2020

Transfunctions and Other Topics in Measure Theory

Jason Bentley
University of Central Florida

TRANSFUNCTIONS AND OTHER TOPICS IN MEASURE THEORY

by

JASON WILLIAM BENTLEY
M.S. University of Central Florida, 2017
B.A. Berea College, 2015

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2020

Major Professor: Piotr Mikusiński

ABSTRACT

Measures are versatile objects which can represent how populations or supplies are distributed within a given space by assigning sizes to subregions (or subsets) of that space. To model how populations or supplies are shifted from one configuration to another, it is natural to use functions between measures, called transfunctions. Any measurable function can be identified with its push-forward transfunction. Other transfunctions exist such as convolution operators. In this manner, transfunctions are treated as generalized functions.

This dissertation serves to build the theory of transfunctions and their connections to other mathematical fields. Transfunctions that identify with continuous or measurable push-forward operators are characterized, and transfunctions that map between measures concentrated in small balls – called localized transfunctions – can be spatially approximated with measurable functions or with continuous functions (depending on the setting). Some localized transfunctions have "fat graphs" in the product space and "fat continuous graphs" are necessarily formed by localized transfunctions.

Any Markov transfunction – a transfunction that is linear, variation-continuous, total-measure-preserving and positive – corresponds to a family of Markov operators and a family of plans (indexed by their marginals) such that all objects have the same "instructions" of transportation between input and output marginals. An example of a Markov transfunction is a push-forward transfunction. In two settings (continuous and measurable), the definition and existence of adjoints of linear transfunctions are formed and simple transfunctions are implemented to approximate linear weakly-continuous transfunctions in the weak sense. Simple Markov transfunctions can be used both to approximate the optimal cost between two marginals with respect to a cost function and to approximate Markov transfunctions in the weak sense. These results suggest implementing

future research to find more applications of transfunctions to optimal transport theory.

Transfunction theory may have potential applications in mathematical biology. Several models are proposed for future research with an emphasis on local spatial factors that affect survivorship, reproducibility and other features. One model of tree population dynamics (without local factors) is presented with basic analysis. Some future directions include the use of multiple numerical implementations through software programs.

ACKNOWLEDGMENTS

I could never imagine getting this far without all of the help I have received in the form of family, friends, teachers and professors.

Words cannot express how grateful I am to my family – my mother Marty, my father Randall, my brother Randy, my grandparents Crystal, Don, Ruth Ann and Truman, and so many others – for they taught me the gift of kindness, encouraged me to follow my passions, and instilled within me a stubborn-like perseverance.

The supportive friendships I have forged with Blake, Chandria, Nikki, Bran, Arielle, Ted, Daniel, Parisa and several others have brought many joyful moments that I hope to never forget. I am especially grateful of Nikki's love, help and support throughout the years as my academic journey was fortified by her diligence and dedication.

As time passes, my progress only emphasizes how significantly my teachers and professors have impacted my career and my approach to academics. With retrospect, I appreciate all of their guidance and support. In particular, I am inspired by my academic experiences with Zi-Xia Song, James Blackburn-Lynch, Kristen Barnard, Jing-Pang Lee, Jan Pierce, and especially with my Ph. D. advisor Piotr Mikusiński whose excellent pedagogy and intriguing storytelling never dulled with correspondence.

Without all of them, I would not be the person I am today. Thank you all.

To my family, friends, teachers, and professors.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 History of Generalized Functions

Functions are mathematical objects which map from points in one space to points in another space. Throughout history, many methods have been developed to extend function spaces to include non-function objects – which the author refers to as “generalized functions” – that hold some properties in common with functions. In this text, there are two types of generalized functions, distinguished by the properties to be maintained. The first type – including Schwartz distributions [4, 6], hyperfunctions [8], Mikusiński operators [14], Boehmians [11, 13], and pseudoquotients [27] – is motivated by generalizing real or complex functions from the perspective of differentiability and integrability. The second type – including Markov operators [10, 25], plans and transfunctions [32] – pertains to generalizing measurable functions from the perspective of transport mappings. Regardless, generalized functions abandon the emphasis of point-to-point mappings as the value of a function at one point does not influence its integral nor its local average values.

The theories of generalized functions were motivated by the task of finding solutions to ordinary or partial differential equations. Although many types of differential equations have straightforward methods to find solutions, some differential equations present obstacles too difficult to surpass without applying one of two types of advanced techniques. The first type of method is to apply an integral transform to the equation, simplify the equation, and invert the integral transform to obtain a solution. Alternatively, the second type of method is to enlarge the space of functions to a space of generalized functions, formulate the weak problem (the weak differential equation), find weak solutions, and verify that the solutions are functions that solve the original problem.

There are several types of integral transforms and several types of generalized functions. Common integral transforms include Fourier and Laplace transforms [2, 5], while common generalized functions include distributions and Mikusiński operators. The most well-known generalized function is the Dirac delta, also known as the point-charge or point-mass measure, which appears in both distribution theory and in Mikusiński operational calculus as the identity with respect to convolution. Some methods have advantages over others; for example, the Laplace integral transform is commonly applied to ordinary differential equations in the current undergraduate setting, yet the operational calculus of Mikusiński is a generalized function technique which is simpler to develop rigorously, nearly equivalent in execution, and applicable to a wider range of ordinary differential equations [12, 30].

Despite the contrasts highlighted above, the connections between integral transforms and generalized functions are fundamental. A linear differential equation can be expressed in the form $Lu = f$, where L is a linear operator, f is the inhomogeneous coefficient and u is the function solution to find. If L have a right inverse K, then K can be expressed as an integral transform by the Schwartz kernel theorem. The kernel of K is a Green's "function" G, which is a distribution satisfying $L(G) = \delta$, leading to $u = K(f) = \int Gf$ as a solution [3].

Recent results concerning Boehmians, pseudoquotients, and their connections to integral transforms have been developed by P. Mikusiński [15], D. Nemzer [16], V. Karunakaran, N. V. Kalpakam [18], D. Loonker, P. K. Banerji [19], J. Burzyk [22], R. Roopkumar [23], S. K. Q. Al-Omari [31], and several others.

1.2 Transfunctions

An alternative approach to generalized functions leads to the theory of transfunctions introduced by P. Mikusiński which has connections to optimal transport theory. To begin, let $\mathcal{M}_X$ and $\mathcal{M}_Y$

denote sets of finite measures on X and Y respectively. A transfunction is a mapping between $\mathcal{M}_X$ and $\mathcal{M}_Y$. Then one can identify any measurable function $f : X \to Y$ as the transfunction $f_\#$, where $f_\#$ is the push-forward operator defined via $(f_\# \mu)(B) = \mu(f^{-1}(B))$ for any measure $\mu \in \mathcal{M}_X$ and any measurable $B \subseteq Y$.

Push-forward operators are the solutions of interest in the Monge formulation of the transportation problem [1, 20]. However, there are Markov transfunctions that are not push-forward operators; an example is $\mu \mapsto \mu * \gamma$, where γ is a probability measure that is absolutely continuous and concentrated at the origin. The convolution transfunction acts by "blurring" μ with filter γ. Markov transfunctions form a convex set of a vector space, and each Markov transfunction corresponds to a family of plans and to a family of Markov operators (indexed by their marginals) in the sense that the transfunction contains the "instructions" needed to transport the input marginal of each Markov operator or plan to the corresponding output marginal. In this sense, the generalized function space of Markov transfunctions are analogous to the convex space of plans used in the Kantorovich formulation of the transportation problem [9, 24]. A similar theory of generalized functions – $\mathfrak{G}$-morphisms – based on extending push-forward operators in a categorical setting has been developed by M.D. Taylor and P. Mikusiński [26].

While the intuition behind transfunctions is similar to that of fuzzy functions, the mathematical formalisms of these two approaches are very different, [21]. On one hand, fuzzy functions do not have to be measurable, yet on the other hand, measures do not have to identify with fuzzy sets. Although neither theory contains the other, one can construct a setting where fuzzy functions can be identified with transfunctions and vice versa.

Another way to think of transfunctions is as follows. A family $\{\Phi_t : t \geq 0\}$ of transfunctions from $\mathcal{M}_X$ to $\mathcal{M}_X$ can be viewed instead as a family $\{\Phi_\bullet(\mu) : \mu \in \mathcal{M}_X\}$ of measure-valued functions on $[0, \infty)$. Then the measure-valued function $\Phi_\bullet(\mu)$ can describe the evolution of μ as time elapses, while the transfunction Φ_t can represent an overall rule on how measures will change from time 0 to time t. When $\{\Phi_t : t \geq 0\}$ form a (C_0) semigroup, many results follow as regular finite measures form Banach spaces [33].

1.3 Chapter Summaries

The following three chapters consist of three manuscripts that are relevant to transfunction theory and have been developed over the course of the author's dissertation. Chapters 2 and 4 are modifications of prepared manuscripts that are currently submitted to journals for review [36, 37]. Chapter 3 is a modification of a published manuscript [34]. The publisher has granted permission to the inclusion of the content of Chapter 3.

Chapter 2 details the construction of measures with nice properties in topological spaces most studied in the theory of transfunctions. Chapter 3 serves to characterize transfunctions that are push-forward operators of continuous and measurable functions, provide spatial approximations of transfunctions that map measures in a local manner, and define a notion of graph for transfunctions. Chapter 4 builds the theory of Markov transfunctions, develops adjoints and approximations of identity, and establishes their connections to Markov operators, plans, and optimal transport theory. Chapter 5 proposes some transfunction models for population dynamics in mathematical biology that are in early stages of development.

CHAPTER 2: CONSTRUCTION OF REGULAR NON-ATOMIC STRICTLY-POSITIVE MEASURES IN SECOND-COUNTABLE NON-ATOMIC LOCALLY COMPACT HAUSDORFF SPACES

2.1 Introduction

One well-known result from measure theory is the construction of the Lebesgue measure on the real line. One construction begins with a set function defined on the ring of finite unions of precompact intervals with rational endpoints which returns the total length. This construction of the Lebesgue measure has several crucial steps to show:

1. The set function has finite additivity on the ring.
2. The outer measure is at most ($\leq$) the set function on open sets in the ring.
3. The outer measure is at least ($\geq$) the set function on compact sets in the ring.
4. The outer measure on boundaries of open sets in the ring (subsets of $\mathbb{Q}$) is zero.
5. The outer measure and the set function agree on the ring, which also shows that the set function is a pre-measure on the ring.
6. Apply Carathéodory's Extension Theorem to extend the set function to the Lebesgue measure on the Borel subsets of the real line.

The Lebesgue measure on Borel sets is known to be regular, non-atomic and strictly-positive. On which topological spaces can measures with such properties be guaranteed?

This chapter answers this question by applying a similar construction to 2nd-countable non-atomic locally compact Hausdorff spaces. The choice of such spaces is motivated by the steps involved in the Lebesgue measure construction. In fact, steps 2 and 3 are guaranteed with the usual choice of outer measure, and once step 4 is established, steps 5 and 6 immediately follow.

Steps 1 and 4 prove to be challenging for these spaces. Indeed, the space is not necessarily a topological group, the ring of sets is only described topologically (or via a metric), and the finitely additive set function is not as easily defined.

This chapter carefully constructs a finitely-additive set function via the limit of a sequence of set functions defined recursively on a growing sequence of rings of sets. The sequence of rings of sets and the sequence of set functions must be coupled together in a non-trivial manner so that the outer measure satisfies step 4. Once the measure is formed, it can be shown to be regular, non-atomic and strictly-positive.

2.2 Notation

Let (X, τ_X) be a topological space and let $A \subseteq X$. Then A°, $\overline{A}$, ∂A and A^e denote the interior, closure, boundary and exterior of A respectively. For sets A and B, the expression $A \uplus B$ is similar to $A \cup B$ but emphasizes that A and B are disjoint, and the expression $A - B$ denotes the relative set complement: that is, $a \in A - B$ if $a \in A$ and $a \notin B$. In this chapter, " $\subset$ " will always mean proper subset, and " $\subseteq$ " will always mean subset or equality. An open set U is called *regular* if U is the interior of $\overline{U}$ and a closed set F is called *regular* if F is the closure of F°.

2.3 Non-Atomic Topologies

Definition 2.1. A topological space (X, τ_X), or a topology τ_X is *non-atomic* if for all $x \in X$ and for every open U containing x, there exists an open set V with $x \in V \subset U$.

The following properties of non-atomic topological spaces can be easily verified by definition.

Proposition 2.2. *Let (X, τ_X) be a topological space.*

(a) If (X, τ_X) is non-atomic, then every non-empty open set in τ_X must be infinite in cardinality. The converse is true if (X, τ_X) is a T_1-space.

(b) If (X, τ_X) is non-atomic, then every $x \in X$ and any open neighborhood U of x yield infinitely many open neighborhoods of x which are proper subsets of U.

Example 2.3. Any normed space over the real or complex field is non-atomic. In particular, $(\mathbb{R}, |\cdot|)$ is non-atomic since $x \in (a, b)$ implies $x \in (a + \varepsilon, b - \varepsilon) \subset (a, b)$ for any ε satisfying $0 < \varepsilon < \min\{x - a, b - x\}$.

Non-atomic spaces have a nice property pertaining to topological bases.

Proposition 2.4. *Let (X, τ_X) be a non-atomic topological space, let $\{U_i : i \in I\}$ be a topological basis for τ_X, and let J be a cofinite subset of I. Then $\{U_j : j \in J\}$ is also a topological basis for τ_X.*

For non-atomic locally compact Hausdorff spaces, the next proposition allows for open sets to be "bored" by compact closures of open subsets. Proposition 2.5 is straightforward to verify and will be utilized in Lemma 2.10 later.

Proposition 2.5. *Let (X, τ_X) be a non-atomic locally compact Hausdorff space. Then every open neighborhood $x \in U$ admits an open neighborhood $x \in V$ with compact closure and with $\overline{V} \subset U$. Additionally, V can be chosen to be regular such that V and $U - \overline{V}$ are disjoint non-empty open subsets of U with $U = V \uplus \partial V \uplus (U - \overline{V})$.*

2.4 Constructing the Desired Measure

Before stating the main theorem of this chapter, we recall some useful properties that measures may or may not possess in general.

Definition 2.6. A measure μ on measurable space (X, Σ_X) is

(a) *regular* if μ is finite on compact sets, if every open U satisfies

$$\mu(U) = \sup\{\mu(K) \ : \ K \subseteq U \ , \ K \text{ compact}\} \tag{2.1}$$

and if every measurable A satisfies

$$\mu(A) = \inf\{\mu(U) \ : \ A \subseteq U \ , \ U \text{ open}\}; \tag{2.2}$$

(b) *non-atomic* if every measurable A with $\mu(A) > 0$ admits $B \subset A$ with $0 < \mu(B) < \mu(A)$;

(c) *strictly-positive* if $\mu(U) > 0$ for every non-empty open U.

Theorem 2.7. *Let (X, Σ_X) be a measurable space generated by a second-countable locally compact Hausdorff non-atomic space (X, τ_X). Then there exists a finite regular non-atomic strictly-positive measure on (X, Σ_X).*

Theorem 2.7 will be proven in four stages. First, we need a candidate for the sequence of rings of sets. Since (X, τ_X) is a second-countable locally compact regular space, there exists a countable basis of nonempty regular open sets in X with compact closure. Let $(V_i) = (V_i)_{i=1}^{\infty}$ be one such basis expressed as a sequence.

Regularity of the open sets in the basis sequence grants several useful properties: V_i being regular implies that $\overline{V_i}$ and V_i^e are also regular; finite intersections of regular open sets are regular open sets; A and B being nonempty regular open sets with A intersecting ∂B implies that $A \cap B$ and $A \cap B^e$ are nonempty open sets.

The following lemma produces a sequence of rings of sets, denoted by $(\mathcal{D}_k)$, and a limit ring of subsets $\mathcal{D}$ which contains basis sets V_i for $i \in \mathbb{N}$ and generates all Borel sets from (X, τ_X).

Lemma 2.8. *Let (X, Σ_X) be a second-countable non-atomic locally compact Hausdorff space with previously established basis sequence $(V_i)_{i=1}^{\infty}$. For each $k \in \mathbb{N}$, define*

$$\begin{aligned}
\mathcal{A}_k &:= \{\cap_{i=1}^{k} R_i \mid R_i = V_i \text{ or } R_i = V_i^e \text{ for all } 1 \leq i \leq k\} - \{\varnothing, \cap_{i=1}^{k} V_i^e\},\\
\mathcal{B}_k &:= \{\uplus_{j=1}^{n} S_j \mid n \in \mathbb{N}_0, S_j \in \mathcal{A}_k \text{ for all } 1 \leq j \leq n\},\\
\mathcal{C}_k &:= \{C \mid C \subseteq \cup_{i=1}^{k} \partial V_i\},\\
\mathcal{D}_k &:= \{B \uplus C \mid B \in \mathcal{B}_k,\ C \in \mathcal{C}_k\}\\
\mathcal{D} &:= \cup_{k=1}^{\infty} \mathcal{D}_k.
\end{aligned} \tag{2.3}$$

Then: $\mathcal{A}_k$ consists of pairwise disjoint nonempty regular open sets; every $A \in \mathcal{A}_{k+1}$ with $A \subseteq \cup_{i=1}^{k} V_i$ has a unique $B \in \mathcal{A}_k$ with $A \subseteq B$; if $S, T \in \mathcal{B}_k$, then $S \cup T, S \cap T,$ and $S \cap T^e \in \mathcal{B}_k$; $(\mathcal{D}_k)$ is an ascending sequence of rings of sets; $\mathcal{D}$ is a ring of sets that generates the σ-algebra Σ_X and is insensitive to permutations on $(V_i)_{i=1}^{\infty}$.

Proof. If $\varnothing \neq A = \cap_{i=1}^{k+1} R_i \in \mathcal{A}_{k+1}$ with $A \subseteq \cup_{i=1}^{k} V_i$, then $B := \cap_{i=1}^{k} R_i$ is the unique set in $\mathcal{A}_k$ containing A.

Let $S := \uplus_{i=1}^{m} S_i$ and $T := \uplus_{j=1}^{n} T_j$ be in $\mathcal{B}_k$. It is fairly straight-forward to show that $\mathcal{B}_k$ is closed under finite unions and intersections. We will show that $S \cap T^e \in \mathcal{B}_k$. If $m = n = 1$, then $S, T \in \mathcal{A}_k$. Since sets in $\mathcal{A}_k$ are contained in each others' exteriors, the open set $\cup\{A \in \mathcal{A}_k - \{T\}\} \in \mathcal{B}_k$ is a subset of T^e. The remaining possible points in T^e are either in $\cup_{i=1}^{k} \partial V_i$ or in $\cap_{i=1}^{k} V_i^e$, which are both disjoint with S. Therefore, $S \cap T^e = S \cap (\cup\{A \in \mathcal{A}_k - \{T\}\}) \in \mathcal{B}_k$. This elementary case, in tandem with the closure of $\mathcal{B}_k$ under finite unions and intersections, proves the general case since $S \cap T^e = (\uplus_{i=1}^{m} S_i) \cap (\uplus_{j=1}^{n} T_j)^e = (\uplus_{i=1}^{m} S_i) \cap (\cap_{j=1}^{n} T_j^e) = \uplus_{i=1}^{m} \cap_{j=1}^{n} (S_i \cap T_j^e)$.

It is clear that $\mathcal{C}_k$ is closed under finite unions. If $D_1 = B_1 \uplus C_1$ and $D_2 = B_2 \uplus C_2$, where $B_1, B_2 \in \mathcal{B}_k$ and $C_1, C_2 \in \mathcal{C}_k$, then it follows that $D_1 \cup D_2 = (B_1 \cup B_2) \uplus (C_1 \cup C_2) \in \mathcal{D}_k$ via closure of $\mathcal{B}_k$ and $\mathcal{C}_k$ under finite unions. Furthermore,

$$
\begin{aligned}
D_1 - D_2 &= (B_1 \uplus C_1) \cap (B_2 \uplus C_2)^c = (B_1 \uplus C_1) \cap (B_2^c \cap C_2^c) \\
&= (B_1 \cap B_2^c \cap C_2^c) \uplus (C_1 \cap B_2^c \cap C_2^c) \\
&= (B_1 \cap B_2^c) \uplus (C_1 \cap B_2^c \cap C_2^c) \\
&= (B_1 \cap B_2^e) \uplus ((B_1 \cap \partial B_2) \cup (C_1 \cap B_2^c \cap C_2^c)) \in \mathcal{D}_k \qquad (2.4)
\end{aligned}
$$

since $B_1 \cap B_2^e \in \mathcal{B}_k$ and $(B_1 \cap \partial B_2) \cup (C_1 \cap B_2^c \cap C_2^c) \in \mathcal{C}_k$. Therefore, $\mathcal{D}_k$ is a ring of sets. It follows easily that $\mathcal{D} = \cup_{k=1}^{\infty} \mathcal{D}_k$ is also a ring of sets. Since $\mathcal{D}$ contains all basis sets and since τ_X is second-countable, $\mathcal{D}$ generates Σ_X.

If $D_k := B_k \uplus C_k \in \mathcal{D}_k$ with $B_k \in \mathcal{B}_k$ and $C_k \in \mathcal{C}_k$, then notice that $(B_k \cap V_{k+1}) \cup (B_k \cap V_{k+1}^e) \in \mathcal{B}_{k+1}$ and that $(B_k \cap \partial V_{k+1}) \cup C_k \in \mathcal{C}_{k+1}$ together ensure that $D_k \in \mathcal{D}_{k+1}$. Therefore, $\mathcal{D}_k \subseteq \mathcal{D}_{k+1}$ for all natural k.

Finally, let $\pi : \mathbb{N} \to \mathbb{N}$ be any permutation and form $\mathcal{A}'_k, \mathcal{B}'_k, \mathcal{C}'_k, \mathcal{D}'_k, \mathcal{D}'$ as in (2.3) with respect to $(V_{\pi(i)})_{i=1}^{\infty}$. Let $K \in \mathcal{D}$. Then $K \in \mathcal{D}_n$ for some natural n. Let m be the smallest natural number such that $\{V_{\pi(1)}, V_{\pi(2)}, ...V_{\pi(m)}\} \supseteq \{V_1, V_2, \ldots V_n\}$. It follows that $K \in \mathcal{D}'_m \subset \mathcal{D}'$. A similar argument shows that $\mathcal{D}' \subseteq \mathcal{D}$. □

For the second stage of proving Theorem 2.7, we need a finitely additive set function defined on $\mathcal{D}$. Given a basis sequence $(V_i)_{i=1}^{\infty}$, we need to intuitively develop a sequence of set functions $\mu_m : \mathcal{A}_m \to [0, 1]$ for $m \in \mathbb{N}$. The crucial idea is that when a regular open set $A \in \mathcal{A}_m$ intersects ∂V_{m+1}, regularity properties of A and V_{m+1} ensure that A is fragmented by ∂V_{m+1} into two nonempty open sets $A' := A \cap V_{m+1}$ and $A'' := A \cap V^e_{m+1}$ in $\mathcal{A}_{m+1}$. Hence, we can evenly divide the size $\mu_m(A)$ in half and distribute each to A' and A'', meaning we insist that $\mu_{m+1}(A') = \mu_{m+1}(A'') = \frac{1}{2}\mu_m(A)$. Of course, we also need to insist that the next set function equals the previous set function for open sets in $\mathcal{A}_m$ that persist in $\mathcal{A}_{m+1}$. Finally, when "new regions" are introduced, we can freely choose that size, and we shall do so in a way to cause all set functions to have maximum size output less than 1. These components correspond to the construction of $(\mu_m)_{m=1}^{\infty}$ in (2.5) from Lemma 2.9. Figure 2.1 illustrates how the set functions on $(\mathcal{A}_k)$ behave.

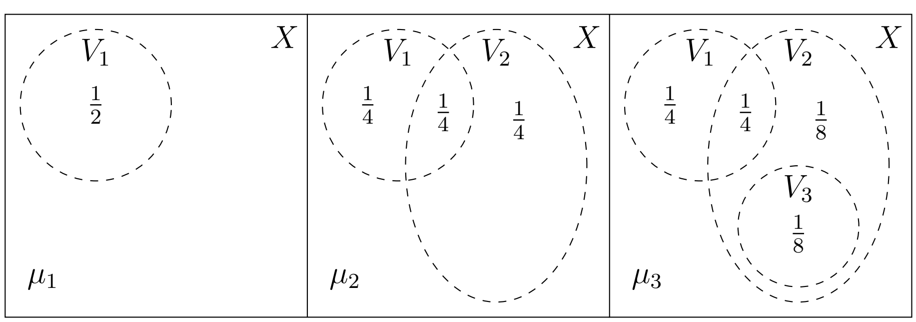

Figure 2.1: An example showing how μ_1, μ_2 and μ_3 are defined based on V_1, V_2 and V_3. Each rule defined in Lemma 2.9 is used at least once in constructing the set functions.

After $(\mu_m)_{m=1}^{\infty}$ is constructed, we can easily develop a sequence of finitely additive set functions $(\kappa_n)_{n=1}^{\infty}$ such that κ_{n+1} is an extension of κ_n for all $n \in \mathbb{N}$, then define a finitely additive set function κ as the overall extension of $(\kappa_n)_{n=1}^{\infty}$ to $\mathcal{D}$.

Lemma 2.9. *Let $(V_i)_{i=1}^{\infty}$ be a basis sequence for (X, τ_X), with $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k$ and $\mathcal{D}$ defined in Lemma 2.8 for all $k \in \mathbb{N}$. Recursively define a sequence of functions $\mu_m : \mathcal{A}_m \cup \{\varnothing\} \to [0,1]$ as follows:*

$$
\begin{aligned}
&\mu_1(\varnothing) := 0; && \\
&\mu_1(V_1) := 1/2; && \\
&\mu_m(A) := 1/2 \cdot \mu_{m-1}(B) && \textit{if } \varnothing \neq A \subset B \in \mathcal{A}_{m-1}; \\
&\mu_m(A) := \mu_{m-1}(A) && \textit{if } A \in \mathcal{A}_{m-1}; \\
&\mu_m(V_m - \cup_{i=1}^{m-1} \overline{V_i}) := 1/2^m && \textit{if } V_m - \cup_{i=1}^{m-1} \overline{V_i} \neq \varnothing.
\end{aligned}
\tag{2.5}
$$

Then there exists a finitely additive set function $\kappa : \mathcal{D} \to [0,1]$ such that $\kappa(A) = \mu_m(A)$ when $A \in \mathcal{A}_m$.

Proof. Define a sequence of functions $\nu_k : \mathcal{B}_k \to [0,1]$ via $\nu_k(\uplus_{j=1}^{n} S_j) := \sum_{j=1}^{n} \mu_k(S_j)$, where $S_j \in \mathcal{A}_k$ for $1 \leq j \leq n$. Then define a sequence of functions $\kappa_n : \mathcal{D}_n \to [0,1]$ via $\kappa_n(S \cup T) = \nu_n(S)$, where $S \in \mathcal{B}_n$ and $T \in \mathcal{C}_n$.

Finite additivity of (ν_k) and finite additivity of (κ_n) are easy to verify. It follows by the definitions of (ν_k) and (κ_n) and by (2.5) that any set in $\mathcal{D}_N$ will have the same value under all functions κ_n with $n \geq N$. Therefore, the set function $\kappa : \mathcal{D} \to [0,1]$ such that $\kappa(T) = \kappa_n(T)$ when $T \in \mathcal{D}_n$ is well defined. Finite additivity of κ follows from finite additivity of (κ_n). □

For the third stage of proving Theorem 2.7, we need to show that $\cup_k \partial V_k$ has zero outer measure.

However, notice that the set function κ we develop *depends on the order of the basis sequence* (V_i). This is important, because without a careful choice made for the ordering of these basis sets, step 4 in Section 2.1 may be difficult or impossible. What kind of sequence do we select? Lemma 2.10 serves two purposes: to provide the crucial properties needed to obtain steps 1 and 4 in Section 2.1, and to help verify non-atomicity of the measure formed at the end.

Lemma 2.10. *Let (X, τ_X) be a second-countable non-atomic locally compact Hausdorff space, let $(V_i)_{i=1}^{\infty}$ be a topological basis sequence, and let $(\mathcal{A}_k)_{k=1}^{\infty}$ be the sequence of collections of sets formed in Lemma 2.8 with respect to $(V_i)_{i=1}^{\infty}$. There exists a permutation $\pi : \mathbb{N} \to \mathbb{N}$ such that the set function κ formed in Lemma 2.9 with respect to $(V_{\pi(i)})_{i=1}^{\infty}$ has the following properties:*

(a) *The set $\cup_{k=1}^{\infty} \partial V_k$ has zero outer measure.*

(b) $\max\{\kappa(A) : A \in \mathcal{A}_n\} \to 0$ *as* $n \to \infty$.

Proof. Note that the boundaries $(\partial V_k)_{k=1}^{\infty}$ are compact, so we can find covers of them using other basis open sets. The utility of the non-atomic topological space is that we can purposefully use the closures of other basis elements to bore "closed holes" into a given cover of a given boundary ∂V_i, then find a *better* cover of the same boundary that does not intersect these holes. The holes should cause the new cover to have outer measure at most half of the previous cover's outer measure when κ is formed. This process is repeated for all ∂V_i countably many times so that the outer measure for each must be zero. Fortunately, the implementation below will also yield that $\max\{\kappa(A) : A \in \mathcal{A}_n\} \to 0$ as $n \to \infty$.

To form a permutation with the desired properties, we form a partition of $\mathbb{N}$ into three double-indexed families $\{F_{i,j} : i, j \in \mathbb{N}\}$, $\{G_{i,j} : i, j \in \mathbb{N}\}$ and $\{H_{i,j} : i, j \in \mathbb{N}\}$ of finite sets such that the following holds:

(i) $\mathcal{G}_{i,j} := \{V_k \mid k \in G_{i,j}\}$ covers ∂V_i for $i, j \in \mathbb{N}$; that is, $\partial V_i \subset \cup\mathcal{G}_{i,j}$.

(ii) $\mathcal{F}_{i,j} := \{\overline{V_k} \mid k \in F_{i,j}\}$ for $i, j \in \mathbb{N}$ satisfies $\cup\mathcal{G}_{i,j} \subseteq \cup\mathcal{G}_{i,j-1} - \cup\mathcal{F}_{i,j}$ for all $i \geq 1, j \geq 2$.

(iii) With $g(i,j) := \max G_{i,j}$, for each $i \geq 1, j \geq 2$, and for each $U \in \mathcal{A}_{g(i+1,j-1)}$, there exists a unique $K \in \mathcal{F}_{i,j}$ such that $K \subset U$.

(iv) $\max G_{i,j} < \min G_{i',j'}$ when $i + j < i' + j'$ or when $i + j = i' + j'$ and $j < j'$. The same is true for $\{F_{i,j}\}_{i,j=1}^{\infty}$ and $\{H_{i,j}\}_{i,j=1}^{\infty}$ whenever the compared sets are both nonempty.

(v) $H_{i,j}$ are remainder sets; that is,
$H_{1,1} = \{1, \ldots, g(1,1)\} - G_{1,1}$,
$H_{i,1} = \{g(1, i-1) + 1, \ldots, g(i,1)\} - G_{i,j}$ for all $i \geq 2$, and
$H_{i,j} = \{g(i-1, j+1) + 1, \ldots, g(i,j)\} - (G_{i,j} \cup F_{i,j})$ for all $i \geq 1, j \geq 2$.

Afterwards, we define the desired permutation $\pi : \mathbb{N} \to \mathbb{N}$, and hence the preferred sequence $(W_i)_{i=1}^{\infty} := (V_{\pi(i)})_{i=1}^{\infty}$ by making the arrangement $R_{i,j} := (F_{i,j}, G_{i,j}, H_{i,j})$ for each $i, j \geq 1$ and then stating that $R_{i,j} \leq R_{i',j'}$ exactly when $i + j < i' + j'$ or when $i + j = i' + j'$ and $j < j'$; that is, make the arrangement $R_{1,1}, R_{2,1}, R_{1,2}, R_{3,1}, R_{2,2}, R_{1,3}, \cdots$. The details are provided below.

$\{V_i\}_{i=1}^{\infty}$ covers the compact set ∂V_1, so there exists some finite subcover. Let $\mathcal{G}_{1,1}$ denote one possible choice, indexed by $G_{1,1} \subset \mathbb{N}$ with maximum index $g(1,1)$. Let $H_{1,1} := \{1, 2, \cdots, g(1,1)\} - G_{1,1}$.

Now $\{V_i : i > g(1,1)\}$ forms a basis by Proposition 2.4, so perform the same procedure on ∂V_2 using the new basis, obtaining the collection $\mathcal{G}_{2,1}$, indexed by $G_{2,1}$ and number $g(2,1)$. Define $H_{2,1} := \{g(1,1)+1, g(1,1)+2, \cdots, g(2,1)\} - G_{2,1}$. Next to construct is $\mathcal{G}_{1,2}, G_{1,2},$ and $g(1,2)$. For each $A \in \mathcal{A}_{g(2,1)}$, there exists a basis set $V_{k(A)}$ such that $\overline{V_{k(A)}} \subset A$ and $k(A) > g(2,1)$. Doing this for every $A \in \mathcal{A}_{g(2,1)}$, let $F_{1,2} := \{k(A) : A \in \mathcal{A}_{g(2,1)}\}$ and let $\mathcal{F}_{1,2} := \{\overline{V_k} : k \in F_{1,2}\}$. The open set $\cup\mathcal{G}_{1,1} - \cup\mathcal{F}_{1,2}$ is the union of some subset from the basis $\{V_i : i > g(2,1) \wedge i \notin F_{1,2}\}$, so choose a finite subcovering $\mathcal{G}_{1,2}$ that covers ∂V_1 from that subset, indexed by $G_{1,2}$ with maximum index $g(1,2)$. Now denote $H_{1,2} := \{g(2,1)+1, g(2,1)+2, \cdots, g(1,2)\} - (G_{1,2} \cup F_{1,2})$.

Inductively, with $\mathcal{G}_{i,j}, G_{i,j}, g(i,j), F_{i,j}, H_{i,j}$ previously determined when $i + j \leq m$, use the basis $\{V_i : i > g(1, m-1)\}$ to select a finite subcover $\mathcal{G}_{m,1}$ of ∂V_m indexed by $G_{m,1}$ with maximum index $g(m,1)$. Construct $\mathcal{F}_{m-1,2}$, $F_{m-1,2}$, $\mathcal{G}_{m-1,2}$, $G_{m-1,2}$, $g(m-1,2)$ and $H_{m-1,2}$ based on the $(m,1)$ step similar to how $\mathcal{F}_{1,2}$, $F_{1,2}$, $\mathcal{G}_{1,2}$, $G_{1,2}$, $g(1,2)$ and $H_{1,2}$ were constructed based on the $(2,1)$ step. Repeat this again by constructing $\mathcal{F}_{m-2,3}$, $F_{m-2,3}$, $\mathcal{G}_{m-2,3}$, $G_{m-2,3}$, $g(m-2,3)$ and $H_{m-2,3}$ based on the $(m-1,2)$ step. Continue in this manner until $\mathcal{F}_{1,m}, F_{1,m}, \mathcal{G}_{1,m}, G_{1,m}, g(1,m)$ and $H_{1,m}$ are constructed. Now all appropriate numbers and collections have been found for when $i + j = m + 1$.

Complete this process via induction. Now let κ be the finitely-additive set function created in Lemma 2.9 with respect to $(V_{\pi(i)})$. It remains to verify the desired properties. For any ∂V_i, the sequence of covers $\{\mathcal{G}_{i,j}\}_{j=1}^{\infty}$ will satisfy (for all natural j)

$$\kappa(\cup\mathcal{G}_{i,j}) \leq \kappa(\cup\mathcal{G}_{i,1}) \cdot \left(\frac{1}{2}\right)^{j-1}. \tag{2.6}$$

This will be shown via induction. (2.6) is obvious when $j = 1$. Assume that (2.6) is true for some natural j. Then (2.5) from Lemma 2.9 ensures that

$$\kappa(\cup\mathcal{G}_{i,j+1}) \leq \kappa(\cup\mathcal{G}_{i,j} - \cup\mathcal{F}_{i,j+1}) = \frac{1}{2}\kappa(\cup\mathcal{G}_{i,j}) \leq \frac{1}{2} \cdot \kappa(\cup\mathcal{G}_{i,1}) \cdot \left(\frac{1}{2}\right)^{j-2}. \tag{2.7}$$

Therefore (2.6) is true for all natural j, and it follows that $\kappa^*(\partial V_i) \leq \kappa(\cup\mathcal{G}_{i,j}) \leq \kappa(\cup\mathcal{G}_{i,1}) \cdot \left(\frac{1}{2}\right)^{j-1}$ for all natural j, which implies that $\kappa^*(\partial V_i) = 0$. Since $i \in \mathbb{N}$ was arbitrary, it follows that all sets in $\mathcal{C}$ have outer measure zero, showing (a).

Let $m \in \mathbb{N}$. The basis sets $\{V_k : k \in F_{1,m}\}$ fragment each of the sets in $\mathcal{A}_{g(2,m-1)}$, resulting in

$$\max\{\kappa(A) : A \in \mathcal{A}_{g(1,m)}\} \leq \frac{1}{2} \max\{\kappa(A) : A \in \mathcal{A}_{g(2,m-1)}\}. \tag{2.8}$$

Since $\max\{\kappa(A) : A \in \mathcal{A}_n\}$ is a non-increasing function of n, (b) follows. □

For the final stage of proving Theorem 2.7, we construct a measure on $\mathcal{D}$ using the steps from Section 2.1 with Lemmas 2.8, 2.9, and 2.10. Let $(\mathcal{A}'_k)$ and $(\mathcal{D}'_k)$ denote the collections of sets formed in Lemma 2.8 when applied to the permuted sequence $(V_{\pi(i)})$ formed in Lemma 2.10.

Step 1. Construct the set function $\kappa : \mathcal{D} \to [0, 1]$ with respect to $(V_{\pi(i)})$ via Lemma 2.9. Consequently, κ is finitely additive.

It follows from finite additivity that κ is σ-superadditive in $\mathcal{D}$; that is, given pairwise disjoint $(A_i)_{i=1}^{\infty}$ in $\mathcal{D}$, we have that $\kappa(\uplus_{i=1}^{\infty} A_i) \geq \sum_{i=1}^{\infty} \kappa(A_i)$. Now we need to show that κ is σ-subadditive in $\mathcal{D}$. To this end, we consider the outer measure

$$\kappa^*(A) := \inf \left\{ \sum_{i=1}^{\infty} \kappa(A_i) : A \subseteq \cup_{i=1}^{\infty} A_i \text{ and } A_i \in \mathcal{D} \text{ are open for } i \in \mathbb{N} \right\}, \tag{2.9}$$

which is known to be finitely additive and σ-subadditive. By showing that $\kappa = \kappa^*$ on the ring $\mathcal{D}$, κ will be a premeasure on $\mathcal{D}$ capable of being extended to a measure on Borel σ-algebra Σ_X, thus we proceed to Step 2.

Step 2. To show that $\kappa^* \leq \kappa$ on open sets in $\mathcal{D}$, let $U \in \mathcal{D}$ be open. Then $U \subseteq U \cup \varnothing \cup \varnothing \cup \dots$, so we obtain that $\kappa^*(U) \leq \kappa(U) + \kappa(\varnothing) + \kappa(\varnothing) + \dots = \kappa(U)$.

Step 3. The following argument shows that $\kappa \leq \kappa^*$ on compact sets in $\mathcal{D}$. Let $C \in \mathcal{D}$ be compact, and let $\varepsilon > 0$. Choose some sequence $(A_i)_{i=1}^{\infty}$ of open sets in $\mathcal{D}$ with $C \subseteq \cup_{i=1}^{\infty} A_i$ and such that $\sum_{i=1}^{\infty} \kappa(A_i) \leq \kappa^*(C) + \varepsilon$. Then there exists some finite subcover of C, meaning there exists $n \in \mathbb{N}$ with $C \subseteq \cup_{i=1}^{n} A_i$ and there exists some $N \in \mathbb{N}$ with $C, A_1, \cdots, A_n \in \mathcal{D}'_N$. Therefore $\kappa(C) = \kappa_N(C) \leq \sum_{i=1}^{n} \kappa_N(A_i) = \sum_{i=1}^{n} \kappa(A_i) \leq \sum_{i=1}^{\infty} \kappa(A_i) \leq \kappa^*(C) + \varepsilon$. With ε arbitrary, it follows that $\kappa \leq \kappa^*$ on compact sets in $\mathcal{D}$.

Step 4. It has been shown in Lemma 2.10 that $\kappa^*(\cup_{k=1}^{\infty} \partial V_k) = 0$, meaning that κ^* is zero on $\mathcal{C}$.

Step 5. To show that $\kappa^* = \kappa$ on $\mathcal{D}$, let $A \in \mathcal{D}$. Note that $\overline{A} \in \mathcal{D}$ is compact, that $A^\circ \in \mathcal{D}$ is open, and that $\overline{A} - A^\circ \in \mathcal{C}$. It follows that

$$\kappa(A) \leq \kappa\left(\overline{A}\right) \leq \kappa^*\left(\overline{A}\right) = \kappa^*(A) = \kappa^*(A^\circ) \leq \kappa(A^\circ) \leq \kappa(A). \tag{2.10}$$

Therefore, $\kappa^* = \kappa$ is a premeasure on the ring $\mathcal{D}$.

Step 6. Now we can apply Theorem 1.3.6 in [29], meaning that the premeasure κ^* on ring $\mathcal{D}$ extends to a measure on a σ-algebra containing $\sigma(\mathcal{D}) = \Sigma_X$, which can be restricted to a measure $\kappa^\dagger$ on Σ_X. It follows that $\kappa^\dagger$ is a finite strictly-positive measure on (X, Σ_X) since each set in the basis sequence (V_i) was assigned a positive measure by κ on $\mathcal{D}$. Since X is a Polish space, we automatically have that $\kappa^\dagger$ is regular by Theorem 8.1.12 in [29]. To show that $\kappa^\dagger$ is non-atomic, we apply the following lemma.

Lemma 2.11. *Let μ be a finite measure on (X, Σ) such that for every $\varepsilon > 0$, there exists a finite partition of X into measurable sets with each set having μ-measure less than ε. Then μ is non-atomic.*

Proof. Let $C \in \Sigma$ with $\mu(C) > 0$ and define $\varepsilon := \mu(C)$. Let $A_1, \ldots, A_k$ be a partition of X with $\max\{\mu(A_i) : 1 \leq i \leq k\} < \varepsilon$. Then consider the sets $B_0 := C$ and $B_i := B_{i-1} - A_i$

for $1 \leq i \leq k$. Of course, each of these sets are subsets of C. Then for $1 \leq i \leq k$, we have that $\mu(B_i) = \mu(B_{i-1} - A_i) \geq \mu(B_{i-1}) - \mu(A_i) > \mu(B_{i-1}) - \varepsilon$, or equivalently that $0 \leq \mu(B_{i-1}) - \mu(B_i) < \varepsilon$ for each $1 \leq i \leq k$. It is true that $\mu(B_0) = \mu(C) = \varepsilon$ and that $\mu(B_k) = \mu(\varnothing) = 0$, so there must be some intermediate set B_j with $0 < \mu(B_j) < \varepsilon$, hence $B_j \subset C$ with $0 < \mu(B_j) < \mu(C)$. □

Using Lemma 2.11, it suffices to show that for all $\varepsilon > 0$, X can be partitioned into a finite collection of measurable sets, each with $\kappa^\dagger$ measure no more than ε. Let $\varepsilon > 0$, and choose some integer m with $\varepsilon \geq 2^{1-m}$. Then according to Lemma 2.10, all sets from $\mathcal{A}'_{g(1,m)} = \mathcal{A}_{g(1,m)}$ have $\kappa^\dagger$ measure no more than ε each because complete fragmentation by closed holes has occurred at least $m-1$ times. Furthermore, $\kappa^\dagger(\cup_{i=1}^{\infty} \partial V_i) = 0$ and

$$\kappa^\dagger\left(\cap_{i=1}^{g(1,m)} V_i^e\right) \leq 2^{-g(1,m)} \sum_{j=1}^{\infty} 2^{-j} \leq \varepsilon. \tag{2.11}$$

Since $X = (\cup_{i=1}^{\infty} \partial V_i) \cup \left(\cap_{i=1}^{g(1,m)} V_i^e\right) \cup (\cup \mathcal{A}_{g(1,m)})$, $\kappa^\dagger$ is non-atomic.

At last, $\kappa^\dagger$ is a measure on (X, Σ_X) with the sought properties.

CHAPTER 3: LOCALIZED TRANSFUNCTIONS

3.1 Introduction

Let (X, Σ_X) and (Y, Σ_Y) be measurable spaces with sets of finite measures $\mathcal{M}_X$ and $\mathcal{M}_Y$, respectively. A *transfunction* is any function $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$, [32]. One can think of transfunctions as maps where the inputs and outputs are "probability clouds" rather than points. While this intuitive interpretation is useful, we are not restricting the domain and range of a transfunction to probability measures. In fact there are situations where it is natural to consider transfunctions on signed measures or vector measures. While formally a transfunction is a map $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$, we are interested in its properties as a "generalized function" from X to Y. To emphasize this point of view we will use the notation $\Phi : X \rightsquigarrow Y$ when the context is clear.

Every measurable function is a transfunction. More precisely, if (X, Σ_X) and (Y, Σ_Y) are measurable spaces and $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ is a measurable function, then the push forward operator $f_\# : \mathcal{M}_X \to \mathcal{M}_Y$ defined by $f_\#(\mu)(B) = \mu(f^{-1}(B))$ is a transfunction. We will say that the transfunction Φ corresponds to f, or simply Φ is f, if $\Phi = f_\#$.

While every measurable function is a transfunction, we are obviously interested in transfunctions that do not necessarily correspond to measurable functions. In this chapter we investigate the following general questions:

- Under what conditions will a tranfunction Φ be a continuous function?
- Under what conditions will a tranfunction Φ be a measurable function?
- If a tranfunction Φ is not a function, under what conditions is Φ "close" to a measurable or continuous function?

We also introduce the notion of a graph of a transfunction, which gives us additional intuition about the nature of transfunctions. The main tool in our investigation is the idea of localization of transfunctions which is introduced in Section 3.3.

3.2 Preliminaries

Unless otherwise specified, all instantiated measures shall be finite and positive and $\mathcal{M}_X$ and $\mathcal{M}_Y$ will contain finite and positive measures. Occasionally, we may sum countably many measures together. When this occurs, the sum may be finite or infinite and we will not determine the finiteness of the measure whenever it is inconsequential to the argument at hand.

If μ is a positive or a vector measure on (X, Σ_X) and $A \in \Sigma_X$, then we say that A is a *carrier* of μ and write $\mu \sqsubset A$ if $|\mu|(A^c) = 0$, where $|\mu|$ denotes the variation measure of μ. If μ is a positive measure, then $\mu \sqsubset A$ is also equivalent to the simpler condition that $\mu(A^c) = 0$.

If $A \subseteq X$, $B \subseteq Y$, and $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ is a transfunction such that $\mu \sqsubset A$ implies $\Phi\mu \sqsubset B$ for every μ, we shall write $\Phi(A) \sqsubset B$.

Let μ be a measure on measurable space (X, Σ_X) and let $A \in \Sigma_X$. Then the *projection* of μ onto A, denoted as $\pi_A\mu$, is the measure defined as $\pi_A\mu(B) = \mu(B \cap A)$ for $B \in \Sigma_X$. If $\mathcal{M}_X$ is a space of measures on (X, Σ_X), then we say that $\mathcal{M}_X$ is *closed under projections* if $\mu \in \mathcal{M}_X$ implies that $\pi_A\mu \in \mathcal{M}_X$ for all $A \in \Sigma_X$.

If μ, ν are measures in $\mathcal{M}_X$, then they are called *orthogonal*, written as $\mu \perp \nu$, if there exists $A \in \Sigma_X$ such that $\mu \sqsubset A$ and $\nu \sqsubset A^c$. A countable sequence of measures $\{\mu_n\}_{n=1}^{\infty}$ is called *(pairwise) orthogonal* if $\mu_i \perp \mu_j$ for $i \neq j$.

If a sequence of measures $(\mu_i)_{i=1}^{\infty}$ satisfies $\sum_{i=1}^{\infty} ||\mu_i|| < \infty$, then we call the finite measure $\mu =$

$\sum_{i=1}^{\infty} \mu_i$ the *bounded sum of* $(\mu_i)_{i=1}^{\infty}$. A bounded sum $\mu = \sum_{i=1}^{\infty} \mu_i$ with $\{\mu_i\}_{i=1}^{\infty}$ being orthogonal will be called a *bounded orthogonal sum*.

Definition 3.1. Let (X, τ) be a topological space. A space of finite positive measures or of vector measures $\mathcal{M}_X$ on (X, Σ_X) is called *ample* if the following conditions hold:

(i) $\mathcal{M}_X$ is closed under projections;

(ii) $\mathcal{M}_X$ is closed under bounded orthogonal sums;

(iii) Every nonempty open set in X carries some nonzero measure in $\mathcal{M}_X$.

If λ is a finite strictly-positive measure on a topological measurable space (X, Σ_X), then it follows that $\{\pi_A \lambda : A \in \Sigma_X\}$ is an ample space of finite measures. Certain spaces (e.g. 2nd-countable non-atomic locally compact Hausdorff spaces from Chapter 2, compact groups) admit finite strictly-positive measures, hence they also admit ample spaces of measures.

Ample spaces will be useful for transfunctions because we will decompose a measure into bounded orthogonal sums of projections and use local properties to determine the behavior of each projection. If the transfunctions have certain properties (e.g. strong or weak σ-additivity as in Definition 3.2), then the collective behavior of the projections will imply some behavior of the overall input measure.

In this chapter we will assume that X and Y are second-countable topological spaces, that Σ_X and Σ_Y are collections of Borel subsets of X and Y, respectively, and that any transfunction $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ will be defined on an ample space $\mathcal{M}_X$ unless otherwise specified.

Definition 3.2. Let $\Phi : X \rightsquigarrow Y$ be a transfunction with $\mathcal{M}_X$ closed under bounded orthogonal sums.

(i) Φ is called *weakly monotone* if $\Phi\mu \leq \Phi(\mu+\nu)$ for each orthogonal pair of positive measures μ and ν.

(ii) Φ is called *weakly σ-additive* if $\Phi(\sum_{i=1}^{\infty} \mu_i) = \sum_{i=1}^{\infty} \Phi\mu_i$ for every bounded orthogonal sum $\sum_{i=1}^{\infty} \mu_i$ in $\mathcal{M}_X$.

(iii) Φ is called *strongly σ-additive* if $\Phi(\sum_{i=1}^{\infty} \mu_i) = \sum_{i=1}^{\infty} \Phi\mu_i$ for every bounded sum $\sum_{i=1}^{\infty} \mu_i$ in $\mathcal{M}_X$.

Notice that strong σ-additivity implies weak σ-additivity, and weak σ-additivity implies weak monotonicity when Φ maps positive measures to positive measures. For any measurable function $f : X \to Y$, the transfunction $f_\# : X \rightsquigarrow Y$ with $\mathcal{M}_X$ closed under bounded orthogonal sums is strongly σ-additive.

The properties of weakly σ-additive transfunctions listed in Proposition 3.3 are straightforward to verify and are often used in arguments later in the chapter.

Proposition 3.3. *Let $\Phi : X \rightsquigarrow Y$ be a weakly σ-additive transfunction. Let A, A', and the sequence $(A_i)_{i=1}^{\infty}$ be from Σ_X and let B, B', and the sequence $(B_j)_{j=1}^{\infty}$ be from Σ_Y.*

(i) *If $\Phi(A) \sqsubset B$, if $A' \subseteq A$ and if $B' \supseteq B$, then $\Phi(A') \sqsubset B'$;*

(ii) *If $\Phi(A_i) \sqsubset B_j$ for all $i, j \in \mathbb{N}$, then $\Phi(\cup_{i=1}^{\infty} A_i) \sqsubset \cap_{j=1}^{\infty} B_j$.*

(iii) *If $\Phi(A_i) \sqsubset B_i$ for all $i \in \mathbb{N}$, then $\Phi(\cap_{i=1}^{\infty} A_i) \sqsubset \cap_{j=1}^{\infty} B_j$ and $\Phi(\cup_{i=1}^{\infty} A_i) \sqsubset \cup_{j=1}^{\infty} B_j$.*

Proposition 3.4 follows easily from Proposition 3.3 and will be useful in the characterization of transfunctions that correspond to continuous functions.

Proposition 3.4. *Let X and Y be topological spaces with X second-countable, and let $\Phi : X \rightsquigarrow Y$ be a weakly σ-additive transfunction. Let U be open in X with open cover $\{S_i : i \in I\}$, and let B*

be measurable in Y. Then $\Phi(S_i) \sqsubset B$ for all $i \in I$ implies that $\Phi(U) \sqsubset B$. In particular, if μ is a measure on X and if $\Phi(\pi_{S_i}\mu) \sqsubset B$ for all $i \in I$, then $\Phi(\pi_U\mu) \sqsubset B$.

A transfunction $\Phi : X \rightsquigarrow Y$ is said to *vanish on an open set* U if $\Phi(U) \sqsubset \varnothing$. Let $\mathcal{V}_\Phi$ denote the collection of all vanishing open sets of Φ. Let $\Phi : X \rightsquigarrow Y$ be a weakly σ-additive transfunction. Then $\cup\mathcal{V}_\Phi$ is called the *null space of* Φ, denoted as $\operatorname{null}\Phi$. Its complement, $(\cup\mathcal{V}_\Phi)^c$, is called the *spatial support of* Φ, denoted as $\operatorname{supp}\Phi$. Note that $\Phi\mu = \Phi\left(\pi_{\operatorname{supp}\Phi}\mu\right)$, which implies that Φ is essentially a transfunction between the subspace $\operatorname{supp}\Phi$ and Y. A transfunction $\Phi : X \rightsquigarrow Y$ is called *non-vanishing* if Φ has no non-empty vanishing open sets, that is, if $\operatorname{supp}\Phi = X$. Furthermore, Φ is *norm-preserving* if $||\Phi\mu|| = ||\mu||$ for all μ on X.

3.3 Localized Transfunctions

In this section and in the following two sections we assume that X and Y are 2nd-countable metric spaces. We use $B(z;\rho)$ to denote the open ball of radius ρ centered at z.

Definition 3.5. Let $x \in X$ and $\varepsilon > 0$. We say that a transfunction $\Phi : X \rightsquigarrow Y$ is *ε-localized at x* if there exist $\delta > 0$ and $y \in Y$ such that $\Phi(B(x,\delta)) \sqsubset B(y,\varepsilon)$. We say that Φ is *0-localized at x* if Φ is ε-localized at x for all $\varepsilon > 0$. If Φ is ε-localized at x for some $\varepsilon > 0$, then we say that Φ is *localized at x*. If Φ is localized at every point in some set $A \in \Sigma_X$, then we say that Φ is *localized on A*. If y needs emphasis, we say that Φ is *ε-localized at (x,y)*. If we need to emphasize δ, we say that Φ is *(δ,ε)-localized at (x,y)*.

Note in the definition of 0-localization that the values for δ and y may depend on ε and x.

Definition 3.6. Let $A \subseteq X$. We say that that a transfunction $\Phi : X \rightsquigarrow Y$ is *uniformly localized on A* if there exist $\varepsilon > 0$ and $\delta > 0$ such that Φ is (δ,ε)-localized on A. If δ and ε are to be emphasized, then we say that Φ is *uniformly* (δ, ε)*-localized on* A. If only ε is to be emphasized, then we say that Φ is *uniformly* ε*-localized on* A.

Definition 3.7. For a transfunction $\Phi : X \rightsquigarrow Y$ we define the function $E_\Phi : X \to [0, \infty]$ via $E_\Phi(x) = \inf\{\varepsilon : \Phi \text{ is } \varepsilon\text{-localized at } x\}$, which describes the extent of localization of Φ throughout X.

Note that Φ is $(E_\Phi(x) + \eta)$-localized at x for all $\eta > 0$ whenever $E_\Phi(x) < \infty$ and that Φ is not localized at x when $E_\Phi(x) = \infty$.

Definition 3.8. Let $A \subseteq X$ and let $f : X \to Y$ be function. We say that Φ is ε*-localized on A via* f or that Φ is ε*-close to* f *on* A if Φ is ε-localized at $(x, f(x))$ for all $x \in A$.

It is worth noting that transfunctions are not necessarily localized anywhere. When verifying whether a transfunction is localized, Proposition 3.9 is often useful.

Proposition 3.9. *Let* f *be a measurable function, let* $\mu \in \mathcal{M}_X$ *be a positive measure or a vector measure and let* $A \in \Sigma_X$*. Then* $|f_\# \mu| \leq f_\# |\mu|$*. Additionally, if* μ *is a positive measure, then* $\mu \sqsubset f^{-1}(A)$ *if and only if* $f_\# \mu \sqsubset A$*, and if* μ *is a vector measure, then* $\mu \sqsubset f^{-1}(A)$ *implies* $f_\# \mu \sqsubset A$*.*

Proof. To show the first claim, notice that for all $B \in \Sigma_X$,

$$\begin{aligned} |f_\# \mu|(B) &= \sup \left\{ \textstyle\sum_{i=1}^n |f_\# \mu(B_i)| : \biguplus_{i=1}^n B_i = B \right\} \\ &= \sup \left\{ \textstyle\sum_{i=1}^n |\mu(f^{-1}(B_i))| : \biguplus_{i=1}^n f^{-1}(B_i) = f^{-1}(B) \right\} \\ &\leq |\mu|(f^{-1}(B)) = f_\# |\mu|(B). \end{aligned} \tag{3.1}$$

If μ is positive, then

$$\mu \sqsubset f^{-1}(A) \Leftrightarrow 0 = \mu((f^{-1}(A))^c) = \mu \circ f^{-1}(A^c) \Leftrightarrow f_\# \mu \sqsubset A. \tag{3.2}$$

If μ is a vector measure, then

$$\mu \sqsubset f^{-1}(A) \Leftrightarrow 0 = |\mu|((f^{-1}(A))^c) = |\mu| \circ f^{-1}(A^c) \Leftrightarrow f_\# |\mu| \sqsubset A \Rightarrow f_\# \mu \sqsubset A, \tag{3.3}$$

where the final implication follows from $|f_\# \mu| \leq f_\# |\mu|$. $\square$

Now we consider some examples.

Example 3.10. If $f : X \to Y$ is a continuous function, then for $\Phi = f_\#$ we have $E_\Phi = 0$.

Example 3.11. Let $H : \mathbb{R} \to \mathbb{R}$ be the Heaviside function centered at 0 and let $\Phi = H_\#$. Since H is continuous everywhere except at 0, it follows that $E_\Phi(x) = 0$ for all $x \neq 0$. However, $E_\Phi(0) = 1/2$.

Example 3.12. Consider the measurable function $g : \mathbb{R} \to \mathbb{R}$ via $g = \sum_{n=0}^{\infty} 2^n H_n$, where H_n is the Heaviside function centered at n, and define $\Phi = g_\#$. Then $E_\Phi(n) = 2^{n-1}$ for each $n \in \mathbb{N}$ and $E_\Phi(x) = 0$ for $x \in \mathbb{R} \setminus \mathbb{N}$, meaning that Φ is localized on $\mathbb{R}$ but that $\sup_{x \in \mathbb{R}} E_\Phi(x) = \infty$.

Example 3.13. Let $A \in \Sigma_X$. Then the projection transfunction $\Phi = \pi_A$ is 0-localized via the identity function since every carrier of μ is also a carrier of $\pi_A \mu$.

Example 3.14. Let $Y = \mathbb{R}$, $\Sigma_Y = \mathcal{B}(\mathbb{R})$, and let ν be a strictly positive finite measure on $\mathbb{R}$. The transfunction $\Phi : \mathbb{R} \rightsquigarrow \mathbb{R}$ defined via $\Phi(\mu) = \|\mu\| \nu$ is not localized anywhere.

Example 3.15. Let $X = Y = \mathbb{R}^d$ and let λ^d be the Lebesgue measure on $\mathbb{R}^d$. For some $\varepsilon > 0$, define $\kappa = \pi_{B(0;\varepsilon)} \lambda^d$. The transfunction $\Phi : \mathbb{R}^d \rightsquigarrow \mathbb{R}^d$ defined via $\Phi(\mu) = \mu * \kappa$, the convolution of measures μ and κ, is ε-localized on X via the identity.

If U denotes the set of points in X where a transfunction Φ is localized, then the function $E_\Phi|_U : U \to [0, \infty)$ does not have to be continuous. However, as Proposition 3.16 states, it does have to be upper-semi-continuous on U, implying that U is an open set.

Proposition 3.16. *Let $\Phi : X \rightsquigarrow Y$ be a transfunction and let U denotes the set of points in X where Φ is localized. Then the function $E_\Phi|_U$ is an upper-semi continuous function and U is an open set.*

Proof. Let $x \in U$ with $E_\Phi(x) = \eta$ and let $\varepsilon > \eta$. Then there exists a $y \in Y$ and a $\delta > 0$ such that $\Phi(B(x;\delta)) \sqsubset B(y;\varepsilon)$. Choose an $x_1 \in B(x;\delta)$ different from x. Then there exists a $\delta_1 > 0$ such that $B(x_1;\delta_1) \subset B(x;\delta)$, so that $\Phi(B(x_1;\delta_1)) \sqsubset B(y;\varepsilon)$. Therefore, Φ is ε-localized at (x_1, y), yielding that $E_\Phi(x_1) \leq \varepsilon$ and that $x_1 \in U$. Since $x_1 \in B(x;\delta)$ was arbitrary, this means that $\sup E_\Phi(B(x;\delta)) \leq \varepsilon$ and that $B(x;\delta) \subseteq U$. Since ε was arbitrary, this means that $\limsup_{\delta\to 0} E_\Phi(B(x;\delta)) \leq \eta = E_\Phi(x)$, meaning that E_Φ is upper-semi-continuous. Since x was arbitrary, this means that U is open. □

3.4 0-Localized Transfunctions

When $f : X \to Y$ is continuous, we know that $f_\#$ is weakly σ-additive, norm-preserving, and 0-localized on X. We will show that these three properties characterize transfunctions that correspond to continuous functions.

Proposition 3.17. *Let X be a metric space with an ample family of measures $\mathcal{M}_X$ and let Y be a complete metric space. For any $A \in \Sigma_X$ and for any non-vanishing transfunction $\Phi : X \rightsquigarrow Y$ which is 0-localized on A there is a unique continuous function $f : A \to Y$ such that Φ is 0-close to f on A.*

Proof. Since Φ is 0-localized on A, it follows that $E_\Phi(x) = 0$ for all $x \in A$. Then for any fixed $x \in A$, there are $\delta_n > 0$ and $y_n \in Y$ indexed by $n \in \mathbb{N}$ such that $\Phi(B(x; \delta_n)) \sqsubset B\left(y_n; \frac{1}{n}\right)$ for every $n \in \mathbb{N}$.

First we show that $d(y_m, y_n) \leq \frac{1}{m} + \frac{1}{n}$ for all $m, n \in \mathbb{N}$. Suppose $d(y_m, y_n) > \frac{1}{m} + \frac{1}{n}$ for some $m, n \in \mathbb{N}$. Then $B\left(y_m; \frac{1}{m}\right) \cap B\left(y_n; \frac{1}{n}\right) = \varnothing$. This implies that $\Phi(B(x; \min\{\delta_m, \delta_n\}) \sqsubset \varnothing$, which is impossible since $\mathcal{M}_X$ is ample and Φ is non-vanishing.

Since $d(y_m, y_n) \leq \frac{1}{m} + \frac{1}{n}$ for all $m, n \in \mathbb{N}$, (y_n) is a Cauchy sequence in the complete metric space Y. So there exists $y \in Y$ with $y_n \to y$. Furthermore, $B\left(y_n; \frac{1}{n}\right) \subseteq B\left(y; \frac{1}{n} + d(y_n, y)\right)$ for $n \in \mathbb{N}$ with $\frac{1}{n} + d(y_n, y) \to 0$ as $n \to \infty$. Indeed, y is the unique point in Y with this property and Φ is 0-localized at (x, y). Now we define $f : A \to Y$ by $f(x) = y$, where $y \in Y$ is the unique point such that Φ is 0-localized at (x, y). Clearly Φ is 0-localized on A via f.

We now show that f is continuous on A. Let $x_n \to x_0$ in A. Define $y_0 = f(x_0)$ and $y_n = f(x_n)$ for $n \in \mathbb{N}$. Let $\varepsilon > 0$. Then there is a $\delta > 0$ such that $\Phi(B(x_0; \delta)) \sqsubset B(y_0; \varepsilon)$. Let $N \in \mathbb{N}$ be such that $d(x_m, x_0) < \delta/2$ for $m \geq N$. For every $m \geq N$ there is a $\delta_m < \delta/2$ such that $\Phi(B(x_m; \delta_m)) \sqsubset B(y_m; \varepsilon)$. Then $B(x_m; \delta_m) \subset B(x_0; \delta)$ implies that $\Phi(B(x_m; \delta_m)) \sqsubset B(y_0; \varepsilon) \cap B(y_m; \varepsilon)$. Consequently, $d(y_m, y_0) \leq 2\varepsilon$, since $\mathcal{M}_X$ is ample and Φ is non-vanishing. Since ε was arbitrary, we have that $y_m \to y_0$ as $m \to \infty$. □

Theorem 3.18. *Let X be a second-countable metric space with an ample family of measures $\mathcal{M}_X$, Y a complete metric space, and $\Phi : X \rightsquigarrow Y$ a non-vanishing transfunction. Then $\Phi = f_\#$ for some continuous function $f : X \to Y$ if and only if Φ is norm-preserving, weakly σ-additive, and 0-localized on X.*

Proof. We only need to show that if Φ is norm-preserving, weakly σ-additive, and 0-localized on X, then $\Phi = f_\#$ for some continuous function $f : X \to Y$. Let $f : X \to Y$ be the unique continuous function guaranteed by Proposition 3.17. We will show that $\Phi = f_{\#}$.

Let $V \subseteq Y$ be an open set. Define $U = f^{-1}(V)$. By 0-localization of Φ via f, there exists an open ball cover $\{B(x;\delta_x) : x \in U\}$ of U such that $\Phi(B(x;\delta_x)) \sqsubset V$ for all $x \in X$. By Proposition 3.4, we have that $\Phi(U) \sqsubset V$. Similarly, if we define $W = f^{-1}\left(\left(\overline{V}\right)^c\right)$, which is also open, we have that $\Phi(W) \sqsubset \left(\overline{V}\right)^c$.

Next, we define $Z = f^{-1}(\partial V) = f^{-1}(\overline{V} \cap V^c)$, which is closed, and for each $n \in \mathbb{N}$, let $L_n = \cup_{y \in \partial V} B(y; 1/n)$, which is an open set. Since Φ is 0-localized, for each $x \in Z$ and $n \in \mathbb{N}$, there exists $\delta_{x,n} > 0$ such that $\Phi(B(x;\delta_{x,n})) \sqsubset L_n$. For $K_n = \cup_{x \in Z} B(x;\delta_{x,n})$ we have $\Phi(K_n) \sqsubset L_n$ for all $n \in \mathbb{N}$, by Proposition 3.4. Noting that $\cap_{n=1}^{\infty} K_n = Z$ and $\cap_{n=1}^{\infty} L_n = \partial V$, it is clear that $\Phi(Z) \sqsubset \partial V$, by Proposition 3.3.

Let $\mu \in \mathcal{M}_X$. Since $\pi_U \mu \sqsubset U$ and $\Phi(U) \sqsubset V$, we have that $\Phi(\pi_U \mu) \sqsubset V$ and we have via norm-preservation of Φ that

$$\Phi(\pi_U \mu)(V) = ||\Phi(\pi_U \mu)|| = ||\pi_U \mu|| = \mu(U) = f_{\#}(\mu)(V). \tag{3.4}$$

Since $\Phi(W) \sqsubset \left(\overline{V}\right)^c$ and $V \cap \left(\overline{V}\right)^c = \varnothing$, it follows that $\Phi(\pi_W \mu)(V) = 0$. Similarly, since $\Phi(Z) \sqsubset \partial V$ and $V \cap \partial V = \varnothing$, it follows that $\Phi(\pi_Z \mu)(V) = 0$. Overall, we obtain

$$\Phi(\mu)(V) = \Phi(\pi_U \mu)(V) + \Phi(\pi_W \mu)(V) + \Phi(\pi_Z \mu)(V) = f_{\#}(\mu)(V). \tag{3.5}$$

Moreover, since $\Phi(\mu)$ and $f_{\#}(\mu)$ are finite measures which agree on open sets, they must agree on all sets in Σ_Y by an application of the π-λ theorem found in Corollary 1.6.3 from [29]. Finally, since $\mu \in \mathcal{M}_X$ is arbitrary, we have $\Phi = f_{\#}$. $\square$

Now we characterize transfunctions which correspond to measurable functions, but under stricter settings. First, we define restrictions of transfunctions.

Definition 3.19. Let $\Phi : X \rightsquigarrow Y$ be a transfunction, and let $A \subseteq X$ be measurable. Then the composition $\Phi \circ \pi_A$ is called the *restriction* of Φ to A.

Note that $\Phi \circ \pi_A = \Phi$ when $\operatorname{supp} \Phi \subseteq A$ and that $\Phi \circ \pi_B = 0$ when $B \subseteq \operatorname{null} \Phi$.

Theorem 3.20. *Let X be locally compact, let λ be a finite regular measure on X, and let $\mathcal{M}_X$ contain exactly the finite positive measures absolutely continuous with respect to λ. Let $\Phi : X \rightsquigarrow Y$ be a weakly σ-additive transfunction. Then Φ corresponds to a measurable function if and only if there exists a sequence of compact sets $\{F_n\}_{n=1}^{\infty}$ such that $\lambda(F_n^c) < \frac{1}{n}$ and that $\Phi \circ \pi_{F_n}$ is identified with some continuous function on F_n for all natural n.*

Proof. The forward direction is a straight-forward consequence of Lusin's theorem, where the measurable and continuous functions are identified with the respective transfunctions; see Theorem 7.4.4 in [29].

We now prove the reverse direction. For each natural n, let $\Phi \circ \pi_{F_n}$ be identified with continuous function $f_n : F_n \to Y$. Let $i \neq j$. If $\lambda(F_i \cap F_j) > 0$, then from Lemma 7.5.2 from [29] there exists some compact subset $G_{i,j} \subseteq F_i \cap F_j$ such that $\lambda(G_{i,j}) = \lambda(F_i \cap F_j)$ and $\lambda(U \cap G_{i,j}) > 0$ whenever $U \cap G_{i,j} \neq \varnothing$ for open U. Otherwise, if $\lambda(F_i \cap F_j) = 0$, then define $G_{i,j} = \varnothing$.

For the latter case, $f_i = f_j$ is vacuously true on $G_{i,j}$. For the former case, let $x \in G_{i,j}$. Suppose that $f_i(x) \neq f_j(x)$. If we let $\varepsilon < d(f_i(x), f_j(x))/2$, this would imply by 0-localization of $\Phi \circ \pi_{F_i}$ and $\Phi \circ \pi_{F_j}$ the existence of $\delta > 0$ such that $\Phi(B(x;\delta) \cap G_{i,j}) \sqsubset B(f_i(x);\varepsilon) \cap B(f_j(x);\varepsilon) = \varnothing$. Choosing μ_0 to be the projection of λ onto $B(x;\delta) \cap G_{i,j}$, we observe that $\mu_0 \neq 0$ and that $\Phi(\mu_0) = \Phi \circ \pi_{F_i}(\mu_0) = 0$, which contradicts the norm-preservation of $\Phi \circ \pi_{F_i}$ on F_i. It follows that $f_i = f_j$ on $G_{i,j} \subseteq F_i \cap F_j$. Having i, j arbitrary, we have that outside the λ-null Borel set

$N = (\cup_{i,j=1}^{\infty} (F_i \cap F_j - G_{i,j})) \cup (\cap_{i=1}^{\infty} F_i^c)$, the functions $(f_i)_{i=1}^{\infty}$ coincide, allowing them to be glued to a measurable function $h : X \to Y$, where $h(N) = \{y_0\}$ for some fixed $y_0 \in Y$.

We now show that $\Phi = h_{\#}$. Let $\mu \in \mathcal{M}_X$ and let $A_n = N^c \cap (F_n - \cup_{m<n} F_m)$. Since $\lambda(N) = 0$ and $\mu \ll \lambda$, $\mu(N) = 0$. This means that

$$\begin{aligned} \Phi(\mu)(B) &= \Phi\left(\pi_N \mu + \textstyle\sum_{n=1}^{\infty} \pi_{A_n} \mu\right)(B) = \textstyle\sum_{n=1}^{\infty} \Phi(\pi_{A_n} \mu)(B) \\ &= \textstyle\sum_{n=1}^{\infty} f_{n\#}(\pi_{A_n} \mu)(B) = \textstyle\sum_{n=1}^{\infty} \mu\left(A_n \cap h^{-1}(B)\right) \\ &= \mu\left(N \cap h^{-1}(B)\right) + \mu\left(N^c \cap h^{-1}(B)\right) = h_{\#}(\mu)(B). \end{aligned} \tag{3.6}$$

Therefore $\Phi = h_{\#}$. □

In summary, we have characterized continuous and measurable push-forward operators.

3.5 ε-Localized Transfunctions

When Φ is not indentifiable with a measurable function, under what condition is it "close" to a measurable function? We consider this question for uniformly localized transfunctions. Given that $\Phi : X \rightsquigarrow Y$ is uniformly ε-localized, can we find a measurable function $f : X \to Y$ such that Φ is uniformly ε-close to f? If we can find such a function, then it gives a rough idea of how the transfunction behaves. In our settings, we can always find such a measurable function: in fact, it can be chosen so that f is σ-simple. Can we choose a continuous f in this way? The answer is also affirmative, but it requires a more demanding setting.

Proposition 3.21. *Let X and Y be metric spaces, with X second-countable. Then every transfunction Φ which is uniformly ε-localized on X is uniformly ε-close to some measurable function $f : X \to Y$.*

Proof. Let $\Phi : X \rightsquigarrow Y$ be a uniformly (δ, ε)-localized transfunction on X. This means that for all $x \in X$, there exists some $y_x \in Y$ with $\Phi(B(x;\delta)) \sqsubset B(y_x;\varepsilon)$. This choice function $y : x \mapsto y_x$ will be used later. Note that the collection $\{B(x;\delta/3) : x \in X\}$ is an open cover of X. It follows from second-countability of X that there is a countable subcover, which shall be indexed as $\{B(x_n;\delta/3) : n \in \mathbb{N}\}$. For each natural n, let $y_n := y_{x_n}$. Next we create a function $f : X \to Y$ given by $f(x) = y_n$ whenever $x \in B(x_n;\delta/3) - \cup_{m<n} B(x_m;\delta/3)$. It follows that f is a σ-simple function, and therefore is measurable. Furthermore, when $f(x) = y_n$, it follows that $x \in B(x;\delta/3) \subseteq B(x_n;\delta)$.

Therefore, it follows that $\Phi(B(x;\delta/3)) \sqsubset B(y_n;\varepsilon) = B(f(x);\varepsilon)$, which shows that Φ is uniformly $(\delta/3, \varepsilon)$-localized on X via f. □

We build upon the proof of Proposition 3.21 to develop the next theorem. First, we define left-translation-invariance of a metric on locally compact groups.

Definition 3.22. Let X be a locally compact group with identity e, and let d be a metric on X. Then d is *left-translation-invariant* if $d(x, y) = d(zx, zy)$ for all $x, y, z \in X$. When the metric is understood by context, the equivalent definition is that $xB(e;\varepsilon) = B(x;\varepsilon)$ for all $x \in X$ and $\varepsilon > 0$.

Theorem 3.23. *Let X be a second-countable metrizable locally compact group with left-translation-invariant metric and let Y be a normed space. If $\Phi : X \rightsquigarrow Y$ is a uniformly ε-localized transfunction on X, then Φ is uniformly ε-close to some continuous function $g : X \to Y$.*

Proof. Let e denote the identity of X and let 0_Y denote the zero in Y. Take from the proof of Proposition 3.21 the measurable $f : X \to Y$. Then there exists $\alpha > 0$ such that for all $x \in X$, $B(x;\alpha)$ has compact closure. Let $x \in X$ be arbitrary. Since $\{B(x_n;\delta/3) : n \in \mathbb{N}\}$ covers $\overline{B(x;\alpha)}$, it follows that there is a finite subcover $\{B(x_n;\delta/3) : n \leq N_x\}$ for some natural number

N_x depending on x. Therefore, $f(B(x;\alpha)) \subseteq \{y_n : n \leq N_x\} \subseteq B(0_Y; M_x)$ for some real M_x depending on x. Since x was arbitrary, this means that f is locally bounded.

Now let $\beta = \min\{\delta/3, \alpha/2\}$. Since X is a locally compact group, there exists a non-zero (uniformly) continuous function $\varphi : X \to [0, \infty)$ with compact support within $B(e;\beta)$. Now choose the unique appropriately scaled left Haar measure κ on X such that $\int \varphi(u^{-1}) d\kappa(u) = 1$.

Now consider the function $g : X \to Y$ given by $g = f * \varphi$, the convolution of $f : X \to Y$ and $\varphi : X \to \mathbb{R}$ using the (vector-valued) integral

$$g(x) = f * \varphi(x) = \int f(t)\varphi(t^{-1}x) d\kappa(t) = \int f(xu)\varphi(u^{-1}) d\kappa(u). \tag{3.7}$$

Note that (3.7) is well-defined, because $t \mapsto \varphi(t^{-1}x)$ is zero outside of $xB(e;\beta) = B(x;\beta)$ and f is bounded and finitely-valued on the set $B(x;\beta)$. Also, the last equality in (3.7) holds due to left-invariance of κ and the substitution $u = x^{-1}t$ which yields $xu = t$ and $u^{-1} = t^{-1}x$.

We shall now show that g is continuous. Let $x \in X$, let $\varepsilon > 0$ and choose some $\eta \in (0, \beta)$ with respect to uniform continuity of φ. Let x' be η-close to x in X: that is, let $x^{-1}x' \in B(e;\eta)$. This implies that $(t^{-1}x)^{-1}(t^{-1}x') = x^{-1}x' \in B(e;\eta)$ for all $t \in X$, so that $t^{-1}x$ and $t^{-1}x'$ are also η-close in X for all $t \in X$. Since $d(x, x') < \alpha/2$, it follows that $B(x';\alpha/2) \subseteq B(x;\alpha)$, which means that $f(B(x';\alpha/2)) \subseteq f(B(x;\alpha)) \subseteq B(0_Y; M_x)$. Therefore M_x bounds the vectors obtained by f in both $B(x;\beta)$ and $B(x';\beta)$. Then it follows that $|\varphi(t^{-1}x) - \varphi(t^{-1}x')| < \varepsilon$ for all $t \in X$ and that

$$\begin{aligned} ||g(x) - g(x')|| &= \left|\left| \int f(t)[\varphi(t^{-1}x) - \varphi(t^{-1}x')] d\kappa(t) \right|\right| \\ &\leq 2\, M_x\, \varepsilon\, \kappa(B(e;\beta)). \end{aligned} \tag{3.8}$$

Continuity of g follows since M_x only depends on x, $\kappa(B(e;\beta))$ is a constant, and ε was arbitrary.

To show that Φ is uniformly (β, ε)-localized via g, let $x \in X$ be arbitrary and let $\mu \sqsubset B(x; \beta)$. Recall that $B(x; \beta)$ is covered by $\cup_{m=1}^{N_x} B(x_m; \delta/3)$. Notice that for every x_m with $B(x_m; \delta/3) \cap B(x; \delta/3) \neq \varnothing$ we have that $B(x; \delta/3) \subseteq B(x_m; \delta)$ which implies that $\Phi(B(x; \delta/3)) \sqsubset B(y_m; \varepsilon)$.

If we denote $R = \{y_m : m \leq N_x$ and $B(x_m; \delta/3) \cap B(x; \delta/3) \neq \varnothing\}$, and if we denote $C = \mathrm{Conv}(R)$, the convex hull of R, this implies that

$$\Phi\mu \sqsubset \bigcap_{y \in R} B(y; \varepsilon) = \bigcap_{y \in C} B(y; \varepsilon). \tag{3.9}$$

If we can show that $g(x) \in C$, then it follows that Φ is (β, ε)-localized at $(x, g(x))$.

For each natural $m \leq N_x$, we define $A_m = B(e; \beta) \cap x^{-1} f^{-1}(y_m)$ which is empty if $y_m \notin R$ and we define $c_m = \displaystyle\int_{A_m} \varphi(u^{-1}) d\kappa(u)$ which is zero if $y_m \notin R$. Then $\sum_{m=1}^{N_x} c_m = \int \varphi(u^{-1}) d\kappa(u) = 1$, and by looking at the convolution function g, we see that $g(x)$ equals

$$\begin{aligned} &\int_{B(e;\beta)} f(xu)\varphi(u^{-1}) d\kappa(u) = \int_{B(e;\beta)} \left[\sum_{m=1}^{N_x} y_m \chi_{A_m}(u) \right] \varphi(u^{-1}) d\kappa(u) \\ &= \sum_{m=1}^{N_x} y_m \int_{A_m} \varphi(u^{-1}) d\kappa(u) = \sum_{m=1}^{N_x} c_m y_m \in C. \end{aligned} \tag{3.10}$$

Therefore, it follows that $\Phi\mu \sqsubset B(g(x); \varepsilon)$, meaning that Φ is uniformly ε-close to g. □

Corollary 3.24. *Give $\mathbb{R}^n$ and $\mathbb{R}^m$ the usual norms. Let $\Phi : \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be a uniformly ε-localized transfunction. Then Φ is uniformly ε-close to some continuous function $g : \mathbb{R}^n \to \mathbb{R}^m$.*

For transfunctions not uniformly localized, there is a result analogous to Proposition 3.21 with appropriate modifications of its proof.

Proposition 3.25. *Let X and Y be metric spaces with X second-countable. Then every transfunction $\Phi : X \rightsquigarrow Y$ which is ε-localized on X is ε-close to some measurable function $f : X \to Y$.*

Proof. We use the same framework as the proof from Proposition 3.21. Each $x \in X$ has some associated $\delta_x > 0$ from definition of ε-localization at x. We form the cover $\{B(x; \delta_x/3) : x \in X\}$ of X which has a countable subcover $\{B(x_n; \delta_n) : n \in \mathbb{N}\}$, where $\delta_n = \delta_{x_n}$. We define $f(x) = y_n$ when $x \in B(x_n; \delta_n/3) - \cup_{m<n} B(x_m; \delta_m/3)$. For x with $f(x) = y_n$, we have that $x \in B(x; \delta_n/3) \subseteq B(x_n; \delta_n)$. This means for all x with $f(x) = y_n$, we have that $\Phi(B(x; \delta_n/3)) \sqsubset B(y_n; \varepsilon) = B(f(x); \varepsilon)$, meaning that Φ is ε-localized on X via f. $\square$

Alternatively, we develop a proposition analogous to the statement that continuous functions on compact sets are uniformly continuous.

Proposition 3.26. *Let $\Phi : X \rightsquigarrow Y$ be a transfunction which is ε-localized on X. Define $D_\varepsilon(x) := \sup\{\delta > 0 : \Phi$ is (δ, ε)-localized at $x\}$. Then $D_\varepsilon : X \to (0, \infty)$ is continuous on X and if X is compact, then Φ is uniformly ε-localized on X.*

Proof. Let $x_0 \in X$. Let $x \in B(x_0; D_\varepsilon(x_0))$. It must follow by definition of D_ε that

$$D_\varepsilon(x_0) - d(x, x_0) \leq D_\varepsilon(x) \leq D_\varepsilon(x_0) + d(x, x_0); \tag{3.11}$$

this is because $B(x; D_\varepsilon(x_0) - d(x, x_0)) \subseteq B(x_0; D_\varepsilon(x_0)) \subseteq B(x; D_\varepsilon(x_0) + d(x, x_0))$.

Therefore, $|D_\varepsilon(x) - D_\varepsilon(x_0)| \leq d(x, x_0) \to 0$ as $x \to x_0$. Hence, D_ε is continuous on X. If X is compact, then D_ε obtains its minimum, positive value on X; call that value δ_X. Then for any positive $\delta < \delta_X$, we have that $\delta < D_\varepsilon(x)$ for all $x \in X$, meaning that Φ is (δ, ε)-localized at every $x \in X$. This precisely means that Φ is uniformly (δ, ε)-localized on X. $\square$

3.6 Graphs of Transfunctions

We introduce a concept analogous to the graph of a function and prove three theorems that shed some light on the nature of localized transfunctions.

Definition 3.27. Let $\Phi : X \rightsquigarrow Y$ be a transfunction, and let $\Gamma \subseteq X \times Y$ be measurable with respect to the product σ-algebra. We say that Γ *carries* Φ, denoted as $\Phi \sqsubset \Gamma$, if every measurable rectangle $A \times B \subseteq X \times Y$ with $(A \times B) \cap \Gamma = \varnothing$ implies that $\Phi(A) \sqsubset B^c$.

Similar to how carriers of a measure describe its support, the carriers of a transfunction describe its graph. This is a generalization of the concept of a graph of a function, as indicated by the following theorem.

Proposition 3.28. *For every measurable function $f : X \to Y$, the graph of f carries $f_\#$, that is, $f_\# \sqsubset \{(x, f(x)) : x \in X\}$.*

Proof. If $(A \times B) \cap \{(x, f(x)) : x \in X\} = \varnothing$, then $A \cap f^{-1}(B) = \varnothing$, so for every $\mu \sqsubset A$,

$$f_\#(\mu)(B) = \mu(f^{-1}(B)) = \mu(A \cap f^{-1}(B)) = 0. \tag{3.12}$$

□

We also have the reverse situation: similar to how functions can be generated by curves in $X \times Y$, transfunctions can be generated by subsets of $X \times Y$.

Proposition 3.29. *Let (X, Σ_X) be a measurable space and let (Y, Σ_Y, λ) be a finite measure space. If $\Gamma \subseteq X \times Y$ is a measurable set with respect to the product σ-algebra, then*

$$\Phi(\mu)(B) := (\mu \times \lambda)(\Gamma \cap (X \times B)) \tag{3.13}$$

defines a strongly σ-additive transfunction from X to Y such that $\Phi \sqsubset \Gamma$.

Proof. If $U_1, U_2, \dots \in \Sigma_Y$ are disjoint, then

$$\begin{aligned}\Phi(\mu)(\cup_{n=1}^{\infty} U_n) &= (\mu \times \lambda)(\Gamma \cap (X \times \cup_{n=1}^{\infty} U_n)) \\ &= (\mu \times \lambda)(\Gamma \cap \cup_{n=1}^{\infty} (X \times U_n)) \\ &= (\mu \times \lambda)(\cup_{n=1}^{\infty} (\Gamma \cap (X \times U_n))) \\ &= \sum_{n=1}^{\infty} (\mu \times \lambda)(\Gamma \cap (X \times U_n)) = \sum_{n=1}^{\infty} \Phi(\mu)(U_n), \end{aligned} \tag{3.14}$$

so $\Phi(\mu)$ is a measure on Y. Strong σ-additivity of Φ follows from $(\sum_{i=1}^{\infty} \mu_i) \times \lambda = \sum_{i=1}^{\infty} (\mu_i \times \lambda)$. Moreover, if $(A \times B) \cap \Gamma = \varnothing$ and $\mu \sqsubset A$, then

$$\begin{aligned}\Phi(\mu)(B) &= (\mu \times \lambda)(\Gamma \cap (X \times B)) \\ &= (\mu \times \lambda)(\Gamma \cap (A \times B)) = 0. \end{aligned} \tag{3.15}$$

□

Some localized transfunctions which are "close" to measurable functions turn out to be carried by what one might call "fat graphs". And if a transfunction has a "fat continuous graph", then it is localized. Proposition 3.30 makes these claims precise.

Proposition 3.30. *Let $f : X \to Y$ be measurable. If Φ is weakly σ-additive and ε-localized on X via f, then $\Phi \sqsubset \Gamma := \bigcup_{x \in X} (\{x\} \times B(f(x), \varepsilon))$.*

If f is continuous and $\Phi \sqsubset \Gamma$, then Φ is localized on X via f with $E_\Phi \leq \varepsilon$.

Proof. If Φ is weakly σ-additive and ε-localized on X via f, then for each $x \in X$ there is a $\delta_x > 0$ such that Φ is (δ_x, ε)-localized at $(x, f(x))$. If $(A \times B) \cap \Gamma = \varnothing$, then $B \cap \left(\bigcup_{a \in A} B(f(a), \varepsilon)\right) = \varnothing$

and thus $\bigcup_{a \in A} B(f(a), \varepsilon) \subseteq B^c$. Note that $\{B(a; \delta_a) : a \in A\}$ is an open cover of A with a countable subcover $\{B(a_n; \delta_n) : n \in \mathbb{N}\}$, where $\delta_n = \delta_{a_n}$. Let $A_n = B(a_n; \delta_n)$. Since $\Phi(A_n) \sqsubset B^c$ for each $n \in \mathbb{N}$, we have $\Phi(\cup_{n=1}^{\infty} A_n) \sqsubset B^c$, by Proposition 3.3. It then follows from $A \subseteq \cup_{n=1}^{\infty} A_n$ that $\Phi(A) \sqsubset B^c$, again by Proposition 3.3.

Now assume that f is continuous and $\Phi \sqsubset \Gamma$. Let $x \in X$ and $n \in \mathbb{N}$ be arbitrary. Then there exists $\delta > 0$ such that $f(B(x; \delta)) \subseteq B(f(x); 2^{-n})$ and it follows by definition of Γ and by our previous argument that $B(x; \delta) \times B(f(x); \varepsilon + 2^{-n})^c \cap \Gamma = \varnothing$. Since $\Phi \sqsubset \Gamma$, it follows that $\Phi(B(x; \delta)) \sqsubset B(f(x); \varepsilon + 2^{-n})$, resulting in Φ being localized on X via f. Moreover $E_\Phi(x) \leq \varepsilon + 2^{-n}$ for all $x \in X$ and $n \in \mathbb{N}$. Since x and n were arbitrary, we have $E_\Phi \leq \varepsilon$. $\square$

CHAPTER 4: TRANSFUNCTIONS AND THEIR CONNECTIONS TO PLANS, MARKOV OPERATORS AND OPTIMAL TRANSPORT

4.1 Introduction

For finite positive measure μ on (X, Σ_X) and real-valued function $f \in \mathcal{L}^1(X, \mu)$, let $f\mu$ denote the measure $A \mapsto \int_A f \, d\mu$ and define $\mathcal{M}_\mu^{p,+} = \{f\mu : f \in \mathcal{L}^p(X, \mu), \; f \geq 0\}$ and $\mathcal{M}_\mu^p := \{f\mu : f \in \mathcal{L}^p(X, \mu)\}$ for $p \in [1, \infty]$. We define $\mathcal{M}_\mu$ to be the set of all signed measures absolutely continuous with respect to μ. By the Radon-Nikodym Theorem, $\mathcal{M}_\mu = \mathcal{M}_\mu^1$. Similarly, for finite positive ν on (Y, Σ_Y) we define $\mathcal{M}_\nu^{p,+}$ and $\mathcal{M}_\nu^p$.

Strongly σ-additive transfunctions are those which are linear and continuous with respect to total variation. We will sometimes call an operator between Banach spaces *strongly σ-additive* if it is linear and norm-continuous.

Plans have applications for finding weak solutions for optimal transport problems [24]. Markov operators, defined in Section 2, have some similarities to stochastic matrices [10]. Plans and Markov operators have a bijective correspondence as described in [25] and in Section 2. We assign to any corresponding Markov operator/plan pair (T, κ) with marginals μ, ν a unique transfunction $\Phi : \mathcal{M}_\mu \to \mathcal{M}_\nu$ – called a *Markov transfunction*. However, each Markov transfunction corresponds to a family of Markov operators (resp. plans) which have different marginals but follow the same "instructions". The triple Φ, T, and κ are related via the equalities

$$\Phi(1_A\mu)(B) = \int_B T(1_A) \, d\nu = \kappa(A \times B) \tag{4.1}$$

and

$$\int_Y g \, d\Phi(f\mu) = \int_Y T(f) \, d(g\nu) = \int_{X\times Y} (f \otimes g) \, d\kappa \tag{4.2}$$

which hold for measurable $A \subseteq X$, $B \subseteq Y$, $f \in \mathcal{L}^\infty(X)$, and $g \in \mathcal{L}^\infty(Y)$. The first set of equalities, although simpler, imply the second set of equalities by strong σ-additivity of Φ, bounded-linearity of T, and σ-additivity of κ.

In this chapter we are motivated by the theory developed for the Monge-Kantorovich transportation problems and their far-reaching outcomes [17, 20, 24, 35]. Let μ, ν be probability measures on the Borel σ-algebra of $\mathbb{R}^d$. Let $\mathcal{M}_\mu$ (resp. $\mathcal{M}_\nu$) denote the space of finite measures on that are absolutely continuous with respect to μ (resp. ν). Let $c : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$ be a cost function. Consider the collection $\Pi(\mu, \nu)$ of all transport plans with μ and ν as their marginals. The goal is to find a transport plan with minimum cost

$$\inf_{\kappa \in \Pi(\mu,\nu)} \left\{ \int_{X\times Y} c \, d\kappa \right\}. \tag{4.3}$$

Since a transport plan "maps" a prior measure μ to a posterior measure ν, it can be described in the framework of transfunctions. There are a few main advantages when using transfunctions. First, all transport plans with the same "instructions" but with different prior and posterior measures correspond to the same transfunction. Second, while transport plans are measure-preserving by definition, it may not be a reasonable assumption in some applications. Finally, while a transport plan optimizes how μ is transformed into ν, it may be more natural to optimize how μ is transformed into one of several acceptable measures. However, describing how a transfunction (not necessarily corresponding to a transport plan) will be optimal with respect to cost function c and prior/posterior measures μ, ν is not as simple as with transport plans.

Let $\mathcal{F}_X$ and $\mathcal{F}_Y$ be spaces of measurable functions which are integrable by measures in $\mathcal{M}_X$ and $\mathcal{M}_Y$, respectively. If $\{\mathcal{F}_X, \mathcal{M}_X\}$ and $\{\mathcal{F}_Y, \mathcal{M}_Y\}$ are separating pairs with respect to integration as defined in Section 3, then we define the *Radon adjoint* of $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ (if it exists) to be the unique linear bounded operator $\Phi^* : \mathcal{F}_Y \to \mathcal{F}_X$ such that

$$\int_X \Phi^*(g)\, d\mu = \int_Y g\, d\Phi(\lambda) \tag{4.4}$$

for all $g \in \mathcal{F}_Y$ and $\lambda \in \mathcal{M}_X$.

If X and Y are second-countable locally compact Hausdorff spaces, if $\mathcal{F}_X$ and $\mathcal{F}_Y$ are Banach spaces of bounded continuous functions (uniform norm) and if $\mathcal{M}_X$ and $\mathcal{M}_Y$ are Banach spaces of finite regular signed measures (total variation), then any strongly σ-additive weakly-continuous transfunction $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ has an adjoint Φ^* which is a linear, uniformly-continuous and bounded-pointwise-continuous operator (and vice versa) such that $||\Phi|| = ||\Phi^*||$. In future research, we wish to develop functional analysis on transfunctions, and adjoints may be utilized to this end. In contexts where operators on functions are more appropriate or preferable, the adjoint may prove crucial.

A simple transfunction $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ is one which has the form

$$\Phi(\lambda) := \sum_{i=1}^{m} \langle f_i, \lambda \rangle \rho_i \tag{4.5}$$

for $f_1, \ldots, f_m \in \mathcal{F}_X$ and $\rho_1, \ldots, \rho_m \in \mathcal{M}_Y$, where $\langle f_i, \lambda \rangle := \int_X f_i\, d\lambda$. Simple transfunctions are weakly-continuous and strongly σ-additive. When working with locally compact Polish (metric) spaces, simple Markov transfunctions have two advantages: they weakly approximate all Markov transfunctions, and a subclass of them can be utilized to approximate the optimal cost between two marginals with respect to a transport cost $c(x, y)$ that is bounded by $\alpha d(x, y)^p$ for constants

$\alpha, p > 0$.

4.2 Markov Transfunctions

In this section, we describe a class of transfunctions in which each transfunction corresponds to a family of plans and a family of Markov operators. First, we introduce these concepts. All measurable or continuous functions shall be real-valued in this text. Note that the following definitions allow for all finite positive measures rather than all probability measures.

Definition 4.1. Let μ and ν be finite positive measures on (X, Σ_X) and (Y, Σ_Y) respectively with $||\mu|| = ||\nu||$. Let κ be a finite positive measure on the product measurable space $(X \times Y, \Sigma_{X\times Y})$. We say that κ is a *plan with marginals μ and ν* if $\kappa(A \times Y) = \mu(A)$ and $\kappa(X \times B) = \nu(B)$ for all $A \in \Sigma_X$ and $B \in \Sigma_Y$. We define $\Pi(\mu, \nu)$ to be the set of all plans with marginals μ and ν.

If random variables X, Y have laws μ, ν, then any coupling of X, Y has a law κ which is a plan in $\Pi(\mu, \nu)$.

Definition 4.2. Let μ and ν be finite positive measures on (X, Σ_X) and (Y, Σ_Y) respectively with $||\mu|| = ||\nu||$, and let $p \in [1, \infty]$. We say that a map $T : \mathcal{L}^p(X, \mu) \to \mathcal{L}^p(Y, \nu)$ is a *Markov operator* if:

(i) T is linear with $T1_X = 1_Y$;

(ii) $f \geq 0$ implies $Tf \geq 0$ for all $f \in \mathcal{L}^p(X, \mu)$;

(iii) $\int_X f \, d\mu = \int_Y Tf \, d\nu$ for all $f \in \mathcal{L}^p(X, \mu)$.

Notice that the definition of Markov operators depends on underlying measures μ and ν on X and Y respectively, even when $p = \infty$. We now define some properties for transfunctions that are analogous to (ii) and (iii) from Definition 4.2.

Definition 4.3. Let $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ be a transfunction.

(i) Φ is *positive* if $\lambda \geq 0$ implies that $\Phi\lambda \geq 0$ for all $\lambda \in \mathcal{M}_X$.

(ii) Φ is *measure-preserving* if $(\Phi\lambda)(Y) = \lambda(X)$ for all $\lambda \in \mathcal{M}_X$.

(iii) Φ is *Markov* if it is strongly σ-additive, positive and measure-preserving.

By [25], there is a bijective relationship between plans and Markov operators. We will show soon that a relationship between Markov operators and Markov transfunctions exists, which will imply that all three concepts are connected.

Lemma 4.4. *Let μ be a finite positive measure on (X, Σ_X). Define $J_\mu : \mathcal{L}^1(X, \mu) \to \mathcal{M}^1_\mu$ via $J_\mu f = f\mu$. Then J_μ (hence J_μ^{-1}) is a positive linear isometry.*

Proof. Positivity and linearity of integrals with respect to μ ensure that J_μ is positive and linear. Surjectivity of J_μ is the statement of the Radon-Nikodym Theorem. Injectivity and isometry hold because

$$||J_\mu f|| = ||J_\mu(f^+) - J_\mu(f^-)|| = \int_X f^+ d\mu + \int_X f^- d\mu = \int_X |f| d\mu = ||f||. \tag{4.6}$$

□

Theorem 4.5. *Let μ and ν be finite positive measures on X and Y respectively, with $||\mu|| = ||\nu||$ and let $s \in [1, \infty]$. For every Markov operator $T : \mathcal{L}^s(X, \mu) \to \mathcal{L}^s(Y, \nu)$, there exists a unique Markov transfunction $\Phi : \mathcal{M}^s_\mu \to \mathcal{M}^s_\nu$ such that*

$$\int_B T(1_A) d\nu = \Phi(1_A \mu)(B) \tag{4.7}$$

for all $A \in \Sigma_X$ and $B \in \Sigma_Y$.

Every Markov transfunction $\Phi : \mathcal{M}^s_\mu \to \mathcal{M}^s_\nu$ corresponds to a family of Markov operators $\{T_{\lambda,\rho} : \mathcal{L}^\infty(X,\lambda) \to \mathcal{L}^\infty(Y,\rho) \mid \lambda \in \mathcal{M}^{s,+}_\mu\ ,\ \rho = \Phi\lambda\}$ which satisfies

$$\int_B T_{\lambda,\rho}(1_A)d\rho = \Phi(1_A\lambda)(B) \tag{4.8}$$

for all $A \in \Sigma_X$ and $B \in \Sigma_Y$.

Proof. We prove the first statements for $s = 1$, then extend the argument to other values of s. Let $T : \mathcal{L}^1(X,\mu) \to \mathcal{L}^1(Y,\nu)$ be a Markov operator. Define $\Phi = J_\nu\, T\, J_\mu^{-1}$. Since all three operators in the definition of Φ are positive and strongly σ-additive, we see that Φ is also positive and strongly σ-additive. Next, if $\lambda \in \mathcal{M}_\mu$, then

$$(\Phi\lambda)(Y) = J_\nu(TJ_\mu^{-1}\lambda)(Y) = \int_Y T(J_\mu^{-1}\lambda)d\nu = \int_X J_\mu^{-1}(\lambda)d\mu = \lambda(X) \tag{4.9}$$

by the definitions of isometries J_μ^{-1} and J_ν, and by property (iii) of T, so Φ is measure-preserving. Finally, notice that

$$\Phi(1_A\mu)(B) = J_\nu T(J_\mu^{-1}(1_A\mu))(B) = J_\nu(T1_A)(B) = \int_B T(1_A)d\nu \tag{4.10}$$

for all $A \in \Sigma_X$ and $B \in \Sigma_Y$, hence the relation holds.

Now let $s \in (1,\infty]$ and let $T : \mathcal{L}^s(X,\mu) \to \mathcal{L}^s(Y,\nu)$ be a Markov operator. By Theorem 1 from [25], T can be uniquely extended to a Markov operator $\widehat{T}$ on $\mathcal{L}^1(X,\mu)$. By our previous argument, $\widehat{T}$ corresponds to a Markov transfunction $\widehat{\Phi}$ defined on $\mathcal{M}_\mu$. We define Φ to be the restriction of $\widehat{\Phi}$ to $\mathcal{M}^s_\mu$. The necessary properties are inherited from the previous argument.

Now we prove the second statement. Let $s \in [1,\infty]$, let $\Phi : \mathcal{M}^s_\mu \to \mathcal{M}^s_\nu$ be a Markov transfunction, let $\lambda \in \mathcal{M}^s_\mu$ be positive, and define $\rho := \Phi(\lambda) \in \mathcal{M}^s_\nu$, which is also positive. Define

$T = T_{\lambda,\rho} := J_\rho^{-1}\ \Phi\ J_\lambda$ with domain $\mathcal{L}^\infty(X, \lambda)$. Then

$$T(1_X) = J_\rho^{-1}\Phi(J_\lambda(1_X)) = J_\rho^{-1}(\Phi\lambda) = J_\rho^{-1}\rho = 1_Y. \tag{4.11}$$

Since all three operators in the definition of T are positive and strongly σ-additive, we see that T is also positive and strongly σ-additive, satisfying parts (i) and (ii) of Definition 4.2. Next, if $f \in \mathcal{L}^\infty(X, \lambda)$, then

$$\int_Y T f d\rho = \int_Y J_\rho^{-1}(\Phi J_\lambda f) d\rho = (\Phi(J_\lambda f))(Y) = (J_\lambda f)(X) = \int_X f d\lambda, \tag{4.12}$$

so (iii) of Definition 4.2 is met. Finally, notice that

$$\int_B T(1_A) d\rho = \int_B J_\rho^{-1}(\Phi J_\lambda(1_A)) d\rho = \Phi(J_\lambda(1_A))(B) = \Phi(1_A\lambda)(B) \tag{4.13}$$

for all $A \in \Sigma_X$ and $B \in \Sigma_Y$, so the relation holds.

□

One consequence from Theorem 4.5 is that any Markov transfunction defined on $\mathcal{M}_\mu^s$ for $s \in [1, \infty]$ uniquely extends or restricts to $\mathcal{M}_\mu^{s'}$ for all $s' \in [1, \infty]$, thus the value of s is insignificant. This is analogous to a similar property held by Markov operators [25].

The remainder of this section aims to emphasize the importance of Theorem 4.5. For any $p \in [1, \infty]$, a transfunction $\Phi : \mathcal{M}_\mu^p \to \mathcal{M}_\nu^p$, a Markov operator $T : \mathcal{L}^p(X, \mu) \to \mathcal{L}^p(Y, \nu)$ and a plan $\kappa \in \Pi(\mu, \nu)$ that satisfy the equalities

$$\Phi(1_A\mu)(B) = \int_B T(1_A)\, d\nu = \kappa(A \times B) \tag{4.14}$$

for all $A \subseteq X$ and $B \subseteq Y$ contain the same information (transportation method), but convey it differently. By extending (4.14) for all $f \in \mathcal{L}^p(X, \mu)$ and $g \in \mathcal{L}^q(Y, \nu)$ with $1/p + 1/q = 1$, we have

$$\int_Y g \, d\Phi(f\mu) = \int_Y T(f) \, d(g\nu) = \int_{X \times Y} (f \otimes g) \, d\kappa. \tag{4.15}$$

Note that if some positive measure μ' also generates $\mathcal{M}_\mu^p$, and if we define $\nu' = \Phi(\mu')$, then the same transfunction $\Phi : \mathcal{M}_\mu^p \to \mathcal{M}_\nu^p$ corresponds to a Markov operator $T' : \mathcal{L}^p(X, \mu') \to \mathcal{L}^p(Y, \nu')$ and it corresponds to a plan κ' with marginals μ' and ν'. Therefore T and T' are different Markov operators, κ and κ' are different plans, yet they follow the same "instructions" encoded by Φ. In this regard, Φ is a global way to describe a transportation method independent of marginals. If μ' instead generates a smaller space than $\mathcal{M}_\mu$, then Φ restricted to $\mathcal{M}_{\mu'}$ contains part but not all of the instructions. Regardless, Φ will be Markov on this restriction. Notably, if $\mu' = h\mu$, then $\Phi : \mathcal{M}_{\mu'} \to \mathcal{M}_{\nu'}$ has associated Markov operator $T_h(f) := T(hf)$ and associated plan $\kappa' = (h \otimes 1_Y)\kappa$.

4.3 Radon Adjoints of Transfunctions

Let (X, Σ_X) be a Borel measurable space, let $\mathcal{F}_X$ be a subset of bounded measurable real-valued functions on X and let $\mathcal{M}_X$ be a subset of finite signed measures on X. Analogously, we have Y, $\mathcal{F}_Y$ and $\mathcal{M}_Y$. For $f \in \mathcal{F}_X$ and $\lambda \in \mathcal{M}_X$, define $\langle f, \lambda \rangle := \int_X f \, d\lambda$. Similarly, for $g \in \mathcal{F}_Y$ and $\rho \in \mathcal{M}_Y$, define $\langle g, \rho \rangle := \int_Y g \, d\rho$. Occasionally, the elements within angular brackets shall be written in reverse order without confusion: for example, see Lemma 4.16.

We say that $\{\mathcal{F}_X, \mathcal{M}_X\}$ *is a separating pair* if $\langle f_1, \lambda \rangle = \langle f_2, \lambda \rangle$ for all $\lambda \in \mathcal{M}_X$ implies that $f_1 = f_2$, and if $\langle f, \lambda_1 \rangle = \langle f, \lambda_2 \rangle$ for all $f \in \mathcal{F}_X$ implies that $\lambda_1 = \lambda_2$. In this section, we shall develop some theory for two choices of the collections $\{\mathcal{F}_X, \mathcal{M}_X\}$ and $\{\mathcal{F}_Y, \mathcal{M}_Y\}$, which we call the *continuous setting* and the *measurable setting*.

Definition 4.6. Let $\{\mathcal{F}_X, \mathcal{M}_X\}$ and $\{\mathcal{F}_Y, \mathcal{M}_Y\}$ be separating pairs. Let $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ be a transfunction and let $S : \mathcal{F}_Y \to \mathcal{F}_X$ be a function. Then Φ and S are *Radon adjoints* of each other if

$$\int_Y g \, d\Phi(\lambda) = \int_X S(g) \, d\lambda, \quad \text{i.e.} \quad \langle g, \Phi(\lambda)\rangle = \langle S(g), \lambda\rangle \tag{4.16}$$

holds for all $g \in \mathcal{F}_Y$ and $\lambda \in \mathcal{M}_X$.

By the separation properties of $\langle \cdot, \cdot \rangle$, Radon adjoints of both kinds are unique if they exist. We shall denote the Radon adjoint of Φ by Φ^* and of S by S^*.

If (Φ, S) is a Radon adjoint pair, then

$$\langle g, \Phi \textstyle\sum_i \lambda_i\rangle = \langle Sg, \sum_i \lambda_i\rangle = \sum_i \langle Sg, \lambda_i\rangle = \sum_i \langle g, \Phi\lambda_i\rangle = \langle g, \sum_i \Phi\lambda_i\rangle, \tag{4.17}$$

meaning that Φ is strongly σ-additive. Similarly,

$$\langle S \textstyle\sum_i g_i, \lambda\rangle = \langle \sum_i g_i, \Phi\lambda\rangle = \sum_i \langle g_i, \Phi\lambda\rangle = \sum_i \langle Sg_i, \lambda\rangle = \langle \sum_i Sg_i, \lambda\rangle, \tag{4.18}$$

meaning that S is linear and uniformly-continuous.

Example 4.7. If $\Phi = f_\#$ (the push-forward operator) for some measurable $f : X \to Y$, then $\Phi^*(g) = g \circ f = f^* g$ (the pull-back operator acting on g). This is because $\int_Y g \, d(f_\# \lambda) = \int_X g \circ f \, d\lambda$ for all $g \in \mathcal{F}_Y$, $\lambda \in \mathcal{M}_X$.

Example 4.8. If $X = Y$ and $\Phi\lambda := f\lambda$ for some continuous (or measurable) $f : X \to \mathbb{R}$, then $\Phi^*(g) = gf$. This is because $\int_X g \, d(f\lambda) = \int_X gf \, d\lambda$ for all $g \in \mathcal{F}_X$, $\lambda \in \mathcal{M}_X$.

Definition 4.9.

(i) (f_n) *weakly converges* to f in $\mathcal{F}_X$, notated as $f_n \xrightarrow{w} f$, if every finite regular measure λ on X yields $\langle f_n, \lambda\rangle \to \langle f, \lambda\rangle$ as $n \to \infty$.

(ii) $(\lambda_n)_{n=1}^{\infty}$ *weakly converges* to λ in $\mathcal{M}_X$, notated as $\lambda_n \xrightarrow{w} \lambda$, if every bounded continuous $f : X \to \mathbb{R}$ yields $\langle f, \lambda_n\rangle \to \langle f, \lambda\rangle$ as $n \to \infty$.

(iii) An operator $S : \mathcal{F}_Y \to \mathcal{F}_X$ is *weakly continuous* if $g_n \xrightarrow{w} g$ in $\mathcal{F}_Y$ implies that $Sg_n \xrightarrow{w} Sg$ in $\mathcal{F}_X$.

(iv) A transfunction $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ is *weakly continuous* if $\lambda_n \xrightarrow{w} \lambda$ in $\mathcal{M}_X$ implies that $\Phi\lambda_n \xrightarrow{w} \Phi\lambda$ in $\mathcal{M}_Y$.

Note that weak convergence of (f_n) in Definition 4.9 (i) is the same notion as bounded-pointwise convergence.

4.4 Approximations of Identity

For the remainder of this chapter, let X and Y be locally-compact Polish spaces, and pick any complete metric for each of them when needed.

Definition 4.10. For a metric space (X, d) with $x \in X$, $A \subseteq X$ and $\delta > 0$, define $B(x; \delta) := \{z \in X : d(x, z) < \delta\}$ to be the *δ-ball around x* and define $B(A; \delta) := \cup_{x \in A} B(x; \delta)$ to be the *δ-inflation around A.*

Lemmas 4.11 and 4.12 aid in proving Proposition 4.14.

Lemma 4.11. *Let (X, d) be a locally compact metric space. The positive function $c : X \to (0, \infty]$ defined via $c(x) := \sup\{\delta > 0 : B(x; \delta)$ is precompact$\}$ is either identically ∞ or it is finite and* *continuous on X. It follows that every compact set K has a precompact inflation $B(K;\delta)$ for some $\delta > 0$ (depending on K).*

Proof. The proof is similar to the proof of Theorem 3.26. $\square$

Lemma 4.12. *Let (X,d) be a locally compact Polish metric space. Then there exists a pair of sequences $(x_i)_{i=1}^{\infty}$ from X and $(\beta_i)_{i=1}^{\infty}$ from $(0,1]$ and there exists a function $p:\mathbb{N}\to\mathbb{N}$ such that for all $n\in\mathbb{N}$,*

$$K_n := \bigcup_{i=1}^{p(n)} \overline{B(x_i,\beta_i/n)} \tag{4.19}$$

is compact with $K_n \supseteq K_n^{\circ} \supseteq K_{n-1}$ for $n\geq 2$ and $\cup_{n=1}^{\infty} K_n = X$.

Proof. Select an increasing sequence $(L_n)_{n=1}^{\infty}$ of compact subsets with $\cup_{n=1}^{\infty} L_n = X$ as guaranteed by Proposition 7.1.5 from [29], and let $c: X\to(0,\infty]$ be the positive function from Lemma 4.11. Let $\beta: X\to(0,1]$ be defined via $\beta(x) := \min\{c(x)/2, 1\}$. Note that $B(x;\beta(x))$ is precompact for all $x\in X$.

The open covering $\{B(x;\beta(x))\}_{x\in X}$ of L_1 has a finite subcover $\{B(x_i;\beta_i)\}_{i=1}^{p(1)}$, where $\beta_i := \beta(x_i)$. Define $K_1 := \cup_{i=1}^{p(1)} \overline{B(x_i;\beta_i)}$. We proceed via induction. Having constructed $K_{n-1} = \cup_{i=1}^{p(n-1)} \overline{B(x_i;\beta_i/(n-1))}$ for $n\geq 2$, create a precompact inflation $B(K_{n-1};\delta_{n-1})$ for some $\delta_{n-1} > 0$ by Lemma 4.11. The open cover $\{B(x;\beta(x)/n)\}_{x\in X}$ of $\overline{B(K_{n-1};\delta_{n-1})} \cup L_n$ admits a finite subcover $\{B(x_i;\beta_i/n)\}_{i=1}^{p(n)}$, where $\beta_i := \beta(x_i)$ and $\{x_i\}_{i=1}^{p(n-1)}$ are the points used to previously describe K_{n-1}. Define $K_n := \cup_{i=1}^{p(n)} \overline{B(x_i;\beta_i/n)}$. Since K_n is a finite union of compact sets, K_n is compact. It is clear by induction that $K_{n-1} \subseteq K_n^{\circ} \subseteq K_n$ and that $\cup_{n=1}^{\infty} K_n = \cup_{n=1}^{\infty} L_n = X$. $\square$

Using the setup from Lemma 4.12, we define the collection of sets

$$C_{n,i} := \overline{B(x_i, \beta_i/n)} - \cup_{j<i}\overline{B(x_j, \beta_j/n)} \tag{4.20}$$

for all $n, i \in \mathbb{N}$. It follows for any $n \in \mathbb{N}$ that $\cup_{i=1}^{p(n)} C_{n,i} = K_n$.

Definition 4.13. A measure μ is called a *point-mass measure at* x if $\mu(A) = 1$ when $x \in A$ and $\mu(A) = 0$ when $x \notin A$. A finite linear combination of point-mass measures is called a *simple measure*.

It is straightforward to show that simple measures are regular. The following proposition suggests a method to create approximations of identity, which shall be discussed in Sections 4.5 and 4.6.

Proposition 4.14. *Simple measures on a second-countable locally compact Hausdorff space form a dense subset of all finite regular measures with respect to weak convergence.*

Proof. Construct sequences $(x_i)_{i=1}^{\infty}$, $(\beta_i)_{i=1}^{\infty}$, $p : \mathbb{N} \to \mathbb{N}$ and $(C_{n,i})$ via Lemma 4.12. Fix some positive measure $\lambda \in \mathcal{M}_X^+$. Construct a sequence $(\lambda_n)_{n=1}^{\infty}$ of positive simple measures via $\lambda_n := \sum_{i=1}^{p(n)} \lambda(C_{n,i})\delta_{x_i}$. We will show that $\lambda_n \xrightarrow{w} \lambda$. In doing so, we fix some function $f \in \mathcal{C}_b(X)$ and show that $\langle f, \lambda_n \rangle \to \langle f, \lambda \rangle$. For density of signed measures, one utilizes the Jordan decomposition and applies a similar argument for each component.

Let $\varepsilon > 0$. Define $\eta := \varepsilon/(3||f|| + 3||\lambda|| + 1)$ so that $||f||\,\eta < \varepsilon/3$ and that $||\lambda||\,\eta < \varepsilon/3$. Choose some natural M such that $\lambda(K_M^c) < \eta$. Apply Lemma 4.11 to obtain some $\alpha > 0$ with $L := \overline{B(K_M; \alpha)}$ being compact. By uniform continuity of $f|_L$, choose some natural $N > M$ such that $2/N < \alpha$ and for all $x \in L$, $f(B(x; 2/N) \cap L) \subseteq B(f(x); \eta)$.

Now let $n > N$. Define $\rho_{n,M} := \sum_{i=1}^{p(n)} \lambda(C_{n,i} \cap K_M)\delta_{x_i}$. Notice that $C_{n,i} \cap K_M \neq \varnothing$ implies that $x_i \in B(K_M; 1/N)$ and that $C_{n,i} \subseteq B(K_M; 2/N) \subseteq L$, resulting in $f(C_{n,i}) \subseteq B(f(x_i); \eta)$. Three observations can be made:

(a) $|\langle f, \lambda - 1_{K_M}\lambda\rangle| \le ||f|| \cdot \lambda(K_M^c) < ||f||\, \eta$;

(b) $|\langle f, 1_{K_M}\lambda - \rho_{n,M}\rangle| \le \left|\int_{K_M} f\, d\lambda - \sum_{i=1}^{p(n)} f(x_i)\lambda(C_{n,i} \cap K_M)\right| < ||\lambda||\, \eta$;

(c) $|\langle f, \rho_{n,M} - \lambda_n\rangle| \le ||f|| \sum_{i=1}^{p(n)} \lambda(C_{n,i} \cap K_M^c) \le ||f||\, \lambda(K_M^c) < ||f||\, \eta$.

Therefore, $|\langle f, \lambda - \lambda_n\rangle| < 3(\varepsilon/3) = \varepsilon$. □

For any finite signed measure λ on X, the sequence (λ_n) of simple measures from Proposition 4.14 weakly converges to λ, hence the sequence of transfunctions (I_n) given by

$$I_n : \lambda \mapsto \lambda_n = \sum_{i=1}^{p(n)} \langle 1_{C_{n,i}}, \lambda\rangle \delta_{x_i} \tag{4.21}$$

is an approximation of identity.

Note that (4.21) is simply described with characteristic functions $(1_{C_{n,i}})$ and point-mass measures (δ_{x_i}). However, in Sections 4.5 and 4.6 either the characteristic functions must be replaced by bounded continuous functions or the point-mass measures must be replaced by compactly-supported measures that are absolutely continuous with respect to some underlying measure. With the correct choice of replacements, the same argument as given in Proposition 4.14 can be applied, yielding valid approximations of identities for the respective settings.

4.5 Continuous Setting: $\mathcal{F}_X = \mathcal{C}_b(X)$, $\mathcal{M}_X = \mathcal{M}_{fr}(X)$

Let $\mathcal{F}_X = \mathcal{C}_b(X)$ denote the Banach space of all bounded continuous functions on X with the uniform norm and let $\mathcal{M}_X = \mathcal{M}_{fr}(X)$ denote the Banach space of all finite (hence, regular) signed measures on X with the total variation norm. Develop Y, $\mathcal{F}_Y$, and $\mathcal{M}_Y$ analogously.

We now show that $\{\mathcal{F}_X, \mathcal{M}_X\}$ (respectively $\{\mathcal{F}_Y, \mathcal{M}_Y\}$) is a separating pair in this setting. First, if $f_1, f_2 \in \mathcal{F}_X$ are different, then they differ at some $x_0 \in X$, meaning $\delta_{x_0} \in \mathcal{M}_X$ distinguishes f_1 from f_2 because $\langle f_1, \delta_{x_0}\rangle = f_1(x_0) \neq f_2(x_0) = \langle f_2, \delta_{x_0}\rangle$. Second, if $\lambda_1, \lambda_2 \in \mathcal{M}_X$ are different, then the Jordan decomposition of $\lambda_1 - \lambda_2$ yields sets $P, N \in \Sigma_X$ with $\lambda_1 \geq \lambda_2$ on P, $\lambda_1 \leq \lambda_2$ on N, and either $\lambda_1(P) > \lambda_2(P)$ or $\lambda_1(N) < \lambda_2(N)$. Then take any sequence $(f_n)_{n=1}^{\infty}$ from $\mathcal{F}_X$ which weakly converges to $1_P - 1_N$. Since $\langle 1_P - 1_N, \lambda_1\rangle > \langle 1_P - 1_N, \lambda_2\rangle$, the Dominated Convergence Theorem yields that there must be some natural m with $\langle f_m, \lambda_1\rangle > \langle f_m, \lambda_2\rangle$. Therefore, f_m distinguishes λ_1 from λ_2.

An approximation of identity can be formed in this setting: keep the point-mass measures $\rho_{n,i} := \delta_{x_i}$, then for each natural n, replace the characteristic functions $\{1_{C_{n,i}} : 1 \leq i \leq p(n)\}$ used in Proposition 4.14 with positive compactly supported continuous functions $\{f_{n,i} : 1 \leq i \leq p(n)\}$ such that $f_{n,i} \leq 1_{B(C_{n,i};1/n)}$ and that $1_{K_n} \leq \sum_{i=1}^{p(n)} f_{n,i} \leq 1_{B(K_n;1/n)}$. Then an approximation of identity in the continuous setting is given by the sequence (I_n), where

$$I_n : \lambda \mapsto \sum_{i=1}^{p(n)} \langle f_{n,i}, \lambda\rangle \, \rho_{n,i} = \sum_{i=1}^{p(n)} \langle f_{n,i}, \lambda\rangle \, \delta_{x_i}. \tag{4.22}$$

Theorem 4.15. *Every strongly σ-additive and weakly-continuous transfunction $\Phi : \mathcal{M}_{fr}(X) \to \mathcal{M}_{fr}(Y)$ has a strongly σ-additive and weakly-continuous Radon adjoint $S : \mathcal{C}_b(Y) \to \mathcal{C}_b(X)$. Conversely, every strongly σ-additive and weakly-continuous operator $S : \mathcal{C}_b(Y) \to \mathcal{C}_b(X)$ has a strongly σ-additive and weakly-continuous Radon adjoint $\Phi : \mathcal{M}_{fr}(X) \to \mathcal{M}_{fr}(Y)$. When the Radon adjoint pair exists, their operator norms equal (with respect to total-variation and uniform-convergence).*

Proof. For the first claim, define $S(g)(x) := \langle g, \Phi(\delta_x)\rangle$ for all $g \in \mathcal{F}_Y$, $x \in X$. It follows that

$\langle S(g), \delta_x \rangle = \langle g, \Phi(\delta_x) \rangle$ for all g, x. Let $x_n \to x$ on X, so that $\delta_{x_n} \xrightarrow{w} \delta_x$, which means that $\Phi(\delta_{x_n}) \xrightarrow{w} \Phi(\delta_x)$. Also let $g_n \to g$ bounded-pointwise in $\mathcal{F}_Y$ (i.e. $g_n \xrightarrow{w} g$). Then

$$S(g)(x_n) = \langle S(g), \delta_{x_n} \rangle = \langle g, \Phi(\delta_{x_n}) \rangle \to \langle g, \Phi(\delta_x) \rangle = \langle S(g), \delta_x \rangle = S(g)(x), \tag{4.23}$$

$$||S(g)|| = \sup_{x \in X} |S(g)(x)| = \sup_{x \in X} |\langle S(g), \delta_x \rangle| = \sup_{x \in X} |\langle g, \Phi(\delta_x) \rangle| \leq ||g|| \cdot ||\Phi||, \tag{4.24}$$

and

$$S(g_n)(x) = \langle S(g_n), \delta_x \rangle = \langle g_n, \Phi\delta_x \rangle \to \langle g, \Phi\delta_x \rangle = \langle S(g), \delta_x \rangle = S(g)(x), \tag{4.25}$$

meaning that $S(g) \in \mathcal{F}_Y$, that S is bounded (hence uniform-continuous) and that S is bounded-pointwise-continuous.

Since countable linear combinations of point-mass measures are weakly dense in $\mathcal{M}_X$, the linearity and weak-continuity of the second coordinate in the $\langle \cdot, \cdot \rangle$ structure and the weak-continuity of Φ yields that $\langle S(g), \lambda \rangle = \langle g, \Phi\lambda \rangle$ for all $g \in \mathcal{F}_Y$ and $\lambda \in \mathcal{M}_X$. Hence, S is the Radon adjoint of Φ with the desired properties.

For the second claim, note that for every $\lambda \in \mathcal{M}_X$, the continuous functional $g \mapsto \langle S(g), \lambda \rangle$ defined on $\mathcal{C}_0(Y)$ has Riesz representation $\langle \cdot, \Phi(\lambda) \rangle$ for some unique signed measure $\Phi(\lambda) \in \mathcal{M}_Y$. Defining Φ in this manner for all λ, we obtain that $\langle S(g), \lambda \rangle = \langle g, \Phi\lambda \rangle$ for all $g \in \mathcal{C}_0(Y)$ and $\lambda \in \mathcal{M}_X$. $C_0(Y)$ is dense in $C_b(Y)$ with respect to bounded-pointwise convergence, so with $\mathcal{C}_0(Y) \ni g_n \xrightarrow{w} g \in \mathcal{C}_b(Y)$, it follows that $\langle g_n, \Phi\lambda \rangle \to \langle g, \Phi\lambda \rangle$ by the Dominated Convergence Theorem. Similarly, bounded-pointwise-continuity of S ensures that $S(g_n) \xrightarrow{w} S(g)$, which means that $\langle S(g_n), \lambda \rangle \to \langle S(g), \lambda \rangle$ by the Dominated Convergence Theorem. Therefore, $\langle S(g), \lambda \rangle = \langle g, \Phi\lambda \rangle$ for all $g \in \mathcal{F}_Y$ and $\lambda \in \mathcal{M}_X$, implying that Φ is the Radon adjoint of S. An earlier remark shows that Φ is strongly σ-additive. To see that Φ is weakly-continuous, let $\lambda_n \xrightarrow{w} \lambda$. Then

$\langle g, \Phi\lambda_n\rangle = \langle S(g), \lambda_n\rangle \to \langle S(g), \lambda\rangle = \langle g, \Phi\lambda\rangle$. Therefore, $\Phi\lambda_n \xrightarrow{w} \Phi\lambda$. Finally, $||\Phi|| \leq ||S||$, hence $||\Phi|| = ||S||$, follows via

$$||\Phi\lambda|| = \sup_{||g||=1} |\langle g, \Phi\lambda\rangle| = \sup_{||g||=1} |\langle S(g), \lambda\rangle| \leq ||S|| \cdot ||\lambda||. \tag{4.26}$$

□

4.6 Measurable Setting: $\mathcal{F}_X = \mathcal{L}^\infty(X, \mu)$, $\mathcal{M}_X = \mathcal{M}_\mu^\infty$

In this setting, let (X, Σ_X, μ) be a finite measure space, let $\mathcal{F}_X := \mathcal{L}^\infty(X, \mu)$ and let $\mathcal{M}_X := \mathcal{M}_\mu^\infty$. Define $(Y, \Sigma_Y, \nu), \mathcal{F}_Y, \mathcal{M}_Y$ analogously.

We now show that $\{\mathcal{F}_X, \mathcal{M}_X\}$ (respectively $\{\mathcal{F}_Y, \mathcal{M}_Y\}$) separate each other. First, if $f_1, f_2 \in \mathcal{F}_X$ are not equal μ-a.e., then there exists some natural number m such that $A_m := \{f_1 - f_2 \geq 1/m\}$ or $B_m := \{f_2 - f_1 \geq 1/m\}$ has positive μ measure. Therefore, the measure $\lambda := (1_{A_m} - 1_{B_m})\mu \in \mathcal{M}_X$ distinguishes f_1 from f_2 because $\langle f_1 - f_2, \lambda\rangle \geq \mu(A_m \cup B_m)/m > 0$. Second, if $\lambda_1, \lambda_2 \in \mathcal{M}_X$ are different, then the Jordan decomposition of $\lambda_1 - \lambda_2$ yields sets $P, N \in \Sigma_X$ with $\lambda_1 \geq \lambda_2$ on P, $\lambda_1 \leq \lambda_2$ on N, and either $\lambda_1(P) > \lambda_2(P)$ or $\lambda_1(N) < \lambda_2(N)$. Then $1_P - 1_N \in \mathcal{F}_X$ distinguishes λ_1 from λ_2 because $\langle 1_P - 1_N, \lambda_1 - \lambda_2\rangle = (\lambda_1(P) - \lambda_2(P)) + (\lambda_2(N) - \lambda_1(N)) > 0$.

An approximation of identity can be formed in this setting: for each natural n and $1 \leq i \leq p(n)$, replace each point-mass measure δ_{x_i} used in Proposition 4.14 with the measure $\rho_{n,i} := 1_{C_{n,i}}\mu$ and define $f_{n,i} := 1_{C_{n,i}}/\mu(C_{n,i})$ when $\mu(C_{n,i}) > 0$; otherwise, define $f_{n,i} = 0$. That is, an approximation of identity in the measurable setting is given by the sequence (I_n), where

$$I_n : \lambda \mapsto \sum_{i=1}^{p(n)} \langle f_{n,i}, \lambda\rangle \, \rho_{n,i} = \sum_{\substack{i=1 \\ \mu(C_{n,i})>0}}^{p(n)} \left\langle \frac{1_{C_{n,i}}}{\mu(C_{n,i})}, \lambda \right\rangle 1_{C_{n,i}}\mu. \tag{4.27}$$

The following lemma will be used in the proof of the next theorem:

Lemma 4.16. *For every strongly σ-additive transfunction $\Phi : \mathcal{M}_\mu^\infty \to \mathcal{M}_\nu^\infty$, there is a unique strongly σ-additive transfunction $\Phi^\dagger : \mathcal{M}_\nu^\infty \to \mathcal{M}_\mu^\infty$ such that*

$$\Phi^\dagger(1_B\nu)(A) = \Phi(1_A\mu)(B) \tag{4.28}$$

for all $A \subseteq X, B \subseteq Y$, which implies that $\langle f, \Phi^\dagger(g\nu)\rangle = \langle \Phi(f\mu), g\rangle$ for all $f \in \mathcal{L}^\infty(X,\mu)$, $g \in \mathcal{L}^\infty(Y,\nu)$. Also, $\Phi^{\dagger\dagger} = \Phi$.

Proof. Let Φ be strongly σ-additive. For fixed $B \subseteq Y$, it follows by strong σ-additivity of Φ that the set function $A \mapsto \Phi(1_A\mu)(B)$ is a measure. Define this measure to be $\Psi(1_B\nu)$. Then Ψ, defined on $\{1_B\nu \mid B \subseteq Y\}$ is a strongly σ-additive transfunction that behaves like $\Phi^\dagger$ in (4.28). Ψ can be linearly extended to $\mathcal{M}_\nu^\infty$ according to the following equalities for $A \subseteq X, g \cong \sum_j \beta_j 1_{B_j}$ with $\sum_j |\beta_j| < \infty$:

$$\Psi(g\nu)(A) = \sum_{j=1}^{\infty} \beta_j \Psi(1_{B_j}\nu)(A) = \sum_{j=1}^{\infty} \beta_j \Phi(1_A\mu)(B_j) = \int_Y g \, d\Phi(1_A\mu). \tag{4.29}$$

The extended Ψ is strongly σ-additive on $\mathcal{M}_\nu^\infty$. A similar calculation shows that

$$\int_X f \, d\Psi(g\nu) = \int_Y g \, d\Phi(f\mu) \tag{4.30}$$

for all $f \in \mathcal{L}^\infty(X,\mu)$ and $g \in \mathcal{L}^\infty(Y,\nu)$. Therefore $\Phi^\dagger$ is uniquely determined to be Ψ. Finally, $\Phi^{\dagger\dagger}(1_A\mu)(B) = \Phi^\dagger(1_B\nu)(A) = \Phi(1_A\mu)(B)$ for measurable $A \subseteq X$ and $B \subseteq Y$, so $\Phi^{\dagger\dagger} = \Phi$. $\square$

If Φ is Markov, then $\Phi^\dagger$ is also Markov. Furthermore, the plans $\kappa, \kappa^\dagger$ corresponding to $\Phi, \Phi^\dagger$ respectively are dual to each other: that is, $\kappa(A \times B) = \kappa^\dagger(B \times A)$ for all measurable sets

$A \subseteq X, B \subseteq Y$. However, $\Phi^\dagger$ is sensitive to the choice of measures μ and ν, which is not ideal when working with non-injective extensions of Φ.

Theorem 4.17. *Every strongly σ-additive $\Phi : \mathcal{M}_\mu^\infty \to \mathcal{M}_\nu^\infty$ has a linear and bounded Radon adjoint $S : \mathcal{L}^\infty(Y, \nu) \to \mathcal{L}^\infty(X, \mu)$. Conversely, every linear and bounded operator $S : \mathcal{L}^\infty(Y, \nu) \to \mathcal{L}^\infty(X, \mu)$ has a strongly σ-additive adjoint $\Phi : \mathcal{M}_\mu^\infty \to \mathcal{M}_\nu^\infty$.*

Proof. Assume that $\Phi : \mathcal{M}_\mu^\infty \to \mathcal{M}_\nu^\infty$ is strongly σ-additive and define $S := J_\mu^{-1}\Phi^\dagger J_\nu$ with domain $\mathcal{L}^\infty(Y, \nu)$. Then S is linear and bounded. $S = \Phi^*$ follows because for any $f \in \mathcal{L}^\infty(X, \mu)$ and $g \in \mathcal{L}^\infty(Y, \nu)$,

$$\langle g, \Phi(f\mu)\rangle = \langle \Phi^\dagger(g\nu), f\rangle = \langle (J_\mu S)g, f\rangle = \langle S(g), f\mu\rangle. \tag{4.31}$$

On the other hand, assume that $S : \mathcal{L}^\infty(Y, \nu) \to \mathcal{L}^\infty(X, \mu)$ is linear and bounded. Then define $\Psi := J_\mu S J_\nu^{-1}$ and $\Phi := \Psi^\dagger$. Then Φ is strongly σ-additive. $\Phi = S^*$ follows because for any $f \in \mathcal{L}^\infty(X, \mu)$ and $g \in \mathcal{L}^\infty(Y, \nu)$,

$$\langle S(g), f\mu\rangle = \langle (J_\mu^{-1}\Psi)(g\nu), f\mu\rangle = \langle \Phi^\dagger(g\nu), f\rangle = \langle g, \Phi(f\mu)\rangle. \tag{4.32}$$

□

4.7 Simple Transfunctions

Let $\mathcal{F}_X$, $\mathcal{F}_Y$, $\mathcal{M}_X$, and $\mathcal{M}_Y$ be defined in either the continuous setting or the measurable setting.

Definition 4.18. A transfunction $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ is *simple* if there exist functions $(f_i)_{i=1}^m$ from

$\mathcal{F}_X$ and there exist measures $(\rho_i)_{i=1}^m$ from $\mathcal{M}_Y$ such that

$$\forall \lambda \in \mathcal{M}_X, \quad \Phi\lambda = \sum_{i=1}^{m} \langle f_i, \lambda\rangle \rho_i. \tag{4.33}$$

It is straightforward to verify that simple transfunctions are strongly σ-additive. In the continuous setting, simple transfunctions are also weakly-continuous. Therefore by Theorem 4.15 in the continuous setting or Theorem 4.17 in the measurable setting, the Radon adjoint Φ^* exists and satisfies

$$\forall g \in \mathcal{F}_Y, \quad \Phi^* g = \sum_{i=1}^{m} \langle g, \rho_i\rangle f_i. \tag{4.34}$$

Note that the approximations of identity covered in the continuous setting in (Subsection 3.2) and in the measurable setting in (Subsection 3.3) involve sequences of simple transfunctions.

Theorem 4.19. *In either continuous or measurable settings, linear weakly-continuous transfunctions can be approximated by simple transfunctions with respect to weak convergence; that is, simple transfunctions form a dense subset of linear weakly-continuous transfunctions with respect to weak convergence.*

Proof. Let $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ be a linear and weakly-continuous transfunction and fix $\lambda \in \mathcal{M}_X$. Define $\Phi_n := \Phi\ I_n$, where $I_n : \lambda \mapsto \sum_{i=1}^{p(n)} \langle f_{n,i}, \lambda\rangle \rho_{n,i}$ forms the approximation of identity as defined in either Subsections 3.2 (continuous setting) or 3.3 (measurable setting). Then $\Phi_n\lambda = \sum_{i=1}^{p(n)} \langle f_{n,i}, \lambda\rangle \Phi\rho_{n,i}$, implying that Φ_n is a simple transfunction. It follows by $I_n\lambda \xrightarrow{w} \lambda$ and by weak-continuity of Φ that $\Phi_n\lambda = \Phi(I_n\lambda) \xrightarrow{w} \Phi\lambda$. □

4.8 Applications: Optimal Transport

Markov transfunctions provide a new perspective to optimal transport theory.

Definition 4.20. Let (X, Σ_X, μ) and (Y, Σ_Y, ν) be Polish measure spaces with finite positive measures μ and ν, respectively, with $||\mu|| = ||\nu||$. A *cost function* is any continuous function $c : X \times Y \to [0, \infty)$. A plan $\kappa \in \Pi(\mu, \nu)$ is *c-optimal* if $\int_{X\times Y} c \, d\kappa \leq \int_{X\times Y} c \, d\pi$ for all $\pi \in \Pi(\mu, \nu)$. A Markov transfunction $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ is *c-optimal on* μ if the corresponding plan κ with marginals μ and $\Phi\mu$ is c-optimal, and Φ is simply *c-optimal* if it is c-optimal on $\mathcal{M}_X$.

The next proposition implies that optimal inputs for Φ form a large class of measures.

Proposition 4.21. *Let $(X, \Sigma_X), (Y, \Sigma_Y)$ be Polish spaces, let c be a cost function, and let $\Phi : \mathcal{M}_X \to \mathcal{M}_Y$ be a Markov transfunction. If Φ is c-optimal on $\mu \in \mathcal{M}_X$, then Φ is c-optimal on $\mathcal{M}_\mu^\infty$.*

Proof. The proof follows easily from Theorem 4.6 in [24] on the inheritance of optimality of plans by restriction. □

In the next theorem, we provide a "warehouse strategy" which approximates the optimal cost between fixed marginals with respect to some cost function. In summary, the input marginal is first subdivided by local regions, and the subdivided measures are sent to point mass measures – warehouses – within their respective regions. Second, mass is transferred between warehouses. Finally, the warehouses locally redistribute to form the output marginal. The overall cost of transport via the warehouse strategy approaches the optimal cost as the size of the regions decreases.

Theorem 4.22. *Let (X, Σ_X) be a locally compact Polish measurable space with complete metric d, let λ and ρ be finite positive compactly-supported measures with $||\lambda|| = ||\rho||$, and let $c : X \times X \to$*

$[0, \infty)$ *be a cost function with* $c(x, y) \leq \alpha d(x, y)^p$ *for some constants* $\alpha, p > 0$. *The optimal cost between marginals* λ, ρ *with respect to* c *can be sufficiently approximated by the costs of simple Markov transfunctions.*

Proof. Assume the continuous setting and consider the approximation of identity (I_n) from Subsection 3.2. For large n, we create a composition of three simple Markov transfunctions: λ first maps to $I_n\lambda = \sum_{i=1}^{p(n)} \langle f_{n,i}, \lambda\rangle \delta_{x_i}$, which maps to $I_n\rho = \sum_{i=1}^{p(n)} \langle f_{n,i}, \rho\rangle \delta_{x_i}$, which finally maps to ρ. These steps are measure-preserving because K_n (from Lemma 4.12) contains the supports of λ and ρ for large n. The most crucial goal is to determine the optimal simple Markov transfunction for the middle step.

The first and last steps cost no more than $\alpha n^{-p}||\lambda||$ each, which reduces to 0 as $n \to \infty$. This means that the optimal cost between marginals λ_n and ρ_n approaches the optimal cost between marginals λ and ρ as $n \to \infty$. By approximating each of the values $\langle f_{n,i}, \lambda\rangle \approx a_{n,i}/z$ and $\langle f_{n,i}, \rho\rangle \approx b_{n,i}/z$ for natural numbers $a_{n,i}, b_{n,i}, z$ with $1 \leq i \leq p(n)$, the middle step can approximately be interpreted as the Assignment Problem on a weighted bipartite graph between vertex sets P and Q, where P denotes a set created by forming $a_{n,i}$ copies of a vertex corresponding to each δ_{x_i} in λ_n, Q denotes the set created by forming $b_{n,j}$ copies of a vertex corresponding to each δ_{x_j} in ρ_n, and drawing edges between these vertices with weight $c(x_i, x_j)$. This problem has been studied, and can be solved in polynomial time of $|P| = \sum_{i=1}^{p(n)} a_{n,i} \approx ||\lambda||z$; the Hungarian method is one well-known algorithm [7]. $\square$

Although Theorem 4.22 provides a sequence of simple transfunctions that approximate the optimal cost between fixed marginals, the sequence is not expected to converge weakly to an optimal Markov transfunction, as the solutions to the middle step could vary greatly as n increases. Consequently, we can find a Markov transfunction whose cost between marginals is sufficiently close to the optimal cost, but Theorem 4.22 does not provide an optimal Markov transfunction.

However, for any Markov transfunction between fixed marginals, the next theorem yields an approximation by simple Markov transfunctions with respect to weak convergence. Consequently, the cost between the marginals of the constructed sequence of simple Markov transfunctions approaches the cost for the original transfunction.

Theorem 4.23. *Let (X, Σ_X, μ) and (Y, Σ_Y, ν) be locally compact Polish measure spaces with finite compactly-supported positive measures μ and ν such that $||\mu|| = ||\nu||$. Any Markov transfunction $\Phi : \mathcal{M}_\mu \to \mathcal{M}_\nu$ can be approximated by simple Markov transfunctions with respect to weak convergence.*

Proof. Assume the measurable setting and consider the approximation of identity (I_n) from Subsection 3.3 with respect to μ. Let n be large so that K_n (from Lemma 4.12) contains the supports of μ and ν.

Let κ be the plan corresponding to Markov transfunction Φ from Theorem 4.5. For $1 \leq i, j \leq p(n)$, the quantity $\kappa(C_{n,i} \times C_{n,j})$ represents how much mass transfers from $1_{C_{n,i}}\mu$ to $1_{C_{n,j}}\nu$. If $\mu(C_{n,i})\nu(C_{n,j}) > 0$, then we can approximate nonzero measure $(1_{C_{n,i}} \otimes 1_{C_{n,j}})\kappa$ with

$$\kappa_{n,i,j} := \kappa(C_{n,i} \times C_{n,j}) \frac{1_{C_{n,i}}\mu}{\mu(C_{n,i})} \times \frac{1_{C_{n,j}}\nu}{\nu(C_{n,j})}. \tag{4.35}$$

Otherwise, we define $\kappa_{n,i,j} := 0$. Then $\kappa_n := \sum_i \sum_j \kappa_{n,i,j}$ is a plan from $\Pi(\mu, \nu)$ which corresponds to a Markov transfunction Φ_n from Theorem 4.5.

Next, we show that $\kappa_n \xrightarrow{w} \kappa$ as $n \to \infty$. Let $c \in \mathcal{C}_b(X \times Y)$ with $||c|| \leq 1$, and for $1 \leq i, j \leq p(n)$, let $\beta_{n,i,j} := \sup c(C_{n,i} \times C_{n,j}) - \inf c(C_{n,i} \times C_{n,j})$. By uniform continuity of c on $K_n \times K_n$, we have that $\beta_n := \max\{\beta_{n,i,j} \; : \; 1 \le i, j \le p(n)\} \to 0$ as $n \to \infty$, which implies that

$$|\langle c, \kappa - \kappa_n \rangle| \le \sum_i \sum_j \beta_{n,i,j} \, \kappa(C_{n,i} \times C_{n,j}) \le \beta_n ||\kappa|| \to 0. \tag{4.36}$$

There are some properties of Φ_n worth noting: Φ_n maps $\mathcal{M}_\mu^\infty$ to span$\{1_{C_{n,j}}\nu\}$; Φ_n behaves as a matrix when applied to span$\{1_{C_{n,i}}\mu\}$; the structure of κ_n guarantees that $\Phi_n = \Phi_n I_n$. If we choose bases $(1_{C_{n,i}}\mu)$ and $(1_{C_{n,j}}\nu)$, the matrix M_n representing Φ_n has entries $M_n(j,i) := \kappa(C_{n,i} \times C_{n,j})/\nu(C_{n,j})$.

Let $\lambda \in \mathcal{M}_\mu^\infty$ and for $1 \le i \le p(n)$. Then

$$\begin{aligned} \Phi_n \lambda = \Phi_n I_n \lambda &= \sum_j \left\langle \sum_i \frac{\kappa(C_{n,i} \times C_{n,j})}{\nu(C_{n,j})} \frac{1_{C_{n,i}}}{\mu(C_{n,i})}, \lambda \right\rangle 1_{C_{n,j}} \nu \\ &= \sum_j \left\langle \sum_i ||\kappa_{n,i,j}|| 1_{C_{n,i}}, \lambda \right\rangle 1_{C_{n,j}} \nu, \end{aligned} \tag{4.37}$$

showing that Φ_n is simple.

We now show that $\Phi_n \lambda \xrightarrow{w} \Phi \lambda$ as $n \to \infty$. Let $g \in \mathcal{C}_b(Y)$ with $||g|| \le 1$. Let $\varepsilon > 0$. Since $\lambda = f\mu$ for some $f \in \mathcal{L}^\infty(X, \mu)$, choose some $\tilde{f} \in \mathcal{C}_b(X)$ such that $||(f - \tilde{f})\mu|| < \varepsilon/3$. Since $||\Phi^*|| = ||\Phi_n^*|| = 1$, we have that

$$|\langle g, \Phi(f - \tilde{f})\mu \rangle| = |\langle \Phi^* g, (f - \tilde{f})\mu \rangle| \le ||\Phi^* g|| \cdot ||(f - \tilde{f})\mu|| < \varepsilon/3, \tag{4.38}$$

and that

$$\langle g, \Phi_n(f - \tilde{f})\mu \rangle| = |\langle \Phi_n^* g, (f - \tilde{f})\mu \rangle| \le ||\Phi_n^* g|| \cdot ||(f - \tilde{f})\mu|| < \varepsilon/3. \tag{4.39}$$

Since $\kappa_n \xrightarrow{w} \kappa$ as $n \to \infty$ and $\tilde{f} \otimes g \in \mathcal{C}_b(X \times Y)$, there is some natural N so that for all $n \geq N$,

$$\langle g, (\Phi - \Phi_n)\tilde{f}\mu\rangle| = |\langle \tilde{f} \otimes g, \kappa - \kappa_n\rangle| < \varepsilon/3. \tag{4.40}$$

It follows that $|\langle g, (\Phi - \Phi_n)\lambda\rangle| < \varepsilon$ for $n \geq N$ by the triangle inequality. □

If μ and ν are not compactly supported, then one can still use the argument from Theorem 4.23 to obtain an approximation of Φ with a sequence of simple transfunctions Φ_n which are Markov when restricted to $K_n \times K_n$. In that case, to show $\kappa_n \xrightarrow{w} \kappa$ as $n \to \infty$, (4.36) should be modified to

$$|\langle c, \kappa - \kappa_n\rangle| \leq \kappa(K_n^c) + \sum_i \sum_j \beta_{n,i,j}\, \kappa(C_{n,i} \times C_{n,j}) \leq \kappa(K_n^c) + \beta_n ||\kappa_n|| \to 0. \tag{4.41}$$

CHAPTER 5: TRANSFUNCTION MODELS IN MATHEMATICAL BIOLOGY

5.1 Examples of Transfunctions as Models for Population Dynamics

One direction for future research is to find applications of transfunction theory within mathematical biology. Transfunctions provide a natural framework for discrete-time population dynamics models. A population can be described as a measure μ on $X \subseteq \mathbb{R}^2$ which contains information about the size of the population and its spatial arrangement. A transfunction captures the dynamics of the population over one unit of time.

Traditional approaches to population dynamics are based on differential equations where the Laplacian operator is used to denote dispersal. In contrast, we will describe dispersal in this chapter by using convolutions of measures within the transfunction models found in Examples 5.1 – 5.7. Convolutions of (positive) measures form a natural candidate for dispersal because they have a blurring or smoothing effect (similar to how one function can be blurred or smoothed out by convolving it with another function). The transfunction models for population dynamics in this chapter were motivated by mathematical considerations and our understanding of how a population would evolve with time. As of now, these models have not been validated or tested with data and have not been compared with classical models.

Example 5.1. Suppose a single tree at the origin (δ_0) produces r (viable) seeds on average per year, dispersed with probability distribution γ (measure), then survives to the next year with probability $s \in (0, 1)$. One year later, the expected distribution of trees is $s\delta_0 + r\gamma$. Convolutions can generalize this behavior: if we now begin with a collection of trees modeled as measure μ, then one year later, the expected distribution of trees is $s\mu + r\mu * \gamma = (s\delta_0 + r\gamma) * \mu$. Therefore, the transfunction model for how the population of trees changes over one year of time is given via

$$\Phi(\mu) = (s\delta_0 + r\gamma) * \mu. \tag{5.1}$$

After n years elapse, we have $\Phi^n(\mu) = (s\delta_0 + r\gamma)^n * \mu$ with expected number of trees expressed as $||(s\delta_0 + r\gamma)^n * \mu|| = (s + r)^n||\mu||$. Therefore, if $s + r < 1$, then the expected number of trees diminishes towards zero as time elapses. If $s + r = 1$ then the expected number of trees stays the same, but the trees disperse and the density thins out as time elapses. If $s + r > 1$ then the expected number of trees grows without bound over time and there is potential for long-term overcrowding. This depends on the dispersal γ; narrow dispersal (meaning $\gamma \approx \delta_0$) encourages faster overcrowding while wide dispersal slows the rate of overcrowding.

Example 5.2. The transfunction model

$$\Phi(\mu) = g \cdot ((f_{\#}\mu) * \gamma) \tag{5.2}$$

describes how a population μ will migrate via a function $f : X \to X$ to become $f_{\#}\mu$, disperse by convolution with measure γ from territorial behavior/offspring to become $(f_{\#}\mu) * \gamma$, and grow/shrink locally via environmental factors $g : X \to [0, \infty)$ (which accounts for food, water, shelter, predators, etc) to become $g \cdot ((f_{\#}\mu) * \gamma)$ after some set amount of time.

Example 5.3. The discrete logistic growth model of an immobile population (e.g. trees) with location-dependent carrying capacity can be described by the transfunction

$$\Phi(\mu) = (\mu * \gamma) + r\left(1 - \frac{d(\mu * \gamma)}{d\nu}\right)(\mu * \gamma) \tag{5.3}$$

where μ is a given population which disperses via probability measure γ to become $\mu * \gamma$, then grows or shrinks according to spatially heterogeneous carrying capacity (measure) ν and growth factor $r > 0$. The function $d(\mu * \gamma)/d\nu$ is the Radon-Nikodym derivative of $\mu * \gamma$ with respect to ν.

Although (5.3) resembles a logistic difference equation (see Section 2.1 in [28]), the transfunction model in Example 5.3 incorporates spatially heterogeneous carrying capacity ν. Figure 5.1 provides an example of long-term behavior of (5.3) computed via MATLAB.

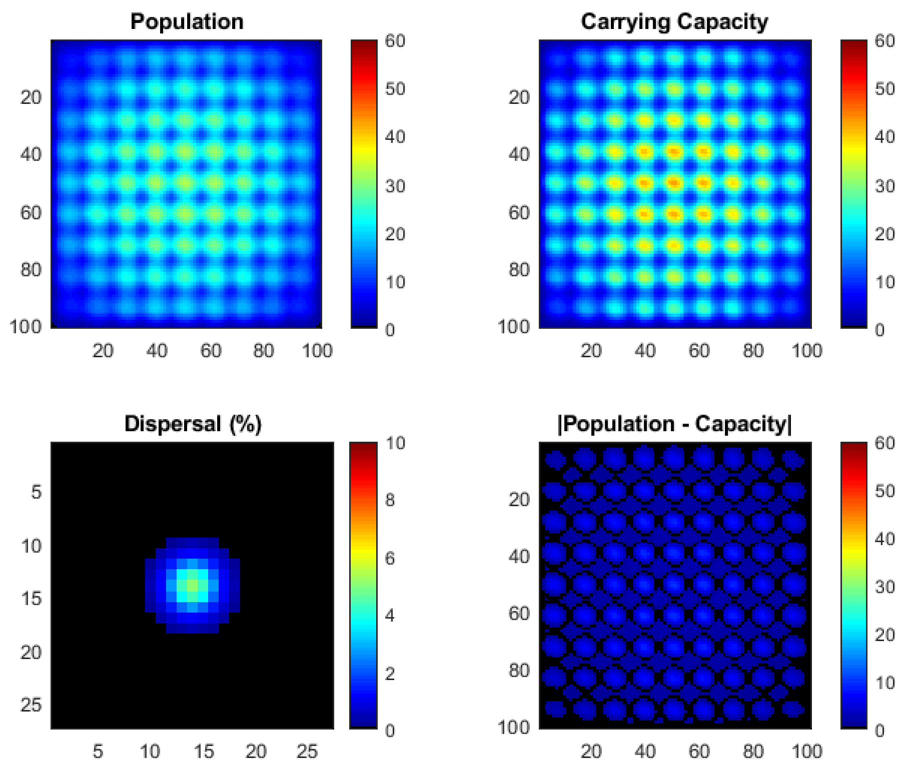

Figure 5.1: Initial population is one tree at the central patch within a 101×101 grid. After a century of time elapses, the population stabilizes (top left) near the carrying capacity (top right) with absolute difference (bottom right) via model (5.3) with dispersal (bottom left).

Example 5.4. A tree population can be described by an l^1-valued vector measure

$$\mu(S) = (\mu_0(S), \mu_1(S), \mu_2(S), \dots), \tag{5.4}$$

where $\mu_j(S)$ represents the number of trees in region S that are j years old. Note that for every $j \in \mathbb{N}_0$, μ_j is a positive measure. Suppose that every year, each j-year old tree (for $j \geq 1$) disperses an average of p_j viable seeds in a distribution γ_j, then survives to become a $(j+1)$-year old tree (the right shift operator A represents aging) with probability v_j. Denote $V := (v_j)$. Then the population of trees changes as time elapses via the linear transfunction

$$\Phi(\mu) = AV\mu + e_0 \sum_{j=1}^{\infty} (p_j \mu_j) * \gamma_j. \tag{5.5}$$

If $(v_j) = (v)$ and $(p_j) = (p)$ are constant sequences (i.e. survivorship and fecundity are age-independent) and if $\mu(\mathbb{R}^2) = e_k$ (i.e. there is only one tree of age k), then it follows that

$$(\Phi^n \mu)_j (\mathbb{R}^2) = \begin{cases} v^j p (v+p)^{n-1-j} & j < n \\ v^n; \quad j = n+k \\ 0; \quad \text{otherwise} \end{cases} \tag{5.6}$$

which yields $||\Phi^n \mu|| = (v+p)^n$. Note that $(\Phi^n \mu)_j$ is proportional to $(v+p)^n$ as $n \to \infty$ for all $j \geq 0$. If $v + p < 1$, trees of each age will become extinct as time passes. When $v + p = 1$, the expected number of trees of each age will remain the same but the trees will disperse as time elapses. When $v + p > 1$, the expected number trees of each age will grow without bound. By linearity of (5.5), having any initial finite positive μ will lead to the same long-term behavior in the cases $v + p < 1$, $v + p = 1$ and $v + p > 1$.

When (v_j) and (p_j) are age-dependent, then $\max\{v_j\} + \max\{p_j\} < 1$ will imply extinction of trees of all ages and $\min\{v_j\} + \min\{p_j\} > 1$ will imply unbounded growth of trees of all ages. Figure 5.2 provides an example of long-term behavior of (5.5) computed via MATLAB.

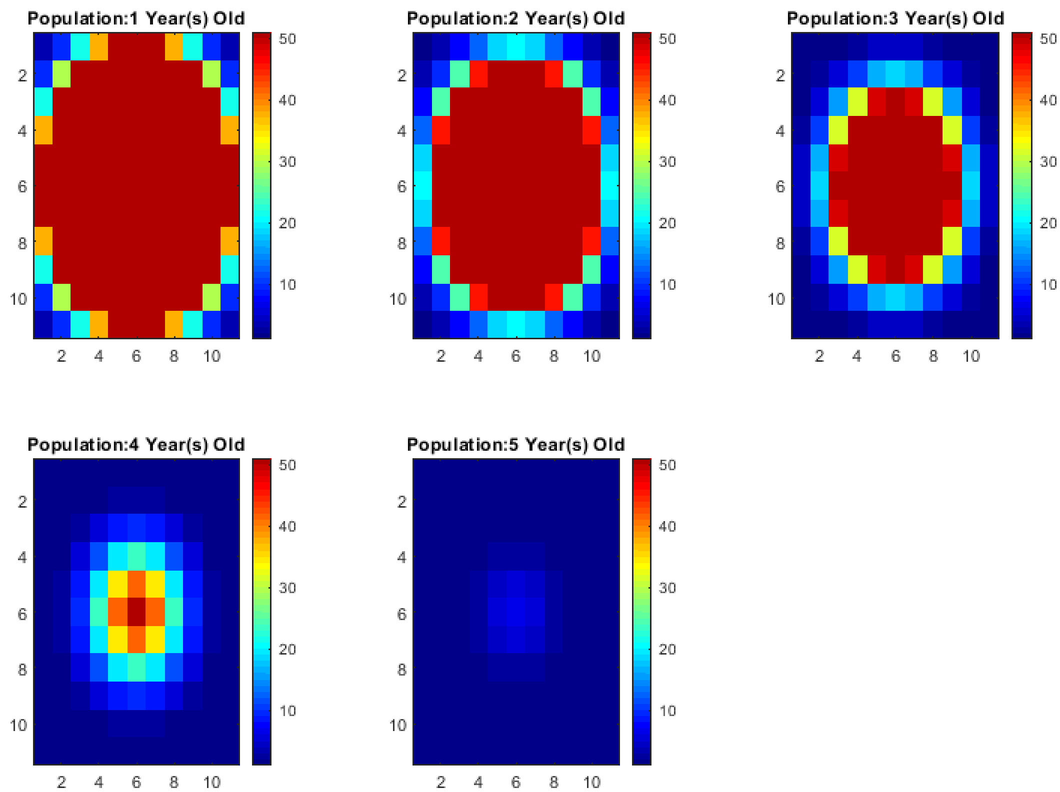

Figure 5.2: Initial population of 10 trees at the center. As trees age, their seed dispersal becomes wider, survivorship worsens with guaranteed death before 6 years, and fecundity is quadratic in age with maximum at 3 years of age and no seed dispersal for 1 or 5 year-old trees. After 180 years, overcrowding develops for ages between 1 and 3 years old. As more time elapses, overcrowding also develops for ages 4 and 5.

The model from Example 5.4 can be generalized so that (v_j) and (p_j) are vectors with bounded measurable function components which contain local behaviors that influence survivorship and fecundity, respectively.

Example 5.5. Now we combine the concepts from Examples 5.3 and 5.4 while generalizing further. Define

$$\Phi(\mu) = AV\mu + e_0 \sum_{j=1}^{\infty} \left[f\left(\frac{d|Q\mu|}{d\kappa}\right) r_j \mu_j \right] * \gamma_j, \tag{5.7}$$

where in addition to previously defined $A, V, \kappa, (r_j)$ and (γ_j), we also introduce a function f which represents local logistic growth and a transfunction Q which represents the usage of resources by μ. The function $d|Q\mu|/d\kappa$ represents the (spatially heterogeneous) proportion of resources used by μ over the resources available by κ, which is then composed into $f : [0, \infty) \to [0, \infty)$ to locally scale the fecundity of μ.

Example 5.6. If there are two species which compete for the same resources, then one can follow similar reasoning from Example 5.5 to develop the pair of transfunctions

$$\begin{cases} \Phi(\mu) = AV_\Phi \mu + e_0 \sum_{j=1}^{\infty} \left[f_\Phi \left(\frac{d|Q\mu + Q'\mu'|}{d\kappa} \right) r_j \mu_j \right] * \gamma_j \\ \Psi(\mu') = AV_\Psi \mu' + e_0 \sum_{j=1}^{\infty} \left[f_\Psi \left(\frac{d|Q\mu + Q'\mu'|}{d\kappa} \right) r'_j \mu'_j \right] * \gamma'_j \end{cases} \tag{5.8}$$

which model how they change over time. Here, $\mu, V_\Phi, f_\Phi, Q, (r_j)$ and (γ_j) pertain to the first species while $\mu', V_\Psi, f_\Psi, Q', (r'_j)$ and (γ'_j) pertain to the second species. Note that $d|Q\mu + Q'\mu'|/d\kappa$ represents the combined and competitive usage of resources.

Example 5.7. To generalize Example 5.6 to n competing species, we have the transfunctions

$$\Phi(\mu^i) = AV^i \mu^i + e_0 \sum_{j=1}^{\infty} \left[f^i \left(\frac{d|\sum_{k=1}^{n} Q^k \mu^k|}{d\kappa} \right) r^i_j \mu^i_j \right] * \gamma^i_j, \quad 1 \leq i \leq n. \tag{5.9}$$

LIST OF REFERENCES

[1] G. Monge, *Mémoire sur la théorie des déblais et des remblais*, Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année (1781), 666–704.

[2] J.B. Fourier, *Théorie analytique de la chaleur*, Paris: Firmin Didot, père et fils (1822).

[3] G. Green, *An Essay on the Application of Mathematial Analysis to the Theories of Electricity and Magnetism*, Nottingham, England: T. Wheelhouse (1828).

[4] S. Soboleff, *Méthode nouvelle à résoudre le problème de Cauchy pour les équations linéaires hyperboliques normales*, Matematicheskii Sbornik (1936), Vol 1(43) No 1, 39–72.

[5] M. F. Gardner, J. L. Barnes, *Transients in Linear Systems studied by the Laplace Transform*, New York: Wiley (1942).

[6] L. Schwartz, *Théorie des distributions I-II*, Paris: Hermann (1951).

[7] H. Kuhn, *The Hungarian Method for the Assignment Problem*. Naval Research Logistics Quarterly (1955), Vol 2, 83-97.

[8] M. Sato, *Theory of Hyperfunctions I*, Journal of the Faculty of Science, University of Tokyo Sect. 1, Mathematics, astronomy, physics, chemistry (1959), Vol 8 No (1), 139–193.

[9] L. Kantorovich, *Mathematical Methods of Organizing and Planning Production*, Management Science (1960), Vol 6 No 4, 366-422.

[10] J. R. Brown, *Approximation Theorems for Markov Operators*. Pacific Journal of Mathematics (1966), Vol 16 No 1, 13-23.

[11] T. K. Boehme, *The support of Mikusiński operators*, Trans. Amer. Math. Soc. 176 (1973), 319–334.

[12] H. Graham Flegg, *Mikusinski's Operational Calculus*, International Journal of Mathematical Education in Science and Technology (1974), Vol 5 No 2, 131-137.

[13] J. Mikusiński, P. Mikusiński, *Quotients de suites et leurs applications dans l'analyse fonctionnelle*, C. R. Acad. Sci. Paris Sr. I Math. 293 (1981), 463-464.

[14] J. Mikusiński, *Hypernumbers, Part I. Algebra*, Studia Math. 77 (1983), 3–16.

[15] P. Mikusiński, *Convergence of Boehmians*, Japan. J. Math. (N.S.) 9 (1983), 159–179.

[16] D. Nemzer, *The Laplace transform on a class of Boehmians*, Bull. Austral. Math. Soc. 46 (1992), 347–352.

[17] P. Jiménez Guerra, B. Rodríguez-Salinas, *A General Solution of the Monge–Kantorovich Mass-Transfer Problem*. Journal of Mathematical Analysis and Applications (1996), Vol 202 No 2, 492-510.

[18] V. Karunakaran, N.V.Kalpakam, *Boehmians and Fourier transforms*, Integral Transforms and Special Functions (2000), Vol 9 No 3, 197–216.

[19] P. K. Banerji, D. Loonker, *Generalized Schwartz distribution: the Boehmians*, J. Rajasthan Acad. Phys. Sci. (2002), Vol 1, 17–24.

[20] L. Ambrosio, *Optimal Transport Maps in Monge-Kantorovich Problem*, Proc. International Conference of Mathematicians, Beijing (2002), Vol 3, 131-140.

[21] J.J. Buckley, E. Eslami, *Fuzzy Functions*, In: *An Introduction to Fuzzy Logic and Fuzzy Sets*. Advances in Soft Computing, Springer-Verlag Berlin Heidelberg (2002), Vol 13.

[22] J. Burzyk, P. Mikusinski, *A characterization of distribution of finite order in the space of Boehmians*, Integral Transform and Special Functions (2005), Vol 16, 639–646.

[23] R. Roopkumar, *Stieltjes transform for Boehmians*, Integral Transforms and Special Functions (2007), Vol 18, 819–827.

[24] C. Villani, *Optimal Transport: Old and New*. Springer-Verlag Berlin Heidelberg (2009).

[25] P. Mikusiński, M.D. Taylor, *Markov operators and n-copulas*. Annales Polonici Mathematici (2009), Vol 96 No 1.

[26] M. D. Taylor, P. Mikusiński, *Theory of $\mathfrak{G}$-Morphisms*, Prepared Manuscript (2009).

[27] P. Mikusinski, *Boehmians and pseudoquotients*, Appl. Math. Inf. Sci. (2011), Vol 5, 192–204.

[28] F. Brauer, C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, 2nd Ed., Texts in Applied Mathematics, Springer (2012).

[29] D. Cohn, *Measure Theory*, 2nd Ed., Birkhäuser, Basel (2013), 213, 245 – 246.

[30] D. Atanasiu, P. Mikusiński, *An elementary operational calculus*, The Mathematical Gazette (2015), Vol 99 No 544, 54-59.

[31] S. K. Q. Al-Omari, *On the extended Fourier transform of generalized functions*, J. Concr. Appl. Math. (2015), Vol 13, 219–231.

[32] P. Mikusiński, *Transfunctions*. Prepared Manuscript (2015), arXiv:1507.03441.

[33] J. Goldstein, *Semigroups of Linear Operators & Applications*, 2nd Ed., Dover Publications, New York (2017).

[34] J. Bentley, P. Mikusiński, *Localized Transfunctions*. International Journal of Applied Mathematics (2018), Vol 31 No 6, 689-707.

[35] S. Campi, P. Gronchi, P. Salani, *A proof of a Loomis–Whitney type inequality via optimal transport*. Journal of Mathematical Analysis and Applications (2019), Vol. 471 No. 1–2, 489-495.

[36] J. Bentley, *Construction of Regular Non-Atomic Strictly-Positive Measures in Second-Countable Locally Compact Non-Atomic Hausdorff Spaces*, Prepared Manuscript (2020), arXiv:1807.05437.

[37] J. Bentley, P. Mikusiński, *Transfunctions and their connections to Plans, Markov Operators and Optimal Transport*, Prepared Manuscript (2020), arXiv:1810.08349.