



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Scuola di Scienze

Corso di Laurea in Matematica

Model-Theoretic Stability and its application to the Szemerédi Regularity Lemma

Tesi di Laurea in Logica Matematica

Relatore:
Chiar.mo Prof.
Martino Lupini

Presentata da:
Alessandro Chiaradia

Anno Accademico 2024/2025

A te, Luca, mia fonte di determinazione

Introduction

In the field of mathematics, the development of its subjects or simply of its theorems is rarely a linear process. Usually it requires the joint work of multiple people, which is often done in different places and especially in different times, as the passage of the years greatly increases the possibility that a new, and sometimes completely unrelated, branch of mathematics finds interest in that specific topic. Furthermore, concepts that are originally created as simple tools to be used in a proof can show themselves to have surprisingly strong properties, which can then transform these tools into the foundations of new branches and methods. This dissertation will show how this happens, as the following pages will describe two separate instances of successful repurposing of what was initially treated as a stepping stone, necessary to reach other results.

The first one, Stability Theory, greatly benefitted from the passage of time, as the story of its foundation spans three decades, and is still benefitting from it, as it is an active and diverse area of research.

The second one, the Regularity Lemma, was originally purposed by Endre Szemerédi as part of the proof for his Regularity Theorem, but then got further developed by the same author. The Lemma provides properties that are valid for any class of simple graphs: the cost behind this generality is the presence of some unavoidable caveats.

At the start of the last decade, Saharon Shelah, one of the main founders of Stability Theory, managed to successfully apply his ideas to the ones of Szemerédi in a joint paper with Maryanthe Malliaris, developing the Stable

Regularity Lemma, which tackles one of the main drawbacks of the original Lemma by restricting it to graphs that live in the realm of stability.

The structure of the dissertation will be the following: there will be a brief introduction to the main concepts of Stability Theory, along with some basic theorems that lead to stability. Then there will be a second quick introduction, this time to the Szemerédi Regularity Lemma, along with its historical background. The Lemma, while very powerful, has some unavoidable shortcomings, one of which is the presence of exceptional pairs; in order to avoid the obstacle, it becomes useful to work with just stable graphs: this simple restriction opens up a variety of different Stable Regularity Lemmas. The one that will be described, along with the process required to reach it, will be the most basic version, which requires only to work under stable graph theory.

The chapter about Stable Theory takes its notation mainly from the Lecture Notes of Artem Chernikov (see [1]), while the Stable Regularity Lemma is from a paper written by Maryanthe Malliaris and Saharon Shelah (see [2] and [3]), with some additional notation and definitions from [4] and [5].

Contents

Introduction	i
1 Model-Theoretic Stability	1
1.1 Basic definitions of Model Theory	2
1.2 Saturation and the monster model	3
1.3 Stability	5
2 The Szemerédi Regularity Lemma on Stable Graphs	10
2.1 Basic definitions of Graph Theory	11
2.2 The Szemerédi Regularity Lemma	12
2.3 Stable Graphs	14
2.4 The Stable Regularity Lemma	21
2.5 Other Stable Regularity Lemmas	24
Bibliography	26

Chapter 1

Model-Theoretic Stability

The concept of stability originally emerged as a tool that Saharon Shelah used to extend the work of Micheal Morley (and previously of Robert Vaught and Jerzy Łoś) on the number of possible models of a certain cardinality. More precisely, if we let $I_T(\kappa)$ be the number of models of a theory T (up to isomorphism) that have cardinality κ , Morley proved that if there is an uncountable cardinal κ and a theory T for which $I_T(\kappa) = 1$, then $I_T(\lambda) = 1$ for any other uncountable cardinal λ . In addition, Morley conjectured that $I_T(\kappa)$ is non-decreasing on uncountable cardinals for any countable theory T . Shelah, thanks to the studies that he did in order to prove Morley's conjecture, managed to show that first-order theories can be separated according to some “dividing lines”: if a theory does not follow the rules dictated by said lines, it can have as many models as possible, but if a theory follows them, then the isomorphism type of its models can be described with some invariants; for example, both sets and vector spaces have a single invariant, which is their cardinality and dimension respectively.

One of the “dividing lines” identified by Shelah was stability, which with the passage of time revealed itself to be widely applicable to larger contexts, some of which are still an active area of research.

1.1 Basic definitions of Model Theory

The following notation will be the one used in the dissertation:

- $\mathcal{M} = \langle M; \mathcal{L} \rangle$ is a first-order structure, where M is a set and $\mathcal{L} = \{R_1, R_2, \dots, f_1, f_2, \dots\}$ is the signature of the structure; furthermore, each R_i is a n_i -ary relation for a certain $n_i \in \omega$ (which means that $R_i \subseteq M^{n_i}$), and each f_j is a function $f_j : M^{n_j} \rightarrow M$ for a certain $n_j \in \omega$; in particular, if $n_k = 0$, then f_k will be called a constant (and the notation $f_k := c_k$ will be used).
- A formula $\varphi(y_1, \dots, y_n)$, where every y_i is a free variable, will be any boolean combination of well-defined concatenations of relations and functions from $\mathcal{L}$, and L will be the set of all formulas; formulas with parameters will be written as $\varphi(y, b)$, where y is a tuple of free variables and b is a tuple of parameters from $B \subseteq M$, and $L(B)$ will be the set of all formulas with parameters from B ; a formula without free variables is called a sentence.
- $\mathcal{M} \models \varphi(a)$, also written as $a \models \varphi(y)$, will mean that a satisfies $\varphi(y)$, or equivalently that a is a solution of $\varphi(y)$; for any set of formulas $\Phi(y)$, the expression $a \models \Phi(y)$ will mean $\forall(\varphi(y) \in \Phi(y))(a \models \varphi(y))$.
- A theory T is any consistent set of sentences in a language L , and T is said to be complete when for every L -sentence, either it is in T or its negation is in T ; $\text{Th}(\mathcal{M})$ will be the theory of $\mathcal{M}$, which is the set of all sentences that are true in $\mathcal{M}$.

In order to simplify reading and avoid excessive repetition, the following writing conventions will be used:

- x and y , when used as dependent or independent variables (like in $\varphi(x)$), are to be treated as tuples of variables.
- $\forall(\psi(x))(\varphi(x))$ will mean $\forall x(\psi(x) \rightarrow \varphi(x))$, and $\forall^{\geq n}x(\varphi(x))$ will mean that the length of the tuple x is at least n ; the same is true for $\exists$.

- If x is a tuple and M is the set of a model, then M_x will refer to the tuples of length $|x|$ (equivalently, $M_x := M^{|x|}$).
- κ and λ will be cardinals, both finite or infinite.

Some theorems that will prove themselves useful are the following:

Theorem 1.1 (Compactness Theorem). *For any language L and any set of formulas Ψ in L , if every finite subset of Ψ is consistent, then Ψ is consistent.*

Theorem 1.2 (Ramsey's Theorem). *For any k -coloring of subsets of $\mathbb{N}$ of size n , there is some infinite subset I of $\mathbb{N}$ such that all subsets of I with n elements have the same color.*

There are many theorems that involve some variant of Ramsey's Theorem; in order to simplify this process, the arrow notation is introduced:

Definition 1.3 (Arrow notation, from [7]). *The notation $\kappa \rightarrow (\lambda)_m^n$ will mean the following: for any m -coloring of subsets of κ of size n , there is some subset u of κ with order type λ such that every subset of u with n elements has the same color.*

Remark. Using the arrow notation it is possible to rewrite Ramsey's Theorem in a more compact way: $\forall(n, k \in \omega)(\aleph_0 \rightarrow (\aleph_0)_k^n)$.

Ramsey's Theorem is used to prove this last Theorem:

Theorem 1.4 (Erdős-Makkai Theorem). *Let $\mathcal{F} \subseteq \mathcal{P}(B)$ be a collection of subsets of B , with $\omega \leq |B| < |\mathcal{F}|$, then there are two sequences $\{b_i\}_{i \in \omega} \subseteq B$ and $\{S_i\}_{i \in \omega} \subseteq \mathcal{F}$ such that one between $\forall(i, j \in \omega)(b_i \in S_j \leftrightarrow j < i)$ and $\forall(i, j \in \omega)(b_i \in S_j \leftrightarrow i < j)$ is true.*

1.2 Saturation and the monster model

From now on, any theory will be complete. Before introducing stability, it would be useful to work in a model large enough to let us assume that every

other model (of the same theory) is embedded in it. Here is the process to do it:

Definition 1.5. *Let $\mathcal{M}$ and $\mathcal{N}$ be two structures of the same language L :*

- *$\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent, also written as $\mathcal{M} \equiv \mathcal{N}$, if $\text{Th}(\mathcal{M}) = \text{Th}(\mathcal{N})$.*
- *A partial map $f : \mathcal{M} \rightarrow \mathcal{N}$ is elementary if for all $a \in \text{Dom}(f)$ and for all $\varphi \in L$ it is true that $(\mathcal{M} \models \varphi(a) \leftrightarrow \mathcal{N} \models \varphi(f(a)))$.*
- *$\mathcal{M}$ is an elementary substructure of $\mathcal{N}$, also written as $\mathcal{M} \preceq \mathcal{N}$, if the embedding map $\iota : \mathcal{M} \hookrightarrow \mathcal{N}$ is elementary.*

Definition 1.6 (Partial type). *Let $\Phi(x) = \{\varphi(x)\}$ be a collection of formulas with parameters from a certain set A in a model $\mathcal{M}$; then $\Phi(x)$ is called a partial type over A if each one of its finite subcollections has a common solution in $\mathcal{M}$.*

The concept of type will be described with higher precision in the following pages.

Definition 1.7 (Saturation and homogeneity). *Given an infinite cardinal κ and a model $\mathcal{M}$:*

- *$\mathcal{M}$ is κ -saturated if for any set of parameters A in $\mathcal{M}$ with $|A| < \kappa$, every partial type $\Phi(x)$ over A with $|x| < \kappa$ can be realized in $\mathcal{M}$.*
- *$\mathcal{M}$ is saturated if it is $|\mathcal{M}|$ -saturated.*
- *$\mathcal{M}$ is κ -homogenous if any partial elementary map from $\mathcal{M}$ to itself with domain of cardinality smaller than κ can be extended to an automorphism of $\mathcal{M}$.*

Through the compactness theorem, it can be proved that for any theory T and infinite cardinal κ there is at least a κ -saturated and κ -homogenous model over T . On the other hand, the existence of saturated models is not guaranteed and heavily depends on the specifics of the set-theoretic assump-

tions.

Consider the following commutative diagram of elementary embeddings:

$$\begin{array}{ccc} \mathcal{M} & \hookrightarrow & \mathcal{N} \\ & \searrow & \downarrow \\ & & \mathcal{O} \end{array}$$

In the diagram, $\mathcal{M}$, $\mathcal{N}$ and $\mathcal{O}$ are models of the same complete L -theory T . Thanks to the completeness of the theory, the diagram can be realized in a single model of the same theory, where the embeddings are treated as inclusions between elementary submodels. In order to apply this line of reasoning to every embedding, the following definition is given:

Definition 1.8 (Monster model). *The monster model $\mathbb{M}$ is a model that satisfies the previously discussed purpose and is $\kappa(\mathbb{M})$ -saturated and $\kappa(\mathbb{M})$ -homogenous for a sufficiently large cardinal $\kappa(\mathbb{M})$. A cardinal λ will be called small if $\lambda < \kappa(\mathbb{M})$. For any $\varphi(x) \in L(\mathbb{M})$ and $a \in \mathbb{M}$, the expression $\models \varphi(a)$ will be used to denote $\mathbb{M} \models \varphi(a)$. For any pair of sets of formulas $\Phi(x)$ and $\Psi(x)$, the expression $\Phi(x) \vdash \Psi(x)$ will mean $\forall (a \in \mathbb{M})(\models \Phi(a) \rightarrow \models \Psi(a))$.*

An immediate consequence of the definition is that every model of size at most $\kappa(\mathbb{M})$ embeds elementarily into $\mathbb{M}$. From now, unless differently stated, expressions like “a model” or “a set of parameters” will mean “an elementary submodel of $\mathbb{M}$ ” and “a set of parameters in $\mathbb{M}$ ”.

1.3 Stability

Definition 1.9 (Order property and stability). *Given $M \models T$ and $k \in \omega$:*

- *A formula $\varphi(x, y)$ has the k -order property if, for $i < k$, there are some $a_i \in M_x$ and $b_i \in M_y$ for which it is true that $M \models \varphi(a_i, b_j) \leftrightarrow i < j$.*
- *A formula $\varphi(x, y)$ has the order property if it has the k -order property for every $k \in \omega$.*
- *A formula $\varphi(x, y)$ is stable if it does not have the order property (equiv-*

alently, there exists at least one $k \in \omega$ for which the formula does not have the k -order property).

- A theory is stable if it implies the stability of all formulas.

When talking about stability, it is useful to take into consideration types, a mathematical object that categorizes formulas. While the concept of partial type has already been given (1.6), there is a high variety of types, each with its own specific definition; the ones that will be used in this dissertation are the following two:

Definition 1.10 (Complete type). $\Phi(x)$ is a complete type over A if it is a partial type that contains either $\varphi(x)$ or $\neg\varphi(x)$, for every $\varphi(x) \in L(A)$. For $b \in \mathcal{M}$, the complete type of b over A is defined as $\text{tp}(b/A) = \{\varphi(x) \in L(A) : b \models \varphi(x)\}$.

Definition 1.11 (Complete φ -type). If $\varphi(x, y)$ is a formula and $A \subseteq M_y$ is a set of parameters, then a complete φ -type over A is a maximal and consistent collection of formulas in the form $\varphi(x, b)$ and $\neg\varphi(x, b)$, with b ranging over A . The space of all complete φ -types over A will be called $S_\varphi(A)$.

An analogous definition to the one of complete φ -types holds for complete Δ -types, where Δ is a set of formulas, in which the same properties have to hold for all formulas in Δ .

In addition to types, in model theory there is a vast array of ranks that are used to obtain orderings from which it is possible to deduce properties, and types are not exempt from this process:

Definition 1.12 (Shelah's local 2-rank). Let Δ be a set of L -formulas and $\theta(x)$ a partial type over $\mathbb{M}$. The Shelah's local 2-rank $R_\Delta(\theta(x))$ is inductively defined on $n \in \omega$ using the following rules:

- If $\theta(x)$ is inconsistent, $R_\Delta(\theta(x)) = -1$.
- If $\theta(x)$ is consistent, $R_\Delta(\theta(x)) \geq 0$.
- If there are $\varphi(x, y) \in \Delta$ and $a \in \mathbb{M}_y$ such that $R_\Delta(\theta(x) \wedge \varphi(x, a)) \geq n$

and $R_\Delta(\theta(x) \wedge \neg\varphi(x, a)) \geq n$ are both true, then $R_\Delta(\theta(x)) \geq n + 1$.

- If $R_\Delta(\theta(x)) \geq n$ and $R_\Delta(\theta(x)) \not\geq n + 1$, then $R_\Delta(\theta(x)) = n$.
- If $R_\Delta(\theta(x)) \geq n$ for all $n \in \omega$, then $R_\Delta(\theta(x)) = \infty$.

Shelah's local 2-rank can be generalized to the following rank:

Definition 1.13. For any type p , any set of formulas Δ and finite cardinal λ , the rank $R(p, \Delta, \lambda)$ is inductively defined:

- If p is inconsistent, $R(p, \Delta, \lambda) = -1$.
- If p is consistent, $R(p, \Delta, \lambda) \geq 0$.
- If δ is a limit ordinal and $R(p, \Delta, \lambda) \geq \alpha$ for all $\alpha < \delta$, then $R(p, \Delta, \lambda) \geq \delta$.
- If for all $\mu < \lambda$ and finite $q \subseteq p$ there is a sequence of Δ -types $\{q_i\}_{i \leq \mu}$ such that for each $i \neq j$ there is at least a formula φ such that $\varphi \in q_i$ and $\neg\varphi \in q_j$, and $R(q \cup q_i, \Delta, \lambda) \geq \alpha$ for all $i \leq \mu$, then $R(p, \Delta, \lambda) \geq \alpha + 1$.
- If $R(p, \Delta, \lambda) \geq \alpha$ and $R(p, \Delta, \lambda) \not\geq \alpha + 1$, then $R(p, \Delta, \lambda) = \alpha$.
- If $R(p, \Delta, \lambda) \geq \alpha$ for all α , then $R(p, \Delta, \lambda) = \infty$.

Remark. If $R_\Delta(p) \in \omega$, then $R(p, \Delta, 2) = R_\Delta(p)$.

Shelah's local 2-rank represents a tool to quickly obtain a necessary and sufficient condition for the stability of a formula. Before showing it, a theorem about linear orders is needed:

Theorem 1.14. Let I be a linear order and $\varphi(x, y)$ an unstable formula in a (necessairly unstable) theory T , then there is a model $\mathcal{M} \models T$ and a sequence $\{(a_i, b_i)\}_{i \in I}$ in M such that $\forall(i, j \in I)(\mathcal{M} \models \varphi(a_i, b_j) \leftrightarrow i < j)$.

Theorem 1.15. For any formula $\varphi(x, y)$ the following are equivalent:

1. $\varphi(x, y)$ is stable.

2. $R_\varphi(x = x) < \omega$, where $x = (x_i)_{i \in I}$ is a tuple of variables, $x = x$ means $(x_1 = x_1 \wedge \dots \wedge x_i = x_i \wedge \dots)_{i \in I}$ and R_φ stands for $R_{\{\varphi\}}$. Notice that $R_\varphi(x = x) < \omega$ implies that $R_\varphi(\theta(x)) < \omega$ for any partial type $\theta(x)$.

Idea of proof. $[2 \Rightarrow 1]$: assume that $\varphi(x, y)$ is unstable, then it has the k -order property for each $k \in \omega$. Thanks to 1.14, there is a $(a_i, b_i)_{i \in [0,1]}$ such that $\models \varphi(a_i, b_j) \leftrightarrow i < j$: this implies that both $\varphi(x, b_{\frac{1}{2}})$ and $\neg\varphi(x, b_{\frac{1}{2}})$ contain dense subsequences of a_i , thanks to which it is possible to increase $R_\varphi(x = x)$. This process can be reiterated, from which the absurd $R_\varphi(x = x) = \infty$ is reached.

$[1 \Rightarrow 2]$: assume that $R_\varphi(x = x) = \infty$, then it is possible to build an infinite tree of parameters B such that every infinite branch maintains the consistency of the set of formulas necessary for the increasing of the rank: this leads to $|S_\varphi(B)| > |B|$. This makes $\varphi(x, y)$ unstable as it is possible to obtain (through the Erdős-Makkai Theorem) the two infinite sequences of parameters necessary to satisfy the requirements for instability. $\square$

As a last proof for this chapter, here is a necessary and sufficient condition for the stability of a theory:

Theorem 1.16. *For any complete theory T the following are equivalent:*

1. T is stable.
2. There is no formula $\varphi(z_1, z_2) \in L(\mathbb{M})$ and sequence $\{c_i\}_{i \in \omega}$ of tuples in $\mathbb{M}$ such that $(\models \varphi(c_i, c_j) \leftrightarrow i < j)$.

Proof. $[1 \Rightarrow 2]$ Let $\varphi(z_1, z_2) \in L(\mathbb{M})$ be a formula for which there is a sequence $\{c_i\}_{i \in \omega}$ of tuples in $\mathbb{M}$ such that $(\models \varphi(c_i, c_j) \leftrightarrow i < j)$; then $\varphi(z_1, z_2)$ has the order property because $\{c_i\}_{i \in \omega}$ satisfies the definition.

$[2 \Rightarrow 1]$ Assume T unstable, then there is a formula $\varphi(z_1, z_2) \in L(\mathbb{M})$ that has the order property, which means that there is at least a sequence (a_i, b_j) that satisfies the k -order property definition for each $k \in \omega$. Now let

$\psi(x_1y_1, x_2y_2) := \varphi(x_1, y_2)$, then, if $c_i := a_ib_i$, the condition $(\models \psi(c_i, c_j) \leftrightarrow i < j)$ is verified. $\square$

Finally, the following are two properties strictly related to stability, which will be extensively used in the second chapter:

Definition 1.17 (Indiscernibility). *Let I be a linear order, A a set of parameters (omitted if $A = M$) and $\{a_i\}_{i \in I}$ a sequence of tuples, then that sequence is said to be indiscernible over A if for all $n \in \omega$ and for all $i_0 < \dots < i_n$ and $j_0 < \dots < j_n$ from I , it is true that $\models \varphi(a_{i_0}, \dots, a_{i_n}) \leftrightarrow \models \psi(a_{j_0}, \dots, a_{j_n})$ for all formulas $\varphi, \psi \in L(A)$. It is possible to take in consideration just the formulas from a specific set Δ : in this case we will talk about Δ -indiscernibility. Additionally, the sequence is totally indiscernible if it is indiscernible independently from the order of the indices.*

Definition 1.18 (Independence). *A formula $\varphi(x, y)$ has the independence property if there are two infinite sequences $\{b_i\}_{i \in \omega}$ and $\{a_s\}_{s \subseteq \omega}$ such that $(\models \varphi(a_s, b_i) \leftrightarrow i \in s)$. A theory has the independence property if some of its formulas have it, otherwise it has the dependence property.*

Both indiscernibility and independence quickly lead to stability or instability:

Theorem 1.19. *A theory is stable if and only if it makes totally indiscernible every indiscernible sequence.*

Theorem 1.20 (Shelah). *If a theory is independent, then it is unstable.*

The reverse implication of the last theorem is not generally true, as there is a second property, called the strict order property, which also implies instability. However, every unstable theory has either the independence property or the strict order property.

Chapter 2

The Szemerédi Regularity Lemma on Stable Graphs

As said in the introduction of the previous chapter, model-theoretic stability can be used in combination with ideas from others fields of mathematics in order to obtain new properties: one such field is Graph Theory.

The following pages will show how the restriction to stability related properties improves the result of what is known as the Szemerédi Regularity Lemma. The story behind the Lemma is surprisingly similar to the one behind stability: originally used to prove a theorem about arithmetic progression, with the passage of time it showed itself to be applicable to other fields, such as algorithmic approximations (see [11] for an example), thanks to the works of many important mathematicians such as Timothy Gowers, János Komlós and Terence Tao.

2.1 Basic definitions of Graph Theory

The concept of graph can be defined in various similar ways. In this dissertation we will follow the model-theoretic definition:

Definition 2.1 (Graph). *A graph is a structure $\mathcal{G} = \langle G; \mathcal{L} \rangle$ whose signature $\mathcal{L} = \{=, R_G\}$ contains the equality relation and an irreflexive, symmetric binary relation R_G , which represents the property of being connected by an edge (aR_Gb means that a and b are connected by an edge). The elements of G are called vertices or nodes, and the cardinality of a graph, written as $|\mathcal{G}|$ or more often as $|G|$, is defined as being the cardinality of the set of its vertices.*

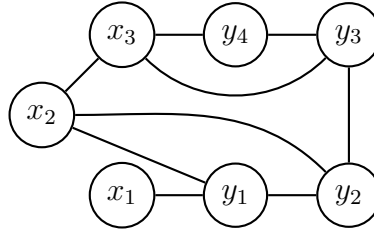
Usually, instead of $\mathcal{G} = \langle G; \{=, R_G\} \rangle$, the less convoluted notation $\mathcal{G} = \langle G; R_G \rangle$ will be used, with $R := R_G$ if it is clear from context.

A graph is called complete if $\forall(x, y \in G)(x \neq y \rightarrow xR_Gy)$ (i.e.: each distinct pair of vertices is connected by an edge) and empty if $\forall(x, y \in G)(\neg(xR_Gy))$ (i.e.: the graph is edgeless). A graph is called bipartite if the set of its nodes can be split into two sets such that no nodes in the same set are connected. If $A \subseteq G$, the notation $\mathcal{G}(A)$ will refer to the restriction of the graph to the nodes from A . Additionally, if $A, B \subseteq G$ are such that $A \cap B = \emptyset$, then the notation $\mathcal{G}(A, B)$ will refer to the bipartite subgraph with node set $A \cup B$ and edges obtained through the rule that two nodes can be connected if and only if they are connected in G and aren't both in A or in B .

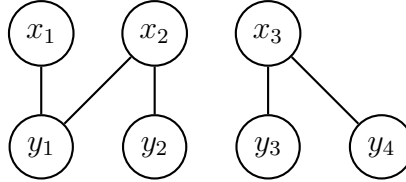
Definition 2.2 (Density). *Let $\mathcal{G}$ be a graph and G the set of its vertices, then for all $X, Y \subseteq G$, the expression $e(X, Y)$ denotes the number of edges between X and Y . If $X \cap Y = \emptyset$ and both X and Y are nonempty, then the following density is defined: $d(X, Y) = \frac{e(X, Y)}{|X| \cdot |Y|}$.*

It is clear that $d(X, Y) \in [0, 1]$, with $d(X, Y) = 0$ if and only if the bipartite subgraph $\mathcal{G}(X, Y)$ is empty and $d(X, Y) = 1$ if and only if $\mathcal{G}(X, Y)$ has the highest possible amount of edges; also, if $\mathcal{G}$ is complete, then $d(X, Y) = 1$, with the converse not being true, as the only edges that are taken in consideration are the ones between X and Y .

Example. Consider the graph $\mathcal{G} = \langle G; R \rangle$, with $G = X \cup Y$, where $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2, y_3, y_4\}$; let R be defined according to the following picture:



Because X, Y are disjoint, the graph $\mathcal{G}(X, Y)$ exists:



Furthermore, it is possible to evaluate the density:

$$d(X, Y) = \frac{e(X, Y)}{|X| \cdot |Y|} = \frac{5}{3 \cdot 4} = \frac{5}{12}$$

2.2 The Szemerédi Regularity Lemma

The following theorem is the reason behind the development of the Szemerédi Regularity Lemma:

Theorem 2.3 (Szemerédi Regularity Theorem). *Let $A \subseteq \mathbb{N}$ and*

$$D^*(A) := \limsup_{N \rightarrow +\infty} \frac{|A \cap \{1, \dots, N\}|}{N};$$

if $D^(A) > 0$, then for any integer $k \geq 3$ it is possible to find an arithmetic progression of length k located entirely in A .*

The proof of this theorem is quite complex and makes use of a diverse set of mathematical tools, one of which is graph theory, in particular the branch about bipartite graphs. To be more precise, using Szemerédi's words, “any

large bipartite graph can be decomposed into nearly regular bipartite sub-graphs” (directly from [10]). To better understand the meaning behind this, it is necessary to introduce ε -regularity:

Definition 2.4 (ε -regularity). *Given $\varepsilon > 0$, a graph $\mathcal{G}$ and $A, B \subseteq G$ disjoint, then the pair (A, B) is said to be ε -regular if, for all $X \subseteq A$ and $Y \subseteq B$ such that $|X| > \varepsilon|A|$ and $|Y| > \varepsilon|B|$, it is true that $|d(X, Y) - d(A, B)| < \varepsilon$.*

The following is one of the many equivalent ways to write the Szemerédi Regularity Lemma, more precisely the version published by Szemerédi in 1978, which is stronger than the one that he used to prove 2.3:

Theorem 2.5 (Szemerédi Regularity Lemma). *For every $\varepsilon \in [0, 1]$ and $m \in \omega$, there are $N = N(\varepsilon, m)$ and $M = M(\varepsilon, m)$ such that for any finite graph G of size $|G| \geq N$, there exists a $k \in \omega$ and a partition $G_1 \cup \dots \cup G_k = G$ that satisfies the following:*

- $m \leq k \leq M$.
- $||G_i| - |G_j|| \leq 1$ for all $i, j \leq k$.
- All but at most $\lfloor \varepsilon k^2 \rfloor$ of the pairs (G_i, G_j) are ε -regular.

Remark. The base form of the Szemerédi Regularity Lemma has some limitations, the main two being:

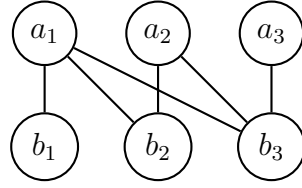
- The size of the partition is very large (proved in [8]), with the lower bound being a tower of 2s with height proportional to $\log(\varepsilon^{-1})$ and the upper bound being a tower of 2s with height proportional to ε^{-5} .
- The condition of ε -regularity is not universal, as there can be up to $\lfloor \varepsilon k^2 \rfloor$ pairs that do not have it: these are called irregular pairs.

The presence of irregular pairs is necessary if we want to cover every possible graph: an example of graph that needs them is the n -half-graph, defined as

$$\mathcal{H}_n := \underbrace{\langle \{a_1, \dots, a_n\} \cup \{b_1, \dots, b_n\} \rangle}_{H=H_n}; \underbrace{\langle \{(a_i, b_j), (b_j, a_i) : i \leq j\} \rangle}_{R=R_H}.$$

In other words, a n -half-graph is a graph whose nodes can be partitioned into two (numbered) empty subgraphs of equal size, both of which can be represented as a sequence in such a way that a node from the first partition set is connected to a node from the second partition set if and only if the index of the first node is lesser or equal than the index of the second node.

Example. The following is the visualization of $\mathcal{H}_3$:



In order to avoid these obstacles, it becomes useful to develop more specific variants of the Lemma, trading the generality of the original one for stronger properties.

2.3 Stable Graphs

From now on, $\zeta = \zeta(x, y)$ will be used to refer to the formula xRy , with R being the edge relation of the graph that is being taken into consideration. In order to apply Stability Theory to graphs, the following definition is introduced:

Definition 2.6 (Non- k -order property). *Let $\mathcal{G}$ be a graph and $R = R_{\mathcal{G}}$ the binary relation that defines the edges. The graph has the non- k -order property if, for each $i < k$, there is no pair $a_i, b_i \in G$ that satisfies $((i < j < k) \rightarrow ((a_i R b_j) \wedge \neg(a_j R b_i)))$. If such configuration exists, then the graph is said to have the k -order property.*

An immediate consequence of this definition is that if a graph $\mathcal{G}$ has the k -order property for all k , then ζ has the order property, making ζ unstable. Also, notice how, for each n , ζ has the $(n+1)$ -order property in the half-graph $\mathcal{H}_n$: this means that the theory of half-graph is unstable. This is the main

reason behind the choice to work in stable conditions, as that eliminates half-graphs, which are one of the major classes of graphs that make exceptional pairs unavoidable in the Szemerédi Regularity Lemma.

From this point onwards, unless differently stated, we will consider G as being a graph with edge relation R and the non- k_* -order property, for a certain $k_* \in \omega$, which means that it makes ζ stable.

Definition 2.7. $\Delta_k = \{x_0 R x_1\} \cup \{\varphi_{k,m}^i : m \leq k \wedge i \in \{1, 2\}\}$, where:

$$\begin{aligned}\varphi_{k,m}^1 &= \varphi_{k,m}^1(x_0, \dots, x_{k-1}) = \exists y \left(\bigwedge_{\ell < m} (x_\ell R y) \right) \\ \varphi_{k,m}^2 &= \varphi_{k,m}^2(x_0, \dots, x_{k-1}) = \exists y \left(\bigwedge_{m \leq \ell < k} (x_\ell R y) \right)\end{aligned}$$

Example. $\Delta_2 = \{x_0 R x_1, \varphi_{2,0}^1, \varphi_{2,1}^1, \varphi_{2,2}^1, \varphi_{2,0}^2, \varphi_{2,1}^2, \varphi_{2,2}^2\}$, with:

$$\begin{aligned}\varphi_{2,0}^1 &= \exists y & \varphi_{2,0}^2 &= \exists y (x_0 R y \wedge x_1 R y \wedge x_2 R y) \\ \varphi_{2,1}^1 &= \exists y (x_0 R y) & \varphi_{2,1}^2 &= \exists y (x_1 R y \wedge x_2 R y) \\ \varphi_{2,2}^1 &= \exists y (x_0 R y \wedge x_1 R y) & \varphi_{2,2}^2 &= \exists y (x_2 R y)\end{aligned}$$

Δ_k will be the set over which the indiscernibility property will be required. The following is a sufficient condition for the k -order property that will show itself to be very useful for building proofs by contradiction:

Theorem 2.8. *Let H be a finite graph and $A = \langle a_i : i < \alpha \rangle$ be a Δ_k -indiscernible sequence in H , with $\alpha > 2k$. If there is a vertex $b \in H$ and a sequence of indices $i_0 < \dots < i_{2k-1} < \alpha$ for which it is true that $b R_H a_{i_\ell}$ for all $0 \leq \ell \leq k-1$ and $\neg b R_H a_{i_\ell}$ for all $k \leq \ell < 2k$, then H has the k -order property.*

Proof. For all $m \in \{0, \dots, k-1\}$, let $c = (c_j)_{0 \leq j \leq k-1}$, with $c_j = a_{m+i}$; then the existence of b guarantees the truth in H of $\varphi_{k,m}^1(c)$. This holds for every increasing subsequence of k elements from A , from which the k -order property is obtained. $\square$

Theorem 2.9. *Let $\langle a_i : i < m \rangle$ be a Δ_{k_*} -indiscernible sequence in a graph G with $m \geq 4k_*$; for $b \in G$, let $X = \{i : a_i R b\}$ and $Y = \{i : \neg(a_i R b)\}$, then one between X and Y has cardinality less than $2k_*$.*

Proof. By contradiction, suppose $|X|, |Y| \geq 2k_*$. By defining i_1 to be the k_*^{th} element of X and i_2 to be the k_*^{th} element of Y , three options are possible:

1. $i_1 = i_2$, which is impossible given the definitions of X and Y .
2. $i_1 < i_2$, which is impossible because, from the hypothesis, there would be a subsequence $a_{j_1} < \dots < a_{j_{2k_*}} \leq a_m$ that satisfies both $\{j_1 < \dots < j_{k_*} = i_1\} \subseteq X$ and $\{i_2 = j_{k_*+1} < \dots < j_{2k_*}\} \subseteq Y$. This would satisfy 2.8, making G have the k_* -order property, which is absurd.
3. $i_1 > i_2$, which is impossible through the same reasoning of the previous option, obtained by replacing R with $\neg R$.

Thus the pair (i_1, i_2) cannot exist. □

The theorem roughly states that each Δ_{k_*} -indiscernible sequence with at least $4k_*$ elements can be split into two sets, defined according to X and Y , of which one is “small” (has cardinality less than $2k_*$) and the other one is “large” (has cardinality more than $2k_*$).

Theorem 2.10. *Let $A = \{a_i\}_{i < s_1}$ and $B = \{b_j\}_{j < s_2}$ be Δ_{k_*} -indiscernible sequences, with $s_1 \geq 2k_*$ and $s_2 \geq (2k_*)^2$. By defining:*

$$U = \{i < s_1 : \exists^{\geq 2k_*} (j < s_2)(a_j R b_i)\}$$

then either $|U| \leq 2k_$ or $|U| \geq s_1 - 2k_*$.*

Proof. Suppose $2k_* < |U| < s_1 - 2k_*$. By defining i_1, i_2 to be the k_*^{th} element of U and the k_*^{th} element of $V := \{0, \dots, s_1 - 1\} \setminus U$ respectively, three options are possible:

1. $i_1 = i_2$, which is impossible given the definitions of i_1 and i_2 .

2. $i_1 < i_2$; in this case there are $\{j_0 < \dots < j_{k_*-1}\}$ from U and $\{j_{k_*} < \dots < j_{2k_*-1}\}$ from V such that $j_{k_*-1} \leq i_1 < i_2 \leq j_{k_*}$ and $j_{2k_*-1} < s_1$. Because $s_2 > 4k_*^2 \geq 4k_*$, thanks to 2.9 it can be said that, for each a_{j_ℓ} , with $\ell < 2k_*$, the set B can be partitioned into two sets, one larger than the other. Let:

$$W_\ell = \{i < s_2 : a_{j_\ell} R b_i \leftrightarrow ((\exists^{\geq 2k_*} i < s_2)(\neg a_{j_\ell} R b_i))\}$$

be the smaller set, which has in fact cardinality $|W_\ell| < 2k_*$. By defining Z to be the union of all the W_ℓ , the following holds:

$$|Z| = \left| \bigcup_{\ell < 2k_*} W_\ell \right| \leq (2k_*)^2 < |B|$$

Let now $n \in \{0, \dots, s_2 - 1\} \setminus Z$, then it is true that $(0 \leq \ell \leq k_* - 1) \rightarrow (b_n R a_\ell)$ and $(k_* \leq \ell < 2k_*) \rightarrow (\neg b_n R a_\ell)$. But then 2.8 would imply that G has the k_* -order property, which is absurd.

3. $i_1 > i_2$, which is impossible through the same reasoning of the previous option, obtained by replacing R with $\neg R$.

Thus the pair (i_1, i_2) cannot exist. $\square$

Definition 2.11 (From [6]). *Given a set of formulas Γ and n_1, n_2 finite ordinals, the expression $n_1 \rightarrow (n_2)_\Gamma$ has the following meaning: for any $\{a_i\}_{i < n_1}$ sequence of elements of G , there is a Γ -indiscernible subsequence $\{a_{i_j}\}_{j < n_2} \subseteq \{a_i\}_{i < n_1}$ that is not constant.*

The definition can be expanded to include any cardinal n_1 and any ordinal n_2 , as well as any theory T and any tuple length $\ell(a_i)$ (in our case, $T = \text{Th}(\mathcal{G})$ and $\ell(a_i) = 1$).

Theorem 2.12. *Using the notation introduced in 1.3:*

$$\left(n_1 \rightarrow (n_2)_{2^{| \Delta_{k_*} |}}^{k_*} \right) \Rightarrow \left(n_1 \rightarrow (n_2)_{\Delta_{k_*}} \right)$$

Proof. Remember that $n_1 \rightarrow (n_2)_{2^{| \Delta_{k_*} |}}^{k_*}$ means that for any $2^{| \Delta_{k_*} |}$ -coloring of subsets of n_1 of size k_* , there is some subset u of n_1 with order type n_2 such

that every subset of u with k_* elements has the same color; the aim is to show that every sequence $\{a_i\}_{i < n_1} \subseteq G$ has a nonconstant Δ_{k_*} -indiscernible subsequence $\{a_{i_j}\}_{j < n_1}$. Consider all the possible increasing subsequences of n_1 that have k_* elements, and color them in such a way that two subsequences $\{i_1, \dots, i_{k_*}\}$ and $\{j_1, \dots, j_{k_*}\}$ have the same color if and only if the property $(G \models \gamma(a_{i_1}, \dots, a_{i_{k_*}}) \leftrightarrow G \models \gamma(a_{j_1}, \dots, a_{j_{k_*}}))$ holds for the same subset of formulas from Δ_{k_*} ; notice that there are exactly $2^{|\Delta_{k_*}|}$ possible colorings. $u \subseteq n_1$, with u being the one described at the start of the proof, will be the set of (less than $n_2 + 1$) indices from which the Δ_{k_*} -indiscernible subsequence can be obtained. $\square$

Definition 2.13 (Tree Order). *Remember that the natural partial order on $\omega^{>\omega}$ (which is the set of all the tuples of natural numbers which have finite length) is given by $\nu \leq \rho \leftrightarrow \nu$ is an initial segment of ρ (for example, $(3, 7, 2) \leq (3, 7, 2, 8)$, but $(3, 7, 2) \not\leq (3, 7, 5, 8)$). Given a finite set of vertices S , a tree order on S is a partial order which is order-isomorphic to some downward closed subset of $\omega^{>\omega}$ under the natural partial order.*

Theorem 2.14. *Let Δ be a set of formulas closed under variable cycling and $k \in \omega$ such that each formula in Δ has at most k free variables. If there is a $k_2 \in \omega$ for which every formula $\varphi(x_1, \dots, x_{k-1}) \in \Delta$ is such that every partition of its free variables in the form $\{\{x_0, \dots, x_\ell\}, \{x_{\ell+1}, \dots, x_{k-1}\}\}$ makes $\varphi(x_1, \dots, x_\ell; x_{\ell+1}, \dots, x_{k-1})$ have the non- k_2 -order property, then the following are true:*

1. *There is a $r \in \omega$ such that for every A subset of G with at least 2 elements, $|S_\Delta(A)| \leq |A|^r$.*
2. *For each $A = \{a_i\}_{i < n} \subseteq G$ there is a $u \subseteq n$ such that $\{a_i\}_{i \in u}$ is Δ -indiscernible and $|u| \geq f^k(n)$, with $f(x) = \left\lfloor \frac{x}{t} \right\rfloor^{\frac{1}{tr+t+1}} - k$, with $r, t, k = r(\Delta), t(\Delta), k(\Delta)$ constants; in particular, r is the constant from the previous statement, t is the R-rank of Δ and k is the one from the hypothesis.*

3. $n_1 \rightarrow (n_2)_{\Delta_{k*}}$ for any $n_1 > (cn_2)^{(2tr)^{k*}}$, with $c = c(\Delta_{k*})$ constant and r, t computed for Δ_{k*} .

Idea of proof. The more combinatorial passages of the proofs will be highly summarized, with major attention being given to the model-theoretic parts:

1. Thanks to [9](4.10(4)), we know that if for every $n \in \omega$ there is a finite set A with $|A| \geq 2$ such that $|S_\varphi^m(A)| \geq |A|^n$, then $\varphi(x, y)$ has the independence property. By remembering 1.20, we get that $\varphi(x, y)$ must be dependent, because its independence would imply instability. The negation of the previous implications gives us the first thesis.
2. Suppose that every $\varphi \in \Delta$ has exactly k free variables $x_0, \dots, x_{k-1}$ (this is done by adding dummy variables if necessary). It will be proven by induction on $m \leq k$ that there is an $u_m \subseteq n$ such that:

- $|u_{m+1}| \geq f(n)$ for $f(x) = \left\lfloor \frac{x}{t} \right\rfloor^{\frac{1}{tr+t+1}} - k$ with a specific r and t
- for every $\varphi \in \Delta$, $i_0 < \dots < i_{k-1}$ and $j_0 < \dots < j_{k-1}$, both from u_m , if $(\ell < k - m \rightarrow i_\ell = j_\ell)$ for each ℓ , then it is true that $\varphi(a_{i_0}, \dots, a_{i_{k-1}}) = \varphi(a_{j_0}, \dots, a_{j_{k-1}})$

The base case $m = 0$ is trivially verified by $u_0 = n$.

For the inductive case, let u_m be given, and define $\ell_m = |u_m|$ and $\Delta^m = \{\varphi(x_0, \dots, x_1, a_{\ell_m-m}, \dots, a_{\ell_m-1})\}_{\varphi \in \Delta}$.

Let $W_\ell \subseteq u_m \setminus \bigcup_{j < \ell} W_j$ be a set and $<_\ell$ a tree order (2.13) on the set

$W_{\leq \ell} := \bigcup_{j \leq \ell} W_j$, both inductively defined for $\ell < \ell_m$ such that:

- $i <_\ell j \rightarrow a_j, a_i$ realize the same Δ -type over the sequence $\{a_k\}_{k <_\ell i}$
- $i \not<_\ell j \not<_\ell i \rightarrow a_j, a_i$ realize two different Δ -types over the sequence $\{a_k\}_{k <_\ell i} \cap \{a_k\}_{k <_\ell j}$

A type tree is defined as $W_* = (u_m, <_*)$, where $<_* = \bigcup_{\ell} <_\ell$ and W_* satisfies the properties described above.

It can be proved that any maximal subset linearly ordered by $<_*$ (equivalently, any branch through a type tree) satisfies the inductive hypothesis on indiscernibility.

The next step is to find a lower bound on the length of any branch: this is done by assuming that a type tree with $|u_m|$ nodes is maximally branching, and then finding a certain height h that bounds said tree. Let $t = R(x = x, \Delta, 2)$ be the stability rank from 1.13: as a consequence of its definition, it is impossible to embed ${}^{t+1}2$ in W_* . For each $s \leq t$, let $S^s = \{i \in W_* : {}^s2 \text{ is embeddable in the tree above } i, \text{ but not } {}^{s+1}2\}$. Notice that $((i_1 \in S^1 \wedge i_2 \in S^2 \wedge i_1 \leq_* i_2) \rightarrow s_1 \geq s_2)$ is true and that $\{S^s\}_{s \leq t}$ is a partition of W_* : thanks to these properties, it is possible to assign a distinct value from 0 to t to any node in W_* : let it be called the $*$ -rank of the node.

Thanks to conclusion (1) of the theorem, each node in the type tree found at height h has at most $(h + m)^r$ immediate successors, where m is the inductive value and r is the constant obtained from conclusion (1). In order to have the shortest tree possible there must be maximal branching. By assuming maximal branching and through a series of combinatorial passages, the number of nodes in a tree of height h reveals itself to be bounded by $t(h + m)^{tr+t+1}$, and if h is such that $t(h + m)^{tr+t+1} \leq |u_m|$, then no type tree of height h can exhaust the nodes of u_m . By solving for h , the last condition becomes $h < \left\lceil \frac{|u_m|}{t} \right\rceil^{\frac{1}{tr+t+1}} - m$; in order to generalize, let $f(x) = \left\lceil \frac{x}{t} \right\rceil^{\frac{1}{tr+t+1}} - k$: this way, $h < f(|u_m|)$ is verified, concluding the inductive step.

Finally, u is obtained by defining $u := u_k$.

3. Let $g(x) = x^{\frac{1}{2tr}}$; by assuming $r \geq 2$, eventually $f^k(x) \geq g^k(x)$: this is a simpler (but less accurate) bound that, combined with the specific definition of Δ_{k*} , permits some final combinatorial steps, which lead directly to the last thesis.

□

Clause 3 of 2.14 will be useful to quickly obtain $n_1 \rightarrow (n_2)_{\Delta_{k_*}}$ in a future proof; this could also be done using Ramsey's theorem, but it would not explicitly convey any information about the size of the numbers, which is one of the aspects that should always be tracked in Graph Theory.

Before the next theorem, one of the most important ones of the chapter, it is useful to introduce the notation of the truth value:

Definition 2.15 (Truth value). *A truth value, written $\mathbf{t}$, is an element of $\{0, 1\}$, where 0 is identified as false and 1 as true. In this dissertation, $\mathbf{t} = \mathbf{t}(X, Y)$, where $X, Y \subseteq G$. Furthermore, the expression $xRy \equiv \mathbf{t}$ will mean $xRy \leftrightarrow \mathbf{t} = 1$.*

2.4 The Stable Regularity Lemma

Theorem 2.16. *Let G be a finite graph with the non- k_* -order property, $A \subseteq G$ such that $|A| = n$ and n_1, n_2 such that $n > n_1 n_2$ and $n_2 > (2k_*)^2$ and $n_1 \rightarrow (n_2)_{\Delta_{k_*}}$, then there are $\overline{A}, m_1, m_2$ that satisfy the following:*

1. $\overline{A} = \langle A_i : i < m_1 \rangle$ is a partition of A .
2. $n = n_2 m_1 + m_2$ and $m_2 < n_1 \leq m_1$.
3. For all i , $|A_i| \in \{n_2, n_2 + 1\}$.
4. With the possible omission of an element, each A_i is either a complete or an empty graph.
5. For every $i \neq j < m_1$, with the possible omission of an element of A_i and A_j respectively, there is a truth value $\mathbf{t}(A_i A_j) \in \{0, 1\}$ that satisfies $aRb \equiv \mathbf{t}(A_i, A_j)$ for all but at most $2k_*$ elements $a \in A_i$ and for all but at most $2k_*$ elements $b \in A_j$.

Proof. Let m_1 be such that $n_2 m_1 \leq n \leq n_2 m_1 + n_1$, from which $m_1 \geq n_1$, and let $<_*$ be a generic linear order on A ; by using the hypothesis $n_1 \rightarrow (n_2)_{\Delta_{k_*}}$, let $\{A'_\ell\}_{\ell < m_1}$ be a collection of disjoint subsets of A , each one of size n_2 , such

that, for each ℓ , the sequence $\{a_{\ell,i}\}_{i < n_2}$, listed following the increasing order obtained from $<_*$, is Δ_{k_*} -indiscernible. Claim (4) is then obtained thanks to the fact that $\{xRy\} \subset \Delta_{k_*}$ and R is symmetric. Let Λ be the union of all the A'_ℓ , from which $|\Lambda| = n_2 m_1 \leq n = |A|$; let $A \setminus \Lambda = \{a_i^*\}_{i < m_2}$ be the list of the elements of A which are not found in any A'_ℓ . Remembering that $m_2 < m_1$, for each $\ell < m_1$ let $A_\ell = A'_\ell \cup \{a_\ell^*\}$ if a_ℓ^* is well-defined, and $A_\ell = A'_\ell$ otherwise. This proves all but the last point of the theorem.

For the final passage, it may be necessary to remove the points that have been added in the previous passage; after doing that, 2.8 can be applied to any pair (A_i, A_j) , which states that either all but at most $2k_*$ points of A_i are connected to all but at most $2k_*$ points of A_j or all but at most $2k_*$ points of A_i are not connected to all but at most $2k_*$ points of A_j , from which it is possible to obtain the desired truth value $\mathbf{t}(A_i, A_j)$. $\square$

Proofs 2.14 and 2.16 quickly lead to the following regularity lemma for stable graphs:

Theorem 2.17. *Let $k_*, n_2 \in \omega$ such that $n_2 > (2k_*)^2$, then there is a $N = N(n_2, k_*)$ such that any finite graph G with $|G| > N$ and with the non- k_* -order property (and thus stable) admits a partition $\overline{G} = \langle G_i \rangle$ that satisfies the following:*

1. $|G_i| \in \{n_2, n_2 + 1\}$ for all $G_i \in \overline{G}$.
2. With the possible omission of an element, every G_i is either a complete or empty graph.
3. For any pair $G_i, G_j \in \overline{G}$, with the possible omission of an element for each, there is a truth value $\mathbf{t}(G_i, G_j) \in \{0, 1\}$ such that $aRb \equiv \mathbf{t}(G_i, G_j)$ for all but at most $2k_*$ elements $a \in G_i$ and for all but at most $2k_*$ elements $b \in G_j$.

Proof. Any graph G from the hypothesis will satisfy the conditions needed for applying 2.16 with $A = G$ as long as both the inequality chain $|G| > n_1 n_2 >$

$n_1(2k_*)^2$ and $n_1 \rightarrow (n_2)_{\Delta_{k_*}}$ hold for a certain n_1 ; in order to do this, assuming $|G| > N \geq n_1 n_2$, with $n_1 > (cn_2)^{(2tr)^{k_*}}$, where c, t, r are computed following 2.14 with the assumption that $\Delta = \Delta_{k_*}$, is sufficient, because $n_1 \rightarrow (n_2)_{\Delta_{k_*}}$ is obtained thanks to clause 3 of 2.14. $\square$

Remark. This theorem surpasses the general Szemerédi Regularity Lemma in terms of regularity, mainly thanks to the last clause, which is stronger than ε -regularity for the following reasons (all the notation used is from 2.16):

- It is valid for every possible pair, which means that there are no exceptional pairs.
- It is hereditary for any $C_i \subseteq A_i$, as long as $|C_i| \geq (2k_*)^2$ (simply reapply the theorem).
- $\frac{|\{(a, b) \in A_i \times A_j : (aRb) \equiv \neg \mathbf{t}_{i,j}\}|}{|A_i||A_j|} \leq \frac{(2k_* + 1)^2}{|A_i||A_j|} < \frac{16k_*^3}{|A_i||A_j|} \leq \frac{2k_*(|A_i| + |A_j|)}{|A_i||A_j|} = \frac{2k_*}{|A_i|} + \frac{2k_*}{|A_j|}$, which means that the density of the nodes that do not follow the truth value described in the last clause is small.
- For large enough $|A_i|$ and $|A_j|$, it is true that $\mathbf{t}_{i,j} = \mathbf{t}_{j,i}$.
- It is possible to omit exceptional points by excluding at most m_2 points from the partition.

2.5 Other Stable Regularity Lemmas

By modifying the restrictions imposed on the types of graphs that are taken in consideration, the Regularity Lemma that emerges can highly vary. Remaining in the realm of stable graphs, two examples of other possible Stable Regularity Lemmas are:

k -edge Stable Regularity Lemma

Definition 2.18 (k -edge stability). *Let $k \in \omega$; a graph G is said to be k -edge stable if there are no pairs of separated subsets of vertices $\{a_1, \dots, a_k\}$ and $\{b_1, \dots, b_k\}$ such that $a_i R b_j \leftrightarrow i < j$. In other words, the graph does not contain subgraphs that are k -half-graphs.*

Remark. The collection of k -edge stable graphs excludes more than just bipartite half-graphs.

Theorem 2.19 (k -edge Stable Regularity Lemma). *Let $k \in \omega$ and $\varepsilon > 0$, then there is a $N = N(k, \varepsilon) \in \omega$ such that any sufficiently large k -edge stable graph admits a partition $\{A_i\}_{1 \leq i \leq \ell}$, with $\ell \leq N$, such that:*

- $||A_i| - |A_j|| \leq 1$ for all i, j .
- all pairs (A_i, A_j) are ε -regular, with density either greater than $1 - \varepsilon$ or lesser than ε .
- $N < \left(\frac{4}{\varepsilon}\right)^{2^{k+3}-7}$.

Stable Excellent Regularity Lemma

Definition 2.20 (ε -good). *A set of nodes $A \subseteq G$ is said to be ε -good if for any $b \in G$ one between $|\{a \in A : bRa\}| < \varepsilon|A|$ and $|\{a \in A : \neg(bRa)\}| < \varepsilon|A|$ is true.*

Definition 2.21 (ε -excellent). *A set of nodes $A \subseteq G$ is said to be ε -excellent if for any ε -good $B \subseteq G$, one between $|\{a \in A : \mathbf{t}(a, B) = 1\}| < \varepsilon|A|$ and*

$|\{a \in A : \mathbf{t}(a, B) = 0\}| < \varepsilon|A|$ is true (where $\mathbf{t}(a, B) := \mathbf{t}(\{a\}, B)$).

Theorem 2.22 (Stable Excellent Regularity Lemma). *Let $k \geq 1$ and $\varepsilon > 0$, then there is a $N = N(k, \varepsilon)$ such that any sufficiently large k -edge stable graph admits a partition $\{A_i\}_{1 \leq i \leq \ell}$, with $\ell \leq N$, such that:*

- $||A_i| - |A_j|| \leq 1$ for all i, j .
- Every A_i is ε -excellent.
- For each pair (A_i, A_j) there is a truth value $\mathbf{t} = \mathbf{t}(A_i, A_j)$ such that $aRb \equiv \mathbf{t}$ for all but at most $\lfloor \varepsilon|A_i| \rfloor$ nodes $a \in A_i$ and all but at most $\lfloor \varepsilon|A_j| \rfloor$ nodes $b \in A_j$.

Bibliography

- [1] A. Chernikov, *Lecture Notes on Stability Theory*; URL: <https://www.math.ucla.edu/~chernikov/teaching/StabilityTheory285D/StabilityNotes.pdf>.
- [2] M. Malliaris and S. Shelah, *Regularity Lemma for Stable Graphs*; Trans. Amer. Math Soc., 366 (2014), 1551-1585.
- [3] M. Malliaris and S. Shelah, *Notes on the Stable Regularity Lemma*; arXiv:2021.09794v2 (2 Jul 2021).
- [4] J. Komlós and M. Simonovits, *Szemerédi's Regularity Lemma and its applications in graph theory*; Lecture Notes in Computer Science 2292:84-112.
- [5] J. Komlós, A. Shokoufandeh, M. Simonovits and E. Szemerédi, *The Regularity Lemma and Its Applications in Graph Theory*; Lecture Notes in Computer Science. 2292. 84-112. 10.1007/3-540-45878-6_3.
- [6] I. Kaplan and S. Shelah, *Examples in Dependent Theories*; The Journal of Symbolic Logic, 79(2), 585–619.
- [7] P. Erdős and R. Rado, *A partition calculus in set theory*; Bull. Amer. Math. Soc. 62(5): 427-489.
- [8] W. T. Gowers, *Lower bounds of tower type for Szemerédi's Uniformity Lemma*; GAFA, Geom. funct. anal. 7, 322–337 (1997).

-
- [9] S. Shelah, *Classification Theory and the number of non-isomorphic models*; rev. ed. North-Holland, 1990.
 - [10] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*; Acta Arith. 27 (1975), 199-245.
 - [11] A. Frieze and R. Kannan, *Quick approximation to matrices and application*; Combinatorica 19, 175–220 (1999).

Ringraziamenti

Un gradissimo ringraziamento al professor Martino Lupini per avermi accompagnato durante la stesura di questa tesi, per essersi mostrato sempre estremamente disponibile a chiarire i miei dubbi, e soprattutto per avermi fatto scoprire e apprezzare il vero valore dietro a questo processo di scrittura. Un abbraccio immenso a mamma, papà, Luca e i miei nonni: sono riuscito a raggiungere questo traguardo grazie a voi che, ascoltandomi sin da quando ho iniziato a fare le mie prime curiose domande, mi avete fatto capire quanto io sia avido di conoscere e quanto adori condividere ciò che imparo.