London Mathematical Society
Lecture Note Series 409

Surveys in Combinatorics 2013

Edited by
Simon R. Blackburn, Stefanie Gerke
and Mark Wildon

CAMBRIDGE

British Combinatorial Committee

London Mathematical Society Lecture Note Series: 409

Surveys in Combinatorics 2013

Edited by

SIMON R. BLACKBURN
Royal Holloway, University of London

STEFANIE GERKE
Royal Holloway, University of London

MARK WILDON
Royal Holloway, University of London

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9781107651951

First published 2013

Printed and bound by CPI Group (UK) Ltd, Croydon CR0 4YY

A catalogue record for this publication is available from the British Library

ISBN 978-1-107-65195-1 Paperback

LONDON MATHEMATICAL SOCIETY LECTURE NOTE SERIES

Managing Editor: Professor M. Reid, Mathematics Institute,
University of Warwick, Coventry CV4 7AL, United Kingdom

The titles below are available from booksellers, or from Cambridge University Press at
www.cambridge.org/mathematics

287 Topics on Riemann surfaces and Fuchsian groups, E. BUJALANCE, A.F. COSTA & E. MARTÍNEZ (eds)
288 Surveys in combinatorics, 2001, J.W.P. HIRSCHFELD (ed)
289 Aspects of Sobolev-type inequalities, L. SALOFF-COSTE
290 Quantum groups and Lie theory, A. PRESSLEY (ed)
291 Tits buildings and the model theory of groups, K. TENT (ed)
292 A quantum groups primer, S. MAJID
293 Second order partial differential equations in Hilbert spaces, G. DA PRATO & J. ZABCZYK
294 Introduction to operator space theory, G. PISIER
295 Geometry and integrability, L. MASON & Y. NUTKU (eds)
296 Lectures on invariant theory, I. DOLGACHEV
297 The homotopy category of simply connected 4-manifolds, H.-J. BAUES
298 Higher operads, higher categories, T. LEINSTER (ed)
299 Kleinian groups and hyperbolic 3-manifolds, Y. KOMORI, V. MARKOVIC & C. SERIES (eds)
300 Introduction to Möbius differential geometry, U. HERTRICH-JEROMIN
301 Stable modules and the D(2)-problem, F.E.A. JOHNSON
302 Discrete and continuous nonlinear Schrödinger systems, M.J. ABLOWITZ, B. PRINARI & A.D. TRUBATCH
303 Number theory and algebraic geometry, M. REID & A. SKOROBOGATOV (eds)
304 Groups St Andrews 2001 in Oxford I, C.M. CAMPBELL, E.F. ROBERTSON & G.C. SMITH (eds)
305 Groups St Andrews 2001 in Oxford II, C.M. CAMPBELL, E.F. ROBERTSON & G.C. SMITH (eds)
306 Geometric mechanics and symmetry, J. MONTALDI & T. RATIU (eds)
307 Surveys in combinatorics 2003, C.D. WENSLEY (ed.)
308 Topology, geometry and quantum field theory, U.L. TILLMANN (ed)
309 Corings and comodules, T. BRZEZINSKI & R. WISBAUER
310 Topics in dynamics and ergodic theory, S. BEZUGLYI & S. KOLYADA (eds)
311 Groups: topological, combinatorial and arithmetic aspects, T.W. MÜLLER (ed)
312 Foundations of computational mathematics, Minneapolis 2002, F. CUCKER *et al* (eds)
313 Transcendental aspects of algebraic cycles, S. MÜLLER-STACH & C. PETERS (eds)
314 Spectral generalizations of line graphs, D. CVETKOVIĆ, P. ROWLINSON & S. SIMIĆ
315 Structured ring spectra, A. BAKER & B. RICHTER (eds)
316 Linear logic in computer science, T. EHRHARD, P. RUET, J.-Y. GIRARD & P. SCOTT (eds)
317 Advances in elliptic curve cryptography, I.F. BLAKE, G. SEROUSSI & N.P. SMART (eds)
318 Perturbation of the boundary in boundary-value problems of partial differential equations, D. HENRY
319 Double affine Hecke algebras, I. CHEREDNIK
320 L-functions and Galois representations, D. BURNS, K. BUZZARD & J. NEKOVÁŘ (eds)
321 Surveys in modern mathematics, V. PRASOLOV & Y. ILYASHENKO (eds)
322 Recent perspectives in random matrix theory and number theory, F. MEZZADRI & N.C. SNAITH (eds)
323 Poisson geometry, deformation quantisation and group representations, S. GUTT *et al* (eds)
324 Singularities and computer algebra, C. LOSSEN & G. PFISTER (eds)
325 Lectures on the Ricci flow, P. TOPPING
326 Modular representations of finite groups of Lie type, J.E. HUMPHREYS
327 Surveys in combinatorics 2005, B.S. WEBB (ed)
328 Fundamentals of hyperbolic manifolds, R. CANARY, D. EPSTEIN & A. MARDEN (eds)
329 Spaces of Kleinian groups, Y. MINSKY, M. SAKUMA & C. SERIES (eds)
330 Noncommutative localization in algebra and topology, A. RANICKI (ed)
331 Foundations of computational mathematics, Santander 2005, L.M PARDO, A. PINKUS, E. SÜLI & M.J. TODD (eds)
332 Handbook of tilting theory, L. ANGELERI HÜGEL, D. HAPPEL & H. KRAUSE (eds)
333 Synthetic differential geometry (2nd Edition), A. KOCK
334 The Navier–Stokes equations, N. RILEY & P. DRAZIN
335 Lectures on the combinatorics of free probability, A. NICA & R. SPEICHER
336 Integral closure of ideals, rings, and modules, I. SWANSON & C. HUNEKE
337 Methods in Banach space theory, J.M.F. CASTILLO & W.B. JOHNSON (eds)
338 Surveys in geometry and number theory, N. YOUNG (ed)
339 Groups St Andrews 2005 I, C.M. CAMPBELL, M.R. QUICK, E.F. ROBERTSON & G.C. SMITH (eds)
340 Groups St Andrews 2005 II, C.M. CAMPBELL, M.R. QUICK, E.F. ROBERTSON & G.C. SMITH (eds)
341 Ranks of elliptic curves and random matrix theory, J.B. CONREY, D.W. FARMER, F. MEZZADRI & N.C. SNAITH (eds)
342 Elliptic cohomology, H.R. MILLER & D.C. RAVENEL (eds)
343 Algebraic cycles and motives I, J. NAGEL & C. PETERS (eds)
344 Algebraic cycles and motives II, J. NAGEL & C. PETERS (eds)
345 Algebraic and analytic geometry, A. NEEMAN
346 Surveys in combinatorics 2007, A. HILTON & J. TALBOT (eds)
347 Surveys in contemporary mathematics, N. YOUNG & Y. CHOI (eds)

348 Transcendental dynamics and complex analysis, P.J. RIPPON & G.M. STALLARD (eds)
349 Model theory with applications to algebra and analysis I, Z. CHATZIDAKIS, D. MACPHERSON, A. PILLAY & A. WILKIE (eds)
350 Model theory with applications to algebra and analysis II, Z. CHATZIDAKIS, D. MACPHERSON, A. PILLAY & A. WILKIE (eds)
351 Finite von Neumann algebras and masas, A.M. SINCLAIR & R.R. SMITH
352 Number theory and polynomials, J. MCKEE & C. SMYTH (eds)
353 Trends in stochastic analysis, J. BLATH, P. MÖRTERS & M. SCHEUTZOW (eds)
354 Groups and analysis, K. TENT (ed)
355 Non-equilibrium statistical mechanics and turbulence, J. CARDY, G. FALKOVICH & K. GAWEDZKI
356 Elliptic curves and big Galois representations, D. DELBOURGO
357 Algebraic theory of differential equations, M.A.H. MACCALLUM & A.V. MIKHAILOV (eds)
358 Geometric and cohomological methods in group theory, M.R. BRIDSON, P.H. KROPHOLLER & I.J. LEARY (eds)
359 Moduli spaces and vector bundles, L. BRAMBILA-PAZ, S.B. BRADLOW, O. GARCÍA-PRADA & S. RAMANAN (eds)
360 Zariski geometries, B. ZILBER
361 Words: Notes on verbal width in groups, D. SEGAL
362 Differential tensor algebras and their module categories, R. BAUTISTA, L. SALMERÓN & R. ZUAZUA
363 Foundations of computational mathematics, Hong Kong 2008, F. CUCKER, A. PINKUS & M.J. TODD (eds)
364 Partial differential equations and fluid mechanics, J.C. ROBINSON & J.L. RODRIGO (eds)
365 Surveys in combinatorics 2009, S. HUCZYNSKA, J.D. MITCHELL & C.M. RONEY-DOUGAL (eds)
366 Highly oscillatory problems, B. ENGQUIST, A. FOKAS, E. HAIRER & A. ISERLES (eds)
367 Random matrices: High dimensional phenomena, G. BLOWER
368 Geometry of Riemann surfaces, F.P. GARDINER, G. GONZÁLEZ-DIEZ & C. KOUROUNIOTIS (eds)
369 Epidemics and rumours in complex networks, M. DRAIEF & L. MASSOULIÉ
370 Theory of p-adic distributions, S. ALBEVERIO, A.YU. KHRENNIKOV & V.M. SHELKOVICH
371 Conformal fractals, F. PRZYTYCKI & M. URBAŃSKI
372 Moonshine: The first quarter century and beyond, J. LEPOWSKY, J. MCKAY & M.P. TUITE (eds)
373 Smoothness, regularity and complete intersection, J. MAJADAS & A. G. RODICIO
374 Geometric analysis of hyperbolic differential equations: An introduction, S. ALINHAC
375 Triangulated categories, T. HOLM, P. JØRGENSEN & R. ROUQUIER (eds)
376 Permutation patterns, S. LINTON, N. RUŠKUC & V. VATTER (eds)
377 An introduction to Galois cohomology and its applications, G. BERHUY
378 Probability and mathematical genetics, N. H. BINGHAM & C. M. GOLDIE (eds)
379 Finite and algorithmic model theory, J. ESPARZA, C. MICHAUX & C. STEINHORN (eds)
380 Real and complex singularities, M. MANOEL, M.C. ROMERO FUSTER & C.T.C WALL (eds)
381 Symmetries and integrability of difference equations, D. LEVI, P. OLVER, Z. THOMOVA & P. WINTERNITZ (eds)
382 Forcing with random variables and proof complexity, J. KRAJÍČEK
383 Motivic integration and its interactions with model theory and non-Archimedean geometry I, R. CLUCKERS, J. NICAISE & J. SEBAG (eds)
384 Motivic integration and its interactions with model theory and non-Archimedean geometry II, R. CLUCKERS, J. NICAISE & J. SEBAG (eds)
385 Entropy of hidden Markov processes and connections to dynamical systems, B. MARCUS, K. PETERSEN & T. WEISSMAN (eds)
386 Independence-friendly logic, A.L. MANN, G. SANDU & M. SEVENSTER
387 Groups St Andrews 2009 in Bath I, C.M. CAMPBELL *et al* (eds)
388 Groups St Andrews 2009 in Bath II, C.M. CAMPBELL *et al* (eds)
389 Random fields on the sphere, D. MARINUCCI & G. PECCATI
390 Localization in periodic potentials, D.E. PELINOVSKY
391 Fusion systems in algebra and topology, M. ASCHBACHER, R. KESSAR & B. OLIVER
392 Surveys in combinatorics 2011, R. CHAPMAN (ed)
393 Non-abelian fundamental groups and Iwasawa theory, J. COATES *et al* (eds)
394 Variational problems in differential geometry, R. BIELAWSKI, K. HOUSTON & M. SPEIGHT (eds)
395 How groups grow, A. MANN
396 Arithmetic dfferential operators over the p-adic Integers, C.C. RALPH & S.R. SIMANCA
397 Hyperbolic geometry and applications in quantum Chaos and cosmology, J. BOLTE & F. STEINER (eds)
398 Mathematical models in contact mechanics, M. SOFONEA & A. MATEI
399 Circuit double cover of graphs, C.-Q. ZHANG
400 Dense sphere packings: a blueprint for formal proofs, T. HALES
401 A double Hall algebra approach to affine quantum Schur-Weyl theory, B. DENG, J. DU & Q. FU
402 Mathematical aspects of fluid mechanics, J. ROBINSON, J.L. RODRIGO & W. SADOWSKI (eds)
403 Foundations of computational mathematics: Budapest 2011, F. CUCKER, T. KRICK, A. SZANTO & A. PINKUS (eds)
404 Operator methods for boundary value problems, S. HASSI, H.S.V. DE SNOO & F.H. SZAFRANIEC (eds)
405 Torsors, étale homotopy and applications to rational points, A.N. SKOROBOGATOV (ed)
406 Appalachian set theory, J. CUMMINGS & E. SCHIMMERLING (eds)
407 The maximal subgroups of the low-dimensional finite classical groups, J.N. BRAY, D.F. HOLT & C.M. RONEY-DOUGAL

Preface

The Twenty-Fourth British Combinatorial Conference was organised by Royal Holloway, University of London. It was held in Egham, Surrey in July 2013. The British Combinatorial Committee had invited nine distinguished combinatorialists to give survey lectures in areas of their expertise, and this volume contains the survey articles on which these lectures were based.

In compiling this volume we are indebted to the authors for preparing their articles so accurately and professionally, and to the referees for their rapid responses and keen eye for detail. We would also like to thank Roger Astley and Sam Harrison at Cambridge University Press for their advice and assistance.

Finally, without the previous efforts of editors of earlier *Surveys* and the guidance of the British Combinatorial Committee, the preparation of this volume would have been daunting: we would like to express our thanks for their support.

Simon R. Blackburn, Stefanie Gerke and Mark Wildon

Royal Holloway, University of London

January 2013

Contents

Graph removal lemmas

David Conlon[1] and Jacob Fox[2]

Abstract

The graph removal lemma states that any graph on n vertices with $o(n^h)$ copies of a fixed graph H on h vertices may be made H-free by removing $o(n^2)$ edges. Despite its innocent appearance, this lemma and its extensions have several important consequences in number theory, discrete geometry, graph theory and computer science. In this survey we discuss these lemmas, focusing in particular on recent improvements to their quantitative aspects.

1 Introduction

The triangle removal lemma states that for every $\varepsilon > 0$ there exists $\delta > 0$ such that any graph on n vertices with at most δn^3 triangles may be made triangle-free by removing at most εn^2 edges. This result, proved by Ruzsa and Szemerédi [94] in 1976, was originally stated in rather different language.

The original formulation was in terms of the $(6, 3)$-problem.[3] This asks for the maximum number of edges $f^{(3)}(n, 6, 3)$ in a 3-uniform hypergraph on n vertices such that no 6 vertices contain 3 edges. Answering a question of Brown, Erdős and Sós [19], Ruzsa and Szemerédi showed that $f^{(3)}(n, 6, 3) = o(n^2)$. Their proof used several iterations of an early version of Szemerédi's regularity lemma [111].

This result, developed by Szemerédi in his proof of the Erdős-Turán conjecture on arithmetic progressions in dense sets [110], states that every graph may be partitioned into a small number of vertex sets so that the graph between almost every pair of vertex sets is random-like. Though this result now occupies a central position in graph theory, its importance only emerged over time. The resolution of the $(6, 3)$-problem was one of the first indications of its strength.

---

[1]Supported by a Royal Society University Research Fellowship.

[2]Supported by a Simons Fellowship and NSF Grant DMS-1069197.

[3]The two results are not exactly equivalent, though the triangle removal lemma may be proved by their method. A weak form of the triangle removal lemma, already sufficient for proving Roth's theorem, is equivalent to the Ruzsa-Szemerédi theorem. This weaker form states that any graph on n vertices in which every edge is contained in exactly one triangle has $o(n^2)$ edges. This is also equivalent to another attractive formulation, known as the induced matching theorem. This states that any graph on n vertices which is the union of at most n induced matchings has $o(n^2)$ edges.

The Ruzsa-Szemerédi theorem was generalized by Erdős, Frankl and Rödl [32], who showed that $f^{(r)}(n, 3r-3, 3) = o(n^2)$, where $f^{(r)}(n, 3r-3, 3)$ is the maximum number of edges in an r-uniform hypergraph such that no $3r-3$ vertices contain 3 edges. One of the tools used by Erdős, Frankl and Rödl in their proof was a striking result stating that if a graph on n vertices contains no copy of a graph H then it may be made K_r-free, where $r = \chi(H)$ is the chromatic number of H, by removing $o(n^2)$ edges. The proof of this result used the modern formulation of Szemerédi's regularity lemma and is already very close, both in proof and statement, to the following generalization of the triangle removal lemma, known as the graph removal lemma.[4] This was first stated explicitly in the literature by Alon, Duke, Lefmann, Rödl and Yuster [4] and by Füredi [47] in 1994.[5]

Theorem 1.1 *For any graph H on h vertices and any $\varepsilon > 0$, there exists $\delta > 0$ such that any graph on n vertices which contains at most δn^h copies of H may be made H-free by removing at most εn^2 edges.*

It was already observed by Ruzsa and Szemerédi that the $(6,3)$-problem (and, thereby, the triangle removal lemma) is related to Roth's theorem on arithmetic progressions [92]. This theorem states that for any $\delta > 0$ there exists an n_0 such that if $n \geq n_0$, then any subset of the set $[n] := \{1, 2, \dots, n\}$ of size at least δn contains an arithmetic progression of length 3. Letting $r_3(n)$ be the largest integer such that there exists a subset of the set $\{1, 2, \dots, n\}$ of size $r_3(n)$ containing no arithmetic progression of length 3, this is equivalent to saying that $r_3(n) = o(n)$. Ruzsa and Szemerédi observed that $f^{(3)}(n, 6, 3) = \Omega(r_3(n)n)$. In particular, since $f^{(3)}(n, 6, 3) = o(n^2)$, this implies that $r_3(n) = o(n)$, yielding a proof of Roth's theorem.

It was further noted by Solymosi [105] that the Ruzsa-Szemerédi theorem yields a stronger result of Ajtai and Szemerédi [1]. This result states that for any $\delta > 0$ there exists an n_0 such that if $n \geq n_0$ then any subset of the set $[n] \times [n]$ of size at least δn^2 contains a set of the form $\{(a, b), (a + d, b), (a, b + d)\}$ with $d > 0$. That is, dense subsets of the

[4]The phrase 'removal lemma' is a comparatively recent coinage. It seems to have come into vogue in about 2005 when the hypergraph removal lemma was first proved (see, for example, [68, 79, 107, 113]).

[5]This was also the first time that the triangle removal lemma was stated explicitly, though the weaker version concerning graphs where every edge is contained in exactly one triangle had already appeared in the literature. The Ruzsa-Szemerédi theorem was usually [40, 41, 46] phrased in the following suggestive form: if a 3-uniform hypergraph is linear, that is, no two edges intersect on more than a single vertex, and triangle-free, then it has $o(n^2)$ edges. A more explicit formulation may be found in [23].

2-dimensional grid contain axis-parallel isosceles triangles. Roth's theorem is a simple corollary of this statement.

Roth's theorem is the first case of a famous result known as Szemerédi's theorem. This result, to which we alluded earlier, states that for any natural number $k \geq 3$ and any $\delta > 0$ there exists n_0 such that if $n \geq n_0$ then any subset of the set $[n]$ of size at least δn contains an arithmetic progression of length k. This was first proved by Szemerédi [110] in the early seventies using combinatorial techniques and since then several further proofs have emerged. The most important of these are that by Furstenberg [48, 50] using ergodic theory and that by Gowers [54, 55], who found a way to extend Roth's original Fourier analytic argument to general k. Both of these methods have been highly influential.

Yet another proof technique was suggested by Frankl and Rödl [42]. They showed that Szemerédi's theorem would follow from the following generalization of Theorem 1.1, referred to as the hypergraph removal lemma. They proved this theorem for the specific case of $K_4^{(3)}$, the complete 3-uniform hypergraph with 4 vertices. This was then extended to all 3-uniform hypergraphs in [78] and to $K_5^{(4)}$ in [90]. Finally, it was proved for all hypergraphs by Gowers [56, 57] and, independently, by Nagle, Rödl, Schacht and Skokan [79, 89]. Both proofs rely on extending Szemerédi's regularity lemma to hypergraphs in an appropriate fashion.

Theorem 1.2 *For any k-uniform hypergraph $\mathcal{H}$ on h vertices and any $\varepsilon > 0$, there exists $\delta > 0$ such that any k-uniform hypergraph on n vertices which contains at most δn^h copies of $\mathcal{H}$ may be made $\mathcal{H}$-free by removing at most εn^k edges.*

As well as reproving Szemerédi's theorem, the hypergraph removal lemma allows one to reprove the multidimensional Szemerédi theorem. This theorem, originally proved by Furstenberg and Katznelson [49], states that for any natural number r, any finite subset S of $\mathbb{Z}^r$ and any $\delta > 0$ there exists n_0 such that if $n \geq n_0$ then any subset of $[n]^r$ of size at least δn^r contains a subset of the form $a \cdot S + d$ with $a > 0$, that is, a dilated and translated copy of S. That it follows from the hypergraph removal lemma was first observed by Solymosi [106]. This was the first non-ergodic proof of this theorem. A new proof of the special case $S = \{(0,0),(1,0),(0,1)\}$, corresponding to the Ajtai-Szemerédi theorem, was given by Shkredov [103] using a Fourier analytic argument. Recently, a combinatorial proof of the density Hales-Jewett theorem, which is an extension of the multidimensional Szemerédi theorem, was discovered as part of the polymath project [82].

As well as its implications in number theory, the removal lemma and its extensions are central to the area of computer science known as property testing. In this area, one would like to find fast algorithms to distinguish between objects which satisfy a certain property and objects which are far from satisfying that property. This field of study was initiated by Rubinfield and Sudan [93] and, subsequently, Goldreich, Goldwasser and Ron [52] started the investigation of such property testers for combinatorial objects. Graph property testing has attracted a particular degree of interest.

A classic example of property testing is to decide whether a given graph G is ε-far from being triangle-free, that is, whether at least εn^2 edges will have to removed in order to make it triangle-free. The triangle removal lemma tells us that if G is ε-far from being triangle free then it must contain at least δn^3 triangles for some $\delta > 0$ depending only on ε. This furnishes a simple probabilistic algorithm for deciding whether G is ε-far from being triangle-free. We choose $t = 2\delta^{-1}$ triples of points from the vertices of G uniformly at random. If G is ε-far from being triangle-free then the probability that none of these randomly chosen triples is a triangle is $(1-\delta)^t < e^{-t\delta} < \frac{1}{3}$. That is, if G is ε-far from being triangle-free, we will find a triangle with probability at least $\frac{2}{3}$, whereas if G is triangle-free, we will clearly find no triangles. The graph removal lemma may be used to derive a similar test for deciding whether G is ε-far from being H-free for any fixed graph H.

In property testing, it is often of interest to decide not only whether a graph is far from being H-free but also whether it is far from being induced H-free. A subgraph H' of a graph G is said to be an induced copy of H if there is a one-to-one map $f : V(H) \to V(H')$ such that $(f(u), f(v))$ is an edge of H' if and only if (u, v) is an edge of H. A graph G is said to be induced H-free if it contains no induced copies of H and ε-far from being induced H-free if we have to add and/or delete at least εn^2 edges to make it induced H-free. Note that it is not enough to delete edges since, for example, if H is the empty graph on two vertices and G is the complete graph minus an edge, then G contains only one induced copy of H, but one cannot simply delete edges from G to make it induced H-free.

By proving an appropriate strengthening of the regularity lemma, Alon, Fischer, Krivelevich and Szegedy [6] showed how to modify the graph removal lemma to this setting. This result, which allows one to test for induced H-freeness, is known as the induced removal lemma.

Theorem 1.3 *For any graph H on h vertices and any $\varepsilon > 0$, there exists a $\delta > 0$ such that any graph on n vertices which contains at most δn^h induced copies of H may be made induced H-free by adding and/or deleting at most εn^2 edges.*

A substantial generalization of this result, known as the infinite removal lemma, was proved by Alon and Shapira [12] (see also [76]). They showed that for each (possibly infinite) family $\mathcal{H}$ of graphs and $\varepsilon > 0$ there is $\delta = \delta_{\mathcal{H}}(\varepsilon) > 0$ and $t = t_{\mathcal{H}}(\varepsilon)$ such that if a graph G on n vertices contains at most δn^h induced copies of H for every graph H in $\mathcal{H}$ on $h \leq t$ vertices, then G may be made induced H-free, for every $H \in \mathcal{H}$, by adding and/or deleting at most εn^2 edges. They then used this result to show that every hereditary graph property is testable, where a graph property is hereditary if it is closed under removal of vertices. These results were extended to 3-uniform hypergraphs by Avart, Rödl and Schacht [14] and to k-uniform hypergraphs by Rödl and Schacht [87].

In this survey we will focus on recent developments, particularly with regard to the quantitative aspects of the removal lemma. In particular, we will discuss recent improvements on the bounds for the graph removal lemma, Theorem 1.1, and the induced graph removal lemma, Theorem 1.3, each of which bypasses a natural impediment.

The usual proof of the graph removal lemma makes use of the regularity lemma and gives bounds for the removal lemma which are of tower-type in ε. To be more specific, let $T(1) = 2$ and, for each $i \geq 1$, $T(i+1) = 2^{T(i)}$. The bounds that come out of applying the regularity lemma to the removal lemma then say that if $\delta^{-1} = T(\varepsilon^{-c_H})$, then any graph on n vertices with at most δn^h copies of a graph H on h vertices may be made H-free by removing at most εn^2 edges. Moreover, this tower-type dependency is inherent in any proof employing regularity. This follows from an important result of Gowers [53] (see also [24]) which states that the bounds that arise in the regularity lemma are necessarily of tower type. We will discuss this in more detail in Section 2.1 below.

Despite this obstacle, the following improvement was made by Fox [38].

Theorem 1.4 *For any graph H on h vertices, there exists a constant a_H such that if $\delta^{-1} = T(a_H \log \varepsilon^{-1})$ then any graph on n vertices which contains at most δn^h copies of H may be made H-free by removing at most εn^2 edges.*

As is implicit in the bounds, the proof of this theorem does not make an explicit appeal to Szemerédi's regularity lemma. However, many of the ideas used are similar to ideas used in the proof of the regularity lemma. The chief difference lies in the fact that the conditions of the removal lemma (containing few copies of a given graph H) allow us to say more about the structure of these partitions. A simplified proof of this theorem will be the main topic of Section 2.2.

Though still of tower-type, Theorem 1.4 improves substantially on the previous bound. However, it remains very far from the best known lower bound on δ^{-1}. The observation of Ruzsa and Szemerédi [94] that $f^{(3)}(n,6,3) = \Omega(r_3(n)n)$ allows one to transfer lower bounds for $r_3(n)$ to a corresponding lower bound for the triangle removal lemma. The best construction of a set containing no arithmetic progression of length 3 is due to Behrend [16] and gives a subset of $[n]$ with density $e^{-c\sqrt{\log n}}$. Transferring this to the graph setting yields a graph containing $\varepsilon^{c \log \varepsilon^{-1}} n^3$ triangles which cannot be made triangle-free by removing fewer than εn^2 edges. This quasi-polynomial lower bound, $\delta^{-1} \geq \varepsilon^{-c \log \varepsilon^{-1}}$, remains the best known.[6]

The standard proof of the induced removal lemma uses the strong regularity lemma of Alon, Fischer, Krivelevich and Szegedy [6]. We will speak at length about this result in Section 3.1. Here it will suffice to say that, like the ordinary regularity lemma, the bounds which an application of this theorem gives for the induced removal lemma are necessarily very large. Let $W(1) = 2$ and, for $i \geq 1$, $W(i+1) = T(W(i))$. This is known as the wowzer function and its values dwarf those of the usual tower function.[7] By using the strong regularity lemma, the standard proof shows that we may take $\delta^{-1} = W(a_H \varepsilon^{-c})$ in the induced removal lemma, Theorem 1.3. Moreover, as with the ordinary removal lemma, such a bound is inherent in the application of the strong regularity lemma. This follows from recent results of Conlon and Fox [24] and, independently, Kalyanasundaram and Shapira [62] showing that the bounds arising in strong regularity are necessarily of wowzer type.

In the other direction, Conlon and Fox [24] showed how to bypass this obstacle and prove that the bounds for δ^{-1} are at worst a tower in a power of ε^{-1}.

Theorem 1.5 *There exists a constant $c > 0$ such that, for any graph H on h vertices, there exists a constant a_H such that if $\delta^{-1} = T(a_H \varepsilon^{-c})$ then any graph on n vertices which contains at most δn^h induced copies of H may be made induced H-free by adding and/or deleting at most εn^2 edges.*

---

[6]It is worth noting that the best known upper bound for Roth's theorem, due to Sanders [96], is considerably better than the best upper bound for $r_3(n)$ that follows from triangle removal. This upper bound is $r_3(n) = O\left(\frac{(\log \log n)^5}{\log n} n\right)$. A recent result of Schoen and Shkredov [100], building on further work of Sanders [97], shows that any subset of $[n]$ of density $e^{-c(\frac{\log n}{\log \log n})^{1/6}}$ contains a solution to the equation $x_1 + \cdots + x_5 = 5x_6$. Since arithmetic progressions correspond to solutions of $x_1 + x_2 = 2x_3$, this suggests that the answer should be closer to the Behrend bound. The bounds for triangle removal are unlikely to impinge on these upper bounds for some time, if at all.

[7]To give some indication, we note that $W(2) = 4$, $W(3) = 65536$ and $W(4)$ is a tower of 2s of height 65536.

A discussion of this theorem will form the subject of Section 3.2. The key observation here is that the strong regularity lemma is used to prove an intermediate statement (Lemma 3.2 below) which then implies the induced removal lemma. This intermediate statement may be proved without recourse to the full strength of the strong regularity lemma. There are also some strong parallels with the proof of Theorem 1.4 which we will draw attention to in due course.

In Section 3.3, we present the proof of Alon and Shapira's infinite removal lemma. In another paper, Alon and Shapira [11] showed that the dependence in the infinite removal lemma can depend heavily on the family $\mathcal{H}$. They proved that for every function $\delta : (0,1) \to (0,1)$, there exists a family $\mathcal{H}$ of graphs such that any $\delta_{\mathcal{H}} : (0,1) \to (0,1)$ which satisfies the infinite removal lemma for $\mathcal{H}$ satisfies $\delta_{\mathcal{H}} = o(\delta)$. However, such examples are rather unusual and the proof presented in Section 3.3 of the infinite removal lemma implies that for many commonly studied families $\mathcal{H}$ of graphs the bound on $\delta_{\mathcal{H}}^{-1}$ is only tower-type, improving the wowzer-type bound from the original proof.

Our discussions of the graph removal lemma and the induced removal lemma will occupy the bulk of this survey but we will also talk about some further recent developments in the study of removal lemmas. These include arithmetic removal lemmas (Section 4) and the recently developed sparse removal lemmas which hold for subgraphs of sparse random and pseudorandom graphs (Section 5). We will conclude with some further comments on related topics.

2 The graph removal lemma

In this section we will discuss the two proofs of the removal lemma, Theorem 1.1, at length. In Section 2.1, we will go through the regularity lemma and the usual proof of the removal lemma. Then, in Section 2.2, we will consider a simplified variant of the second author's recent proof [38], showing how it connects to the weak regularity lemma of Frieze and Kannan [44, 45].

2.1 The standard proof

We begin with the proof of the regularity lemma and then deduce the removal lemma. For vertex subsets S, T of a graph G, we let $e_G(S,T)$ denote the number of pairs in $S \times T$ that are edges of G and $d_G(S,T) = \frac{e_G(S,T)}{|S||T|}$ denote the fraction of pairs in $S \times T$ that are edges of G. For simplicity of notation, we drop the subscript if the graph G is clear from context. Although non-standard, it will be convenient to define the *edge density* of a graph $G = (V,E)$ to be $d(G) = d(V,V) = \frac{2e(G)}{|V|^2}$, which is the fraction of all ordered pairs of (not necessarily distinct) vertices which are edges. A pair (S, T) of subsets is ε*-regular* if, for all subsets $S' \subset S$ and $T' \subset T$ with $|S'| \geq \varepsilon|S|$ and $|T'| \geq \varepsilon|T|$, we have $|d(S', T') - d(S, T)| \leq \varepsilon$. Informally, a pair of subsets is ε-regular with a small ε if the edges between S and T are uniformly distributed among large subsets.

Let $G = (V, E)$ be a graph and $P : V = V_1 \cup \ldots \cup V_k$ be a vertex partition of G. The partition of P is *equitable* if each pair of parts differ in size by at most 1. The partition P is ε*-regular* if all but at most εk^2 pairs of parts (V_i, V_j) are ε-regular. Note that we are considering all k^2 ordered pairs (V_i, V_j), including those with $i = j$. We next state Szemerédi's regularity lemma [111].

Lemma 2.1 *For every $\varepsilon > 0$, there is $K = K(\varepsilon)$ such that every graph $G = (V, E)$ has an equitable, ε-regular vertex partition into at most K parts. Moreover, we may take K to be a tower of height $O(\varepsilon^{-5})$.*

Let $q : [0, 1] \to \mathbb{R}$ be a convex function. For vertex subsets $S, T \subset V$ of a graph G, let $q(S, T) = q(d(S, T))|S||T|/|V|^2$. For partitions $\mathcal{S} : S = S_1 \cup \ldots \cup S_a$ and $\mathcal{T} : T = T_1 \cup \ldots \cup T_b$, let $q(\mathcal{S}, \mathcal{T}) = \sum_{1 \leq i \leq a, 1 \leq j \leq b} q(S_i, T_j)$. For a vertex partition $P : V = V_1 \cup \ldots \cup V_k$ of G, define the mean-q density to be

$$q(P) = q(P, P) = \sum_{1 \leq i, j \leq k} q(V_i, V_j).$$

We next state some simple properties which follow from Jensen's inequality using the convexity of q. A *refinement* of a partition P of a vertex set V is another partition Q of V such that every part of Q is a subset of a part of P.

Proposition 2.2 *1. For partitions $\mathcal{S}$ and $\mathcal{T}$ of vertex subsets S and T, we have $q(\mathcal{S}, \mathcal{T}) \geq q(S, T)$.*

2. If Q is a refinement of P, then $q(Q) \geq q(P)$.

3. If $d = d(G) = d(V, V)$ is the edge density of G, then, for any vertex partition P,

$$q(d) \leq q(P) \leq dq(1) + (1 - d)q(0).$$

The first and second part of Proposition 2.2 show that by refining a vertex partition the mean-q density cannot decrease, while the last part gives the range of possible values for $q(P)$ if we only know the edge density d of G.

The convex function $q(x) = x^2$ for $x \in [0, 1]$ is chosen in the standard proof of the graph regularity lemma and we will do the same for the rest