The stable regularity lemma revisited

Maryanthe Malliaris* University of Chicago

Anand Pillay† University of Notre Dame

April 23, 2015

Abstract

We prove a regularity lemma with respect to arbitrary Keisler measures μ on V, ν on W where the bipartite graph (V, W, R) is definable in a saturated structure $\bar{M}$ and the formula $R(x, y)$ is stable. The proof is rather quick, making use of local stability theory. The special case where (V, W, R) is pseudofinite, μ, ν are the counting measures, and $\bar{M}$ is suitably chosen (for example a nonstandard model of set theory), yields the stable regularity theorem of [3], though without explicit bounds or equitability.

1 Introduction

We refer to [3] for a discussion of Szemerédi's Regularity Lemma and various elaborations on it. Our context is a saturated model $\bar{M}$ of an arbitrary theory, and a definable (with parameters) bipartite graph (V, W, R) where the edge relation $R(x, y)$ is stable. Stability of R means that for some k there do not exist $a_i \in V$, $b_i \in W$ for $i \leq k$ such that $R(a_i, b_j)$ holds iff $i \leq j$. We also have global Keisler measures μ_x on V and ν_y on W (i.e. finitely additive probability measures on the Boolean algebras of definable subsets of V, W respectively). By a Δ-*formula* we mean a finite Boolean combination of things like $R(x, b)$ and $x = a$, and by a $\Delta^*(x, y)$-formula we mean a finite Boolean combination of things like $R(a, y)$ and $y = b$. We don't actually

*Partially supported by NSF grant DMS-1300634 and a Sloan fellowship

†Partially supported by NSF grant DMS-1360702

need the = formulas but it is convenient to include them. By convention the x variable is restricted to V and y variable to W so by definition a Δ-formula (Δ^*-formula) defines a subset of V (W).

Theorem 1.1. *In the context above, namely a definable graph (V, W, R) in a saturated structure $\bar{M}$, where $R(x, y)$ is stable, and Keisler measures μ on V, ν on W, we have the following: Given $\epsilon > 0$, we can partition V into finitely many definable sets $V_1, .., V_m$ each defined by a Δ-formula, and also partition W into finitely many $W_1, .., W_m$ each defined by a Δ^*-formula, such that for each V_i, W_j exactly one of the following holds:*
(i) for all $a \in V_i$ outside a set of μ measure $\leq \epsilon\mu(V_i)$, for all $b \in W_j$ outside a set of ν measure $\leq \epsilon\nu(W_j)$, we have $R(a, b)$, AND DUALLY for all $b \in W_j$ outside a set of ν measure $\leq \epsilon\nu(W_j)$, for all $a \in V_i$ outside a set of μ measure $\leq \epsilon\mu(V_i)$, we have $R(a, b)$ holds, or
(ii) for all $a \in V_i$ outside a set of μ measure $\leq \epsilon\mu(V_i)$, for all $b \in W_j$ outside a subset of ν measure $\leq \epsilon\nu(W_j)$, we have $\neg R(a, b)$, AND DUALLY for all $b \in W_j$ outside a subset of ν measure $\leq \epsilon\nu(W_j)$, for all $a \in V_i$ outside a set of μ measure $\leq \epsilon\mu(V_i)$, $\neg R(a, b)$.

Remark 1.2. *We obtain δ-regularity for suitable δ and in a suitable sense, of each of the $(V_i, W_j, R|(V_i \times W_j))$ (as essentially in Claim 5.17 of [3]): Take $\delta = \sqrt{2\epsilon}$, and suppose we are in case (i) of the conclusion of the theorem. Suppose $A \subseteq V_i$ is definable (not necessarily by a Δ-formula), and $B \subseteq W_j$ is definable (not necessarily by a Δ^*-formula), with $\mu(A) \geq \delta\mu(V_i)$ and $\nu(B) \geq \delta\nu(W_j)$. Then for all $a \in A$ outside a set of μ-measure $\leq \delta\mu(A)$, for all $b \in B$ outside a set of ν-measure $\leq \delta\nu(B)$ we have $R(a, b)$.*

Remark 1.3. *We also recover the regularity lemma for finite stable graphs in the sense of [3] by considering the counting measure on pseudofinite graphs. (But without explicit bounds or equitability of the partition.)*

Explanation. Let (V_i, W_i, R_i) be a family of finite graphs where the relation R_i does not have the k-order property for fixed k. Take an infinite ultraproduct, to obtain a definable nonstandard finite graph (V, W, R) in a (saturated) nonstandard model of set theory $\bar{M}$. Then $R(x, y)$ is stable. Let μ be the counting measure on V, namely for definable (in $\bar{M}$) $A \subseteq V$, define $\mu(A)$ to be the standard part of $|A|/|V|$, and likewise for ν on W. Then μ, ν are global Keisler measures and we can apply the theorem. Note that the definable sets in $\bar{M}$ are the internal sets in the sense of nonstandard analysis, so the standard compactness arguments give regularity in the finite situation with respect to arbitrary subsets (remembering the parenthesis).

Our model theory notation is standard. Stability theory is due to Shelah and the basic reference is [5]. For the purposes of the current paper it is convenient to refer to Chapter 1 of [4]. We will be in particular making use of "local stability theory" namely the theory of forking for complete Δ-type $p(x)$ where Δ is a finite set of stable formulas $\phi(x,y)$, and where a complete Δ-type over a model M is given (axiomatized) by a maximal consistent set of formulas of the form $\phi(x,b)$, $\neg\phi(x,b)$ for $\phi(x,y) \in \Delta$ and $b \in M$. The Δ-rank refers to $R_\Delta(-)$ as in Section 1.3 of [4], which is denoted $R(-,\Delta,\omega)$ in [5]. By a Δ-formula we mean a finite Boolean combination of formulas $\phi(x,b)$ for $\phi \in \Delta$. Recall that a formula $\phi(x,b)$ divides over A if there an A-indiscernible sequence $b = b_0, b_1, b_2...$ such that $\{\phi(x,b_i) : i < \omega\}$ is inconsistent. As in this context we will be concerned with stable formulas, we use the expression "forking" to mean dividing.

Keisler measures are of course important in this paper. If $\bar{M}$ is a saturated model and V a sort, a Keisler measure μ on V is a finitely additive probability measure on the Boolean algebra of definable (with parameters) subsets of V. If we are only concerned with subsets definable over M we call μ a Keisler measure on V over M. If Δ is a collection of L-formulas $\phi(x,y)$ (x fixed of sort V, y arbitrary), we can restrict μ to Δ formulas and call it $\mu|\Delta$. Likewise $\mu|(\Delta, M)$ is the restriction of μ to Δ-formulas over M. We will use freely facts (from [2]) such that if μ is a global Keisler measure on V say, then for any type-definable subset X of V, type-defined over a small set of parameters (equivalently any partial type Σ over a small set of parameters), $\mu(X)$ is defined and is approximated from above by the μ- measure of definable sets (formulas in Σ).

We depend somewhat on Keisler's seminal paper [2], which is very much concerned with Keisler measures $\mu|\Delta$ for Δ a finite set of stable formulas. A key observation (Lemma 1.7 of [2]), is that such $\mu|\Delta$ is a weighted sum of complete Δ-types (the proof of which is more or less repeated in Lemma 2.1 in the next section).

We will need the following, which after a suitable translation is Proposition 1.20 of [2]:

Fact 1.4. *Suppose Δ is a finite set of stable formulas and μ_x is a global Keisler measure. Then there is a small model M such that $\mu|\Delta$ does not fork over M in the sense that each Δ-formula $\psi(x)$ of positive μ-measure does* *not fork over M.*

In fact $\mu|\Delta$ will be the unique nonforking extension of $\mu|(\Delta, M)$, so $\mu|\Delta$ will be a "generically stable" measure in the sense of [1]. But all we will really need is the

Fact 1.5. *A complete Δ-type over a model M has a unique nonforking extension to a global complete Δ-type.*

As in the statement of the theorem we will take Δ to be $\{R(x,y), x = z\}$ and likewise for Δ^*.

2 Proof of Theorem 1.1

First by Fact 1.4, let M be a small model such that (V, W, R) is definable over M, and both $\mu|\Delta$, $\nu|\Delta^*$ do not fork over M. We may assume M to be reasonably saturated.

Lemma 2.1. *Given $\epsilon > 0$, we can write V as a disjoint union of Δ-formulas (restricted to V) over M, $V = V_1 \cup ... \cup V_m$, such that for each i there is a complete Δ-type p_i over M such that $V_i \in p_i$ (or $p_i \subseteq V_i$) and $\mu(V_i \setminus p_i) \leq \epsilon\mu(V_i)$.*

Proof. By induction on $R_\Delta(V)$ (which is a finite number). If it is 0, then V is finite and consists of finitely many types realized in M, so there is nothing to do. Suppose $R_\Delta(V) = n > 0$. Let $p_1, .., p_k$ be the complete Δ-types over M of $R_\Delta = n$. Let $\mu(p_i) = \alpha_i$. Then we can find $V_i \in p_i$ for $i = 1, .., k$ such that $\mu(V_i) \leq \alpha_i/(1-\epsilon)$ whereby $\mu(V_i \setminus p_i) \leq \epsilon\mu(V_i)$. We may assume that the V_i are disjoint. Let $U = V_1 \cup .. \cup V_k$. Then $V \setminus U$ has Δ-rank $< n$ so we can apply the induction hypothesis to it, to obtain the lemma for V. $\square$

We can do the same thing for W to write W as a disjoint union of Δ^*-definable (over M) sets such that for each j there is a complete Δ^*-type q_j over M such that $\nu(W_j \setminus q_j) \leq \epsilon\nu(W_j)$.

Lemma 2.2. *For each $i = 1, .., m$ and $j = 1, .., n$, we have (i) or (ii) of Theorem 1.1*

Proof. Fix i, j and we have complete Δ-type $p_i(x)$ over M and complete Δ^*-type q_j over M. Now p_i is definable, and its R-definition is given by a Δ^*-formula (1.27 of [4]). Namely there is a Δ^*-formula $\psi(y)$ over M such that for $b \in W(M)$, $R(x, b) \in p_i(x)$ iff $\models \psi(b)$. Likewise if p'_i is the unique nonforking extension of p_i to a complete global Δ-type, $\psi(y)$ is the $R(x, y)$-definition of p'_i. We have two cases:
Case (i): $\psi(y) \in q_j$.
Hence for all $b \in W_j$ other than a set a set of ν-measure $\leq \epsilon\nu(W_j)$ we have $\models \psi(b)$. Now suppose that $\models \psi(b)$, hence $R(x, b) \in p'$, whereby $p \cup \{\neg R(x, b)\}$ divides over M (by Fact 1.5) so as μ does not divide over M, $\mu(p \cup \{\neg(R(x, b)\}) = 0$, so for all $a \in V_j$ outside a set of μ-measure $\leq \epsilon\mu(V_i)$, we have $\models R(a, b)$. We have actually proved the second clause of (i) of the theorem. To obtain the first clause, let $\chi(x)$ be the Δ^*-definition of q_j, so by Lemma 2.8 of [4], and our case hypothesis, $\chi(x)$ (a Δ-formula over M) is in p_i. Continue as above.

Case (ii): $\neg\psi(y) \in q_j$. In which case we obtain as in Case (i) that for all $a \in V_i$ outside a set of μ-measure $\leq \epsilon\mu(V_i)$, for all $b \in W_j$ outside a set of ν-measure $\leq \epsilon\nu(W_j)$ we have $\neg R(a, b)$.

This completes the proof.

$\square$

Theorem 1.1 follows directly from Lemmas 2.1 and 2.2.

References

[1] E. Hrushovski, A. Pillay, P. Simon, Generically stable and smooth measures in NIP theories, 365(2013),2341-2366.

[2] H. J. Keisler, Measures and forking, Annals of Pure and Applied Logic 34 (1987), 119-169.

[3] M. Malliaris and S. Shelah, Regularity lemmas for stable graphs, Trans. Amer. Math Soc, 366 (2014), 1551-1585.

[4] A. Pillay, *Geometric Stability Theory*, OUP, 1996.

[5] S. Shelah, *Classification Theory and the number of nonisomorphic models*, North Holland, 1990.