On the Duality Theory for the Monge–Kantorovich Transport Problem

Mathias Beiglböck, Christian Léonard, Walter Schachermayer*

1 Introduction

This article, which is an accompanying paper to [BLS09], consists of two parts: In section 2 we present a version of Fenchel's perturbation method for the duality theory of the Monge–Kantorovich problem of optimal transport. The treatment is elementary as we suppose that the spaces $(X, \mu), (Y, \nu)$, on which the optimal transport problem [Vil03, Vil09] is defined, simply equal the finite set $\{1, \ldots, N\}$ equipped with uniform measure. In this setting the optimal transport problem reduces to a finite-dimensional linear programming problem.

The purpose of this first part of the paper is rather didactic: it should stress some features of the linear programming nature of the optimal transport problem, which carry over also to the case of general polish spaces X, Y equipped with Borel probability measures μ, ν, and general Borel measurable cost functions $c : X \times Y \to [0, \infty]$. This general setting is analyzed in detail in [BLS09]; section 2 below may serve as a motivation for the arguments in the proof of Theorems 1.2 and 1.7 of [BLS09] which pertain to the general duality theory.

The second — and longer — part of the paper, consisting of sections 3 and 4 is of a quite different nature.

Section 3 is devoted to illustrate a technical feature of [BLS09, Theorem 4.2] by an explicit example. The technical feature is the appearance of the singular part $\widehat{h}^s$ of the dual optimizer $\widehat{h} \in L^1(X \times Y, \pi)^{**}$ obtained in ([BLS09, Theorem 4.2]). In Example 3.1 below we show that, in general, the dual optimizer $\widehat{h}$ does indeed contain a non-trivial singular part. In addition, this example allows to observe in a rather explicit way how this singular part "builds up", for an optimizing sequence $(\varphi_n \oplus \psi_n)_{n=1}^\infty \in L^1(X \times Y, \pi)$ which converges to $\widehat{h}$ with respect to the weak-star topology. The construction of this example, which is a variant of an example due to L. Ambrosio and A. Pratelli [AP03], is rather longish and technical. Some motivation for this construction will be given at the end of Section 2.

Section 4 pertains to a modified version of the duality relation in the Monge-Kantorovich transport problem. Trivial counterexamples such as [BLS09, Example 1.1] show that in the case of a measurable cost function $c : X \times Y \to [0, \infty]$ there may be a duality gap. The main result (Theorem 1.2) of [BLS09] asserts that one may avoid this difficulty by considering a suitable relaxed form of the primal problem; if one does so, duality holds true in complete generality. In a different vein, one may leave the primal problem unchanged, and overcome the difficulties encountered in the above mentioned simple example by considering a slightly modified dual problem (cf. [BLS09, Remark 3.4]). In the last part of the article we consider

*The first author acknowledges financial support from the Austrian Science Fund (FWF) under grant P21209. The third author acknowledges support from the Austrian Science Fund (FWF) under grant P19456, from the Vienna Science and Technology Fund (WWTF) under grant MA13 and by the Christian Doppler Research Association (CDG). All authors thank A. Pratelli for helpful discussions on the topic of this paper. We also thank L. Summerer for his advice.

a certain twist of the construction given in section 3, which allows us to prove that this dual relaxation does not lead to a general duality result.

2 The finite case

In this section we present the duality theory of optimal transport for the finite case: Let $X = Y = \{1, \dots, N\}$ and let $\mu = \nu$ assign probability N^{-1} to each of the points $1, \dots, N$. Let $c = (c(i,j))_{i,j=1}^N$ be an $\mathbb{R}_+$-valued $N \times N$ matrix.

The problem of optimal transport then becomes the subsequent linear optimization problem

$$\langle c, \pi \rangle := \sum_{i=1}^N \sum_{j=1}^N \pi(i,j)\, c(i,j) \to \min, \qquad \pi \in \mathbb{R}^{N^2}, \tag{1}$$

under the constraints

$$\begin{aligned}
&\sum_{j=1}^N \pi(i,j) = N^{-1}, && i = 1, \dots, N, \\
&\sum_{i=1}^N \pi(i,j) = N^{-1}, && j = 1, \dots, N, \\
&\pi(i,j) \geq 0, && i,j = 1, \dots, N.
\end{aligned}$$

Of course, this is an easy and standard problem of linear optimization; yet we want to treat it in some detail in order to develop intuition and concepts for the general case considered in [BLS09] as well as in section 3 .

For the two sets of *equality* constraints we introduce $2N$ Lagrange multipliers $(\varphi(i))_{i=1}^N$ and $(\psi(j))_{j=1}^N$ taking values in $\mathbb{R}$, and for the inequality constraints (4) we introduce Lagrange multipliers $(\varrho_{ij})_{i,j=1}^N$ taking values in $\mathbb{R}_+$. The Lagrangian functional $L(\pi, \varphi, \psi, \varrho)$ then is given by

$$\begin{aligned}
L(\pi, \varphi, \psi, \varrho) = &\sum_{i=1}^N \sum_{j=1}^N c(i,j)\, \pi(i,j) \\
&- \sum_{i=1}^N \varphi(i) \left(\sum_{j=1}^N \pi(i,j) - N^{-1} \right) \\
&- \sum_{j=1}^N \psi(j) \left(\sum_{i=1}^N \pi(i,j) - N^{-1} \right) \\
&- \sum_{i=1}^N \sum_{j=1}^N \varrho(i,j) \pi(i,j),
\end{aligned}$$

where the $\pi(i,j), \varphi(i)$ and $\psi(j)$ range in $\mathbb{R}$, while the $\varrho(i,j)$ range in $\mathbb{R}_+$.

It is designed in such a way that

$$C(\pi) := \sup_{\varphi, \psi, \varrho} L(\pi, \varphi, \psi, \varrho) = \langle c, \pi \rangle + \chi_{\Pi(\mu,\nu)}(\pi),$$

where $\Pi(\mu, \nu)$ denotes the admissible set of π's, i.e., the probability measures on $X \times Y$ with marginals μ and ν, and $\chi_A(\ . \ )$ denotes the indicator function of a set A in the sense of convex function theory, i.e., taking the value 0 on A, and the value $+\infty$ outside of A.

In particular, we have

$$P := \inf_{\pi \in \mathbb{R}^{N^2}} C(\pi) = \inf_{\pi \in \mathbb{R}^{N^2}} \sup_{\varphi,\psi,\varrho} L(\pi, \varphi, \psi, \varrho),$$

where P is the optimal value of the primal optimization problem (1).

To develop the duality theory of the primal problem (1) we pass from *inf sup* L to *sup inf* L. Denote by $D(\varphi, \psi, \varrho)$ the dual function

$$\begin{aligned} D(\varphi, \psi, \varrho) &= \inf_{\pi \in \mathbb{R}^{N^2}} L(\pi, \varphi, \psi, \varrho) \\ &= \inf_{\pi \in \mathbb{R}^{N^2}} \sum_{i=1}^{N} \sum_{j=1}^{N} \pi(i,j)[c(i,j) - \varphi(i) - \psi(j) - \varrho(i,j)] \\ &\quad + N^{-1} \left[\sum_{i=1}^{N} \varphi(i) + \sum_{j=1}^{N} \psi(j) \right]. \end{aligned}$$

Hence we obtain as the optimal value of the dual problem

$$D := \sup_{\varphi,\psi,\varrho} D(\varphi, \psi, \varrho) = (\mathbb{E}_\mu[\varphi] + \mathbb{E}_\nu[\psi]) - \chi_\Psi(\varphi, \psi) \tag{2}$$

where Ψ denotes the admissible set of φ, ψ, i.e. satisfying

$$\varphi(i) + \psi(j) + \varrho(i,j) = c(i,j), \qquad 1 \leq i, j \leq N,$$

for some non-negative "slack variables" $\varrho_{i,j}$.

Let us show that there is no duality gap, i.e., the values of P and D coincide. Of course, in the present finite dimensional case, this equality as well as the fact that the *inf sup* (resp. *sup inf*) above is a *min max* (resp. *a max min*) easily follows from general compactness arguments. Yet we want to verify things directly using the idea of "complementary slackness" of the primal and the dual constraints (good references are, e.g. [PSU88, ET99, AE06]).

We apply "Fenchel's perturbation map" to explicitly show the equality $P = D$. Let $T : \mathbb{R}^{N^2} \to \mathbb{R}^N \times \mathbb{R}^N$ be the linear map defined as

$$T\left(\left(\pi(i,j)\right)_{1 \leq i,j \leq N} \right) = \left(\left(\sum_{j=1}^{N} \pi(i,j) \right)_{i=1}^{N}, \left(\sum_{i=1}^{N} \pi(i,j) \right)_{j=1}^{N} \right)$$

so that the problem (1) now can be phrased as

$$\langle c, \pi \rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} c(i,j)\, \pi(i,j) \to min, \qquad \pi \in \mathbb{R}_+^{N^2},$$

under the constraint

$$T(\pi) = \left((N^{-1}, \ldots, N^{-1}), (N^{-1}, \ldots, N^{-1}) \right).$$

The range of the linear map T is the subspace $E \subseteq \mathbb{R}^N \times \mathbb{R}^N$, of codimension 1, formed by the pairs (f, g) such that $\sum_{i=1}^{N} f(i) = \sum_{j=1}^{N} g(j)$, in other words $\mathbb{E}_\mu[f] = \mathbb{E}_\nu[g]$. We consider T as a map from $\mathbb{R}^{N^2}$ to E and denote by E_+ the positive orthant of E.

Let $\Phi : E_+ \to [0, \infty]$ be the map

$$\Phi(f,g) = \inf \left\{ \langle c, \pi \rangle, \ \pi \in \mathbb{R}_+^{N^2}, \ T(\pi) = (f,g) \right\}.$$

We shall verify explicitly that Φ is an $\mathbb{R}_+$-valued, convex, lower semi-continuous, positively homogeneous map on E_+.

The finiteness and positivity of Φ follow from the fact that, for $(f,g) \in E_+$, the set of $\pi \in \mathbb{R}_+^{N^2}$ with $T(\pi) = (f,g)$ is non-empty and from the non-negativity of c. As regards the convexity of Φ, let $(f_1, g_1), (f_2, g_2) \in E_+$ and find $\pi_1, \pi_2 \in \mathbb{R}_+^{N^2}$ such that $T(\pi_1) = (f_1, g_1)$, $T(\pi_2) = (f_2, g_2)$ and $\langle c, \pi_1 \rangle < \Phi(f_1, g_1) + \varepsilon$ as well as $\langle c, \pi_2 \rangle < \Phi(f_2, g_2) + \varepsilon$. Then

$$\Phi\left(\frac{(f_1, g_1) + (f_2, g_2)}{2}\right) \leq \left\langle c, \frac{\pi_1 + \pi_2}{2} \right\rangle < \frac{\Phi(f_1, g_1) + \Phi(f_2, g_2)}{2} + \varepsilon,$$

which proves the convexity of Φ.

If $((f_n, g_n))_{n=1}^\infty \in E_+$ converges to (f,g) find $(\pi_n)_{n=1}^\infty$ in $\mathbb{R}_+^{N^2}$ such that $T(\pi_n) = (f_n, g_n)$ and $\langle c, \pi_n \rangle < \Phi(f_n, g_n) + n^{-1}$. Note that $(\pi_n)_{n=1}^\infty$ is bounded in $\mathbb{R}_+^{N^2}$, so that there is a subsequence $(\pi_{n_k})_{k=1}^\infty$ converging to $\pi \in \mathbb{R}_+^{N^2}$. Hence $\Phi(f, g,) \leq \langle c, \pi \rangle$ showing the lower semi-continuity of Φ. Finally note that Φ is positively homogeneous, i.e., $\Phi(\lambda f, \lambda g) = \lambda \Phi(f,g)$, for $\lambda \geq 0$.

The point (f_0, g_0) with $f_0 = g_0 = (N^{-1}, \ldots, N^{-1})$ is in E_+ and Φ is bounded in a neighbourhood V of (f_0, g_0). Indeed, fixing any $0 < a < N^{-1}$ the subsequent set V does the job

$$V = \{(f,g) \in E \ : |f(i) - N^{-1}| < a, \ |g(j) - N^{-1}| < a, \quad \text{for } 1 \leq i,j \leq N\}.$$

The boundedness of the lower semi-continuous convex function Φ on V implies that the subdifferential of Φ at (f_0, g_0) is non-empty. Considering Φ as a function on $\mathbb{R}^{2N}$ (by defining it to equal $+\infty$ on $\mathbb{R}^{2N} \backslash E_+$) we may find an element $(\widehat{\varphi}, \widehat{\psi}) \in \mathbb{R}^N \times \mathbb{R}^N$ in this subdifferential. By the positive homogeneity of Φ we have

$$\Phi(f,g) \geq \langle (\widehat{\varphi}, \widehat{\psi}), (f,g) \rangle = \langle \widehat{\varphi}, f \rangle + \langle \widehat{\psi}, g \rangle, \quad \text{for } (f,g) \in \mathbb{R}^N \times \mathbb{R}^N,$$

and

$$P = \Phi(f_0, g_0) = \langle \widehat{\varphi}, f_0 \rangle + \langle \widehat{\psi}, g_0 \rangle.$$

By the definition of Φ we therefore have, for each $\pi \in \mathbb{R}_+^{N^2}$,

$$\begin{aligned}
\langle c, \pi \rangle &\geq \inf_{\tilde{\pi} \in \mathbb{R}_+^{N^2}} \{ \langle c, \tilde{\pi} \rangle : T(\pi) = T(\tilde{\pi}) \} \\
&= \Phi(T(\pi)) \\
&\geq \langle T(\pi), (\hat{\varphi}, \hat{\psi}) \rangle \\
&= \sum_{i=1}^N \sum_{j=1}^N \pi(i,j) \, [\widehat{\varphi}(i) + \widehat{\psi}(j)]
\end{aligned}$$

so that

$$c(i,j) \geq \widehat{\varphi}(i) + \widehat{\psi}(j), \qquad \text{for } 1 \leq i,j \leq n. \tag{3}$$

By compactness, there is $\widehat{\pi} \in \Pi(\mu, \nu)$, i.e., there is an element $\widehat{\pi} \in \mathbb{R}_+^{N^2}$ verifying $T(\widehat{\pi}) = (f_0, g_0)$ such that

$$\langle c, \widehat{\pi} \rangle = \langle \widehat{\varphi} + \widehat{\psi}, \widehat{\pi} \rangle. \tag{4}$$

Summing up, we have shown that $\widehat{\pi}$ and $(\widehat{\varphi}, \widehat{\psi})$ are primal and dual optimizers and that the value of the primal problem equals the value of the dual problem, namely $\langle \widehat{\varphi} + \widehat{\psi}, \widehat{\pi} \rangle$.

To finish this elementary treatment of the finite case, let us consider the case when we allow the cost function c to take values in $[0, \infty]$ rather than in $[0, \infty[$. In this case the primal problem simply loses some dimensions: for the (i, j)'s where $c(i, j) = \infty$ we must have $\pi(i, j) = 0$ so that we consider

$$\langle c, \pi \rangle := \sum_{i=1}^{N} \sum_{j=1}^{N} \pi(i, j)\, c(i, j) \to \min, \qquad \pi \in \mathbb{R}_{+}^{N^2},$$

where we now optimize over $\pi \in \mathbb{R}_{+}^{N^2}$ with $\pi(i, j) = 0$ if $c(i, j) = \infty$. For the problem to make sense we clearly must have that there is at least one $\pi \in \Pi(\mu, \nu)$ with $\langle c, \pi \rangle < \infty$. If this non-triviality condition is satisfied, the above arguments carry over without any non-trivial modification.

We now analyze explicitly the well-known "complementary slackness conditions" and interpret them in the present context. For a pair $\widehat{\pi}$ and $(\widehat{\varphi}, \widehat{\psi})$ of primal and dual optimizers we have

$$c(i, j) > \widehat{\varphi}(i) + \widehat{\psi}(j) \;\Rightarrow\; \widehat{\pi}(i, j) = 0,$$

and

$$\widehat{\pi}(i, j) > 0 \;\Rightarrow\; c(i, j) = \widehat{\varphi}(i) + \widehat{\psi}(j).$$

Indeed, these relations follow from the admissibility condition $c \geq \hat{\varphi} + \hat{\psi}$ and the duality relation $\langle \widehat{\pi}, c - (\widehat{\varphi} + \widehat{\psi}) \rangle = 0$.

This motivates the following definitions in the theory of optimal transport (see, e.g., [RR96] for (a) and [ST08] for (b).)

Definition 2.1. *Let $X = Y = \{1, \ldots, N\}$ and $\mu = \nu$ the uniform distribution on X and Y respectively, and let $c : X \times Y \to \mathbb{R}_{+}$ be given.*

(a) *A subset $\Gamma \subseteq X \times Y$ is called "cyclically c-monotone" if, for $(i_1, j_1), \ldots, (i_n, j_n) \in \Gamma$ we have*

$$\sum_{k=1}^{n} c(i_k, j_k) \leq \sum_{k=1}^{n} c(i_k, j_{k+1}), \tag{5}$$

where $j_{n+1} = j_1$.

(b) *A subset $\Gamma \subseteq X \times Y$ is called "strongly cyclically c-monotone" if there are functions φ, ψ such that $\varphi(i) + \psi(j) \leq c(i, j)$, for all $(i, j) \in X \times Y$, with equality holding true for $(i, j) \in \Gamma$.*

In the present finite setting, the following facts are rather obvious (assertion (iii) following from the above discussion):

(i) The support of each primal optimizer $\widehat{\pi}$ is cyclically c-monotone.

(ii) Every $\pi \in \Pi(\mu, \nu)$ which is supported by a cyclically c-monotone set Γ, is a primal optimizer.

(iii) A set $\Gamma \subseteq X \times Y$ is cyclically c-monotone iff it is strongly cyclically c-monotone.

In general, one may ask, for a given Monge–Kantorivich transport optimization problem, defined on polish spaces X, Y, equipped with Borel probability measures μ, ν, and a Borel measurable cost function $c : X \times Y \to [0, \infty]$, the following natural questions:

(P) Does there exist a primal optimizer to (1), i.e. a Borel measure $\widehat{\pi} \in \Pi(\mu, \nu)$ with marginals μ, ν, such that

$$\int_{X\times Y} c \, d\widehat{\pi} = \inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c \, d\pi =: P$$

holds true?

(D) Do there exist dual optimizers to (2), i.e. Borel functions $(\widehat{\varphi}, \widehat{\psi})$ in $\Psi(\mu, \nu)$ such that

$$\int_X \widehat{\varphi} \, d\mu + \int_Y \widehat{\psi} \, d\nu = \sup_{(\varphi,\psi)\in\Psi(\mu,\nu)} \left(\int_X \varphi \, d\mu + \int_Y \psi \, d\nu \right) =: D, \tag{6}$$

where $\Psi(\mu, \nu)$ denotes the set of all pairs of $[-\infty, +\infty[$-valued integrable Borel functions (φ, ψ) on X, Y such that $\varphi(x) + \psi(y) \leq c(x, y)$, for all $(x, y) \in X \times Y$?

(DG) Is there a duality gap, or do we have $P = D$, as it should − morally speaking − hold true?

These are three natural questions which arise in every convex optimization problem. In addition, one may ask the following two questions pertaining to the special features of the Monge–Kantorovich transport problem.

(CC) Is every cyclically c-monotone transport plan $\pi \in \Pi(\mu, \nu)$ optimal, where we call $\pi \in \Pi(\mu, \nu)$ cyclically c-monotone if there is a Borel subset $\Gamma \subseteq X \times Y$ of full support $\pi(\Gamma) = 1$, verifying condition (5), for any $(x_1, y_1), \ldots, (x_n, y_n) \in \Gamma$?

(SCC) Is every strongly cyclically c-monotone transport plan $\pi \in \Pi(\mu, \nu)$ optimal, where we call $\pi \in \Pi(\mu, \nu)$ strongly cyclically c-monotone if there are Borel functions $\varphi : X \to [-\infty, +\infty[$ and $\psi : Y \to [-\infty, +\infty[$, satisfying $\varphi(x) + \psi(y) \leq c(x, y)$, for all $(x, y) \in X \times Y$, and $\pi\{\varphi + \psi = c\} = 1$?

Much effort has been made over the past decades to provide increasingly general answers to the questions above. We mention the work of Rüschendorf [Rüs96] who adapted the notion of cyclical monotonicity from Rockafellar [Roc66]. Rockafellar's work pertains to the case $c(x, y) = -\langle x, y\rangle$, for $x, y \in \mathbb{R}^n$, while Rüschendorf's work pertains to the present setting of general cost functions c, thus arriving at the notion of cyclical c-monotonicity. Intimately related is the notion of the c-conjugate φ^c of a function φ.

We also mention G. Kellerer's fundamental work on the duality theory; in [Kel84] he established that $P = D$ provided that $c : X \times Y \to [0, \infty]$ is lower semi-continous, or merely Borel-measurable and uniformly bounded.

The seminal paper [GM96] proves (among many other results) that we have a positive answer to question (CC) above in the following situation: every cyclically c-monotone transport plan is optimal provided that the cost function c is continuous and X, Y are compact subsets of $\mathbb{R}^n$. In [Vil03, Problem 2.25] it is asked whether this extends to the case $X = Y = \mathbb{R}^n$ with the squared euclidian distance as cost function. This was answered independently in [Pra08] and [ST08]: the answer to (CC) is positive for general polish spaces X and Y, provided that the cost function $c : X \times Y \to [0, \infty]$ is continuous ([Pra08]) or lower semi-continuous and finitely valued ([ST08]). Indeed, in the latter case, a transport plan is optimal if and only if it is strongly c-monotone.

Let us briefly resume the state of the art pertaining to the five questions above.

As regards the most basic issue, namely (DG) pertaining to the question whether duality makes sense at all, this is analyzed in detail — building on a lot of previous literature — in section 2 of the accompanying paper [BLS09]: it is shown there that, for a properly relaxed version of the primal problem, question (DG) has an affirmative answer in a perfectly general setting, i.e. for arbitrary Borel-measurable cost functions $c : X \times Y \to [0, \infty]$ defined on the product of two polish spaces X, Y, equipped with Borel probability measures μ, ν.

As regards question (P) we find the following situation: if the cost function $c : X \times Y \to [0, \infty]$ is *lower semi-continuous*, the answer to question (P) is always positive. Indeed, for an optimizing sequence $(\pi_n)_{n=1}^\infty$ in $\Pi(\mu, \nu)$, one may apply Prokhorov's theorem to find a weak limit $\widehat{\pi} = \lim_{k\to\infty} \pi_{n_k}$. If c is lower semi-continuous, we get

$$\int\limits_{X\times Y} c \, d\widehat{\pi} \leq \lim_{k\to\infty} \int\limits_{X\times Y} c \, d\pi_{n_k},$$

which yields the optimality of $\widehat{\pi}$.

On the other hand, if c fails to be lower semi-continuous, there is little reason why a primal optimizer should exist (see, e.g., [Kel84, Example 2.20]).

As regards (D), the question of the existence of a dual optimizer is more delicate than for the primal case (P): it was shown in [AP03, Theorem 3.2] that, for $c : X \times Y \to \mathbb{R}_+$, satisfying a certain moment condition, one may assert the existence of *integrable* optimizers $(\widehat{\varphi}, \widehat{\psi})$. However, if one drops this moment condition, there is little reason why, for an optimizing sequence $(\varphi_n, \psi_n)_{n=1}^\infty$ in (D) above, the L^1-norms should remain bounded. Hence there is little reason why one should be able to find *integrable* optimizers $(\widehat{\varphi}, \widehat{\psi})$ as shown by easy examples (e.g. [BS09, Examples 4.4, 4.5]), arising in rather regular situations.

Yet one would like to be able to pass to *some kind of limit* $(\widehat{\varphi}, \widehat{\psi})$, whether these functions are integrable or not. In the case when $\widehat{\varphi}$ and/or $\widehat{\psi}$ fail to be integrable, special care then has to be taken to give a proper sense to (6).

This situation was the motivation for the introduction of the notion of *strong cyclical c-monotonicity* in [ST08]: this notion (see (SCC) above) characterizes the optimality of a given $\pi \in \Pi(\mu, \nu)$ in terms of a "complementary slackness condition", involving some $(\varphi, \psi) \in \Psi(\mu, \nu)$, playing the role of a dual optimizer $(\widehat{\varphi}, \widehat{\psi})$. The crucial feature is that *we do not need any integrability of the functions φ and ψ* for this notion to make sense. It was shown in [BS09] that, also in situations where there are no integrable optimizers $(\widehat{\varphi}, \widehat{\psi})$, one may find Borel measurables functions (φ, ψ), taking their roles in the setting of (SCC) above.

This theme was further developed in [BS09], where it was shown that, for $\mu \otimes \nu$-a.s. finite, Borel measurable $c : X \times Y \to [0, \infty]$, one may find Borel functions $\widehat{\varphi} : X \to [-\infty, +\infty)$ and $\widehat{\psi} : Y \to [-\infty, \infty)$, which are dual optimizers if we interpret (6) properly: instead of considering

$$\int\limits_X \widehat{\varphi} \, d\mu + \int\limits_Y \widehat{\psi} \, d\nu, \tag{7}$$

which needs integrability of $\widehat{\psi}$ and $\widehat{\psi}$ in order to make sense, we consider

$$\int\limits_{X\times Y} (\widehat{\varphi}(x) + \widehat{\psi}(y)) \, d\pi(x, y), \tag{8}$$

where the transport plan $\pi \in \Pi(\mu,\nu)$ is assumed to have finite transport cost $\int_{X\times Y} c(x,y)d\pi(x,y) < \infty$. If (7) makes sense, then its value coincides with the value of (8); the crucial feature is that, (8) also makes sense in cases when (7) does not make sense any more as shown in [BS09, Lemma 1.1]. In particular, the value of (8) does not depend on the choice of the transport plan $\pi \in \Pi(\mu,\nu)$, provided π has finite transport cost $\int_{X\times Y} c(x,y)d\pi(x,y) < \infty$.

Summing up the preceding discussion on the existence (D) of a dual optimizer $(\widehat{\varphi},\widehat{\psi})$: this question has a – properly interpreted – positive answer provided that the cost function $c : X\times Y \to [0,\infty]$ is $\mu\otimes\nu$-a.s. finite ([BS09, Theorem 2]).

But things become much more complicated if we pass to cost functions $c : X\times Y \to [0,\infty]$ assuming the value $+\infty$ on possibly "large" subsets of $X\times Y$.

In [BLS09, Example 4.1] we exhibit an example, which is a variant of an example due to G. Ambrosio and A. Pratelli [AP03, Example 3.5], of a lower semicontinuous cost function $c : [0,1)\times[0,1)\to[0,\infty]$, where $(X,\mu)=(Y,\nu)$ equals $[0,1)$ equipped with Lebesgue measure, for which there are no *Borel measurable* functions $\widehat{\varphi},\widehat{\psi}$ verifying $\widehat{\varphi}(x)+\widehat{\psi}(y)\le c(x,y)$, maximizing (8) above.

In this example, the cost function c equals the value $+\infty$ on "many" points of $X\times Y = [0,1)\times[0,1)$. In fact, for each $x\in[0,1[$, there are precisely two points $y_1,y_2\in[0,1[$ such that $c(x,y_1)<\infty$ and $c(x,y_2)<\infty$, while for all other $y\in[0,1[$, we have $c(x,y)=\infty$. In addition, there is an optimal transport plan $\widehat{\pi}\in\Pi(\mu,\nu)$ whose support equals the set $\{(x,y)\in[0,1)\times[0,1) : c(x,y)<\infty\}$.

In this example one may observe the following phenomenon: while there *do not* exist Borel measurable functions $\widehat{\varphi} : [0,1)\to[-\infty,+\infty)$ and $\widehat{\psi} : [0,1)\to[-\infty,\infty)$ such that $\widehat{\varphi}(x)+\widehat{\psi}(y) = c(x,y)$ on $\{c(x,y)<\infty\}$, there *does* exist a Borel function $\widehat{h} : [0,1)\times[0,1)\to[-\infty,\infty)$ such that $\widehat{h}(x,y)=c(x,y)$ on $\{c(x,y)<\infty\}$ and such that $\widehat{h}(x,y) = \lim_{n\to\infty}(\varphi_n(x)+\psi_n(y))$ where $(\varphi_n,\psi_n)_{n=1}^{\infty}$ are properly chosen, bounded Borel functions. The point is that the limit holds true (only) in the norm of $L^1([0,1[\times[0,1[,\widehat{\pi})$ as well as $\widehat{\pi}$-a.s.

In other words, in this example we are able to identify some kind of dual optimizer $\widehat{h}\in L^1([0,1)\times[0,1),\widehat{\pi})$ which, however, is not of the form $\widehat{h}(x,y)=\widehat{\varphi}(x)+\widehat{\psi}(y)$ for some Borel functions $(\widehat{\varphi},\widehat{\psi})$, but only a $\widehat{\pi}$-a.s. limit of such functions $(\varphi_n(x)+\psi_n(y))_{n=1}^{\infty}$.

In [BLS09, Theorem 4.2] we established a result which shows that much of the positive aspect of this phenomenon, i.e. the existence of an optimal $\widehat{h}\in L^1(\widehat{\pi})$, encountered in the context of the above example, can be carried over to a general setting. For the convenience of the reader we restate this theorem and the notations required to formulate it.

Fix a finite transport plan $\pi_0\in\Pi(\mu,\nu,c) := \Big\{\pi\in\Pi(\mu,\nu) : \int_{X\times Y} c\,d\pi<\infty\Big\}$. We denote by $\Pi^{(\pi_0)}(\mu,\nu)$ the set of elements $\pi\in\Pi(\mu,\nu)$ such that $\pi\ll\pi_0$ and $\big\|\frac{d\pi}{d\pi_0}\big\|_{L^\infty(\pi_0)}<\infty$. Note that $\Pi^{(\pi_0)}(\mu,\nu)=\Pi(\mu,\nu)\cap L^\infty(\pi_0)\subseteq\Pi(\mu,\nu,c)$. We shall replace the usual Kantorovich optimization problem over the set $\Pi(\mu,\nu,c)$ by the optimization over the smaller set $\Pi^{(\pi_0)}(\mu,\nu)$. Its value is

$$P^{(\pi_0)} = \inf\{\langle c,\pi\rangle = \textstyle\int c\,d\pi : \pi\in\Pi^{(\pi_0)}(\mu,\nu)\}. \tag{9}$$

As regards the dual problem, we define, for $\varepsilon>0$,

$$\begin{aligned} D^{(\pi_0,\varepsilon)} &= \sup\Big\{\int\varphi\,d\mu+\int\psi\,d\nu :\ \varphi\in L^1(\mu),\psi\in L^1(\nu), \\ &\qquad\qquad \int_{X\times Y}(\varphi(x)+\psi(y)-c(x,y))_+\,d\pi_0\le\varepsilon\Big\} \quad \text{and} \\ D^{(\pi_0)} &= \lim_{\varepsilon\to 0} D^{(\pi_0,\varepsilon)}. \end{aligned} \tag{10}$$

Define the "summing" map S by

$$S : L^1(X,\mu) \times L^1(Y,\nu) \to L^1(X \times Y, \pi_0)$$
$$(\varphi,\psi) \mapsto \varphi \oplus \psi,$$

where $\varphi \oplus \psi$ denotes the function $\varphi(x) + \psi(y)$ on $X \times Y$. Denote by $L^1_S(X \times Y, \pi_0)$ the $\|.\|_1$-closed linear subspace of $L^1(X \times Y, \pi_0)$ spanned by $S(L^1(X,\mu) \times L^1(Y,\nu))$. Clearly $L^1_S(X \times Y, \pi_0)$ is a Banach space under the norm $\|.\|_1$ induced by $L^1(X \times Y, \pi_0)$.

We shall also need the bi-dual $L^1_S(X \times Y, \pi_0)^{**}$ which may be identified with a subspace of $L^1(X \times Y, \pi_0)^{**}$. In particular, an element $h \in L^1_S(X \times Y, \pi_0)^{**}$ can be decomposed into $h = h^r + h^s$, where $h^r \in L^1(X \times Y, \pi_0)$ is the regular part of the finitely additive measure h and h^s its purely singular part.

Theorem 2.2. *Let $c : X \times Y \to [0,\infty]$ be Borel measurable, and let $\pi_0 \in \Pi(\mu,\nu,c)$ be a finite transport plan. We have*

$$P^{(\pi_0)} = D^{(\pi_0)}. \tag{11}$$

There is an element $\hat{h} \in L^1_S(X \times Y, \pi_0)^{**}$ such that $\hat{h} \le c$ and

$$D^{(\pi_0)} = \langle \hat{h}, \pi_0 \rangle.$$

If $\pi \in \Pi^{(\pi_0)}(\mu,\nu)$ (identifying π with $\frac{d\pi}{d\pi_0}$) satisfies $\int c\, d\pi \le P^{(\pi_0)} + \alpha$ for some $\alpha \ge 0$, then

$$|\langle \hat{h}^s, \pi \rangle| \le \alpha. \tag{12}$$

In particular, if π is an optimizer of (9), then $\hat{h}^s$ vanishes on the set $\{\frac{d\pi}{d\pi_0} > 0\}$.
In addition, we may find a sequence of elements $(\varphi_n,\psi_n) \in L^1(\mu) \times L^1(\nu)$ such that

$$\varphi_n \oplus \psi_n \to \hat{h}^r,\ \pi_0\text{-}a.s., \qquad \|(\varphi_n \oplus \psi_n - \hat{h}^r)_+\|_{L_1(\pi_0)} \to 0$$

and

$$\lim_{\delta \to 0} \sup_{A \subseteq X \times Y, \pi_0(A) < \delta} \lim_{n \to \infty} -\langle (\varphi_n \oplus \psi_n) \mathbf{1}_A, \pi_0 \rangle = \|\hat{h}^s\|_{L_1(\pi_0)^{**}}. \tag{13}$$

The assertion of the theorem extends the phenomenon of [BLS09, Example 4.1] to a general setting. There is, however, one additional complication, as compared to the situation of this specific example: in the above theorem we only can assert that we find the optimizer $\widehat{h}$ in $L^1(\widehat{\pi})^{**}$ rather than in $L^1(\widehat{\pi})$. The question arises whether this complication is indeed unavoidable. The purpose of the subsequent section is to construct an example showing that the phenomenon of a non-vanishing singular part $\widehat{h}^s$ of $\widehat{h} = \widehat{h}^r + \widehat{h}^s$ may indeed arise in the above setting. In addition, the example gives a good illustration of the subtleties of the situation described by the theorem above.

3 The singular part of the dual optimizer

In this section we refine the construction of Examples 4.1 and 4.3 in [BLS09] (which in turn are variants of an example due to G. Ambrosio and A. Pratelli [AP03, Example 3.2]). We assume that the reader is familiar with these examples and freely use the notation from this paper.
In particular, for an irrational $\alpha \in [0,1)$ we write, for $k \in \mathbb{Z}$, [1]

$$\begin{aligned} \varrho_k(x) = 1 &+ \#\{0 \le i < k : x \oplus i\alpha \in [0, \tfrac{1}{2})\} \\ &- \#\{0 \le i < k : x \oplus i\alpha \in [\tfrac{1}{2}, 1)\}, \end{aligned} \tag{14}$$

[1] In [BLS09] the constructions are carried out for $\mathbb{N}$ instead of $\mathbb{Z}$, but for our purposes the latter choice turns out to be better suited.

where, for $k < 0$, we mean by $0 \le i < k$ the set $\{k+1, k+2, \dots, 0\}$ and $\oplus$ denotes addition modulo 1. We also recall that the function $h : [0,1) \times [0,1) \to \mathbb{Z}$ is defined in [BLS09, Example 4.3] as

$$h(x,y) = \begin{cases} \varrho_k(x), & k \in \mathbb{Z} \text{ and } y = x \oplus k\alpha \\ \infty, & \text{otherwise.} \end{cases} \tag{15}$$

In [BLS09, Example 4.3] we considered the $[0,\infty]$-valued cost function $c(x,y) := h_+(x,y)$, the positive part of the function h. We now construct an example restricting $h_+(x,y)$ to a certain subset of $[0,1) \times [0,1)$.

Example 3.1. *Consider $X = Y = [0,1)$ and denote by μ resp. ν the Lebesgue measure on X, resp. Y. There is an irrational $\alpha \in [0,1)$ and a map $\tau : [0,1) \to \mathbb{Z}$ such that, for*

$$\begin{aligned} \Gamma_0 &= \{(x,x), x \in [0,1)\}, \\ \Gamma_1 &= \{(x, x \oplus \alpha) : x \in [0,1)\}, \\ \Gamma_\tau &= \{(x, x \oplus \tau(x)\alpha) : x \in [0,1)\} \end{aligned}$$

and letting

$$c(x,y) = \begin{cases} h_+(x,y), & \textit{for } x \in \Gamma_0 \cup \Gamma_1 \cup \Gamma_\tau \\ \infty, & \textit{otherwise} \end{cases}$$

the following properties are satisfied.

(i) The maps

$$T^0_\alpha(x) = x, \qquad T^1_\alpha(x) = x \oplus \alpha, \qquad T^{(\tau)}_\alpha(x) = x \oplus (\tau(x)\,\alpha)$$

are measure preserving bijections from $[0,1)$ to $[0,1)$, with respect Denote by π_0, π_1, π_τ the corresponding transport plans in $\Pi(\mu,\nu)$, i.e.

$$\pi_0 = (id, id)_\# \mu, \quad \pi_1 = (id, T_\alpha)_\# \mu, \quad \pi_\tau = (id, T^{(\tau)}_\alpha)_\# \mu,$$

and let $\pi = (\pi_0 + \pi_1 + \pi_\tau)/3$.

(ii) The transport plans π_0 and π_1 are optimal while π_τ is not. In fact, we have

$$\langle c, \pi_0 \rangle = \langle c, \pi_1 \rangle = 1 \textit{ while } \langle c, \pi_\tau \rangle \ge \langle h, \pi_\tau \rangle > 1. \tag{16}$$

(iii) There is a sequence $(\varphi_n, \psi_n)_{n=1}^\infty$ of bounded Borel functions such that

$$(a)\ \varphi_n(x) + \psi_n(y) \le c(x,y), \quad \textit{for } x \in X, y \in Y, \tag{17}$$

$$(b)\ \lim_{n\to\infty} \Big(\int_X \varphi_n(x)\, d\mu(x) + \int_Y \psi_n(y), d\nu(y) \Big) = 1, \tag{18}$$

$$(c)\ \lim_{n\to\infty} (\varphi_n(x) + \psi_n(y)) = h(x,y),\ \pi\textit{-almost surely.} \tag{19}$$

(iv) Using the notation of [BLS09, Theorem 4.2] we find that for each dual optimizer $\widehat{h} \in L^1(\pi)^{**}$, which decomposes as $\widehat{h} = \widehat{h}^r + \widehat{h}^s$ into its regular part $\widehat{h}^r \in L^1(\pi)$ and its purely singular part $\widehat{h}^s \in L^1(\pi)^{**}$, we have

$$\widehat{h}^r = h,\ \pi\textit{-a.s.}, \tag{20}$$

and the singular part $\widehat{h}^s$ satisfies $\|\widehat{h}^s\|_{L^1(\pi)^{**}} = \langle h, \pi_\tau \rangle - 1 > 0$. In particular, the singular part $\widehat{h}^s$ of $\widehat{h}$ does not vanish. The finitely additive measure $\widehat{h}^s$ is supported by Γ_τ, i.e. $\langle \widehat{h}^s, \mathbf{1}_{\Gamma_0} + \mathbf{1}_{\Gamma_1} \rangle = 0$.

We shall use a special irrational $\alpha \in [0,1)$, namely

$$\alpha = \sum_{j=1}^{\infty} \frac{1}{M_j},$$

where $M_j = m_1 m_2 \dots m_j = M_{j-1} m_j$, and $(m_j)_{j=1}^{\infty}$ is a sequence of prime numbers $m_j \geq 5$ tending sufficiently fast to infinity, to be specified below. We let

$$\alpha_n := \sum_{j=1}^{n} \frac{1}{M_j},$$

which, of course, is a rational number.

We will need the following lemma. We thank Leonhard Summerer for showing us the proof of Lemma 3.2.

Lemma 3.2. *It is possible to choose a sequence $m_1, m_2, \dots$ of primes growing arbitrarily fast to infinity, such that with $M_1 = m_1, M_2 = m_1 \cdot m_2, \dots, M_n = m_1 \cdots m_n, \dots$ we have, for each $n \in \mathbb{N}$,*

$$\sum_{j=1}^{n} \frac{1}{M_j} = \frac{P_n}{M_n},$$

with P_n and M_n relatively prime.

Proof. We have

$$\sum_{j=1}^{n} \frac{1}{M_j} = \frac{m_2 \dots m_n + \dots + m_n + 1}{M_n} =: \frac{P_n}{M_n},$$

thus P_n and M_n are relatively prime, if and only if

$$\begin{aligned}
m_1 \nmid \quad & m_2 \cdots m_n + m_3 \cdots m_n + \dots + \quad m_n + 1 && (21)\\
m_2 \nmid \quad & m_3 \cdots m_n + \dots + \quad m_n + 1 && (22)\\
\vdots \quad & \qquad\qquad\qquad \vdots && (23)\\
m_{n-1} \nmid \quad & m_n + 1. && (24)
\end{aligned}$$

We claim that these conditions are, e.g., satisfied provided that we choose $m_1, m_2, \dots$ such that $m_i \geq 3$ and

$$\begin{aligned}
m_{i+1} &\equiv +1 \ (m_i) && (25)\\
m_{i+j} &\equiv -1 \ (m_i) \text{ if } j \geq 2. && (26)
\end{aligned}$$

for all $i \geq 1$. Indeed (25), (26) imply that for $k \in \{1, \dots, n-1\}$ we have modulo (m_k)

$$\begin{array}{cccccc}
m_{k+1} \cdots m_n + & m_{k+2} \cdots m_n + & m_{k+3} \cdots m_n + & \dots + & m_n + & 1 \equiv \\
(\pm 1) + & (\pm 1) + & (\mp 1) + & \dots + & (-1) + & (+1),
\end{array}$$

where in the second line the $(n-k+1)$ summands start to alternate after the second term. Thus, for even $n-k$, this amounts to

$$\begin{array}{cccccc}
m_{k+1} \cdots m_n + & m_{k+2} \cdots m_n + & m_{k+3} \cdots m_n + & \dots + & m_n + & 1 \equiv \\
(-1) + & (-1) + & (+1) + & \dots + & (-1) + & (+1) \equiv -1,
\end{array}$$

while we obtain, for odd $n-k$,

$$\begin{array}{cccccc} m_{k+1}\cdots m_n+ & m_{k+2}\cdots m_n+ & m_{k+3}\cdots m_n+ & \ldots+ & m_n+ & 1\equiv \\ (+1)+ & (+1)+ & (-1)+ & \ldots+ & (-1)+ & (+1)\equiv +2 \end{array}$$

Hence (21)-(24) are satisfied as the m_n where chosen such that $m_n > 2$.

We use induction to construct a sequence of primes satisfying (25) and (26). Assume that $m_1, \ldots, m_i$ have been defined. By the chinese remainder theorem the system of congruences

$$x \equiv -1\ (m_1), \ldots, \quad x \equiv -1\ (m_{i-1}), \quad x \equiv +1\ (m_i)$$

has a solution $x_0 \in \{1, \ldots, m_1 \ldots m_i\}$. By Dirichlet's theorem, the arithmetic progression $x_0 + km_1 \ldots m_i, k \in \mathbb{N}$ contains infinitely many primes, so we may pick one which is as large as we please. The induction continues. □

For $\beta \in [0,1)$, denote by $T_\beta : [0,1) \to [0,1), T_\beta(x) := x \oplus \beta$ the addition of β modulo 1. With this notation we have $T_{\alpha_n}^{M_n} = id$ and, by Lemma 3.2, it is possible to choose $m_1, \ldots, m_n$ in such a way that M_n is the smallest such number in $\mathbb{N}$. Our aim is to construct a function $\tau : [0,1) \to \mathbb{Z}$ such that the map

$$T_\alpha^{(\tau)} : \begin{cases} [0,1) & \to & [0,1) \\ x & \mapsto & T_\alpha^{(\tau)}(x) = T_\alpha^{\tau(x)}(x) \end{cases},$$

defines, up to a μ-null set, a measure preserving bijection on $[0,1)$, and such that the corresponding transport plan $\pi_\tau \in \Pi(\mu,\nu)$, given by $\pi_\tau = (id, T_\alpha^{(\tau)})_\# \mu$, has the properties listed above with respect to the cost function $c(x,y)$ which is the restriction of the function $h_+(x,y)$ to $\Gamma_0 \cup \Gamma_1 \cup \Gamma_\tau$. We shall do so by an inductive procedure, defining bounded $\mathbb{Z}$-valued functions τ_n on $[0,1)$ such that the maps $T_{\alpha_n}^{(\tau_n)}$ are measure preserving bijections on $[0,1)$. The map $T_\alpha^{(\tau)}$ then will be the limit of these $T_{\alpha_n}^{(\tau_n)}$.

Step n=1: Fix a prime $M_1 = m_1 \geq 5$, so that $\alpha_1 = \frac{1}{M_1}$. Define

$$I_{k_1} := \left[\tfrac{k_1-1}{M_1}, \tfrac{k_1}{M_1}\right),\ k_1 = 1, \ldots, M_1,$$

so that $(I_{k_1})_{k_1=1}^{M_1}$ forms a partition of $[0,1)$ and T_{α_1} maps I_{k_1} to I_{k_1+1}, with the convention $I_{M_1+1} = I_1$. We also introduce the notations

$$L^1 := [0, \tfrac{1}{2} - \tfrac{1}{2M_1}) \text{ and } R^1 := [\tfrac{1}{2} + \tfrac{1}{2M_1}, 1)$$

for the segments left and right of the middle interval

$$I^1_{\text{middle}} := I_{(M_1+1)/2} = [\tfrac{1}{2} - \tfrac{1}{2M_1}, \tfrac{1}{2} + \tfrac{1}{2M_1}).$$

We define the functions φ^1, ψ^1 on $[0,1)$ such that $\varphi^1(x) + \psi^1(x) \equiv 1$ and

$$\varphi^1(x) + \psi^1(T_{\alpha_1}(x)) = \begin{cases} 0 & x \in L^1 \\ 1 & x \in I^1_{\text{middle}} \\ 2 & x \in R^1 \end{cases}$$

which leads to the relation

$$\varphi^1(T_{\alpha_1}(x)) = \varphi^1(x) + \begin{cases} 1, & x \in L^1, \\ 0, & x \in I^1_{\text{middle}}, \\ -1, & x \in R^1. \end{cases}$$

Making the choice $\varphi^1 \equiv 0$ on I_1 this leads to

$$\varphi^1(x) = \begin{cases} k_1 - 1, & x \in I_{k_1}, k_1 \in \{1, \dots, (M_1+1)/2\}, \\ M_1 + 1 - k_1, & x \in I_{k_1}, k_1 \in \{(M_1+3)/2, M_1\}, \end{cases} \tag{27}$$
$$\psi^1(x) = 1 - \varphi^1(x).$$

The function φ^1 starts at 0, increases until the middle interval, stays constant when stepping to the interval right of the middle, and then decreases, reaching 1 on the final interval I_{M_1}.

The idea is to define the map $\tau_1 : [0,1) \to \mathbb{Z}$ in such a way that the map

$$T_{\alpha_1}^{(\tau_1)} : \begin{cases} [0,1) & \to & [0,1) \\ x & \mapsto & T_{\alpha_1}^{\tau_1(x)}(x) \end{cases},$$

is a measure preserving bijection enjoying the following property: the map

$$x \mapsto \varphi^1(x) + \psi^1(T_{\alpha_1}^{(\tau_1)}(x)),$$

equals the value two on a large set while it has concentrated a negative mass which is close to -1 on a small set.

This can be done, e.g., by shifting the first interval I_1 to the interval $I_{(M_1-1)/2}$, which is left of the middle one, while we shift the intervals $I_2, \dots, I_{(M_1-1)/2}$ by one interval to the left. On the right hand side of $[0,1)$ we proceed symmetrically while the middle interval simply is not moved.

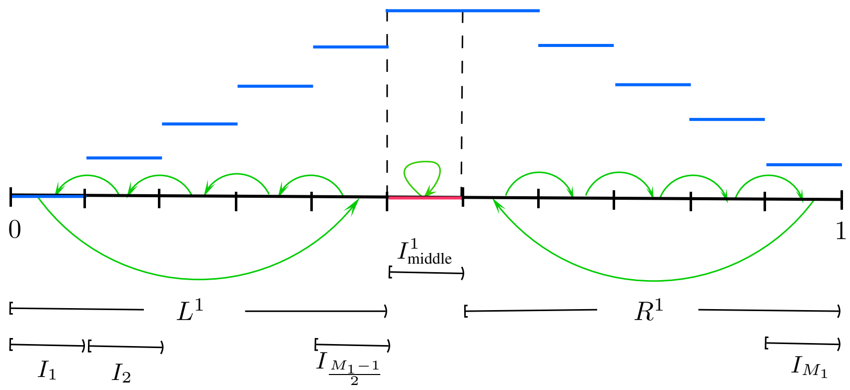

Fig. 1. Representations of φ^1 and τ^1.

The step function is φ^1 and the arrows indicate the action of $T_{\alpha_1}^{(\tau_1)}$. This figure corresponds to the value $M_1 = 11$.

More precisely, we set

$$\tau_1(x) = \begin{cases} \frac{M_1-3}{2}, & x \in I_1, \\ -1, & x \in I_{k_1}, k_1 \in \{2, \dots, (M_1-1)/2\}, \\ 0, & x \in I_{(M_1+1)/2}, \\ 1, & x \in I_{k_1}, k_1 \in \{(M_1+3)/2, \dots, M_1\}, \\ -\frac{M_1-3}{2}, & x \in I_{M_1}. \end{cases} \tag{28}$$

Then $T_{\alpha_1}^{(\tau_1)}$ induces a permutation of the intervals $(I_{k_1})_{k_1=1}^{M_1}$ and a short calculation shows that

$$\varphi^1(x) + \psi^1(T_{\alpha_1}^{(\tau_1)}(x)) = \begin{cases} 2, & x \in I_{k_1}, k_1 \in \{2, \dots, (M_1-1)/2, \\ & \qquad (M_1+3)/2, \dots, M_1 - 1\}, \\ -\frac{M_1-5}{2}, & x \in I_{k_1}, k_1 = 1, M_1, \\ 1, & x \in I_{(M_1+1)/2}. \end{cases} \tag{29}$$

Next figure is a representation of this "quasi-cost" at level $n = 1$, with the same value $M_1 = 11$ as in Figure 1.

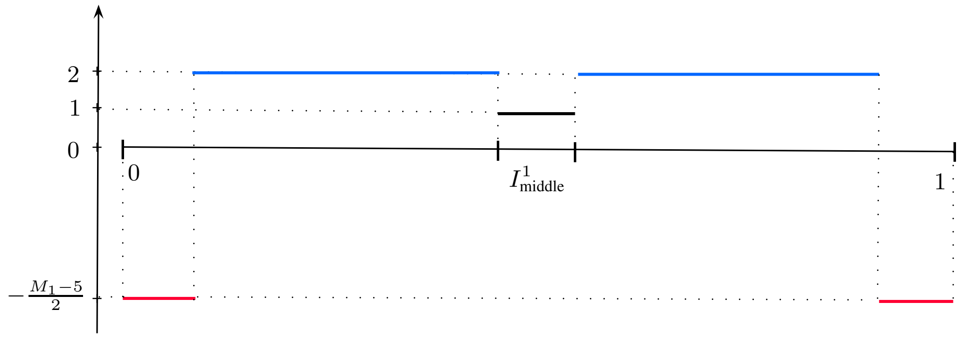

Fig. 2. Representation of $\varphi^1 + \psi^1 \circ T_{\alpha_1}^{(\tau_1)}$.

Assessment of Step $n = 1$. Let us resume what we have achieved in the first induction step. For later use we formulate things only in terms of $\varphi^1(\cdot)$ rather than $\psi^1(\cdot) = 1 - \varphi^1(\cdot)$. For the set $J_1^g = \{2, \dots, \frac{M-1}{2}\} \cup \{\frac{M+3}{2}, \dots, M_1 - 1\}$ of *"good*[2] indices" we have

$$\varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) = 1, \qquad x \in I_{k_1}, k_1 \in J_1^g, \tag{30}$$

while for the set $J_1^s = \{1, M_1\}$ of *"singular indices"* we have

$$\varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) = -\frac{M_1 - 3}{2}, \qquad x \in I_{k_1},\ k_1 \in J^s, \tag{31}$$

so that

$$\sum_{k_1 \in J_1^s} \int_{I_{k_1}} [\varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_1)}(x))]\, dx = -\frac{M_1-3}{2}\ \frac{2}{M_1} = -1 + \frac{3}{M_1}.$$

For the middle interval $I^1_{\text{middle}} = I_{(M_1+1)/2}$ we have $\varphi^1(x) - \varphi^1(T_{\alpha_1}^{\tau_1}(x)) = 0$.
We also note for later use that, for $x \in [0, 1)$, the orbit $(T_{\alpha_1}^i(x))_{i=1}^{\tau_1(x)}$ never visits I^1_{middle}. Here we mean that i runs through $\{\tau_1(x), \tau_1(x)+1, \dots, -1\}$ when $\tau_1(x) < 0$ and runs through the empty set when $\tau_1(x) = 0$.

Step n=2: We now pass from $\alpha_1 = \frac{1}{M_1}$ to $\alpha_2 = \frac{1}{M_1} + \frac{1}{M_2}$, where $M_2 = M_1 m_2 = m_1 m_2$ and where m_2, to be specified below, satisfies the relations of Lemma 3.2 and is large compared to M_1. For $1 \le k_1 \le M_1$ and $1 \le k_2 \le m_2$ we denote by I_{k_1,k_2} the interval

$$I_{k_1,k_2} = \left[\tfrac{k_1-1}{M_1} + \tfrac{k_2-1}{M_2}, \tfrac{k_1-1}{M_1} + \tfrac{k_2}{M_2}\right).$$

[2] We use the term "good" rather than "regular" as the abbreviation r is already taken by the word "right".

Similarly as above we will also use the notations $L^2 = [0, \frac{1}{2} - \frac{1}{2M_2})$, $R^2 = [\frac{1}{2} + \frac{1}{2M_2}, 1)$, and $I^2_{\text{middle}} = I_{(M_1+1)/2,(m_2+1)/2} = [\frac{1}{2} - \frac{1}{2M_2}, \frac{1}{2} + \frac{1}{2M_2})$.
We now define functions φ^2, ψ^2 such that $\varphi^2(x) + \psi^2(x) \equiv 1$ and

$$\varphi^2(x) + \psi^2(T_{\alpha_2}(x)) = \begin{cases} 0, & x \in L^2, \\ 1, & x \in I^2_{\text{middle}}, \\ 2, & x \in R^2. \end{cases}$$

This is achieved if we set, e.g., $\varphi^2 \equiv 0$ on $I_{1,1}$, and

$$\begin{aligned} \varphi^2(T_{\alpha_2}(x)) &= \varphi^2(x) + \begin{cases} 1 & x \in L^2, \\ 0 & x \in I^2_{\text{middle}}, \\ -1 & x \in R^2, \end{cases} \\ \psi^2(x) &= 1 - \varphi^2(x). \end{aligned} \tag{32}$$

Yet another way to express this is to say that for $j \in \{0, \dots, M_2 - 1\}$ we have

$$\begin{aligned} \varphi^2(T^j_{\alpha_2}(x)) = \#\{i \in \{0, \dots, j-1\} : T^i_{\alpha_2}(x) \in L^2\} \\ - \#\{i \in \{0, \dots, j-1\} : T^i_{\alpha_2}(x) \in R^2\} \end{aligned}, \quad x \in I_{1,1}, \tag{33}$$

in analogy to (14).

While the function $\varphi^1(x)$ in the first induction step was increasing from I_1 to $I_{(M_1+1)/2}$ and then decreasing from $I_{(M_1+3)/2}$ to I_{M_1}, the function $\varphi^2(x)$ displays a similar feature on each of the intervals I_{k_1}: roughly speaking, i.e. up to terms controlled by M_1, it increases on the left half of each such interval and then decreases again on the right half. The next lemma makes this fact precise. We keep in mind, of course, that m_2 will be much bigger than M_1.

Lemma 3.3 (Oscillations of φ^2)**.** *The function φ^2 defined in* (32) *has the following properties.*

(i) $|\varphi^2(x) - \varphi^2(x \oplus \frac{1}{M_2})| \leq 4M_1^2, \quad x \in [0, 1).$

(ii) For each $1 \leq k_1', k_1'' \leq M_1$ *we have*

$$\varphi^2{}_{|I_{k_1',(m_2+1)/2}} - \varphi^2{}_{|I_{k_1'',1}} \geq \tfrac{m_2}{2M_1} - 10M_1^3.$$

Proof. Let us begin with the proof of (i).
• *Proof of (i).* While $T^{M_1}_{\alpha_1} = id$ holds true, we have that $T^{M_1}_{\alpha_2}$ is only close to the identity map. In fact, as $T_{\alpha_2}(x) = x \oplus \frac{m_2+1}{M_2}$, we have

$$T^{M_1}_{\alpha_2}(x) = x \oplus \tfrac{M_1}{M_2}. \tag{34}$$

Somewhat less obvious is the fact that $T^{m_2-2}_{\alpha_2}$ also is close to the identity map. In fact

$$T^{m_2-2}_{\alpha_2}(x) = x \ominus \tfrac{2}{M_2}. \tag{35}$$

Indeed, by (25) applied to $i = 1$, there is $c \in \mathbb{N}$ such that $m_2 = cM_1 + 1$. Hence

$$\begin{aligned} T^{m_2-2}_{\alpha_2}(x) &= x \oplus (m_2 - 2)\frac{m_2+1}{M_2} \\ &= x \oplus (cM_1 - 1)\frac{m_2+1}{M_2} \\ &= x \oplus \frac{cM_2 - m_2 + (m_2 - 2)}{M_2} = x \ominus \frac{2}{M_2}. \end{aligned}$$

Here is one more remarkable feature of the map $T_{\alpha_2}^{m_2-2}$.

Claim: For $x \in [0,1)$ the orbit $(T_\alpha^i(x))_{i=1}^{m_2-2}$ visits the intervals $L^2 = [0, \frac{1}{2} - \frac{1}{2M_2})$ and $R^2 = [\frac{1}{2} + \frac{1}{2M_2}, 1)$ approximately equally often. More precisely, the difference of the visits of these two intervals is bounded in absolute value by $4M_1$.

Indeed, by Lemma 3.2, the orbit $(T_{\alpha_2}^i(x))_{i=1}^{M_2}$ visits each of the intervals I_{k_1,k_2} exactly one time so that it visits L^2 and R^2 equally often, namely $\frac{M_2-1}{2}$ times. The M_1 many disjoint subsets $\left(T_{\alpha_2}^{j(m_2-2)}\left(T_{\alpha_2}^i(x)\right)_{i=1}^{m_2-2}\right)_{j=1}^{M_1}$ of this orbit are obtained by shifting them successively by $2/M_2$ to the left (35). As the difference $(T_{\alpha_2}^i(x))_{i=1}^{M_2} \setminus \left(T_{\alpha_2}^{j(m_2-2)}\left(T_{\alpha_2}^i(x)\right)_{i=1}^{m_2-2}\right)_{j=1}^{M_1}$ consists only of $2M_1$ many points we have that the difference of the visits of $\left(T_{\alpha_2}^{j(m_2-2)}\left(T_{\alpha_2}^i(x)\right)_{i=1}^{m_2-2}\right)_{j=1}^{M_1}$ to L^2 and R^2 is bounded by $4M_1$. This implies that the difference of the visits of $(T_{\alpha_2}^i(x))_{i=1}^{m_2-2}$ to L^2 and R^2 can be estimated by $4M_1$ too: indeed, if this orbit visits $4M_1+k$ many times L^2 more often then R^2 (or vice versa) for some $k \geq 0$, then $(T_{\alpha_2}^{m_2-2}(T_{\alpha_2}^i(x)))_{i=1}^{m_2-2}$ visits L^2 at least $4M_1+k-4$ many times more often than R^2 etc. and finally $(T_{\alpha_2}^{M_1(m_2-2)}(T_{\alpha_2}^i(x)))_{i=1}^{m_2-2}$ visits L^2 at least k many times more often than R^2 which yields a contradiction. Hence we have proved the claim.

To prove assertion (i) note that by (34) and (35)

$$T_{\alpha_2}^{\frac{M_1-1}{2}(m_2-2)} \circ T_{\alpha_2}^{M_1}(x) = x \oplus \tfrac{1}{M_2} \tag{36}$$

We deduce from the claim that the difference of the visits of the orbit $(T_{\alpha_2}^i)_{i=0}^{\frac{M_1-1}{2}(m_2-2)+M_1}$ to L^2 and R^2 is bounded in absolute value by $\frac{M_1-1}{2}(4M_1) + M_1$ which proves (i).

• *Proof of (ii).* As regards (ii) suppose first $k_1' = k_1'' =: k_1$. Note that, for $x \in I_{k_1}^{\text{left}} := [\frac{k_1-1}{M_1}, \frac{k_1-1}{M_1} + \frac{1}{2M_1} - \frac{2M_1+1}{2M_2})$, we have that the orbit $(T_{\alpha_2}^i(x))_{x=0}^{M_1-1}$ visits L^2 one time more often than R^2, namely $\frac{M_1+1}{2}$ versus $\frac{M_1-1}{2}$ times. If we start with $x \in I_{k_1,1}$ then, for $1 \leq j < \frac{m_2}{2M_1} - 1$ we have that $T_{\alpha_2}^{jM_1}(x) \in I_{k_1}^{\text{left}}$. Hence, for the orbit $(T_{\alpha_2}^i)_{i=0}^{(\lfloor \frac{m_2}{2M_1} \rfloor - 1)M_1}$, the difference of the visits to the interval L^2 and R^2 equals $\lfloor \frac{m_2}{2M_1} \rfloor - 1$, the integer part of $\frac{m_2}{2M_1} - 1$. Combining this estimate with the estimate (i) as well as the fact that the distance between $x \oplus \left(\lfloor \frac{m_2}{2M_1} \rfloor - 1\right) \frac{M_1}{M_2}$ and $x \oplus \frac{m_2-1}{2M_2}$ is bounded by $\frac{2M_1-1}{M_2}$, we obtain, for $x \in I_{k_1,1}$ and $y \in I_{k_1, \frac{m_2+1}{2}}$, that

$$\begin{aligned}
\varphi^2(y) - \varphi^2(x) &\geq \varphi^2(T_{\alpha_2}^{(\lfloor \frac{m_2}{2M_1} \rfloor - 1)M_1}(x)) - \varphi^2(x) - \left|\varphi^2(y) - \varphi^2(T_{\alpha_2}^{(\lfloor \frac{m_2}{2M_1} \rfloor - 1)M_1}(x))\right| \\
&\geq (\lfloor \tfrac{m_2}{2M_1} \rfloor - 1) - (2M_1 - 1)(4M_1^2) \\
&\geq \tfrac{m_2}{2M_1} - 8M_1^3.
\end{aligned}$$

Passing to the general case $1 \leq k_1', k_1'' \leq M_1$ observe that $T_{\alpha_2}^{k_1''-k_1'}$ maps $I_{k_1', \frac{m_2+1}{2}}$ to $I_{k_1'', \frac{m_2+1}{2}+k_1''-k_1'}$. Using again (i) we obtain estimate (ii). □

We now are ready to do the inductive construction for $n = 2$. For m_2 satisfying the conditions of Lemma 3.1 and to be specified below, we shall define $\tau_2 : [0,1) \to \{-\frac{M_2-1}{2}, \ldots, 0, \ldots, \frac{M_2-1}{2}\}$, where $M_2 = m_2 m_1$, such that the map

$$T_{\alpha_2}^{(\tau_2)} : \left\{ \begin{array}{rcl} [0,1) & \to & [0,1) \\ x & \mapsto & T_{\alpha_2}^{(\tau_2)}(x) := T_{\alpha_2}^{\tau_2(x)}(x) \end{array} \right.$$

has the following properties.

(i) The measure-preserving bijection $T_{\alpha_2}^{(\tau_2)} : [0,1) \to [0,1)$ maps each interval I_{k_1} onto $T_{\alpha_1}^{(\tau_1)}(I_{k_1})$. It induces a permutation of the intervals I_{k_1,k_2}, where $1 \le k_1 \le M_1, 1 \le k_2 \le m_2$.

(ii) When $\tau_2(x) > 0$, we have

$$T_{\alpha_2}^{i}(x) \notin I^2_{\text{middle}}, \qquad i = 0, \dots, \tau_2(x), \tag{37}$$

and, when $\tau_2(x) < 0$, we have

$$T_{\alpha_2}^{i}(x) \notin I^2_{\text{middle}}, \qquad i = \tau_2(x), \dots, 0. \tag{38}$$

(iii) On the *"good"* intervals I_{k_1}, where $k_1 \in J_1^g = \{2, \dots, \frac{M_1-1}{2}\} \cup \{\frac{M_1+3}{2}, \dots, M_1 - 1\}$, for which we have, by (30),

$$\varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) = 1,$$

the function τ_2 will satisfy the estimates

$$\mu[I_{k_1} \cap \{\tau_2 \neq \tau_1\}] \le \tfrac{M_1}{m_2} \mu[I_{k_1}], \tag{39}$$

and

$$\sum_{k_1 \in J_1^g} \int_{I_{k_1}} |1 - \varphi^2(x) + \varphi^2(T_{\alpha_2}^{(\tau_2)}(x))| \, dx < \frac{4M_1^2}{m_2}. \tag{40}$$

(iv) On the *"singular"* intervals I_{k_1}, where $k_1 \in J_1^s = \{1, M_1\}$, for which we have , by (31),

$$\varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) = -\frac{M_1 - 3}{2},$$

we split $\{1, \dots, m_2\}$ into a set $J^{k_1,g}$ of *"good"* indices, and a set $J^{k_1,s}$ of *"singular"* indices, such that

$$\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x)) = 0, \quad \text{for } x \in I_{k_1,k_2}, k_2 \in J^{k_1,g},$$

while

$$\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x)) < -\tfrac{m_2}{2M_1} + 20M_1^3 \quad \text{for } x \in I_{k_1,k_2}, k_2 \in J^{k_1,s}$$

where $J^{k_1,s}$ consists of $M_1(M_1-3)$ many elements of $\{1, \dots, m_2\}$.
Hence we have a total "singular mass" of

$$\sum_{k_1 \in J_1^s} \sum_{k_2 \in J^{k_1,s}} \int_{I_{k_1,k_2}} [\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x))] \, dx < -1 + \tfrac{3}{M_1} + \tfrac{c(M_1)}{m_2}, \tag{41}$$

where $c(M_1)$ is a constant depending only on M_1.

(v) On the middle interval $I^1_{\text{middle}} = I_{\frac{M_1+1}{2}}$ we simply let $\tau_2 = \tau_1 = 0$.

Let us illustrate graphically an interesting property of this construction, namely the shape of the quasi-cost function $\varphi^2 + \psi^2 \circ T_{\alpha_2}^{(\tau_2)}$.

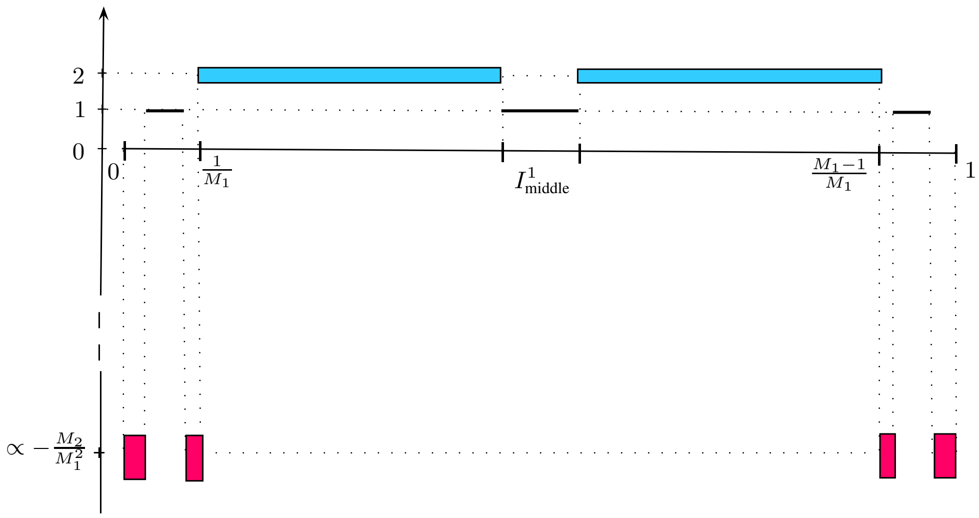

Fig. 3. Shape of the quasi-cost $\varphi^2 + \psi^2 \circ T_{\alpha_2}^{(\tau_2)}$.

The strips in this graphic representation symbolize the oscillations of the function $\varphi^2 + \psi^2 \circ T_{\alpha_2}^{(\tau_2)}$. *On the "singular" set, it achieves values of order* $-M_2/M_1^2$.

It will sometimes be more convenient to specify to which interval I_{l_1,l_2} the interval I_{k_1,k_2} is mapped under $T_{\alpha_2}^{(\tau_2)}$, instead of spelling out the value of τ_2 on the interval I_{k_1,k_2}. Note that by Lemma 3.2, for each map associating to (k_1,k_2) a pair (l_1,l_2), there corresponds precisely one value $\tau_2|_{I_{k_1,k_2}} : I_{k_1,k_2} \to \{-M_2+1,\dots,0,\dots,M_2-1\}$ such that (37) (resp. (38)) is satisfied and $T_{\alpha_2}^{(\tau_2)}(I_{k_1,k_2}) = I_{l_1,l_2}$.

Let us start with a "good" interval I_{k_1}, with $k_1 \in J_1^g$ as in (iii) above, say $k_1 \in \{2,\dots,\frac{M_1-1}{2}\}$, for which we have $\tau_1(x) = -1$. Then the intervals $I_{k_1,2},\dots,I_{k_1,m_2}$ are mapped under $T_{\alpha_2}^{\tau_1(x)}(x) = T_{\alpha_2}^{-1}(x)$ onto the intervals $I_{k_1-1,1},\dots,I_{k_1-1,m_2-1}$. Defining $\tau_2(x) = \tau_1(x)$ on these intervals we get for $x \in I_{k_1,k_2}$, where $2 \le k_1 \le \frac{M-1}{2}, 2 \le k_2 \le m_2$,

$$1 = \varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) = \varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x)). \tag{42}$$

We still have to define the value of $\tau_2(x)$, for $x \in I_{k_1,1}$. The map $T_{\alpha_2}^{(\tau_2)}$ has to map $I_{k_1,1}$ to the remaining gap I_{k_1-1,m_2}, which happens to be its left neighbour. We do not explicitly calculate the unique number $\tau_2|_{I_{k_1,1}} \in \{-M_2+1,\dots,M_2-1\}$, satisfying (37) (resp. (38)), which does the job, but only use the conclusion of Lemma 3.3 to find that, for $x \in I_{k_1,1}$ such that $T_{\alpha_2}^{(\tau_2)}(x) \in I_{k_1-1,m_2}$,

$$|1 - [\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x))]| \le 4M_1^2 + 1. \tag{43}$$

This takes care of the "good" intervals I_{k_1}, where $k_1 \in \{2,\dots,\frac{M_1-1}{2}\}$.

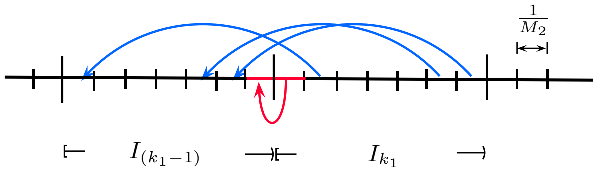

Fig. 4-a. $k_1 \in J_1^g$ on the left side.[3]

For the "good" intervals I_{k_1}, where $k_1 \in \{\frac{M_1+3}{2}, \ldots, M_1 - 1\}$ we have $\tau_1(x) = 1$ so that $T_{\alpha_2}^{(\tau_1)}$ maps the intervals $I_{k_1,1}, \ldots, I_{k_1,m_2-1}$ to $I_{k_1+1,2}, \ldots, I_{k_1+1,m_2}$. Again we define $\tau_2(x) = \tau_1(x) = 1$, for x in these intervals so that we obtain the identity (42), for $\frac{M_1+3}{2} \leq k_1 \leq M_1 - 1$ and $1 \leq k_2 \leq m_2 - 1$. Finally, $T_{\alpha_2}^{(\tau_2)}$ has to map I_{k_1,m_2} to the interval $I_{k_1+1,1}$ so that again we derive an estimate as in (43).

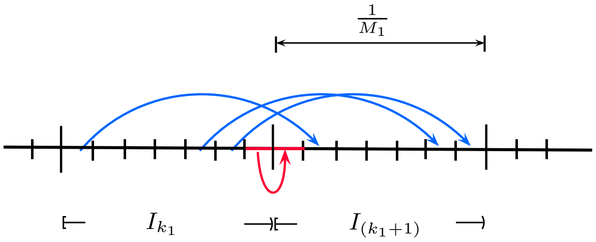

Fig. 4-b. $k_1 \in J_1^g$ on the right side.

This finishes item (iii) i.e. the definition of τ_2 on the "good" intervals I_{k_1}. Noting that on this set we have $\tau_1 \neq \tau_2$ only on $M_1 - 3$ many intervals of length $\frac{1}{M_2}$ we obtain the estimate (40).

To show (iv) let us first consider the "singular" interval I_1, on which we have $\tau_1(x) = \frac{M_1-3}{2}$ and $\varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) = \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) - \varphi^1(x) = \frac{M_1-3}{2}$. For the subintervals I_{1,k_2} of I_1, define the set of good indices as $J^{1,g} = J^{1,g,l} \cup J^{1,g,r}$ where

$$J^{1,g,l} = \{\tfrac{(M_1-3)(M_1-1)}{2} + 1, \ldots, \tfrac{m_2-1}{2}\}, \quad J^{1,g,r} = \{\tfrac{m_2+1}{2}, \ldots, m_2 - \tfrac{(M_1-3)(M_1+1)}{2}\}.$$

Let us start by considering $k_2 \in J^{1,g,r}$. We define

$$\tau_2(x) = \tau_1(x) + \frac{M_1-3}{2}M_1 = \frac{(M_1-3)(M_1+1)}{2}, \quad x \in I_{1,k_2}, k_2 \in J^{1,g,r}.$$

First note that $T_{\alpha_2}^{(\tau_2)}$ then maps the intervals I_{1,k_2}, for $k_2 \in J^{1,g,r}$, to the intervals

$$I_{\frac{M_1-1}{2}, \frac{m_2+1}{2} + \frac{(M_1-3)(M_1+1)}{2}}, \ \ldots \ , I_{\frac{M_1-1}{2}, m_2}.$$

Observe that, for x as above, the orbit $(T_{\alpha_2}^i(x))_{i=0}^{\tau_2(x)-1}$ always lies in the right halfs of the respective intervals I_{k_1}.

Let us count how often the orbit $(T_{\alpha_2}^i(x))_{i=0}^{\tau_2(x)-1}$ visits L^2 and R^2 respectively, for $x \in I_{1,k_2}$ and $k_2 \in J^{1,g,r}$. The first $\tau_1(x) = \frac{M_1-3}{2}$ elements of this orbit are all in L^2 which yields, similarly as in the induction step $n = 1$,

$$\varphi^2(T_{\alpha_2}^{(\tau_1)}(x)) - \varphi^2(x) = \varphi^1(T_{\alpha_1}^{(\tau_1)}(x)) - \varphi^1(x) = \frac{M_1-3}{2}.$$

[3]Figure 3 is built with the small value $m_2 = 7$ for the sake of clarity of the drawing. But this value is not feasible since with the lowest $m_1 = 5$, (25) implies that m_2 is at least equal to 11; other requirements of the construction imply that it has to be even larger.

But the next M_1 many elements of this orbit, namely

$$(T^i_{\alpha_2}(x))_{i=\tau_1(x)}^{\tau_1(x)+M_1-1}$$

visit R^2 one time more often than L^2 as the unique element of this orbit which lies in I^1_{middle} belongs to the right half of I^1_{middle}.

This phenomenon repeats on the orbit $(T^i_{\alpha_2}(x))_{i=0}^{\tau_1(x)+\frac{M_1-3}{2}M_1-1}$ for $\frac{M_1-3}{2}$ many times so that

$$\begin{aligned}\varphi^2(x)-\varphi^2(T^{(\tau_2)}_{\alpha_2}(x)) &= \varphi^2(x)-\varphi^2(T^{(\tau_1)}_{\alpha_2}(x))) + \varphi^2(T^{(\tau_1)}_{\alpha_2}(x)) - \varphi^2(T^{(\tau_2)}_{\alpha_2}(x) \\ &= -\frac{M_1-3}{2}+\frac{M_1-3}{2} \\ &= 0, \quad \text{for } x\in I_{1,k_2} \text{ and } k_2\in J^{1,g,r}.\end{aligned} \tag{44}$$

This takes care of I_{1,k_2} with $k_2 \in J^{1,g,r}$.

For $x\in I_{1,k_2}$ with $k_2\in J^{1,g,l}$, the left half of the "good" intervals, we define symmetrically

$$\tau_2(x)=\tau_1(x)-\tfrac{M_1-3}{2}M_1=-\tfrac{(M_1-3)(M_1-1)}{2}.$$

A similar analysis as above shows that $T^{(\tau_2)}_{\alpha_2}$ maps the intervals I_{1,k_2}, where $k_2\in J^{1,g,l}$, to the intervals $I_{\frac{M_1-1}{2},1},\dots,I_{\frac{M_1-1}{2},\frac{m_2-1}{2}-\frac{(M_1-3)(M_1-1)}{2}}$. Hence by a symmetric reasoning we again obtain equality (44) for x in the intervals I_{1,k_2}, and for $k_2\in J^{1,g,r}$ too.

Now we have to deal with the "singular" subintervals I_{1,k_2}, where $k_2\in J^{1,s}$, and the singular indices are given by

$$\begin{aligned} J^{1,s} &= \{1,\dots,m_2\}\setminus J^{1,g} \\ &= \{1,\dots,\tfrac{(M_1-3)(M_1-1)}{2}\}\cup\{m_2-\tfrac{(M_1-3)(M_1+1)}{2}+1,\dots,m_2\},\end{aligned}$$

which consists of $M_1(M_1-3)$ many indices.

The map $T^{(\tau_2)}_{\alpha_2}$ has to map these intervals I_{1,k_2}, where $k_2\in J^{1,s}$, to the "remaining gaps" $I_{\frac{M_1-1}{2},l_2}$ in the interval $I_{\frac{M_1-1}{2}}$, where $l_2\in\{\frac{m_2+1}{2}-\frac{(M_1-3)(M_1-1)}{2},\dots,\frac{m_2+1}{2}+\frac{(M_1-3)(M_1+1)}{2}-1\}$. Note that the corresponding intervals $I_{\frac{M_1-1}{2},l_2}$ are – roughly speaking – in the middle of the interval $I_{\frac{M_1-1}{2}}$, while the intervals I_{1,k_2}, with $k_2\in J^{1,s}$, are at the boundary of I_1.

To define τ_2 on I_{1,k_2}, for $k_2\in J^{1,s}$, choose any function τ_2 taking values in $\{-M_2+1,\dots,M_2-1\}$, satisfying (37) (resp. (38)) as above, which induces a bijection between the intervals $(I_{1,k_2})_{k_2\in J^{1,s}}$ and the intervals $I_{\frac{M_1-1}{2},l_2}$ considered above.

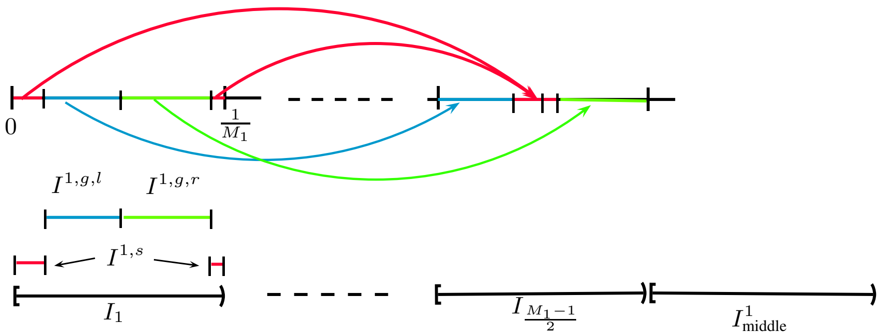

Fig. 5. τ_2 for the "singular" indices on the left side.

In this drawing, the interval $I^{1,g,l}$ is the union of the intervals I_{1,k_2} with $k_2 \in J^{1,g,l}$. A similar convention holds for $I^{1,g,r}$ and $I^{1,s}$ (which is not an interval anymore).

For each such τ_2 we obtain, for $x \in I_{1,k_2}, k_2 \in J^{1,s}$, from Lemma 3.3

$$\begin{aligned} \varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x)) &\leq -\frac{m_2}{2M_1} + 10M_1^3 + 2\frac{(M_1-3)(M_1-1)}{2}4M_1^2 \\ &\leq -\frac{m_2}{2M_1} + 20M_1^4. \end{aligned} \tag{45}$$

Indeed, the leading term $\frac{-m_2}{2M_1}$ and the first error term $10M_1^3$ in the first line above come from Lemma 3.3-(ii) when comparing the difference of the value of φ^2 on the interval $I_{1,1}$ to that of $I_{\frac{M_1-1}{2},\frac{m_2+1}{2}}$. For the difference of the value of φ^2 on I_{1,k_2} and $I_{\frac{M_1-1}{2},l_2}$, for arbitrary $k_2 \in J^{1,s}$ and $l_2 \in \{\frac{m_2+1}{2} - \frac{(M_1-3)(M_1-1)}{2}, \ldots, \frac{m_2+1}{2} + \frac{(M_1-3)(M_1-1)}{2}\}$ we apply for both cases at most $\frac{(M_1-3)(M_1+1)}{2}$ times estimate (i) of Lemma 3.3 which gives (45).
In particular, for $m_2 > 40M_1^5$, which of course we shall assume, we have that

$$\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x)) \leq 0, \qquad \text{for } x \in I_{1,k_2}, k_2 \in J^{1,s}.$$

There are $M_1(M_1-3) = M_1^2 - 3M_1$ many intervals I_{1,k_2} with $k_2 \in J^{1,s}$ each of length $1/M_2$. Hence we may estimate the "singular mass" on the interval I_1 by

$$\begin{aligned} \sum_{k_2 \in J^{1,s}} \int_{I_{1,k_2}} [\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_2)}(x))]\,dx &\leq \big(-\frac{m_2}{2M_1} + 20M_1^4\big)(M_1^2 - 3M_1)\frac{1}{M_2} \\ &\leq -\frac{1}{2} + \frac{3}{2M_1} + \frac{c(M_1)}{2m_2}. \end{aligned} \tag{46}$$

where $c(M_1)$ is a constant depending on M_1 only.[4]

We still have another "singular" interval at the present induction step $n = 2$, namely I_{M_1}. The analysis for this case is symmetric to the analysis of I_1 and – after properly defining τ_2 on this interval I_{M_1} – we arrive at the same estimate (46). In total, the thus obtain (41) by doubling the right hand side of (46), showing that the "singular mass" essentially equals -1.

Finally define the sets J_2^g (resp. J_2^s) of "good" (resp. "singular") indices at level 2 as

$$\begin{aligned} J_2^g &= \{(k_1,k_2) : (k_1 \in J_1^g \text{ and } 1 \leq k_2 \leq m_2), \text{ or } (k_1 \in J_1^s \text{ and } k_2 \in J^{k_1,g})\}, \\ J_2^s &= \{(k_1,k_2) : k_1 \in J_1^s \text{ and } k_2 \in J^{k_1,2}\}. \end{aligned}$$

This finishes the inductive step for $n = 2$.

General inductive step. Suppose that the prime numbers $m_1, \ldots, m_{n-1}$ have been defined. We use the notation $\alpha_{n-1} = \frac{1}{M_1} + \cdots + \frac{1}{M_{n-1}}$, where $M_{n-1} = m_1 \cdot m_2 \cdot \ldots \cdot m_{n-1}$. For a prime m_n satisfying the condition of Lemma 3.2, and to be specified below, let $M_n = m_1 \cdot \ldots \cdot m_n$ and

$$L^n = \left[0, \frac{1}{2} - \frac{1}{2M_n}\right), \; R^n = \left[\frac{1}{2} + \frac{1}{2M_n}, 1\right), \; I^n_{\text{middle}} = \left[\frac{1}{2} - \frac{1}{2M_n}, \frac{1}{2} + \frac{1}{2M_n}\right).$$

[4] We shall find it convenient in the sequel to write $c(M_1, M_2, \ldots, M_i)$ for constants depending only on the choice of the numbers $M_1, M_2, \ldots, M_i$. The concrete numerical value of this expression may change, i.e. become bigger, from one line of reasoning to the next one, but at every stage it will be clear that an explicit bound for the respective meaning of the constant $c(M_1, M_2, \ldots, M_i)$ could be given, at least in principle. In fact, we shall always have that the constants $c(M_1, M_2, \ldots, M_i)$ used in the sequel are dominated by a polynomial in the variables $M_1, M_2, \ldots, M_i$.

For $1 \le k_1 \le m_1, \dots, 1 \le k_n \le m_n$, let

$$I_{k_1,\dots,k_n} = [\tfrac{k_1-1}{M_1} + \tfrac{k_2-1}{M_2} + \cdots + \tfrac{k_n-1}{M_n}\ ,\tfrac{k_1-1}{M_1} + \tfrac{k_2-1}{M_2} + \cdots + \tfrac{k_n}{M_n}).$$

For $x \in I_{1,\dots,1}$ and $j \in \{0,\dots,M_n\}$ we define, similarly as in (33), $\varphi^n(x) = 0$ and

$$\begin{aligned} \varphi^n(T^j_{\alpha_n}(x)) = \quad & \#\{i \in \{0,\dots,j-1\} : T^i_{\alpha_2}(x) \in L^n\} \\ & -\#\{i \in \{0,\dots,j-1\} : T^i_{\alpha_2}(x) \in R^n\}, \end{aligned} \tag{47}$$

where $\alpha_n = \alpha_{n-1} + \frac{1}{M_n}$ and $M_n = M_{n-1}m_n$. We also let $\psi^n(x) = 1 - \varphi^n(x)$, for $x \in [0,1)$.

Lemma 3.4 (Oscillations of φ^n)**.** *For given $M_1,\dots,M_{n-1}$ there is a constant $c(M_1,\dots,M_{n-1})$ depending only on $M_1,\dots,M_{n-1}$, such that for all m_n as above we have*

(i) $|\varphi^n(x) - \varphi^n(x \oplus \frac{1}{M_n})| \le c(M_1,\dots,M_{n-1})$,

(ii) for each $1 \le k_1', k_1'' \le M_1, \dots, 1 \le k_{n-1}', k_{n-1}'' \le m_{n-1}$,

$$\varphi^n{}_{|I_{k_1',\dots,k_{n-1}',(m_n+1)/2}} - \varphi^n{}_{|I_{k_1'',\dots,k_{n-1}'',1}} \ge \tfrac{m_n}{2M_{n-1}} - c(M_1,\dots,M_{n-1}),$$

(iii) for each $1 \le k_1', k_1'' \le M_1, \dots, 1 \le k_{n-1}', k_{n-1}'' \le m_{n-1}$, and $1 \le k_n', k_n'' \le m_n$, with $\min\{k_n', m_n - k_n'\} < M_{n-1}$ and $\min\{k_n'', m_n - k_n''\} < M_{n-1}$ we have

$$\left| \varphi^n{}_{|I_{k_1',\dots,k_{n-1}',k_n'}} - \varphi^n{}_{|I_{k_1'',\dots,k_{n-1}'',k_n''}} \right| \le c(M_1,\dots,M_{n-1}). \tag{48}$$

Proof. We may and do assume that $m_n \ge 5M_{n-1}$.
• Proof of (i). We have $T_{\alpha_n}(x) = T_{\alpha_{n-1}}(T_{1/M_n}(x))$ so that

$$T^{M_{n-1}}_{\alpha_n}(x) = x \oplus \tfrac{M_{n-1}}{M_n} = x \oplus \tfrac{1}{m_n}, \tag{49}$$

in perfect analogy to (34). As regards the analogue to (35) things now are somewhat more complicated. First note that there is a unique number $1 \le q_{n-1} \le M_{n-1} - 1$ such that

$$T^{q_{n-1}}_{\alpha_{n-1}}(x) = x \ominus \frac{1}{M_{n-1}}, \qquad x \in [0,1). \tag{50}$$

Indeed, by Lemma 3.2, when q_{n-1} runs through $\{1,\dots,M_{n-1}-1\}$, the left hand side assumes the values $x \ominus \frac{l_{n-1}}{M_{n-1}}$, where l_{n-1} also runs through $\{1,\dots,M_{n-1}-1\}$.

Claim: Letting $r_n = \lfloor \frac{m_n}{M_{n-1}} \rfloor$, the integer part of $\frac{m_n}{M_{n-1}}$, and taking q_{n-1} as in (50)*, we have*

$$T^{r_nM_{n-1}+q_{n-1}}_{\alpha_n}(x) = x \oplus \frac{d_{n-1}}{M_n},$$

where $|d_{n-1}| < M_{n-1}$.

Indeed, write m_n as $m_n = r_nM_{n-1} + e_{n-1}$, for some $1 \le e_{n-1} \le M_{n-1}$ to obtain

$$\begin{aligned} T^{r_nM_{n-1}+q_{n-1}}_{\alpha_n}(x) &= (T^{M_{n-1}}_{\alpha_n})^{r_n} \circ T^{q_{n-1}}_{\alpha_{n-1}} \circ T^{q_{n-1}}_{\frac{1}{M_n}}(x) \\ &= x \oplus r_n\frac{M_{n-1}}{M_n} \ominus \frac{1}{M_{n-1}} \oplus \frac{q_{n-1}}{M_n} \\ &= x \oplus \frac{m_n}{M_n} \ominus \frac{e_{n-1}}{M_n} \ominus \frac{1}{M_{n-1}} \oplus \frac{q_{n-1}}{M_n} \\ &= x \oplus \frac{q_{n-1} - e_{n-1}}{M_n} =: x \oplus \frac{d_{n-1}}{M_n} \end{aligned}$$

which proves the claim.
Define $s_{n-1}^{(1)} = q_{n-1}$ if $d_{n-1} = q_{n-1} - e_{n-1} > 0$ and $s_{n-1}^{(1)} = q_{n-1} + M_{n-1}$ otherwise, to obtain by (49) and (50) that

$$T_{\alpha_n}^{r_n M_{n-1} + s_{n-1}^{(1)}}(x) = x \oplus \tfrac{l_{n-1}^{(1)}}{M_n},$$

for some $l_{n-1}^{(1)} \in \{1, \ldots, M_{n-1}\}$. We also deduce from (49) that $l_{n-1}^{(1)}$ must actually be in $\{1, \ldots, M_{n-1} - 1\}$.
Repeat the above argument to find $s_{n-1}^{(2)}$ with $-2M_{n-1} < s_{n-1}^{(2)} < 2M_{n-1}$ such that

$$T_{\alpha_n}^{2r_n M_{n-1} + s_{n-1}^{(2)}}(x) = x \oplus \tfrac{l_{n-1}^{(2)}}{M_n},$$

for some $l_{n-1}^{(2)} \in \{1, \ldots, M_{n-1} - 1\}$. Continuing in the same way, we find numbers $s_{n-1}^{(j)}$, for $j = 1, 2, \ldots, M_{n-1} - 1$ verifying $-jM_{n-1} < s_{n-1}^{(j)} < jM_{n-1}$ such that

$$T_{\alpha_n}^{jr_n M_{n-1} + s_{n-1}^{(j)}}(x) = x \oplus \tfrac{l_{n-1}^{(j)}}{M_n}, \tag{51}$$

for some $l_{n-1}^{(j)} \in \{1, \ldots, M_{n-1} - 1\}$. Note that, under the assumption $m_n \gg M_{n-1}$ so that $r_n \gg M_{n-1}$, the elements in (51) are all different. Therefore $(l_{n-1}^{(j)})_{j=1}^{M_{n-1}-1}$ runs through all elements of $\{1, \ldots, M_{n-1} - 1\}$ when j runs through $\{1, \ldots, M_{n-1} - 1\}$; in particular there must be some j_0 such that

$$T_{\alpha_n}^{j_0 r_n M_{n-1} + s_{n-1}^{(j_0)}}(x) = x \oplus \tfrac{1}{M_n},$$

in analogy to (36).

Now observe that there is a constant $c(M_1, \ldots, M_{n-1})$, depending only on $M_1, \ldots, M_{n-1}$, such that, for $x \in [0, 1)$, the difference of the number of visits of the orbit $(T_{\alpha_n}^i(x))_{i=0}^{r_n M_{n-1} + q_{n-1}}$ to L^n and R^n is bounded in absolute value by the constant $c(M_1, \ldots, M_{n-1})$. The argument is analogous to the corresponding one in the proof of the claim which is part of the proof of Lemma 3.3-(i), and therefore skipped.
The numbers j_0 as well as $s_{n-1}^{(j_0)}$ are bounded in absolute value by M_{n-1}^2 so that the difference of the visits of the orbits $(T_{\alpha_n}^i(x))_{i=0}^{j_0 r_n M_{n-1} + s_{n-1}^{(j_0)}}$ to L^n and R^n are bounded in absolute value by some constant $c(M_1, \ldots, M_{n-1})$. This finishes the proof of assertion (i).

• Proof of (ii). Suppose first, as in the proof of Lemma 3.3-(ii), that $(k'_1, \ldots, k'_{n-1}) = (k''_1, \ldots, k''_{n-1}) =: (k_1, \ldots, k_{n-1})$. For $x \in I_{k_1, \ldots, k_{n-1}, 1}$ we have that each of the orbits $(T_{\alpha_n}^{jM_{n-1}+i}(x))_{i=0}^{M_{n-1}-1}$, for $j = 0, \ldots, \lfloor \frac{m_n}{2M_{n-1}} \rfloor - 1$ visits L^n one time more often than R^n. Hence

$$\varphi^n(T_{\alpha_n}^{\lfloor \frac{m_n}{2M_{n-1}} \rfloor M_{n-1}}(x)) - \varphi^n(x) = \lfloor \tfrac{m_n}{2M_{n-1}} \rfloor > \tfrac{m_n}{2M_{n-1}} - 1.$$

Noting that

$$T_{\alpha_n}^{\lfloor \frac{m_n}{2M_{n-1}} \rfloor M_{n-1}}(x) = x \oplus \lfloor \tfrac{m_n}{2M_{n-1}} \rfloor \tfrac{M_{n-1}}{M_n}$$

and

$$\tfrac{m_n+1}{2M_n} - \lfloor \tfrac{m_n}{2M_{n-1}} \rfloor \tfrac{M_{n-1}}{M_n} \leq \tfrac{M_{n-1}}{M_n},$$

we obtain (ii) by using assertion (i), and possibly passing to a bigger constant $c(M_1, \ldots, M_{n-1})$. Finally the passage to general $(k'_1, \ldots, k'_{n-1})$ and $(k''_1, \ldots, k''_{n-1})$ is done again, similarly as in the proof of Lemma 3.3, by repeated application of (i) and by passing once more to a bigger constant $c(M_1, \ldots, M_{n-1})$.

• Proof of (iii). Fix $1 \le k_1', k_1'' \le M_1, \ldots, 1 \le k_{n-1}', k_{n-1}'' \le m_{n-1}$ and $1 \le k_n', k_n'' \le m_n$ as above. Suppose, e.g., $k_n' \le M_{n-1}$ and $m_n - k_n'' \le M_{n-1}$, the other three cases being similar. Denote by $(k_1''', \ldots, k_{n-1}''')$ the index so that $I_{k_1''',\ldots,k_{n-1}'''} = I_{k_1'',\ldots,k_{n-1}'',k_n''} \oplus \frac{1}{M_{n-1}}$, i.e. $I_{k_1''',\ldots,k_{n-1}'''}$ is the right neighbour of $I_{k_1'',\ldots,k_{n-1}''}$. Now find $0 \le q_{n-1} < M_{n-1}$ such that $T_{\alpha_{n-1}}^{q_{n-1}}$ maps $I_{k_1',\ldots,k_{n-1}'}$ onto $I_{k_1''',\ldots,k_{n-1}'''}$. Hence $T_{\alpha_n}^{q_{n-1}}$ maps $I_{k_1',\ldots,k_{n-1}',k_n'}$ onto $I_{k_1''',\ldots,k_{n-1}''',k_n'+q_{n-1}}$.
Finally note that the distance from the latter interval to $I_{k_1'',\ldots,k_{n-1}'',k_n''}$ is bounded by $(2M_{n-1} + M_{n-1})\frac{1}{M_n}$. Hence we obtain (48) by applying $2M_{n-1} + M_{n-1}$ times assertion (i) and using $0 \le q_{n-1} < M_{n-1}$. $\square$

After this preparation we are ready for the inductive step from $n-1$ to n. Suppose that the following inductive hypotheses are satisfied, for $1 \le l \le n-1$, functions $\tau_l : [0,1) \to \{-M_l+1, \ldots, M_l-1\}$ and index sets J_l^g, J_l^s contained in $\{(k_1, \ldots, k_l) : 1 \le k_1 \le m_1, \ldots, 1 \le k_l \le m_l\}$.

(i) The measure preserving bijection $T_{\alpha_{n-1}}^{(\tau_{n-1})} : [0,1) \to [0,1)$ maps the intervals $I_{k_1,\ldots,k_l}$, for $1 \le l < n-1$, and $1 \le k_1 \le m_1, \ldots, 1 \le k_l \le m_l$, onto the intervals $T_{\alpha_l}^{(\tau_l)}(I_{k_1,\ldots,k_l})$. It induces a permutation of the intervals $I_{k_1,\ldots,k_{n-1}}$, where $1 \le k_1 \le m_1, \ldots, 1 \le k_{n-1} \le m_{n-1}$.

(ii) When $\tau_{n-1}(x) > 0$, we have
$$T_{\alpha_{n-1}}^i(x) \notin I_{\text{middle}}^{n-1}, \qquad i = 0, \ldots, \tau_{n-1}(x), \tag{52}$$
and, when $\tau_{n-1}(x) < 0$, we have
$$T_{\alpha_{n-1}}^i(x) \notin I_{\text{middle}}^{n-1}, \qquad i = \tau_{n-1}(x), \ldots, 0. \tag{53}$$

(iii) There is a set of "good" indices $J_{n-1}^g \subseteq \{1 \le k_1 \le m_1, \ldots, 1 \le k_{n-1} \le m_{n-1}\}$. For $(k_1, \ldots, k_{n-2}) \in J_{n-2}^g$ we have that $(k_1, \ldots, k_{n-2}, k_{n-1}) \in J_{n-1}^g$ as well as
$$\mu[I_{k_1,\ldots,k_{n-2}} \cap \{\tau_{n-2} \ne \tau_{n-1}\}] \le \tfrac{M_{n-2}}{m_{n-1}} \mu[I_{k_1,\ldots,k_{n-2}}], \tag{54}$$
and
$$\begin{aligned} \sum_{(k_1,\ldots,k_{n-2}) \in J_{n-2}^g} \int_{I_{k_1,\ldots,k_{n-2}}} \Big| &[\varphi^{n-2}(x) - \varphi^{n-2}(T_{\alpha_{n-2}}^{(\tau_{n-2})}(x))] \\ &- [\varphi^{n-1}(x) - \varphi^{n-1}(T_{\alpha_{n-1}}^{(\tau_{n-1})}(x))]\Big|\, dx \\ \le \tfrac{c(M_1,\ldots,M_{n-2})}{m_{n-1}}. \end{aligned} \tag{55}$$

(iv) There is a set of "singular" indices $J_{n-1}^s \subseteq \{(k_1, \ldots, k_{n-1}) : 1 \le k_1 \le m_1, \ldots, 1 \le k_{n-1} \le m_{n-1}\}$, disjoint from J_{n-1}^g, such that J_{n-1}^s consists of less than $2M_{n-1}^2$ many elements and such that
$$\begin{aligned} \varphi^{n-1}(x) - \varphi^{n-1}(T_{\alpha_{n-1}}^{(\tau_{n-1})}(x)) \ \le\ 0, \quad &\text{for } x \in I_{k_1,\ldots,k_{n-1}} \\ &\text{and } (k_1, \ldots, k_{n-1}) \in J_{n-1}^s, \end{aligned} \tag{56}$$
and
$$\begin{aligned} \sum_{(k_1,\ldots,k_{n-1}) \in J_{n-1}^s} \int_{I_{k_1,\ldots,k_{n-1}}} &[\varphi^{n-1}(x) - \varphi^{n-1}(T_{\alpha_{n-1}}^{(\tau_{n-1})}(x))]\, dx \\ &\le -1 + \tfrac{3}{m_1} + \tfrac{c(M_1)}{m_2} + \cdots + \tfrac{c(M_1,\ldots,M_{n-2})}{m_{n-1}}, \end{aligned} \tag{57}$$
where $c(\cdot)$ are constants depending only on $(\cdot)$.

(v) On the middle interval $I^1_{\text{middle}} = I^1_{\frac{M_1+1}{2}}$ we have $\tau_1 = \tau_2 = \cdots = \tau_{n-1} = 0$ and I^1_{middle} together with the intervals $(I_{k_1,\ldots,k_{n-1}})_{(k_1,\ldots,k_{n-1})\in J^g_{n-1}\cup J^s_{n-1}}$ form a partition of $[0,1)$.

We have to define τ_n as well as J^g_n and J^s_n so that the above list is satisfied with $n-1$ replaced by n.

Let us illustrate graphically some features of this construction. Namely, the fractal structure of the singular set and the resulting quasi-cost.

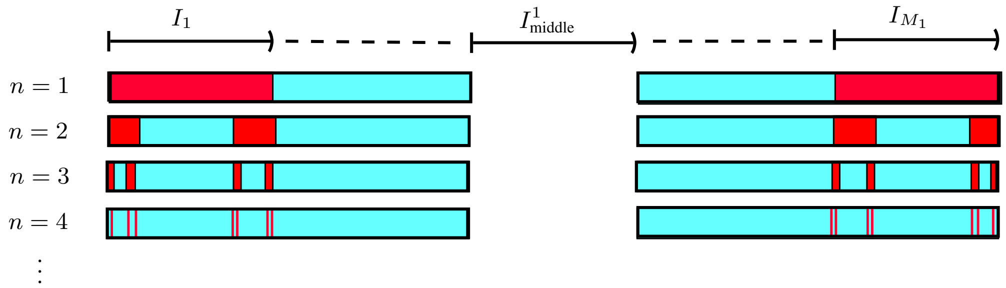

Fig. 6. The fractal structure of the "singular" set.

For the sake of simplicity of the drawing, the red area which represents the singular set is thicker than it should be. Note also that the effective singular set is not perfectly balanced.

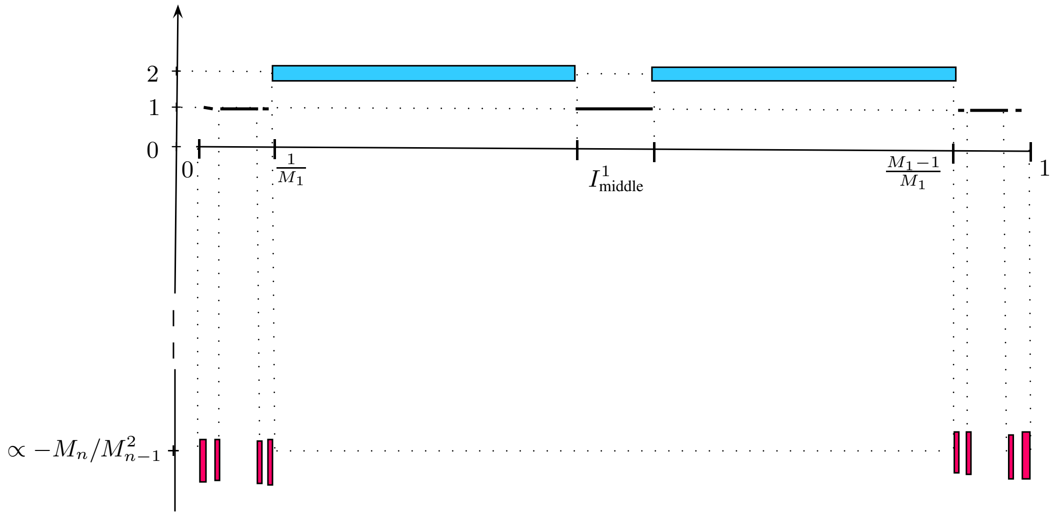

Fig. 7. Shape of the quasi-cost $\varphi^n + \psi^n \circ T^{(\tau_n)}_{\alpha_n}$.

The strips on this graphic representation symbolize the oscillations of the function $\varphi^n + \psi^n \circ T^{(\tau_n)}_{\alpha_n}$. On the "singular" set, this finction achieves values of order $-M_n/M^2_{n-1}$. Of course, the effective singular set is much more fragmented than it appears on this figure.

We start with a "good" interval $I_{k_1,\ldots,k_{n-1}}$, i.e. $(k_1,\ldots,k_{n-1}) \in J^g_{n-1}$ and simply write τ for $\tau_{n-1}|_{I_{k_1,\ldots,k_{n-1}}}$. If $\tau > 0$, define $J^{k_1,\ldots,k_{n-1},c}$, where c stands for "change", as $\{m_n - \tau + 1, \ldots, m_n\}$. This set consists of those indices k_n such that the interval $I_{k_1,\ldots,k_n}$ is not mapped into $T^{(\tau_{n-1})}_{\alpha_{n-1}}(I_{k_1,\ldots,k_{n-1}})$ under $T^{(\tau_{n-1})}_{\alpha_n}$. If $\tau < 0$, we define $J^{k_1,\ldots,k_{n-1},c}$ as $\{1,\ldots,|\tau|\}$.

The complement $\{1,\dots,m_n\}\backslash J^{k_1,\dots,k_{n-1},c}$ is denoted by $J^{k_1,\dots,k_{n-1},u}$, where u stands for "unchanged".

Define $\tau_n := \tau_{n-1} = \tau$ on the intervals $I_{k_1,\dots,k_{n-1},k_n}$, for $k_n \in J^{k_1,\dots,k_{n-1},u}$. For x in one of those intervals we have by (52), (53) and (47) that

$$\varphi^n(x) - \varphi^n(T^{(\tau_n)}_{\alpha_n}(x)) = \varphi^{n-1}(x) - \varphi^{n-1}(T^{(\tau_{n-1})}_{\alpha_{n-1}}(x)),$$

which yields (54) with $n-1$ replaced by n.

On the remaining intervals $I_{k_1,\dots,k_n}$ with $k_n \in J^{k_1,\dots,k_{n-1},c}$ we define τ_n such that it takes constant values in $\{-M_n+1,\dots,M_n-1\}$ on each of these intervals, such that (52) (resp. (53)) is satisfied, and such that these intervals $I_{k_1,\dots,k_n}$ are mapped onto the "remaining gaps" in $T^{(\tau_{n-1})}_{\alpha_{n-1}}(I_{k_1,\dots,k_{n-1}})$.

The crucial observation is that the intervals $I_{k_1,\dots,k_{n-1},k_n}$ where we have $\tau_n \neq \tau_{n-1}$, i.e. where $k_n \in J^{k_1,\dots,k_{n-1},c}$, are all on the "boundary" of $I_{k_1,\dots,k_{n-1}}$: they are the $|\tau|$ many intervals on the left or right end of $I_{k_1,\dots,k_{n-1}}$, depending on the sign of τ. Similarly, the "remaining gaps" in $T^{(\tau_{n-1})}_{\alpha_{n-1}}(I_{k_1,\dots,k_{n-1}})$ are the $|\tau|$ many intervals on the opposite end of $T^{(\tau_{n-1})}_{\alpha_{n-1}}(I_{k_1,\dots,k_{n-1}})$. Hence we may apply assertion (iii) of Lemma 3.4 to conclude that

$$|\varphi^n(x) - \varphi^n(T^{(\tau_n)}_{\alpha_n}(x))| \leq c(M_1,\dots,M_{n-1}),$$

for those $x \in I_{k_1,\dots,k_{n-1}}$ where $\tau_n(x) \neq \tau_{n-1}(x)$. Summing over all "good intervals" $I_{k_1,\dots,k_{n-1}}$, where $(k_1,\dots,k_{n-1}) \in J^g_{n-1}$, we conclude that the contribution to (55), with $n-1$ replaced by n, is controlled by the following factors: M_{n-1}, which is a bound for the number of elements in J^g_{n-1}, times M_{n-1}, which is a bound for $|\tau|$, times $\frac{1}{M_n}$, which is the length of the intervals $I_{k_1,\dots,k_n}$, times the above found constant $c(M_1,\dots,M_{n-1})$. In total, this implies the estimate (55), with $n-1$ replaced by n.

We now turn to item (iv), i.e. to the "singular" indices: fix $k_1,\dots,k_{n-1} \in J^s_{n-1}$ and let $\Delta\varphi$ denote the constant

$$\Delta\varphi := \varphi^{n-1}(T^{(\tau_{n-1})}_{\alpha_{n-1}}(x)) - \varphi^{n-1}(x), \quad x \in I_{k_1,\dots,k_{n-1}},$$

and again τ the constant $\tau_{n-1|I_{k_1,\dots,k_{n-1}}}$, so that $0 \leq \Delta\varphi \leq |\tau| < M_{n-1}$.
Similarly as for the case $n = 2$ define

$$J^{k_1,\dots,k_{n-1},g,l} = \{k_n^l, k_n^l + 1 \dots, \tfrac{m_n-1}{2}\}, \quad J^{k_1,\dots,k_{n-1},g,r} = \{\tfrac{m_n+1}{2},\dots,k_n^r\}.$$

Here k_n^r is the largest number such that, for the orbit $(T^i_{\alpha_n}(x))_{i=\tau}^{\tau+\Delta\varphi M_{n-1}-1}$ and for $x \in I_{k_1,\dots,k_{n-1},k_n^r}$, all its members lie in the right half of the respective intervals $I_{k'_1,\dots,k'_{n-1}}$. In fact, we get as in the step $n=2$ that $k_n^r = m_n - (\tau + \Delta\varphi M_{n-1})$.
Similarly k_n^l is the smallest number such that, for the orbit $(T^i_{\alpha_n}(x))_{i=\tau}^{\tau-\Delta\varphi M_{n-1}+1}$ and for $x \in I_{k_1,\dots,k_{n-1},k_n^l}$, all its members are in the left half of the respective intervals $I_{k'_1,\dots,k'_{n-1}}$.
We get $k_n^l = \tau - \Delta\varphi M_{n-1} + 1$.
Now we define τ_n as

$$\tau_n(x) = \tau + \Delta\varphi M_{n-1}, \quad \text{for } x \in I_{k_1,\dots,k_{n-1},k_n}, k_n \in J^{k_1,\dots,k_{n-1},g,r},$$

and

$$\tau_n(x) = \tau - \Delta\varphi M_{n-1}, \quad \text{for } x \in I_{k_1,\dots,k_{n-1},k_n}, k_n \in J^{k_1,\dots,k_{n-1},g,l}.$$

Similarly as in (44) at step $n=2$, we get for $k_n \in J^{k_1,\dots,k_{n-1},g} := J^{k_1,\dots,k_{n-1},g,l} \cup J^{k_1,\dots,k_{n-1},g,r}$, and $x \in I_{k_1,\dots,k_{n-1},k_n}$ that

$$\begin{aligned}\varphi^n(x) &- \varphi^n(T_{\alpha_n}^{(\tau_n)}(x)) \\ &= [\varphi^n(x) - \varphi^n(T_{\alpha_n}^{(\tau_{n-1})}(x))] + [\varphi^n(T_{\alpha_n}^{(\tau_{n-1})}(x)) - \varphi^n(T_{\alpha_n}^{(\tau_n)}(x))] \\ &= [\varphi^{n-1}(x) - \varphi^{n-1}(T_{\alpha_{n-1}}^{(\tau_{n-1})}(x))] + [\varphi^n(T_{\alpha_n}^{(\tau_{n-1})}(x)) - \varphi^n(T_{\alpha_n}^{(\tau_n)}(x))] \\ &= -\Delta\varphi + \Delta\varphi = 0.\end{aligned}$$

We still have to deal with the "singular" indices

$$J^{k_1,\dots,k_{n-1},s} := \{1,\dots,m_n\} \setminus J^{k_1,\dots,k_{n-1},g} = \{1,\dots,k_n^l - 1\} \cup \{k_n^r + 1,\dots,m_n\},$$

which consists of $2\Delta\varphi M_{n-1}$ many indices. This number is bounded by $2M_{n-1}^2$ as $\Delta\varphi \le |\tau| < M_{n-1}$. These intervals have to be mapped onto the "remaining gaps" in the interval $T_{\alpha_{n-1}}^{(\tau_{n-1})}(I_{k_1,\dots,k_{n-1}})$. Make the crucial observation that, while the intervals $I_{k_1,\dots,k_{n-1},k_n}$, for $k_n \in J^{k_1,\dots,k_{n-1},s}$, are at the boundary of $I_{k_1,\dots,k_{n-1}}$, the "remaining gaps" are in the middle of the interval $T_{\alpha_{n-1}}^{(\tau_{n-1})}(I_{k_1,\dots,k_{n-1}})$. This fact is analogous to the situation for $n = 1$ and $n = 2$.

Now define τ_n on the intervals $I_{k_1,\dots,k_{n-1},k_n}$ for $k_n \in J^{k_1,\dots,k_{n-1},s}$, in such a way that $T_{\alpha_n}^{(\tau_n)}$ maps these intervals onto the "remaining gaps" in $T_{\alpha_{n-1}}^{(\tau_{n-1})}(I_{k_1,\dots,k_{n-1}})$ and such that τ_n is constant on each of these intervals, takes values in $\{-M_n+1,\dots,M_n-1\}$ and such that (52) (resp. (53)) is satisfied with $n-1$ replaced by n. Applying Lemma 3.4, assertion (ii) as well as $2(M_{n-1}+1)|\tau|$ many times assertion (i) we obtain, for $x \in I_{k_1,\dots,k_{n-1},k_n}$ and $k_n \in J^{k_1,\dots,k_{n-1},s}$,

$$\varphi^n(x) - \varphi^n(T_{\alpha_n}^{(\tau_n)}(x)) \le -\tfrac{m_n}{2M_{n-1}} + c(M_1,\dots,M_{n-1}).$$

Assuming that m_n is sufficiently large as compared to M_{n-1} we have that the right hand side is negative.
Keeping in mind that there are $2\Delta\varphi M_{n-1}$ many indices in $J^{k_1,\dots,k_{n-1},s}$, we may estimate the "singular mass" on the interval $I_{k_1,\dots,k_{n-1}}$ by

$$\begin{aligned}\sum_{k_n \in J^{k_1,\dots,k_{n-1},s}} \int_{I_{k_1,\dots,k_{n-1},k_n}} &[\varphi^n(x) - \varphi^n(T_{\alpha_n}^{(\tau_n)}(x))]\, dx \\ &\le 2\Delta\varphi M_{n-1}[-\tfrac{m_n}{2M_{n-1}} + c(M_1,\dots,M_{n-1})] \ \tfrac{1}{M_n} \\ &= -\tfrac{\Delta\varphi}{M_{n-1}} \ [1 - \tfrac{c(M_1,\dots,M_{n-1})}{m_n}].\end{aligned} \tag{58}$$

We have by the inductive hypothesis that

$$\sum_{k_1,\dots,k_{n-1} \in J_{n-1}^s} \int_{I_{k_1,\dots,k_{n-1}}} [\varphi^{n-1}(x) - \varphi^{n-1}(T_{\alpha_n}^{(\tau_{n-1})}(x))]\, dx$$

$$\le -1 + \tfrac{3}{m_1} + \tfrac{c(M_1)}{m_2} + \dots + \tfrac{c(M_1,\dots,M_{n-2})}{m_{n-1}},$$

or, writing now $\Delta\varphi_{k_1,\dots,k_{n-1}}$ for the above value of $\Delta\varphi$ on the interval $I_{k_1,\dots,k_{n-1}}$,

$$\tfrac{1}{M_{n-1}} \sum_{k_1,\dots,k_{n-1} \in J_{n-1}^s} \Delta\varphi_{k_1,\dots,k_{n-1}} \le -1 + \tfrac{3}{m_1} + \tfrac{c(M_1)}{m_2} + \dots + \tfrac{c(M_1,\dots,M_{n-2})}{m_{n-1}}.$$

Letting $J_n^s := \bigcup_{k_1,\dots,k_{n-1}\in J_{n-1}^s} \{(k_1,\dots,k_{n-1},k_n) : k_n \in J^{k_1,\dots,k_{n-1},s}\}$ we obtain from (58)

$$\begin{aligned}\sum_{k_1,\dots,k_n\in J_n^s} \int_{I_{k_1,\dots,k_n}} &[\varphi^n(x) - \varphi^n(T_{\alpha_n}^{(\tau_n)}(x))]\,dx \\ &\le (-1 + \tfrac{3}{m-1} + \dots + \tfrac{c(M_1,\dots,M_{n-2})}{m_{n-1}})(1 - \tfrac{c(M_1,\dots,M_{n-1})}{m_n}) \\ &= -1 + \tfrac{3}{m_1} + \dots + \tfrac{c(M_1,\dots,M_{n-2})}{m_{n-1}} + \tfrac{c(M_1,\dots,M_{n-1})}{m_n}.\end{aligned}$$

where we may have increased the constant $c(1,\dots,M_{n-1})$ in the last line. This concludes the inductive step.

Construction of the Example: Let $\alpha = \lim_{n\to\infty} \alpha_n$ so that $T_\alpha = \lim_{n\to\infty} T_{\alpha_n}$ is the shift by the irrational number α.

The sequence $(\tau_n)_{n=1}^\infty$ of functions $\tau_n : [0,1) \to \mathbb{Z}$ converges, by (54), almost surely to a $\mathbb{Z}$-valued function $\tau = \lim_{n\to\infty} \tau_n$. Hence the maps $(T_{\alpha_n}^{(\tau_n)})_{n=1}^\infty$ converge almost surely to a map

$$T_\alpha^{(\tau)} : \begin{cases} [0,1) & \to \quad [0,1) \\ \quad x & \mapsto \quad T_\alpha^{(\tau)}(x) = T_\alpha^{\tau(x)}(x). \end{cases}$$

Using the fact that each $T_{\alpha_n}^{(\tau_n)}$ is a measure preserving almost sure bijection on $[0,1)$, it is straightforward to check that $T_\alpha^{(\tau)}$ is so too.

Letting $\Gamma_\tau = \{(x, T_\alpha^{(\tau)}(x)),\ x \in [0,1)\}$ in analogy to the notations $\Gamma_0 = \{(x,x),\ x \in [0,1)\}$ and $\Gamma_1 = \{(x, T_\alpha(x)),\ x \in [0,1)\}$, we define

$$c(x,y) = \begin{cases} h_+(x,y), & \text{if } (x,y) \in \Gamma_0 \cup \Gamma_1 \cup \Gamma_\tau, \\ \infty & \text{otherwise,} \end{cases}$$

where h is defined in (15) above. From this definition we deduce the almost sure identity, for $\tau(x) > 0$,

$$\begin{aligned} h(x, T_\alpha^{(\tau)}(x)) &= \#\{i \in \{0,\dots,\tau(x)-1\} : T_\alpha^i(x) \in [0,\tfrac12)\} \\ &\quad - \#\{i \in \{0,\dots,\tau(x)-1\} : T_\alpha^i(x) \in [\tfrac12,1)\} + 1 \\ &= \lim_{n\to\infty} [\varphi^n(x) - \varphi^n(T_{\alpha_n}^{(\tau_n)}(x))] + 1, \end{aligned} \tag{59}$$

a similar formula holding true for $\tau(x) < 0$.

As regards the Borel functions $(\varphi_n, \psi_n)_{n=1}^\infty$ announced in (17), (18) and (19) above, we need to slightly modify the functions $(\varphi^n, \psi^n)_{n=1}^\infty$ constructed in the above induction to make sure that they satisfy the inequality

$$\varphi_n(x) + \psi_n(y) \le c(x,y), \quad \text{for } x \in X, y \in Y. \tag{60}$$

As $c = \infty$ outside of $\Gamma_0 \cup \Gamma_1 \cup \Gamma_\tau$ it is sufficient to make sure that the following inequalities hold true almost surely, for $x \in [0,1)$:

(0) $\quad \varphi_n(x) + \psi_n(x) \le c(x,x) = 1,$

(1) $\quad \varphi_n(x) + \psi_n(T_\alpha(x)) \le c(x, T_\alpha(x)) = \begin{cases} 2, & \text{for } x \in [0,\tfrac12), \\ 0, & \text{for } x \in [\tfrac12, 1), \end{cases}$

(τ) $\quad \varphi_n(x) + \psi_n(T_\alpha^{(\tau)}(x)) \le c(x, T_\alpha^{(\tau)}(x)).$

The above constructed $(\varphi^n, \psi^n)_{n=1}^\infty$ only satisfy condition (0). We still have to pass from φ^n to a smaller function φ_n – while leaving $\psi_n := \psi^n$ unchanged – to satisfy (1) and (τ) too. Let

$$\begin{aligned}\varphi_n(x) := \varphi^n(x) - [\varphi^n(x) + \psi^n(T_\alpha(x)) - c(x, T_\alpha(x))]_+ \\ - [\varphi^n(x) + \psi^n(T_\alpha^{(\tau)}(x)) - c(x, T_\alpha^{(\tau)}(x)]_+.\end{aligned} \tag{61}$$

Clearly $\varphi_n \leq \varphi^n$ and the functions (φ_n, ψ_n) satisfy the inequality (60).

We have to show that the functions φ_n defined in (61) *satisfy that $\varphi^n - \varphi_n$ is small in the norm of $L^1(\mu)$, as $n \to \infty$, that is*

$$\lim_{n\to\infty} \int_{[0,1)} (\varphi^n(x) - \varphi_n(x))\, dx = 0, \tag{62}$$

provided that $(m_n)_{n=1}^\infty$ increases sufficiently fast to infinity.

We may estimate the first correction term in (61) by

$$\begin{aligned}&[\varphi^n(x) + \psi^n(T_\alpha(x)) - c(x, T_\alpha(x))]_+ \\ &\qquad \leq [\psi^n(T_\alpha(x)) - \psi^n(T_{\alpha_n}(x))]_+ + [\varphi^n(x) + \psi^n(T_{\alpha_n}(x)) - c(x, T_\alpha(x))]_+.\end{aligned}$$

The second term above is dominated by $\mathbb{1}_{I^n_{\text{middle}}}$ which is harmless as $\|\mathbb{1}_{I^n_{\text{middle}}}\|_{L^1(\mu)} = \frac{1}{M_n}$. As regards the first term, note that $T_\alpha(x) \ominus T_{\alpha_n}(x) = \alpha - \alpha_n = \sum_{j=n+1}^\infty \frac{1}{M_j}$ which we may bound by $\frac{2}{M_{n+1}}$ by assuming that $(m_n)_{n=1}^\infty$ increases sufficiently fast to infinity. As ψ^n is constant on each of the M_n many intervals $I_{k_1,\ldots,k_n}$ we get

$$\mu\{x \in [0,1) : \psi^n(T_\alpha(x)) \neq \psi^n(T_{\alpha_n}(x)\} \;\leq\; M_n(\alpha - \alpha_n) < \tfrac{2}{m_{n+1}}.$$

On this set we may estimate, using only the obvious bound $|\psi_n(x)| < M_n$, that

$$|\psi^n(T_\alpha(x)) - \psi^n(T_{\alpha_n}(x))| \leq 2M_n, \quad x \in [0,1),$$

to obtain

$$\|\psi^n(T_\alpha(x)) - \psi^n(T_{\alpha_n}(x))\|_{L^1(\mu)} < \tfrac{4M_n}{m_{n+1}}.$$

Hence for $(m_n)_{n=1}^\infty$ growing sufficiently fast to infinity, the first correction term in (61) is also small in L^1-norm.

To estimate the second correction term in (61) note that

$$\varphi^n(x) + \psi^n(T_\alpha^{(\tau)}(x)) = \varphi^n(x) + \psi^n(T_{\alpha_n}^{(\tau_n)}(x)), \quad \text{for } x \in [0,1). \tag{63}$$

Indeed, $T_{\alpha_n}^{(\tau_n)}$ induces a permutation between the intervals $I_{k_1,\ldots,k_n}$ and, by assertion (i) preceding the formula (52), we have that $T_{\alpha_{n+j}}^{(\tau_{n+j})}$ maps the intervals $I_{k_1,\ldots,k_n}$ onto the intervals $T_{\alpha_n}^{(\tau_n)}(I_{k_1,\ldots,k_n})$, for each $j \geq 0$. Noting that ψ^n is constant on each of the intervals $I_{k_1,\ldots,k_n}$ we obtain (63), by letting j tend to infinity.

By (47), $\varphi^n(x) + \psi^n(T_{\alpha_n}^{(\tau_n)}(x))$ is the number of visits to L^n minus the number of visits to R^n plus one, of the orbit $(T_{\alpha_n}^j)_{j=0}^{\tau_n(x)-1}$. Similarly, by (15), $h(x, T_\alpha^{\tau(x)}(x))$ is the number of visits to L minus the number of visits to R plus one, of the orbit $(T_\alpha^j)_{j=0}^{\tau(x)-1}$. We have to show that the positive part of the difference

$$f_n(x) := [\varphi^n(x) + \psi^n(T_{\alpha_n}^{(\tau_n)}(x)) - h_+(x, T_\alpha^{(\tau)}(x))]_+, \quad x \in [0,1), \tag{64}$$

is small in L^1-norm, as $n \to \infty$. To do so, we argue separately on $I^1_{\text{middle}} = [\frac{1}{2} - \frac{1}{2M_1}, \frac{1}{2} + \frac{1}{2M_1}]$, on the union of the "good" intervals at level n : $G_n = \bigcup_{(k_1,\ldots,k_n)\in J_n^g} I_{k_1,\ldots,k_n}$, and the union of the "singular" intervals at level n, $S_n = \bigcup_{(k_1,\ldots,k_n)\in J_n^s} I_{k_1,\ldots,k_n}$.

- For $x \in I^1_{\text{middle}}$, the correction term $f_n(x)$ in (64) simply equals zero as $\tau_n(x) = \tau(x) = 0$.

- For $x \in S_n$, we have by (56) that $\varphi^n(x) + \psi^n(T^{\tau_n(x)}_{\alpha_n}(x)) \leq 1$ so that $f_n(x) \leq 1$ too; hence $\lim_{n\to\infty} \|f_n \mathbb{1}_{S_n}\|_{L^1(\mu)} = 0$.

- For $x \in G_n$, we use

$$\begin{aligned} f_n(x) &\leq [\varphi^n(x) + \psi^n(T^{(\tau_n)}_{\alpha_n}(x)) - h(x, T^{\tau(x)}_\alpha(x))]_+ \\ &\leq \sum_{k=n+1}^{\infty} [(\varphi^{k-1}(x) + \psi^{k-1}(T^{(\tau_{k-1})}_{\alpha_{k-1}}(x))) - (\varphi^k(x) + \psi^k(T^{(\tau_k)}_{\alpha_k}(x)))]_+ \end{aligned}$$

  and (55) to conclude that

$$\lim_{n\to\infty} \|f_n \mathbb{1}_{G_n}\|_{L^1(\mu)} \leq \lim_{n\to\infty} \sum_{k=n+1}^{\infty} \tfrac{c(M_1,\ldots,M_{k-1})}{m_k} = 0.$$

This proves (62).
Hence (17), (18) and (19) are satisfied.

As regards assertion (16), let us verify that π_0 and π_1 are optimal transport plans. Indeed, it follows from (17) and (18) that the dual value of the present transport problem is greater than or equal to one which implies that $\langle c, \pi_0\rangle = \langle c, \pi_1\rangle = 1$ is the optimal primal value.

The fact that $\langle c, \pi_\tau\rangle > 1$ should be rather obvious to a reader who has made it up to this point of the construction. It follows from rough estimates. The set $\{[0, \frac{1}{2}) \cap \{\tau = -1\}\} \cup \{[\frac{1}{2}, 1) \cap \{\tau = 1\}\}$ has measure bigger than $1 - \frac{3}{M_1} + \sum_{i=2}^{\infty} \frac{c(M_1,\ldots,M_{i-1})}{m_i}$, which is bigger than, say, $\frac{3}{4}$, for $(m_n)_{n=1}^{\infty}$ tending sufficiently quick to infinity. As $c(x, T^{(\tau)}_\alpha(x))$ equals 2 on this set we get

$$\langle c, \pi_\tau\rangle \geq \tfrac{3}{2} > 1.$$

A slightly more involved argument, whose verification is left to the energetic reader, shows that, for $\varepsilon > 0$, we may choose $(m_n)_{n=1}^{\infty}$ such that

$$\langle h, \pi_\tau\rangle \geq 2 - \varepsilon. \tag{65}$$

Finally, we show assertion (iv) at the beginning of this section (see (20)). Let $\widehat{h} \in L^1(\pi)^{**}$ be a dual optimizer in the sense of [BLS09, Theorem 4.2]. We know from this theorem that there is a sequence $(\varphi_n, \psi_n)_{n=1}^{\infty}$ of bounded Borel functions[5] such that

$$(\alpha)\ \lim_{n\to\infty} \|[\varphi_n \oplus \psi_n - c]_+\|_{L^1(\pi)} = 0 \tag{66}$$

$$(\beta)\ \lim_{n\to\infty} \Big(\int_X \varphi_n(x)\, d\mu(x) + \int_Y \psi_n(y)\ d\nu(y)\Big) = 1, \tag{67}$$

$$(\gamma)\ \lim_{n\to\infty} \varphi_n \oplus \psi_n = \widehat{h}^r, \quad \pi\text{-a.s.}, \tag{68}$$

$$(\delta)\ \hat{h} \text{ is a } \sigma(L^1(\pi)^{**}, L^\infty(\pi)) \text{ cluster point of } (\varphi_n \oplus \psi_n)_{n=1}^{\infty}. \tag{69}$$

Here $\widehat{h} = \widehat{h}^r + \widehat{h}^s$ is the decomposition of $\widehat{h} \in L^1(\pi)^{**}$ into its regular part $\widehat{h}^r \in L^1(\pi)$ and into its purely singular part $\widehat{h}^s \in L^1(\pi)^{**}$.

[5] The (φ_n, ψ_n) need not be the same as the special sequence constructed above; still we find it convenient to use the same notation.

We shall show that $\widehat{h}^r$ equals h, π-almost surely. Indeed by assertions (66) and (67) above we have that, for $x \in [0,1)$,

$$\lim_{n\to\infty} (\varphi_n(x) + \psi_n(x)) = c(x,x) = h(x,x) = 1,$$

and

$$\lim_{n\to\infty} (\varphi_n(x) + \psi_n(T_\alpha(x))) = c(x,T_\alpha(x)) = h(x,T_\alpha(x)) = \begin{cases} 2, & \text{for } x \in [0,\frac{1}{2}), \\ 0, & \text{for } x \in [\frac{1}{2},1), \end{cases}$$

the limit holding true in $L^1([0,1],\mu)$ as well as for μ-a.e. $x \in [0,1)$, possibly after passing to a subsequence. As in the discussion following [BLS09, Theorem 4.2] this implies that, for each fixed $i \in \mathbb{Z}$,

$$\lim_{n\to\infty} (\varphi_n(x) + \psi_n(T_\alpha^i(x))) = h(x,T_\alpha^i(x)), \quad i \in \mathbb{Z},$$

the limit again holding true in $L^1(\mu)$ and μ-a.s., after possibly passing to a diagonal subsequence. Whence, we obtain with (68) that

$$\lim_{n\to\infty} (\varphi_n(x) + \psi_n(T_\alpha^{(\tau)}(x))) = h(x,T_\alpha^{(\tau)}(x)) = \widehat{h}^r(x,T_\alpha^{(\tau)}(x)),$$

convergence now holding true for μ-a.e. $x \in [0,1]$.
As $x \to T_\alpha^{(\tau)}(x)$ is a measure preserving bijection we get

$$\int_{[0,1)} [\varphi_n(x) + \psi_n(T_\alpha^{(\tau)}(x))]\,dx = \int_{[0,1)} (\varphi_n(x) + \psi_n(x))\,dx = 1,$$

so that, using (65) we get

$$\begin{aligned} & \lim_{n\to\infty} \int_{[0,1)} [\varphi_n(x) + \psi_n(T_\alpha^{(\tau)}(x))] \mathbb{1}_{\{\varphi_n(x)+\psi_n(T_\alpha^{(\tau)}(x))<h(x,T_\alpha^{(\tau)}(x))\}}(x)\,dx \\ = {} & 1 - \lim_{n\to\infty} \int_{[0,1)} [\varphi_n(x) + \psi_n(T_\alpha^{(\tau)}(x))] \mathbb{1}_{\{\varphi_n(x)+\psi_n(T_\alpha^{(\tau)}(x))\geq h(x,T_\alpha^{(\tau)}(x))\}}(x)\,dx \\ = {} & 1 - \langle h, \pi_\tau \rangle \\ < {} & 0. \end{aligned}$$

From $\lim_{n\to\infty} \mu\big\{x : \varphi_n(x) + \psi_n(T_\alpha^{(\tau)}(x)) < h(x,T_\alpha^{(\tau)}(x))\big\} = 0$ we conclude that each σ^*-cluster point of $([\varphi_n(\cdot) + \psi_n(T_\alpha^{(\tau)}(\cdot))]_-)_{n=1}^\infty$ is a purely singular element of $L^1(\pi)^{**}$ of norm equal to $\langle h, \pi_\tau \rangle - 1$.

Finally, we still have to specify the prime numbers $(m_n)_{n=1}^\infty$ in the above induction. It is now clear what we need: apart from satisfying the conditions of Lemma 3.1 as well as the requirements whenever we wrote "*for m_n tending sufficiently fast to infinity*", we choose the $(m_n)_{n=1}^\infty$ inductively such that in (54) we have $\frac{M_{n-2}}{m_{n-1}} < 2^{-n}$, that in (55) we have $\frac{c(M_1,\ldots,M_{n-2})}{m_{n-1}} < 2^{-n}$ and in (57) we have $\frac{3}{m_1} < \frac{1}{4}$ as well as again $\frac{c(M_1,\ldots,M_{n-2})}{m_{n-1}} < 2^{-n}$.

Hence we have shown all the assertions (i)-(iv) of Example 3.1 and the construction of the example is complete. ☐

4 A Relaxation of the Dual Problem

As in [BLS09, Remark 3.4], for a given cost function $c : X \times Y \to [0,\infty]$, we consider the family of pairs of functions

$$\Psi^{\mathrm{rel}}(\mu,\nu) = \left\{ \begin{array}{l} (\varphi,\psi) : \varphi,\psi \text{ Borel, integrable and} \\ \varphi(x) + \psi(y) \leq c(x,y),\ \pi\text{-a.s,} \\ \text{for each finite transport plan } \pi \in \Pi(\mu,\nu,c) \end{array} \right\}$$

and define the relaxed value of the dual problem as

$$D^{\mathrm{rel}} = \sup\Big\{ \int_X \varphi \, d\mu + \int_Y \psi \, d\nu : (\varphi,\psi) \in \Psi^{\mathrm{rel}}(\mu,\nu)\Big\}. \tag{70}$$

Using the notation of [BLS09] it is obvious that $D \leq D^{\mathrm{rel}}$ and it is straightforward to verify that the trivial duality inequality $D^{\mathrm{rel}} \leq P$ still is satisfied. One might conjecture – and the present authors did so for some time – that $D^{\mathrm{rel}} = P$ holds true in full generality, i.e. for arbitrary Borel measurable cost functions $c : X \times Y \to [0,\infty]$, defined on the product of two polish spaces X and Y. In this section we construct a counterexample showing that this is not the case, i.e. it may happen that we have a duality gap $P - D^{\mathrm{rel}} > 0$. The example will be a variant of the example in the previous section, i.e. the $(n+1)$'th variation of [AP03, Example 3.2].

In section 3 we constructed a measure preserving bijection $T_\alpha^{(\tau)} : [0,1) \to [0,1)$ having certain properties; we now shall construct a sequence $(T_\alpha^{(\tau_n)})_{n=0}^\infty$ of such maps and consider as cost function the restriction of h_+, where h is defined in (15) to the graphs $(\Gamma_n)_{n=0}^\infty$ of the maps $(T_\alpha^{(\tau_n)})_{n=0}^\infty$. This sequence also "builds up a singular mass", which now is positive as opposed to the negative singular mass in the previous section, but it does so in a different way. We resume the properties of these maps which we shall construct in the following proposition.

Proposition 4.1. *With the notation of section 3 there is an irrational $\alpha \in [0,1)$ and a sequence $(\tau_n)_{n=0}^\infty$ of maps $\tau_n : [0,1) \to \mathbb{Z}$, with $\tau_0 = 0$ and $\tau_1 = 1$, such that the transformations $T_\alpha^{(\tau_n)} : [0,1) \to [0,1)$, defined by*

$$T_\alpha^{(\tau_n)}(x) = T_\alpha^{\tau_n(x)}(x), \qquad x \in [0,1),$$

have the following properties.

(i) Each τ_n is constant on a countable collection of disjoint, half open intervals in $[0,1)$ whose union has full measure. For $n \geq 0$, the map $T_\alpha^{(\tau_n)}$ defines a measure preserving almost sure bijection of $([0,1),\mu)$ onto itself, where $\mu = \nu$ denotes Lebesgue measure on $[0,1)$. We have, for each $n \geq 0$,

$$\int_{[0,1)} h(x, T_\alpha^{(\tau_n)}(x))\, dx = 1. \tag{71}$$

(ii) The function

$$f_n(x) := h(x, T_\alpha^{(\tau_n)}(x)), \qquad x \in [0,1),$$

where h is defined in (15), satisfies

$$\|f_n - g_n\|_{L^1(\mu)} < 2^{-n} \tag{72}$$

where g_n is a Borel function on $[0,1)$ such that

$$\mu\{g_n = 0\} = 1 - \eta_n, \qquad \mu\{g_n = \tfrac{1-\eta_n}{\eta_n}\} = \eta_n \tag{73}$$

for some sequence $(\eta_n)_{n=1}^\infty$ tending to zero.

(iii) There is a sequence $(\varphi_n, \psi_n)_{n=1}^\infty$ of bounded Borel functions such that, for every fixed $n \in \mathbb{N}$,

$$\lim_{m\to\infty} \|h(x, T_\alpha^{(\tau_n)}(x)) - [\varphi_m(x) + \psi_m(T_\alpha^{(\tau_n)}(x))]\|_{L^1(\mu)} = 0,$$

and

$$\lim_{n\to\infty} \Big[\int_{[0,1)} \varphi_n(x)\, dx + \int_{[0,1)} \psi_n(y)\, dy \Big] = 1.$$

(iv) The sequence $(T_\alpha^{(\tau_n)})_{n=1}^\infty$ converges to the identity map in the following sense:

$$\delta(x, T_\alpha^{(\tau_n)}(x)) < 2^{-n}, \qquad x \in [0,1),\ n \geq 1, \tag{74}$$

where $\delta(\cdot,\cdot)$ denotes the Riemannian metric on $\mathbb{T} = [0,1)$.

We postpone the proof of the proposition and first draw some consequences. Suppose that α as well as $(T_\alpha^{(\tau_n)})_{n=0}^\infty$ have been defined and satisfy the assertions of Proposition 4.1.

Proposition 4.2. *Fix $M \geq 2$ and define the cost function $c_M : [0,1) \times [0,1) \to [0,\infty]$ by*

$$c_M(x,y) = \begin{cases} h_+(x,y), & \text{for } (x,y) \text{ in the graph of } T_\alpha^0, T_\alpha^1, T_\alpha^{(\tau_2)}, T_\alpha^{(\tau_3)}, \dots, T_\alpha^{(\tau_M)}, \\ \infty, & \text{otherwise.} \end{cases}$$

For this cost function c_M we find that the primal value, denoted by P^M, as well as the dual value, denoted by D^M, of the Monge–Kantorovich problem both are equal to 1.

In addition, there is $\beta = \beta(M) > 0$, such that, for every partial transport

$$\sigma \in \Pi^{\text{part}}(\mu,\nu) := \{\sigma : \mathcal{M}(X \times Y) : p_X(\pi) \leq \mu, p_Y(\pi) \leq \nu\}$$

with

$$\|\sigma\| \geq \tfrac{2}{3} \text{ and } \int_{X\times Y} c_M(x,y)\, d\sigma(x,y) \leq \tfrac{1}{2},$$

there is no partial transport $\varrho \in \Pi^{\text{part}}(\mu,\nu)$ with

$$\|\sigma + \varrho\| = 1 \text{ and } \sigma + \varrho \in \Pi(\mu,\nu)$$

with the property that ϱ is supported by

$$\Delta^\beta = \{(x,y) \in [0,1)^2 : \delta(x,y) < \beta\}.$$

Proof. First note that there is an open and dense subset $G \subseteq [0,1)$ of full measure $\mu(G) = 1$ such that c_M, restricted to $G \times G$ is lower semi-continuous. This follows from assertion (i) of Proposition 4.1 by replacing the half open intervals by their open interior. Noting that G is polish we may apply the general duality theory [Kel84] to the cost function c_M restricted to $G \times G$ to conclude that there is no duality gap for the cost function $c_M|_{G\times G}$. It follows that there is also no duality gap for the original setting of c_M, defined on $[0,1) \times [0,1)$, either.

We claim that, for every $M \geq 0$, the value D^M of the dual problem equals 1. Indeed, let $(\varphi_n, \psi_n)_{n=1}^\infty$ be a sequence as in Proposition 4.1 (iii). Defining

$$\tilde{\varphi}_n := \varphi_n - \sum_{j=0}^{M} [\varphi_n(x) + \psi_n(T_\alpha^{(\tau_j)}(x)) - h(x, T_\alpha^{(\tau_j)}(x))]_+$$

and $\tilde{\psi}_n = \psi_n$, we have that

$$\tilde{\varphi}_n(x) + \tilde{\psi}_n(y) \leq h(x,y) \leq h_+(x,y),$$

for all (x,y) in the graph of $T_\alpha^0, T_\alpha^1, T_\alpha^{(\tau_2)}, \dots, T_\alpha^{(\tau_M)}$, and

$$\lim_{n\to\infty} \Big[\int_X \tilde{\varphi}_n(x)\, dx + \int_Y \tilde{\psi}_n(y)\, dy \Big] = 1,$$

showing that $D^M \geq 1$. It follows that $D^M = P^M = 1$.

Now suppose that the final assertion of the proposition is wrong to find a sequence $(\sigma_n)_{n=1}^\infty \in \Pi^{\mathrm{part}}(\mu,\nu)$ with $\|\sigma_n\| \geq \frac{2}{3}$ and $\int_{X\times Y} c_M(x,y)\, d\sigma_n(x,y) \leq \frac{1}{2}$, as well as a sequence $(\varrho_n)_{n=1}^\infty \in \Pi^{\mathrm{part}}(\mu,\nu)$ with $\|\pi_n + \varrho_n\| = 1$ and $\pi_n + \varrho_n \in \Pi(\mu,\nu)$ such that ϱ_n is supported by

$$\Delta^{1/n} = \{(x,y) \in [0,1)^2 : \delta(x,y) < \tfrac{1}{n}\}. \tag{75}$$

Considering $(\sigma_n)_{n=1}^\infty$ as measures on the product $G\times G$ of the polish space G, we then can find by Prokhorov's theorem a subsequence $(\sigma_{n_k})_{k=1}^\infty$ converging weakly on $G\times G$ to some $\sigma \in \Pi^{\mathrm{part}}(\mu,\nu)$, for which we find $\|\sigma\| \geq \frac{2}{3}$ and $\int_{X\times Y} c(x,y)\; d\sigma(x,y) \leq \frac{1}{2}$. By passing once more to a subsequence, we may also suppose that $(\varrho_{n_k})_{k=1}^\infty$ weakly converges (as measures on $G\times G$ or $[0,1)\times[0,1)$; here it does not matter) to some $\varrho \in \Pi^{\mathrm{part}}(\mu,\nu)$ for which we get $\|\sigma+\varrho\| = 1$ and $\sigma + \varrho \in \Pi(\mu,\nu)$. By (75) we conclude that ϱ induces the identity transport from its marginal $p_X(\varrho)$ onto its marginal $p_Y(\varrho) = p_X(\varrho)$. As $c_M(x,x) = 1$, for $x \in [0,1)$ we find that $\int_{X\times Y} c_M(x,y)\; d\varrho(x,y) = \|\varrho\| \leq \frac{1}{3}$, which implies that

$$\int c_M(x,y)\; d(\pi+\varrho)(x,y) \leq \tfrac{1}{2} + \tfrac{1}{3},$$

a contradiction to the fact that $P^M = 1$ which finishes the proof. ☐

We now can proceed to the construction of the example.

Proposition 4.3. *Assume the setting of Proposition 4.1. For a subsequence $(i_j)_{j=2}^\infty$ of $\{2,3,\dots\}$ we define the cost function $c : [0,1)\times[0,1) \to [0,\infty]$ by*

$$c(x,y) = \begin{cases} h_+(x,y), & \text{for } (x,y) \text{ in the support of } T_\alpha^0, T_\alpha^1, T_\alpha^{(\tau_{i_2})}, T_\alpha^{(\tau_{i_3})}, \dots, T_\alpha^{(\tau_{i_j})}, \dots \\ \infty, & \text{otherwise.} \end{cases} \tag{76}$$

If $(i_j)_{j=2}^\infty$ tends sufficiently fast to infinity we have that, for this cost function c, the primal value P is strictly positive, while the relaxed primal value P^{rel} (see [BLS09, Example 4.3]) as well as the dual value D and the relaxed dual value D^{rel} (see (70)) all are equal to 0.

In particular there is a duality gap $P - D^{\mathrm{rel}} > 0$, disproving the conjecture mentioned at the beginning of this section.

Proof. We proceed inductively: let $j \geq 2$ and suppose that $i_0 = 0, i_1 = 1, i_2, \dots, i_j$ have been defined. Apply Proposition 4.2 to

$$c_j(x,y) = \begin{cases} h_+(x,y), & \text{for } (x,y) \text{ in the support of } T_\alpha^0, T_\alpha^1, T_\alpha^{(\tau_{i_2})}, T_\alpha^{(\tau_{i_3})}, \dots, T_\alpha^{(\tau_{i_j})}, \\ \infty, & \text{otherwise,} \end{cases}$$

to find $\beta_j > 0$ satisfying the conclusion of Proposition 4.2. We may and do assume that $\beta_j \leq \min(\beta_1, \dots, \beta_{j-1})$. Now choose i_{j+1} such that

$$\delta(x, T_\alpha^{(\tau_{i_{j+1}})}(x)) < \beta_j, \quad x \in [0,1). \tag{77}$$

This finishes the inductive step and well-defines the cost function $c(x,y)$ in (76).

By (71) each $T_\alpha^{(\tau_{i_j})}$ induces a Monge transport $\pi_{i_j} \in \Pi(\mu,\nu)$ which satisfies

$$\int_{X\times Y} h(x,y)\, d\pi_{i_j}(x,y) = \int_X h(x, T_\alpha^{(\tau_{i_j})})\, dx = 1.$$

The fact that the relaxed primal value P^{rel} for the cost function c equals zero, directly follows from the definition of P^{rel} [BLS09, Section 1.1], (72) and (73) by transporting the measure $\mu\mathbb{1}_{\{g_n=0\}}$, which has mass $1-\eta_n$, via the Monge transport map $T_\alpha^{(\tau_n)}$ where n is a large element of the sequence $(i_j)_{j=1}^\infty$. Hence we conclude from [BLS09, Theorem 1.2] that the dual value D of the Monge–Kantorovich problem for the cost function c defined in (76) also equals zero.

Finally observe that we have $D = D^{\mathrm{rel}}$ in the present example: indeed, the set $\{(x,y) \in [0,1)^2 : c(x,y) < \infty\}$ is the countable union of the supports of the finite cost Monge transport plans $T_\alpha^0, T_\alpha^1, T_\alpha^{(\tau_{i_1})}, T_\alpha^{(\tau_{i_2})}, \ldots, T_\alpha^{(\tau_{i_j})}, \ldots$, so that the requirements $\varphi(x) + \psi(y) \le c(x,y)$, for all $(x,y) \in [0,1)^2$, and $\varphi(x) + \psi(y) \le c(x,y)$, π-a.s., for each finite transport plan $\pi \in \Pi(\mu,\nu)$, coincide (after possibly modifying $\varphi(x)$ on a μ-null set).

What remains to prove is that the primal value P satisfies $P > 0$. We shall show that, for every transport plan $\pi \in \Pi(\mu,\nu)$, we have $\int_{X\times Y} c(x,y)\ d\pi(x,y) \ge \frac{1}{2}$. Assume to the contrary that there is $\pi \in \Pi(\mu,\nu)$ such that

$$\int_{X\times Y} c(x,y)\ d\pi(x,y) < \tfrac{1}{2}.$$

Denoting by σ_j the restriction of π to the union of the graphs of the maps T_α^0, T_α^1, $T_\alpha^{(\tau_{i_1})}$, $T_\alpha^{(\tau_{i_2})}$, $\ldots, T_\alpha^{(\tau_{i_j})}$, each σ_j is a partial transport in $\Pi^{\mathrm{part}}(\mu,\nu)$ and the norms $(\|\sigma_j\|)_{j=1}^\infty$ increase to one. Choose j such that

$$\|\sigma_j\| > \tfrac{2}{3}.$$

We apply Proposition 4.2 to conclude that there is no partial transport plan ϱ_j such that $\pi_j + \varrho_j \in \Pi(\mu,\nu)$, and such that ϱ_j is supported by Δ^{β_j}. But this is a contradiction as $\varrho_j = \pi - \sigma_j$ has precisely these properties by (77). □

Proof of Proposition 4.1: The construction of the example described by Proposition 4.1 will be an extension of the construction in the previous section from which we freely use the notation.

We shall proceed by induction on $j \in \mathbb{N}$ and define a double-indexed family of maps $\tau_{n,j} : [0,1) \to \mathbb{Z}$, where $1 \le n \le j$.

Step $j = 1$**:** Define

$$\tau_{1,1} : [0,1) \to \mathbb{Z}$$

as

$$\tau_{1,1} = -\tau_1,$$

where we have $m_1 = M_1, \alpha_1 = \frac{1}{M_1}$ and τ_1 as in (28) above. At this stage the only difference to the previous section is that we change the sign of τ_1 as we now shall build up a "positive singular mass", as opposed to the "negative singular mass" which we constructed in the previous section. More precisely, defining φ^1, ψ^1 as in (27), we obtain, similarly as in (29)

$$\varphi^1(x) + \psi^1(T_{\alpha_1}^{(\tau_{1,1})}(x)) = \begin{cases} 0, & \text{for } x \in I_{k_1}, k_1 \in \{2, \ldots, (M_1-1)/2, \\ & \qquad (M_1+3)/2, \ldots, M_1 - 1\}, \\ (M_1-1)/2, & \text{for } x \in I_{k_1}, k_1 = 1, M_1, \\ 1, & \text{for } x \in I_{(M_1+1)/2}. \end{cases}$$

This finishes the inductive step for $j = 1$.

Step $j = 2$**:** Let m_2 and $M_2 = M_1 m_2$ be as in section 3, where m_2 satisfies the requirements of Lemma 3.1, and still is free to be eventually specified. To define $\tau_{1,2} : [0,1) \to \mathbb{Z}$ we want to make sure that the map $T_{\alpha_2}^{(\tau_{1,2})}$ maps the intervals I_{k_1} bijectively onto $T_{\alpha_1}^{(\tau_{1,1})}(I_{k_1})$. Using the notation of the previous section, we consider *all* the intervals I_{k_1} as "*good*" intervals so that we do not have to take extra care of some *"singular"* intervals.

More precisely, fix $1 \leq k_1 \leq M_1$, and write τ for $\tau_{1,1}|_{I_{k_1}}$. If $\tau > 0$, define $J^{k_1,c}$ as $\{m_2 - \tau + 1, \ldots, m_2\}$, i.e. the set of those indices k_2 such that the interval I_{k_1,k_2} is not mapped into $T_{\alpha_1}^{(\tau_{1,1})}(I_{k_1})$ under $T_{\alpha_2}^{(\tau_{1,1})}$. If $\tau < 0$, we define $J^{k_1,c}$ as $\{1, \ldots, |\tau|\}$, and if $\tau = 0$, we define $J^{k_1,c}$ as the empty set. The complement $\{1, \ldots, m_2\} \backslash J^{k_1,c}$ is denoted by $J^{k_1,u}$.

Define $\tau_{1,2} := \tau_{1,1} = \tau$ on the intervals I_{k_1,k_2}, for $k_2 \in J^{k_1,u}$. On the remaining intervals I_{k_1,k_2} with $k_2 \in J^{k_1,c}$ we define $\tau_{1,2}$ such that it takes constant values in $\{-M_2+1, \ldots, M_2-1\}$ on each of these intervals, such that (37) (resp. (38) is satisfied, and such that these intervals I_{k_1,k_2} are mapped onto the "remaining gaps" in $T_{\alpha_1}^{(\tau_{1,1})}(I_{k_1})$.

Using again Lemma 3.3 we resume the properties of the thus constructed map $T_{\alpha_2}^{(\tau_{1,2})}$: $[0, 1) \to [0, 1)$.

(i) The measure-preserving bijection $T_{\alpha_2}^{(\tau_{1,2})}$ maps each interval I_{k_1} onto $T_{\alpha_1}^{(\tau_{1,1})}(I_{k_1})$. It induces a permutation of the intervals I_{k_1,k_2}, where $1 \leq k_1 \leq M_1, 1 \leq k_2 \leq m_2$.

(ii) Defining φ^2, ψ^2 as in (32) we get, for each $1 \leq k_1 \leq M_1$, similarly as in (39) and (40)

$$\mu[I_{k_1} \cap \{\tau_{1,2} \neq \tau_{1,1}\}] \leq \tfrac{M_1}{m_2} \mu[I_{k_1}],$$

as well as

$$\sum_{k_1=1}^{M_1} \int_{I_{k_1}} |(\varphi^1(x) - \varphi^1(T_{\alpha_1}^{(\tau_{1,1})}(x)) - (\varphi^2(x) - \varphi^2(T_{\alpha_2}^{(\tau_{1,2})}(x))| dx < \frac{4M_1^2}{m_2}.$$

(iii) On the middle interval $I^1_{\text{middle}} = I_{\frac{M_1+1}{2}}$ we have $\tau_{1,2} = \tau_{1,1} = 0$.

We now pass to the construction of the map $\tau_{2,2} : [0, 1) \to \mathbb{Z}$. We define, for each $1 \leq k_1 \leq M_1$, and $x \in I_{k_1,k_2}$,

$$\tau_{2,2}(x) = \begin{cases} a_2(k_2), & \text{for } k_2 \in \{1, \ldots, M_1\} \\ -M_1, & \text{for } k_2 \in \{M_1 + 1, \ldots, (m_2 - 1)/2\}, \\ 0, & \text{for } k_2 = (m_2 + 1)/2, \\ M_1, & \text{for } k_2 \in \{(m_2 + 3)/2, \ldots, m_2 - M_1\}, \\ a_2(k_2), & \text{for } k_2 \in \{m_2 - M_1 + 1, \ldots, m_2\}. \end{cases}$$

The definition of the function a_2 on the "singular" intervals I_{k_1,k_2}, where $k_2 \in \{1, \ldots, M_1\} \cup \{m_2 - M_1 + 1, \ldots, m_2\}$ is done such that $T_{\alpha_2}^{(\tau_{2,2})}$ maps these intervals onto "remaining gaps" I_{k_1,l_2}, where l_2 runs through the set

$$\{(m_2 - 1)/2 - M_1 + 1, \ldots, (m_2 - 1)/2\} \cup \{(m_2 + 3)/2, \ldots, (m_2 + 3)/2 + M_1 - 1\}$$

in the middle region of the interval I_{k_1}. As above we require in addition that a_2 on each I_{k_1,k_2} takes constant values in $\{-M_2 + 1, \ldots, M_2 - 1\}$ and that (37) (resp. (38)) is satisfied.

The function $\tau_{2,2}$ mimics the construction of $\tau_{1,1}$ above, with the role of $[0, 1)$ replaced by each of the intervals I_{k_1}, for $1 \leq k_1 \leq M_1$. The idea is that, $T_{\alpha_1}^{M_1}$ being the identity map, we have that $T_{\alpha_2}^{M_1}$ satisfies $T_{\alpha_2}^{M_1}(x) = x \oplus \frac{M_1}{M_2}$ and $\frac{M_1}{M_2} = \frac{1}{m_2}$ is small. Hence the role of T_{α_1} in the previous section now is taken by $T_{\alpha_2}^{M_1}$.

More precisely, we have, for each $k_1 = 1, \ldots, M_1$, and $x \in I_{k_1,k_2}$

$$\varphi^2(x) + \psi^2(T_{\alpha_2}^{(\tau_{2,2})}(x)) =$$
$$= \begin{cases} 0, & \text{for } k_2 \in \{M_1 + 1, \ldots, (m_2 - 1)/2, \\ & \qquad (m_2 + 3)/2, \ldots, m_2 - M_1\}, \\ \frac{m_2}{2M_1} + \gamma(M_1), & \text{for } k_2 \in \{1, \ldots, M_1\} \cup \{m_2 - M_1 + 1, \ldots, m_2\}, \\ 1, & \text{for } k_2 = (m_2 + 1)/2. \end{cases} \tag{78}$$

The notation $\gamma(M_1)$ denotes a quantity verifying $|\gamma(M_1)| \leq c(M_1)$ for some constant $c(M_1)$, depending only on M_1. The verification of (78) uses Lemma 3.3 and is analogous as in section 3.

As $T_{\alpha_2}^{(\tau_{2,2})}$ defines a measure preserving bijection on $[0,1)$, we get

$$\int_0^1 (\varphi^2(x) + \psi^2(T_{\alpha_2}^{(\tau_{2,2})}(x))\ dx = \int_0^1 (\varphi^2(x) + \psi^2(x))\, dx = 1. \tag{79}$$

This finishes the inductive step for $j = 2$.

General Inductive step: For prime numbers $m_1, \ldots, m_{j-1}$ as in the previous section suppose that we have defined, for $1 \leq n \leq j-1$ maps $\tau_{n,j} : [0,1) \to \mathbb{Z}$ such that the following inductive hypotheses are satisfied.

(i) For $1 \leq n \leq j-1$, the measure preserving bijection $T_{\alpha_i}^{(\tau_{n,j-1})} : [0,1) \to [0,1)$ maps the intervals $I_{k_1,\ldots,k_{n-1}}$ onto themselves. It induces a permutation of the intervals $I_{k_1,\ldots,k_{j-1}}$, where $1 \leq k_1 \leq m_1, \ldots, 1 \leq k_{j-1} \leq m_{j-1}$.

(ii) For $1 \leq n < j-1$ we have, for $1 \leq k_1 \leq m_1, \ldots, 1 \leq k_{j-2} \leq m_{j-2}$,

$$\mu[I_{k_1,\ldots,k_{j-2}} \cap \{\tau_{n,j-2} \neq \tau_{n,j-1}\}] \leq \tfrac{M_{j-2}}{m_{j-1}} \mu[I_{k_1,\ldots,k_{j-2}}], \tag{80}$$

and

$$\begin{aligned}\sum_{1\leq k_1 \leq m_1,\ldots,1\leq k_{j-2}\leq m_{j-2}} \int_{I_{k_1,\ldots,k_{j-2}}} &\Big|\Big(\varphi^{j-2}(x) - \varphi^{j-2}(T_{\alpha_{j-2}}^{(\tau_{n,j-2})}(x))\Big) \\ &- \Big(\varphi^{j-1}(x) - \varphi^{j-1}(T_{\alpha_{j-1}}^{(\tau_{n,j-1})}(x))\Big)\Big| dx < \tfrac{c(M_1,\ldots,M_{j-2})}{m_{j-1}}.\end{aligned} \tag{81}$$

We now shall define $\tau_{n,j} : [0,1) \to \mathbb{Z}$, for $1 \leq n \leq j$ and $\tau_{j,j} : [0,1) \to \mathbb{Z}$.

Fix $1 \leq n \leq j-1$ as well as $1 \leq k_1 \leq m_1, \ldots, 1 \leq k_{j-1} \leq m_{j-1}$. Denote by τ the constant value $\tau_{n,j-1}|_{I_{k_1,\ldots,k_{j-1}}}$. If $\tau > 0$ define $J^{k_1,\ldots,k_{j-1},c}$ as $\{m_j - \tau + 1, \ldots, m_j\}$, similarly as for the case $j = 2$ above. If $\tau \leq 0$ define $J^{k_1,\ldots,k_{j-1},c}$ as $\{1, \ldots, |\tau|\}$ which, for $\tau = 0$, equals the empty set. On the intervals $I_{k_1,\ldots,k_{j-1},k_j}$ where k_j lies in the complement $J^{k_1,\ldots,k_{j-1},u} = \{1, \ldots, m_j\} \backslash J^{k_1,\ldots,k_{j-1},c}$ we define $\tau_{n,j} := \tau_{n,j-1}$. On the remaining intervals $I_{k_1,\ldots,k_{j-1},k_j}$, where $k_j \in J^{k_1,\ldots,k_{j-1},c}$, we define $\tau_{n,j}$ in such a way that it takes constant values in $\{-M_j + 1, \ldots, M_j - 1\}$ on each of these intervals, such that (37) (resp. (38)) is satisfied, and such that these intervals $I_{k_1,\ldots,k_{j-1},k_j}$ are mapped onto the "remaining gaps" in $T_{\alpha_{j-1}}^{(\tau_{n,j-1})}(I_{k_1,\ldots,k_{j-1}})$.

Similarly as in the previous section we thus well-define the function $\tau_{n,j}$ which then verifies (80) and (81), with $j-1$ replaced by j.

We still have to define $\tau_{j,j} : [0,1) \to \mathbb{Z}$. For $1 \leq k_1 \leq m_1, \ldots, 1 \leq k_{j-1} \leq m_{j-1}$, we define $\tau_{j,j}(x)$ on the intervals $I_{k_1,\ldots,k_{j-1},k_j}$ by

$$\tau_{j,j}(x) = \begin{cases} a_j(k_j), & \text{for } k_j \in \{1, \ldots, M_{j-1}\} \\ -M_j, & \text{for } k_j \in \{M_{j-1}+1, \ldots, (m_j-1)/2\}, \\ 0, & \text{for } k_j = (m_j+1)/2, \\ M_j, & \text{for } k_j \in \{(m_j+3)/2, \ldots, m_j - M_{j-1}\}, \\ a_j(k_j), & \text{for } k_j \in \{m_j - M_{j-1} + 1, \ldots, m_j\}. \end{cases}$$

Similarly as in step $j = 2$ the $\{-M_j + 1, \dots, M_j - 1\}$-valued function $a_j(k_j)$ is defined in such a way that $T^{(\tau_j)}_{\alpha_j}$ maps the intervals $I_{k_1,\dots,k_{j-1},k_j}$ with $k_j \in \{1, \dots, M_{j-1}\} \cup \{m_j - M_{j-1} + 1, \dots, m_j\}$ to the intervals $I_{k_1,\dots,k_{j-1},k_j}$, where k_j runs through the "middle region"

$$\{(m_j - 1)/2 - M_{j-1} + 1, \dots, (m_j - 1)/2\} \cup \{(m_j + 3)/2, \dots, (m_j + 3)/2 + M_{j-1} - 1\}.$$

We now deduce from Lemma 3.3 that, for $x \in I_{k_1,\dots,k_{j-1},k_j}$

$$\begin{aligned} &\varphi^j(x) + \varphi^j(T^{(\tau_{j,j})}_{\alpha_j}(x)) = \\ &= \begin{cases} 0, & \text{for } k_j \in \{M_{j-1} + 1, \dots, (m_j - 1)/2\} \\ & \qquad \cup \{(m_j + 3)/2, \dots, m_j - M_{j-1}\}, \\ \frac{m_j}{2M_{j-1}} + \gamma(M_1, \dots, M_{j-1}), & \text{for } k_j \in \{1, \dots, M_{j-1}\} \cup \{m_j - M_{j-1} + 1, \dots, m_j\}, \\ 1, & \text{for } k_j = (m_j + 1)/2, \end{cases} \end{aligned}$$

where $\gamma(M_1, \dots, M_{j-1})$ denotes a quantity which is bounded in absolute value by a constant $c(M_1, \dots, M_{j-1})$ depending only on $M_1, \dots, M_{j-1}$.

This completes the inductive step.

We now define $\tau_0 = 0, \tau_1 = 1$ and, for $n \geq 2$

$$\tau_n = \lim_{j \to \infty} \tau_{n-1,j}. \tag{82}$$

It follows from (80) that, for each $n \geq 2$, the limit (82) exists almost surely provided the sequence $(m_n)_{n=1}^\infty$ converges sufficiently fast to infinity, similarly as in section 3 above. The $(\tau_n)_{n=0}^\infty$ and the above constructed functions $(\varphi_n, \psi_n)_{n=1}^\infty$ satisfy the assertions of Proposition 4.1. The verification of items (i), (ii), and (iii) is analogous to the arguments of section 3 and therefore skipped. As regards assertions (iv) note that, for $1 \leq n \leq j$ the function $T^{(\tau_{n,j})}_{\alpha_j}$ maps the intervals $I_{k_1,\dots,k_{n-1}}$ onto themselves. It follows that $T^{(\tau_n)}_\alpha$ does so too, whence

$$\delta(x, T^{(\tau_n)}_\alpha(x)) < M_n^{-1},$$

which readily shows (74). □

References

[AE06] J.-P. Aubin and I. Ekeland. *Applied nonlinear analysis*. Dover Publications Inc., Mineola, NY, 2006. Reprint of the 1984 original.

[AP03] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In *Optimal transportation and applications (Martina Franca, 2001)*, volume 1813 of *Lecture Notes in Math.*, pages 123–160. Springer, Berlin, 2003.

[BLS09] M. Beiglböck, C. Léonard, and W. Schachermayer. A general duality theorem for the monge–kantorovich transport problem. *submitted, preprint available on* www.arxiv.org, 2009.

[BS09] M. Beiglböck and W. Schachermayer. Duality for Borel measurable cost functions. *Trans. Amer. Math. Soc.*, to appear, 2009.

[ET99] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[GM96] W. Gangbo and R.J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[Kel84] H.G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984.

[Pra08] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 2008. in press.

[PSU88] A.L. Peressini, F.E. Sullivan, and J.J. Uhl, Jr. *The mathematics of nonlinear programming*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1988.

[Roc66] Ralph T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17:497–510, 1966.

[RR96] D. Ramachandran and L. Rüschendorf. Duality and perfect probability spaces. *Proc. Amer. Math. Soc.*, 124(7):2223–2228, 1996.

[Rüs96] L. Rüschendorf. On c-optimal random variables. *Statist. Probab. Lett.*, 27(3):267–270, 1996.

[ST08] W. Schachermayer and J. Teichmann. Characterization of optimal transport plans for the Monge-Kantorovich problem. *Proceedings of the A.M.S.*, 2008. in press.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil09] C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.