be assumed to be computable.

Brenier's theorem provides an answer to the first question: if μ is regular enough, then the answer is positive. The second question will be answered with the help of the Effective Brenier Theorem from our previous chapter and a regularity result which will be discussed later.

With the above in mind, we can formulate a provisional generalization of the Zahorski construction to $\mathbb{R}^n$:

The Zahorski construction on $\mathbb{R}^n$ via transport maps

Obtain an absolutely continuous measure μ not differentiable at $z \in (0,1)^n$. Then the monotone transport map from μ onto λ_n is not differentiable at z.

It is worth making it clear, that when we write about *the effective Zahorski construction on* $\mathbb{R}^n$, we actually mean the part of the proof of Theorem 6.0.1 where we exhibit a computable monotone transport map not differentiable at a given (not computably random) point.

6.2 Effective Zahorski construction on $\mathbb{R}^n$

The proof of the main result is a short one, but it does rely on three other results: the effective Brenier theorem from Chapter 5, and two other results which will be discussed later in this section. The effective Brenier theorem gives us an a.e. computable monotone transport map. The other two results will ensure the transport map is actually Hölder continuous and hence computable.

Theorem 6.0.1 (Main theorem)**.**
Let $z \in [0,1]^n$ be not computably random. There exists a computable convex function $u : [0,1]^n \to \mathbb{R}$ and a computable function $g : [0,1]^n \to \mathbb{R}^n$ such that

- $\nabla u = g$ *on* $(0,1)^n$ *and,*
- ∇u *is not differentiable at* z*.*

Proof. Let $z \in [0,1]^n$ be not computably random. We may assume it lies in the interior of $[0,1]^n$.
By Proposition 6.2.6, there exists a computable absolutely continuous probability measure μ on $\mathbb{R}^n$ such that its support is equal to $[0,1]^n$, its density is bounded away from 0 and $+\infty$ and it is not differentiable at z.
By Theorem 5.0.2, there exists a computable convex function $\phi : [0,1]^n \to \mathbb{R}$ such that $\nabla\phi$ transports μ onto λ_n.

By Theorem 6.2.5, we know that $\nabla\phi : (0,1)^n \to \mathbb{R}^n$ is Hölder continuous. There is a unique Hölder continuous extension of $\nabla\phi$ to $[0,1]^n$, which we will call $g : [0,1]^n \to \mathbb{R}^n$. Since ϕ is a computable convex function, g is a.e. computable. Since g is Hölder continuous and a.e. computable, it must be computable.
Since $\nabla\phi(B_r(z)) \subset [0,1]^n$ for sufficiently small r, we have

$$\lim_{r\to 0} \frac{\lambda\left(\nabla\phi(B(z,r))\right)}{\lambda\left(B(z,r)\right)} = \lim_{r\to 0} \frac{\lambda_n\left(\nabla\phi(B(z,r))\right)}{\lambda\left(B(z,r)\right)} = \lim_{r\to 0} \frac{\mu(B(z,r))}{\lambda\left(B(z,r)\right)} = D_\lambda\mu(z).$$

We know that $D_\lambda\mu(z)$ does not exists. Thus, by Theorem 6.1.3, $\nabla\phi$ is not differentiable at z. □

6.2.1 The Monge-Ampère equation and regularity of optimal transport

The Monge-Ampère equation is a famous fully nonlinear elliptic partial differential equation of the following form:

$$\det D^2\phi(x) = F(x, \phi(x), \nabla\phi(x)), \tag{6.3}$$

where by $D^2\phi$ we denote the second derivative of ϕ (that is, $D\nabla\phi$, when ϕ is convex).

In order to ensure this equation is elliptic (so that the theory of uniformly elliptic equations is applicable), Eq. (6.3) is restricted to the set of convex ϕ's. Note that when ϕ is convex, $\det D^2\phi(x)$ is an L^1 function defined almost everywhere. In order for ϕ to satisfy Eq. (6.3) for all $x \in \mathbb{R}^n$, ϕ would have to be twice-differentiable. Such functions are called *classical solutions* of Eq. (6.3). Since we have no interest in twice-differentiable functions in this chapter, we will consider *weak solutions* of Eq. (6.3) of a particular type.

> **Definition 6.2.1.** We say ϕ is *an Aleksandrov solution* of Eq. (6.3), if the Monge-Ampère measure $M\phi$ is absolutely continuous and its density coincides with the right-hand side of Eq. (6.3) almost everywhere.

It is known that the following particular type of the Monge-Ampére equation is closely related to the theory of optimal transport:

$$\det D^2\phi(x) = \frac{f(x)}{g(\nabla\phi(x))}. \tag{6.4}$$

It can be shown (see Section 1.7.6 in [San15]) that when ϕ is C^2, f is the density of μ and g is the density of ν, (6.4) is equivalent to

$$\nu = \nabla\phi\#\mu.$$

This motivates the following notion of weak solutions:

> **Definition 6.2.2.** Let μ and ν be two probability measures on $\mathbb{R}^n$. A convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be *a Brenier solution* of Eq. (6.4) if
>
> 1. $\nu = \nabla\phi\#\mu$, and
>
> 2. f is the density of μ and g is the density of ν.

This notion of weak solutions is strictly weaker than the notion of Aleksandrov solutions — a Brenier solution is not necessarily an Aleksandrov solution.

In order to prove Theorem 6.0.1 we needed to show that the Brenier map ($\nabla\phi$ from the proof) is Hölder continuous. In general, Brenier maps are not necessarily continuous. This leads us to the following question of *regularity* of optimal maps: under what conditions on the measures μ and ν, is the Brenier map Hölder continuous? There is a regularity theory developed by Caffarelli and Urbas for (solutions of) the Monge-Ampère equation which naturally provides regularity results for Brenier maps. Below we present the only two results (relating the Monge-Ampère equation to the theory of optimal transport) needed in this thesis. For more information on the Monge-Ampère equation, please consult Chapter 4 in [Vil03], a book by Gutiérrez [Gut01] and lecture notes by Urbas [Urb97].

The following result by Caffarelli shows that under certain conditions on μ and ν, Brenier maps (and hence, Brenier solutions) are Aleksandrov solutions of Eq. (6.4).

Theorem 6.2.3 (cf. Theorem 4.10 in [Vil03]). *Let μ and ν be two absolutely continuous probability measures on $\mathbb{R}^n$. Let f and g be their respective densities, and let X and Y be their respective supports. Let ϕ be a convex function such that $\nu = \nabla\phi\#\mu$. Assume that Y is convex and that g is positive a.e. on Y. Then $M\phi$ has no singular part on X. In particular, ϕ is an Aleksandrov solution of Eq.* (6.4).

Remark 6.2.4. Maximizers for the dual problem are called *Kantorovich potentials*. Theorem 5.1.2, Lemma 5.1.5 and Theorem 5.1.4 imply that Kantorovich potentials are as regular as corresponding Brenier solutions (which, under conditions of Theorem 6.2.3, are also Aleksandrov solutions). This allows us to apply Caffarelli's regularity theory to Brenier maps. In particular, we will require the following regularity result:

Theorem 6.2.5 (cf. Theorem 12.50 in [Vil09]). *Let Ω, Δ be connected bounded open subsets of $\mathbb{R}^n$. Let f, g be probability densities on Ω and Δ respectively, with f and g bounded from above and below. Let ϕ be a convex function such that its gradient is the Brenier map between measures corresponding to f and g. If Δ is convex, then $\phi_{|\Omega} \in C^{1,\alpha}(\Omega)$ for some $\alpha \in (0,1)$.*

6.2.2 Exhibiting a non-differentiable computable measure

In this section we will prove the following result.

Proposition 6.2.6. *Suppose $z \in [0,1]^n$ is not computably random. There exists an absolutely continuous computable probability measure μ on $\mathbb{R}^n$ such that:*

1. *its support is equal to $[0,1]^n$,*

2. *its density, $h : [0,1]^n \to \mathbb{R}$, is bounded away from 0 and $+\infty$, that is*

$$0 < C_1 < h < C_2$$

   for some $C_1, C_2 \in \mathbb{R}$ and

3. *$D_\lambda \mu(z)$ does not exist.*

Notation 6.2.7. For a given computable martingale M, we define *the corresponding computable probability measure μ_M^n* on $[0,1]^n$ by

$$\mu_M^n([\sigma]) = \lambda([\sigma]) \cdot M(\sigma)/M(\emptyset), \text{ for all } \sigma \in 2^{<\omega}.$$

When the dimension n is clear from the context, we may omit the superscript and write μ_M instead of μ_M^n.

A proof of a one-dimensional (weaker) variant of this Lemma is contained in the proof of Theorem 4.2 in [FKHNS14]. Below we summarize the relevant construction and then discuss how it can be adjusted so that it is useful in our proof of Proposition 6.2.6.
The main idea of the proof of Theorem 4.2 in [FKHNS14] is that it is possible to turn the success of one martingale into oscillations of another bounded martingale. In particular, let $z \in [0,1]$ be not computably random and let $Z \in 2^\omega$ be its binary expansion. There is a computable martingale B succeeding on Z. Freer et al. defined a computable martingale M, based on B, which repeats the following pattern of betting: M adds the capital that B risks, until its capital reaches 3, then M subtracts the capital that B risks, until its capital reaches 2, and so on. The resulting computable martingale is bounded from above and its value "oscillates" infinitely often along Z. That is, the following two conditions hold:

1. $1 \leq M(\sigma) \leq 4$ for all $\sigma \in 2^{<\omega}$, and

2. $\liminf M(Z\restriction_i) \leq 2$ and $\limsup M(Z\restriction_i) \geq 3$.

It is not difficult to see that the corresponding measure μ_M^1 on the unit interval satisfies the first two properties listed in Proposition 6.2.6. However, the third property (being not differentiable at z), does not necessarily hold. Consider quotients of the form $\frac{\mu_M^1(A)}{\lambda(A)}$, where $A \subseteq \mathbb{R}$ is Borel. The dyadic derivative of μ_M^1 at z is defined as the limit of quotients $\frac{\mu_M^1(A)}{\lambda(A)}$, where A ranges over the family of basic dyadic cubes containing z. This limit, clearly, does not exists. On the other hand, the symmetric derivative of μ_M^1 at z is defined as the limit of quotients $\frac{\mu_M^1(A)}{\lambda(A)}$, where A ranges over the family of (non-degenerate) balls centered at z. This limit might exist and therefore μ_M^1, as defined, satisfies only a weaker form of the third property: its dyadic derivative at z does not exist.

Now suppose $z \in [0,1]^n$ and let $Z \in 2^\omega$ be its binary expansion. Using the described construction from [FKHNS14], we can obtain a computable martingale M oscillating on Z and use it to define the corresponding probability measure μ_M^n. The measure μ_M^n has the same issue as μ_M^1 — it could be differentiable at z.

> *Remark* 6.2.8. Using the notion of *nicely shrinking sets* and Theorem 7.10 (see Section 7.9 in [Rud87]), it can be shown that z is not a Lebesgue point of the density of μ_M^n. However, this property is also not strong enough for our purposes.

In order to see how to fix this problem, consider what happens to μ_M^n when M does not bet for some stretch of Z. That is, $M(Z\restriction_i) = q$ for i ranging between j and $j+k$ for some j, k. Suppose $d(z, \partial[Z\restriction_j]) = \epsilon > 0$. If k is large enough, then (the value of) $\frac{\mu_M^n(B(z,\epsilon))}{\lambda(B(z,\epsilon))}$ can not be too different from (the value of) $\frac{\mu_M^n([Z\restriction_j])}{\lambda([Z\restriction_j])}$. The precise meaning of "not too different" will be given in the proof of Proposition 6.2.6. This observation suggests a conceptually straightforward fix: from time to time, M should stop betting and wait until $\frac{\mu_M^n([Z\restriction_j])}{\lambda([Z\restriction_j])}$ is not too different from $\frac{\mu_M^n(B(z,\epsilon))}{\lambda(B(z,\epsilon))}$ for some $\epsilon \in (0, 2^{-j}]$. We describe the formal details below.

Suppose $z \in [0,1]^n$ is contained in a basic dyadic cube D. If z is relatively far from the boundary of D, then there is a relatively large ball B centered at z also contained in D. In the following two proofs we will be interested in situations like that. For this reason we need to introduce a special notation, which we define below.

Notation 6.2.9. Fix n. Let $z \in [0,1]^n$ be an element with no dyadic components and let $Z \in 2^\omega$ be its binary expansion. For every $i \in \mathbb{N}$, define

$$d_{Z,i} = 2^i \cdot d\left(z, \partial[Z{\upharpoonright}_{ni}]\right).$$

$\partial[Z{\upharpoonright}_{ni}]$ is the boundary of the basic dyadic cube containing z with its side length equal to 2^{-i}. Thus $d_{Z,i}$ measures relative closeness of z to $\partial[Z{\upharpoonright}_{ni}]$.
For every $\epsilon \in \mathbb{R}$, define

$$l_{Z,\epsilon} = \{n \cdot j \mid d_{Z,j} > \epsilon\}.$$

Similarly, for all $\sigma \in 2^{<\omega}$ and all $i \in \mathbb{N}$ with $0 \leq ni \leq |\sigma|$, define

$$d_{\sigma,i} = 2^i \cdot d\left([\sigma], \partial[\sigma{\upharpoonright}_{ni}]\right),$$

$$l_{\sigma,\epsilon} = \{n \cdot j \mid n \cdot j \leq |\sigma| \text{ and } d_{\sigma,j} > \epsilon\}.$$

Observe that when ϵ is rational, the function $\sigma \mapsto l_{\sigma,\epsilon}$ is computable. Moreover, $l_{Z,\epsilon} = \cup_i l_{Z{\upharpoonright}_i,\epsilon}$ for all $Z \in 2^\omega$.

Proof of Proposition 6.2.6. Suppose $z \in [0,1]^n$ is not computably random. Let $(B_i)_{i\in\mathbb{N}}$ be a sequence of open balls defined by $B_j = B(z, 2^{-j})$. By Proposition 2.4.2, there exists $\hat{t} \in \{0, 1/3, 2/3\}^n$ and an infinite subsequence of $(B_i)_{i\in\mathbb{N}}$ such that whenever B_i is an element of this subsequence, $B_i - \hat{t}$ is contained in a basic dyadic cube belonging to $\mathcal{D}^n(k_i)$ for some k_i with $2^{-k_i} \leq 12 \cdot 2^{-i}$.
For every $t \in \{0, 1/3, 2/3\}^n$, define $\nu_t : [0,1]^n \to [0,1]^n$ by

$$\nu_t(x) = (\{(x-t)_1\}, \ldots, \{(x-t)_n\}), \text{ where } \{\cdot\} \text{ denotes the fractional part.}$$

Note that ν_t is an a.e. computable isomorphism.
Fix $t \in \{0, 1/3, 2/3\}^n$ and consider $\hat{z} = \nu_t(z)$. Suppose Proposition 6.2.6 holds for $\hat{z}$. That is, there exists a computable, absolutely continuous probability measure μ_t such that (1)-(3) hold. Define a measure μ by

$$\mu(A) = \mu_t(\nu_t(A)) \text{ for all Borel } A \subseteq [0,1]^n.$$

Then μ is a computable, absolutely continuous probability measure on $[0,1]^n$ with its support equal to $[0,1]^n$. Furthermore, μ is not differentiable at z. Therefore, Proposition 6.2.6 holds for z if it holds for $\nu_t(z)$ for any $t \in \{0, 1/3, 2/3\}^n$. In particular, it is sufficient to prove Proposition 6.2.6 for $\nu_{\hat{t}}(z)$ instead of z.

Let $\hat{z} = \nu_{\hat{t}}(z)$ and let $\hat{Z}$ be its binary expansion. Via considerations from the first paragraph of this proof, there exists a sequence of basic balls $(\hat{B}_i)_{i\in\mathbb{N}}$ centered at $\hat{z}$ such that for every i, $\hat{B}_i$ is contained in a basic dyadic cube belonging to $\mathcal{D}^n(k_i)$ with $2^{-k_i} \leq 12 \cdot r_i$, where r_i is the radius of $\hat{B}_i$. It follows that $l_{\hat{Z}} = l_{\hat{Z},2^{-4}}$ is infinite.
Pick $k \in \mathbb{N}$ such that $2^{k-4} > 5\sqrt{n}$ and define $q = 2^{kn+2}$. Let V_n be the volume of the unit ball in $\mathbb{R}^n$, so that $\lambda\left(B(x,r)\right) = V_n r^n$ for all $x \in \mathbb{R}^n$ and $r \geq 0$.
Suppose there exists a computable martingale M with the following properties:

(A1) $M(\emptyset) = 2$ and for all $\sigma \in 2^{<\omega}$, $1 \leq M(\sigma) \leq q + 1$,

(A2) for infinitely many $j \in l_{\hat{Z}}$,

$$M(\hat{Z}\upharpoonright_j) \leq 2, \text{ and} \tag{6.5}$$

(A3) for infinitely many $j \in l_{\hat{Z}}$,

$$M(\hat{Z}\upharpoonright_{j+i}) \geq q \text{ for all } i \text{ with } 0 \leq i \leq k \cdot n. \tag{6.6}$$

Let $\mu = \mu_M$ be the corresponding probability measure. Clearly, it is computable and absolutely continuous. Moreover, its support is equal to $[0,1]^n$ and its density is bounded away from 0 and from ∞.
Let $j \in l_{\hat{Z}}$ be such that (6.5) holds. Define $B = B(\hat{z}, 2^{-j/n-4})$ and $D = [\hat{Z}\upharpoonright_j]$, so that $B \subset D$. Then

$$\frac{\mu(B)}{\lambda(B)} \leq \frac{\lambda(D)}{\lambda(B)}\frac{\mu(D)}{\lambda(D)} = \frac{\lambda(D)}{\lambda(B)}M(\hat{Z}\upharpoonright_j) \leq 2\frac{2^{-j}}{V_n(2^{-j/n-4})^n} = \frac{2}{V_n}(16)^n.$$

Let $j \in l_{\hat{Z}}$ be such that (6.6) holds. Define $B = B(\hat{z}, 2^{-j/n-4})$ and $D = [\hat{Z}\upharpoonright_j]$, so that $B \subset D$. By Lemma 2.4.3, B contains a basic dyadic cube D_B with $l(D_B) \geq \frac{2^{-j/n-4}}{5\sqrt{n}} > 2^{-j/n-k}$ and $\hat{z} \in D_B$. Then $D_B = [\hat{Z}\upharpoonright_{j+k_1 n}]$ for some k_1 with $0 \leq k_1 \leq k$. We get

$$\frac{\mu(B)}{\lambda(B)} \geq \frac{\mu(D_B)}{\lambda(B)} = \frac{\lambda(D_B)}{\lambda(B)}\frac{\mu(D_B)}{\lambda(D_B)} \geq \frac{2^{-j-k_1 n}}{V_n(2^{-j/n-4})^n}M(\hat{Z}\upharpoonright_{j+k_1 n}) \geq \frac{2^{(4-k)n}q}{V_n} = \frac{2^{(4-k)n}2^{kn+2}}{V_n} = \frac{4}{V_n}(16)^n.$$

It follows that μ is not differentiable at $\hat{z}$. In order to complete the proof, we need to show that a computable martingale M satisfying (A1)-(A3) exists. Proposition 6.2.10, below, shows this.

□

Proposition 6.2.10. *Let $z \in [0,1]^n$ be an element with no dyadic components and let $Z \in 2^\omega$ be its binary expansion. Let ϵ be a positive rational number. Suppose z is not computably random and $l_Z = l_{Z,\epsilon}$ is infinite. Let q be a dyadic rational with $2 < q$ and let $k \in \mathbb{N}$. There exists a computable martingale $M_{q,k}$ such that:*

1. *for all $\sigma \in 2^{<\omega}$, $1 \leq M_{q,k}(\sigma) \leq q + 1$,*
2. *for infinitely many $j \in l_Z$, for all $i \in \mathbb{N}$ with $0 \leq i < k$, $M_{q,k}(Z\upharpoonright_{j+i}) \leq 2$, and*
3. *for infinitely many $j \in l_Z$, for all $i \in \mathbb{N}$ with $0 \leq i < k$, $M_{q,k}(Z\upharpoonright_{j+i}) \geq q$.*

Proof. Since z is not computably random, some computable martingale M succeeds on Z. We build $M_{q,k}$ from M.

As mentioned in Remark 2.5.4, we may assume that M only takes positive rational values which can be computed in a single output from the input string. We may also assume that M has the savings property

$$M(\sigma\eta) \geq M(\sigma) - 1 \text{ for each strings } \sigma, \eta.$$

The martingale $M_{q,k}$ has to satisfy the three conditions from the statement of this result. To satisfy the first condition, we turn the success of M into oscillation of $M_{q,k}$ so that the capital of $M_{q,k}$ is always between 1 and $q+1$. At each σ, the martingale $M_{q,k}$ is in one of three possible phases. In the *up phase*, it adds the capital that M risks, until its value $M_{q,k}(\sigma)$ reaches q (if this value would exceed q, $M_{q,k}$ adds less in order to ensure the value equals q). In the *waiting phase*, $M_{q,k}$ does not bet until a certain condition is satisfied. In the *down phase*, $M_{q,k}$ subtracts the capital that M risks, until the value $M_{q,k}(\sigma)$ reaches 2.

To satisfy the second (respectively, third) condition, our construction ensures that $M_{q,k}$ infinitely often is in the waiting phase with its capital fixed at 2 (respectively, q).

The construction of $M_{q,k}$ is as follows.

We will make use of the following notation. For $m, i \in \mathbb{N}$ and $\sigma \in 2^{<\omega}$, define

$$l_\sigma[m,i] = \{j \in l_{\sigma,\epsilon} \mid m \leq j \leq i\}.$$

Observe that $l_\sigma[m,i]$ is computable uniformly in m, i and σ.

Inductively, we show that if $M_{q,k}$ is in the up phase at σ, then $M_{q,k}(\sigma) < q$, if $M_{q,k}$ is in the down phase at σ, then $M_{q,k}(\sigma) > 2$ and if it is in the waiting phase, either $M_{q,k}(\sigma) = q$ or $M_{q,k}(\sigma) = 2$. At the empty string $\emptyset$, the martingale $M_{q,k}$ is in the up phase and $M_{q,k}(\emptyset) = 2$. Thus the inductive condition holds at the empty string.
For every $\sigma \in 2^{<\omega}$, let phase$(\sigma) \in \{$up,wait,down$\}$ denote the phase of $M_{q,k}$ at σ. By phase(σ) we denote the longest prefix of σ where $M_{q,k}$ changed from one phase to another:

$$\text{start}(\sigma) = \sigma\restriction_{\max\{i \leq |\sigma| \ \mid \ \text{phase}(\sigma\restriction_i) \neq \text{phase}(\sigma\restriction_{i-1})\}}.$$

Suppose now that $M_{q,k}(\sigma)$ has been defined.

Case 1: $M_{q,k}$ is in the up phase at σ. For all $j \in \{0,1\}$, let

$$r_j = M_{q,k}(\sigma) + M(\sigma j) - M(\sigma).$$

If $r_0, r_1 < q$ then let $M_{q,k}(\sigma j) = r_j$; stay in the up phase at both $\sigma 0$ and $\sigma 1$; let start$(\sigma 0) =$ start$(\sigma 1) = \sigma$. Otherwise, since M is a martingale and $M_{q,k}(\sigma) < q$, there is a unique j such that $r_j \geq q$. Let $M_{q,k}(\sigma j) = q$ and $M_{q,k}(\sigma(1-j)) = 2M_{q,k}(\sigma) - q$.

Go into the waiting phase at σj, but stay in the up phase at $\sigma(1-j)$. Note that the inductive condition is maintained at both $\sigma 0$ and $\sigma 1$.

Case 2: $M_{q,k}(\sigma)$ is in the down phase.
For all $j \in \{0,1\}$, let

$$r_j = M_{q,k}(\sigma) - (M(\sigma j) - M(\sigma)).$$

If $r_0, r_1 > 2$ then let $M_{q,k}(\sigma 1) = r_1$ and $M_{q,k}(\sigma 0) = r_0$; stay in the down phase at both $\sigma 0$ and $\sigma 1$. Otherwise, since M is a martingale and $M_{q,k}(\sigma) > 2$, there is a unique j such that $r_j \leq 2$. Let $M_{q,k}(\sigma j) = 2$ and $M_{q,k}(\sigma(1-j)) = 2M_{q,k}(\sigma) - 2$. Go into the waiting phase at σj, but stay in the down phase at $\sigma(1-j)$. The inductive condition is maintained at both $\sigma 0$ and $\sigma 1$.

Case 3: $M_{q,k}(\sigma)$ is in the waiting phase. Let $\Delta = l_\sigma[|\text{start}(\sigma)|, |\sigma|]$. If Δ is not empty, let j be its least element. If the *waiting condition*

$$W(\sigma) = \Delta \text{ is empty or } j < |\sigma| - k, \tag{6.7}$$

is satisfied, stay in the waiting phase and don't bet. That is, let

$$M_{q,k}(\sigma 1) = M_{q,k}(\sigma 0) = M_{q,k}(\sigma).$$

Otherwise,

- if $M_{q,k}(\sigma) = q$, then go into the down phase at $\sigma 0$ and $\sigma 1$,
- else go into the up phase at $\sigma 0$ and $\sigma 1$.

Since in the waiting phase $M_{q,k}$ does not bet and $M_{q,k}$ can only go into the waiting phase when its capital is equal either to 2 or q, the inductive condition is maintained at both $\sigma 0$ and $\sigma 1$.

Claim 6.2.11. *For each string τ,*

$$1 \leq M_{q,k}(\tau) \leq q + 1.$$

If $M_{q,k}$ is in the waiting phase at τ, then either $M_{q,k}(\tau) = q$ or $M_{q,k}(\tau) = 2$.

Suppose that $M_{q,k}$ is in the up phase at τ. Then $M_{q,k}(\tau) \leq q$. For the lower bound on $M_{q,k}(\tau)$, suppose that $M_{q,k}$ entered the up phase at the string $\sigma \preceq \tau$ with $|\sigma|$ maximal. Then $M_{q,k}(\sigma) = 2$. By the savings property we have $M(\tau) - M(\sigma) \geq -1$. Therefore $M_{q,k}(\tau) = M_{q,k}(\sigma) + M(\tau) - M(\sigma) \geq 1$.

Next suppose that $M_{q,k}$ is in the down phase at τ. Then $M_{q,k}(\tau) \geq 2$. For the upper bound on $M_{q,k}(\tau)$, suppose that $M_{q,k}$ entered the down phase at the string $\sigma \preceq \tau$ with $|\sigma|$ maximal. Then $M_{q,k}(\sigma) = q$. By the savings property we have $M_{q,k}(\tau) = M_{q,k}(\sigma) - (M(\tau) - M(\sigma)) \leq q + 1$. □

Claim 6.2.12. *$M_{q,k}$ oscillates between up and down phases infinitely often along Z. That is, for infinitely many j, $M_{q,k}$ is in the down phase at $Z\restriction_j$ and for infinitely many i, $M_{q,k}$ is in the up phase at $Z\restriction_i$.*

To prove the claim, it is sufficient to demonstrate that $M_{q,k}$, along Z, does not stay in either of the three phases indefinitely.
Firstly suppose $M_{q,k}$ is in the up phase at $Z\restriction_i$ for some i. In the up phase, $M_{q,k}$ mimics M until its capital reaches q and then goes into the waiting phase. By our assumption, M succeeds on Z, hence for some j, $M_{q,k}(Z\restriction_{i+j}) = q$ with $M_{q,k}$ going into the waiting phase at $Z\restriction_{i+j}$.
Next suppose $M_{q,k}(Z\restriction_i)$ is in the waiting phase for some i with $\text{start}(Z\restriction_i) = Z\restriction_i$. Observe that

$$\{j \in l_Z \ : \ j \geq i\} = \bigcup_{j \in \mathbb{N}} l_{Z\restriction_{i+j}}[i, i+j],$$

and

$$l_{Z\restriction_{i+j}}[i, i+j] \subseteq l_{Z\restriction_{i+j+m}}[i, i+j+m] \text{ for all } i, m, j.$$

It is easy to see that $l_{Z\restriction_i}[i, i]$ is empty and $W(Z\restriction_i)$ is satisfied. Since l_Z is infinite, there exist the least j such that $l_{Z\restriction_{i+j}}[i, i+j]$ is not empty. Then for some $k_1 \leq k$, $W(Z\restriction_{i+j+k_1})$ does not hold and hence $M_{q,k}$ will leave the waiting phase eventually.
Finally, suppose $M_{q,k}(Z\restriction_i)$ is in the down phase for some i. In the down phase, $M_{q,k}$ subtracts the capital that M risks, until the value $M_{q,k}(\sigma)$ reaches 2 and then goes into the waiting phase. Since M succeeds on Z, it is clear that every time $M_{q,k}$ is in the down phase at $Z\restriction_i$, there exists j such that, $M_{q,k}$ goes into the waiting phase at $Z\restriction_{i+j}$. □

Claim 6.2.13. *For infinitely many $j \in l_Z$, for all $t \in \mathbb{N}$ with $0 \leq t < k$,*

$$M_{q,k}(Z\restriction_{j+t}) = q.$$

From Claim 6.2.12 we know that $M_{q,k}$ goes from the up phase into the waiting phase infinitely often along Z. Suppose $M_{q,k}(Z\restriction_i)$ is in the waiting phase for some i with $\text{start}(Z\restriction_i) = Z\restriction_i$ and $M_{q,k} = q$. Then $M_{q,k}$ waits (with its capital fixed at q) until $l_{Z\restriction_j}[i, j]$ is not empty for some $j > i$. Let j be the minimal such number and let $m = \min l_{Z\restriction_j}[i, j]$. Then $m \in l_Z, m \geq i$ and, by our construction, for all $t \in \mathbb{N}$ with $0 \leq t < k$, $M_{q,k}(Z\restriction_{m+t}) = q$. □
By analogous considerations the following claim holds.

Claim 6.2.14. *For infinitely many $j \in l_Z$, for all $t \in \mathbb{N}$ with $0 \leq t < k$,*

$$M_{q,k}(Z\restriction_{j+t}) = 2.$$

This complete the proof of Proposition 6.2.10. □

6.3 Consequences of the main theorem

6.3.1 Monotone and convex functions

We start with a result that follows straightforwardly from Theorem 6.0.1.

Theorem 6.3.1. *Let $z \in \mathbb{R}^n$. The following are equivalent:*

1. *z is computably random,*
2. *every a.e. computable monotone function $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable at z, and*
3. *every computable convex function $f : \mathbb{R}^n \to \mathbb{R}$ is twice-differentiable at z.*

Proof. (1) $\implies$ (2) follows from Theorem 4.6.3. (1) $\implies$ (3) follows from Theorem 4.5.5.

To prove (2) $\implies$ (1) and (3) $\implies$ (1), let $z \in [0,1]^n$ be not computably random. Via Theorem 6.0.1, there is a computable convex function $\phi : [0,1]^n \to \mathbb{R}$ such that $\nabla\phi$ is not differentiable at z. Moreover, the continuous extension of $\nabla\phi$ to $[0,1]^n$ (which, for the sake of simplicity, we will also denote by $\nabla\phi$) is computable.
Define $\hat{\phi} : \mathbb{R}^n \to \mathbb{R}$ by

$$\hat{\phi}(x) = \sup_{y\in[0,1]^n} \{\phi(y) + \langle\nabla\phi(y), x - y\rangle\} = \max_{y\in[0,1]^n} \{\phi(y) + \langle\nabla\phi(y), x - y\rangle\}.$$

Let K be the Lipschitz constant of ϕ on $[0,1]^n$. Then functions

$$x \mapsto \phi(y) + \langle\nabla\phi(y), x - y\rangle$$

are affine and $K-$Lipschitz. Hence $\hat{\phi}$, being a sup of a family of affine $K-$Lipschitz functions, is both convex and $K-$Lipschitz on $\mathbb{R}^n$. Moreover, since $\hat{\phi}(x)$ is computable uniformly in $x \in \mathbb{Q}^n$ and $\hat{\phi}$ is Lipschitz, $\hat{\phi}$ is a computable convex function. Let us show that $\hat{\phi}$ coincides with ϕ on $[0,1]^n$. Via Proposition 5.4 [ET99], we have

$$\phi(x) - \phi(y) \geq \langle\nabla\phi(y), x - y\rangle \quad \text{for all } x, y \in [0,1]^n. \tag{6.8}$$

Rewriting (6.8) we have

$$\phi(x) \geq \max_{y\in[0,1]^n} \{\phi(y) + \langle\nabla\phi(y), x - y\rangle\} \quad \text{for all } x \in [0,1]^n.$$

It follows that $\phi \geq \hat{\phi}$ on $[0,1]^n$. From the definition of $\hat{\phi}$ we also know that $\phi \leq \hat{\phi}$ on $[0,1]^n$. Hence $\hat{\phi} = \phi$ on $[0,1]^n$. This shows that $\hat{\phi} : \mathbb{R}^n \to \mathbb{R}$ is a computable convex function not twice-differentiable at z. Then, by Proposition 4.5.1, $\nabla\hat{\phi}$ is a monotone a.e. computable function not differentiable at z. □

6.3.2 Probability measures on $[0,1]^n$

Theorem 6.1.3 can not be applied to arbitrary monotone functions, since those are not necessarily continuous. However, there is an analogous result specifically for subgradients of convex functions (it is Theorem A.2 in [McC97]):

Theorem 6.3.2 (Jacobian theorem for monotone maps)**.** *Let ϕ be a convex function on $\mathbb{R}^n$ and suppose it is twice differentiable at $x \in \mathbb{R}^n$. Then*

$$\lim_{r\to 0} \frac{\lambda\left(\partial\phi(B(x,r))\right)}{\lambda\left(B(x,r)\right)} = \det D^2\phi(x).$$

With the results we have proven so far, we can easily get the following characterisation of computable randomness in $[0,1]^n$ in terms of differentiability of absolutely continuous computable probability measures on $[0,1]^n$.

Theorem 6.3.3. *Let $z \in [0,1]^n$.*

z is computably random $\iff$

every absolutely continuous computable probability measure on $[0,1]^n$ is differentiable at z.

Proof $\Rightarrow$. Let μ be an absolutely continuous computable probability measure on $[0,1]^n$ and let $z \in [0,1]^n$. We may assume z belongs to the interior of $[0,1]^n$. Suppose $D_\lambda\mu(z)$ does not exist.

Define a probability measure $\hat{\mu}$ on $[0,1]^n$ by $\hat{\mu} = \frac{1}{2}(\mu + \lambda_n)$. It is computable, absolutely continuous and its support is equal to $[0,1]^n$. Moreover, $D_\lambda\hat{\mu}(z)$ does not exist.

By the effective Brenier theorem, there exists a computable convex function $\phi : [0,1]^n \to \mathbb{R}$ such that $\lambda_n = \nabla\phi\#\hat{\mu}$. Let f be the density of $\hat{\mu}$. By Theorem 6.2.3, ϕ is an Aleksandrov solution of

$$\det D^2\phi(x) = \frac{f(x)}{1_{[0,1]^n} \circ \nabla\phi(x)}.$$

This means that the density of $M\phi$ coincides with f on $[0,1]^n$. Hence $M\phi$ coincides with $\hat{\mu}$ on $[0,1]^n$. Since $z \in (0,1)^n$, we have

$$\lim_{r\to 0} \frac{\lambda\left(\partial\phi(B(z,r))\right)}{\lambda\left(B(z,r)\right)} = \lim_{r\to 0} \frac{M\phi(B(z,r))}{\lambda\left(B(z,r)\right)} = \lim_{r\to 0} \frac{\hat{\mu}(B(z,r))}{\lambda\left(B(z,r)\right)} = D_\lambda\hat{\mu}(z).$$

We know $D_\lambda\hat{\mu}(z)$ does not exists and thus by Theorem 6.3.2 ϕ is not differentiable at z. Just like in the proof of Theorem 6.3.1, ϕ can be extended to a computable convex function on $\mathbb{R}^n$ and then by Theorem 4.5.5 z is not computably random. □

Proof ⇐. Follows from Proposition 6.2.6. □

Remark 6.3.4. This theorem can be reformulated as an effective version of the Lebesgue differentiation theorem for functions that are densities of computable absolutely continuous measures on $[0,1]^n$. It is not known if there is a more natural (from computability theory point of view) characterization of such functions.

An analogous characterisation of Schnorr randomness has been proven in [Rut13] (for measures) and [PRS14] (for functions on $\mathbb{R}^n$). In that case the corresponding class of effective functions is very natural (L^1-computability), while the class of corresponding probability measures is somewhat artificial.

6.3.3 Critical values of computable monotone Lipschitz functions

Theorem 6.3.5. *Let $z \in \mathbb{R}^n$. The following are equivalent:*

1. *z is computably random, and*
2. *z is not a critical value of any computable monotone Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^n$.*

Proof ⇒. Follows from Theorem 4.3.2. □

Proof ⇐. Suppose $z \in \mathbb{R}^n$ is not computably random. By Theorem 6.3.1, there is an a.e. computable monotone function $u : \mathbb{R}^n \to \mathbb{R}^n$ not differentiable at z. Then

$$f = (u + I)^{-1}$$

is a computable Lipschitz function. Theorem 2.7.23 implies that $f(u(z) + z) = z$ is a critical value of f.

□

6.3.4 The Monge-Ampère equation

Robert McCann in [McC97] demonstrated that the Monge-Ampère equation of the type associated with optimal transport holds true almost everywhere. The following theorem is an effective variant of this result.

Theorem 6.3.6. *Consider the following Monge-Ampère equation:*

$$\det D^2\phi = \frac{D_\lambda \mu}{1_{[0,1]^n} \circ \nabla \phi}. \tag{6.9}$$

Let $z \in [0,1]^n$. The following are equivalent:

1. *z is computably random, and*
2. *for every computable absolutely continuous probability measure μ on $\mathbb{R}^n$ with supp $(\mu) = [0,1]^n$, if $\phi : [0,1]^n \to \mathbb{R}$ is a Brenier solution of (6.9), then the following holds:*

$$\det D^2\phi(z) = \frac{D_\lambda \mu}{1_{[0,1]^n} \circ \nabla \phi}(z), \tag{6.10}$$

3. *for every computable absolutely continuous probability measure μ on $\mathbb{R}^n$ with supp $(\mu) = [0,1]^n$, if $\phi : [0,1]^n \to \mathbb{R}$ is a computable Brenier solution of (6.9), then (6.10) holds.*

Proof (1) $\implies$ (2). Let $z \in [0,1]^n$. Let μ be a computable absolutely continuous probability measure with supp $(\mu) = [0,1]^n$. Let ϕ be a Brenier solution of (6.9). Suppose (6.10) doesn't hold. We will show that z is not computably random. We may assume z belongs to the interior of $[0,1]^n$.

By the effective Brenier theorem and by Lemma 5.2.1, we may assume ϕ is computable on $[0,1]^n$. Since ϕ is computable and convex, $\nabla\phi$ and $\det D^2\phi$ are well defined on all computably random elements of $[0,1]^n$. Hence we may assume $\nabla\phi(z)$ and $\det D^2\phi(z)$ are well defined. Since μ is absolutely continuous and computable, if $D_\lambda\mu(z)$ does not exist, then by Theorem 6.3.3 z is not computably random. Suppose it does exist.

Since ϕ is a Brenier solution of (6.9), by Theorem 6.2.3, ϕ is also an Aleksandrov solution of (6.9) hence $\frac{D_\lambda \mu}{1_{[0,1]^n} \circ \nabla \phi}$ coincides with the density of $M\phi$ on the support of μ. Moreover, $\nabla\phi(\text{supp}\,(\mu)) \subseteq [0,1]^n$ and hence $\frac{D_\lambda \mu}{1_{[0,1]^n} \circ \nabla \phi} = D_\lambda\mu$ and then $M\phi = \mu$ on the support of μ. Then

$$\frac{D_\lambda \mu}{1_{[0,1]^n} \circ \nabla \phi}(z) = D_\lambda\mu(z) = \lim_{r\to 0} \frac{\mu(B(z,r))}{\lambda\,(B(z,r))} = \lim_{r\to 0} \frac{M\phi(B(z,r))}{\lambda\,(B(z,r))} = \lim_{r\to 0} \frac{\lambda(\partial\phi(B(z,r)))}{\lambda\,(B(z,r))}.$$

By our assumption,

$$\lim_{r\to 0} \frac{\lambda(\partial\phi(B(z,r)))}{\lambda\,(B(z,r))} \neq \det D^2\phi(z).$$

Hence, by Theorem 6.3.2, ϕ is not twice differentiable at z and thus z is not computably random.

□

Proof (2) $\implies$ (3). Trivial. □

Proof (3) $\implies$ (1). Suppose $z \in [0,1]^n$ is not computably random. By Theorem 6.3.3 there exists a computable absolutely continuous probability measure μ on $[0,1]^n$ not differentiable at z. We may assume $\operatorname{supp}(\mu) = [0,1]^n$. Since μ is not differentiable at z, (6.10) doesn't hold.

□

Notation index

General

$2^{<\omega}$	set of strings of zeros and ones
2^{ω}	Cantor space, the space of infinite binary strings with the product topology
$\sigma\tau$	concatenation of σ and τ
σa	σ followed by the symbol a
$\sigma \preceq \tau$	σ is a prefix of τ
$\lvert\sigma\rvert$	the length of σ
$Z\restriction_n$	$Z(0)Z(1)\dots Z(n-1)$
$\emptyset$	the empty string
$\emptyset$	the empty set
$\#X$	cardinality of the set X
$A \oplus B$	$\{2n\colon\ n \in A\} \cup \{2n+1\colon\ n \in B\}$, effective disjoint union
$\mathbb{N}$	set of natural numbers $0, 1, 2\dots$
$\mathbb{N}^{+}$	set of positive natural numbers
$\mathbb{R}$	set of real numbers
$\mathbb{Q}$	set of rational numbers
$\mathbb{R}_0^{+}$	set of non-negative real numbers
λ	Lebesgue measure (either on 2^{ω} or on $\mathbb{R}^n$)
λ_n	restriction of the Lebesgue measure to $[0,1]^n$
$\mathrm{supp}\,(\mu)$	support of the measure μ
$\overline{X}$	closure of X
∂X	boundary of X
$int(A)$	interior of A
$B(x,r)$	open metric ball centered at x with radius r
$\overline{B}(x,r)$	closed metric ball centered at x with radius r
$\langle x, y\rangle$	dot product
Γf	the graph of function f, for monotone set-valued functions see Notation 2.7.8
$\lvert x\rvert$	euclidean norm of $x \in \mathbb{R}^n$
$d(x,A)$	distance between $x \in \mathbb{R}^n$ and $A \subset \mathbb{R}^n$, that is $\inf_{y\in A} \lvert x-y\rvert$
$d(A,B)$	distance between two subsets of $\mathbb{R}^n$, that is $\inf_{y\in A, x\in B} \lvert x-y\rvert$
S^n	$n-$sphere
$e_1, \dots, e_n$	canonical orthonormal basis of $\mathbb{R}^n$
$\lVert f\rVert_\infty$	supremum norm
1_A	indicator function of a set A

Preliminaries

$(x)_i$	$i-$th element of a Cauchy name, see Notation 2.1.3
CN_x	the set of all Cauchy names for x, see Definition 2.1.1
$(f(x))_t$	approximate value of $f(x)$, see Notation 2.1.7
$P(X)$	set of all probability measures on X, see Notation 2.3.1
$\pi(\mu,\nu)$	Prokhorov distance between μ and ν, see Definition 2.3.2
$\mathbb{D}_i^n$	set of points in $\mathbb{R}^n$ with all coordinates of the form $k2^{-i}$ for some integer k, see Section 2.4
$\mathbb{D}_i^{n,S}$	the set $\mathbb{D}_i^n \cap S^n$, see Section 2.4
$\mathbb{D}_*^n$	$\cup_i \mathbb{D}_i^n$, see Section 2.4
$\mathbb{D}_*^{n,S}$	$\cup_i \mathbb{D}_i^{n,S}$, see Section 2.4
$\mathcal{D}^n$	set of all half-open basic dyadic cubes in $\mathbb{R}^n$, see Section 2.4
$l(D)$	side length of a cube D, see Section 2.4
$\mathcal{D}^n(k)$	those cubes D in $\mathcal{D}^n$ with $l(D)=2^{-k}$, see Section 2.4
$\mathcal{D}^n(i,x)$	the unique element of $\mathcal{D}^n(i)$ containing x, see Section 2.4
$p_i^n(Z)$	binary subsequence of Z, defined as $\{Z(kn+i)\ :\ k\in\mathbb{N}\}$, see Definition 2.5.9
$0.Z$	an element of $[0,1]^n$ whose binary expansion is Z, see Definition 2.5.9
$[\sigma]$	depending on the context, either a clopen in 2^ω, or a finite union of dyadic cubes in $[0,1]^n$, see Notation 2.5.10
$[\sigma]_p$	a shifted dyadic cube, that is $[\sigma]+p$, see Notation 2.5.12
M^A	oracle martingale M with oracle A, see Definition 2.5.5
$Df(x)$	derivative of f at x, see Definition 2.6.1, Definition 2.6.3 and Definition 2.6.6
$\overline{D}f(x)$	upper derivative of f at x, see Definition 2.6.2
$\underline{D}f(x)$	lower derivative of f at x, see Definition 2.6.2
$\overline{D}_2f(x)$	upper dyadic derivative of f at x, see Definition 2.6.2
$\underline{D}_2f(x)$	lower dyadic derivative of f at x, see Definition 2.6.2
$\overline{D}_+f(x)$	right-sided derivative of f at x, see Definition 2.6.2
$\underline{D}_-f(x)$	left-sided derivative of f at x, see Definition 2.6.2
$D_if(x)$	ith partial derivative of f at x, see Definition 2.6.4
$\nabla f(x)$	gradient of f at x, see Definition 2.6.4
$Df(x;v)$	directional derivative of f at x in the direction of v, see Definition 2.6.5
$Df_+(x;v)$	one-sided directional derivative of f at x in the direction of v, see Definition 2.6.5
N_f	non-differentiability set of f, see Notation 2.6.8
$\mathbf{Lip}(f)$	Lipschitz constant of f, see Definition 2.7.1
$\text{epi}f$	epigraph of f, see Definition 2.7.4
$\text{Dm}\ u$	domain of a monotone function u, see Notation 2.7.8
∂u	subdifferential of u, see Definition 2.7.16
I	identity function on $\mathbb{R}^n$
f^*	Fenchel conjugate of f, see Definition 2.7.26
Mu	Monge-Ampère measure associated with u, see Section 2.7.7
$D_\lambda\mu(x)$	symmetric derivative of μ, see Definition 2.7.28
$D_2\mu(x)$	dyadic derivative of μ, see Definition 2.7.28
$\overline{D}_2\mu(x)$	upper dyadic derivative of μ, see Definition 2.7.28
$\underline{D}_2\mu(x)$	lower dyadic derivative of μ, see Definition 2.7.28

Chapter 3

$S_f(\sigma)$	slope of f at the dyadic interval $[\sigma]$, see Notation 3.1.1
$S_f(a,b)$	slope of f at the interval (a,b), see Notation 3.1.1
μ_M	measure corresponding to martingale M, see Notation 3.1.2
M_μ	martingale corresponding to measure μ, see Notation 3.1.2
M_f	martingale corresponding to function f, see Notation 3.1.3
$\mathbf{por}_2(A,B)$	dyadic porosity constant, see Definition 3.1.4
$\mathcal{A}(\mu)$	set of atoms of measure μ, see Notation 3.3.21

Chapter 4

$\Theta_{u\to v}$	change of basis map with $\Theta_{u\to v}(u) = v$, see Section 4.1.2

Chapter 5

$T\#\mu$	measure ν defined by $\nu(A) = \mu(T^{-1}(A))$, see Definition 5.0.1
$\Pi(\mu,\nu)$	set of all admissible transference plans between μ and ν, see Section 5.1.1
$I_c[\pi]$	total transportation cost associated with measure π, defined in Eq. (5.2)
$I_c[T]$	total transportation cost associated with map T, defined in Eq. (5.3)
$\mathbb{I}_c[\mu,\nu]$	total transportation cost between μ and ν, defined as $\inf_{\pi\in\Pi(\mu,\nu)} I_c[\pi]$, see Section 5.1.1
$M_2(\mu,\nu)$	defined as $\int_{\mathbb{R}^n} \frac{\lvert x\rvert^2}{2}\,\mathrm{d}\mu(x) + \int_{\mathbb{R}^n} \frac{\lvert x\rvert^2}{2}\,\mathrm{d}\nu(x)$, see Notation 5.1.1
$J_{\mu,\nu}(f,g)$	defined as $\int_{\mathbb{R}^n} f\,\mathrm{d}\mu + \int_{\mathbb{R}^n} g\,\mathrm{d}\nu$, see Notation 5.1.1
$J(f,g)$	short for $J_{\mu,\nu}(f,g)$, see Notation 5.1.1
Φ_c	set of pairs of integrable functions ϕ,ψ with $\phi(x)+\psi(y)\le c(x,y)$ for almost all x,y, defined in Theorem 5.1.4
L_K	$\{f\in C([0,1]^n,\mathbb{R}) : \mathbf{Lip} f\le K \text{ and } \lvert f\rvert\le K\}$, see Section 5.3
$\underline{f}$	lower interpolant of f, defined in Definition 5.3.3
$(\hat\gamma_i)_{i\in\mathbb{N}}$	computable sequence of piecewise affine functions defined in Definition 5.3.6
$\mathbf{L}_K$	computable metric space $(L_K, d_\infty, (\hat\gamma_i)_{i\in\mathbb{N}})$ defined in Proposition 5.3.7
$W_p(\mu,\nu)$	Wasserstein distance between μ and ν, defined as $\mathbb{I}_p(\mu,\nu)^{1/p}$, see Definition 5.4.1

Chapter 6

$D^2\phi$	second derivative of ϕ
μ_M^n	probability measure on $[0,1]^n$ corresponding to martingale M, defined by $\mu_M^n([\sigma]) = \lambda([\sigma])\cdot M(\sigma)/M(\emptyset)$, see Notation 6.2.7

Bibliography

[AA99] G. Alberti and L. Ambrosio. A geometrical approach to monotone functions in $\mathbb{R}^n$. *Mathematische Zeitschrift*, 230:259–316, 1999. [25, 31, and 55]

[AG13] L. Ambrosio and N. Gigli. *A User's Guide to Optimal Transport*, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. [73]

[Ale39] A. D. Aleksandrov. Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. *Leningrad Univ. Ann. (Math. ser.)*, (6):3–35, 1939. [54]

[APC11] G. Alberti, D. Preiss, and M. Csörnyei. Differentiability of Lipschitz functions, structure of null sets, and other problems. In *Proceedings of the International Congress of Mathematicians 2010*, pages 1379–1394. World Scientific Publishing, 2011. [7]

[BDH+12] L. Bienvenu, A. R. Day, M. Hoyrup, I. Mezhirov, and A. Shen. A constructive version of Birkhoff's ergodic theorem for martin-löf random points. *Inf. Comput.*, 210:21–30, 2012. [4]

[Bel07] G. Beliakov. Smoothing Lipschitz functions. *Optimization Methods and Software*, 22(6):901–916, 2007. [81]

[BHMN14] L. Bienvenu, R. Hölzl, J. Miller, and A. Nies. Denjoy, Demuth and density. *Journal of Mathematical Logic*, 14(1), 2014. [6]

[BLPT86] A. M. Bruckner, M. Laczkovich, G. Petruska, and B. S. Thomson. Porosity and approximate derivati ves. *Canadian Journal of Mathematics*, 38:1149–1180, 1986. [35]

[BMN16] V. Brattka, J. Miller, and A. Nies. Randomness and differentiability. *Transactions of the AMS*, 368:581–605, 2016. arXiv version at arxiv.org/abs/1104.4465. [1, 6, 7, 35, 38, 40, 46, 50, 54, and 90]

[Bog07] V.I. Bogachev. *Measure Theory*. Number v. 1 in Measure Theory. Springer Berlin Heidelberg, 2007. [51]

[BP03] V. Brattka and G. Presser. Computability on subsets of metric spaces. *Theoretical Computer Science*, 305(1):43 – 76, 2003. Topology in Computer Science. [13]

[Bra01] V. Brattka. *Computable Versions of Baire's Category Theorem*, pages 224–235. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001. [12]

[Bra08] V. Brattka. Plottable real number functions and the computable graph theorem. *SIAM J. Comput.*, 38:303–328, 2008. [13 and 69]

[Bre87] Y. Brenier. Polar decomposition and increasing rearrangement of vector fields. *C.R. Acad. Sci. Paris Sér. I Math.*, (305):805–808, 1987. [73]

[Bru78] A. M. Bruckner. *Differentiation of real functions*, volume 659 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978. [25 and 32]

[BT84] A. M. Bruckner and B. S. Thomson. Porosity estimates for the Dini derivatives. *Real Anal. Exch.*, 9:508–538, 1984. [35]

[BV10] J. M. Borwein and J. D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010. [25, 27, and 32]

[CAS07] J.P. Crouzeix, E. Ocaña Anaya, and W. Sosa Sandoval. A construction of a maximal monotone extension of a monotone map. *ESAIM, Proc.*, 20:93–104, 2007. [70]

[CC87] Y. Chabrillac and J.-P. Crouzeix. *Continuity and differentiability properties of monotone real functions of several real variables*, pages 1–16. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987. [5]

[DH10] R. Downey and D. Hirschfeldt. *Algorithmic randomness and complexity*. Springer-Verlag, Berlin, 2010. 855 pages. [1 and 17]

[DK89] Ding-Zhu Du and K. Ko. Computational complexity of integration and differentiation of convex functions. *Systems Science and Mathematical Sciences*, 2(1):70–79, 1989. [41 and 46]

[EG92] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, Florida, 1992. [25 and 26]

[ET99] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*, volume 28 of *Classics in Applied Mathematics*. SIAM, 1999. [25 and 103]

[Eva98] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998. [77]

[FGMN12] J. Franklin, N. Greenberg, J. Miller, and K. M. Ng. Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. *Proc. Amer. Math. Soc.*, 140(10):3623–3628, 2012. [4]

[FKHNS14] C. Freer, B. Kjos-Hanssen, A. Nies, and F. Stephan. Algorithmic aspects of Lipschitz functions. *Computability*, 3(1):45–61, 2014. [1, 6, 8, 41, 54, 96, and 97]

[FP09] T. Fowler and D. Preiss. A simple proof of Zahorski's description of non-differentiability sets of Lipschitz functions. *Real Anal. Exchange*, 34(1):127–138, 2009. [39]

[Gác05] P. Gács. Uniform test of algorithmic randomness over a general space. *Theoret. Comput. Sci.*, 341(1-3):91–137, 2005. [14]

[Gut01] C. E. Gutiérrez. *The Monge-Ampère Equation*, volume 44 of *Progress in Nonlinear Differential Equations and Their Applications*. Springer-Verlag, 2001. [33 and 95]

[Hei05] J. Heinonen. Lectures on Lipschitz analysis. *Rep. Univ. Jyväskylä Dept. Math. Stat.*, 100:1–77, 2005. [25]

[HR09] M. Hoyrup and C. Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Inform. and Comput.*, 207(7):830–847, 2009. [14, 15, and 87]

[Ko91] Ker-I Ko. *Complexity theory of real functions*. Birkhauser Boston Inc., 1991. [2, 11, and 12]

[LPT12] J. Lindenstrauss, D. Preiss, and J. Tišer. *Fréchet Differentiability of Lipschitz Functions and Porous Sets in Banach Spaces*. Annals of Mathematics Studies. Princeton University Press, 2012. [25]

[McC97] R. J. McCann. A convexity principle for interacting gases. *Advances in Mathematics*, 128(1):153–179, 1997. [104 and 105]

[Mig76] F. Mignot. Contrôle optimal dans les inéquations variationelles elliptiques. *J. Funct. Anal.*, (22):130–185, 1976. [5, 8, 54, 66, and 68]

[Mil02] J. S. Miller. Effectiveness for embedded spheres and balls. *Electronic Notes in Theoretical Computer Science*, 66(1):127 – 138, 2002. CCA 2002, Computability and Complexity in Analysis (ICALP 2002 Satellite Workshop). [83]

[Min60] G. Minty. Monotone networks. *Proceedings of the Royal Society of London (Series A)*, 257:194–212, 1960. [27]

[Min62] G. Minty. Monotone nonlinear operators on a Hilbert space. *Duke Math J.*, (29):341–346, 1962. [31 and 55]

[Miy13] K. Miyabe. Characterization of Kurtz randomness by a differentiation theorem. *Theor. Comp. Sys.*, 52(1):113–132, January 2013. [6]

[MMS06] W. Merkle, N. Mihailovic, and T. A. Slaman. Some results on effective randomness. *Theory of Computing Systems*, 39(5):707–721, 2006. [9 and 21]

[MR13] K. Miyabe and J. Rute. Van Lambalgen's Theorem for uniformly relative Schnorr and computable randomness. In *Proceedings of the 12th Asian Logic Conference*, pages 251–270, 2013. [17 and 18]

[Nie09] A. Nies. *Computability and randomness*, volume 51 of *Oxford Logic Guides*. Oxford University Press, Oxford, 2009. [1, 16, 17, and 20]

[Nie14] A. Nies. Differentiability of polynomial time computable functions. In Ernst W. Mayr and Natacha Portier, editors, *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014)*, volume 25 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 602–613, Dagstuhl, Germany, 2014. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. [6, 9, 35, and 38]

[PER89] M. B. Pour-El and J. I. Richards. *Computability in analysis and physics*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1989. [2, 11, and 12]

[Phe93] R. R. Phelps. *Convex Functions, Monotone Operators and Differentiability*. Lecture Notes in Mathematics. Springer-Verlag, Berlin Heidelberg, 1993. [25]

[PRS14] N. Pathak, C. Rojas, and S. G. Simpson. Schnorr randomness and the Lebesgue differentiation theorem. *Proc. Amer. Math. Soc.*, 142(1):335–349, 2014. [4, 5, 6, and 105]

[PS15] D. Preiss and G. Speight. Differentiability of Lipschitz functions in Lebesgue null sets. *Inventiones mathematicae*, 199(2):517–559, 2015. [7]

[Qi83] Liqun Qi. Uniqueness of the maximal extension of a monotone operator. *Nonlinear Analysis: Theory, Methods & Applications*, 7(4):325–332, 1983. [70]

[Rad19] H. Rademacher. Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale. *Math. Ann.*, 79(1):340–359, 1919. [54]

[Roc66] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17(3):497–510, 1966. [30 and 55]

[Roc70] R. T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific Journal of Mathematics*, 33(1):209–216, 1970. [30 and 55]

[Rog04] L. G. Rogers. *A Degree-independent Sobolev Extension Operator. PhD thesis*. Yale University, 2004. [16]

[Rud87] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987. [40, 92, and 97]

[Rut13] J. Rute. Topics in algorithmic randomness and computable analysis. *Dissertations*, (260), 2013. [6 and 105]

[Rut16] J. Rute. Computable randomness and betting for computable probability spaces. *Mathematical Logic Quarterly*, 62(4-5):335–366, 2016. [18, 20, 21, 46, 47, 58, and 59]

[RW97] R. T. Rockafellar and R. J-B Wets. *Variational Analysis*. Grundlehren der Mathematischen Wissenschaften 317. Springer-Verlag, 1997. [25, 29, 30, 32, 55, 66, and 70]

[San15] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser Basel, 2015. [73 and 94]

[Tap12] O. Tapiola. Random and non-random dyadic systems in doubling metric spaces. *MSc thesis*, 2012. [16]

[Tho85] B. S. Thomson. Real functions. Lecture Notes in Mathematics. 1170. Berlin etc.: Springer-Verlag. VII, 229 p. DM 31.50 (1985)., 1985. [35]

[Urb97] J. Urbas. Mass transfer problems. Lecture notes from a course given in Univ. Bonn. 1997. [95]

[Ves92] L. Veselý. A connectedness property of maximal monotone operators and its application to approximation theory. *Proceedings of the American Mathematical Society*, 115(3):663–667, 1992. [63]

[Vil03] C. Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, 2003. [73, 75, 77, 88, 91, and 95]

[Vil09] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 2009. [73, 75, and 95]

[Vuo82] M. Vuorinen. On the Harnack constant and the boundary behavior of Harnack funcions. *Annales AcademiæScientiarum Fennicæ*, 7:259–277, 1982. [25]

[Wei93] K. Weihrauch. Computability on computable metric spaces. *Theoretical Computer Science*, 113(2):191 – 210, 1993. [13]

[Wei00] Klaus Weihrauch. *Computable analysis*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2000. An introduction. [2, 11, and 12]

[Zah46] Z. Zahorski. Sur l'ensemble des points de non-dérivabilité d'une fonction continue. *Bull. Soc. Math. France*, 74:147–178, 1946. [7 and 39]

[Zaj79] L. Zajíček. On the differentiation of convex functions in finite and infinite dimensional spaces. *Czechoslovak Mathematical Journal*, 29(3):340–348, 1979. [7]

[Zar60] E. H. Zarantonello. Solving functional equations by contractive averaging. *U.S. Army Mathematics Research Center*, Technical Report No. 160, 1960. [27]

[Zho98] N. Zhong. Derivatives of computable functions. *Mathematival Logic Quaterly*, 44:304–316, 1998. [41]