

**Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig**

**A study in quantitative equidistribution
on the unit square**

(revised version: May 2022)

by

Max Goering and Christian Weiss

Preprint no.: 2

2022



A STUDY IN QUANTITATIVE EQUIDISTRIBUTION ON THE UNIT SQUARE

MAX GOERING AND CHRISTIAN WEISS

ABSTRACT. The distributional properties of the translation flow on the unit square have been considered in different fields of mathematics, including algebraic geometry and discrepancy theory. One method to quantify equidistribution is to compare the error between the actual time the translation flow spent in specific sets $E \subset [0, 1]^2$ to the expected time. In this article, we prove that when E is in the algebra generated by convex sets the error is of order at most $\log(T)^{1+\varepsilon}$ for almost every direction. For all but countably many badly approximable directions, the bound can be sharpened to $\log(T)^{1/2+\varepsilon}$. The error estimates we produce are smaller than for general measurable sets as proved by Beck, while our class of examples is larger than in the work of Grepstad-Larcher who obtained the bounded remainder property for their sets. Our proof relies on the duality between local convexity of the boundary and regularity of sections of the flow.

1. INTRODUCTION

In this paper, we are interested in the two-dimensional translation flow, Y_α^x , in direction α on the two-dimensional torus \mathbb{T}^2 , i.e. the unit square $[0, 1]^2$ with opposite sides glued. Given a starting point $x = (x_1, x_2) \in [0, 1]^2$, $Y_\alpha^x : [0, \infty) \rightarrow [0, 1]^2$ is defined by

$$Y_\alpha^x(t) = (\{x_1 + t\}, \{x_2 + \alpha t\}),$$

where $\{\cdot\}$ denotes the fractional part of a real number. Our focus here is on distributional properties of the flow. The two-dimensional Kronecker-Weyl equidistribution theorem states that the orbit of a translation flow on the unit square is uniformly distributed if α is irrational and otherwise it is periodic. The translation flow on the square may also be regarded as a flow on the corresponding Riemann surface which is obtained by gluing opposite edges of the square.

In order to describe the distributional properties of the translation flow quantitatively, it is a natural approach to compare the actual time spent in specific sets $E \subset [0, 1]^2$ to the expected time. According to Beck, [Bec15, Theorem 1], the following superuniformity result holds for arbitrary measurable sets.

Theorem 1.1 (Beck, 2015). *Let $E \subset [0, 1]^2$ be an arbitrary Lebesgue measurable set with $\lambda_2(E) > 0$. Then for every $\varepsilon > 0$, almost all $\alpha > 0$ and every starting point $(x_1, x_2) \in [0, 1]^2$, we have*

$$\int_0^T \chi_E(Y_\alpha^x(t)) dt - T\lambda_2(E) = o(\log(T)^{3+\varepsilon}).$$

Note that Beck's result does not impose any further geometric restrictions on E besides being measurable. In face of the generality, the error term is surprisingly, or in Beck's words *shockingly* small. In fact, Beck even proved a more general version of the theorem, extending it to the case of arbitrary real-valued square-integrable functions. The proof covers more than 30 pages in [Bec15] and uses techniques from Fourier analysis, the proof of Weyl's criterion, lattice-point counting, and averaging arguments.

If geometric restrictions are enforced, stronger results with even smaller error terms can be derived. To describe them we introduce the following notion, compare [Fer92, GL16].

Definition 1.2. *Let $E \subset [0, 1]^2$ be an arbitrary measurable subset of the unit square. We say that E is a **bounded remainder set** for the two-dimensional translation flow with starting slope α and starting point $x \in [0, 1]^2$ if the error satisfies*

$$\int_0^T \chi_E(Y_\alpha^x(t)) dt - T\lambda_2(E) = O(1),$$

i.e. it is uniformly bounded for all $T > 0$.

It is worth noting that the constant in Definition 1.2 may depend on E as well as on α and x . The following bounded remainder sets were discovered in [GL16].

Theorem 1.3 ([GL16], Theorem 1.3 and Theorem 1.4).

- *For almost all $\alpha > 0$ and every $x \in [0, 1]^2$, every polygon $E \subset [0, 1]^2$ with no edge of slope α is a bounded remainder set for the translation flow $Y_\alpha^x(t)$.*

- For almost all $\alpha > 0$ and every $x \in [0, 1]^2$, every convex set $E \subset [0, 1]^2$ whose boundary ∂E is a twice continuously differentiable curve with positive curvature at every point is a bounded remainder set for the translation flow $Y_\alpha^x(t)$.

The proof of Theorem 1.3 is based on reducing the calculation of the error term to a one-dimensional problem and a clever application of the classical Koksma-Hlawka inequality involving a rational approximation of Kronecker sequences by means of the Ostrowski expansion. In this paper, we propose another approach to bound the error term of the time discrepancy. Unlike the methods in [GL16], our technique depends upon the local geometry of the set. This gives more flexibility and allows us to consider any set $E \subset [0, 1]^2$ such that E can be written as the finite union of convex sets and their complements, that is, whenever E is in the **algebra of convex sets**, denoted $\mathcal{E}([0, 1]^2)$. Similarly we let $\mathcal{E}_\sigma([0, 1]^2)$ denote the algebra of sets generated by σ -convex sets, see Definition 3.6. Under this assumption, we discover new results analogous to [Bec15] and [GL16] with an error term of intermediate order.

Theorem 1.4. *If $E \in \mathcal{E}([0, 1]^2)$ then for almost every slope $\alpha \in \mathbb{R}$:*

(A) *If $\varepsilon > 0$ then,*

$$(1.1) \quad \int_0^T \chi_E(Y_\alpha^x(t)) dt - T\lambda_2(E) = o(\log(T)^{1+\varepsilon}).$$

(B) *If $\alpha \notin \mathcal{D}_2(E)$ ¹ is badly-approximable then (1.1) can be sharpened to*

$$(1.2) \quad \int_0^T \chi_E(Y_\alpha^x(t)) dt - T\lambda_2(E) = o(\log(T)^{\frac{1}{2}+\varepsilon}) \quad \forall \varepsilon > 0.$$

If $E \in \mathcal{E}_\sigma([0, 1]^2)$ for some $\sigma \in [2, \infty)$,

(C) *For almost every $\alpha \in \mathbb{R}$ and all $\varepsilon > 0$, (1.1) holds.*

(D) *If α is badly-approximable, then for all $\varepsilon > 0$*

$$\int_0^T \chi_E(Y_\alpha^x(t)) dt - T\lambda_2(E) = \begin{cases} o(\log(T)^{\frac{1}{2}+\varepsilon}) & \alpha \notin \mathcal{D}_2(E) \\ o(\log(T)^{\frac{1}{\sigma'}+\varepsilon}) & \alpha \in \mathcal{D}_2(E), \end{cases}$$

where σ' denotes the Hölder conjugate of σ .

¹See definition 3.10

We note that all of the errors in Theorem 1.4 are smaller than in Theorem 1.1 (from [Bec15]) but larger than in Theorem 1.3 (from [GL16]). This is reasonable because the classes $\mathcal{E}([0, 1]^2)$ and $\mathcal{E}_\sigma([0, 1]^2)$ are more general than those considered in [GL16], but less general than those in [Bec15]. In particular, Theorem 1.1 holds for Lebesgue measurable sets, which can be recognized as the completion (with respect to the Lebesgue measure) of the σ -algebra generated by convex sets, instead of the *finite* algebra generated by convex sets, $\mathcal{E}([0, 1]^2)$. In early preparation of these works, we expected to strengthen Theorem 1.4 to hold for all irrational directions for which the convexity of ∂E was not too degenerate. Hence, we expected all but countably many directions to be covered. However, this is not possible to achieve, the reasons for which are outlined in Remark 2.8.

On the other hand, in [BDY20a, Theorem 1.4.1] it is shown that for convex sets (1.1) holds for some constant *independent of the convex set*. Compared to the techniques herein, bounds independent of the convex set are striking. But, one can easily show that the independence on the set in [BDY20a, Theorem 1.4.1] cannot be extended to the class of sets $\mathcal{E}([0, 1]^2)^2$. Hence we ask,

Open Question 1: If $E \in \mathcal{E}([0, 1]^2)$ does (1.1) hold with a constant depending on $\#\mathcal{C}$ ³ and α , but not on E itself?

Finally, we note that in [BDY20a, Corollary 2 of Theorem 1.4.1] it is shown that when α is a quadratic irrational slope the error term is like $O(\log(T))$, again, independent of the convex set. Since the class of badly-approximable α contains the set of quadratic irrationals (see [EW13, Chapter 3.3] and references therein) it follows that, after throwing out all degenerate slopes $\alpha \in \mathcal{D}_2(E)$ —merely a countable set of directions— Theorem 1.4(B) is a strict improvement on both the asymptotics and the class of sets. At first glance, parts (B) and (D) of Theorem 1.4 might seem to contradict the lower bound given in [BDY20a, Theorem 1.4.1], where it is claimed that for any direction α there exist a convex set C_0 such that the error term is of order at least $\log(T)$. However, the choice of C_0 (a parallelogram for which two sides are parallel to α) depends on α and in particular has $\alpha \in \mathcal{D}_2(C_0)$.

²For instance, by taking E to be the finite union of line segments with slope α .

³See Definition 3.8 and Lemma 3.9

The strict asymptotic improvement in Theorem 1.4(B) comes from applying an L^p -Koksma-Hlawka inequality instead of the classical Koksma-Hlawka inequality. For this reason, we need to be able to verify the technical assumption that $\tau_{E,\alpha}$ is in the Sobolev space $W^{1,s}(\mathbb{T})$, see Section 2 for details. This result is stated formally in Theorem 1.5 due to its independent interest.

Theorem 1.5. *Fix $E \in \mathcal{E}([0, 1]^2)$. For any $\sigma \in [2, \infty)$ and any $\alpha \notin \mathcal{D}_\sigma(E)$ ⁴ it holds $\tau_{E,\alpha} \in W^{1,s}(\mathbb{T})$ for all $s < \frac{1}{\sigma'}$. In particular, if there exists $\sigma \in [2, \infty)$ so that $E \in \mathcal{E}_\sigma([0, 1]^2)$ then for all $\alpha \in \mathbb{R}$ and all $s < \frac{1}{\sigma'}$ it holds $\tau_{E,\alpha} \in W^{1,s}(\mathbb{T})$.*

If $\sigma < \tilde{\sigma}$ then $\mathcal{D}_{\tilde{\sigma}} \subset \mathcal{D}_\sigma$ and Corollary 3.12 ensures that for all $\sigma \geq 2$, the set $\mathcal{D}_\sigma(E)$ is a countable set for all $E \in \mathcal{E}([0, 1]^2)$.

In light of Lemma 3.9, one can read Theorem 1.5 as a statement about quantitative duality between local convexity (or concavity) of ∂E and regularity on $\tau_{E,\alpha}$. This sort of duality is typical in convex analysis, for instance: there is qualitative equivalence between strict convexity of f and the smoothness of the sub-differential map associated to f , and vice-versa (see [Roc15, Theorems 26.1, 26.3]), and there is quantitative equivalence of convexity of a Banach space and smoothness of its dual space, and vice-versa (see [ZR91, Page 193] and the references therein). Nonetheless, to the best of the authors' knowledge, this is the first time this type of duality has been used to address the quantitative equidistribution properties of flows.

Finally, the method used here is flexible enough to easily extend Theorem 1.4 to the class of measures with density $f \in W^{1,p}(\mathbb{T}^2)$ for any $p \geq 1$. This result is of the same style as the extension in [Bec15] to measures with square-integrable density, except, we attain a better bound by looking at a smaller class of measures.

Theorem 1.6. *Fix a measure $\mu \ll \lambda_2$, such that $d\mu = fd\lambda_2$. For each $\alpha \in \mathbb{R}$ and $x = (1, h) \in [0, 1]^2$, consider*

$$(1.3) \quad \tau_{\mu,\alpha}(x) := \int_{-1}^0 f(Y_\alpha^x(t))dt.$$

If there exists $p \in [2, \infty)$ so that $f \in W^{1,p}(\mathbb{T}^2)$ then for almost every $\alpha \in \mathbb{R}$

$$(1.4) \quad \int_0^T f(Y_\alpha^x(t))dt - T\mu(E) \leq o(\log(T)^{1+\epsilon}),$$

⁴See Definition 3.10.

and for all badly-approximable α

$$\int_0^T f(Y_\alpha^x(t))dt - T\mu(E) \leq o\left(\log(T)^{\frac{1}{p'}+\varepsilon}\right).$$

Note that measures of the form $d\mu = \chi_E d\lambda_2$ do not satisfy $\chi_E \in W^{1,p}(\mathbb{T}^2)$, so Theorems 1.4 and 1.6 are entirely distinct. However, the simpler proof of Theorem 1.6 illuminates why the bound on the error term in Theorems 1.4 and 1.6 depend upon the one-dimensional projections of the measure behaving nicely from an analytic point of view.

In a series of papers, [BDY20a, BDY20b, BDY21a, BDY21b], Beck et al. provide an in-depth discussion of the quantitative behavior of the translation flow on higher genus surfaces, and prominently on square-tiled surfaces of genus 2. A full summary of these extensive results (almost 500 pages) is beyond the scope of the present article. Nevertheless,

Open Question 2: Does an error bound like (1.1) hold for a non-trivial extension of the class of convex sets on higher-genus surfaces?

ACKNOWLEDGMENT

The authors would like to thank Christoph Aistleitner, József Beck, Dmitriy Bilyk, Samantha Fairchild, and Stefan Steinerberger for useful discussions on important aspects of this article.

2. PRELIMINARIES AND PROOFS OF THEOREMS 1.4 AND 1.6

In this section we provide background information and then deduce Theorem 1.4 from Theorem 1.5. All our techniques prominently use the function on the right edge of the square obtained from measuring the time which the flow in direction α ending at x spends in E within one time unit: More precisely, if $E \subset [0, 1]^2$ is a measurable set, $\alpha \in \mathbb{R}$, and $x := (1, h)$, we define the function

$$(2.1) \quad \tau_{E,\alpha}(h) := \int_{-1}^0 \chi_E(Y_\alpha^x(t))dt.$$

It follows from Fubini's Theorem that

$$\lambda_2(E) = \int_0^1 \tau_{E,\alpha}(h) dh.$$

2.1. Discretization. The core idea of the proof of Theorem 1.4 is the following: First we reduce from a continuous problem in $[0, 1]^2$ to a discrete problem in $[0, 1]$ by considering the specific points in time $N \in \mathbb{N}$. If $x = (0, x_0)$, then at these times the amount of time spent by the flow Y_α^x in E up to time N is exactly $\sum_{k=1}^N \tau_{E,\alpha}(\{k\alpha + x_0\})$. So, instead of considering the continuous-time translation flow we consider the discrete-time setting. Indeed, when $N := \lfloor T \rfloor$, the discretization error satisfies

$$0 \leq \int_0^T \chi_E(Y(t)) dt - \int_0^N \chi_E(Y(t)) dt \leq (T - N) \sqrt{1 + \alpha^2} \leq \sqrt{1 + \alpha^2},$$

because the error is maximized if $\{Y_\alpha^{(0,x_0)}(s) : s \in (N, T)\} \subset E$. Therefore the discretization error is $O(1)$ justifying our ability to neglect it and only discuss the discrete setting.

The gain from this simplification is that we only need to evaluate $\tau_{E,\alpha}$ for the **shifted finite Kronecker sequence** $\{k\alpha + x_0\}_{k=1}^N$ to understand $\int_0^N \chi_E(Y_\alpha^x(t)) dt$. Moreover, the technical assumption $\tau_{E,\alpha} \in W^{1,p}(\mathbb{T})$ allows us to make use of the L^p -Koksma Hlawka inequality (Theorem 2.3) in its particularly simple 1-dimensional setting (Corollary 2.5). This makes our proof both shorter and simpler than the ones in [Bec15] and [GL16] while allowing for a better error term than in [Bec15] and a richer class of sets than in [GL16].

2.2. Background on L^p -discrepancy. Recall that the **Sobolev space** $W^{1,p}(\mathbb{T})$ is defined as the subspace of $\tau \in L^p(\mathbb{T})$, with the additional property that there exists a (suggestively denoted) $\tau' \in L^p(\mathbb{T})$, so that

$$\int_{[0,1]} \tau f' = - \int_0^1 f \tau' \quad \forall f \in C_c^1(\mathbb{T}).$$

The function τ' is called the **Sobolev** or **weak derivative** of τ . On a first read, one could now skip directly to Remark 2.4 and understand the 1-dimensional version of the L^p -Koksma-Hlawka inequality, Corollary 2.5. But, for completeness we introduce the necessary terminology to state the L^p -Koksma-Hlawka inequality in \mathbb{R}^s , c.f.,

[Hic98, Pil20]. For all $\emptyset \neq u \subset S := \{1, \dots, s\}$ and $x \in \mathbb{R}^s$ we write x_u for the point $(x_{u_1}, \dots, x_{u_{|u|}})$ and embed $C^u = [0, 1]^u$ naturally in \mathbb{R}^s . Given any finite collection of points $P \subset [0, 1]^s$ define P_u as the orthogonal projection of P onto C^u . Moreover, let $[0, x_u)$ denote the rectangle inside C^u whose diagonal corners are the origin and x_u . Finally, if $P \subset [0, 1]^s$ is so that $\#|P| < \infty$, we write $\mu_P = \frac{1}{|P|} \sum_{x \in P} \delta_x$ where δ_x is the Dirac mass centered at x . Using this notation we define the L^p -discrepancy and the L^p -variation for $1 \leq p \leq \infty$ by:

Definition 2.1 (L^p -Discrepancy). *For a point set $P \subset I^s$ with $|P| = N$ the L^p -discrepancy of P is defined as*

$$D_p^*(P) = \max_{u \neq \emptyset} \left\| \left(\frac{\#|P_u \cap [0, x_u)|}{N} - \lambda_{|u|}([0, x_u)) \right) \right\|_{L^p(C^u)}.$$

For $p = \infty$ the L^∞ -discrepancy is just the usual star-discrepancy, compare e.g. [KN74].

Definition 2.2 ($L^{p'}$ -variation). *We define the $L^{p'}$ -variation of a sufficiently smooth⁵ function $f : [0, 1]^s \rightarrow \mathbb{R}$ as*

$$V_{p'}(f) = \max_{u \neq \emptyset} \left\| \left(\frac{\partial^{|u|} f}{\partial x_u} \Big|_{x_{S-u}=(1, \dots, 1)} \right) \right\|_{L^{p'}(C^u)}.$$

The following result from [Hic98] or more precisely its Corollary 2.5 is essential for our proof of Theorem 1.4.

Theorem 2.3 (Hickernell). *Given a point set $P = \{x_1, \dots, x_N\}$, we have*

$$\left| \int_{[0,1]^s} f d\lambda_s - \int_{[0,1]^s} f d\mu_P \right| \leq D_p^*(P) V_{p'}(f)$$

whenever p and p' are Hölder conjugates.

Remark 2.4. *When $s = 1$, the definition of L^p discrepancy and $L^{p'}$ -variation simplify dramatically, most noticeably because one need never project onto cubes C^u , but conveniently one also never needs to consider $|u| > 1$. The latter implies that, if $f \in W^{1,p}(\mathbb{T})$ then*

$$V_{p'}(f) = \|f'\|_{L^{p'}([0,1])},$$

⁵For instance, $f \in W^{s,p}([0, 1]^s)$.

is the Sobolev seminorm. On the other hand, the point set P has L^p -discrepancy

$$D_p^*(P) = \|\mu_P([0, \cdot]) - \cdot\|_{L^p([0,1])}.$$

From Remark 2.4 we obtain the following

Corollary 2.5 (1-dimensional L^p -Koksma-Hlawka inequality). *If $P \subset [0, 1]$ is a collection of N points, and $\mu_P = N^{-1} \sum_{x \in P} \delta_x$ is the corresponding empirical measure, then*

$$\left| \int_0^1 f(x) dx - \int_0^1 f d\mu_P \right| \leq D_p^*(P) \|f'\|_{L^{p'}([0,1])} \quad \text{for all } p, p' \in \mathbb{N} \text{ with } \frac{1}{p} + \frac{1}{p'} = 1.$$

Proof of Theorem 1.4. In light of Corollary 2.5 and Corollary 3.12, Theorem 1.4(A) and (C) follows immediately from Theorem 1.5 and the fact that for almost all $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ we have

$$(2.2) \quad D_{p'}^*(P_N(\alpha)) \lesssim D^*(P_N(\alpha)) \lesssim \frac{\log(N)^{1+\varepsilon}}{N},$$

compare [KN74], p.128, while Theorem 1.4(B) and (C) follows from Theorem 1.5 and Lemma 2.6. \square

Lemma 2.6. *If α is badly approximable and $p \geq 2$, then*

$$D_p^*(\{\alpha k\}_{k=1}^N) \lesssim \frac{\log(N)^{\frac{1}{p'}}}{N}.$$

Proof. By [Gra20, Remark 1]

$$\|\mu - \lambda_1\|_{W^{-1,p}(\mathbb{T})} = D_p^*(\mu, \lambda_1)$$

and by [Gra20, Proposition 4] for all $p \in [2, \infty]$, there exists $C_p < \infty$ so that for all $n \in \mathbb{N}$

$$(2.3) \quad \|\mu - \lambda_1\|_{W^{1,-p}(\mathbb{T})} \leq \frac{C_p}{n} + \left(\sum_{k=1}^{n-1} \frac{|\widehat{\mu}(k)|^{p'}}{k^{p'}} \right)^{\frac{1}{p'}}.$$

We write $\mu_N = N^{-1} \sum_{k=1}^N \delta_{\{\alpha k\}}$. According to [Ste21, Page 16] it holds that

$$|\widehat{\mu}_N(k)| \lesssim N^{-1} \{k\alpha\}$$

and $\{k\alpha\}$, the decimal part of $k\alpha$, is approximately $2^{-\ell}$ -separated. Using these facts we mimic the computation in [Ste21] to obtain

$$\begin{aligned} \sum_{k=1}^{n-1} \frac{|\widehat{\mu}_N(k)|^{p'}}{k^{p'}} &\lesssim \sum_{\ell=1}^{\log(n-1)} \sum_{2^\ell \leq k \leq 2^{\ell+1}} \frac{|\widehat{\mu}_N(k)|^{p'}}{k^{p'}} \lesssim \sum_{\ell=1}^{\log(n-1)} \sum_{2^\ell \leq k \leq 2^{\ell+1}} \frac{|N^{-1}\{k\alpha\}|^{p'}}{k^{p'}} \\ &\lesssim \frac{1}{N^{p'}} \sum_{\ell=1}^{\log(n-1)} 2^{-\ell p'} \sum_{m=1}^{2^\ell} \frac{1}{(m/2^\ell)^{p'}} \lesssim \frac{1}{N^{p'}} \sum_{\ell=1}^{\log(n-1)} \sum_{m=1}^{\infty} \frac{1}{m^{p'}} \\ &\lesssim \frac{\log(n-1)}{N^{p'}}. \end{aligned}$$

Consequently, (2.3) says that for all $n \in \mathbb{N}$, we have $D_p^*(\mu_N) \lesssim \frac{1}{n} + \frac{\log(n-1)^{\frac{1}{p'}}}{N}$. Choosing $n = N$ completes the proof. \square

It has been informally claimed in the literature, see e.g., the discussion following [Ste21, Theorem 3] that for badly approximable α , $D_p^*(\{\alpha k\}_{k=1}^N) \lesssim \frac{\log(N)^{\frac{1}{2}}}{N}$ whenever $p \geq 2$. Yet, finding a formal statement and/or proof evaded the authors. But, if this claim were true, Lemma 6.3 is a particularly bad estimate, and Theorem 1.4(2) can be improved, because $\varepsilon = 0$ would be allowed. Similarly Theorem 1.4(4) would be improved. Therefore, we ask:

Open Question 3: For badly-approximable α is the L^p -discrepancy of $\{\alpha k\}_{k=1}^N$ like $O(\log(N)^{\frac{1}{2}}/N)$?

Remark 2.7. *The proof of Theorem 1.3 relies on the specific geometric shapes of the sets $E \subset [0, 1]^2$ and on approximating the (finite) Kronecker sequences $\{n\alpha\}_{n=1}^N$ by a discretized version stemming from the Ostrowski expansion of N with respect to α . Readers who are familiar with the proof in [GL16] might wonder whether our estimate of the error term can be significantly improved by using their approximation. In fact, it is not difficult to show that also the star-discrepancy of the discretized Kronecker sequence grows with at least $\log(N)N^{-1}$. So, while that technique might yield an improvement, we prefer to present the (in our eyes) much simpler proof here and accept the slightly larger error term.*

Remark 2.8. *One might wonder whether the almost everywhere condition in Theorem 1.4 can be sharpened to hold for all but a countable collection of directions as*

we originally expected. This is untenable as can be seen from the following argument: Let $[a_0, a_1, a_2, \dots]$ be the continued fraction expansion of α with convergents $(p_i/q_i)_{i \in \mathbb{N}}$. For $q_i < n < q_{i+1}$ we have

$$|n\alpha - (np_i/q_i)| < \frac{n}{a_{i+1}q_i}$$

by a basic result from the theory of continued fraction. If $a_{i+1} \gg q_i$, then the irrational rotation is close to the rational rotation for $\frac{p_i}{q_i}$ for a long time t_i . Therefore, there exists a rectangle in which the flow spends no time for $t \in [q_i, q_i + t_i]$. If a_{i+1}^* is chosen big enough, then the error term is of order $> T^{1/2}$ implying that a logarithmic error term cannot be obtained. However, for all $a_{i+1} > a_{i+1}^*$ the error term is thus also of order $> T^{1/2}$. Although a corresponding condition needs to be imposed for all $i \in \mathbb{N}$, it still allows for an uncountable number of choices for α with

$$\int_0^T \chi_E(Y_\alpha^x(t)) dt - T\lambda_2(E) > T_n^{1/2}$$

for a sequence $(T_n)_{n \in \mathbb{N}} \rightarrow \infty$.

Remark 2.9. The situation described in Remark 2.8 should be compared to the condition under which Theorem 1.3 was proven in [GL16]. Lemma 2.1 of [GL16] states that for fixed $m > 0$,

$$\sum_{i=0}^s \frac{a_{i+1}}{q_i^m} \sum_{k=1}^{i+1} a_k$$

is uniformly bounded in s for almost every irrational $\alpha \in (0, 1)$. This implies in particular that $\frac{a_{i+1}}{q_i}$ converges to 0 for almost every irrational $\alpha \in (0, 1)$ and prevents that an exploding error like in Remark 2.8 occurs, where $\frac{a_{i+1}}{q_i}$ grows very fast.

We now prove Theorem 1.6. Hence, we assume $f \in W^{1,p}(\mathbb{T}^2)$ and will show this implies $\tau_{\mu,\alpha} \in W^{1,p}(\mathbb{T})$ for all α .

Proof of Theorem 1.6. Let λ_2 denote the Lebesgue measure on $[0, 1]^2$ and fix some $\mu \ll \lambda_2$ an absolutely continuous Borel measure on $[0, 1]^2$ with density f , i.e., $d\mu = f d\lambda_2$. By Corollary 2.5, equation (2.2), and Lemma 2.6 it suffices to show that for all $\alpha \in \mathbb{R}$, $\tau_{\mu,\alpha} \in W^{1,p}([0, 1])$.

Since $f \in W^{1,p}(\mathbb{T}^2)$, the function f is absolutely continuous with respect to almost every line, see [EG18, Chapter 4]. We now identify \mathbb{T}^2 and \mathbb{T} with $[0, 1]^2$ and $[0, 1]$.

In particular $f|_{L^h}$, where

$$(2.4) \quad L^h := \{Y_\alpha^{(1,h)}(t) : -1 \leq t \leq 0\},$$

is continuous for almost every $h \in I$. We claim this implies that difference quotients of $\tau_{\mu,\alpha}$ converge for almost every h . Since α is fixed, we abuse notation and write $Y^h = Y_\alpha^{(1,h)}$. By dominated convergence theorem and a standard mollification argument

$$\begin{aligned} \lim_{\delta \downarrow 0} \frac{\tau_{\mu,\alpha}(h+\delta) - \tau_{\mu,\alpha}(h)}{\delta} &= \lim_{\delta \downarrow 0} \int_0^1 \frac{f(Y_\alpha^{h+\delta}(t)) - f(Y_\alpha^h(t))}{\delta} dt \\ &= \int_0^1 \lim_{\delta \downarrow 0} \frac{f(Y_\alpha^{h+\delta}(t)) - f(Y_\alpha^h(t))}{\delta} dt, \end{aligned}$$

when the final limit exists and is integrable. Since $\{Y_\alpha^{h+\delta}(t) - Y_\alpha^h(t)\} = \delta e_2$, it follows that

$$\lim_{\delta \downarrow 0} \frac{f(Y_\alpha^{h+\delta}(t)) - f(Y_\alpha^h(t))}{\delta} = \partial_2 f(Y^h(t)) \quad a.e. (h, t).$$

In particular, Fubini's Theorem guarantees that

$$\int_0^1 |\tau'_{\mu,\alpha}(h)| dh \leq \int_0^1 \int_0^1 |\partial_2 f(Y^h(t))| dt dh = \int_{[0,1]^2} |\partial_2 f| < \infty$$

verifying $\tau_{\mu,\alpha} \in W^{1,p}([0,1])$. Since f can be extended to $\bar{f} \in W_{loc}^{1,p}(\mathbb{R}^2)$ by periodicity, it follows that we can again recognize $[0,1]$ as \mathbb{T} and conclude $\tau_{\mu,\alpha} \in W^{1,p}(\mathbb{T})$. \square

3. PROOF OF THEOREM 1.5

We now return our attention to the more classic case where $\mu = \chi_E d\lambda$ to study sets. Recall that $\mathcal{E}_\sigma([0,1]^2) \subset \mathcal{E}([0,1]^2)$ are the algebras generated by σ -convex and convex sets respectively. For E in either class, in order to produce the desired bound on the remainder of E in the direction α we need to show $\tau_{E,\alpha} \in W^{1,p}(\mathbb{T})$, for some $p \in (1, \infty)$. We first look at some toy examples to build an understanding of what causes $\tau_{E,\alpha}$ to gain or lose regularity.

Example 3.1. (*Power-growth graphical case*) Fix $0 < p' < \infty$, $f(x) = cx^{p'}$, and

$$E_{p'} = \{(x, y) \in [0, 1]^2 : y > f(x)\}.$$

Then the function $\tau = \tau_{E,p,0}$ defined as in (2.1) is given by

$$\tau(t) = \left(\frac{t}{c}\right)^{\frac{1}{p'}} = f^{-1}(t).$$

Consequently, for all $t \neq 0$,

$$\tau'(t) = \frac{c^{-1/p'}}{p'} t^{\frac{1-p'}{p'}} = (f^{-1})'(t).$$

In particular,

$$\tau \in \begin{cases} C^{1, \frac{1-p'}{p'}}([0, 1]) & p' < 1 \\ Lip([0, 1]) & p' = 1 \\ C^{\frac{1}{p'}} \cap W^{1,s}([0, 1]) & \forall s < p, \quad p' > 1 \end{cases}$$

The example can be readily verified by observing that since $\alpha = 0$, the function $\tau_{E,p',0}(t)$ is the length of the level-sets of the function $f(x) = cx^{p'}$. See Remark 3.13 for a more in depth discussion of the simplifications made within this example. We also compare Example 3.1 with the “flat case” when $p' = 0$.

Example 3.2. (*Horizontal constant case*) If $E_0 = \{(x, y) \in [0, 1]^2 : y > c\}$, then $\tau_0(t) = \chi_{\{t \geq c\}}(t)$. In particular, the distributional derivative $\tau'_0 = \delta_c \in \mathcal{M}([0, 1])$ is a measure which cannot be recognized as any function whenever $0 < c < 1$.

Examples 3.1 and 3.2 provide substantial theoretical insight into regularity of $\tau_{E,\alpha}$. Vaguely, they indicate the regularity of $\tau_{E,\alpha}$ is most threatened at any h , where L^h (see (2.4)) is tangent to ∂E . The danger being, as demonstrated in Example 3.2, if L^h is tangent to E on a piece where ∂E is flat, then the distributional derivative $\tau'_{E,\alpha}$ is a measure, and in particular $\tau_{E,\alpha} \in W^{1,p}(\mathbb{T})$ cannot be achieved. Nonetheless, Example 3.1 indicates that tangency of ∂E with L^h can be handled whenever ∂E separates from L^h fast enough. In fact, example 3.1 demonstrates that there should be a quantitative relationship between the regularity of $\tau_{E,\alpha}$ and how quickly ∂E separates from L^h at points of tangency. The desire to take advantage of this quantitative relationship between $\tau_{E,\alpha}$ and ∂E leads naturally to desiring the boundary decomposition of Lemma 3.9.

Example 3.3. (*Linear case*) For non-zero slope m , consider $E = \{(x, y) \in [0, 1]^2 : y > \alpha x\}$. Then, $\tau'_0(h) = \alpha^{-1}$ for all $h \in (0, 1)$.

Example 3.3 shows that flat pieces of ∂E only cause problems when they are tangent to L^h . We now recall some facts about convex functions.

Definition 3.4. *If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, a direction x^* is called a **sub-gradient to f at x** if,*

$$f(y) \geq f(x) + \langle x^*, y - x \rangle \quad \forall y \in \mathbb{R}^n.$$

We define the multi-valued map $\partial f(x) = \{x^ : x^* \text{ is a sub-gradient to } f \text{ at } x\}$. The map $\partial f(x)$ is called the **sub-differential of f** . When f is differentiable at x , it follows $\partial f(x) = \{\nabla f(x)\}$.*

The following proposition is well-known. For instance, (3.1) can be found in [Roc15, Chapter 24] while the moreover statement, due to Alexandrov, can be found in [EG18, 6.4 Theorem 1].

Proposition 3.5. *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is **convex** if and only if for all $x, y \in \mathbb{R}^n$, and all $x^* \in \partial f(x), y^* \in \partial f(y)$ we have*

$$(3.1) \quad \langle x^* - y^*, x - y \rangle \geq 0.$$

Moreover, if f is convex, there exists a set $U \subset \mathbb{R}^n$ so that $f \in C^2(U)$ and $\lambda^n(\mathbb{R}^n \setminus U) = 0$.

We also often abuse notation and simply write $\partial f(x)$ instead of writing both x^* and for all $x^* \in \partial f(x)$. For example, (3.1) would be written $\langle \partial f(x) - \partial f(y), x - y \rangle \geq 0$ for all $x, y \in \mathbb{R}^n$.

Definition 3.6. *If, for some $\sigma \in [2, \infty)$ and $c > 0$, the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is uniformly convex with*

$$(3.2) \quad \langle \partial f(x) - \partial f(y), x - y \rangle \geq c|x - y|^\sigma,$$

*then f is called **σ -convex with constant c** . When the precise constant c is not particularly important, we just say f is σ -convex.*

Convexity is a global property of functions. Nonetheless, convexity also implies a lot of local structure and vice-versa. A σ -convex set is a set which is locally the graph of a σ -convex function at every point in the boundary. Roughly, σ -convexity of a set means that at every point in the boundary, the boundary of the set separates from

the tangent plane at least as quickly as $c|x|^\sigma$, or in other words, that the boundary is never too flat. If $p > 2$, then the set $\{x \in \mathbb{R}^n : \|x\|_{\ell^p} \leq 1\}$ is an example of a p -convex set which is not s convex for any $2 \leq s < p$.

Regarding the regularity of $\tau_{E,\alpha}$, Examples 3.1 - 3.3 indicate restrictions on the global geometry, e.g., convexity, of E may not be strictly necessary. Instead we focus on requiring that the boundary separates from tangent planes sufficiently fast (or, when this fails, that the boundary is not in the direction of the flow $Y_\alpha^x(t)$). As preparation for the proof of Theorem 1.5, we collect some important properties of σ -convex functions in a separate Proposition.

Proposition 3.7 (Properties of σ -convex functions). *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is σ -convex with constant c , then:*

- (A) *The right derivative $f'_+(x) = \lim_{h \downarrow 0} \frac{f(x+h) - f(x)}{h}$ is an increasing right-continuous function, the left derivative $f'_-(x)$ is an increasing left-continuous function.*
- (B) *For all $x, y \in \mathbb{R}$, the inequality $|\partial f(x) - \partial f(y)| \geq c|x - y|^{\sigma-1}$ holds.*
- (C) *If $0 \in \partial f(x_0)$ then $g = (f|_{[x_0, \infty)})^{-1}$ is well-defined. Moreover, for all $h > x_0$,*

$$(3.3) \quad \limsup_{\delta \rightarrow 0} \left| \frac{g(h + \delta) - g(h)}{\delta} \right| \leq c^{1-\frac{1}{\sigma}} \sigma |h - x_0|^{-(1-\frac{1}{\sigma})}.$$

Proof. (1) is known to hold for any convex function, see for instance [Roc15, Chapter 24]. (2) follows by applying Cauchy-Schwarz on the left side of (3.2) then dividing by $|x - y|$. For (3), since f is σ -convex, $x > y$ implies $\partial f(x) > \partial f(y)$. Since $0 \in \partial f(x_0)$ it follows $f|_{[x_0, \infty)}$ is strictly increasing and hence has an inverse. So, it only remains to show (3.3).

Without loss of generality, $x_0 = 0$. The general result can be recovered by translation. Define $\psi(x) = c \frac{|x|^\sigma}{\sigma}$. Since f is σ -convex and $0 \in \partial f(0)$, it follows $f - \psi$ is a convex function and $0 \in \partial(f - \psi)$. In particular, $f(x) \geq \psi(x)$ and $f'(x) \geq \psi'(x)$. Consequently, if $\phi = (\psi|_{\mathbb{R}_+})^{-1}$ then for all $h > 0$, $g(h) \leq \phi(h)$, $g'(h) \leq \phi'(h)$.

Evidently $\phi(x) = (\sigma c^{-1})^{\frac{1}{\sigma}} x^{\frac{1}{\sigma}}$ and $\phi'(x) = c^{-\frac{1}{\sigma}} (\sigma x)^{\frac{-1}{\sigma}}$. Hence, if $g'(h)$ is well-defined,

$$\lim_{\delta \rightarrow 0} \frac{g(h + \delta) - g(h)}{\delta} = g'(h) \leq \phi'(h) = c^{-\frac{1}{\sigma}} (\sigma h)^{-\frac{1}{\sigma}}.$$

Since $-g$ is a convex function, it follows $g'_-(x) \geq \partial g(x)$ and g'_- is left-continuous and increasing, by part (1). Since g is differentiable almost everywhere, choose $h_k \uparrow h$

so that $g'(h_k)$ exists for each k . Then, by applying the differentiable case to this sequence, we achieve (3.3) by observing

$$g'(h_k) = g'_-(h_k) \rightarrow g'_-(h) \geq \partial g(h).$$

□

We denote the algebra generated by convex (σ -convex resp.) sets in $[0, 1]^2$ by $\mathcal{E}([0, 1]^2)$ ($\mathcal{E}_\sigma([0, 1]^2)$ resp.), and as usual we call it the **algebra of convex (σ -convex) sets** respectively.

Since a convex set E_0 is locally the epigraph of a convex function, it natural to associate to $x \in \partial E_0$, some special line L_x and a function ψ_x , not necessarily unique, for which near x , the boundary ∂E_0 is the graph of ψ_x over the line L_x with some appropriate choice of orientation. Morally, the same should be true for any $E \in \mathcal{E}([0, 1]^2)$, but one has to be careful about places where the boundaries of the various sets from which E was constructed intersect. We handle these complications by introducing convexity and concavity of boundary pieces for sets in $\mathcal{E}([0, 1]^2)$.

Definition 3.8 (Convexity and concavity of boundary pieces). *Fix a measurable $E \subset \mathbb{R}^2$, and a closed connected $C \subset \partial E$. Then C is called a **convex piece of ∂E** if:*

- (1) *For all $x \in C$, there exists a line L_x containing x and some $\varepsilon > 0$ so that $C \cap B(x, \varepsilon)$ is the graph of a convex function ψ defined on a subset of L_x , and*
- (2) *the convex hull of C is a subset of E .*

*We say $C \subset \partial E$ is a **concave piece of ∂E** if C is a convex piece of $\partial(E^c)$, the complement of E .*

*Given $\sigma \in [2, \infty)$ we say $C \subset \partial E$ is a **σ -convex (with constant c) piece of ∂E** if C is a convex piece of ∂E and the convex function ψ is σ -convex. Similarly a **σ -concave piece of ∂E** is a σ -convex piece of $\partial(E^c)$ for E^c .*

The following lemma is now evident

Lemma 3.9. *If $E \in \mathcal{E}([0, 1]^2)$ then there exists some finite collection of boundary pieces, \mathcal{C} , so that $\partial E = \cup_{C \in \mathcal{C}} C$ and each $C \in \mathcal{C}$ is either a convex or concave boundary piece.*

For $\sigma \geq 2$, if $E \in \mathcal{E}_\sigma([0, 1]^2)$ then there exists some finite collection of boundary pieces so that $\partial E = \cup_{C \in \mathcal{C}} C$ and each $C \in \mathcal{C}$ is either a σ -convex or σ -concave boundary piece.

Definition 3.10. Fix $\sigma \in [2, \infty)$, and any $E \in \mathcal{E}([0, 1]^2)$ with \mathcal{C} as in Lemma 3.9.

For each $C \in \mathcal{C}$, there exists directions θ_C, θ_C^\perp , a line L_C an interval $D \subset L_C$ and a convex function $\psi_C : D \rightarrow \mathbb{R}$ so that $C = \{x + \theta_C^\perp \psi_C(x) : x \in D\}$. We define the set of σ -degenerate slopes of ψ from the left or right by

$$\mathcal{D}_\sigma^\pm(\psi) = \left\{ m = \psi'_\pm(x) \mid \lim_{\pm(y-x) \downarrow 0} \langle m - \partial f(y), x - y \rangle = o(|x - y|^\sigma) \right\},$$

and the set of σ -degenerate slopes of ψ by

$$\mathcal{D}_\sigma(\psi) = \mathcal{D}_\sigma^+(\psi) \cup \mathcal{D}_\sigma^-(\psi).$$

For each boundary piece $C \in \mathcal{C}$ define the set of σ -degenerate directions of C as

$$\mathcal{A}_\sigma(C) = \left\{ \theta = \frac{\theta_C + m\theta_C^\perp}{|\theta_C + m\theta_C^\perp|} \mid m \in \mathcal{D}_\sigma(\psi) \right\}.$$

This gives rise to the set of σ -degenerate directions of E

$$\mathcal{A}_\sigma(E) = \bigcup_{C \in \mathcal{C}} \mathcal{A}_\sigma(C)$$

and the corresponding set of σ -degenerate slopes of E

$$\mathcal{D}_\sigma(E) = \bigcup_{C \in \mathcal{C}} \mathcal{D}_\sigma(\psi_C).$$

We now prove Proposition 3.11 and Corollary 3.12 which will guarantee that for any $\sigma \in [2, \infty)$ and any $E \in \mathcal{E}([0, 1]^2)$ the set of σ -degenerate directions of E is countable.

Proposition 3.11. For any convex $f : \mathbb{R} \rightarrow \mathbb{R}$, the set of 2-degenerate slopes $\mathcal{D}_2(f)$ is countable.

This proposition is analogous to Sard's theorem applied to the one-sided derivative of a convex function. Indeed, if f is a C^2 convex function and x is so that $\langle m - \partial f(y), x - y \rangle = o(|x - y|^2)$ for y near x , then $\partial f(x) = f'(x) = m$ and $f''(x) = 0$. So, the sets $\mathcal{D}_2^\pm(f)$ are like image of the critical set of the one sided derivatives of f .

MG: Instead of defining $\mathcal{D}_2(f)$ inside the proposition, I defined $\mathcal{D}_\sigma(\psi)$ in the definition... this then allowed some simplification of the remaining things because I didn't have to keep writing $\mathcal{D}^+ \cup \mathcal{D}^-$, etc. I also then removed the extra definition from the proposition.

Proof. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be convex. For each choice of sign \pm we observe we can write

$$\mathcal{D}_2^\pm(f) = \left\{ m = f'_\pm(x) \mid \left\langle m - \partial f(y), \frac{x-y}{|x-y|} \right\rangle \leq |x-y|\omega_x(|x-y|) \quad \forall y > x \right\},$$

for some gauge function ω_x , namely $\omega_x : [0, \infty) \rightarrow [0, \infty)$ is non-decreasing and continuous with $\omega_x(0) = 0$. We will show that $\mathcal{D}_2^\pm(f)$ are each countable. Consider $m_1, m_2 \in \mathcal{D}_2^+(f)$ so that, without loss of generality $f'_+(x_i) = m_i$ implies $x_2 > x_1$. By convexity of f and $m_1 \in \mathcal{D}_2^+(f)$, it follows

$$0 \leq \left\langle m - \partial f(y), \frac{x_1-y}{|x_1-y|} \right\rangle = |x_1-y|\omega_{x_1}(|x_1-y|) \quad \forall y > x_1.$$

Equivalently,

$$|m - \partial f(y)| = |x_1-y|\omega_{x_1}(|x_1-y|).$$

Choose δ small enough that $|x-y|\omega(|x-y|) < m_2 - m_1$ whenever $0 < y - x < \delta$. Then $0 < y - x < \delta$ implies $f'_+(y) < m_2$. In particular, if $I_i = \{x : f'_+(x) = m_i\}$ there is a positive distance between I_1 and I_2 . Thus, $\mathcal{D}_2^+(f)$ (and similarly $\mathcal{D}_2^-(f)$) are countable. \square

Corollary 3.12. *If $\sigma \in [2, \infty)$, $\Omega \subset \mathbb{R}^2$, and $E \in \mathcal{E}([0, 1]^2)$ then $\mathcal{A}_\sigma(E)$, the set of σ -degenerate directions of E , is countable.*

The corollary follows immediately from the finiteness of \mathcal{C} in Lemma 3.9, the definition of degenerate direction (Definition 3.10), Proposition 3.11 and the observation that if $\sigma < \sigma'$ then $D_{\sigma'}(f) \subset D_\sigma(f)$.

To prepare for the proof of Theorem 1.5 we provide a more in-depth discussion of Examples 3.1-3.2 and how they can be seen as models for the more general setting with $E \in \mathcal{E}_\sigma([0, 1]^2)$.

Remark 3.13. *We start by discussing the dramatic simplifications.*

1) *As f is convex, we know that $\{f = t\}$ has 0, 1, or 2 points. Example 3.1 conveniently has $\{f = t\}$ is always a set with 1 point. This is the first key observation for why $\tau_{E,0}$ acts identically to f^{-1} in this example.*

2) *The domain of f is tangent to the flow $Y_0^x(t)$. This is what allows $\tau_{E,0}$ to be recognized as the length of the level sets of f , and not the length of the intersection of $\text{epi}(f)$ with the graph of $y = m_\alpha x$, where m_α represents the slope of the*

line in the direction α , as a graph over the same domain as f , see (3.6). By the Pythagorean Theorem writing $F(x) = f(x) - m_\alpha x$ the length of this intersection is $\sqrt{1 + |m_\alpha|^2} \lambda_1(F^{-1}(\cdot))$.

3) The domain of f is perpendicular to the direction e_2 . The direction e_2 is special for the function $\tau_{E,\alpha}$ since e_2 is the direction the flow $Y_\alpha^{(1,h)}(t)$ is translated as h changes. Therefore, this simplification allows $\tau_{E,\alpha}(h)$ to correspond to the h -level set of F instead of $\tau_{E,\alpha}(h + \delta)$ and $\tau_{E,\alpha}(h)$ corresponding to two level sets of $\{F = y_1\}$ and $\{F = y_2\}$ for some $|y_1 - y_2| = \frac{\delta}{(e_2 \cdot \theta_L^\perp)}$. This means when θ_L is not perpendicular to e_2 , and while avoiding dividing by zero, one could write $G(e_2 \cdot \theta_L^\perp h) = F(h)$ and up to translation identify

$$\tau_{E,\alpha}(h) = \sqrt{1 + |m_\alpha|^2} \lambda_1 \left(\left\{ F = \frac{h}{e_2 \cdot \theta_L^\perp} + y_0 \right\} \right),$$

for some vertical translation y_0 of the graph of G . Therefore, if ∂E is the graph of a C^1 -bijection f over the line L , then

$$(3.4) \quad \tau'_{E,\alpha}(h) = \frac{\sqrt{1 + |m_\alpha|^2}}{e_2 \cdot \theta_L^\perp} (G^{-1})'(h + y_0).$$

In particular, if the slope of G can be bounded away from zero for some range of h , then on that interval, $\tau_{E,\alpha}$ will be Lipschitz.

If f were not injective, formally, the above analysis still indicates

$$(3.5) \quad \tau'_{E,\alpha}(h) = \sum_{x:f(x)-m_\alpha(x)=c_h} \frac{\sqrt{1 + |m_\alpha|^2}}{(\theta_L^\perp \cdot e_2) M(f'(x), \alpha)}$$

where $M(f'(x), \alpha)$ is the slope of L^h with respect to the tangent line to f at x . That is,

$$M(f'(x), \alpha) = \frac{(e_1 + \alpha e_2) \cdot (\theta_L + f'(x) \theta_L^\perp)}{(e_1 + \alpha e_2) \cdot (-f'(x) \theta_L + \theta_L^\perp)}.$$

In theory, cancellation can occur in (3.5), allowing very specific sets and directions to satisfy $\tau_{E,\alpha} \in W^{1,s}$ despite the hypotheses of Theorem 1.5 failing, see Figure 1. The loss in the proof arises from putting absolute values inside the summation of (3.5), preventing multiple singular parts of $\tau'_{E,\alpha}$ from potentially canceling out.

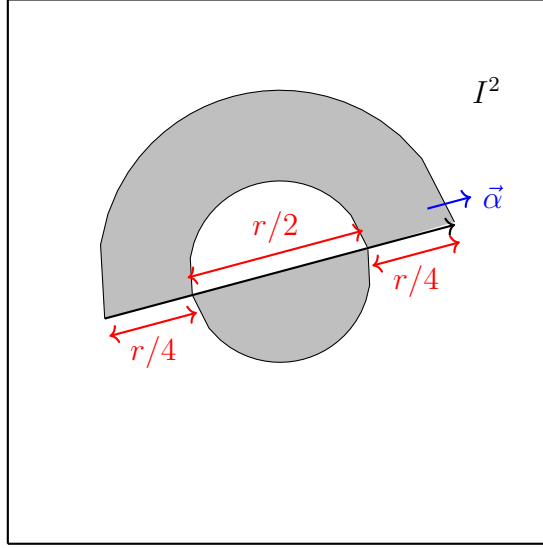


FIGURE 1. An example of a set E and direction α so that, despite the existence of flat pieces of the boundary in the direction α , the function $\tau_{E,\alpha}$ would behave nicely. The sort of cancellation occurring here is lost in our proof, because we ultimately put absolute values on each term inside the summation in (3.5).

We now set some terminology and notation that will be used to maintain succinctness of exposition without losing precision.

Throughout the proof we will be discussing $x_0 \in \partial E$. Since $E \in \mathcal{E}([0, 1]^2)$ there are $C \in \mathcal{C}$ so that $x_0 \in C$ and C is the graph of some ψ over some $L_C \ni x_0$. We will let θ_C denote a choice of the positive direction of the line L_C . We choose θ_C^\perp to denote the **positive “vertical” direction**. Since L_C is a supporting hyperplane of C , this is equivalent to choosing the direction so that $\langle y - x_0, \theta_C^\perp \rangle \geq 0$ for all $y \in C$ when C is a convex piece of ∂E and $\langle y - x_0, \theta_C^\perp \rangle \leq 0$ for all $y \in C$ when C is a concave piece of ∂E .

When we talk about ψ being a function over the line L_C , we keep L_C embedded in $[0, 1]^2$. Then the graph of ψ means the set of points $\{x + \psi(x)\theta_C^\perp : x \in L_C \cap \text{dom}(\psi)\}$ and the set of $x \in L_C$ is parametrized by $\{x_0 + t\theta_C\}$. We say $x_0 + t\theta_C$ is left or right of x_0 depending on whether $t < 0$ or $t > 0$ respectively.

Once we fix L_C, θ_C and θ_C^\perp we also define $m_\alpha = m_\alpha(C)$ by

$$(3.6) \quad m_\alpha = \frac{\theta_C^\perp \cdot (e_1 + \alpha e_2)}{\theta_C \cdot (e_1 + \alpha e_2)},$$

whenever the denominator is non-zero. In this case, if h_0 is so that $x_0 \in L^{h_0} \cap \partial E$, then the graph over L_C of the function $f(x_0 + t\theta_C) = tm_\alpha$ is contained in L^{h_0} , where L^{h_0} is as in (2.4). Moreover, $L^{h_0+\delta}$ is contained in the graph of $f_\delta(x_0 + t\theta_{x_0}) = tm_\alpha + \frac{\delta}{(e_2 \cdot \theta_C^\perp)}$. Indeed,

$$(3.7) \quad e_2 \cdot [f_\delta(x_0 + t\theta_{x_0}) - f(x_0 + t\theta_{x_0})] = e_2 \cdot \left[\frac{\delta}{(e_2 \cdot \theta_{x_0}^\perp)} (\theta_C^\perp) \right] = \delta.$$

We recall

$$(3.8) \quad \theta_\alpha = \frac{e_1 + \alpha e_2}{\sqrt{1 + |\alpha|^2}}$$

is the normalized direction of our trajectory.

Proof of Theorem 1.5. Fix $E \in \mathcal{E}([0, 1]^2)$, $h_0 \in [0, 1]$, $\sigma \in [2, \infty)$, $\alpha \notin \mathcal{D}_\sigma(E)$, and $x_0 \in L^{h_0} \cap \partial E$. Choose some $C \in \mathcal{C}$ with $x_0 \in C$. Then C is the graph over some line L_C by some convex (or concave) ψ . Let H_C denote the collection of $h \in [0, 1]$ so that $L^h \cap C \neq \emptyset$. Then H_C is an interval and $C \cap \partial E$ contains 1 or 2 points for every $h \in H_C$. We write $C \cap \partial E = \{x_\pm(h)\}$ where $x_+(h) \cdot \theta_C \geq x_-(h) \cdot \theta_C$. Since C is the graph of a convex function, it turns out that $h \mapsto x_\pm(h)$ are continuous functions on their domain of definition.

Define $f(t) = \psi(x + t\theta_{x_0})$. By assumption f is convex or concave. Without loss of generality, suppose f is convex. We split into several cases depending on C and ψ .

Case 0: $\theta_C^\perp \cdot e_2 = 0$.

In this case, the discussion within Remark 3.13 does not apply, because L^h and $L^{\tilde{h}}$ are translations of each other in the θ_C -direction. In fact, this makes our job easier. Since C is the graph of ψ , for all $h, \tilde{h} \in H_C$ the inequality

$$\begin{aligned} \left| \theta_\alpha \cdot (x_\pm(h) - x_\pm(\tilde{h})) \right|^2 &= \left| \theta_\alpha \cdot \left((h - \tilde{h}) \theta_C + (\psi(x) - \psi(\tilde{x})) \theta_C^\perp \right) \right|^2 \\ &\leq |h - \tilde{h}|^2 + |\psi(x) - \psi(\tilde{x})|^2 \end{aligned}$$

holds, where x, \tilde{x} are the points in L_C with heights h, \tilde{h} respectively. In particular $|x - \tilde{x}| = |h - \tilde{h}|$. Since ψ is convex, it is locally Lipschitz. Thus, there exists $\delta(h) > 0$

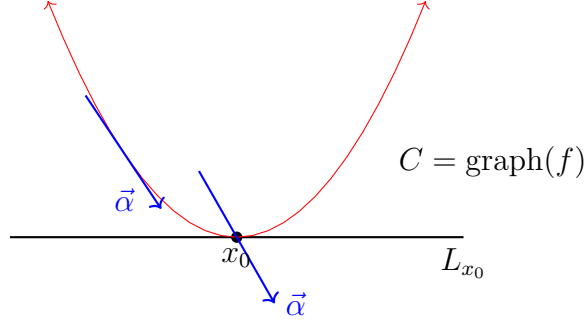


FIGURE 2. To the right of x_0 we see Case 1. As one moves right, the slope of f only gets further from the slope of α . So, the distance between m_α and $f'_+(0)$ easily bounds the distance between ∂f and m_α as one moves right.

To the left of x_0 we see Case 2. A priori, we have no idea how far we can go before the slope of the graph agrees with m_α . Nonetheless, continuity of f'_- from the left ensures there is some distance one can move left before this happens.

small enough that $h, \tilde{h} \in B_\delta(h_0)$ ensures $|\psi(x) - \psi(\tilde{x})| \leq L|h - \tilde{h}|$ for some (local) Lipschitz constant L . In particular, $|\theta_\alpha \cdot (x_\pm(h) - x_\pm(\tilde{h}))| \leq \sqrt{1 + L^2}|h - \tilde{h}|$ for all $h \in I_C(x_0) := (h_0 - \delta(h), h_0 + \delta(h))$.

Case 1: $f'_-(0) < m_\alpha$ or $m_\alpha < f'_+(0)$ and $e_2 \cdot \theta_C^\perp \neq 0$.

The main idea for this case is to use monotonicity of $\partial\psi$ and hence of ∂f to demonstrate that the accessible tangent directions for C , i.e., the directions $(1, f'_\pm(t))$ remain far from the direction θ_α . Consequently, the epigraph of ψ , denoted $\text{epi}(\psi)$, is contained in some closed cone, and the direction θ_α “lies below” this cone. A bound on the α -directional movement of the boundary points, similar to (3.4), where G' will be bounded uniformly below. See Figures 2, 3.

By convexity of f , it follows $f'_+(0) \leq \partial f(t)$ for all $t > 0$ and therefore

$$(3.9) \quad \partial f(t) - m_\alpha \geq f'_+(0) - m_\alpha > 0 \quad \forall t \geq 0.$$

For $h, \tilde{h} \in H_C$ we use the graphicality of ψ to define x, \tilde{x} by

$$(3.10) \quad x_+(h) = x\theta_C + \psi(x)\theta_C^\perp \quad \text{and} \quad x_+(\tilde{h}) = \tilde{x}\theta_C + \psi(\tilde{x})\theta_C^\perp,$$

and without loss of generality suppose $x_0 \cdot \theta_C \leq x < \tilde{x}$. As in Remark 3.13(3) or (3.7) we know $(\psi(x) - \psi(\tilde{x}) - m_\alpha(x - \tilde{x})) = \frac{h - \tilde{h}}{\theta_C^\perp \cdot e_2}$. Then, from (3.10) it follows

$$(3.11) \quad \begin{aligned} & \left| \theta_\alpha \cdot \left(\frac{x_+(h) - x_+(\tilde{h})}{h - \tilde{h}} \right) \right|^2 = \left| \theta_\alpha \cdot \left(\theta_C \frac{x - \tilde{x}}{h - \tilde{h}} + \theta_C^\perp \frac{\psi(x) - \psi(\tilde{x})}{h - \tilde{h}} \right) \right|^2 \\ & \leq \left| \frac{1}{\theta_C^\perp \cdot e_2} \right|^2 \left[\left| \frac{x - \tilde{x}}{\psi(x) - \psi(\tilde{x}) - m_\alpha(x - \tilde{x})} \right|^2 + \left| \frac{\psi(x) - \psi(\tilde{x})}{\psi(x) - \psi(\tilde{x}) - m_\alpha(x - \tilde{x})} \right|^2 \right] \\ & = \left| \frac{1}{\theta_C^\perp \cdot e_2} \right|^2 \left[\left| \frac{1}{\frac{\psi(x) - \psi(\tilde{x})}{x - \tilde{x}} - m_\alpha} \right|^2 + \left| \frac{\psi(x) - \psi(\tilde{x})}{\frac{\psi(x) - \psi(\tilde{x})}{x - \tilde{x}} - m_\alpha} \right|^2 \right] \\ & \leq \frac{c_0}{|\theta_C^\perp \cdot e_2|^2} \frac{1}{|f'_+(0) - m_\alpha|^2}, \end{aligned}$$

where the final line used the convexity and boundedness of ψ in addition to (3.9). Note $c_0 = 9 \geq 1 + 4\|\psi\|_{L^\infty}^2$ suffices. In particular, we have proven a Lipschitz bound of the form

$$(3.12) \quad \left| \theta_\alpha \cdot (x_+(h) - x_+(\tilde{h})) \right| \leq c_1 |h - \tilde{h}| \quad \forall h, \tilde{h} \in H_C.$$

Therefore, the bound (3.12) holds throughout H_C , which either contains an open interval around h_0 , or a half-open interval around h_0 . In the latter case, we can symmetrize the interval around h_0 by replacing the natural bound of 0 on the displacement when $x_+(h)$ does not exist, by the bound (3.12) so that the bound (3.12) holds on some open interval, denoted $I_C(x_0)$, containing h_0 .

When $f'_-(0) < m_\alpha$ then $\theta_\alpha \cdot (x_-(h) - x_-(\tilde{h}))$ can be bounded in the same fashion.

Case 2: $m_\alpha < f'_-(0)$ or $f'_+(0) < m_\alpha$ and $e_2 \cdot \theta_C^\perp \neq 0$.

The main idea for this case is to use the continuity of $f'_+(t)$ from the right and of $f'_-(t)$ from the left to show that near enough to $t = 0$ the absolute value $|\partial f(t)|$ remains bounded above by $|m_\alpha|$. So, $\text{epi}(\{y = m_\alpha x\})$ is contained in a cone which stays a positive distance away from $\partial f(t)$ for t small enough, producing the desired bound, analogously to the previous case, see Figure 2.

We will first consider the instance $f'_+(0) < m_\alpha$.

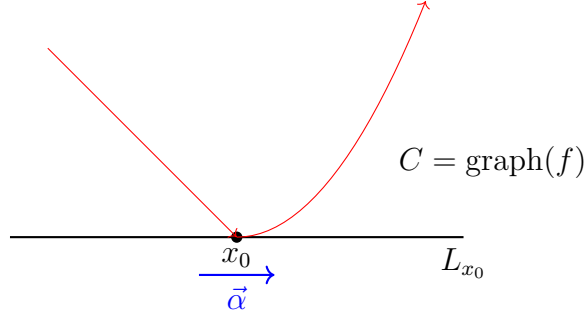


FIGURE 3. To the right of x_0 we see Case 3. At x_0 holds $f'_+(0) = m_\alpha$. If $\theta_\alpha \notin \mathcal{A}_\sigma(E)$, the quantitative convexity guarantees that the slope of f grows quickly. To the left of x_0 , we again see Case 1.

Since f'_+ is continuous from the right, it follows there exists $\delta > 0$ so that

$$(3.13) \quad 0 < \frac{m_\alpha - f'_+(0)}{2} \leq m_\alpha - \partial f(t) \leq m_\alpha f'_+(0) \quad \forall 0 \leq t \leq \delta.$$

Define $I_C(x_0) = \{h \in H_C : |(x_+(h) - x_0) \cdot \theta_C| < \delta\}$. For $h, \tilde{h} \in H_C$ write $x_+(h), x_+(\tilde{h})$ as in (3.10), where without loss of generality $0 \leq (x - x_0 \cdot \theta_C) \leq (\tilde{x} - x_0 \cdot \theta_C) < \delta$.

Then, proceeding identically to Case 1, we get inequality (3.11) from which we

$$\left| \theta_\alpha \cdot \left(\frac{x_+(h) - x_+(\tilde{h})}{h - \tilde{h}} \right) \right|^2 \leq \frac{c_0}{|\theta_C^\perp \cdot e_2|^2} \frac{2}{|f'_+(\delta) - m_\alpha|^2},$$

where the extra 2 comes from the $\frac{1}{2}$ in (3.13). If necessary, we can again symmetrize $I_C(x_0)$ to create an open interval around h_0 where (3.12) holds. The case $f'_-(0) > m_\alpha$ follows similarly, using left continuity of f'_- in place of right continuity of f'_+ .

Case 3: $m_\alpha = f'_-(0)$ or $f'_+(0) = m_\alpha$ and $e_2 \cdot \theta_C^\perp \neq 0$.

For this case, the main idea is starkly different from the others. We cannot bound $|\partial f(t) - m_\alpha|$ below as in Case 1 and 2. Instead, we use the quantitative convexity of f to show that, as t moves away from 0, the derivative $f'(t)$ moves away from m_α at least as fast as some fixed rate. The inverse of this rate then produces the desired upper-bounds on the rate of movement of $x(h) \cdot \alpha$, see Figure 3.

We start with the case $f'_+(0) = m_\alpha$. Since $\alpha \notin \mathcal{D}_\sigma(E)$ by hypothesis, it follows that ψ is σ -convex in a neighborhood of x_0 .

Define $\tilde{f}(t) = f(t) - tm_\alpha$. Then \tilde{f} is also σ -convex in a neighborhood of 0 by convexity of $t \mapsto -tm_\alpha$. Since $m_\alpha = f'_+(0)$ it follows $0 \in \partial\tilde{f}(0)$. Moreover, $x_0 \in L_{x_0} \cap C$ implies $f(0) = \tilde{f}(0) = 0$. Let $g = \tilde{f}|_{\{t \geq 0\}}^{-1}$. As f is strictly increasing and continuous, also the function g is continuous and defined on some half-open interval. Let $x_+(h)$ be as in the previous two subcases. Without loss of generality, by shrinking the domain of ψ if necessary, we can suppose ψ is σ -convex. Let $I_C(x_0) \subset H_C$ correspond to those h so that $L^h \cap \text{graph}(\psi(x_0 + t\theta_C)|_{t \geq 0})$ has one point. Then $I_C(x_0)$ is a half-open interval containing h_0 . Moreover, on $I_C(x_0)$ ⁶

$$(3.14) \quad \alpha \cdot (x_+(h) - x_+(\tilde{h})) = \sqrt{1 + |m_\alpha|^2} \left[g \left(\frac{h - h_0}{e_2 \cdot \theta_{x_0}^\perp} \right) - g \left(\frac{\tilde{h} - h_0}{e_2 \cdot \theta_{x_0}^\perp} \right) \right].$$

By (3.3) of Proposition 3.7 it follows

$$(3.15) \quad \limsup_{h \rightarrow \tilde{h}} \left| \frac{g \left(\frac{h - h_0}{e_2 \cdot \theta_{x_0}^\perp} \right) - g \left(\frac{\tilde{h} - h_0}{e_2 \cdot \theta_{x_0}^\perp} \right)}{\frac{h_1 - h_2}{e_2 \cdot \theta_{x_0}^\perp}} \right| \leq c^{-\frac{1}{\sigma}} \sigma^{-\frac{1}{\sigma'}} \left(\frac{|\tilde{h} - h_0|}{e_2 \cdot \theta_{x_0}^\perp} \right)^{\frac{-1}{\sigma'}},$$

where σ' is the Hölder conjugate of σ . Combining (3.14) and (3.15) yields

$$(3.16) \quad \limsup_{h \rightarrow \tilde{h}} \left| \alpha \cdot \left(\frac{x_+(h) - x_+(\tilde{h})}{h - \tilde{h}} \right) \right| \leq c_1 |\tilde{h} - h_0|^{\frac{-1}{\sigma'}},$$

for some constant $c_1 = c_1(C, \alpha, \sigma)$. We can suppose $I_C(x_0)$ is an open interval by symmetrization and perhaps making our estimates worse. In the case $f'_-(x) = m_\alpha$, we bound x_- similarly, and most readily by defining $\tilde{f}(-t) = f(t) + tm_\alpha$ for $t > 0$.

We now translate our bounds on the individual points in $L^{h_0} \cap \partial E$ into information about the regularity of $\tau_{E, \alpha}$. Fix $h_0 \in [0, 1]$, $C \in \mathcal{C}$ so that $C \cap L^{h_0} \neq \emptyset$. Consider the function

$$F_C(h) = \sum_{x(\delta) \in C \cap L^h} \min_{x_0 \in C \cap L^{h_0}} \theta_\alpha \cdot (x(\delta) - x_0)$$

which measures the displacement in the θ_α -direction of points in $C \cap L^h$ from the closest points in $C \cap L^{h_0}$. By the preceding casework, each $x_0 \in C \cap L^{h_0}$ has some interval $I_C(x_0)$ on which $|\theta_\alpha \cdot (x_\pm(h) - x_0)|$ is bounded by (3.12) or (3.16).

⁶For comparison with Remark 3.13(3), $g = G^{-1}$.

Consider the open interval $I_C(h_0) := \cap_{x_0 \in C \cap L^{h_0}} I_C(x_0)$. It follows from (3.12) and (3.16) that F_C is locally Lipschitz on $I_C(h_0) \setminus \{h_0\}$ and in particular F_C is differentiable almost everywhere on $I_C(h_0)$. Moreover, (3.12) and (3.16) imply

$$(3.17) \quad |F'_C(h)| \leq \sum_{x_0 \in C \cap L^h} A_{C,x_0} + B_{C,x_0} |h - h_0|^{-\frac{1}{\sigma'}} \quad a.e. \ h \in I_C(h_0).$$

where A_{C,x_0} and B_{C,x_0} depend upon which case C, x_0 fall into.

We now define the interval $I(h_0) = \cap_{C \in \mathcal{C}} I_C(h_0)$ which is open by finiteness of \mathcal{C} . Moreover, $h, \tilde{h} \in I(h_0)$ implies

$$|\tau_{E,\alpha}(h) - \tau_{E,\alpha}(\tilde{h})| \leq \sum_{\substack{C \in \mathcal{C} \\ C \cap L^{h_0} \neq \emptyset}} |F_C(h) - F_C(\tilde{h})|.$$

Consequently, $\tau_{E,\alpha}|_{I(h)}$ is differentiable almost everywhere, and by (3.17),

$$(3.18) \quad |\tau'_{E,\alpha}(h)| \leq \sum_{\substack{C \in \mathcal{C} \cap \mathcal{L} \\ C \cap L^h \neq \emptyset}} |F'_C(h)| \leq C_{h_0} (1 + |h - h_0|^{-\frac{1}{\sigma'}})$$

holds for almost every $h \in I(h_0)$. The dependence on h_0 on the right hand side of (3.18) comes from the dependence of A_{C,x_0} and B_{C,x_0} for each $x_0 \in C \cap L^{h_0}$ and for each C with $C \cap L^{h_0} \neq \emptyset$. Still, the dependence of (3.18) on h_0 does not prevent $\tau'_{E,\alpha} \in L^s(\mathbb{T})$. Indeed, since each $h_0 \in [0, 1]$ has some open interval on which (3.18) holds, and $[0, 1]$ is compact, there exist finitely many $\{h_k\}_{k=1}^M$ so that $\{I(h_k)\}_{k=1}^M$ covers $[0, 1]$. Since $\tau_{E,\alpha}$ is differentiable almost everywhere on each $I(h_k)$ it is differentiable almost everywhere on \mathbb{T} . Consequently, writing $\bar{C} := \max_k C_{h_k}$ it follows

$$\int_0^1 |\tau'_{E,\alpha}(h)|^s dh \leq \sum_{k=1}^M \int_{I(h_k)} C_{h_k}^s (1 + |h - h_0|^{-\frac{1}{\sigma'}})^s dh \leq 2^s \bar{C}^s M \int_{-1}^1 (1 + |h|^{-\frac{s}{\sigma'}}) dh,$$

which is finite if and only if $s < \sigma'$. Since $|\tau_{E,\alpha}| \leq 1$, this completes the proof that $\tau_{E,\alpha} \in W^{1,s}(\mathbb{T})$ for all $s < \sigma'$. \square

REFERENCES

- [BDY20a] J. Beck, M. Donders, and Y. Yang. Quantitative behavior of non-integrable systems. I. *Acta Math. Hungar.*, 161(1):66–184, 2020.
- [BDY20b] J. Beck, M. Donders, and Y. Yang. Quantitative behavior of non-integrable systems. II. *Acta Math. Hungar.*, 162(1):220–324, 2020.

- [BDY21a] J. Beck, M. Donders, and Y. Yang. Quantitative behavior of non-integrable systems (iii), 2021.
- [BDY21b] J. Beck, M. Donders, and Y. Yang. Quantitative behavior of non-integrable systems (iv), 2021.
- [Bec15] J. Beck. From Khinchin’s conjecture on strong uniformity to superuniform motions. *Mathematika*, 61:591–707, 2015.
- [EG18] L. Evans and R. Gariepy. *Measure theory and fine properties of functions*. Routledge, 2018.
- [EW13] M. Einsiedler and T. Ward. *Ergodic theory*. Springer, 2013.
- [Fer92] S. Ferenczi. Bounded remainder sets. *Acta Arithmetica*, 61 (4):319–326, 1992.
- [GL16] S. Grepstad and G. Larcher. Sets of bounded remainder for a continuous irrational rotation on $[0, 1]^2$. *Acta Arith.*, 176(4):365–395, 2016.
- [Gra20] C. Graham. Irregularity of distribution in wasserstein distance. *Journal of Fourier Analysis and Applications*, 26(5):1–21, 2020.
- [Hic98] F. Hickernell. A generalized discrepancy and quadrature error bound. *Math. of Comp.*, 67(221):299–322, 1998.
- [KN74] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. John Wiley & Sons, New York, 1974.
- [Pil20] F. Pillichshammer. Discrepancy of digital sequences: New results on a classical qmc topic. In B. Tuffin and P. L’Ecuyer, editors, *Monte Carlo and Quasi-Monte Carlo Methods. MCQMC 2018*, pages 81–103. Proceedings in Mathematics & Statistics, vol 324. Springer, 2020.
- [Roc15] R. Rockafellar. *Convex analysis*. Princeton university press, 2015.
- [Ste21] S. Steinerberger. Wasserstein distance, fourier series and applications. *Monatshefte für Mathematik*, 194(2):305–338, 2021.
- [ZR91] Z.-B. Zong and G. Roach. Characteristic inequalities of uniformly convex and uniformly smooth banach spaces. *Journal of Mathematical Analysis and Applications*, 157(1):189–210, 1991.

MAX-PLANCK-INSTITUT FÜR MATHEMATIK IN DEN NATURWISSENSCHAFTEN, INSELSTR. 22,
04103 LEIPZIG, GERMANY

Email address: `goering@mis.mpg.de`

RUHR WEST UNIVERSITY OF APPLIED SCIENCES, DUISBURGER STR. 100, 45479 MÜLHEIM AN
DER RUHR, GERMANY

Email address: `christian.weiss@hs-ruhrwest.de`