Limits, Regularity and Removal for Relational and Weighted Structures

by

Ashwini Aroskar

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Mathematical Sciences
Pittsburgh, Pennsylvania 15213

Thesis Committee:
James Cummings (Chair)
Jeremy Avigad
Steven Awodey
Simcha Haber

May 2012

Abstract

The Szemerédi Regularity Lemma states that any graph can be well-approximated by graphs that are almost random. A well-known application of the Szemerédi Regularity Lemma is in the proof of the Removal Lemma for graphs. There are several extensions of the Regularity Lemma to hypergraphs. Our work builds on known results for k-uniform hypergraphs including the existence of limits, a Regularity Lemma and a Removal Lemma.

Our main tool here is a theory of measures on ultraproduct spaces which establishes a correspondence between ultraproduct spaces and Euclidean spaces. We show the existence of a limit object for sequences of relational structures and as a special case, we retrieve the known limits for graphs and digraphs. We also state and prove a Regularity Lemma, a Removal Lemma and a Strong Removal Lemma for relational structures. The Strong Removal Lemma deals with the removal of a family of relational structures and has applications in property testing.

We have also extended the above correspondence to measurable functions on the ultraproduct and Euclidean spaces. This enabled us to find limit objects for sequences of weighted structures and these can be seen as generalizations of the limits we have obtained for relational structures. We also formulate and prove Regularity and Removal Lemmas for weighted structures.

Acknowledgements

My time at Carnegie Mellon has been one of the most rewarding experiences of my life, both professionally and personally. I have met incredible people who have changed my life and shaped my career. I wish to thank them all for making the work in this thesis possible.

Foremost, I would like to express my deepest gratitude to my advisor, Prof. James Cummings for his infallible guidance, immense knowledge and infinite patience. With his vast experience and encouraging nature, he has been an incredible mentor who has shaped me into the mathematician I am today. I admire his insatiable appetite for learning and I would be honored to be an academic like him some day.

I would also like to thank the rest of my thesis committee – Prof. Jeremy Avigad, Prof. Steven Awodey and Prof. Simcha Haber – for their insightful comments and the encouragement they have given me. I am also grateful to Prof. Oleg Pikhurko for helpful discussions and suggestions regarding interesting combinatorial problems to pursue.

The Mathematics Department is fortunate to have a wonderful administrative staff in the form of Stella Andreoletti, Jeff Moreci, Patsy McCarthy, Ferna Hartman, Nancy Watson, Florin Manoleche and Adrian Dumitrasc. I am grateful for their help and assistance.

I am lucky to have many wonderful friends who have always taken great pride in my achievements and I am grateful for their support. I am especially thankful for my friends at CMU who have navigated the doctoral path with me. We have been there for each other, shared our worries and soothed away anxieties. Pittsburgh wouldn't have seemed half as fun without my comrades-in-arms.

I owe a debt of gratitude to my 'cricket team' back home. My family has been a limitless source of love and support through the years. Nurturing parents are the most influential force in a child's life and in this regard, I have been doubly blessed. None of my successes would have meant anything if it hadn't been for the joy and pride in their eyes. I wouldn't have had the strength to overcome my shortcomings if it had not been for their encouraging words. My special thanks go to my foremost champion – my mother – who always dreamed big for me and never rested till she gave me everything she possibly could to help me achieve my goals. As I fulfil her biggest dream, I thank her from my heart for her selflessness and I want her to know she is truly appreciated.

I also want to thank my grandmother who taught me to be the person I am today. She is a woman who has always been ahead of her times and it is from her that I have learnt all my core values. Her inner strength has been a lifelong inspiration to me. To her, I dedicate my thesis.

Last but by no means the least, I would like to thank the man who has been by my side every step of this journey. He believes in me when I've lost faith in myself. He keeps me calm when I'm closest to despair. He is my pillar of strength. He is my rock. To him, I dedicate all my success.

Siddharth Shenoy, the man I love,

I am forever grateful to have you in my life.

Contents

Chapter 1

Introduction

The Szemerédi Regularity Lemma [15] is an important tool in graph theory. It states that all large graphs can be well-approximated, in a precise sense, by random graphs. It is extremely useful in proving theorems for large graphs when the corresponding statement for random graphs can be easily verified. Szemerédi first invented this result as a lemma in his proof of a famous conjecture of Erdös and Turán involving arithmetic progressions.

A well-known application of this Regularity Lemma is in the proof of the Triangle Removal Lemma, and more generally the Graph Removal Lemma. The Triangle Removal Lemma states that if a large graph is nearly triangle-free, then we can delete a small number of edges to remove all triangles. Similarly, the Graph Removal Lemma asserts that given a fixed graph F, any large graph that has very few copies of F can be made F-free by deleting a small number of edges.

Many variants of the Regularity Lemma as well as generalizations have since been proved. The most prominent of these is the Hypergraph Regularity Lemma proved independently by Rödl-Skokan [14] and Gowers [7].

The Hypergraph Regularity Lemma can also be used to prove a removal lemma for hypergraphs, first proved by Nagle, Rödl and Schacht [11].

Benjamini and Schramm [2] constructed limit objects for sequences of graphs with bounded degree. This was extended to sequences of graphs with bounded average degree by Lyons [10].

Lovász and Szegedy [9] showed that a "limit" exists for any convergent sequence of graphs. A sequence of finite graphs G_n is *convergent* if $|V(G_n)| \to \infty$ and for every finite graph F, the density of homomorphic copies of F in G_n, denoted as $t(F, G_n)$, converges. A limit object for a convergent sequence of graphs is such that it determines all the limits of subgraph densities in the sequence. Lovász and Szegedy showed that a natural limit in the case of graphs is a *graphon*, which is a symmetric measurable function $W : [0,1]^2 \to [0,1]$. For instance, a limit of a sequence of random graphs with density p, $p \in [0,1]$, is the constant function $W = p$.

In general, given $n \in \mathbb{N}$ and any finite graph F on $[n]$,

$$t(F, W) = \int_{[0,1]^n} \cdots \int \prod_{(i,j)\in E(F)} W(x_i, x_j)\, d(x_1, \ldots, x_n)$$

and $\lim_{n\to\infty} t(F, G_n) = t(F, W)$.

Conversely, they show that every graphon arises as a limit of some convergent sequence of graphs. Their work introduces a random graph construction by sampling from the graphon W which is quite useful.

Non-standard analysts have developed measures on ultraproducts. Given a sequence of finite sets of increasing size, equipped with the normalized counting measure, we can obtain an ultraproduct measure space using the *Loeb measure* construction [8].

Elek and Szegedy [6, 5] built separable algebras on ultraproducts and using an instance of Maharam's theorem, they showed these separable spaces are equivalent to the standard Lebesgue measure space. This established an equivalence between ultraproduct spaces with Loeb measure and Euclidean spaces with the familiar Lebesgue measure. They used this to prove a correspondence between sequences of k-uniform hypergraphs on finite sets, a measurable k-uniform hypergraph on the ultraproduct and a *Euclidean hypergraph*, which is an $S_{[k]}$ symmetric Lebesgue-measurable subset of $[0,1]^{2^k-1}$. In this context, a graph limit is a Lebesgue-measurable subset of $[0,1]^3$. However, we can use this new graph limit to recover the graphons we mentioned above.

We can convert theorems of finite combinatorics to measure-theoretic statements on the ultraproduct spaces. Using *separable realizations*, we can translate these measure-theoretic statements to well-known facts about the Lebesgue measure. Elek and Szegedy used this correspondence to prove a Regularity Lemma, a Removal Lemma and a Strong Removal Lemma for k-uniform hypergraphs. The Strong Removal Lemma has applications in property testing and has been used to show that *hereditary* properties are testable [1] [13].

Our work builds on the results of Elek and Szegedy that we have described above. In Chapter 2, we introduce the theory of measures on ultraproduct spaces used by Elek and Szegedy. We give alternate proofs of theorems involving the Loeb measure on the ultraproduct, using $\aleph_1$ *saturation.* This gives us a cleaner look at the ultraproduct mechanism behind the first piece of the Correspondence Principle we want to prove. Any measurable subset of the ultraproduct differs from an internal set by a null set. In the case of functions, a bounded measurable function on the ultraproduct differs on a null set from the standard part of a bounded internal function. The standard part function is necessary in this correspondence because the ultralimit of a sequence of bounded measurable functions on finite sets yields a function on the ultraproduct that takes values that are non-standard reals. Each non-standard real is infinitesimally close to a real number that is returned by the standard part function.

Section 2.4 presents the concept of separable realizations developed by Elek and Szegedy that is critical to the second piece of the Correspondence Principle. The measure algebra constructed on the ultraproduct in Section 2.1 is non-separable. Separable realizations help to establish an equivalence between separable algebras on ultraproducts with the Loeb measure and the familiar Euclidean spaces equipped with the Lebesgue measure.

Elek and Szegedy's work studies k-uniform hypergraphs and the natural question to ask is whether their results generalize to other structures, such as directed hypergraphs. In Chapter 3, we examine many interesting combinatorial results in the more general setting of relational structures. A relational structure consists of finitely many relations on an underlying set. We first extend the Elek-Szegedy Correspondence Principle to relational structures. For technical reasons, we represent relational structures by a system of directed hypergraphs. Then we establish a correspondence between :

- A sequence of relational structures on finite sets that are increasing in size
- A relational structure on the ultraproduct such that each directed hypergraph in its representation is a measurable set
- A system of Lebesgue-measurable subsets that we will refer to as a *Euclidean structure*

We define what it means to be a limit object for relational structures on finite sets and using the above correspondence, we see that limits of *convergent* sequences of relational structures are Euclidean structures as described above. We further show that any such Euclidean structure is a limit of a sequence of relational structures on finite sets. These limit objects generalize the Elek-Szegedy limits for k-uniform hypergraphs, so in particular we can retrieve graphons as graph limits. We also recover digraph limits in the style of Offner and Pikhurko [12]. Their limit object for a sequence of finite digraphs is a set of four graphon-like functions that satisfy certain conditions. In Section 3.1, we shall describe these digraph limits further and show how we obtain the four measurable functions from our limits. We also discuss uniqueness of our limits.

Elek and Szegedy used familiar facts about the Lebesgue measure to prove Regularity and Removal Lemmas for k-uniform hypergraphs. In a similar manner, we use our Correspondence Principle to extend these results to relational structures. The Regularity Lemma that we formulate and prove for relational structures in Section 3.2 depends on the fact that Lebesgue measurable sets can be well-approximated by a union of uniform hypercubes. Our Removal Lemma relies on the Lebesgue Density theorem. We also prove a Strong Removal Lemma that deals with simultaneously removing all copies of a family of relational structures from a large relational structure by changing a small number of relations. This is possible if the original large structure has only a few copies of some initial subset of the family. We shall make this more precise in Section 3.3. We also describe the applications of the Strong Removal Lemma in the area of property testing.

A more challenging question to answer is regarding the generalization of these techniques and results to the case of weighted structures. A *weighted structure* consists of finitely many bounded weight functions on an underlying set. The Correspondence Principle above involved a correspondence between measurable sets on the three spaces. We prove a similar correspondence between bounded measurable functions on these spaces. This led to a Weighted Correspondence Principle relating

- A sequence of weighted structures on finite sets
- A weighted structure on the ultraproduct such that the bounded weight functions in

its representation are measurable

- A system of Lebesgue-measurable bounded weight functions that we call a *Euclidean weighted structure*

We can now use this correspondence in a similar manner as the previous correspondence in the case of relational structures. In Chapter 4, we define the notion of homomorphisms for weighted structures. We define a notion of limits for *convergent* sequences of weighted structures on finite sets and show that Euclidean weighted structures are limits of such sequences. We also formulate and prove a Regularity Lemma that states all weighted structures on large sets are "close" to having a nice uniform structure. We make this precise in Section 4.3. In Section 4.4, we also formulate and prove a Removal Lemma for weighted structures. If we represent relational structures from Chapter 3 as 0 -1 weighted structures, then our results for weighted structures in Chapter 4 can easily be seen as extensions of analogous results for relational structures.

Regularity Lemmas for graphs usually define regularity in terms of edge densities of subgraphs. Csaba and Pluhár [4] proved a Weighted Regularity Lemma for weighted digraphs where regularity is defined with respect to weighted edge densities. We prove a similar result for weighted digraphs using our Weighted Correspondence Principle.

Chapter 2

Ultraproduct Spaces

In this chapter, we introduce the theory of measures on ultraproduct spaces used by Elek and Szegedy. We give alternate proofs of theorems involving the Loeb measure on the ultraproduct, using Łoś' theorem and $\aleph_1$-saturation. We present their main results on the measure theory of ultraproducts here, most of them without proof, in order to develop the key ideas necessary for the central correspondence.

2.1 Measure Algebras on Ultraproduct Spaces

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of finite sets of increasing size and U be a non-principal ultrafilter on ω. For the ease of proofs that follow, we will often work within the ultrapower $\mathbf{V}$ of the universe V of set theory, modulo the ultrafilter U. Take the ultraproduct $\mathbf{X} = (\prod_i X_i)/U = [X_i]$ of this sequence modulo U. Given $A_i \subseteq X_i$ for each $i \in \omega$, $[A_i]$ is an internal subset of $\mathbf{X}$. Internal sets form a Boolean algebra $\boldsymbol{\mathcal{A}}$ over $\mathbf{X}$.

It is useful to recall the following important theorem about ultraproducts.

Theorem 2.1.1 (Łoś' theorem). Given a first-order language L and an ω-sequence of

L-structures M_i, let $\mathbf{M} = (\prod_i M_i)/U$. Given a first-order L-formula $\phi(x_1, \ldots, x_k)$ and $\mathbf{a}_1, \ldots, \mathbf{a}_k \in |\mathbf{M}|$,
$\mathbf{M} \models \phi(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ if and only if $M_i \models \phi(a_{1,i}, \ldots, a_{k,i})$ for U-almost every i, that is, $\{i : M_i \models \phi(a_{1,i}, \ldots, a_{k,i})\} \in U$. $\square$

Another useful fact is that as an ultrapower, $\mathbf{V}$ is $\aleph_1$-saturated, that is, every countable type of first-order formulae over $|\mathbf{V}|$ is realized in $\mathbf{V}$.

If each X_i is equipped with the normalized counting measure μ_i, then for any internal set $\mathbf{A} = [A_i]$, we can define $\nu(\mathbf{A}) = [\mu_i(A_i)]$. ν takes values in $[0,1]^{\mathbf{V}}$ and ν is finitely additive on the algebra of internal sets. In fact, using Łoś' theorem we see that $\mathbf{V}$ thinks $\mathbf{X}$ is finite and ν is the counting measure on $\mathbf{X}$.

Given $b \in \mathbb{R}$, let $b^* = [\{b\}_{i=1}^{\infty}]$.

Define $\mathbb{S} = \bigcup_{m=1}^{\infty} [-m, m]^{\mathbf{V}} = \bigcup_{m=1}^{\infty} [-m^*, m^*] \subsetneq \mathbb{R}^{\mathbf{V}}$. Since $[-m, m]$ is compact for every $m \in \mathbb{N}$, given any $\mathbf{a} \in \mathbb{S}$, there is a unique $b \in \mathbb{R}$ such that for all $\epsilon > 0$, $\{i : |a_i - b| < \epsilon\} \in U$. In other words, $b = \lim_U a_i$.

We define the *standard part* function by $std(\mathbf{a}) = b$.

It is easy to see that $std(b^*) = b$ and that the standard part function commutes with mathematical operations, such as, addition, multiplication, absolute value. Given $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{S}$ such that $\mathbf{a}_1 \leq \mathbf{a}_2$, $std(\mathbf{a}_1) \leq std(\mathbf{a}_2)$.

Let $c, d \in \mathbb{R}$ and let $\mathbf{a} \in \mathbb{S}$. If $c \leq std(\mathbf{a}) \leq d$, then $(c - \frac{1}{n})^* \leq \mathbf{a} \leq (d + \frac{1}{n})^*$, that is, $\{i : c - \frac{1}{n} \leq a_i \leq d + \frac{1}{n}\} \in U$, for each $n \in \mathbb{N}$.

Define $\mu = std \circ \nu$. Then for every internal set $\mathbf{A} = [A_i]$, $\mu(\mathbf{A}) = \lim_U \mu_i(A_i)$. μ is real-valued and finitely additive on the algebra of internal sets.

$\mathbf{N} \subseteq \mathbf{X}$ is a *null set* if it can be covered by internal sets of arbitrarily small μ-measure. Define $\boldsymbol{\mathcal{B}} = \{\mathbf{A}\Delta\mathbf{N}|\mathbf{A}$ is an internal set, $\mathbf{N}$ is a null set$\}$.

Lemma 2.1.2. *$\boldsymbol{\mathcal{B}}$ is a σ-algebra over $\mathbf{X}$ and μ extends as a σ-additive measure over $\boldsymbol{\mathcal{B}}$, where $\mu(\mathbf{A}\Delta\mathbf{N}) = \mu(\mathbf{A})$.*

Proof. It is enough to prove that for a sequence of disjoint internal sets $\mathbf{A}_n$, $\bigcup_{n=1}^{\infty} \mathbf{A}_n \in \boldsymbol{\mathcal{B}}$ and $\mu(\bigcup_{n=1}^{\infty} \mathbf{A}_n) = \sum_{n=1}^{\infty} \mu(\mathbf{A}_n)$. Let $a = \sum_{n=1}^{\infty} \mu(\mathbf{A}_n)$. (Since the internal sets $\mathbf{A}_n$ are disjoint, μ is finitely additive on internal sets and $\mu(\mathbf{X}) = 1$, the bounded partial sums converge.)

We define for each $n \in \mathbb{N}$, the first-order formula $\phi_n(\mathbf{B})$ as

$$(\mathbf{B} \in \boldsymbol{\mathcal{A}}) \wedge \left(\bigcup_{k=1}^{n} \mathbf{A}_k \subseteq \mathbf{B}\right) \wedge \left(\nu(\mathbf{B}) \leq \left(a + \frac{1}{n}\right)^*\right)$$

We claim that $\{\phi_n(\mathbf{B}) : n \in \omega\}$ is a countable type over $|\mathbf{V}|$. Take any finite subset and let n be the largest natural number such that $\phi_n(\mathbf{B})$ is in this subset. All formulae in this subset are now easily satisfied by $\mathbf{B} = \bigcup_{k=1}^{n} A_k$.

As the ultrapower is $\aleph_1$-saturated, this type is satisfiable in $\mathbf{V}$. There must exist $\mathbf{B} \in \mathbf{A}$ such that $\bigcup_{n=1}^{\infty} \mathbf{A}_n \subseteq \mathbf{B}$ and $\nu(\mathbf{B}) \leq (a + \frac{1}{n})^*$ for each n.
So $\mu(\mathbf{B}) = std \circ \nu(\mathbf{B}) \leq a + \frac{1}{n}$ for each n. Thus, $\mu(\mathbf{B}) \leq a$.

Also, $\bigcup_{k=1}^{n} \mathbf{A}_k \subseteq \mathbf{B}$ implies $\sum_{k=1}^{n} \mu(\mathbf{A}_k) \leq \mu(\mathbf{B})$, for each n. We can now conclude $\mu(\mathbf{B}) = a = \sum_{n=1}^{\infty} \mu(\mathbf{A}_n)$.

Consider the sequence of internal sets $\mathbf{C}_n = \mathbf{B} \setminus (\bigcup_{k=1}^{n} \mathbf{A}_k)$, $n \in \mathbb{N}$. Then $\mu(\mathbf{C}_n) = a - \sum_{k=1}^{n} \mu(\mathbf{A}_k) \to 0$ as $n \to \infty$. Each $\mathbf{C}_n$ covers $\mathbf{B} \setminus \bigcup_{n=1}^{\infty} \mathbf{A}_n$. Since we can now cover $\mathbf{B} \setminus \bigcup_{n=1}^{\infty} \mathbf{A}_n$ by internal sets of increasingly small measure, $\mathbf{B} \setminus \bigcup_{n=1}^{\infty} \mathbf{A}_n$ is a null set and $\bigcup_{n=1}^{\infty} \mathbf{A}_n \in \boldsymbol{\mathcal{B}}$. Also $\mu(\bigcup_{n=1}^{\infty} \mathbf{A}_n) = \mu(\mathbf{B}) = \sum_{n=1}^{\infty} \mu(\mathbf{A}_n)$.

This shows that $\boldsymbol{\mathcal{B}}$ is closed under countable unions and $\boldsymbol{\mathcal{B}}$ is a σ-algebra with the countably additive measure μ. $\square$

Suppose $f_i : \mathbf{X} \to [-M, M]$ for U-almost every i. We say $\mathbf{f} = [f_i] : \mathbf{X} \to [-M^*, M^*]$ is a *bounded internal function*. Then $std \circ \mathbf{f} : \mathbf{X} \to [-M, M]$ is a bounded real-valued function.

Lemma 2.1.3. *Let $\mathbf{f} = [f_i] : \mathbf{X} \to [-M^*, M^*]$ be a bounded internal function. Then $std \circ \mathbf{f}$ is $\boldsymbol{\mathcal{B}}$-measurable, and*

$$\int_{\mathbf{X}} (std \circ \mathbf{f})\, d\mu = std\left(\int_{\mathbf{X}} \mathbf{f}\, d\nu\right) = std\left(\left[\int_{X_i} f_i\, d\mu_i\right]\right)$$

Proof.

$$(std \circ \mathbf{f})^{-1}([a, b]) = \bigcap_{n=1}^{\infty} \mathbf{f}^{-1}\left(\left[\left(a - \frac{1}{n}\right)^*, \left(b + \frac{1}{n}\right)^*\right]\right)$$

The intersection of countably many internal subsets of $\mathbf{X}$ is in $\boldsymbol{\mathcal{B}}$, so $std \circ \mathbf{f}$ is a $\boldsymbol{\mathcal{B}}$-measurable function.

We can approximate $std \circ \mathbf{f}$ using $\boldsymbol{\mathcal{B}}$-measurable simple functions $g_n = \sum_{k=1}^{m_n} b_k\, \chi_{\mathbf{B}_k}$, where each $b_k \in \mathbb{R}$ and $\mathbf{B}_k \in \boldsymbol{\mathcal{B}}$. There are internal sets $\mathbf{A}_k$ such that $\mu(\mathbf{A}_k \Delta \mathbf{B}_k) = 0$.

Define internal functions $\mathbf{h}_n = \sum_{k=1}^{m_n} b_k^*\, \chi^{\mathbf{V}}_{\mathbf{A}_k}$. Then $g_n = std \circ \mathbf{h}_n$ a.e. and we can approximate $std \circ \mathbf{f}$ a.e. using $std \circ \mathbf{h}_n$. That is, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$, $|std \circ \mathbf{f}(x) - std \circ \mathbf{h}_n(x)| < \epsilon$ for almost every x and $\int_{\mathbf{X}} (std \circ \mathbf{f})\, d\mu = \lim_{n\to\infty} \int_{\mathbf{X}} (std \circ \mathbf{h}_n)\, d\mu$.

Also, if $|std(\mathbf{f}(x)) - std(\mathbf{h}_n(x))| < \epsilon$, then for all $m \in \mathbb{N}$, $|\mathbf{f}(x) - \mathbf{h}_n(x)| < (\epsilon + \frac{1}{m})^*$.

Using properties of the standard part function, we see that

$$std\left(\int_{\mathbf{X}} \mathbf{f}\, d\nu\right) = \lim_{n\to\infty} std\left(\int_{\mathbf{X}} \mathbf{h}_n\, d\nu\right).$$

To prove the integral formula above for all bounded internal functions, it would now be enough to prove it for simple internal functions like $\mathbf{h}_n$.

First let us prove the integral formula for $\chi^{\mathbf{V}}_{\mathbf{A}}$, where $\mathbf{A}$ is an internal set. Then $std \circ \chi^{\mathbf{V}}_{\mathbf{A}} = \chi_{\mathbf{A}}$.

$$\int_{\mathbf{X}} \chi_{\mathbf{A}}\, d\mu = \mu(\mathbf{A}) = std(\nu(\mathbf{A})) = std\left(\int_{\mathbf{X}} \chi^{\mathbf{V}}_{\mathbf{A}}\, d\nu\right)$$

As every $\mathbf{h}_n$ above is a finite linear combination of these internal characteristic functions, we can use simple properties of the standard part function to extend the integral formula to functions $\mathbf{h}_n$. □

Lemma 2.1.4. *Given any real-valued bounded $\boldsymbol{\mathcal{B}}$-measurable function $\mathbf{g}$ on $\mathbf{X}$, there is a bounded internal function $\mathbf{f}$ on $\mathbf{X}$ such that $\mathbf{g} = std \circ \mathbf{f}$ a.e.*

Proof. We may assume that $\mathbf{g} : \mathbf{X} \to [0, 1]$. For the purposes of this proof, it is easier to identify $[0, 1]$ with the space of binary sequences with the following topology : The basic open sets are $N_s = \{y : \omega \to \{0, 1\} \mid y \restriction_{l(s)} = s\}$ and the measure is defined by $\lambda(N_s) = 2^{-l(s)}$ for any finite binary sequence s. Given $\mathbf{y} : \omega^{\mathbf{V}} \to \{0, 1\}$, $std(\mathbf{y})$ is simply the restriction of $\mathbf{y}$ to its initial ω-segment.

Since $\mathbf{g}$ is measurable, for every finite binary sequence s, $\mathbf{B}_s = \mathbf{g}^{-1}(N_s) \in \boldsymbol{\mathcal{B}}$. Note that for any n, $\{\mathbf{B}_s | \, l(s) = n\}$ is a partition of $\mathbf{X}$. Furthermore, for any sequence s with length

n and given any $m > n$, $\{\mathbf{B}_t \,|\, l(t) = m,\ t \restriction_n = s\}$ is a partition of $\mathbf{B}_s$. By definition of $\boldsymbol{\mathcal{B}}$, there must exist internal sets $\mathbf{A}_s$ such that $\mu(\mathbf{A}_s \Delta \mathbf{B}_s) = 0$. We may assume that the sets $\mathbf{A}_s$ also form a system of nested partitions of $\mathbf{X}$ as above.

For every finite sequence s, define $\phi_s(\mathbf{f})$ as the first-order formula

$$\forall \mathbf{x} \in \mathbf{A}_s,\ \mathbf{f}(\mathbf{x}) : \omega^{\mathbf{V}} \to \{0,1\} \ \wedge\ \mathbf{f}(\mathbf{x}) \restriction_{l(s)} = s$$

Any finite subset of $\{\phi_s(\mathbf{f})\}$ is easily satisfiable. Take the index s of longest length and build $\mathbf{f}(x)$ as the formulae ask, up to length $l(s)$ (beyond that map to 0).

Since the ultraproduct is $\aleph_1$-saturated and $\phi_s(\mathbf{f})$ is a countable type over $|\mathbf{V}|$, there exists an internal function $\mathbf{f}$ that satisfies each $\phi_s(\mathbf{f})$. We have ensured that $\mathbf{f}^{-1}(N_s^{\mathbf{V}}) = \mathbf{A}_s$. Then $(\mathit{std} \circ \mathbf{f})^{-1}(N_s) = \mathbf{A}_s$. Thus $\mathit{std} \circ \mathbf{f}$ and $\mathbf{g}$ may differ only on $\bigcup_s (\mathbf{A}_s \Delta \mathbf{B}_s)$, which is a null set. So $\mathbf{g} = \mathit{std} \circ \mathbf{f}$ a.e. and $\mathbf{f}$ is clearly a bounded internal function. $\square$

2.2 Fubini's Theorem and Total Independence

Let X_i^j be copies of X_i for $j \in [k]$ and for every $A \subseteq [k]$, let $X_i^A = \prod_{j \in A} X_i^j$ with counting measure $\mu_{i,A}$. Let $r \leq k$ and let e be an ordered r-tuple of distinct elements from $[k]$, then $X_i^e = \prod_{j \in dom(e)} X_i^j$.

For every $A \subseteq [k]$, we may take the ultraproduct $\mathbf{X}^A$ of the sequence of finite sets X_i^A and define ν_A, $\boldsymbol{\mathcal{B}}_A$ and μ_A as before. We can repeat everything above for this ultraproduct and Lemmas 2.1.2, 2.1.3 and 2.1.4 hold. There are natural projection maps $\pi_A : \mathbf{X}^{[k]} \to \mathbf{X}^A$, using which we define $\sigma_A = \pi_A^{-1}(\boldsymbol{\mathcal{B}}_A)$, the σ-algebra on $\mathbf{X}^{[k]}$ depending only on coordinates of A.

Let $\mathbf{f}$ be a bounded internal function on $\mathbf{X}^{[k]}$ that depends only on A-coordinates, it is easy to see that $std \circ \mathbf{f}$ is σ_A-measurable.

On the other hand, if $\mathbf{g}$ on $\mathbf{X}^{[k]}$ is a bounded σ_A-measurable function, then there exists a bounded internal function $\mathbf{f}$ on $\mathbf{X}^{[k]}$ depending only on A coordinates such that $\mathbf{g} = std \circ \mathbf{f}$ a.e.

Lemma 2.2.1. *Let $A, B \subseteq [k]$ and let $\mathbf{g} : \mathbf{X}^{[k]} \to \mathbb{R}$ be a bounded σ_B-measurable function. Then for all $y \in \mathbf{X}^{A^c}$, the function $\mathbf{g}_y : \mathbf{X}^A \to [-M^*, M^*]$, defined by $\mathbf{g}_y(x) = \mathbf{g}(x, y)$, is a bounded $\pi_A[\sigma_{A\cap B}]$-measurable function.*

Equivalently $\mathbf{g}_y \circ \pi_A$ is a $\sigma_{A\cap B}$-measurable function on $\mathbf{X}^{[k]}$.

(Here, $\pi_A[\sigma_{A\cap B}] = \{\pi_A[S] : S \in \sigma_{A\cap B}\}$ is the σ-algebra on $\mathbf{X}^A$.) □

In the future, we will often treat $\mathbf{X}^A$ as a subspace of $\mathbf{X}^{[k]}$ and won't make a distinction between $\boldsymbol{\mathcal{B}}_A$ and σ_A or between $\mathbf{g}_y$ and $\mathbf{g}_y \circ \pi_A$.

We define σ_A^* as the σ-algebra generated by σ_B, for all $B \subsetneq A$ such that $|B| = |A| - 1$.

σ_A^* is, in general, a strictly smaller algebra than σ_A. For instance, if $k = 2$, the σ-algebra on the 2-dimensional ultraproduct space $\mathbf{X}^{[2]}$ is not the product of the two 1-dimensional algebras on either coordinate. Theorem 2.4.1 states there are measurable 2-dimensional sets independent of the σ-algebra of measurable rectangles.

However, a Fubini-like theorem is still true in this setting because Fubini's theorem holds trivially for the finite sets $X_i^{[k]}$.

Theorem 2.2.2. *Let $\mathbf{g}$ be a bounded $\sigma_{[k]}$-measurable real-valued function on $\mathbf{X}^{[k]}$. Then*

1. *for a.e.* $y \in \mathbf{X}^{A^c}$, $\mathbf{g}_y$ *is a* σ_A*-measurable function on* $\mathbf{X}^A$.

2. $y \to \int_{\mathbf{X}^A} \mathbf{g}_y(x)\, d\mu_A(x)$ *is a* σ_{A^c}*-measurable function on* $\mathbf{X}^{A^c}$.

3. *Also* $\int_{\mathbf{X}^{[k]}} \mathbf{g}\, d\mu_{[k]} = \int_{\mathbf{X}^{A^c}} (\int_{\mathbf{X}^A} \mathbf{g}_y(x)\, d\mu_A(x))\, d\mu_{A^c}(y)$

Proof. Part 1 follows from Lemma 2.2.1.

From Lemma 2.1.4, we know that there is a bounded internal function $\mathbf{f}$ such that $\mathbf{g} = std \circ \mathbf{f}$ a.e. Then $\mathbf{g}_y = std \circ \mathbf{f}_y$ a.e. and

$$\int_{\mathbf{X}^A} \mathbf{f}_y(x)\, d\nu_A(x) = \left[\int_{X_i^A} f_{i,y_i}\, d\mu_{i,A} \right],$$

and $y \to \int_{\mathbf{X}^A} \mathbf{f}_y(x)\, d\nu_A(x)$ is a bounded internal function.

Using Lemma 2.1.3, we know $\int_{\mathbf{X}^A} \mathbf{g}_y\, d\mu_A = std(\int_{\mathbf{X}^A} \mathbf{f}_y\, d\nu_A)$, hence proving part 2 above.

We remarked earlier that in $\mathbf{V}$, $\mathbf{X}^A$ appears to be finite and ν_A is the counting measure on $\mathbf{X}^A$. So a Fubini-like statement is trivially true, that is,

$$\int_{\mathbf{X}^{[k]}} \mathbf{f}\, d\nu_{[k]} = \int_{\mathbf{X}^{A^c}} \left(\int_{\mathbf{X}^A} \mathbf{f}_y\, d\nu_A \right) d\nu_{A^c}$$

Now taking standard part and using the integral formula from Lemma 2.1.3, we prove part 3 of our theorem. $\square$

Let $\mathbf{N} \in \sigma_{[k]}$. For any $y \in \mathbf{X}^{A^c}$, define $\mathbf{N}_y = \{x \in \mathbf{X}^A : (x, y) \in \mathbf{N}\}$.

Applying Fubini's theorem to characteristic functions of null sets, in particular, we can make the following observation.

Corollary 2.2.3. *Almost every slice of a set is null if and only if the set is null, that is,* *for all* $\mathbf{N} \in \sigma_{[k]}$, $\mu_{[k]}(\mathbf{N}) = 0$ *if and only if* $\mu_{A^c}(\{y \in \mathbf{X}^{A^c} : \mu_A(\mathbf{N}_y) = 0\}) = 1$. $\square$

We now state some important results from [6, 5], without proof, to illustrate the construction of separable realizations that are critical to the second piece of the correspondence we want to establish.

Let $\mathcal{B} \leq \mathcal{A}$ be σ-algebras and let μ be a countably additive measure on $\mathcal{A}$. For every $\mathcal{A}$-measurable function f, $E(f|\mathcal{B})$ is the *conditional expectation of* f over $\mathcal{B}$. It is the unique $\mathcal{B}$-measurable function (within measure 0) such that $\int_Y E(f|\mathcal{B})\, d\mu = \int_Y f\, d\mu$, for all $Y \in \mathcal{B}$.

The *conditional expectation of a set* $A \in \mathcal{A}$ over $\mathcal{B}$ is $E(A|\mathcal{B}) = E(\chi_A|\mathcal{B})$, where χ_A denotes the characteristic function of set A.

$A \in \mathcal{A}$ is *independent* of $\mathcal{B}$ if $E(A|\mathcal{B})$ is a constant function. Then $E(A|\mathcal{B}) = \mu(A)$ and $\mu(A \cap B) = \mu(A)\mu(B)$ for every $B \in \mathcal{B}$.

Sub-σ-algebras $\mathcal{B}$ and $\mathcal{C}$ of $\mathcal{A}$ are said to be *independent* σ*-algebras* if each $B \in \mathcal{B}$ is independent of $\mathcal{C}$ and vice-versa. Equivalently, $\mu(B \cap C) = \mu(B)\mu(C)$ for each $B \in \mathcal{B}$ and $C \in \mathcal{C}$.

Theorem 2.2.4 (**Total Independence Lemma**). *Let* $A_1, \ldots, A_m$ *be a list of distinct subsets of* $[k]$ *and* $S_i \in \sigma_{A_i}$ *such that* $E(S_i|\sigma^*_{A_i})$ *is constant. Then* $\mu(S_1 \cap \ldots \cap S_m) = \mu(S_1) \ldots \mu(S_m)$. $\square$

2.3 Random Partitions and Independent Complements

Let $\mathcal{B} \leq \mathcal{A}$ be σ-algebras and let μ be a countably additive measure on $\mathcal{A}$. A $\mathcal{B}$*-random* k*-partition* in $\mathcal{A}$ is a partition $A_1, ..., A_k$ of X into $\mathcal{A}$-measurable sets such that $E(A_i|\mathcal{B}) = \frac{1}{k}$

for each i.

Theorem 2.3.1. *Let $A \subseteq [k]$ and $n \in \mathbb{N}$. Then there exists a σ_A^*-random n-partition of $\mathbf{X}^A$ in σ_A, that is, there is a partition of $\mathbf{X}^A = S_1 \cup \ldots \cup S_n$ such that for each i, $S_i \in \sigma_A$ and $E(S_i|\sigma_A^*) = \frac{1}{n}$.* □

Let $\mathcal{B} \leq \mathcal{A}$ be σ-algebras and let μ be a countably additive measure on $\mathcal{A}$. $\mathcal{B}$ is said to be *dense* in $\mathcal{A}$ if for every $A \in \mathcal{A}$, there exists a sequence $\{B_n\} \subseteq \mathcal{B}$ such that $\mu(A \Delta B_n) \to 0$.

Sub-σ-algebras $\mathcal{B}$ and $\mathcal{C}$ of $\mathcal{A}$ are said to be *independent complements* of each other in $\mathcal{A}$ if they are independent and the σ-algebra generated by $\mathcal{B}$ and $\mathcal{C}$ is dense in $\mathcal{A}$.

Theorem 2.3.2. *Suppose $\mathcal{A} \geq \mathcal{B}$ are separable σ-algebras on a set X, μ is a probability measure on $\mathcal{A}$ and there exists a $\mathcal{B}$-random k-partition $A_{1,k}, ..., A_{k,k}$ in $\mathcal{A}$, for every $k \in \mathbb{N}$. Then there exists an independent complement $\mathcal{C}$ of $\mathcal{B}$ in $\mathcal{A}$.* □

Let G be a group that acts on a set X. For any $g \in G$ and $S \subseteq X$, let S^g denote the image of S under the action of g. Let $\mathcal{S}$ be a subset of the powerset of X.

We say $\mathcal{S}$ is *setwise G-invariant* if for each $g \in G$ and $S \in \mathcal{S}$, $S^g \in \mathbf{S}$.

$\mathcal{S}$ is *elementwise G-invariant* if for each $g \in G$ and $S \in \mathcal{S}$, $S^g = S$.

Let $(X, \mathcal{A}, \mu)$ be a probability space and G be a finite group that acts on X such that $\mathcal{A}$ is setwise G-invariant. We say the *action of G* on this space is *independent* if there exists $S \in \mathcal{A}$ such that $\mu(S) = \frac{1}{|G|}$ and $S^{g_1} \cap S^{g_2} = \emptyset$ whenever $g_1 \neq g_2 \in G$.

We need the following consequence of Theorem 2.3.2 and it plays a critical role in obtaining separable realizations that will be defined in the following section. Separable realizations are the key to the second piece of the correspondence we are trying to establish.

Theorem 2.3.3. *Let $\mathcal{A} \geq \mathcal{B}$ be separable σ-algebras with a probability measure μ. Let a finite group G act on X such that $\mathcal{A}, \mathcal{B}, \mu$ are G-invariant and the action of G on $(X, \mathcal{B}, \mu)$ be independent. Let there also exist a $\mathcal{B}$-random k-partition of X in $\mathcal{A}$ for all k.*

Then there is an independent complement $\mathcal{C}$ of $\mathcal{B}$ in $\mathcal{A}$ and $\mathcal{C}$ is elementwise G-invariant.

$\square$

2.4 Separable Realizations

The symmetric group $S_{[k]}$ acts on $\mathbf{X}^{[k]}$ by permuting the coordinates.

$$(x_1, \ldots, x_k)^\pi = (x_{\pi^{-1}(1)}, \ldots, x_{\pi^{-1}(k)})$$

$S_{[k]}$ also acts on the power set of $[k]$. We let A^π denote the image of a set $A \subseteq [k]$ under $\pi \in S_{[k]}$. $(\sigma_A)^\pi = \sigma_{A^\pi}$.

For any set A and $k \leq |A|$, let $r(A, k) = \{B \subseteq A : B \neq \emptyset, |B| \leq k\}$ and $r(A) = r(A, |A|)$.

A *separable system* on $\mathbf{X}^{[k]}$ is a system of atomless separable σ-algebras $\{l_A : A \in r([k])\}$ such that

1. l_A is a subalgebra of σ_A, independent of σ_A^*.

2. $l_A^\pi = l_{A^\pi}$ for every permutation $\pi \in S_{[k]}$.

3. $Y^\pi = Y$ for every $Y \in l_A$ and $\pi \in S_A$.

For instance, when $k = 2$, a separable system on $\mathbf{X}^{[2]}$ consists of atomless separable σ-algebras $l_{\{1\}}, l_{\{2\}}, l_{\{1,2\}}$ on $\mathbf{X}^{[2]}$ such that

1. $l_{\{1\}}$ depends only on the first coordinate, $l_{\{2\}}$ depends only on the second and they are essentially the same algebra on one coordinate.

2. $l_{\{1,2\}} \leq \sigma_{\{1,2\}}$ and it is independent from $\sigma^*_{[2]}$, the σ-algebra generated by measurable rectangles.

3. $l_{\{1,2\}}$ contains symmetric sets.

Random partitions and independent complements play an important role in showing the existence of separable systems. Elek and Szegedy used Theorem 2.3.1 and Theorem 2.3.3 to prove the following theorem about the existence of separable systems.

Theorem 2.4.1. *Given $J \in r([k])$ and separable sub-σ-algebras $\boldsymbol{\mathcal{A}}_j$ of $\sigma_{[j]}$ for all $j \in J$, there exists a separable system such that for every $\boldsymbol{\mathcal{A}}_j$-measurable set $\mathbf{M}$, there is $\mathbf{N}$ in $\langle l_A : A \in r([j]) \rangle$ and $\mu_{[j]}(\mathbf{M}\Delta\mathbf{N}) = 0$.* □

Let $(X, \mathcal{A}, \mu)$ be a measure space. We say $A, A' \in \mathcal{A}$ are equivalent if $\mu(A\Delta A') = 0$ and $\mu([A]) = \mu(A)$. The equivalence classes form a Boolean algebra, with a countably additive measure μ, called the *measure algebra* $\mathbb{M}(X, \mathcal{A}, \mu)$.

Using a particular case of Maharam's theorem, we know the following :

Lemma 2.4.2. *Let $(X, \mathcal{A}, \mu)$ be a separable, atomless probability measure space. Then*

1. *The measure algebra $\mathbb{M}(X, \mathcal{A}, \mu)$ is isomorphic to $\mathbb{M}([0, 1], \mathbf{B}, \lambda)$, the measure algebra of the standard Lebesgue space on the unit interval.*

2. *There exists a map $f : X \to [0, 1]$ that defines a measure algebra isomorphism between the measure algebras, that is,*

(a) *Given any Lebesgue-measurable set* $U \subseteq [0,1]$*,* $f^{-1}(U) \in \mathcal{A}$ *and* $\mu(f^{-1}(U)) = \lambda(U)$*.*

(b) *For every* $A \in \mathcal{A}$*, there exists a Lebesgue-measurable set* B*, such that* A *and* $f^{-1}(B)$ *differ by a null set.* $\square$

A *separable realization* is a measure-preserving map $\phi : \mathbf{X}^{[k]} \to [0,1]^{r([k])}$ such that

1. ϕ commutes with any permutation in $S_{[k]}$.
2. Given any Lebesgue-measurable set $B \subseteq [0,1]$ and any $A \in r([k])$, $\phi_A^{-1}(B) \in \sigma_A$ and it is independent of σ_A^*.

In the case that $k = 2$, a separable realization on $\mathbf{X}^{[2]}$ is a measure-preserving map $\phi = (\phi_{\{1\}}, \phi_{\{2\}}, \phi_{[2]}) : \mathbf{X}^{[2]} \to [0,1]^{r([2])}$ such that

- $\phi_A : \mathbf{X}^{[2]} \to [0,1]$ is a σ_A-measurable function.
- $\phi_{\{1\}}(x, -) = \phi_{\{2\}}(-, x)$.
- $\phi_{[2]}(x, y) = \phi_{[2]}(y, x)$.
- $\phi_{[2]}^{-1}(B)$ is independent of $\sigma_{[2]}^*$, for all Lebesgue measurable sets $B \subseteq [0,1]$.

It is easy to see that a separable realization on $\mathbf{X}^{[k]}$ induces a separable system on $\mathbf{X}^{[k]}$.

Elek and Szegedy used Theorem 2.4.1 and Theorem 2.4.2 to prove the existence of separable realizations in the following theorem.

Theorem 2.4.3. *For every* $J \in r([k])$ *and separable sub-*σ*-algebras* $\boldsymbol{\mathcal{A}}_j$ *of* $\sigma_{[j]}$ *for* $j \in J$*, there exists a separable realization* ϕ *such that for every* $\boldsymbol{\mathcal{A}}_j$*-measurable set* $\mathbf{A}$*, there exists* *a Lebesgue-measurable set* $B \subseteq [0,1]^{r([j])}$ *and* $\mu_{[j]}(\phi^{-1}(P^{-1}_{r([j])}(B))\Delta\mathbf{A}) = 0$*, where* $P_{r([j])} : [0,1]^{r([k])} \to [0,1]^{r([j])}$ *is the projection map.* $\square$

Let F_A be the measurable functions obtained by using Theorem 2.4.2 on the separable algebras from Theorem 2.4.1. Using the Total Independence Lemma, it can be verified that ϕ defined on $\mathbf{X}^{[k]}$ by $\phi(\mathbf{x}) = (F_A(\mathbf{x}) : A \in r([n],k))$ is a separable realization on $\mathbf{X}^{[k]}$.

In the case that $k = 2$, $\phi(x,y) = (F_{\{1\}}(x,-), F_{\{2\}}(-,y), F_{\{1,2\}}(x,y))$.

It is useful to observe that given an $S_{[j]}$-symmetric $\boldsymbol{\mathcal{A}}_j$-measurable set $\mathbf{A}$, the symmetric properties of the functions F_S ensure that $\mathbf{A}$ differs on a null set from an $S_{[j]}$-symmetric subset of $[0,1]^{r([j])}$.

Remark 2.4.4. The restriction of any separable system on $\mathbf{X}^{[k]}$ to $\mathbf{X}^{[j]}$ for $j < k$, using the projection map $\pi_{[j]}$, also yields a separable system on $\mathbf{X}^{[j]}$ and a corresponding separable realization $\phi_j = P_{r([j])} \circ \phi \circ \pi^{-1}_{[j]} : \mathbf{X}^{[j]} \to [0,1]^{r([j])}$. In this scheme, ϕ is in fact ϕ_k.

We can also extend or *lift* our separable realizations on $\mathbf{X}^{[k]}$ to $\mathbf{X}^{[n]}$ for $n > k$.

Given $n \geq k$ and $\phi : \mathbf{X}^{[k]} \to [0,1]^{r([k])}$, then a measure-preserving map $\psi : \mathbf{X}^{[n]} \to [0,1]^{r([n],k)}$ is called a *degree n lifting of* ϕ if $P_{r([k])} \circ \psi = \phi \circ \pi_{[k]}$ and $\psi(x)^\tau = \psi(x^\tau)$ for all $x \in \mathbf{X}^{[n]}$ and all $\tau \in S_n$.

Theorem 2.4.5. *Let* $\phi : \mathbf{X}^{[k]} \to [0,1]^{r([k])}$ *be a separable realization and let* $n \geq k$ *be a natural number. Then a degree n lifting* ψ *of* ϕ *exists.*

Proof. Let $A \subseteq [n]$ and $|A| = j \leq k$ and let $\tau \in S_{[n]}$ such that $A^\tau = [j]$. As one may expect, we define $\psi_A(x) = \phi_{[j]}(\pi_{[k]}(x^\tau))$. Since the separable realization ϕ commutes with every permutation in $S_{[k]}$, we find that $\psi_A = \phi_A \circ \pi_{[k]}$ for every $A \subseteq [k]$.

For $\tau \in S_{[n]}$, the A-coordinate of $\psi(x)$ is the $A^{\tau^{-1}}$-coordinate of $\psi(x)$, which in turn is the A-coordinate of $\psi(x^\tau)$. Thus we see ψ commutes with the action of $S_{[n]}$ on $\mathbf{X}^{[k]}$.

To see that ψ is measure-preserving, we will need to use the Total Independence Lemma. Each ψ_A is clearly measure-preserving. Consider a set $W = \prod_{A \in r([n],k)} I_A$, where each $I_A \subseteq [0,1]$ is an interval. Since every measurable subset of $[0,1]^{r([n],k)}$ can be approximated by such hypercubes, it is enough to check that ψ^{-1} preserves the measure of such a set W.

Note that $\psi^{-1}(W) = \bigcap_{A \in r([n],k)} \psi_A^{-1}(I_A)$ and each $\psi_A^{-1}(I_A) \in \sigma_A$ is independent of σ_A^* due to the nature of the separable realization. The Total Independence Lemma now completes the proof. $\square$

Remark 2.4.6. $\psi_j = P_{r([j])} \circ \psi$ is a degree n lifting of ϕ_j, for each $j \in [k]$ and $\psi_k = \psi$.

We will need Theorem 2.4.5 and Remark 2.4.6 in the chapters to follow, most importantly to show the existence of limits.

Chapter 3

Relational Structures

3.1 General Relational Structures and Limits

A *relational type* τ is a first-order signature that has finitely many relation symbols, but no function symbols or constant symbols. Let ar be the function that assigns an arity to each relation symbol in τ. A *relational structure* $\mathcal{R}$ *of type* τ on an underlying set X interprets each relation symbol $R \in \tau$ with $ar(R) = k$ as a k-ary relation $R^{\mathcal{R}}$ on X.

An interesting problem arises when we deal with relations or relational structures on an ultraproduct space. Consider a binary relation $\mathbf{R}$ on $\mathbf{X}$. Then the set of loops $\{(x, x) : (x, x) \in \mathbf{R}\}$ always has measure 0. This makes it difficult to say anything interesting about the set of loops in the ultraproduct or to translate any property of loops in the sequence of structures on finite sets to the ultraproduct setting in a meaningful way. However, $\{x : (x, x) \in \mathbf{R}\} \subseteq \mathbf{X}$ could be a set of appreciable measure.

Therefore, we need to address tuples with repetition separately in every relation.

A hypergraph consists of a vertex set and a set of hyperedges. In a *directed r-uniform hypergraph*, each hyperedge is an ordered r-tuple (no repeated elements). We will say that an $S_{[r]}$-symmetric directed r-uniform hypergraph is an *r-uniform hypergraph*.

We will represent any k-ary relation R on a set X as a system of directed r-uniform hypergraphs, where r varies from 1 to k.

Let $\mathcal{P}_k$ be the set of all partitions of $[k]$. For any partition $p \in \mathcal{P}_k$, let $\|p\|$ denote the number of parts in p. Given $p_1, p_2 \in \mathcal{P}_k$, we say $p_1 \leq p_2$ if p_1 is a refinement of p_2. $p_1 < p_2$ if and only if $p_1 \neq p_2$ and $p_1 \leq p_2$, that is, p_1 is a strict refinement of p_2.

Given any $\bar{x} = (x_1, \ldots, x_k) \in X^{[k]}$, there is a unique $p = p(\bar{x}) \in \mathcal{P}_k$, such that $x_i = x_j$ if and only if i and j occur in the same part in p.

Let $R_p = \{\bar{x} \in R : p(x) = p\}$. Then $R = \bigcup_{p \in \mathcal{P}_k} R_p$.

For $\bar{x} \in R_p$, let $\eta_p(\bar{x})$ be the tuple in $X^{[\|p\|]}$ obtained from $\bar{x}$ by retaining only the first occurrence of every element. $\eta_p : R_p \to X^{[\|p\|]}$ is an injective function. $H_p = H_p(R) = \eta_p[R_p]$ is a directed $\|p\|$-uniform hypergraph on X and $R = \bigcup_{p \in \mathcal{P}_k} \eta_p^{-1}(H_p)$.

Thus, the relation R can be represented by the system of directed hypergraphs $\{H_p(R) : p \in \mathcal{P}_k\}$.

Every relational structure $\mathcal{R}$ of type τ can be represented by the system of hypergraphs $\{H_p(R^{\mathcal{R}}) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$.

In the case that $k = 2$, a binary relation can be represented using $H_{(1,2)} \subseteq X$ and a digraph $H_{(1)(2)}$ on X. $H_{(1,2)} = \{x : (x, x) \in R\}$ and $H_{(1)(2)} = \{(x, y) : x \neq y, (x, y) \in R\}$, that is, a binary relation is essentially a digraph with loops.

In the case that $k = 3$, a ternary relation $R = R_{(1,2,3)} \cup R_{(1,2)(3)} \cup R_{(1,3)(2)} \cup R_{(1)(2,3)} \cup R_{(1)(2)(3)}$. R can now be represented by 5 directed hypergraphs : $H_{(1,2,3)} \subseteq X$, 3 digraphs $H_{(1,2)(3)}$, $H_{(1,3)(2)}$ and $H_{(1)(2,3)}$ on X, and a directed 3-uniform hypergraph $H_{(1)(2)(3)}$ on X. As an example, consider $(x, y, x) \in R$ and $x \neq y$. Then $(x, y, x) \in R_{(1,3)(2)}$ and $\eta_{(1,3)(2)}(x, y, x) = (x, y)$. Note that $\eta^{-1}_{(1,3)(2)}(x, y) = (x, y, x)$, however $\eta^{-1}_{(1)(2,3)}(x, y) = (x, y, y)$.

Consider the relational type $\tau = \{R, S, P\}$ with $ar(R) = ar(S) = 2$ and $ar(P) = 1$. A relational structure $\mathcal{R}$ of type τ can be represented by the system $\{H_{(1,2)}(R^{\mathcal{R}}), H_{(1)(2)}(R^{\mathcal{R}}), H_{(1,2)}(S^{\mathcal{R}}), H_{(1)(2)}(S^{\mathcal{R}}), H_{(1)}(P^{\mathcal{R}})\}$.

Now let us consider an ω-sequence of finite sets X_i that are increasing in size and a non-principal ultrafilter U on ω. Suppose we have k-ary relations $R_i \subseteq X_i^{[k]}$. We can now take ultraproducts modulo U to obtain a k-ary relation $\mathbf{R} = [R_i] \subseteq \mathbf{X}^{[k]}$. Clearly $\mathbf{R}$ is a $\sigma_{[k]}$-measurable set and we can now repeat the above procedure to obtain for each $p \in \mathcal{P}_k$, measurable sets $H_p(\mathbf{R}) = \eta_p[\mathbf{R}_p] \subseteq \mathbf{X}^{[\|p\|]}$ such that $\mathbf{R} = \bigcup_{p \in \mathcal{P}_k} \eta_p^{-1}(H_p(\mathbf{R}))$.

In fact, $\mathbf{R}_p = [R_{i,p}]$, $H_p(\mathbf{R}) = [H_p(R_i)]$ and $\eta_p = [\eta_{i,p}]$.

Fix a relational type τ. Consider an increasing sequence of finite sets X_i and relational structures $\mathcal{R}_i$ of type τ on X_i. We will now define a relational structure $\boldsymbol{\mathcal{R}}$ of type τ on $\mathbf{X}$ as the ultraproduct of this sequence. For each $R \in \tau$ with $k = ar(R)$, take the ultraproduct of each sequence of relations $R^{\mathcal{R}_i}$ to obtain a k-ary relation $R^{\boldsymbol{\mathcal{R}}}$ on $\mathbf{X}$. $\boldsymbol{\mathcal{R}}$ can now be represented by the system $\{H_p(R^{\boldsymbol{\mathcal{R}}}) = [H_p(R^{\mathcal{R}_i})] : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$.

Then as a consequence of Theorem 2.4.3, there exists a separable realization ϕ such that for each $R \in \tau$ and each $p \in \mathcal{P}_k$ where $k = ar(R)$, there exist measurable sets $W_p(R^{\boldsymbol{\mathcal{R}}}) \subseteq [0,1]^{r([\|p\|])}$ such that

$$\mu_{[\|p\|]}(\phi_{\|p\|}^{-1}(W_p(R^{\boldsymbol{\mathcal{R}}}))\Delta H_p(R^{\boldsymbol{\mathcal{R}}})) = 0.$$

Let $W(\boldsymbol{\mathcal{R}}) = \{W_p(R^{\boldsymbol{\mathcal{R}}}) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$. We call this the corresponding *Euclidean structure of type* τ.

Let $\mathcal{R}$ and $\mathcal{S}$ be relational structures of type τ on underlying sets X and Y respectively. A map $\phi : Y \to X$ is a *homomorphism* from $\mathcal{S}$ to $\mathcal{R}$ if for each relation $R \in \tau$ with $k = ar(R)$ and for all $y_1, \ldots, y_k \in Y$,

$$(y_1, \ldots, y_k) \in R^{\mathcal{S}} \implies (\phi(y_1), \ldots, \phi(y_k)) \in R^{\mathcal{R}}.$$

Let $T(\mathcal{S}, \mathcal{R})$ denote the set of homomorphisms from $\mathcal{S}$ to $\mathcal{R}$ and let $t(\mathcal{S}, \mathcal{R})$ denote the *homomorphism density* of $\mathcal{S}$ in $\mathcal{R}$. If X and Y are finite sets, define $t(\mathcal{S}, \mathcal{R}) = \frac{|T(\mathcal{S},\mathcal{R})|}{|X|^{|Y|}}$.

Similarly, $T_o(\mathcal{S}, \mathcal{R})$ denotes the set of injective homomorphisms and the *injective homomorphism density* $t_o(\mathcal{S}, \mathcal{R}) = \frac{|T_o(\mathcal{S},\mathcal{R})|}{\binom{|X|}{|Y|}}$.

A map $\phi : Y \to X$ is an *induced homomorphism* from $\mathcal{S}$ to $\mathcal{R}$ if for each relation $R \in \tau$ with $k = ar(R)$ and for all $y_1, \ldots, y_k \in Y$,

$$(y_1, \ldots, y_k) \in R^{\mathcal{S}} \iff (\phi(y_1), \ldots, \phi(y_k)) \in R^{\mathcal{R}}.$$

We can also look at the sets $T_{ind}(\mathcal{S}, \mathcal{R})$ and $T_{o,ind}(\mathcal{S}, \mathcal{R})$ of induced homomorphisms and injective induced homomorphisms respectively, from $\mathcal{S}$ to $\mathcal{R}$, and analogously define the *induced homomorphism density* $t_{ind}(\mathcal{S}, \mathcal{R})$ and the *injective induced homomorphism density* $t_{o,ind}(\mathcal{S}, \mathcal{R})$.

Now a homomorphism need not be an injective map. Consider $\tau = \{R\}$ with $ar(R) = 3$. Let $Y = \{y, z\}$, $X = \{x\}$ and $\phi(y) = \phi(z) = x$. Suppose $R^{\mathcal{S}} = (y, z, y)$. Since ϕ maps (y, z, y) to (x, x, x), $\phi \in T(\mathcal{S}, \mathcal{R})$ if and only if $(x, x, x) \in R^{\mathcal{R}}$. Note that, in the context of the directed hypergraphs representing these relations, the homomorphism ϕ mapped $(y, z) \in H_{(1,3)(2)}(R^{\mathcal{S}})$ to $(x) \in H_{(1,2,3)}(R^{\mathcal{R}})$. We must account for such cases when calculating the (induced) homomorphism density, especially when $\mathcal{R}$ is a relational structure on a finite set. Interestingly, in the ultraproduct case the set of non-injective homomorphisms into $\boldsymbol{\mathcal{R}}$ has measure 0. This is what we would expect since the probability of two elements from a finite set Y being mapped to the same point in the ultraproduct $\mathbf{X}$ is 0.

Let $k_{max} = \max\{ar(R) : R \in \tau\}$ and let $E_r(X)$ denote the complete r-uniform hypergraph on X, for any $r \in \mathbb{N}$. Recall that we have previously defined r-uniform hypergraphs as $S_{[r]}$-symmetric directed r-uniform hypergraphs.

Assume $Y = [n]$. We can represent a homomorphism $\phi : \mathcal{S} \to \mathcal{R}$ as an n-tuple $(\phi(1), \ldots, \phi(n)) \in X^{[n]}$. Then $T(\mathcal{S}, \mathcal{R}), T_o(\mathcal{S}, \mathcal{R}), T_{ind}(\mathcal{S}, \mathcal{R})$ and $T_{o,ind}(\mathcal{S}, \mathcal{R})$ are subsets of $X^{[n]}$. For a relational structure $\boldsymbol{\mathcal{R}}$ on the ultraproduct $\mathbf{X}$ that is represented by a system of measurable directed hypergraphs, all homomorphism sets are measurable. Then $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \mu_{[n]}(T(\mathcal{S}, \boldsymbol{\mathcal{R}}))$, $t_o(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \mu_{[n]}(T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}))$, $t_{ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \mu_{[n]}(T_{ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}))$ and $t_{o,ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \mu_{[n]}(T_{o,ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}))$.

Since the copies of X in the product $X^{[n]}$ are indexed, we use the following maps to ensure that any r-tuple below lives in the appropriate isomorphic image of $X^{[r]}$. For every $e \in E_r([n])$, let $\pi_e : X^{[n]} \to X^e$ be the natural projection which can be defined as $\pi_e(\mathbf{x}) = (x_{e(1)}, \ldots, x_{e(r)})$. Let $\theta_e : X^{[r]} \to X^e$ be the natural bijection induced by the ordered tuple e. Given $\mathbf{x} = (x_1, \ldots, x_r) \in X^{[r]}$ where each $x_i \in X^{\{i\}}$, θ_e maps $\mathbf{x}$ to a copy of itself in X^e by sending each x_i to a copy of itself in $X^{\{e(i)\}}$.

For example, given $Y = [3]$, $e = (3, 2)$ and $\mathbf{x} = (x_1, x_2, x_3) \in X^{[3]}$, $\pi_e(\mathbf{x}) = (x_3, x_2) \in X^{(3,2)}$ and $\theta_e^{-1}(\pi_e(\mathbf{x})) = (x_3, x_2) \in X^{[2]}$.

Consider $\tau = \{R\}$ with $ar(R) = 3$. We need the set $\mathcal{P}_3$ of partitions of [3]. Let $Y = [4]$ and $R^{\mathcal{S}} = \{(4, 2, 2)\}$. This implies $(4, 2) \in H_{(1)(2,3)}$. Now $\mathbf{x} = (x_1, x_2, x_3, x_4) \in T(\mathcal{S}, \mathcal{R})$ if and only if $(x_4, x_2, x_2) \in R^{\mathcal{R}}$. Let $e = (4, 2)$ and $p = (1)(2, 3)$. Then $\|p\| = 2$ and $\theta_e^{-1}(\pi_e(\mathbf{x})) = (x_4, x_2) \in X^{[2]}$.

Either $x_4 \neq x_2$ and $\mathbf{x} \in T(\mathcal{S}, \mathcal{R})$ if and only if $\theta_e^{-1}(\pi_e(\mathbf{x})) = (x_4, x_2) \in H_{(1)(2,3)}(R^{\mathcal{R}})$, or $x_4 = x_2$ and $\mathbf{x} \in T(\mathcal{S}, \mathcal{R})$ if and only if $(x_4) \in H_{(1,2,3)}(R^{\mathcal{R}})$. In the latter case, observe that there is a unique partition $q = p(\theta_e^{-1}(\pi_e(\mathbf{x}))) = p(x_4, x_2) = (1, 2) \in \mathcal{P}_2 = \mathcal{P}_{\|p\|}$ such that $\|q\| < \|p\|$ since $x_4 \neq x_2$. Also there is a unique partition $p' = (1, 2, 3) \in \mathcal{P}_3$ such that $\eta_q(\theta_e^{-1}(\pi_e(\mathbf{x}))) = \eta_q(x_4, x_2) = (x_4) \in H_{p'}(R^{\mathcal{S}})$ and $\|p'\| = \|q\|$.

In general, given a tuple $\mathbf{x} \in X^{[n]}$ and a directed hyperedge $e \in E_k(X)$, $\theta_e^{-1}(\pi_e(\mathbf{x})) = (x_{e(1)}, \ldots, x_{e(k)}) \in X^{[k]}$. $\mathbf{x}$ represents a homomorphism from $\mathcal{S}$ to $\mathcal{R}$, that is, $\mathbf{x} \in T(\mathcal{S}, \mathcal{R})$ if for all $R \in \tau$ with $k = ar(R)$, all $p \in \mathcal{P}_k$ and all $e \in H_p(R^{\mathcal{S}})$, either

Case 1 : $\theta_e^{-1}(\pi_e(\mathbf{x})) \in R_p^{\mathcal{R}}$ has $\|p\|$ distinct elements. Then $\theta_e^{-1}(\pi_e(\mathbf{x})) \in H_p(R^{\mathcal{R}})$.

or

Case 2 : There exists a unique partition $q = p(\theta_e^{-1}(\pi_e(\mathbf{x}))) \in \mathcal{P}_{\|p\|}$ and $\|q\| < \|p\|$. Then there exists a unique partition $p' \in \mathcal{P}_k$ such that $\eta_q(\theta_e^{-1}(\pi_e(\mathbf{x}))) \in H_{p'}(R^{\mathcal{R}})$. $\|p'\| = \|q\|$

and $p \leq p'$.

Let $Z_p(R^{\mathcal{R}}) = \{\mathbf{y} \in X^{[\|p\|]} \mid q = p(\mathbf{y}) \in \mathcal{P}_{\|p\|} : \|q\| < \|p\| \text{ and } \exists!\, p' \in \mathcal{P}_k : \eta_q(\mathbf{y}) \in H_{p'}(R^{\mathcal{R}})\}$. Then

$$T(\mathcal{S}, \mathcal{R}) = \bigcap_{R \in \tau} \bigcap_{p \in \mathcal{P}_k} \bigcap_{e \in H_p(R^{\mathcal{S}})} \pi_e^{-1}(\theta_e(H_p(R^{\mathcal{R}}) \cup Z_p(R^{\mathcal{R}})))$$

If $\mathbf{x} \in X^{[n]}$ represents an injective homomorphism, then each $\theta_e^{-1}(\pi_e(\mathbf{x}))$ always has distinct elements. Then the set of injective homomorphisms from $\mathcal{S}$ to $\mathcal{R}$ is

$$T_o(\mathcal{S}, \mathcal{R}) = \bigcap_{R \in \tau} \bigcap_{p \in \mathcal{P}_k} \bigcap_{e \in H_p(R^{\mathcal{S}})} \pi_e^{-1}(\theta_e(H_p(R^{\mathcal{R}})))$$

A sequence of relational structures $\mathcal{R}_i$ of type τ on finite sets X_i is said to be *convergent* if for any finite set Y and a relational structure $\mathcal{S}$ of type τ on Y, $\lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i)$ exists.

The sequence is said to be *increasingly convergent* if it is convergent and $|X_i| \to \infty$.

Given a sequence of relational structures $\mathcal{R}_i$ of type τ on finite sets X_i of increasing size, we can take the ultraproduct $\boldsymbol{\mathcal{R}}$ and obtain a system $\{H_p(R^{\boldsymbol{\mathcal{R}}}) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$ that represents $\boldsymbol{\mathcal{R}}$ and we can also obtain the corresponding Euclidean structure $W(\boldsymbol{\mathcal{R}})$ as above.

Let $\mathcal{S}$ be a relational structure of type τ on $[n]$. Then

$$T(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \bigcap_{R\in\tau} \bigcap_{p\in\mathcal{P}_k} \bigcap_{e\in H_p(R^{\mathcal{S}})} \pi_e^{-1}(\theta_e(H_p(R^{\boldsymbol{\mathcal{R}}}) \cup Z_p(R^{\boldsymbol{\mathcal{R}}}))) = [T(\mathcal{S}, \mathcal{R}_i)].$$

Note that $Z_p(R^{\boldsymbol{\mathcal{R}}}) \subseteq \mathbf{X}^{[\|p\|]}$ and $\mu_{[\|p\|]}(Z_p(R^{\boldsymbol{\mathcal{R}}})) = 0$.

$T(\mathcal{S}, \boldsymbol{\mathcal{R}}) \subseteq \mathbf{X}^{[n]}$ is $\sigma_{[n]}$-measurable and

$$t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \mu_{[n]}(T(\mathcal{S}, \boldsymbol{\mathcal{R}})) = \lim_{U} t(\mathcal{S}, \mathcal{R}_i).$$

Further, if the sequence $\{\mathcal{R}_i\}$ is increasingly convergent, then $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i)$.

Given a relational structure $\mathcal{S}$ of type τ on $[n]$ and a Euclidean structure $W = \{W_p(R) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$, define the *homomorphism density of $\mathcal{S}$ in W* as :

$$t(\mathcal{S}, W) \;\; = \;\; \int_0^1 \cdots \int_0^1 \prod_{R\in\tau} \prod_{p\in\mathcal{P}_k} \prod_{e\in H_p(R^{\mathcal{S}})} \mathcal{W}_p(R)(\mathbf{x}_e) \prod_{A\in r([n],k_{max})} dx_A$$

where $\mathcal{W}_p(R)$ is the characteristic function of $W_p(R)$ and for every $e \in E_k([n])$, $\mathbf{x}_e = (x_{e[A]} : A \in r([k])$.

Theorem 3.1.1. *Let $\mathcal{S}$ be a relational structure of type τ on $[n]$ and let R_i form a sequence of relational structures of type τ on finite sets X_i of increasing size. Let $\boldsymbol{\mathcal{R}}$ be the relational structure obtained by taking ultraproducts and let $W(\boldsymbol{\mathcal{R}})$ be the corresponding Euclidean structure. Then $t(S, W(\boldsymbol{\mathcal{R}})) = t(S, \boldsymbol{\mathcal{R}})$.*

Proof. For every $e \in E_{\|p\|}([n])$, let $L_e : [0,1]^{r([n],\|p\|)} \to [0,1]^{r(e)}$ be the natural measurable projection induced by π_e and let $F_e : [0,1]^{r[\|p\|]} \to [0,1]^{r(e)}$ be the measurable isomorphism induced by the bijection θ_e.

A separable realization ϕ on $\mathbf{X}^{[k_{max}]}$ exists as in Theorem 2.4.3 and by Theorem 2.4.5, a lifting ψ of ϕ to $\mathbf{X}^{[n]}$ exists.

We know that $\mu_{[\|p\|]}(\phi^{-1}_{\|p\|}(W_p(R^{\boldsymbol{\mathcal{R}}})\Delta H_p(R^{\boldsymbol{\mathcal{R}}})) = 0$ and $\mu_{[\|p\|]}(Z_p(R^{\boldsymbol{\mathcal{R}}})) = 0$, for each $R \in \tau$ and $p \in \mathcal{P}_k$ where $k = ar(R)$.

$$
\begin{aligned}
&\mu_{[n]}(T(\mathcal{S},\boldsymbol{\mathcal{R}})) \\
= \;& \mu_{[n]}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k}\bigcap_{e\in H_p(R^{\mathcal{S}})} \pi_e^{-1}(\theta_e(H_p(R^{\boldsymbol{\mathcal{R}}})))\right) \\
= \;& \mu_{[n]}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k}\bigcap_{e\in H_p(R^{\mathcal{S}})} \pi_e^{-1}(\theta_e(\phi^{-1}_{\|p\|}(W_p(R^{\boldsymbol{\mathcal{R}}}))))\right) \\
= \;& \mu_{[n]}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k}\bigcap_{e\in H_p(R^{\mathcal{S}})} \psi^{-1}_{\|p\|}(L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}}))))\right) \\
= \;& \mu_{[n]}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} \psi^{-1}_{\|p\|}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right) \\
= \;& \mu_{[n]}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} \psi^{-1}\left(P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right)\right) \\
= \;& \mu_{[n]}\left(\psi^{-1}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right)\right) \\
= \;& Vol\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right) \\
= \;& \int_0^1 \cdots \int_0^1 \prod_{R\in\tau}\prod_{p\in\mathcal{P}_k}\prod_{e\in H_p(R^{\mathcal{S}})} \mathcal{W}_p(R^{\boldsymbol{\mathcal{R}}})(\mathbf{x}_e) \prod_{A\in r([n],k_{max})} dx_A
\end{aligned}
$$

We have thus proved, $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = t(\mathcal{S}, W(\boldsymbol{\mathcal{R}}))$. □

If the sequence $\{\mathcal{R}_i\}$ is increasingly convergent, then for each structure $\mathcal{S}$ on some finite set, $t(\mathcal{S}, W(\boldsymbol{\mathcal{R}})) = \lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i)$ and the Euclidean structure $W(\boldsymbol{\mathcal{R}})$ is said to be a *limit* for the sequence $\{\mathcal{R}_i\}$.

Remark 3.1.2. In the course of the proof of Theorem 3.1.1, we have also proved $t_o(\mathcal{S}, \boldsymbol{\mathcal{R}}) = t(\mathcal{S}, W(\boldsymbol{\mathcal{R}}))$.

We shall view graphs as digraphs with symmetric edge sets. Digraphs can be represented as relational structures of type $\tau = \{E\}$ such that $ar(E) = 2$. Given digraphs G, $H_{(1)(2)}(E^G) = E^G$ and $H_{(1,2)}(E^G) = \emptyset$. Given a convergent sequence of finite digraphs G_n, we can obtain, as a limit, a Euclidean structure $W = \{W_{(1)(2)}, W_{(1,2)}\}$, where $W_{(1)(2)}$ is a Lebesgue-measurable subset of $[0,1]^3$, while $W_{(1,2)} = \emptyset$. Then $t(F, W) = \lim_{n\to\infty} t(F, G_n)$ for all finite digraphs F. Essentially, the three-dimensional set $W_{(1)(2)}$ is our digraph limit.

Let G be a graph. Then $W_{(1)(2)}$ is $S_{[2]}$-symmetric. How are these graph limits related to graphons that were shown to be graph limits in [9] by Lovász and Szegedy?

Let $\widetilde{W} : [0,1]^2 \to [0,1]$ be defined by

$$\widetilde{W}(x, y) = \int_0^1 \mathcal{W}_{(1)(2)}(x, y, z)\, dz.$$

Since $W_{(1)(2)}$ is $S_{[2]}$-symmetric and measurable, $\widetilde{W}$ is symmetric and measurable. So $\widetilde{W}$ is a graphon. Using the classical Fubini's theorem, we can see that for every $n \in \mathbb{N}$ and every graph F on $[n]$,

$$\begin{aligned} t(F,W) &= \int_0^1 \dots \int_0^1 \prod_{(i,j)\in E^F} \widetilde{W}(x_i,x_j)\,dx_1\dots dx_n \\ &= t(F,\widetilde{W}) \end{aligned}$$

as defined in [9]. Given any sequence of finite graphs G_n that converges to W, $t(F,\widetilde{W}) = \lim_{n\to\infty} t(F,G_n)$ for all finite graphs F and $\widetilde{W}$ is a limit for the sequence G_n. We have recovered the traditional graph limit from our form of the limit object.

The advantage of the three-dimensional limit is perhaps more apparent in comparison to the digraph limit, which Offner and Pikhurko showed to be a set of four measurable functions $\{\overleftrightarrow{W}, \overrightarrow{W}, \overleftarrow{W}, \overline{W}\}$ from $[0,1]^2$ to $[0,1]$ such that for all $x,y \in [0,1]$

1. $\overleftrightarrow{W}(x,y) + \overrightarrow{W}(x,y) + \overleftarrow{W}(x,y) + \overline{W}(x,y) = 1$,
2. $\overleftrightarrow{W}(x,y) = \overleftrightarrow{W}(y,x)$, and
3. $\overrightarrow{W}(x,y) = \overleftarrow{W}(y,x)$.

As the notation would indicate, the functions $\overleftrightarrow{W}, \overrightarrow{W}, \overleftarrow{W}, \overline{W}$ are intended to correspond to the probability of edges going both ways, an edge from left to right only, an edge from right to left only and no edges, respectively. Clearly $\overleftrightarrow{W}$ and $\overrightarrow{W}$ are enough.

Given a digraph F on a finite set, the set of directed edges $E(F) = \overleftrightarrow{F} \cup \overrightarrow{F}$ where $\overleftrightarrow{F} = \{(x,y) : (x,y) \in E(F), (y,x) \in E(F)\}$ and $\overrightarrow{F} = \{(x,y) : (x,y) \in E(F), (y,x) \notin E(F)\}$.

$$t(F,\overleftrightarrow{W},\overrightarrow{W}) = \int_0^1 \dots \int_0^1 \prod_{(i,j)\in\overleftrightarrow{F}} \overleftrightarrow{W}(x_i,x_j) \prod_{(i,j)\in\overrightarrow{F}} (\overrightarrow{W}(x_i,x_j) + \overleftrightarrow{W}(x_i,x_j))\,dx_1\dots dx_n$$

We can retrieve these functions from our digraph limit object $W_{(1)(2)}$. Recall that in the case of digraphs, $W_{(1)(2)}$ need not be $S_{[2]}$-symmetric. For all $x, y \in [0,1]$, define

$$\begin{aligned}
\overleftrightarrow{W}(x,y) &= \int_0^1 \mathcal{W}_{(1)(2)}(x,y,z)\mathcal{W}_{(1)(2)}(y,x,z)\,dz \\
\overrightarrow{W}(x,y) &= \int_0^1 \mathcal{W}_{(1)(2)}(x,y,z)(1-\mathcal{W}_{(1)(2)}(y,x,z))\,dz \\
\overleftarrow{W}(x,y) &= \int_0^1 (1-\mathcal{W}_{(1)(2)}(x,y,z))\mathcal{W}_{(1)(2)}(y,x,z)\,dz \\
\overline{W}(x,y) &= \int_0^1 (1-\mathcal{W}_{(1)(2)}(x,y,z))(1-\mathcal{W}_{(1)(2)}(y,x,z))\,dz
\end{aligned}$$

Using Fubini's theorem, we can once again see that $t(F, W_{(1)(2)}) = t(F, \overleftrightarrow{W}, \overrightarrow{W})$, for all digraphs F on finite sets.

Let us look at the nature of the limit object for a few other simple relational types.

Let $\tau = \{P, Q\}$ such that $ar(P) = 1$ and $ar(Q) = 1$. A limit object for two unary relations is a Euclidean structure W of type τ that consists of Lebesgue-measurable subsets $W_{(1)}(P)$ and $W_{(1)}(Q)$ of $[0,1]$.

Let $\tau = \{R\}$ such that $ar(R) = 2$. This is the type for binary relations and we have seen that our limit object for binary relations consists of Lebesgue-measurable sets $W_{(1)(2)}(R) \subseteq [0,1]^3$ and $W_{(1,2)}(R) \subseteq [0,1]$.

Let $\tau = \{R, S\}$ such that $ar(R) = 2$ and $ar(S) = 2$. A limit object is a Euclidean structure of type τ, represented by a system of Lebesgue-measurable subsets $\{W_{(1)(2)}(R), W_{(1,2)}(R), W_{(1)(2)}(S), W_{(1,2)}(S)\}$ such that $W_{(1)(2)}(R), W_{(1)(2)}(S) \subseteq [0,1]^3$ and

$W_{(1,2)}(R), W_{(1,2)}(S) \subseteq [0,1]$.

These examples indicate how rapidly the complexity of the limit object increases with the complexity of the relational type.

Let $H_p^c(R^{\mathcal{S}}) = \{e \in E_{\|p\|}([n]) : e \notin H_p(R^{\mathcal{S}})\}$. Define the *induced homomorphism density of $\mathcal{S}$ in $W(\boldsymbol{\mathcal{R}})$* as

$$\begin{aligned} t_{ind}(\mathcal{S}, W(\boldsymbol{\mathcal{R}})) &= \int_0^1 \cdots \int_0^1 \prod_{R \in \tau} \prod_{p \in \mathcal{P}_k} \prod_{e \in H_p(R^{\mathcal{S}})} \mathcal{W}_p(R^{\boldsymbol{\mathcal{R}}})(\mathbf{x}_e) \\ &\quad \prod_{e \in H_p^c(R^{\mathcal{S}})} (1 - \mathcal{W}_p(R^{\boldsymbol{\mathcal{R}}})(\mathbf{x}_e)) \prod_{A \in r([n], k_{max})} dx_A. \end{aligned}$$

Remark 3.1.3. We can repeat the proof of Theorem 3.1.1 to show that

$$t_{ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}) = t_{o,ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}) = t_{ind}(\mathcal{S}, W(\boldsymbol{\mathcal{R}})).$$

Let τ be a relational type. Let $W = \{W_p(R) : p \in \mathcal{P}_k, k = ar(R)\}$ be a Euclidean structure of type τ. Each $W_p(R)$ is a measurable subset of $[0,1]^{r([\|p\|])}$. We can generate a random relational structure $\mathcal{R}_m = \mathcal{R}(W, m)$ on $[m]$ as follows. Choose x_A uniformly at random in $[0,1]$, for $A \in r([m], k_{max})$. Fix relation $R \in \tau$ and let $k = ar(R)$. For each $p \in \mathcal{P}_k$, generate $H_p(R)$ as follows :

$$e \in H_p(R) \iff \mathbf{x}_e \in W_p(R).$$

The following result is similar to Lemma 2.4 in [9] and the proof follows along the same lines.

Lemma 3.1.4. *For every relational structure S on $[n]$*

1. $E(t_o(\mathcal{S}, \mathcal{R}_m)) = t(\mathcal{S}, W)$.

2. $|E(t(\mathcal{S}, \mathcal{R}_m)) - t(\mathcal{S}, W)| \leq \frac{n^2}{m}$.

3. $Var(t(\mathcal{S}, \mathcal{R}_m)) \leq \frac{3n^2}{m}$. □

Remark 3.1.5. We can also prove an analog of Lemma 3.1.4 for t_{ind} and $t_{o,ind}$. For a future result, we will use the first part : $E(t_{o,ind}(\mathcal{S}, \mathcal{R}_m)) = t_{ind}(\mathcal{S}, W)$.

Theorem 3.1.6. *The sequence $\{R_{m^2}\}$ is increasingly convergent and W is its limit with probability 1.*

Proof. Let $\epsilon > 0$. Fix n and a relation S on $[n]$. Assume m is large enough that $\frac{n^2}{m} \leq \frac{\epsilon}{2}$.

Using Chebyshev's inequality, $Prob(|t(\mathcal{S}, \mathcal{R}_{m^2}) - t(\mathcal{S}, W)| \geq \epsilon) \leq \frac{4n^2}{3\epsilon^2}\frac{1}{m^2}$.

Let $E_{m,\epsilon}$ be the event that $|t(\mathcal{S}, \mathcal{R}_{m^2}) - t(\mathcal{S}, W)| \geq \epsilon$. Then $\sum_{m=1}^{\infty} Prob(E_{m,\epsilon}) < \infty$. By the Borel-Cantelli Lemma, $Prob(\limsup E_{m,\epsilon}) = 0$.

Then $Prob(\cup_{\epsilon \in \mathbb{Q}^+} \limsup E_{m,\epsilon}) = 0$, that is, $Prob(\exists \epsilon > 0 : |t(\mathcal{S}, \mathcal{R}_{m^2}) - t(\mathcal{S}, W)| \geq \epsilon, i.o.) = 0$.

Thus we have proved $Prob(t(\mathcal{S}, \mathcal{R}_{m^2}) \to t(\mathcal{S}, W)) = 1$ for each relational structure $\mathcal{S}$ on a finite set. Since there are only countably many such $\mathcal{S}$, $\{\mathcal{R}_{m^2}\}$ is increasingly convergent and W is a limit for the random sequence $\{\mathcal{R}_{m^2}\}$ with probability 1. □

We can repeat the above proof for t_{ind} to prove that $t_{ind}(\mathcal{S}, \mathcal{R}_{m^2}) \to t_{ind}(\mathcal{S}, W)$ with probability 1.

3.2 Uniqueness

In this section, we address the question of uniqueness of limits. It is known that graphons are unique as graph limits up to measure preserving transformations [3]. Elek and Szegedy found limits for k-uniform hypergraphs and showed they were unique up to "structure preserving" maps. We extend their results to determine uniqueness of our limits for relational structures and show that they are also, in fact, unique up to structure preserving maps.

Let U and W be two Euclidean structures of type τ. Let

$$d_1(U,W) = \max\{\mathrm{Vol}(U_p(R)\Delta W_p(R)) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}.$$

$d_1(U,W) = 0$ if and only if U and W are *0-close*, that is, for all $R \in \tau$ with $k = ar(R)$, for all $p \in \mathcal{P}_k$, $U_p(R)\Delta W_p(R)$ is a null set. d_1 is clearly symmetric and satisfies the triangle inequality.

Let $\mathcal{S}$ be a relational structure of type τ on a finite set. Let $|\mathcal{S}| = \sum\limits_{R\in\tau}\sum\limits_{p\in\mathcal{P}_k} |H_p(R^{\mathcal{S}})|$. It is easy to verify that $|t(\mathcal{S},U) - t(\mathcal{S},W)| \leq |\mathcal{S}|\, d_1(U,W)$.

Let $\delta_w(U,W) = \inf\{\delta : \forall \mathcal{S},\ |t(\mathcal{S},U) - t(\mathcal{S},W)| \leq |\mathcal{S}|\, \delta\}$.

It is easy to see that $\delta_w(U,W) \leq d_1(U,W)$ and it satisfies the triangle inequality. If $\delta_w(U,W) = 0$, then $t(\mathcal{S},U) = t(\mathcal{S},W)$ for all relational structures $\mathcal{S}$ of type τ on finite sets. We are interested in the relation between U and W that ensures $\delta_w(U,W) = 0$.

Fix $k \in \mathbb{N}$. For every $S \in r([k])$, let $\mathcal{A}_S$ be the σ-algebra generated by the projection

$[0,1]^{r([k])} \to [0,1]^{r(S)}$. Let $\mathcal{A}_S^* = \langle \mathcal{A}_T : T \subsetneq S \rangle$.

A Lebesgue measurable function $\phi = (\phi_S : S \in r([k])) : [0,1]^{r([k])} \to [0,1]^{r([k])}$ is a *structure preserving map* on $[0,1]^{r([k])}$ if

1. ϕ is measure preserving.
2. $\phi^{-1}(\mathcal{A}_S) \subseteq \mathcal{A}_S$, for each $S \in r([k])$.
3. Given a Lebesgue measurable set $I \subseteq [0,1]$ and $S \in r([k])$, $\phi_S^{-1}(I)$ is independent from $\mathcal{A}_S^*$.
4. $\phi \circ \pi = \pi \circ \phi$ for all $\pi \in S_{[k]}$.

Remark 3.2.1. Given $m \leq k$, let $\phi_m : [0,1]^{r([m])} \to [0,1]^{r([m])}$ be the natural restriction of the structure preserving map ϕ. Then ϕ_m is also structure preserving.

Here, $\phi = \phi_k$.

Remark 3.2.2. Let ϕ, $\widetilde{\phi}$ be two structure preserving maps on $[0,1]^{r([k])}$. Then $\phi \circ \widetilde{\phi}$ and ϕ^{-1} are also structure preserving maps on $[0,1]^{r([k])}$.

Let ψ be a separable realization on $\mathbf{X}^{[k]}$. Then $\phi \circ \psi$ is also a separable realization on $\mathbf{X}^{[k]}$.

Let $\phi^{-1}(W)$ denote the Euclidean structure represented by $\{\phi_{\|p\|}^{-1}(W_p(R)) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$.

Lemma 3.2.3. *Given a Euclidean structure W of type τ and a structure preserving map ϕ on $[0,1]^{k_{max}}$, $\delta_w(W, \phi^{-1}(W)) = 0$.*

Proof. It is enough to show that $t(\mathcal{S}, W) = t(\mathcal{S}, \phi^{-1}(W)$ for all relational structures $\mathcal{S}$ of type τ on finite sets.

Just as in the proof of Theorem 2.4.5, we can lift structure preserving maps. We can obtain a *lifting* ψ of ϕ such that $\psi : [0,1]^{r([n],k)} \to [0,1]^{r([n],k)}$ is measure-preserving, $\psi \circ \tau = \tau \circ \psi$ for all $\tau \in S_{[n]}$ and $\phi \circ P_{r([k])} = P_{r([k])} \circ \psi$.

Let $\psi_m : [0,1]^{r([n],m)} \to [0,1]^{r([n],m)}$ be the natural restriction of ψ. Then $\psi = \psi_k$ and we can verify that ψ_m is a lifting of ϕ_m using the properties of structure preserving maps. In particular, each ψ_m is measure-preserving.

Let $\mathcal{S}$ be a relational structure of type τ on a finite set.

$$\begin{aligned}
t(\mathcal{S}, \phi^{-1}(W)) &= Vol\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(\phi^{-1}_{\|p\|}(W_p(R^{\boldsymbol{\mathcal{R}}}))))\right)\right)\\
&= Vol\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} \psi^{-1}_{\|p\|}(L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}}))))\right)\right)\\
&= Vol\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\circ\psi^{-1}_{\|p\|}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right)\\
&= Vol\left(\psi^{-1}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right)\right)\\
&= Vol\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k} P^{-1}_{r([\|p\|])}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(W_p(R^{\boldsymbol{\mathcal{R}}})))\right)\right)\\
&= t(\mathcal{S}, W)
\end{aligned}$$

□

A *random coordinate system* $\zeta_n : E_k([n]) \to [0,1]^{r([n],k))}$ can be defined as follows : Let x_A, $A \in r([n],k)$, be random variables chosen uniformly at random in $[0,1]$. Then for any $(n_1,\ldots,n_k) \in E_k([n])$ and $S = \{i_1,\ldots,i_t) \in r([k])$, $\zeta_{n,S}(n_1,\ldots,n_k) = x_{\{n_{i_1},\ldots,n_{i_t}\}}$.

For any Euclidean structure W of type τ, $(\zeta_n)^{-1}(W)$ is a relational structure-valued random variable that has the same distribution as $\mathcal{R}(W,n)$.

Let $\zeta_S = std \circ [\zeta_{n,S}]$ for all $S \in r([k])$. Then $\zeta = (\zeta_S : S \in r([k])) : \mathbf{X}^{[k]} \to [0,1]^{r([k])}$ is a random coordinate system on $\mathbf{X}^{[k]}$.

The random coordinate system ζ is a separable realization on $\mathbf{X}^{[k]}$ with probability 1.

Lemma 3.2.4. *Let W be a Euclidean structure of type τ. Let $\boldsymbol{\mathcal{R}}$ be the ultraproduct of the random sequence $\mathcal{R}(W,n)$. Then there exists a separable realization $\phi : \mathbf{X}^{k_{max}} \to [0,1]^{r([k_{max}])}$ such that $\boldsymbol{\mathcal{R}}$ is 0-close to $\phi^{-1}(W)$, with probability 1.*

The random relational structure $\boldsymbol{\mathcal{R}} = [\mathcal{R}(W,n)] = [\zeta_n^{-1}(W)]$, since $\mathcal{R}(W,n)$ has the same distribution as $\zeta_n^{-1}(W)$. Then $\boldsymbol{\mathcal{R}}$ and $\zeta^{-1}(W)$ are 0-close.

Theorem 3.2.5. *Let U and W be Euclidean structures of type τ. $\delta_w(U,W) = 0$ if and only if there exist structure-preserving maps ψ_U and ψ_W on $[0,1]^{r([k_{max}])}$ such that*

$$d_1(\psi_U^{-1}(U), \psi_W^{-1}(W)) = 0.$$

Proof. If there exist structure-preserving maps ϕ and ψ on $[0,1]^{r([k_{max}])}$ such that $d_1(\phi^{-1}(U), \psi^{-1}(W)) = 0$, then using Lemma 3.2.3 we can easily see that $\delta_w(U,W) = 0$.

For the converse, let $\delta_w(U,W) = 0$. Since $t(\mathcal{S},U) = t(\mathcal{S},W)$ for all relational structures $\mathcal{S}$ on finite sets, $t_{ind}(\mathcal{S},U) = t_{ind}(\mathcal{S},W)$. Then $\mathcal{R}(U,n)$ and $\mathcal{R}(W,n)$ have the same distribution, say, $\mathcal{Z}_n$. Let $\boldsymbol{\mathcal{R}} = [\mathcal{Z}_n]$. Then by Lemma 3.2.4, there exist separable realizations ϕ_U and ϕ_W on $\mathbf{X}^{k_{max}}$ such that $\phi_U^{-1}(U)$, $\phi_W^{-1}(W)$ and $\boldsymbol{\mathcal{R}}$ are all 0-close, with probability 1. We can find the separable systems $\{l_{A,U} : A \in r([k_{max}])\}$ and $\{l_{A,W} : A \in r([k_{max}])\}$ induced by ϕ_U and ϕ_W respectively. For each A, let l_A be the separable σ-algebra generated by $l_{A,U}$ and $l_{A,W}$. Using Theorem 2.4.3, we can find a separable realization ϕ on $\mathbf{X}^{[k_{max}]}$ that corresponds to the separable system $\{l_A : A \in r([k])\}$. We can now find structure preserving maps ψ_U and ψ_W on $[0,1]^{r([k_{max}])}$ such that for all m, all Lebesgue measurable $B \subseteq [0,1]^{r([m])}$, $\phi_m^{-1}(\psi_{U,m}^{-1}(B))$ and $\phi_{U,m}^{-1}(B)$ as well as $\phi_m^{-1}(\psi_{W,m}^{-1}(B))$ and $\phi_{W,m}^{-1}(B)$ differ on null sets. Then $\psi_U^{-1}(U)$ and $\psi_W^{-1}(W)$ are 0-close. $\square$

3.3 Regularity for Relational Structures

In the discussions to follow, if X is a finite set, then the measure μ is the normalized counting measure on X. In case of the ultraproduct $\mathbf{X}$, μ is the Loeb measure we defined in Section 2.

A *k-level ℓ-hyperpartition* $\mathscr{H}$ on a set X is a family of partitions of $E_r(X) = \bigcup_{j=1}^{\ell} P_j^r$, where $E_r(X)$ is the complete r-uniform hypergraph on X and P_j^r is an r-uniform hypergraph on X, for each $r \in [k]$ and $j \in [\ell]$. $\mathbf{a}, \mathbf{b} \in E_r(\mathbf{X})$ are in the same *$\mathscr{H}$-r-cell* if for each non-empty $A \subseteq [r]$, $\pi_A(\mathbf{a})$ and $\pi_A(\mathbf{b})$ are in the same $P_j^{|A|}$. This induces a partition of each $E_r(X)$ into $\mathscr{H}$-r-cells.

$\mathscr{H}$ is a *δ-equitable* hyperpartition if for all $r \in [k]$ and $i, j \in [l]$,

$$|\mu_{[r]}(P_i^r) - \mu_{[r]}(P_j^r)| < \delta.$$

A *cylindric intersection set* L in X^A is a set $L = \bigcap_{\{B: B \subsetneq A\}} \pi_B^{-1}(Y_B)$, where each $Y_B \subseteq X^B$ is measurable. Recall that if X is finite, all subsets are measurable.

If $\mathbf{L}$ is a cylindric intersection in the ultraproduct $\mathbf{X}^{[r]}$, $\mathbf{L} \in \sigma_{[r]}^*$.

In the case that $r = 2$, every cylindric intersection in $X^{[2]}$ is a measurable rectangle, whether X is a finite set or the ultraproduct.

A directed r-uniform hypergraph H on X is *ϵ-regular* if for any cylindric intersection set L in $X^{[r]}$ with $\mu_{[r]}(L) \geq \epsilon$,

$$|\mu_{[r]}(H \cap L) - \mu_{[r]}(H)\mu_{[r]}(L)| < \epsilon\, \mu_{[r]}(L).$$

Theorem 3.3.1. *Let $\mathcal{R}$ be a relational structure of type τ on an ultraproduct $\mathbf{X}$ of an increasing sequence of finite sets X_i. For every $\epsilon > 0$, there exists $\ell \in \mathbb{N}$ and a 0-equitable k_{max}-level ℓ-hyperpartition $\mathscr{H}$ such that*

1. *Each $\mathbf{P}^{\mathbf{r}}_{\mathbf{j}} \in \sigma_{[r]}$ and it is independent of $\sigma^*_{[r]}$.*
2. *For all $R \in \tau$, $\mu_{[\|p\|]}(H_p(R^{\mathcal{R}}) \Delta T_p(R)) < \epsilon$, where $T_p(R)$ is a union of some $\mathscr{H}$-$\|p\|$-cells.*

Proof. Let ϕ be a separable realization and let W be the corresponding Euclidean structure such that $\boldsymbol{\mathcal{R}}$ and $\phi^{-1}(W)$ are 0-close. Then there exist measurable $W_p(R^{\mathcal{R}}) \subseteq [0,1]^{r([\|p\|])}$ such that $\mu_{[\|p\|]}(H_p(R^{\mathcal{R}}) \Delta \phi^{-1}_{\|p\|}(W_p(R^{\mathcal{R}}))) = 0$. An *$\ell$-box* in $[0,1]^{r([\|p\|])}$ is $\prod_{A \in r([\|p\|])}[\frac{j_A - 1}{\ell}, \frac{j_A}{\ell})$ where each $j_A \in [\ell]$. There exists $\ell \in \mathbb{N}$ such that for each $R \in \tau$ and each $p \in \mathcal{P}_k$ where $k = ar(R)$, there exists $W'_p(R)$, a union of ℓ-boxes in $[0,1]^{r([\|p\|])}$, and $\mathrm{Vol}(W_p(R^{\mathcal{R}}) \Delta W'_p(R)) < \epsilon$. So $\mu_{[\|p\|]}(H_p(R^{\mathcal{R}}) \Delta \phi^{-1}_{\|p\|}(W'_p(R))) < \epsilon$. Now consider the hyperpartition $\mathscr{H}$ formed by $\mathbf{P}^{\mathbf{r}}_{\mathbf{j}} = \phi_r^{-1}([\frac{j-1}{\ell}, \frac{j}{\ell}))$, for each $r \in [k_{max}]$ and $j \in [\ell]$. Clearly each $\mathbf{P}^{\mathbf{r}}_{\mathbf{j}}$ is in the separable algebra $l_{[r]} \leq \sigma_{[r]}$ and thus it is independent of $\sigma^*_{[r]}$. We may assume $\mathbf{P}^{\mathbf{r}}_{\mathbf{j}} \subseteq E_r(\mathbf{X})$. So $\mathbf{P}^{\mathbf{r}}_{\mathbf{j}}$ is an r-uniform hypergraph on $\mathbf{X}$. Also $\mu_{[r]}(\mathbf{P}^{\mathbf{r}}_{\mathbf{j}}) = \frac{1}{\ell}$, so the partition is 0-equitable.

Since $W'_p(R)$ is a union of ℓ-boxes, $T_p(R) = \phi^{-1}_{\|p\|}(W'_p(R))$ is a union of $\mathscr{H}$-$\|p\|$-cells. $\square$

Note that if $\mathbf{E} \in \sigma_{[r]}$ is independent of $\sigma^*_{[r]}$, then $\mathbf{E}$ is, in fact, δ-regular for all $\delta > 0$.

Theorem 3.3.2. *Given $\epsilon > 0$ and $m \in \mathbb{N}$, there exist $M, N \in \mathbb{N}$ such that given any relational structure $\mathcal{R}$ on a finite set X with $|X| \geq N$, there exists an ϵ-equitable k_{max}-level ℓ-hyperpartition $\mathscr{H}$ on X for some ℓ such that $m \leq \ell \leq M$ and*

1. *Each P^r_j is ϵ-regular.*
2. *For each $H_p(R^{\mathcal{R}})$, there exists $T_p(R)$, a union of some $\mathscr{H}$-$\|p\|$-cells, such that*

$$\mu_{[\|p\|]}(H_p(R^{\mathcal{R}}) \Delta T_p(R)) < \epsilon$$

Proof. Suppose, for contradiction, there exist $\epsilon > 0$, $m \in \mathbb{N}$ and relational structures $\mathcal{R}_i$ on X_i such that $|X_i| \to \infty$ and there is no ϵ-equitable ℓ-hyperpartition, for $m \leq \ell \leq i$ satisfying above conditions for any $\mathcal{R}_i$.

Now take ultraproducts to obtain the relational structure $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$. There exists a 0-equitable ℓ-hyperpartition $\boldsymbol{\mathscr{H}}$ on $\mathbf{X}$ satisfying the conditions in Theorem 3.3.1 for $\frac{\epsilon}{2}$. Let $\ell \geq m$. Let $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}} = [P^r_{j,i}]$ be an internal set that differs from $\mathbf{P}^{\mathbf{r}}_{\mathbf{j}}$ by a null set. We may assume each $P^r_{j,i}$ is an r-uniform hypergraph on X_i. The hypergraphs $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}}$ form a 0-equitable ℓ-hyperpartition $\widetilde{\boldsymbol{\mathscr{H}}}$.

For U-almost every i, the hypergraphs $P^r_{j,i}$ form an ℓ-hyperpartition $\mathscr{H}_i$ on X_i. Since $\widetilde{\boldsymbol{\mathscr{H}}}$ is a 0-equitable hyperpartition, $\mu_{[r]}(\mathbf{P}^{\mathbf{r}}_{\mathbf{j}}) = \mu_{[r]}(\mathbf{P}^{\mathbf{r}}_{\mathbf{j'}})$ for every $j, j' \in [\ell]$ and $r \in [k_{max}]$. Then $\lim_U |\mu_{i,[r]}(P^r_{j,i}) - \mu_{i,[r]}(P^r_{j',i})| = 0$.

For each $R \in \tau$, there must exist internal sets $T'_p(R) = [T_{p,i}(R)]$, such that each $T_{p,i}(R)$ is a union of $\mathscr{H}_i$-$\|p\|$-cells and $\mu_{[\|p\|]}(H_p(R^{\boldsymbol{\mathcal{R}}}) \Delta T'_p(R)) < \frac{\epsilon}{2}$.

Then $\lim_U \mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}))\Delta T_{p,i}(R)) < \frac{\epsilon}{2}$ and $\mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}))\Delta T_{p,i}(R)) < \epsilon$ for U-almost every i.

For U-almost every i, for each $r \in [k_{max}]$ and $j, j' \in [\ell]$, $\lim_U |\mu_{i,[r]}(P^r_{j,i}) - \mu_{i,[r]}(P^r_{j',i})| < \epsilon$, that is, $\mathscr{H}_i$ is an ϵ-equitable ℓ-hyperpartition, and for each $R \in \tau$ and $p \in \mathcal{P}_k$ where $k = ar(R)$, $\mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}))\Delta T_{p,i}(R)) < \epsilon$.

Now we show that for U-almost every i, $P^j_{r,i}$ is ϵ-regular. Suppose there is $r \in [k_{max}]$ and $j \in [\ell]$ such that for U-almost every i, there exists a cylindric intersection set $L_i \subseteq E_r(X_i)$ with $\mu_{i,[r]}(L_i) \geq \epsilon$ and $|\mu_{i,[r]}(P^r_{j,i} \cap L_i) - \mu_{i,[r]}(P^r_{j,i})\mu_{i,[r]}(L_i)| \geq \epsilon\, \mu_{i,[r]}(L_i)$.

Consider the cylindric intersection set $\mathbf{L} = [L_i]$. So $\mathbf{L} \in \sigma^*_{[r]}$ and it is independent of $\mathbf{P^r_j}$. Therefore, $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}}$ and $\mathbf{L}$ are also independent. However $\mu_{[r]}(\mathbf{L}) \geq \epsilon$ and $|\mu_{[r]}(\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}} \cap \mathbf{L}) - \mu_{[r]}(\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}})\mu_{[r]}(\mathbf{L})| = \lim_U |\mu_{i,[r]}(P^r_{j,i} \cap L_i) - \mu_{i,[r]}(P^r_{j,i})\mu_{i,[r]}(L_i)| \geq \epsilon\, \mu_{[r]}(\mathbf{L})$. This implies $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}}$ and $\mathbf{L}$ cannot be independent, leading to a contradiction. Thus, $P^r_{j,i}$ is ϵ-regular, for each $r \in [k_{max}]$, $j \in [\ell]$ and U-almost every i.

Let $\mathcal{R}'_i$ be a new relational structure of type τ on X_i, where each $R^{\mathcal{R}'_i}$ is represented by the system $\{T_{p,i}(R) : p \in \mathcal{P}_k, k = ar(R)\}$. Since there are only finitely many pairs (r, j) with $r \in [k_{max}]$, $j \in [\ell]$ and finitely many relations in the type τ, we have proved, for U-almost every i :

1. for each $r \in [k_{max}]$ and $j \in [\ell]$, $P^r_{j,i}$ is ϵ-regular,
2. $\mathscr{H}_i$ is an ϵ-equitable ℓ-hyperpartition, and
3. for each $R \in \tau$ with $k = ar(R)$ and each $p \in \mathcal{P}_k$,

$$\mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}))\Delta T_{p,i}(R)) < \epsilon.$$

Since $|X_i| \to \infty$, there exists i such that $|X_i| > \ell$ and then $\mathcal{R}'_i$ contradicts our initial assumption. $\square$

Let us consider the type $\tau = \{E\}$ with $ar(E) = 2$ that we use for graphs and digraphs to see how our results extend the familiar notions of regularity for graphs and digraphs. In particular, let us examine how our regular edge cells imply the regularity conditions involving edge densities of pairs of sets. Theorem 3.3.2 transfers regularity from the ultraproduct to the sequence, so let us consider $\mathbf{G}$ on the ultraproduct $\mathbf{X}$. We know $H_{(1,2)}(E^{\mathbf{G}}) = \emptyset$. By Theorem 3.3.1, there exists a partition $\{\mathbf{V}_1, \ldots, \mathbf{V}_\ell\}$ of $\mathbf{X}$ and a partition $\{\mathbf{E}_1, \ldots, \mathbf{E}_\ell\}$ of $\mathbf{X}^{[2]}$ such that

1. For each $i \in [\ell]$, $\mu_{[1]}(\mathbf{V}_i) = \frac{1}{\ell}$ and $\mu_{[2]}(\mathbf{E}_i) = \frac{1}{\ell}$.

2. Each $\mathbf{E}_k$ is independent of measurable rectangles, that is, for all measurable $\mathbf{A}, \mathbf{B} \subseteq \mathbf{X}$,
$$\mu_{[2]}(\mathbf{E}_k \cap \mathbf{A} \times \mathbf{B}) = \mu_{[2]}(\mathbf{E}_k)\mu_{[2]}(\mathbf{A} \times \mathbf{B}) = \frac{1}{\ell}\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B}).$$

3. There exists $C \subseteq [\ell]^3$ such that $\mu_{[2]}(E^{\mathbf{G}} \Delta \bigcup_{(i,j,k)\in C}(\mathbf{V}_i \times \mathbf{V}_j) \cap \mathbf{E}_k) < \epsilon$.

Given $\mathbf{A} \subseteq \mathbf{V}_i$ and $\mathbf{B} \subseteq \mathbf{V}_j$, let $E^{\mathbf{G}}(\mathbf{A}, \mathbf{B}) = E^{\mathbf{G}} \cap (\mathbf{A} \times \mathbf{B})$.

$\mu_{[2]}(E^{\mathbf{G}}(\mathbf{A}, \mathbf{B}) \Delta \bigcup_{\{k:(i,j,k)\in C\}}(\mathbf{A} \times \mathbf{B}) \cap \mathbf{E}_k) < \epsilon$, since all 2-cells also form a partition of $\mathbf{X}^{[2]}$. If $\mu_{[2]}(\mathbf{A} \times \mathbf{B}) > 0$ and $d_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) = \frac{\mu_{[2]}(E^{\mathbf{G}}(\mathbf{A},\mathbf{B}))}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})}$, then

$$\left| d_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) - \sum_{\{k:(i,j,k)\in C\}} \frac{1}{\ell} \right| < \epsilon.$$

This is true for all $\mathbf{A} \subseteq \mathbf{V}_i$ and $\mathbf{B} \subseteq \mathbf{V}_j$, and in particular for $\mathbf{V}_i$ and $\mathbf{V}_j$. So

$|d_{\mathbf{G}}(\mathbf{A},\mathbf{B}) - d_{\mathbf{G}}(\mathbf{V}_i,\mathbf{V}_j)| < 2\epsilon$.

In particular for graphs, $E^{\mathbf{G}}$ is symmetric, so we can assume C has $S_{[2]}$-symmetry, that is, in the approximation for $E^{\mathbf{G}}$ using 2-cells of the partition, we include $(\mathbf{V}_i \times \mathbf{V}_j) \cap \mathbf{E}_k$ if and only if we include $(\mathbf{V}_j \times \mathbf{V}_i) \cap \mathbf{E}_k$. This is similar to the usual graph regularity condition involving edge densities.

In the case of digraphs, typically four edge densities are used. In general, $E^{\mathbf{G}}$ need not be symmetric. Let $(E^{\mathbf{G}})^c$ denote the edge set that is the complement of $E^{\mathbf{G}}$. Let $\pi = (1\ 2) \in S_{[2]}$.

$$\begin{aligned}\overleftrightarrow{d}_{\mathbf{G}}(\mathbf{A},\mathbf{B}) &= \frac{\mu_{[2]}(E^{\mathbf{G}}(\mathbf{A},\mathbf{B}) \cap E^{\mathbf{G}}(\mathbf{B},\mathbf{A}))}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})} \\ &= \frac{\mu_{[2]}((E^{\mathbf{G}} \cap (E^{\mathbf{G}})^{\pi})(\mathbf{A},\mathbf{B}))}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})}\end{aligned}$$

$$\begin{aligned}\overrightarrow{d}_{\mathbf{G}}(\mathbf{A},\mathbf{B}) &= \frac{\mu_{[2]}(E^{\mathbf{G}}(\mathbf{A},\mathbf{B}) \cap ((E^{\mathbf{G}})^c(\mathbf{B},\mathbf{A}))^{\pi})}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})} \\ &= \frac{\mu_{[2]}((E^{\mathbf{G}} \cap ((E^{\mathbf{G}})^c)^{\pi})(\mathbf{A},\mathbf{B}))}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})}\end{aligned}$$

Similarly,

$$\overleftarrow{d}_{\mathbf{G}}(\mathbf{A},\mathbf{B}) = \frac{\mu_{[2]}(((E^{\mathbf{G}})^{\pi} \cap (E^{\mathbf{G}})^c)(\mathbf{A},\mathbf{B}))}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})}$$

$$\bar{d}_{\mathbf{G}}(\mathbf{A},\mathbf{B}) = \frac{\mu_{[2]}((((E^{\mathbf{G}})^c \cap ((E^{\mathbf{G}})^c)^{\pi})(\mathbf{A},\mathbf{B}))}{\mu_{[1]}(\mathbf{A})\mu_{[1]}(\mathbf{B})}$$

Also

$$\mu_{[2]}((E^{\mathbf{G}})^c \,\Delta \bigcup_{(i,j,k)\notin C} (\mathbf{V}_i \times \mathbf{V}_j) \cap \mathbf{E}_k) < \epsilon$$

$$\mu_{[2]}((E^{\mathbf{G}})^\pi \,\Delta \bigcup_{(i,j,k)\in C} (\mathbf{V}_j \times \mathbf{V}_i) \cap \mathbf{E}_k) < \epsilon$$

Thus, all four edge sets $(E^{\mathbf{G}} \cap (E^{\mathbf{G}})^\pi)$, $(E^{\mathbf{G}} \cap ((E^{\mathbf{G}})^c)^\pi)$, $(E^{\mathbf{G}})^\pi \cap (E^{\mathbf{G}})^c)$ and $((E^{\mathbf{G}})^c \cap ((E^{\mathbf{G}})^c)^\pi)$ can be approximated by unions of 2-cells. Once again using the fact that edge cells are independent of measurable rectangles, we can show that given $\mathbf{A} \subseteq \mathbf{V}_i$, $\mathbf{B} \subseteq \mathbf{V}_j$ such that $\mu_{[2]}(\mathbf{A} \times \mathbf{B}) > 0$, $|\overleftrightarrow{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) - \overleftrightarrow{d}_{\mathbf{G}}(\mathbf{V}_i, \mathbf{V}_j)| < 2\epsilon$, $|\overrightarrow{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) - \overrightarrow{d}_{\mathbf{G}}(\mathbf{V}_i, \mathbf{V}_j)| < 2\epsilon$, $|\overleftarrow{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) - \overleftarrow{d}_{\mathbf{G}}(\mathbf{V}_i, \mathbf{V}_j)| < 2\epsilon$ and $|\overline{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) - \overline{d}_{\mathbf{G}}(\mathbf{V}_i, \mathbf{V}_j)| < 2\epsilon$.

This yields a regularity condition for digraphs using four edge densities, similar to the one commonly used for the Digraph Regularity lemma.

Of course if we are dealing with graphs, then $E^{\mathbf{G}}$ are $(E^{\mathbf{G}})^c$ are symmetric and $\overleftrightarrow{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) = d_{\mathbf{G}}(\mathbf{A}, \mathbf{B})$ and $\overrightarrow{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) = \overleftarrow{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) = 0$ and $\overline{d}_{\mathbf{G}}(\mathbf{A}, \mathbf{B}) = 1 - d_{\mathbf{G}}(\mathbf{A}, \mathbf{B})$. Thus we retrieve the regularity condition we examined above in the special case of graphs.

3.4 Removal Lemmas for Relational Structures

Let $\mathcal{R}$ and $\mathcal{S}$ be relational structures of type τ on a set X. $\mathcal{R}$ and $\mathcal{S}$ are *ϵ-close* if for each $R \in \tau$ with $k = ar(R)$, $\mu_{[||p||]}(H_p(R^{\mathcal{R}}) \Delta H_p(R^{\mathcal{S}})) < \epsilon$ for each $p \in \mathcal{P}_k$.

$\mathcal{S}$ is a *substructure* of $\mathcal{R}$ if for each $R \in \tau$, $R^{\mathcal{S}} \subseteq R^{\mathcal{R}}$. Equivalently, $H_p(R^{\mathcal{S}}) \subseteq H_p(R^{\mathcal{R}})$ for each $p \in \mathcal{P}_k$, where $k = ar(R)$.

Let $\mathcal{R} \backslash \mathcal{S}$ denote the relational structure represented by $\{H_p(R^{\mathcal{R}}) \backslash H_p(R^{\mathcal{S}}) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$.

First we prove an infinite version of Removal for relational structures on an ultraproduct space.

Theorem 3.4.1. *Let $\boldsymbol{\mathcal{R}}$ be a relational structure on $\mathbf{X}$ such that for each $R \in \tau$, each $p \in \mathcal{P}_k$ where $k = ar(R)$, $H_p(R^{\boldsymbol{\mathcal{R}}})$ is a measurable subset of $\mathbf{X}^{\|p\|}$. Then there exists a relational structure $\boldsymbol{\mathcal{N}}$ on $\mathbf{X}$ such that for each R and each $p \in \mathcal{P}_k$ where $k = ar(R)$, $\mu_{[\|p\|]}(H_p(R^{\boldsymbol{\mathcal{N}}})) = 0$ and for every n and every relational structure $\mathcal{S}$ on $[n]$, $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}}) = \emptyset$ or $\mu_{[n]}(T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}})) > 0$.*

Proof. We know there exists a separable realization ϕ on $\mathbf{X}^{[k_{max}]}$ and a corresponding Euclidean structure $W(\boldsymbol{\mathcal{R}})$ such that $\mu_{[\|p\|]}(H_p(R^{\boldsymbol{\mathcal{R}}})\Delta\phi^{-1}_{\|p\|}(W_p(R^{\boldsymbol{\mathcal{R}}}))) = 0$ for each $R \in \tau$ and $p \in \mathcal{P}_k$ where $k = ar(R)$. Let $D_p(R) \subseteq [0,1]^{r([\|p\|])}$ be the set of density points of $W_p(R^{\boldsymbol{\mathcal{R}}})$. By Lebesgue's density theorem, $\mathrm{Vol}(W_p(R^{\boldsymbol{\mathcal{R}}})\Delta D_p(R)) = 0$. There exists a system of directed hypergraphs $\{H'_p(R) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$ such that each $H'_p(R) \subseteq \phi^{-1}_{\|p\|}(D_p(R))$ and $\mu_{[\|p\|]}(\phi^{-1}_{\|p\|}(D_p(R)) \setminus H'_p(R)) = 0$. Let $\boldsymbol{\mathcal{R}}'$ be the relational structure on $\mathbf{X}$ represented by this system.

For any n, there exists a lifting ψ of ϕ to $\mathbf{X}^{[n]}$, by Theorem 2.4.5. Let $\mathcal{S}$ be a relational structure on $[n]$. Since every point of $D_p(R)$ is a density point of the set, either $D_p(R) = \emptyset$ or $\mathrm{Vol}(D_p(R)) > 0$. Then for each R, every $p \in \mathcal{P}_k$ and every $e \in H_p(R^{\mathcal{S}})$, $L_e^{-1}(F_e(D_p(R)))$ is empty or has positive measure. As a result, $\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k}\psi^{-1}_{\|p\|}(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(D_p(R))))$ is empty or has positive measure.

$$T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}') \subseteq \bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k}\psi^{-1}_{\|p\|}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(D_p(R)))\right)$$

and

$$\mu_{[n]}(T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}')) = \mu_{[n]}\left(\bigcap_{R\in\tau}\bigcap_{p\in\mathcal{P}_k}\psi^{-1}_{\|p\|}\left(\bigcap_{e\in H_p(R^{\mathcal{S}})} L_e^{-1}(F_e(D_p(R)))\right)\right)$$

Thus, $T_o(\mathcal{S}, \mathcal{R}')$ is either empty or has positive measure. Let $\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{R}}'$, so $R^{\boldsymbol{\mathcal{N}}} = R^{\boldsymbol{\mathcal{R}}} \setminus R^{\boldsymbol{\mathcal{R}}'}$. Then $H_p(R^{\boldsymbol{\mathcal{N}}}) = H_p(R^{\boldsymbol{\mathcal{R}}}) \setminus H_p(R^{\boldsymbol{\mathcal{R}}'})$ for each $R \in \tau$ and $p \in \mathcal{P}_k$ and $\mu_{[\|p\|]}(H_p(R^{\boldsymbol{\mathcal{N}}})) = 0$.

Now $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}}) = T_o(S, \boldsymbol{\mathcal{R}} \cap \boldsymbol{\mathcal{R}}')$. Note that $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \cap \boldsymbol{\mathcal{R}}') \subseteq T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}')$ and $\mu_{[n]}(T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \cap \boldsymbol{\mathcal{R}}')) = \mu_{[n]}(T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}'))$.

Therefore, $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}}) = \emptyset$ or $\mu_{[n]}(T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}})) > 0$. $\square$

Theorem 3.4.2. *Given $n \in \mathbb{N}$, a relational structure $\mathcal{S}$ of type τ on $[n]$ and given $\epsilon > 0$, there exists $\delta > 0$, such that for any relational structure $\mathcal{R}$ of type τ on a finite set X with $t(\mathcal{S}, \mathcal{R}) < \delta$, there exists a substructure $\mathcal{R}'$ of $\mathcal{R}$ such that $t_o(\mathcal{S}, \mathcal{R}') = 0$ and $\mathcal{R}'$ is ϵ-close to $\mathcal{R}$.*

Proof. Suppose, for contradiction, there exists $n \in \mathbb{N}$, a relational structure $\mathcal{S}$ on $[n]$ and $\epsilon > 0$ such that for all i, there is a relational structure $\mathcal{R}_i$ on a finite set X_i with $t(\mathcal{S}, \mathcal{R}_i) < \frac{1}{i}$ but no such $\mathcal{R}'_i$ exists. So $\lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i) = 0$.

We may assume $|X_i| \to \infty$. Then we can take the ultraproduct $\mathbf{X}$ and obtain the relational structure $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$, represented by the system of directed hypergraphs $\{H_p(R^{\boldsymbol{\mathcal{R}}}) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$. There exists a relational structure $\boldsymbol{\mathcal{N}}$ on $\mathbf{X}$ as described in Theorem 3.4.1 .

Now $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t_o(\mathcal{S}, \mathcal{R}_i) = 0$. This implies $t_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}}) = 0$. By Theorem 3.4.1 , $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}}) = \emptyset$ and each $H_p(R^{\boldsymbol{\mathcal{N}}})$ is a null set in $\mathbf{X}^{[\|p\|]}$. By the definition of a null set, for each $\delta > 0$ there is an internal set $\mathbf{I}_p(R) = [I_{p,i}(R)]$ such that $H_p(R^{\boldsymbol{\mathcal{N}}}) \subseteq \mathbf{I}_p(R)$ and $\mu_{[\|p\|]}(\mathbf{I}_p(R)) < \delta$. Let $\delta = \frac{\epsilon}{2}$.

Now consider the structure $\mathcal{R}_i'$ on X_i represented by $\{H_p(R^{\mathcal{R}_i}) \setminus I_{p,i}(R) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$. Let $\boldsymbol{\mathcal{R}}'$ be the ultraproduct of the sequence $\{\mathcal{R}_i'\}$. Since $H_p(R^{\boldsymbol{\mathcal{R}}'}) \subseteq (H_p(R^{\boldsymbol{\mathcal{R}}}) \setminus \mathbf{I}_p(R)) \subseteq (H_p(R^{\boldsymbol{\mathcal{R}}}) \setminus H_p(R^{\boldsymbol{\mathcal{N}}}))$, we have $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}') \subseteq T_o(\mathcal{S}, \boldsymbol{\mathcal{R}} \setminus \boldsymbol{\mathcal{N}})$. This implies $T_o(\mathcal{S}, \boldsymbol{\mathcal{R}}') = [T_o(\mathcal{S}, \mathcal{R}_i')] = \emptyset$. Thus, for U-almost every i, $t_o(\mathcal{S}, \mathcal{R}_i') = 0$.

Now for each $R \in \tau$ with $k = ar(R)$ and each $p \in \mathcal{P}_k$, $\mu_{[\|p\|]}(\mathbf{I}_p(R)) = std(\nu_{[\|p\|]}(\mathbf{I}_p(R))) < \delta$. So for each $m \in \mathbb{N}$, $\nu_{[\|p\|]}(\mathbf{I}_p(R)) < (\delta + \frac{1}{m})^*$. Choose m such that $\frac{1}{m} < \frac{\epsilon}{2}$, then $\delta + \frac{1}{m} < \epsilon$. For U-almost every i, $\mu_{i,[\|p\|]}(I_{p,i}) \leq \epsilon$. We know that $H_p(R^{\mathcal{R}_i}) \setminus H_p(R^{\mathcal{R}_i'}) \subseteq I_{p,i}(R)$. Since the type τ is finite and each $\mathcal{P}_k$ is finite, we can now prove for U-almost every i, for each $R \in \tau$ and $p \in \mathcal{P}_k$, $\mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}) \setminus H_p(R^{\mathcal{R}_i'})) \leq \mu_{i,[\|p\|]}(I_{p,i}(R)) \leq \epsilon$.

This contradicts our supposition, thus proving our result. $\square$

Let us apply Theorem 3.4.2 to the case of graphs using type $\tau = \{E\}$. Recall that $H_{(1,2)}(E^G) = \emptyset$ for all graphs G. Given a graph F on $[n]$ and a large finite graph G on X, the condition $t(F, G) < \delta$ implies that there are fewer than $\delta|X|^n$ copies of F in G. Theorem 3.4.2 tells us there exists an F-free subgraph G' of G that is ϵ-close to G. This simply means G' was obtained by deleting less than $\epsilon|X|^2$ edges from G. This is the familiar Graph Removal Lemma.

Now we prove a Strong Removal Lemma for relational structures that allows us to simultaneously remove copies of a family of relational structures. This result also implies the testability of hereditary properties of relational structures. We shall give a brief introduction to these ideas from the area of property testing following the proof of the theorem.

Theorem 3.4.3. *Let $\mathscr{F}$ be a family of relational structures of type τ on finite sets. For every $\epsilon > 0$, there exists $\delta > 0$ and $n \in \mathbb{N}$ such that for every relational structure $\mathcal{R}$ on some finite set X satisfying $t_{ind}(\mathcal{S}, \mathcal{R}) \leq \delta$ for each $\mathcal{S} \in \mathscr{F}$ on some set Y with $|Y| \leq n$, there exists $\mathcal{R}'$ on X such that $t_{o,ind}(\mathcal{S}, \mathcal{R}') = 0$ for each $\mathcal{S} \in \mathscr{F}$ and $\mathcal{R}'$ is ϵ-close to $\mathcal{R}$.*

Proof. Suppose, for contradiction, there exists $\epsilon > 0$ and a sequence $\{\mathcal{R}_i\}$ such that $\mathcal{R}_i$ is a relational structure on X_i, $|X_i| \to \infty$ and $t_{ind}(\mathcal{S}, \mathcal{R}_i) \to 0$ for all $\mathcal{S} \in \mathcal{F}$, but no $\mathcal{R}_i'$ as described above exists. Then we can take the ultraproduct $\mathbf{X}$ and obtain the relational structure $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$, represented by the system of directed hypergraphs $\{H_p(R^{\boldsymbol{\mathcal{R}}}) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$. We know there exists a separable realization ϕ on $\mathbf{X}^{[k_{max}]}$ and a Euclidean structure $W(\boldsymbol{\mathcal{R}})$ such that

$$\mu_{[||p||]}(\phi^{-1}_{||p||}(W_p(R^{\boldsymbol{\mathcal{R}}}))\Delta H_p(R^{\boldsymbol{\mathcal{R}}})) = 0.$$

Then $t_{ind}(\mathcal{S}, W) = t_{ind}(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_U t_{ind}(\mathcal{S}, \mathcal{R}_i) = 0$ for each $\mathcal{S} \in \mathscr{F}$.

For every $R \in \tau$ and $p \in \mathcal{P}_k$ where $k = ar(R)$, $W_p(R^{\boldsymbol{\mathcal{R}}}) \subseteq [0,1]^{r[||p||]}$ is measurable. Given $\epsilon > 0$, there exists $\ell \in \mathbb{N}$ such that for each $R \in \tau$ and $p \in \mathcal{P}_k$, there is $W_p'(R)$, a union of ℓ-boxes in $[0,1]^{r([||p||])}$ such that $\mathrm{Vol}(W_p(R^{\boldsymbol{\mathcal{R}}})\Delta W_p'(R)) < \frac{\epsilon}{4}$. Let W' be the Euclidean structure represented by $\{W_p'(R) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$. Recall that an ℓ-box in $[0,1]^{r[||p||]}$ is of the form $\prod_{A \in r([||p||])} [\frac{j_A - 1}{\ell}, \frac{j_A}{\ell})$ where each $j_A \in [\ell]$. We can represent this ℓ-box by a function $f : r([||p||]) \to [\ell]$ where $f(A) = j_A$. Let $C_p(R)$ be the set containing functions that represent the ℓ-boxes in $W'_p(R)$.

Let $\mathbf{P^r_j} = \phi^{-1}_{[r]}([\frac{j-1}{\ell}, \frac{j}{\ell}))$. We may assume that every tuple in this set has distinct coordinates. Then $\mathbf{P^r_j}$ is an r-uniform hypergraph on $\mathbf{X}$. $\mathbf{P^r_1}, \ldots, \mathbf{P^r_\ell}$ form a partition of $\mathbf{X}^{[r]}$, within measure 0, for each $r \in [k_{max}]$. We call the resulting ℓ-hyperpartition $\boldsymbol{\mathscr{H}}$.

Since each $\mathbf{P^r_j}$ is measurable, there is an internal set $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}} = [P^r_{j,i}]$ that differs from $\mathbf{P^r_j}$ by a null set. These internal sets form an ℓ-hyperpartition $\widetilde{\boldsymbol{\mathscr{H}}}$. We may assume each $P^r_{j,i}$ is an r-uniform hypergraph on $X_i^{[r]}$. Let us call the resulting ℓ-hyperpartition $\mathscr{H}_i$. For convenience, let us assume $X_i = [n_i]$. Let $Q_{i,p}(R)$ be the union of $\mathscr{H}_i$-r-cells indexed by $C_p(R)$ and let $\widetilde{\mathcal{R}}_i$ be the relational structure on $[n_i]$ represented by $\{Q_{i,p}(R) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$.

Consider a random relational structure $\mathcal{R}(W', \mathscr{H}_i, n_i)$ of type τ on $X_i = [n_i]$ generated using a hyperpartition sampling. It differs from $\mathcal{R}(W', n_i)$ in that every x_A, $A \in r([n_i], k_{max})$, is chosen uniformly at random in $[\frac{g(A)-1}{\ell}, \frac{g(A)}{\ell})$, where $g(A) = j$ if A is in $P^{|A|}_{j,i}$. Recall that each $P^r_{j,i}$ is an r-uniform hypergraph. Although $\mathcal{R}(W', \mathscr{H}_i, n_i)$ is a random structure, it always takes the same value : $\mathcal{R}(W', \mathscr{H}_i, n_i) = \widetilde{\mathcal{R}}_i$. Let $\widetilde{\boldsymbol{\mathcal{R}}}$ be the ultraproduct of the sequence $\widetilde{\mathcal{R}}_i$.

Let $Q_p(R) = \phi^{-1}_{\|p\|}(W'_p(R))$ and let $\boldsymbol{\mathcal{Q}}$ be the relational structure on $\mathbf{X}$ represented by $\{Q_p(R) : p \in \mathcal{P}_k, k = ar(R), R \in \tau\}$. $Q_p(R)$ is a union of $\widetilde{\boldsymbol{\mathscr{H}}}$-$\|p\|$-cells indexed by $C_p(R)$. Then $\mu_{[\|p\|]}(Q_p(R)\Delta H_p(R^{\widetilde{\boldsymbol{\mathcal{R}}}})) = 0$. Since $Vol(W_p(R^{\boldsymbol{\mathcal{R}}})\Delta W'_p(R)) < \frac{\epsilon}{4}$, we have $\mu_{[\|p\|]}(H_p(R^{\boldsymbol{\mathcal{R}}})\Delta H_p(R^{\widetilde{\boldsymbol{\mathcal{R}}}})) < \frac{\epsilon}{4}$.

Finally, let us consider the random relation $\mathcal{R}'_i = \mathcal{R}(W, \mathscr{H}_i, n_i)$.

We know $E(t_{o,ind}(\mathcal{S}, \mathcal{R}(W, n_i))) = t_{ind}(\mathcal{S}, W) = 0$ for each $\mathcal{S} \in \mathcal{F}$, and this implies $t_{o,ind}(\mathcal{S}, \mathcal{R}(W, n_i)) = 0$ and $T_{o,ind}(\mathcal{S}, \mathcal{R}(W, n_i)) = \emptyset$ with probability 1.

Since $T_{o,ind}(\mathcal{S}, \mathcal{R}(W, \mathscr{H}_i, n_i)) \subseteq T_{o,ind}(\mathcal{S}, \mathcal{R}(W, n_i))$, we infer that $t_{o,ind}(\mathcal{S}, \mathcal{R}'_i) = 0$ for each $\mathcal{S} \in \mathcal{F}$ with probability 1.

Now $\lim_U E(\mu_{i,[\|p\|]}(Q_{i,p}(R) \Delta H_p(R^{\mathcal{R}'_i}))) = \mathrm{Vol}(W'_p(R) \Delta W_p(R^{\boldsymbol{\mathcal{R}}})) < \frac{\epsilon}{4}$.

Also $\lim_U \mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}) \Delta Q_{i,p}(R)) = \mu_{[\|p\|]}(H_p(R^{\boldsymbol{\mathcal{R}}}) \Delta H_p(R^{\widetilde{\boldsymbol{\mathcal{R}}}})) < \frac{\epsilon}{4}$.

So $\lim_U \mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}) \Delta H_p(R^{\mathcal{R}'_i})) < \frac{\epsilon}{2}$. Then for U-almost every i, for each $R \in \tau$ and $p \in \mathcal{P}_k$ where $k = ar(R)$, $\mu_{i,[\|p\|]}(H_p(R^{\mathcal{R}_i}) \Delta H_p(R^{\mathcal{R}'_i})) < \epsilon$ and $t_{o,ind}(\mathcal{S}, \mathcal{R}'_i) = 0$ for each $\mathcal{S} \in \mathscr{F}$ with probability 1, which contradicts our initial assumption. $\square$

Fix a relational type τ. Let $\mathscr{P}$ be a property of relational structures of type τ on finite sets. The property $\mathscr{P}$ is often identified with the family of relational structures on finite sets, up to isomorphism, that satisfy $\mathscr{P}$. A relational structure $\mathcal{R}$ of type τ is said to be *ϵ-far* from satisfying $\mathscr{P}$ if there is no relational structure $\widetilde{\mathcal{R}}$ of type τ that satisfies $\mathscr{P}$ and is ϵ-close to $\mathcal{R}$.

We assume there exists an oracle that given n and a relational structure $\mathcal{R}$ on an underlying set X such that $|X| = n$, tells us for every $R \in \tau$ with $k = ar(R)$, whether any k-tuple of vertices is in $R^{\mathcal{R}}$ or not. An *ϵ-test* for $\mathscr{P}$ is a (randomized) algorithm which, given the ability to query the oracle, distinguishes with high probability between the case of $\mathcal{R}$ satisfying $\mathscr{P}$ and the case of $\mathcal{R}$ being ϵ-far from satisfying $\mathscr{P}$.

A property $\mathscr{P}$ is said to be *testable* if for every $\epsilon > 0$ there exists an ϵ-test for $\mathscr{P}$ whose total number of queries is bounded only by a function of ϵ and is independent of the size of the input relational structure.

A property $\mathscr{P}$ of relational structures of type τ is said to be *hereditary* if it is closed under the removal of vertices, that is, it is closed under taking induced substructures. If a relational structure $\mathcal{R}$ does not satisfy some hereditary property $\mathscr{P}$, then any superstructure of $\mathcal{R}$ also fails to satisfy the property $\mathscr{P}$. We say $\mathcal{R}$ is a forbidden induced substructure for the property $\mathscr{P}$. Any hereditary property is definable by its family of forbidden induced substructures. Conversely, given any family of structures $\mathscr{F}$, the property of not containing any structure from $\mathscr{F}$ as an induced substructure is a hereditary property. Thus, a hereditary property $\mathscr{P}$ can be characterized by its family of forbidden induced substructures.

Given a hereditary property $\mathscr{P}$, let $\mathscr{F}$ be the family of forbidden substructures for $\mathscr{P}$. If a relational structure $\mathcal{R}$, on a large set X, is ϵ-far from satisfying $\mathscr{P}$, then $\mathcal{R}$ cannot satisfy the conditions of Theorem 3.4.3 for the given ϵ. Otherwise, the conclusion of Theorem 3.4.3 would imply $\mathcal{R}$ is ϵ-close to satisfying $\mathscr{P}$. So there exist $\delta > 0$, $n \in \mathbb{N}$ and $\mathcal{S} \in \mathscr{F}$ on an underlying set Y such that $|Y| \leq n$ and $t_{ind}(\mathcal{S}, \mathcal{R}) > \delta$. Then we can randomly sample N vertices, for a large enough N, independent of $|X|$, and we expect to find an induced copy of $\mathcal{S}$ with a high probability. This yields an ϵ-test for $\mathscr{P}$. Thus, using the Strong Removal Lemma we can easily show that hereditary properties are testable.

Chapter 4

Weighted Structures

4.1 Weighted Structures and the Correspondence Principle

A *weighted signature* τ is a first-order signature with finitely many function symbols and a function ar that assigns an arity to each function symbol.

$\mathcal{R}$ is a *weighted structure of type* τ on an underlying set X if each $\rho \in \tau$ is interpreted in $\mathcal{R}$ as a bounded weight function as follows :
Let $k = ar(\rho)$. Then $\rho^{\mathcal{R}} : E_k(X) \to [0,1]$ is a weight function. We may view $\rho^{\mathcal{R}}$ as a weight function on $X^{[k]}$ by assuming it takes the value 0 outside $E_k(X)$.

Recall that in the case of finite sets X, $\mu_{[k]}$ is the counting measure on $X^{[k]}$. In case X is the ultraproduct $\mathbf{X}$, $\mu_{[k]}$ is the Loeb measure on the ultraproduct space $\mathbf{X}^{[k]}$ defined in Chapter 2.

Let $\rho^{\mathcal{R}}(A) = \int\limits_A \rho^{\mathcal{R}}(x)\, d\mu_{[k]}(x)$, for any measurable $A \subseteq X^{[k]}$.

Example : A digraph $G = (V, E)$ can be viewed as a weighted structure of type $\{\rho_e, \rho_v\}$, where $ar(\rho_e) = 2$ and $ar(\rho_v) = 1$, if we interpret the function symbols as $\rho_e^G = \chi_E$ and $\rho_v^G = 1$ on V.

If G is a graph, then E is a symmetric set of ordered tuples.

Example : Let $G = (V, f_e, f_v)$ be a weighted graph with edge-weights and vertex-weights given by weight functions f_e, f_v respectively. It is a weighted structure of type $\{\rho_e, \rho_v\}$, where $\rho_e^G = f_e$ and $\rho_v^G = f_v$.

Consider a sequence of weighted structures $\mathcal{R}_i$ of type τ on finite sets X_i and let $|X_i| \to \infty$. For each $\rho \in \tau$ with $k = ar(\rho)$, the functions $\rho^{\mathcal{R}_i}$ are uniformly bounded in $[0, 1]$, so $[\rho^{\mathcal{R}_i}]$ is a bounded internal function on $\mathbf{X}^{[k]}$ while $std \circ [\rho^{\mathcal{R}_i}]$ is a real-valued function that takes values in $[0, 1]$. Let $\boldsymbol{\mathcal{R}}$ be the weighted structure of type τ on $\mathbf{X}$ that interprets each $\rho \in \tau$ as $\rho^{\boldsymbol{\mathcal{R}}} = std \circ [\rho^{\mathcal{R}_i}]$.

Given any measurable $\mathbf{A} \subseteq \mathbf{X}^{[k]}$ and $A_i \subseteq X_i^{[k]}$ for each i, such that, $\mu_{[k]}(\mathbf{A}\Delta[A_i]) = 0$, Lemma 2.1.3 tells us

$$\begin{aligned}
\rho^{\boldsymbol{\mathcal{R}}}(\mathbf{A}) &= \int_{\mathbf{A}} \rho^{\boldsymbol{\mathcal{R}}}(x)\, d\mu_{[k]}(x) = \int_{\mathbf{A}} std([\rho^{\mathcal{R}_i}(x)])\, d\mu_{[k]}(x) \\
&= std\left(\left[\int_{A_i} \rho^{\mathcal{R}_i}(x)\, d\mu_{i,[k]}(x)\right]\right) = std([\rho^{\boldsymbol{\mathcal{R}}_i}(A_i)]).
\end{aligned}$$

Fix a weighted type τ and let $k_{max} = \max\{ar(\rho) : \rho \in \tau\}$.

Theorem 4.1.1. *Let $\boldsymbol{\mathcal{R}}$ be a weighted structure of type τ on the ultraproduct $\mathbf{X}$. There exists a separable realization ϕ on $\mathbf{X}^{[k_{max}]}$ such that for each $\rho \in \tau$ with $k = ar(\rho)$,*

1. *$\rho^{\boldsymbol{\mathcal{R}}}$ is $\langle \ell_A : A \in r([k])\rangle$-measurable, and*

2. *there exists a Lebesgue-measurable function $w_\rho : [0,1]^{r([k])} \to [0,1]$ such that $\rho^{\boldsymbol{\mathcal{R}}} = w_\rho \circ \phi_k$ a.e. and*

3. *$\int_{\phi_k^{-1}(B)} \rho^{\boldsymbol{\mathcal{R}}}\, d\mu_{[k]} = \int_B w_\rho\, d\lambda$, that is, $\rho^{\boldsymbol{\mathcal{R}}}(\phi_k^{-1}(B)) = w_\rho(B)$ for all Lebesgue-measurable subsets B of $[0,1]^{r([k])}$.*

We say $W(\boldsymbol{\mathcal{R}}) = \{w_\rho : \rho \in \tau\}$, is the corresponding *weighted Euclidean structure of type τ.*

Proof. Given a binary sequence $s \in \mathbf{2}^n$, let $m_s = \sum_{i=0}^{n-1} s(i)2^i$ and let $I_s = [\frac{m_s}{2^n}, \frac{m_s+1}{2^n}) \subseteq [0,1]$. For each $k \in [k_{max}]$, let $\boldsymbol{\mathcal{A}}_k$ be the σ-algebra on $\mathbf{X}^{[k]}$ generated by $\{(\rho^{\boldsymbol{\mathcal{R}}})^{-1}(I_s) : s \in \mathbf{2}^{<\omega}, \rho \in \tau, ar(\rho) = k\}$. Using Theorem 2.4.3 for these σ-algebras $\boldsymbol{\mathcal{A}}_k$, we know that there exists a separable realization ϕ on $\mathbf{X}^{[k_{max}]}$. So each $(\rho^{\boldsymbol{\mathcal{R}}})^{-1}(I_s) \in \langle \ell_A : A \in r([k])\rangle$ where $k = ar(\rho)$. Therefore each $\rho^{\boldsymbol{\mathcal{R}}}$ is measurable with respect to $\langle \ell_A : A \in r([k])\rangle$.

Fix $\rho \in \tau$ and let $k = ar(R)$. For each $s \in \mathbf{2}^{<\omega}$, there exists Lebesgue-measurable $U_{\rho,s} \subseteq [0,1]^{r([k])}$ such that $\mu_{[k]}(\phi_k^{-1}(U_{\rho,s})\Delta(\rho^{\boldsymbol{\mathcal{R}}})^{-1}(I_s)) = 0$. We know that $\{I_s : s \in \mathbf{2}^n\}$ form a family of nested partitions of $[0,1]$. Since ϕ_k is a measurable isomorphism, we may assume $\{U_{\rho,s} : s \in \mathbf{2}^n\}$ form a family of nested partitions of $[0,1]^{r([k])}$ modulo a null set N.

Given $y \in [0,1]^{r([k])}$, let $w_\rho(y) = 0$ if $y \in N$, otherwise let $w_\rho(y) = z$ where $\{z\} = \bigcap_{\{s : y \in U_{\rho,s}\}} I_s$. Then $w_\rho : [0,1]^{r([k])} \to [0,1]$ is a Lebesgue-measurable function.

$N' = \bigcup_{s \in \mathbf{2}^{<\omega}} (\phi_k^{-1}(U_{\rho,s}) \Delta (\rho^{\boldsymbol{\mathcal{R}}})^{-1}(I_s))$ is a null set. For all $x \notin N'$,

$$\begin{aligned}
\{w_\rho \circ \phi_k(x)\} = \{w_\rho(\phi_k(x))\} &= \bigcap_{\{s:\phi_k(x)\in U_{\rho,s}\}} I_s = \bigcap_{\{s:x\in\phi_k^{-1}(U_{\rho,s})\}} I_s \\
&= \bigcap_{\{s:x\in(\rho^{\boldsymbol{\mathcal{R}}})^{-1}(I_s)\}} = \bigcap_{\{s:\rho^{\boldsymbol{\mathcal{R}}}(x)\in I_s\}} I_s \\
&= \{\rho^{\boldsymbol{\mathcal{R}}}(x)\}
\end{aligned}$$

So $w_\rho \circ \phi_k = \rho^{\boldsymbol{\mathcal{R}}}$ a.e. on $\mathbf{X}^{[k]}$. $W(\boldsymbol{\mathcal{R}}) = \{w_\rho : \rho \in \tau\}$ is the corresponding Euclidean structure.

Let A, B be any Lebesgue-measurable subsets of $[0,1]^{r([k])}$ and let $w = \chi_A$. Then $w \circ \phi_k = \chi_{\phi_k^{-1}(A)}$. Since ϕ_k is a measure-preserving measurable isomorphism, $\phi_k^{-1}(A)$ is measurable in $\sigma_{[k]}$ and $\mu_{[k]}(\phi_k^{-1}(A \cap B)) = \lambda(A \cap B)$.

So $\int_{\phi_k^{-1}(B)} w \circ \phi_k \, d\mu_{[k]} = \int_B w \, d\lambda$ for all Lebesgue-measurable subsets A of $[0,1]^{r([k])}$. We can now inductively prove the same for any Lebesgue-measurable function $w : [0,1]^{r([k])} \to [0,1]$, in particular, for each w_ρ.

Since $w_\rho \circ \phi_k = \rho^{\boldsymbol{\mathcal{R}}}$ a.e., we have now proved $\int_{\phi_k^{-1}(B)} \rho^{\boldsymbol{\mathcal{R}}} \, d\mu_{[k]} = \int_B w_\rho \, d\lambda$ for any B. $\square$

4.2 Limits for Weighted Structures

We now define homomorphisms between weighted structures that will help us define a notion of limits for weighted structures.

Let $\mathcal{R}, \mathcal{S}$ be weighted structures of type τ on sets X, Y respectively and let $\epsilon \geq 0$. A map $f : Y \to X$ is an *ϵ-homomorphism* from $\mathcal{S}$ to $\mathcal{R}$ if for each $\rho \in \tau$ with $k = ar(\rho)$ and for all $y_1, \ldots, y_k \in Y$,

$$\rho^{\mathcal{S}}(y_1, \ldots, y_k) - \epsilon \leq \rho^{\mathcal{R}}(f(y_1), \ldots, f(y_k)).$$

A *homomorphism* is a 0-homomorphism. We are mainly interested in 0-homomorphisms, but we need ϵ-homomorphisms for technical reasons in the course of finding limit objects.

If we view a graph G as a weighted digraph with $\rho_e = \chi_{E(G)}$ and $\rho_v = 1$, as described in the previous section, then a weighted homomorphism is the same as a standard graph homomorphism. Similarly, we can view a relational structure as a 0-1 weighted structure and the notion of weighted homomorphisms once again coincides with the notion of standard homomorphisms.

Let $T_\epsilon(\mathcal{S}, \mathcal{R})$ denote the *set of ϵ-homomorphisms* from $\mathcal{S}$ to $\mathcal{R}$. Let $t_\epsilon(\mathcal{S}, \mathcal{R}) = \frac{|T_\epsilon(\mathcal{S},\mathcal{R})|}{|X|^{|Y|}}$ denote the *ϵ-homomorphism density* of $\mathcal{S}$ in $\mathcal{R}$. Then $T(\mathcal{S}, \mathcal{R}) = T_0(\mathcal{S}, \mathcal{R})$ denotes the *set of homomorphisms* from $\mathcal{S}$ to $\mathcal{R}$ and $t(\mathcal{S}, \mathcal{R}) = t_0(\mathcal{S}, \mathcal{R})$ denotes the *homomorphism density* of $\mathcal{S}$ in $\mathcal{R}$.

Let us assume $Y = [n]$. Then for any $\epsilon \geq 0$, $T_\epsilon(\mathcal{S}, \mathcal{R}) \subseteq X^{[n]}$ for any weighted structure $\mathcal{R}$ on a set X and $t_\epsilon(\mathcal{S}, \mathcal{R}) = \mu_{[n]}(T_\epsilon(\mathcal{S}, \mathcal{R}))$.

Recall that for every $e \in E_k([n])$, $\pi_e : X^{[n]} \to X^e$ is the natural projection defined by $\pi_e(\mathbf{x}) = (x_{e(1)}, \ldots, x_{e(k)})$. Also $\theta_e : X^{[k]} \to X^e$ is the natural bijection induced by the ordered tuple e. Given $\mathbf{x} = (x_1, \ldots, x_k) \in X^{[k]}$ where each $x_i \in X^{\{i\}}$, θ_e maps $\mathbf{x}$ to a copy of itself in X^e by sending each x_i to a copy of itself in $X^{\{e(i)\}}$.

Also recall that given a tuple $\mathbf{x} \in X^{[n]}$ and $e \in E_k(X)$, $\theta_e^{-1}(\pi_e(\mathbf{x})) = (x_{e(1)}, \ldots, x_{e(k)}) \in$

$X^{[k]}$. Given $\epsilon > 0$, $\mathbf{x}$ represents an ϵ-homomorphism from $\mathcal{S}$ to $\mathcal{R}$, that is, $\mathbf{x} \in T(\mathcal{S},\mathcal{R})$ if for all $\rho \in \tau$ with $k = ar(\rho)$, for all $e \in E_k([n])$, $\rho^{\mathcal{R}}(x_{e(1)},\ldots,x_{e(k)}) = \rho^{\mathcal{R}}(\theta_e^{-1}(\pi_e(\mathbf{x}))) \geq \rho^{\mathcal{S}}(e) - \epsilon$, that is, $\rho^{\mathcal{R}}(\theta_e^{-1}(\pi_e(\mathbf{x}))) \in [\rho^{\mathcal{S}}(e) - \epsilon, 1]$. Therefore,

$$T_\epsilon(\mathcal{S},\mathcal{R}) = \bigcap_{\rho\in\tau} \bigcap_{e\in E_k([n])} \pi_e^{-1}(\theta_e((\rho^{\mathcal{R}})^{-1}([\rho^{\mathcal{S}}(e) - \epsilon, 1])))$$

Remark 4.2.1. If $0 \leq \epsilon_1 \leq \epsilon_2$, then $T_{\epsilon_1}(\mathcal{S},\mathcal{R}) \subseteq T_{\epsilon_2}(\mathcal{S},\mathcal{R})$ and $t_{\epsilon_1}(\mathcal{S},\mathcal{R}) \leq t_{\epsilon_2}(\mathcal{S},\mathcal{R})$. In fact for all $\epsilon \geq 0$, $T_\epsilon(\mathcal{S},\mathcal{R}) = \bigcap_{\delta\to 0^+} T_{\epsilon+\delta}(\mathcal{S},\mathcal{R})$ and $t_\epsilon(\mathcal{S},\mathcal{R}) = \inf_{\delta\to 0^+} t_{\epsilon+\delta}(\mathcal{S},\mathcal{R})$.

Consider the type $\tau = \{\rho\}$ such that $ar(\rho) = 1$. Let $\mathcal{R}_i$ be the weighted structure of type τ on $X_i = [i]$ such that $\rho^{\mathcal{R}_i} = 1 - \frac{1}{i}$. Since $\rho^{\mathcal{R}_i}$ is constant, for any $\epsilon \geq 0$ and any $\mathcal{S}$, $t_\epsilon(\mathcal{S},\mathcal{R}_i) = 1$ or 0.

Take ultraproducts to obtain $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$. Since the weight functions in the sequence are constant, so is the weight function on the ultraproduct. In this case, $\rho^{\boldsymbol{\mathcal{R}}} = 1$. Similarly, for all $\epsilon \geq 0$, $t_\epsilon(\mathcal{S},\boldsymbol{\mathcal{R}}) = 1$ or 0.

Let $\mathcal{S}$ be the weighted structure of type τ on $[1]$ such that $\rho^{\mathcal{S}}(1) = 1$. Then $t_\epsilon(\mathcal{S},\mathcal{R}_i) = 1$ if and only if $i \geq \frac{1}{\epsilon}$ for $\epsilon > 0$ while $t(\mathcal{S},\mathcal{R}_i) = 0$ for all i.

However, $t_\epsilon(\mathcal{S},\boldsymbol{\mathcal{R}}) = 1$ for all $\epsilon \geq 0$. In this example we see that $\lim_{i\to\infty} t_\epsilon(\mathcal{S},\mathcal{R}_i)$ exists for all $\epsilon \geq 0$, but $T(\mathcal{S},\boldsymbol{\mathcal{R}})$ does not differ from $[T(\mathcal{S},\mathcal{R}_i)]$ on a null set and $t(\mathcal{S},\boldsymbol{\mathcal{R}}) \neq \lim_{i\to\infty} t(\mathcal{S},\mathcal{R}_i)$. So we don't always have the kind of correspondence between homomorphism sets and densities in the sequence and in the ultraproduct that we had obtained for relational structures in Section 3.1.

Pointwise convergence of the homomorphism densities in a finite sequence of weighted structures does not appear to be a sufficient condition to help us find an analytic object as a limit. If we hope to find a limit object for all convergent sequences, we must define a stronger notion of convergence. We use a very natural strengthening that asks for uniform convergence of the homomorphism densities.

A sequence of weighted structures $\mathcal{R}_i$ of type τ on finite sets X_i is said to be *convergent* if for any weighted structure $\mathcal{S}$ of type τ on any finite set, for any $\delta > 0$, there exists $N = N(\delta) \in \mathbb{N}$, such that for all $\epsilon \geq 0$,

$$i, j \geq N \Rightarrow |t_\epsilon(\mathcal{S}, \mathcal{R}_i) - t_\epsilon(\mathcal{S}, \mathcal{R}_j)| < \delta.$$

For any $\mathcal{S}$, define $f_i^{\mathcal{S}} : [0, \infty) \to [0, 1]$ as $f_i^{\mathcal{S}}(\epsilon) = t_\epsilon(\mathcal{S}, \mathcal{R}_i)$. Then we can restate the above definition as follows :

A sequence of weighted structures $\mathcal{R}_i$ of type τ on finite sets X_i is said to be *convergent* if for any weighted structure $\mathcal{S}$ of type τ on any finite set, the sequence of functions $f_i^{\mathcal{S}}$ is uniformly Cauchy in ϵ, that is, for any $\delta > 0$, there exists $N = N(\delta) \in \mathbb{N}$, such that for all $\epsilon \geq 0$,

$$i, j \geq N \Rightarrow |f_i^{\mathcal{S}}(\epsilon) - f_j^{\mathcal{S}}(\epsilon)| < \delta.$$

Then the functions $f_i^{\mathcal{S}}$ form a uniformly convergent sequence. Furthermore, $f_i^{\mathcal{S}}$ is pointwise convergent for all $\epsilon \geq 0$, that is, for any $\epsilon \geq 0$ and any relational structure $\mathcal{S}$ on any finite set, $\lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$ exists. Equivalently, for any relational structure $\mathcal{S}$ on any finite set, $\lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i)$ exists.

The sequence $\mathcal{R}_i$ is said to be *increasingly convergent* if it is convergent and $|X_i| \to \infty$.

One of the key technical difficulties in proving the existence of limits in the case of weighted structures is that the homomorphism set in the ultraproduct $T(\mathcal{S}, \boldsymbol{\mathcal{R}})$ is not necessarily an internal set. Unlike the case of relational structures, $T(\mathcal{S}, \boldsymbol{\mathcal{R}})$ need not equal $[T(\mathcal{S}, \mathcal{R}_i)]$ and these sets may not differ by a null set, in general. However, given an increasingly convergent sequence $\mathcal{R}_i$, these two sets only differ by a null set . We use ϵ-homomorphisms to prove this by showing $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i)$ for any $\mathcal{S}$ on a finite set.

Theorem 4.2.2. *Let $\mathcal{R}_i$ be weighted structures of type τ on finite sets X_i of increasing size and take ultraproducts to obtain the weighted structure $\boldsymbol{\mathcal{R}}$ of type τ on $\mathbf{X}$. Let $n \in \mathbb{N}$ and let $\mathcal{S}$ be a weighted structure of type τ on $[n]$. Then for all $\epsilon \geq 0$ and $\delta > 0$, $T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) \subseteq \mathbf{X}^{[n]}$ is a $\sigma_{[n]}$-measurable set and $[T_\epsilon(\mathcal{S}, \mathcal{R}_i)] \subseteq T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) \subseteq [T_{\epsilon+\delta}(\mathcal{S}, \mathcal{R}_i)]$.*

If the sequence $\mathcal{R}_i$ is also convergent, that is, the given sequence is increasingly convergent, then $t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$.

Proof. Given a sequence of weighted structures $\mathcal{R}_i$ on finite sets X_i of increasing size, we can take ultraproducts and obtain the relational structure $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$. Given $\epsilon \geq 0$, we know that

$$T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \bigcap_{\rho\in\tau} \bigcap_{e\in E_k([n])} \pi_e^{-1}(\theta_e((\rho^{\boldsymbol{\mathcal{R}}})^{-1}([\rho^{\mathcal{S}}(e) - \epsilon, 1]))).$$

Since each $\rho^{\boldsymbol{\mathcal{R}}}$ is a measurable function, each θ_e is a measurable bijection and each π_e is a measurable projection, $T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}})$ is a $\sigma_{[n]}$-measurable set.

Let $\bar{\mathbf{x}} = [\bar{x}_i] \in [T_\epsilon(\mathcal{S}, \mathcal{R}_i)]$. Then for U-almost all i, for all $\rho \in \tau$ with $k = ar(\rho)$, for all $e \in E_k([n])$, $\rho^{\mathcal{S}}(e) - \epsilon \leq \rho^{\mathcal{R}_i}(\theta_e^{-1}(\pi_e(\bar{x}_i)))$. Therefore, for all $\rho \in \tau$ with $k = ar(\rho)$, for all $e \in E_k([n])$, $(\rho^{\mathcal{S}}(e) - \epsilon)^* \leq [\rho^{\mathcal{R}_i}](\theta_e^{-1}(\pi_e(\bar{\mathbf{x}})))$ and taking standard parts, $\rho^{\mathcal{S}}(e) - \epsilon \leq \rho^{\boldsymbol{\mathcal{R}}}(\theta_e^{-1}(\pi_e(\bar{\mathbf{x}})))$.

Thus, we have proved $[T_\epsilon(\mathcal{S}, \mathcal{R}_i)] \subseteq T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}})$ for all $\epsilon \geq 0$.

Given $\epsilon \geq 0$, let $f = [f_i] \in T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}})$. Then for all $\rho \in \tau$ with $k = ar(\rho)$ and for all $e \in E_k([n])$,

$$\rho^{\mathcal{S}}(e) - \epsilon \leq \rho^{\boldsymbol{\mathcal{R}}}(\theta_e^{-1}(\pi_e(\bar{\mathbf{x}})))$$

Recall that

$$\begin{aligned}
\rho^{\boldsymbol{\mathcal{R}}}(\theta_e^{-1}(\pi_e(\bar{\mathbf{x}}))) &= std([\rho^{\mathcal{R}_i}]([\theta_e^{-1}(\pi_e(\bar{x}_i))])) \\
&= std([\rho^{\mathcal{R}_i}(\theta_e^{-1}(\pi_e(\bar{x}_i)))])
\end{aligned}$$

Then for each $\delta > 0$, each $\rho \in \tau$, each $e \in E_k([n])$, $\{i : \rho^{\mathcal{S}}(e) - \epsilon - \delta \leq \rho^{\mathcal{R}_i}(\theta_e^{-1}(\pi_e(\bar{x}_i)))\} \in U$. There are only finitely many function symbols ρ in the type τ and $E_k([n])$ is finite for all $k \in [k_{max}]$. Therefore, for all $\delta > 0$, for U-almost every i, $f_i \in T_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i)$ and $f \in [T_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i)]$.

So $[T_\epsilon(\mathcal{S}, \mathcal{R}_i)] \subseteq T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) \subseteq [T_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i)]$, for all $\epsilon \geq 0$ and $\delta > 0$.

Taking measures of these subsets of $\mathbf{X}^{[n]}$, for all $\epsilon \geq 0$ and $\delta > 0$,

$$\lim_U t_\epsilon(\mathcal{S}, \mathcal{R}_i) \leq t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) \leq \lim_U t_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i).$$

If the given sequence $\mathcal{R}_i$ is convergent, then for all $\epsilon \geq 0$, $t_\epsilon(\mathcal{S}, \mathcal{R}_i)$ is convergent and

$\lim_U t_\epsilon(\mathcal{S}, \mathcal{R}_i) = \lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$. Therefore, for all $\epsilon \geq 0$ and $\delta > 0$,

$$\lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i) \leq t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) \leq \lim_{i\to\infty} t_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i).$$

Also $f_i^{\mathcal{S}}$ as defined above is uniformly convergent. Therefore,

$$\lim_{i\to\infty} \lim_{\delta\to 0^+} f_i^{\mathcal{S}}(\epsilon+\delta) = \lim_{\delta\to 0^+} \lim_{i\to\infty} f_i^{\mathcal{S}}(\epsilon+\delta),$$

that is,

$$\lim_{i\to\infty} \lim_{\delta\to 0^+} t_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i) = \lim_{\delta\to 0^+} \lim_{i\to\infty} t_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i).$$

For all i, $T_\epsilon(\mathcal{S}, \mathcal{R}_i) = \bigcap_{\delta>0} T_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i)$ and $t_\epsilon(\mathcal{S}, \mathcal{R}_i) = \lim_{\delta\to 0^+} t_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i)$. Therefore, $\lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i) = \lim_{\delta\to 0^+} \lim_{i\to\infty} t_{(\epsilon+\delta)}(\mathcal{S}, \mathcal{R}_i)$.

Thus we have proved for all $\epsilon \geq 0$, $t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$, if the given sequence is increasingly convergent.

So, $T(\mathcal{S}, \boldsymbol{\mathcal{R}})$ differs from $[T(\mathcal{S}, \mathcal{R}_i)]$ by a null set and $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t(\mathcal{S}, \mathcal{R}_i)$. □

Given $\epsilon \geq 0$, we define the *ϵ-homomorphism density of a weighted structure $\mathcal{S}$ on $[n]$ in a weighted Euclidean structure W* as :

$$t_\epsilon(\mathcal{S}, W) \quad = \quad \int_0^1 \cdots \int_0^1 \prod_{\rho\in\tau} \prod_{e\in E_k([n])} \chi_{w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon,1])}(\mathbf{x}_e) \prod_{A\in r([n],k_{max})} dx_A$$

The *homomorphism density of $\mathcal{S}$ in W* is the 0-homomorphism density of $\mathcal{S}$ in W.

Theorem 4.2.3. *Let $\boldsymbol{\mathcal{R}}$ be a weighted structure of type τ on the ultraproduct $\mathbf{X}$ and let*

$W(\boldsymbol{\mathcal{R}})$ be the corresponding weighted Euclidean structure. Then for every n, every weighted structure $\mathcal{S}$ of type τ on $[n]$ and every $\epsilon > 0$,

$$t_\epsilon(\mathcal{S}, W(\boldsymbol{\mathcal{R}})) = t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}).$$

Proof. Let $n \in \mathbb{N}$ and let $\mathcal{S}$ be a weighted structure of type τ on $[n]$.

$$\begin{aligned}
& t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) \\
= \; & \mu_{[n]}(T_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}})) \\
= \; & \mu_{[n]}\left(\bigcap_{\rho\in\tau}\bigcap_{e\in E_k([n])} \pi_e^{-1}(\theta_e((\rho^{\boldsymbol{\mathcal{R}}})^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1])))\right) \\
= \; & \mu_{[n]}\left(\bigcap_{\rho\in\tau}\bigcap_{e\in E_k([n])} \pi_e^{-1}(\theta_e(\phi_k^{-1}(w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1]))))\right) \\
= \; & \mu_{[n]}\left(\bigcap_{\rho\in\tau}\bigcap_{e\in E_k([n])} \psi_k^{-1}(L_e^{-1}(F_e(w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1]))))\right) \\
= \; & \mu_{[n]}\left(\bigcap_{\rho\in\tau}\psi_k^{-1}\left(\bigcap_{e\in E_k([n])} L_e^{-1}(F_e(w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1])))\right)\right) \\
= \; & \mu_{[n]}\left(\bigcap_{\rho\in\tau}\psi^{-1}\left(P_{r([k])}^{-1}\left(\bigcap_{e\in E_k([n])} L_e^{-1}(F_e(w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1])))\right)\right)\right) \\
= \; & \mu_{[n]}\left(\psi^{-1}\left(\bigcap_{\rho\in\tau}P_{r([k])}^{-1}\left(\bigcap_{e\in E_k([n])} L_e^{-1}(F_e(w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1])))\right)\right)\right) \\
= \; & Vol\left(\bigcap_{\rho\in\tau}P_{r([k])}^{-1}\left(\bigcap_{e\in E_k([n])} L_e^{-1}(F_e(w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon, 1])))\right)\right) \\
= \; & \int_0^1 \cdots \int_0^1 \prod_{\rho\in\tau}\prod_{e\in E_k([n])} \chi_{w_\rho^{-1}([\rho^{\mathcal{S}}(e)-\epsilon,1])}(\mathbf{x}_e) \prod_{A\in r([n],k_{max})} dx_A \\
= \; & t_\epsilon(\mathcal{S}, W(\boldsymbol{\mathcal{R}})). \qquad \square
\end{aligned}$$

We say that a weighted Euclidean structure $W = \{w_\rho : \rho \in \tau\}$ is a *limit* of weighted structures $\mathcal{R}_i$ on sets X_i if for any weighted structure $\mathcal{S}$ on any finite set and for any $\epsilon \geq 0$, $t_\epsilon(\mathcal{S}, W) = \lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$.

Corollary 4.2.4. *If $\mathcal{R}_i$ is an increasingly convergent sequence, then $W(\boldsymbol{\mathcal{R}})$ is a limit for the sequence $\{\mathcal{R}_i\}$.*

Proof. We know from the second part of Theorem 4.2.2 that for any weighted structure $\mathcal{S}$ on a finite set and any $\epsilon \geq 0$, $t_\epsilon(\mathcal{S}, \boldsymbol{\mathcal{R}}) = \lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$. Theorem 4.2.3 implies $t_\epsilon(\mathcal{S}, W(\boldsymbol{\mathcal{R}})) = \lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i)$ for all $\mathcal{S}$ and all $\epsilon \geq 0$. Then W is a limit for the increasingly convergent sequence $\mathcal{R}_i$. □

Let τ be a weighted type. Let $W = \{w_\rho : \rho \in \tau\}$ be a weighted Euclidean structure of type τ. Each w_ρ is a measurable function on $[0,1]^{r([\|p\|])}$. By a random sampling on W, we can generate random weighted structures $\mathcal{R}_m = \mathcal{R}(W, m)$ on $[m]$ as follows : Choose x_A uniformly at random in $[0,1]$, for $A \in r([m], k_{max})$. For each $\rho \in \tau$ with $k = ar(\rho)$ and for each $e \in E_k([m])$

$$\rho^{\mathcal{R}_m}(e) = w_\rho(\mathbf{x}_e).$$

We believe we can show, similar to the case of relational structures, a subsequence of this random sequence converges to W, with probability 1.

Remark 4.2.5. Given a weighted Euclidean structure W of type τ and a structure preserving map ϕ on $[0,1]^{r([k_{max}])}$, let $\phi^{-1}(W)$ denote the weighted Euclidean structure represented by $\{w_\rho \circ \phi_k : k = ar(\rho), \rho \in \tau\}$. Using a similar proof to that of Theorem 3.2.3, we can prove for all weighted structures $\mathcal{S}$ on finite sets, for all $\epsilon \geq 0$, $t_\epsilon(\mathcal{S}, \phi^{-1}(W)) = t_\epsilon(\mathcal{S}, W)$.

We believe we can use a proof similar to that of Theorem 3.2.5 to show that weighted Euclidean structures are unique as limits up to structure preserving maps.

4.3 Regularity for Weighted Structures

Let $f : X^{[k]} \to [0,1]$ and $g : X^{[k]} \to [0,1]$ be measurable functions. We say f and g are ϵ-*close* if

$$\mu_{[k]}(\{\mathbf{x} \in X^{[k]} : |f(\mathbf{x}) - g(\mathbf{x})| > \epsilon\}) \leq \epsilon.$$

Let $\mathcal{R}$ and $\mathcal{R}'$ be two weighted structures of type τ on a set X. They are said to be ϵ-*close* to each other if for all $\rho \in \tau$, $\rho^{\mathcal{R}}$ and $\rho^{\mathcal{R}'}$ are ϵ-close.

$\mathcal{R}$ and $\mathcal{R}'$ are said to be 0-*close* if for all $\rho \in \tau$, $\rho^{\mathcal{R}} = \rho^{\mathcal{R}'}$ a.e.

Theorem 4.3.1. *Let $\boldsymbol{\mathcal{R}}$ be a weighted structure of type τ on an ultraproduct $\mathbf{X}$ of an increasing sequence of finite sets X_i. For every $\epsilon > 0$, there exists $\ell \in \mathbb{N}$ and a 0-equitable k_{max}-level ℓ-hyperpartition $\mathscr{H}$ such that*

1. *Each $\mathbf{P^r_j} \in \sigma_{[r]}$ and it is independent of $\sigma^*_{[r]}$.*
2. *There exists a weighted structure $\widetilde{\boldsymbol{\mathcal{R}}}$ of type τ on $\mathbf{X}$ that is ϵ-close to $\boldsymbol{\mathcal{R}}$ and for each $\rho \in \tau$ with $k = ar(\rho)$, $\rho^{\widetilde{\boldsymbol{\mathcal{R}}}}$ takes constant values on each $\mathscr{H}$-k-cell.*

Proof. We know there exists a separable realization ϕ on $\mathbf{X}^{k_{max}}$ and a weighted Euclidean structure $W(\boldsymbol{\mathcal{R}}) = \{w_\rho : \rho \in \tau\}$ such that for every $\rho \in \tau$ with $k = ar(\rho)$, $\rho^{\boldsymbol{\mathcal{R}}} = w_\rho \circ \phi_k$, as proved by Theorem 4.1.1.

Fix $\epsilon > 0$. Choose a natural number K such that $\frac{1}{K} < \epsilon$. Let intervals $I_1, \ldots, I_K$ form a partition of $[0,1]$ such that $|I_j| = \frac{1}{K} < \epsilon$ for each $j \in [K]$. Choose $c_j \in I_j$ for each $j \in [K]$.

Let $A_{\rho,j} = w_\rho^{-1}(I_j)$, for each $\rho \in \tau$ and $j \in [K]$. Since $w_\rho : [0,1]^{r([k])} \to [0,1]$ is a measurable function for each $\rho \in \tau$ with $k = ar(\rho)$, each $A_{\rho,j}$ is a measurable subset of $[0,1]^{r([k])}$. There exists a natural number ℓ large enough such that for each $\rho \in \tau$ and each $j \in [K]$, there is a union $B_{\rho,j}$ of ℓ-boxes in $[0,1]^{r([k])}$ such that $Vol(A_{\rho,j} \Delta B_{\rho,j}) < \frac{\epsilon}{K}$. Note that for each ρ, $\{A_{\rho,j} : j \in [K]\}$ is a partition of $[0,1]^{r([k])}$. We may assume ℓ is large enough to ensure that each ℓ-box in $[0,1]^{r([k])}$ appears in at most one $B_{\rho,j}$.

Now consider the hyperpartition $\mathscr{H}$ formed by $\mathbf{P}^{\mathbf{r}}_{\mathbf{i}} = \phi_r^{-1}([\frac{i-1}{\ell}, \frac{i}{\ell}))$, for each $r \in [k_{max}]$ and $i \in [\ell]$. Clearly each $\mathbf{P}^{\mathbf{r}}_{\mathbf{i}}$ is in the separable algebra $l_{[r]} \leq \sigma_{[r]}$ and is, therefore, independent of $\sigma^*_{[r]}$. We may assume $\mathbf{P}^{\mathbf{r}}_{\mathbf{i}} \subseteq E_r(\mathbf{X})$. So $\mathbf{P}^{\mathbf{r}}_{\mathbf{i}}$ is an r-uniform hypergraph on $\mathbf{X}$. Also $\mu_{[r]}(\mathbf{P}^{\mathbf{r}}_{\mathbf{i}}) = \frac{1}{\ell}$ for each $i \in [\ell]$, so the partition is 0-equitable.

Since each $B_{\rho,j}$ is a union of ℓ-boxes in $[0,1]^{r([k])}$, $\phi_k^{-1}(B_{\rho,j})$ is a union of $\mathscr{H}$-k cells.

Fix $\rho \in \tau$. Let $k = ar(\rho)$. Define $\tilde{\rho} : \mathbf{X}^{[k]} \to [0,1]$ as follows :

For each $\mathbf{x} \in \phi_k^{-1}(B_{\rho,j})$, $\tilde{\rho}(\mathbf{x}) = c_j$. For all $\mathbf{x} \in \mathbf{X}^{[k]} \setminus \bigcup_{j \in [K]} \phi_k^{-1}(B_{\rho,j})$, $\tilde{\rho}(\mathbf{x}) = 0$.

Clearly each $\tilde{\rho}$ is a measurable function that takes a constant value on every $\mathscr{H}$-k-cell.

For every $\rho \in \tau$ with $k = ar(\rho)$, every $j \in [K]$,

$$\begin{aligned}\mathbf{x} \in \phi_k^{-1}(A_{\rho,j}) \cap \phi_k^{-1}(B_{\rho,j}) &\Rightarrow \phi_k(\mathbf{x}) \in A_{\rho,j} \cap B_{\rho,j} \\ &\Rightarrow \tilde{\rho}(\mathbf{x}) = c_j \wedge w_\rho(\phi_k(\mathbf{x})) \in I_j \\ &\Rightarrow |w_\rho \circ \phi_k(\mathbf{x}) - c_j| < \epsilon.\end{aligned}$$

Recall that for each $\rho \in \tau$ with $k = ar(\rho)$, $\mu_{[k]}\left(\bigcup_{j\in[K]}(A_{\rho,j}\Delta B_{\rho,j})\right) < \epsilon$ and $\rho^{\boldsymbol{\mathcal{R}}} = w_\rho \circ \phi_k$ a.e. Therefore, $\rho^{\boldsymbol{\mathcal{R}}}$ and $\tilde{\rho}$ are ϵ-close.

Let $\widetilde{\boldsymbol{\mathcal{R}}}$ be the weighted structure of type τ represented by $\{\tilde{\rho} : \rho \in \tau\}$. Then $\widetilde{\boldsymbol{\mathcal{R}}}$ is ϵ-close to $\boldsymbol{\mathcal{R}}$. □

Theorem 4.3.2. *Given $\epsilon > 0$ and $m \in \mathbb{N}$, there exist $M, N \in \mathbb{N}$ such that for any weighted structure $\mathcal{R}$ on a finite set X with $|X| \geq N$, there exists an ϵ-equitable k_{max}-level ℓ-hyperpartition $\mathscr{H}$ on X for some ℓ such that $m \leq \ell \leq M$ and*

1. *Each P_j^r is ϵ-regular.*
2. *There exists a weighted structure $\widetilde{\mathcal{R}}$ of type τ on X that is ϵ-close to $\mathcal{R}$ and for each $\rho \in \tau$ with $k = ar(\rho)$, $\rho^{\widetilde{\mathcal{R}}}$ takes constant values on each $\mathscr{H}$-k-cell.*

Proof. Suppose, for contradiction, there exist $\epsilon > 0$, $m \in \mathbb{N}$, weighted structures $\mathcal{R}_i$ of type τ on X_i and $|X_i| \to \infty$ such that no ϵ-equitable ℓ-hyperpartition for $m \leq \ell \leq i$ satisfying above conditions exists for any $\mathcal{R}_i$.

We can take ultraproducts and obtain the weighted structure $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$ as before. There exists a 0-equitable ℓ-hyperpartition $\boldsymbol{\mathscr{H}}$ on $\mathbf{X}$ satisfying the conditions in Theorem 4.3.1. We may assume $\ell \geq m$. For each $r \in [k_{max}]$, each $j \in [K]$, let $\widetilde{\mathbf{P}}_{\mathbf{j}}^{\mathbf{r}} = [P_{j,i}^r]$ be an internal set that differs from $\mathbf{P^r_j}$ by a null set. We may assume each $P^r_{j,i}$ is an r-uniform hypergraph on $X_i^{[r]}$. The hypergraphs $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}}$ form a 0-equitable ℓ-hyperpartition $\widetilde{\mathscr{H}}$ that also satisfies the conditions in Theorem 4.3.1. So there is a weighted structure $\widetilde{\mathcal{R}}$ on $\mathbf{X}$ such that for each $\rho \in \tau$, $\rho^{\widetilde{\mathcal{R}}}$ is ϵ-close to $\rho^{\mathcal{R}}$ and takes constant values on each $\widetilde{\mathscr{H}}$-k-cell.

The hypergraphs $P^r_{j,i}$ form an ℓ-hyperpartition $\mathscr{H}_i$ on X_i for U-almost all i. Since $\widetilde{\mathscr{H}}$ is a 0-equitable hyperpartition, for every $j, j' \in [\ell]$ and $r \in [k_{max}]$, $\lim_U |\mu_{i,[r]}(P^r_{j,i}) - \mu_{i,[r]}(P^r_{j',i})| = 0$. Then for U-almost all i, $\mathscr{H}_i$ is an ϵ-equitable ℓ-hyperpartition on X_i.

We claim that for U-almost every i, $P^r_{j,i}$ is ϵ-regular. Suppose not, that is, there exist $r \in [k_{max}]$ and $j \in [\ell]$ such that for U-almost every i, there exists a cylindric intersection set $L_i \subseteq E_r(X_i)$ with $\mu_{i,[r]}(L_i) \geq \epsilon$ and $|\mu_{i,[r]}(P^r_{j,i} \cap L_i) - \mu_{i,[r]}(P^r_{j,i})\mu_{i,[r]}(L_i)| \geq \epsilon\, \mu_{i,[r]}(L_i)$.

Consider the cylindric intersection set $\mathbf{L} = [L_i]$. $\mathbf{L} \in \sigma^*_{[r]}$ and it is independent of $\mathbf{P^r_j}$. Therefore, $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}}$ and $\mathbf{L}$ are also independent. However $|\mu_{[r]}(\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}} \cap \mathbf{L}) - \mu_{[r]}(\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}})\mu_{[r]}(\mathbf{L})| = \lim_U |\mu_{i,[r]}(P^r_{j,i} \cap L_i) - \mu_{i,[r]}(P^r_{j,i})\mu_{i,[r]}(L_i)| \geq \epsilon\, \mu_{[r]}(\mathbf{L})$ and $\mu_{[r]}(\mathbf{L}) \geq \epsilon$, which contradicts the fact that $\widetilde{\mathbf{P}}^{\mathbf{r}}_{\mathbf{j}}$ and $\mathbf{L}$ are independent. Thus, $P^r_{j,i}$ is ϵ-regular, for each $r \in [k_{max}]$, $j \in [\ell]$ and U-almost every i.

Let $r \in [k_{max}]$ and let $\mathbf{C}$ be any $\widetilde{\mathscr{H}}$-r-cell. Note that $\mathbf{C}$ is an internal set. Then $\mathbf{C} = [C_i]$ such that for U-almost all i, C_i is an $\mathscr{H}_i$-r-cell. There are only finitely many r-cells for any $r \in [k_{max}]$. Then for U-almost all i, for all $r \in [k_{max}]$, for all $\widetilde{\mathscr{H}}$-r-cells $\mathbf{C}$, there is an $\mathscr{H}_i$-r-cell C_i such that C_i is an $\mathscr{H}_i$-r-cell. For all such i, define $\widetilde{\mathcal{R}}_i$ as follows :
For each $\rho \in \tau$ with $k = ar(\rho)$, let $\rho^{\widetilde{\mathcal{R}}_i}$ be the weight function on $X_i^{[k]}$ such that $\rho^{\widetilde{\mathcal{R}}_i}[C_i] = \rho^{\widetilde{\mathcal{R}}}[C]$. So $\rho^{\widetilde{\mathcal{R}}_i}$ takes constant values on each $\mathscr{H}_i$-k-cell and $\rho^{\widetilde{\mathcal{R}}} = [\rho^{\widetilde{\mathcal{R}}_i}]$.

Suppose for all i, there exists $\rho \in \tau$ with $k = ar(\rho)$ such that

$$\mu_{i,[k]}(\{x \in X_i^{[k]} : |\rho^{\mathcal{R}_i}(x) - \rho^{\widetilde{\mathcal{R}}_i}(x)| > \epsilon\}) > \epsilon.$$

Since there are only finitely many function symbols in τ, there exists $\rho \in \tau$ such that for U-almost all i, $\mu_{i,[k]}(\{x \in X_i^{[k]} : |\rho^{\mathcal{R}_i}(x) - \rho^{\widetilde{\mathcal{R}}_i}(x)| > \epsilon\}) > \epsilon$.

Then $\mu_{[k]}(\{\mathbf{x} \in X^{[k]} : |\rho^{\boldsymbol{\mathcal{R}}}(\mathbf{x}) - \rho^{\widetilde{\boldsymbol{\mathcal{R}}}}(\mathbf{x})| \geq \epsilon\}) \geq \epsilon$, which contradicts the fact that $\boldsymbol{\mathcal{R}}$ and $\widetilde{\boldsymbol{\mathcal{R}}}$ are ϵ-close. $\square$

4.4 Removal Lemma for Weighted Structures

Theorem 4.4.1. *Given $n \in \mathbb{N}$, let $\mathcal{S}$ be a weighted structure of type τ on $[n]$. Let $\boldsymbol{\mathcal{R}}$ be a weighted structure of type τ on the ultraproduct $\mathbf{X}$ of finite sets X_i that are increasing in size. Then there exists a weighted structure $\widetilde{\boldsymbol{\mathcal{R}}}$ of type τ on $\mathbf{X}$ that is 0-close to $\boldsymbol{\mathcal{R}}$ and $T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = \emptyset$ or $t(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) > 0$.*

Proof. By Theorem 4.1.1, we know there exists a separable realization ϕ on $\mathbf{X}^{[k_{max}]}$ and a weighted Euclidean structure $W(\boldsymbol{\mathcal{R}}) = \{w_\rho : \rho \in \tau\}$ such that for every $\rho \in \tau$ with $k = ar(\rho)$, $\rho^{\boldsymbol{\mathcal{R}}} = w_\rho \circ \phi_k$ a.e.

Fix $\rho \in \tau$ and let $k = ar(\rho)$. Arrange the weights $\{\rho^{\mathcal{S}}(e) : e \in E_k([n])\}$ in increasing order, without repetition, into a finite strictly increasing sequence $c_1 < c_2 < \ldots < c_M$. For each $i \in [M]$, $A_i = w_\rho^{-1}([c_i, 1])$ is a Lebesgue-measurable subset of $[0,1]^{r([k])}$. Then $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_M$. For each $i \in [M]$, let D_i be the set of density points of A_i. By Lebesgue's density theorem, D_i is Lebesgue-measurable and $Vol(A_i \Delta D_i) = 0$ for each $i \in [M]$.

Let $S = A_1^c \cup (D_1 \cap A_2^c) \cup \ldots \cup (D_{M-1} \cap A_M^c) \cup D_M$.

Since $[0,1]^{r([k])} = A_1^c \cup (A_1 \cap A_2^c) \cup \ldots \cup (A_{M-1} \cap A_M^c) \cup A_M$, clearly $Vol(S) = 1$. Let $\widetilde{w}_\rho = w \cdot \chi_S$. So $w_\rho = \widetilde{w}_\rho$ a.e.

For each $i \in [M]$, $x \in \widetilde{w}_\rho^{-1}([c_i, 1])$ if and only if $x \in A_i \cap S = D_i \cap S$, that is, $\widetilde{w}_\rho^{-1}([c_i, 1]) = D_i \cap S$. Since $D_i \cap S \subseteq D_i$, each point of $D_i \cap S$ is a density point of D_i. Also $Vol(D_i \setminus (D_i \cap S)) = 0$. Given any $x \in D_i \cap S$ and $\epsilon > 0$ let $B_\epsilon(x)$ denote the ϵ-ball centered at x in $[0,1]^{r([k])}$. Then for all $\epsilon > 0$, $Vol((D_i \cap B_\epsilon(x))\Delta((D_i \cap S) \cap B_\epsilon(x))) = 0$. So

$$\lim_{\epsilon \to 0} \frac{Vol((D_i \cap S) \cap B_\epsilon(x))}{Vol(B_\epsilon(x)} = \lim_{\epsilon \to 0} \frac{Vol(D_i \cap B_\epsilon(x))}{Vol(B_\epsilon(x))} = 1$$

Therefore, every point of $D_i \cap S$ is a density point of the set.

Let $\widetilde{\boldsymbol{\mathcal{R}}}$ be the weighted structure of type τ on $\mathbf{X}$ represented by the system of weight functions $\{\rho^{\widetilde{\boldsymbol{\mathcal{R}}}} = \widetilde{w}_\rho \circ \phi_k : k = ar(\rho), \rho \in \tau\}$. Then $\boldsymbol{\mathcal{R}}$ and $\widetilde{\boldsymbol{\mathcal{R}}}$ are 0-close.

$$T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = \bigcap_{\rho \in \tau} \psi_k^{-1} \left(\bigcap_{e \in E_k([n])} L_e^{-1}(F_e(\widetilde{w}_\rho^{-1}([\rho^{\mathcal{S}}(e), 1]))) \right).$$

We have just shown that every $\widetilde{w}_\rho^{-1}([\rho^{\mathcal{S}}(e), 1])$ is the set of its density points, therefore so is $T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}})$. If $t(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = \mu_{[n]}(T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}})) = 0$, then $T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = \emptyset$. □

Theorem 4.4.2. *Given $n \in \mathbb{N}$, a weighted structure $\mathcal{S}$ of type τ on $[n]$ and given $\epsilon > 0$, there exist $\delta > 0$ and $N \in \mathbb{N}$ such that for any weighted structure $\mathcal{R}$ of type τ on a finite set X satisfying $|X| \geq N$ and $t_\epsilon(\mathcal{S}, \mathcal{R}) < \delta$, there exists a weighted structure $\widetilde{\mathcal{R}}$ of type τ on X such that $T_0(\mathcal{S}, \widetilde{\mathcal{R}}) = \emptyset$ and $\widetilde{\mathcal{R}}$ is ϵ-close to $\mathcal{R}$.*

Proof. Suppose, for contradiction, there exists $n \in \mathbb{N}$, a weighted structure $\mathcal{S}$ on $[n]$ and $\epsilon > 0$ such that there is a sequence of weighted structures $\mathcal{R}_i$ on finite sets X_i such that $|X_i| \to \infty$ and $t_\epsilon(\mathcal{S}, \mathcal{R}_i) < \frac{1}{i}$ but no such $\mathcal{R}'_i$ exists. So $\lim_{i\to\infty} t_\epsilon(\mathcal{S}, \mathcal{R}_i) = 0$.

We can now take ultraproducts and obtain the weighted structure $\boldsymbol{\mathcal{R}}$ on $\mathbf{X}$. By Theorem 4.2.2, $T(\mathcal{S}, \boldsymbol{\mathcal{R}}) \subseteq [T_\epsilon(\mathcal{S}, \mathcal{R}_i)]$. Then $t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = 0$.

By Theorem 4.4.1, there exists a weighted structure $\widetilde{\boldsymbol{\mathcal{R}}}$ on $\mathbf{X}$ that is 0-close to $\boldsymbol{\mathcal{R}}$ and $T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = \emptyset$ or $t(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) > 0$. Since $\widetilde{\boldsymbol{\mathcal{R}}}$ is 0-close to $\boldsymbol{\mathcal{R}}$, $t(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = t(\mathcal{S}, \boldsymbol{\mathcal{R}})$. Then $\mu_{[n]}(T(\mathcal{S}, \boldsymbol{\mathcal{R}})) = t(\mathcal{S}, \boldsymbol{\mathcal{R}}) = 0$ and $(T(\mathcal{S}, \boldsymbol{\mathcal{R}})) = \emptyset$.

For each $\rho \in \tau$, there exists an internal function $[\tilde{\rho}_i]$ such that $\rho^{\widetilde{\boldsymbol{\mathcal{R}}}} = std([\tilde{\rho}_i])$ a.e. We may assume that for U-almost all i, $\tilde{\rho}_i : X_i^{[k]} \to [0,1]$ and $\tilde{\rho}_i$ takes the value 0 on $X_i^{[k]} \setminus E_k(X_i)$. Let $\widetilde{\mathcal{R}}_i$ be the weighted structure of type τ represented by $\{\rho^{\widetilde{\mathcal{R}}_i} = \tilde{\rho}_i : \rho \in \tau\}$. Since $[T(\mathcal{S}, \widetilde{\mathcal{R}}_i)] \subseteq T(\mathcal{S}, \widetilde{\boldsymbol{\mathcal{R}}}) = \emptyset$, we know that for U-almost all i, $T(\mathcal{S}, \widetilde{\mathcal{R}}_i) = \emptyset$.

To contradict our original assumption and prove the theorem, it is enough to prove that for U-almost all i, $\widetilde{\mathcal{R}}_i$ is ϵ-close to $\mathcal{R}_i$. Suppose not, that is, for U-almost all i, there exists $\rho \in \tau$ such that $\rho^{\widetilde{\mathcal{R}}_i}$ is not ϵ-close to $\rho^{\widetilde{\mathcal{R}}_i}$. Therefore there exists $\rho \in \tau$ with $k = ar(\rho)$ such that for U-almost all i,

$$\mu_{i,[k]}(\{x \in X_i^{[k]} : |\rho^{\mathcal{R}_i}(x) - \rho^{\widetilde{\mathcal{R}}_i}(x)| \geq \epsilon\}) \geq \epsilon.$$

Then $\mu_{[k]}(\{\mathbf{x} \in X^{[k]} : |\rho^{\boldsymbol{\mathcal{R}}}(\mathbf{x}) - \rho^{\widetilde{\boldsymbol{\mathcal{R}}}}(\mathbf{x})| \geq \epsilon\}) \geq \epsilon$, which contradicts the fact that $\boldsymbol{\mathcal{R}}$ and $\widetilde{\boldsymbol{\mathcal{R}}}$ are 0-close. $\square$

4.5 Weighted Version of Regularity for Digraphs

In this section, we look for regular partitions where the condition for regularity depends on a notion of weighted edge density, instead of the usual edge density. We shall deal with the case of weighted digraphs where there is a natural notion of weighted edge density. A weighted digraph is such that there are weights on the directed edges as well as on the vertices. As discussed before, a weighted digraph G is a weighted structure of type $\tau = \{\rho_e, \rho_v\}$ where $ar(\rho_e) = 2$ and $ar(\rho_v) = 1$.

Let $\mathscr{H}$ be a 2-level ℓ-hyperpartition on X and let G be a weighted digraph on X with strictly positive vertex weights. A weighted digraph G is *weighted ϵ-regular* with respect to $\mathscr{H}$ if for any $i, j \in [\ell]$, for all $A_i \subseteq P_i^1$, $A_j \subseteq P_j^1$ such that $\rho_v^G(A_i) \geq \epsilon\rho_v^G(X)$ and $\rho_v^G(A_j) \geq \epsilon\rho_v^G(X)$, we have

$$\left| \frac{\rho_e^G(A_i \times A_j)}{\rho_v^G(A_i)\rho_v^G(A_j)} - \frac{\rho_e^G(P_i^1 \times P_j^1)}{\rho_v^G(P_i^1)\rho_v^G(P_j^1)} \right| < \epsilon.$$

Theorem 4.5.1. *Let $\mathbf{G}$ be a weighted digraph on the ultraproduct $\mathbf{X}$ of a sequence of finite sets X_n such that $|X_n| \to \infty$. Then for every $\epsilon > 0$, there exists $\ell \in \mathbb{N}$ and a 0-equitable 2-level ℓ-hyperpartition $\boldsymbol{\mathscr{H}}$ such that*

1. *Each $\mathbf{P_j^2} \in \sigma_{[2]}$ and it is independent of $\sigma_{[2]}^*$, that is, each edge cell is independent of measurable rectangles.*

2. *There exists a weighted digraph $\widetilde{\mathbf{G}}$ ϵ-close to $\mathbf{G}$ such that $\rho_v^{\widetilde{\mathbf{G}}} > 0$ and for all $A_i \subseteq P_i^1$ and $A_j \subseteq P_j^1$,*
$$\frac{\rho_e^{\widetilde{\mathbf{G}}}(A_i \times A_j)}{\rho_v^{\widetilde{\mathbf{G}}}(A_i)\rho_v^{\widetilde{\mathbf{G}}}(A_j)} = \frac{\rho_e^{\widetilde{\mathbf{G}}}(P_i^1 \times P_j^1)}{\rho_v^{\widetilde{\mathbf{G}}}(P_i^1)\rho_v^{\widetilde{\mathbf{G}}}(P_j^1)},$$
that is, $\widetilde{\mathbf{G}}$ is weighted δ-regular with respect to $\boldsymbol{\mathscr{H}}$, for all $\delta > 0$.

Proof. We can repeat the proof of Theorem 4.3.1 for the type $\tau = \{\rho_e, \rho_v\}$ and the weighted digraph $\mathbf{G}$. Then there exists $\ell \in \mathbb{N}$, a 0-equitable 2-level ℓ-hyperpartition $\mathscr{H}$ and a weighted digraph $\widetilde{\mathbf{G}}$ on $\mathbf{X}$ such that

- Each $\mathbf{P}^{\mathbf{2}}_{\mathbf{j}} \in \sigma_{[2]}$ and it is independent of $\sigma^*_{[2]}$.
- $\rho_v^{\widetilde{\mathbf{G}}}$ takes constant values on each $\mathscr{H}$-1-cell, that is, each $\mathbf{P}^{\mathbf{1}}_{\mathbf{i}}$ for $i \in [\ell]$.
- $\rho_e^{\widetilde{\mathbf{G}}}$ takes constant values on each $\mathscr{H}$-2-cell, which is of the form $(\mathbf{P}^{\mathbf{1}}_{\mathbf{i}} \times \mathbf{P}^{\mathbf{1}}_{\mathbf{j}}) \cap \mathbf{P}^{\mathbf{2}}_{\mathbf{k}}$ for $i, j, k \in [\ell]$.
- $\widetilde{\mathbf{G}}$ is ϵ-close to $\mathbf{G}$.

In the course of the proof of Theorem 4.5.1, we choose the constant value that $\rho_v^{\widetilde{\mathbf{G}}}$ takes on each vertex cell from an interval in $[0, 1]$. So we can further ensure that $\rho_v^{\widetilde{\mathbf{G}}} > 0$.

Given $i, j, k \in [\ell]$, let $\rho_v^{\widetilde{\mathbf{G}}} = c_i > 0$ on $\mathbf{P}^{\mathbf{1}}_{\mathbf{i}}$, $\rho_v^{\widetilde{\mathbf{G}}} = c_j > 0$ on $\mathbf{P}^{\mathbf{1}}_{\mathbf{j}}$ and $\rho_e^{\widetilde{\mathbf{G}}} = d_{i,j,k}$ on $(\mathbf{P}^{\mathbf{1}}_{\mathbf{i}} \times \mathbf{P}^{\mathbf{1}}_{\mathbf{j}}) \cap \mathbf{P}^{\mathbf{2}}_{\mathbf{k}}$. Let $\mathbf{A}_i \subseteq \mathbf{P}^{\mathbf{1}}_{\mathbf{i}}$, $\mathbf{A}_j \subseteq \mathbf{P}^{\mathbf{1}}_{\mathbf{j}}$ be any measurable sets.

Then $\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_i) = c_i \mu_{[1]}(\mathbf{A}_i)$ and $\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_j) = c_j \mu_{[1]}(\mathbf{A}_j)$.

Also $\rho_e^{\widetilde{\mathbf{G}}}((\mathbf{A}_i \times \mathbf{A}_j) \cap \mathbf{P}^{\mathbf{2}}_{\mathbf{k}}) = d_{i,j,k}\, \mu_{[2]}((\mathbf{A}_i \times \mathbf{A}_j) \cap \mathbf{P}^{\mathbf{2}}_{\mathbf{k}})$. Since $\mathbf{P}^{\mathbf{2}}_{\mathbf{k}}$ is independent of rectangles, $\rho_e^{\widetilde{\mathbf{G}}}((\mathbf{A}_i \times \mathbf{A}_j) \cap \mathbf{P}^{\mathbf{2}}_{\mathbf{k}}) = d_{i,j,k}\, \mu_{[1]}(\mathbf{A}_i) \mu_{[1]}(\mathbf{A}_j) \mu_{[2]}(\mathbf{P}^{\mathbf{2}}_{\mathbf{k}})$.

Since the ℓ-hyperpartition $\mathscr{H}$ is 0-equitable, $\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_i) = c_i \mu_{[1]}(\mathbf{A}_i)$ and $\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_j) = c_j \mu_{[1]}(\mathbf{A}_j)$, we have $\frac{\rho_e^{\widetilde{\mathbf{G}}}((\mathbf{A}_i \times \mathbf{A}_j) \cap \mathbf{P}^{\mathbf{2}}_{\mathbf{k}})}{\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_i) \rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_j)} = \frac{d_{i,j,k}}{c_i c_j \ell}$. Since the edge cells form a partition of $\mathbf{X}^{[2]}$,

$$\frac{\rho_e^{\widetilde{\mathbf{G}}}(\mathbf{A}_i \times \mathbf{A}_j)}{\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_i) \rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_j)} = \frac{\sum_{k \in [\ell]} d_{i,j,k}}{c_i c_j \ell}.$$

This proves our claim. □

Let us consider the special case of digraphs represented as weighted structures, then $\rho_e^G = \chi_{_{EG}}$ and $\rho_v^G = 1$ for each graph G. Theorem 4.5.1 extends Theorem 3.3.1 for the relational type $\tau = \{E\}$.

In Section 3.2 we applied Theorem 3.3.1 to a digraph $\mathbf{G}$ on $\mathbf{X}$ and found a 0-equitable 2-level partition on $\mathbf{X}$ and $C \subseteq [\ell]^3$ such that $\mu_{[2]}\left(E^{\mathbf{G}} \Delta \bigcup_{(i,j,k)\in C}(\mathbf{V}_i \times \mathbf{V}_j) \cap \mathbf{E}_k\right) < \epsilon$. We would obtain the same partition using Theorem 4.5.1. Also for all $i \in [\ell]$, $c_i = 0$ and $d_{i,j,k} = 1$ for $(i,j,k) \in C$ and $d_{i,j,k} = 0$ for $(i,j,k) \notin C$. Thus, as a special case, we obtain the familiar regularity condition for graphs involving edge densities.

Theorem 4.5.2. *Given $\epsilon > 0$ and $m \in \mathbb{N}$, there exist $M, N \in \mathbb{N}$ such that for any weighted digraph G on a finite set X with $|X| \geq N$, there exists an ϵ-equitable 2-level ℓ-hyperpartition $\mathscr{H}$ on X for some ℓ such that $m \leq \ell \leq M$ and*

1. *Each P_j^2 is ϵ-regular, and*

2. *There exists a weighted digraph $\widetilde{G}$ that is ϵ-close to G and is weighted ϵ-regular with respect to $\mathscr{H}$.*

Proof. Suppose, for contradiction, there exist $\epsilon > 0$, $m \in \mathbb{N}$ and weighted digraphs G_n on finite sets X_n such that $|X_n| \to \infty$ and no ϵ-equitable ℓ-hyperpartition for $m \leq \ell \leq n$ and no weighted digraph $\widetilde{G}_n$ satisfying above conditions exist for any G_n. We can take ultraproducts and obtain the weighted digraph $\mathbf{G}$ on $\mathbf{X}$ as above. There exists a 0-equitable ℓ-hyperpartition $\boldsymbol{\mathscr{H}}$ and a weighted digraph $\widetilde{\mathbf{G}}$ on $\mathbf{X}$ satisfying the conditions in Theorem 4.5.1.

We can repeat the proof of Theorem 4.3.2 to obtain an ϵ-equitable ℓ-hyperpartition $\mathscr{H}_n$ on X_n such that each $P_{k,n}^2$ is ϵ-regular, for U-almost all n. We also obtain weighted digraphs $\widetilde{G}_n$ ϵ-close to G_n that satisfy the conditions in Theorem 4.3.2, for U-almost all n.

Suppose for U-almost all n, $\widetilde{G}_n$ is not weighted ϵ-regular in $\mathscr{H}_n$. Since there are only finitely many vertex cells, there exist $i, j \in [\ell]$ such that for U-almost all n, there exist $A_{n,i} \subseteq P^1_{n,i}$, $A_{n,j} \subseteq P^1_{n,j}$ such that $\rho_v^{\widetilde{G}_n}(A_{n,i}) \geq \epsilon \rho_v^{\widetilde{G}_n}(X_n)$, $\rho_v^{\widetilde{G}_n}(A_{n,j}) \geq \epsilon \rho_v^{\widetilde{G}_n}(X_n)$ and

$$\left| \frac{\rho_e^{\widetilde{G}_n}(A_{n,i} \times A_{n,j})}{\rho_v^{\widetilde{G}_n}(A_{n,i})\rho_v^{\widetilde{G}_n}(A_{n,j})} - \frac{\rho_e^{\widetilde{G}_n}(P^1_{n,i} \times P^1_{n,j})}{\rho_v^{\widetilde{G}_n}(P^1_{n,i})\rho_v^{\widetilde{G}_n}(P^1_{n,j})} \right| \geq \epsilon.$$

Then $\mathbf{A}_i = [A_{n,i}]$ and $\mathbf{A}_j = [A_{n,j}]$ are $\sigma_{[2]}$-measurable sets such that $\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_i) \geq \epsilon \rho_v^{\widetilde{\mathbf{G}}_\mathbf{n}}(\mathbf{X})$, $\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_j) \geq \epsilon \rho_v^{\widetilde{\mathbf{G}}_\mathbf{n}}(\mathbf{X})$ and

$$\left| \frac{\rho_e^{\widetilde{\mathbf{G}}}(\mathbf{A}_i \times \mathbf{A}_j)}{\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_i)\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{A}_j)} - \frac{\rho_e^{\widetilde{\mathbf{G}}}(\mathbf{P^1_i} \times \mathbf{P^1_j}))}{\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{P^1_i})\rho_v^{\widetilde{\mathbf{G}}}(\mathbf{P^1_j})} \right| \geq \epsilon.$$

This implies that $\mathbf{P^2_j}$ is not weighted ϵ-regular in $\boldsymbol{\mathscr{H}}$ relative to $\widetilde{\mathbf{G}}$, which contradicts the conclusion from Theorem 4.5.1.

Thus we prove for U-almost all n, $\widetilde{G}_n$ is weighted-ϵ-regular with respect to $\mathscr{H}_n$, contradicting our initial assumption. □

Let $D > 1$ and $\beta > 0$. A weighted digraph G on X is (D, β)-*quasi-random* if $\rho_v^G(X) > 0$ and for any $A, B \subseteq X$ such that $\rho_v^G(A), \rho_v^G(B) \geq \beta \rho_v^G(X)$, we have

$$\frac{1}{D}\frac{\rho_e^G(X^{[2]})}{(\rho_v^G(X))^2} \leq \frac{\rho_e^G(A \times B)}{\rho_v^G(A)\rho_v^G(B)} \leq D\frac{\rho_e^G(X^{[2]})}{(\rho_v^G(X))^2}.$$

Csaba and Pluhár [4] proved a Weighted Regularity Lemma for (D, β)-quasi-random weighted digraphs where regularity is defined in terms of the weighted edge density. Their result examines the condition for regularity of sparse sets of directed edges. Below, we see how our methods and Theorems 4.5.1 and 4.5.2 yield a special case of their result.

Note that the ultraproduct of a sequence of weighted digraphs, on finite sets of increasing size, with strictly positive vertex weights will be a weighted digraph, but the vertex weight function need not be strictly positive. So the ultraproduct need not be (D, β)-quasi-random even if the weighted digraphs in the sequence are (D, β)-quasi-random. However Theorem 4.5.2 helps us prove the following result.

Theorem 4.5.3. *Given $D > 1$, $\beta, \delta > 0$ such that $\beta << \delta << \frac{1}{D}$ and $m \in \mathbb{N}$, there exist $M, N \in \mathbb{N}$ such that for any (D, β)-quasi-random weighted digraph G on a finite set X with $|X| \geq N$, there exists a δ-equitable 2-level ℓ-hyperpartition $\mathscr{H}$ on X for some ℓ such that $m \leq \ell \leq M$ and*

1. *Each P_j^2 is δ-regular, and*

2. *G is weighted δ-regular with respect to $\mathscr{H}$.*

Proof. Let $A, B \subseteq X$ such that $\rho_v^G(A), \rho_v^G(B) \geq \delta\rho_v^G(X)$. Let $x = \rho_e^G(A \times B)$, $y = \rho_v^G(A)$ and $z = \rho_v^G(B)$. Given a weighted digraph $\widetilde{G}$, let $\tilde{x} = \rho_e^{\widetilde{G}}(A \times B)$, $\tilde{y} = \rho_v^{\widetilde{G}}(A)$ and $\tilde{z} = \rho_v^{\widetilde{G}}(B)$. If G and $\widetilde{G}$ are $\frac{\epsilon}{2}$-close, then $|\tilde{x} - x| < \epsilon$, $|\tilde{y} - y| < \epsilon$ and $|\tilde{z} - z| < \epsilon$. We would like to estimate $|\frac{\tilde{x}}{\tilde{y}\tilde{z}} - \frac{x}{yz}|$.

Let $\alpha_G = \frac{\rho_e^G(X^{[2]})}{(\rho_v^G(X))^2}$ and $\beta_G = \beta\rho_v^G(X)$. Let $f(x, y, z) = \frac{x}{yz}$. Since $y, z \geq \delta\rho_v^G(X) >> \beta\rho_v^G(X)$ and G is (D, β)-quasi-random, $\frac{1}{D}\alpha_G \leq \frac{\rho_e^G(A\times B)}{\rho_v^G(A)\rho_v^G(B)} \leq D\alpha_G$. $f_x = \frac{1}{yz}$, so $1 \leq f_x(x, y, z) \leq \frac{1}{(\beta_G)^2}$. $f_y = -\frac{x}{y^2 z}$, $f_z = -\frac{x}{yz^2}$ and $\frac{\alpha_G}{D} \leq |f_y(x, y, z)|, |f_z(x, y, z)| \leq \frac{D\alpha_G}{\beta_G}$. $f_{xx} = 0$. $f_{xy} = -\frac{1}{y^2 z}$ and $f_{xz} = -\frac{1}{yz^2}$. Therefore, $1 \leq |f_{xy}(x, y, z)|, |f_{xz}(x, y, z)| \leq \frac{1}{(\beta_G)^3}$. $f_{yy} = \frac{2x}{y^3 z}$ and $f_{zz} = \frac{2x}{yz^3}$. Therefore $\frac{2\alpha_G}{D} \leq f_{yy}, f_{zz} \leq \frac{2D\alpha_G}{(\beta_G)^2}$. Also $f_{yz} = \frac{x}{y^2 z^2}$ and

$\frac{\alpha_G}{D} \leq f_{yz} \leq \frac{D\alpha_G}{(\beta_G)^2}$. We have bound all first-order and second-order derivatives using D, α_G and β_G. If $\delta \geq \epsilon >> \beta$, then using a (first-order) Taylor expansion, we can bound the difference $|\frac{\tilde{x}}{\tilde{y}\tilde{z}} - \frac{x}{yz}| \leq \epsilon\, C$, where $C > 1$ is a constant that does not depend on ϵ.

Then we can obtain $\mathscr{H}$ and $\widetilde{G}$ satisfying the conditions in Theorem 4.5.2 for $\epsilon = \frac{\delta}{2C}$ and use the above estimate to observe G is weighted δ-regular with respect to $\mathscr{H}$. $\square$

In a similar manner, we can formulate and prove extensions of these results for general weighted structures that use an appropriate definition of weighted density and weighted regularity.

Bibliography

[1] N. Alon and A. Shapira. Every Monotone Graph Property is Testable. In *Proceedings of the thirty-seventh annual ACM Symposium on Theory of Computing*, pages 128–137. ACM Press, New York, NY, USA, 2005.

[2] I. Benjamini and O. Schramm. Recurrence of Distributional Limits of Finite Planar Graphs. *Electronic Journal of Probability*, 6(23):1–13, 2001.

[3] C. Borgs, J. Chayes, and L. Lovász. Moments of Two-Variable Functions and the Uniqueness of Graph Limits. *Geometric and Functional Analysis*, 19(6):1597–1619, 2010.

[4] B. Csaba and A. Pluhár. Weighted Regularity Lemma with Applications. Preprint `arXiv:0907.0245v2`.

[5] G. Elek and B. Szegedy. A Measure-Theoretic Approach to the Theory of Dense Hypergraphs. Preprint `arXiv:0810.4062v2`.

[6] G. Elek and B. Szegedy. Limits of Hypergraphs, Removal and Regularity Lemmas. A Non-standard Approach. Preprint `arXiv:0705.2179v1`.

[7] T. Gowers. Quasirandomness, Counting and Regularity for 3-Uniform Hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):143–184, 2006.

[8] P. Loeb. Conversion from Nonstandard to Standard Measure Spaces and Applications in Probability Theory. *Transactions of the American Mathematical Society*, 211:113–122, 1975.

[9] L. Lovász and B. Szegedy. Limits of Dense Graph Sequences. *Journal of Combinatorial Theory Ser. B*, 96:933–957, 2006.

[10] R. Lyons. Asymptotic Enumeration of Spanning Trees. *Combinatorics, Probability and Computing*, 14:491–522, 2005.

[11] B. Nagle, V. Rödl, and M. Schacht. The Counting Lemma for Regular K-Uniform Hypergraphs. *Random Structures Algorithms*, 28(2):113–179, 2006.

[12] D. Offner. *Extremal Problems on the Hypercube.* PhD thesis, Carnegie Mellon University, 2009.

[13] V. Rödl and M. Schacht. Property Testing in Hypergraphs and the Removal Lemma. 39^{th} *ACM Symposium on Theory of Computing, STOC (San Diego, CA), Proceedings of STOC 2007*, pages 488–495, 2007.

[14] V. Rödl and J. Skokan. Regularity Lemma for K-Uniform Hypergraphs. *Random Structures Algorithms*, 25(1):1–42, 2004.

[15] E. Szemerédi. Regular Partitions of Graphs. *Colloques Internationaux C.N.R.S. 260 - Problèmes Combinatoires et Théorie des Graphes, Orsay*, pages 399–401, 1976.