CORE
Metadata, citation and similar papers at core.ac.uk
Provided by University of Birmingham Research Archive, E-theses Repository

ON CYCLES IN DIRECTED GRAPHS

by

LUKE TRISTAN KELLY

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Mathematics

The University of Birmingham

May 2010

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

The main results of this thesis are the following. We show that for each $\alpha > 0$ every sufficiently large oriented graph G with minimum indegree and minimum outdegree at least $3|G|/8 + \alpha|G|$ contains a Hamilton cycle. This gives an approximate solution to a problem of Thomassen. Furthermore, answering completely a conjecture of Häggkvist and Thomason, we show that we get every possible orientation of a Hamilton cycle.

We also deal extensively with short cycles, showing that for each $\ell \geq 4$ every sufficiently large oriented graph G with minimum indegree and minimum outdegree at least $\geq |G|/3+1$ contains an ℓ-cycle. This is best possible for all those $\ell \geq 4$ which are not divisible by 3. Surprisingly, for some other values of ℓ, an ℓ-cycle is forced by a much weaker minimum degree condition. We propose and discuss a conjecture regarding the precise minimum degree which forces an ℓ-cycle (with $\ell \geq 4$ divisible by 3) in an oriented graph.

We also give an application of our results to pancyclicity.

This thesis is dedicated to Rachel.

I can live with doubt and uncertainty and not knowing. I think it's much more interesting to live not knowing than to have answers which might be wrong. I have approximate answers and possible beliefs and different degrees of certainty about different things, but I'm not absolutely sure of anything and there are many things I don't know anything about, such as whether it means anything to ask why we're here. I don't have to know the answer. I don't feel frightened by not knowing things, by being lost in a mysterious universe without any purpose, which is the way it really is as far as I can tell. It doesn't frighten me.

The Pleasure of Finding Things Out

Richard Feynman, 1983

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors Deryk Osthus and Daniela Kühn for all the help they have given me over the last three years, including, but far from limited to, their coauthorship of many of the results contained within.

I would like to acknowledge those who shaped my earlier mathematical education: Laurent Storoni and Mark Hennings at Sidney Sussex College and Aydin Önaç at the Chase High School.

Thanks are due to Peter Keevash who pointed out a simplification of one of the main lemmas and Oliver Cooley, Richard Mycroft and Andrew Treglown for their comments on drafts, conversations on ideas and hours in pubs.

Finally I would like to thank my family for the support they have provided throughout not just the last years but my entire life. Most importantly though, I must acknowledge Rachel, without whose endless encouragement, patience and love I could not have completed this thesis.

To each of the above, I extend my deepest appreciation.

TABLE OF CONTENTS

CHAPTER 1

PREFACE

The study of cycles, both Hamilton and short, is one of the most important and most studied areas of graph theory. There are many papers published every year seeking more sufficient conditions for a graph to contain a Hamilton cycle, looking at the behaviour of Hamilton cycles in various models of random graphs and examining refinements of the idea of Hamiltonicity. The Caccetta-Häggkvist conjecture has inspired years of research into sufficient conditions for short cycles in digraphs.

A simple graph $G = (V(G), E(G))$ is a set of *vertices* $V(G)$ with a set of edges $E(G) \subseteq V^{(2)}$. The number of vertices in a graph is called its *order* and is often denoted $|G|$. The number of edges in a graph is denoted by $e(G)$. A cycle (or k-cycle) is a sequence of distinct vertices $x_1, x_2, \ldots, x_k$ where $x_1x_2, \ldots, x_{k-1}x_k, x_kx_1 \in E(G)$. A *Hamilton cycle* is a cycle containing every vertex. The *degree* $d(x)$ of a vertex $x \in V(G)$ is the number of vertices sharing an edge with x. The *minimum degree* $\delta(G)$ of a graph is the minimum of the degrees of the vertices of G.

A *digraph* or *directed graph* is a graph in which all the edges are assigned a direction and there are no multiple edges of the same direction. I.e. we allow an edge in each direction between two vertices, but no other multiple edges are allowed. An *oriented graph* is a (simple) graph in which every edge is assigned a direction. Equivalently, an oriented graph is a digraph with no multiple edges.

In the following four sections we summarise the main results of this thesis. In the corresponding chapters there are introductions giving more details on these results, the history of the area and related results. Finally in Chapter 8 we summarise the open problems coming from this work. All of the work detailed in Chapters 4 and 5 is joint work with Deryk Osthus and Daniela Kühn.

1.1 Hamilton Cycles in Oriented Graphs

A fundamental result of Dirac states that a minimum degree of $|G|/2$ guarantees a Hamilton cycle in an undirected graph G on at least 3 vertices. Ore in 1960 gave a stronger sufficient condition: if the sum of the degrees of every pair of non-adjacent vertices is at least $|G|$, then the graph is Hamiltonian [66].

There is an obvious analogue of a Hamilton cycle for digraphs. That is, an ordering $x_1, \ldots, x_n$ of the vertices of a digraph D such that $x_i x_{i+1}$ is a directed edge for all i. A *cycle* (or k-cycle) is a sequence of distinct vertices $x_1, x_2, \ldots, x_k$ where at least one of the ordered pairs $x_i x_{i+1}$ and $x_{i+1} x_i$ is in $E(G)$ (for all $1 \leq i \leq k$, counting module k). It is a *directed cycle* if $x_i x_{i+1} \in E(G)$ for all i. We will use the convention that a cycle in an oriented graph or digraph is directed unless otherwise stated. The *minimum semi-degree* $\delta^0(G)$ of an oriented graph G (or of a digraph) is the minimum of its minimum outdegree $\delta^+(G)$ and its minimum indegree $\delta^-(G)$. See Chapter 2 for more precise definitions. There are corresponding versions of the famous theorems of Dirac and Ore for digraphs. Ghouila-Houri [34] proved in 1960 that every digraph D with minimum semi-degree at least $|D|/2$ contains a Hamilton cycle. Meyniel [60] showed that an analogue of Ore's theorem holds for digraphs, that is a digraph on at least 4 vertices is either Hamiltonian or the sum of the indegrees and outdegrees of a pair of non-adjacent vertices is less than $2|D| - 1$. All these bounds are best possible.

It is natural to ask for the (smallest) minimum semi-degree which guarantees a Hamilton cycle in an oriented graph G. This question was first raised by Thomassen [73], who [75] showed that a minimum semi-degree of $|G|/2 - \sqrt{|G|/1000}$ suffices (see also [74]).

Note that this degree requirement means that G is not far from being complete. Häggkvist [37] improved the bound further to $|G|/2 - 2^{-15}|G|$ and conjectured that the actual value lies close to $3|G|/8$. The best previously known bound is due to Häggkvist and Thomason [39], who showed that for each $\alpha > 0$ every sufficiently large oriented graph G with minimum semi-degree at least $(5/12+\alpha)|G|$ has a Hamilton cycle. Our first result implies that the actual value is indeed close to $3|G|/8$.

Theorem 1.1 (Kelly, Kühn and Osthus [48]). *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that every oriented graph G of order $|G| \geq N$ with $\delta^0(G) \geq (3/8 + \alpha)|G|$ contains a Hamilton cycle.*

A construction of Häggkvist [37] shows that the bound in Theorem 1.1 is asymptotically best possible (see Proposition 4.6).

We also give two stronger sufficient conditions for a large oriented graph to contain a Hamilton cycle. We show that the property $\delta^*(G) := \delta(G)+\delta^+(G)+\delta^-(G) \geq (3|G|-3)/2$ suffices and we prove an Ore-type result, where $\delta(G)$ is the minimum of $|N(x)|$ over all $x \in V(G)$. Since this work was originally published, Keevash, Kühn and Osthus [46] have improved upon Theorem 1.1, giving an exact minimum semi-degree bound (Theorem 1.9) forcing a Hamilton cycle.

1.2 Cycles of Given Length

A central problem in digraph theory is the Caccetta-Häggkvist conjecture [18]:

Conjecture 1.2. *An oriented graph on n vertices with minimum outdegree d contains a cycle of length at most $\lceil n/d \rceil$.*

Note that in Conjecture 1.2 it does not matter whether we consider oriented graphs or general digraphs. Chvátal and Szemerédi [22] showed that a minimum outdegree of at least d guarantees a cycle of length at most $\lceil 2n/(d+1) \rceil$. For most values of n and d, this is improved by a result of Shen [68], which guarantees a cycle of length at most $3\lceil 0.44n/d \rceil$.

Chvátal and Szemerédi [22] also showed that Conjecture 1.2 holds if we increase the bound on the cycle length by adding a constant c. They showed that $c := 2500$ will do. Nishimura [65] refined their argument to show that one can take $c := 304$. The next result of Shen gives the best known constant.

Theorem 1.3 (Shen [67])**.** *An oriented graph on n vertices with minimum outdegree d contains a cycle of length at most $\lceil n/d \rceil + 73$.*

The special case of Conjecture 1.2 that has attracted most interest is when $d = \lceil n/3 \rceil$. Here the conjecture is that a minimum outdegree of $\lceil n/3 \rceil$ implies a cycle of length 3, that is, a directed triangle. The following bound towards this case improves an earlier one of Caccetta and Häggkvist [18].

Theorem 1.4 (Shen [67])**.** *If G is any oriented graph on n vertices with $\delta^+(G) \geq 0.355n$ then G contains a directed triangle.*

If one considers the *minimum semi-degree* $\delta^0(G) := \min\{\delta^+(G), \delta^-(G)\}$ instead of the minimum outdegree $\delta^+(G)$, then the constant can be improved slightly. The best known value for the constant in this case is currently 0.346. [43] See the monograph [5] or the survey [64] for further partial results on Conjecture 1.2.

We consider the natural and related question of which minimum semi-degree forces cycles of length exactly $\ell \geq 4$ in an oriented graph. We will often refer to cycles of length ℓ as *ℓ-cycles*. Our main result answers this question completely when ℓ is not a multiple of 3.

Theorem 1.5 (Kelly, Kühn and Osthus [49])**.** *Let $\ell \geq 4$. If G is an oriented graph on $n \geq 10^{10}\ell$ vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then G contains an ℓ-cycle. Moreover for any vertex $u \in V(G)$ there is an ℓ-cycle containing u.*

The extremal example showing this to be best possible for $\ell \geq 4$, $\ell \not\equiv 0 \mod 3$ is given by the blow-up of a 3-cycle. More precisely, let G be the oriented graph on n vertices formed by dividing $V(G)$ into 3 vertex classes V_1, V_2, V_3 of as equal size as possible and adding all possible edges from V_i to V_{i+1}, counting modulo 3. Then this oriented graph contains no ℓ-cycle and has minimum semi-degree $\lfloor n/3 \rfloor$.

Also, for all those $\ell \geq 4$ which are multiples of 3, the 'moreover' part is best possible for infinitely many n. To see this, consider the modification of the above example formed by deleting a vertex from the largest vertex class and adding an extra vertex u with $N^+(u) = V_2$ and $N^-(u) = V_1$. This gives an oriented graph with minimum semi-degree $\lfloor (n-1)/3 \rfloor$. For $\ell \equiv 0 \mod 3$ it contains no ℓ-cycle through u.

Perhaps surprisingly, we can do much better than Theorem 1.5 for some cycle lengths (if we do not ask for a cycle through a given vertex). Indeed, we conjecture that the correct bounds are those given by the obvious extremal example: when we seek an ℓ-cycle, the extremal example is probably the blow-up of a k-cycle, where $k \geq 3$ is the smallest integer which is not a divisor of ℓ.

Conjecture 1.6. *Let $\ell \geq 4$ be a positive integer and let $k \geq 3$ be minimal such that k does not divide ℓ. Then there exists an integer $n_0 = n_0(\ell)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lfloor n/k \rfloor + 1$ contains an ℓ-cycle.*

In Chapter 5 we discuss this conjecture in some detail and provide a series of partial results in support of it.

1.3 Arbitrary Orientations of Cycles

It is natural to ask whether the bound in Theorem 1.1 gives only directed Hamilton cycles or whether it gives every possible orientation of a Hamilton cycle. Indeed this question was answered for digraphs, asymptotically at least, by Häggkvist and Thomason in 1995.

Theorem 1.7 (Häggkvist and Thomason [38])**.** *There exists n_0 such that every digraph D on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(D) \geq n/2 + n^{5/6}$ contains every orientation of a Hamilton cycle.*

For oriented graphs this question was asked originally by Häggkvist and Thomason [39] who proved that for all $\alpha > 0$ and all sufficiently large oriented graphs G

a minimum semi-degree of $(5/12 + \alpha)|G|$ suffices to give *any* orientation of a Hamilton cycle. They conjectured that $(3/8 + \alpha)|G|$ suffices, the same bound as for the directed Hamilton cycle up to the error term $\alpha|G|$. Whilst not asked explicitly before Häggkvist and Thomason's paper, there is some previous work of Thomason and Grant relevant to this area. Grant [36] proved in 1980 that any digraph D with minimum semi-degree $\delta^0(D) \geq 2|D|/3 + \sqrt{|D| \log |D|}$ contains an anti-directed Hamilton cycle, provided that n is even. (An anti-directed cycle is one in which the edge orientations alternate.) Thomason [71] showed in 1986 that every sufficiently large tournament contains every possible orientation of a Hamilton cycle (except possibly the directed Hamilton cycle: the transitive tournament with vertices $\{1, 2, \ldots, n\}$ and the edge ij when $i < j$ clearly has no Hamilton cycle as there is no edge out of n). The following theorem confirms the conjecture of Häggkvist and Thomason.

Theorem 1.8 (Kelly [47]). *For every $\alpha > 0$ there exists an integer $n_0 = n_0(\alpha)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3/8+\alpha)n$ contains every orientation of a Hamilton cycle.*

1.4 Pancyclicity

Building on the proof of Theorem 1.1, Keevash, Kühn and Osthus [46] recently gave an exact minimum semi-degree bound which forces a Hamilton cycle in an oriented graph.

Theorem 1.9 (Keevash, Kühn and Osthus [46]). *There exists n_0 such that every oriented graph G on $n \geq n_0$ vertices with $\delta^0(G) \geq (3n-4)/8$ contains a directed Hamilton cycle.*

This is best possible for all $n \geq n_0$. The arguments in [46] can easily be modified to show that G even contains an ℓ-cycle for every $\ell \geq n/10^{10}$ through any given vertex. Details of the changes needed can be found in Chapter 7. Together with Theorems 1.4 and 1.5 this implies that G is *pancyclic*, i.e. it contains cycles of all possible lengths.

Theorem 1.10. *There exists an integer n_0 such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3n-4)/8$ contains an ℓ-cycle for all $3 \leq$*

$\ell \leq n$. Moreover, if $4 \leq \ell \leq n$ and if u is any vertex of G then G contains an ℓ-cycle through u.

We can also extend our result on arbitrary orientations of Hamilton cycles, Theorem 1.8, to a pancyclicity result for arbitrary orientations: If an oriented graph G on n vertices contains every possible orientation of an ℓ-cycle for all $3 \leq \ell \leq n$ we say that G is *universally pancyclic*. The following theorem says that asymptotically universal pancyclicity requires the same minimum semi-degree as pancyclicity.

Theorem 1.11. *For all $\alpha > 0$ there exists an integer $n_0 = n_0(\alpha)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3/8 + \alpha)n$ is universally pancyclic.*

CHAPTER 2

NOTATION AND TERMINOLOGY

This chapter contains the main terminology used in this thesis. For completeness we repeat any definitions given in the introduction.

A simple graph $G = (V(G), E(G))$ is a set of *vertices*, $V(G)$ (or V if this is unambiguous), often taken to be $[n] := \{1, \dots, n\}$, with a set of edges $E(G) \subseteq V^{(2)}$ (or E). The number of vertices in a graph is called its *order* and is often denoted $|G|$. The number of edges in a graph is denoted by $e(G)$. A *multigraph* is a graph in which edges are given a multiplicity.

A *digraph* or *directed graph* is a multigraph in which all the edges are assigned a direction and there are no multiple edges of the same direction. I.e. we allow an edge in each direction between two vertices, but no other multiple edges are allowed. An *oriented graph* is a (simple) graph in which every edge is assigned a direction. Equivalently, an oriented graph is a digraph with no multiple edges.

Given two vertices x and y of an oriented graph G, we write xy for the edge directed from x to y. We write $N^+_G(x)$ for the *outneighbourhood* of a vertex x and $d^+(x) := |N^+_G(x)|$ for its *outdegree*. Similarly, we write $N^-_G(x)$ for the *inneighbourhood* of x and $d^-(x) := |N^-_G(x)|$ for its *indegree*. Given $X \subseteq V(G)$ we denote $|N^+_G(x) \cap X|$ by $d^+_X(x)$, and define $d^-_X(x)$ similarly. We write $N_G(x) := N^+_G(x) \cup N^-_G(x)$ for the *neighbourhood* of x. We use $N^+(x)$ etc. whenever this is unambiguous. The *maximum degree* $\Delta(G)$ of G is the maximum of $|N(x)|$ over all vertices $x \in G$. We write $\delta(G)$, $\delta^+(G)$ and $\delta^-(G)$

respectively for the minimum of $|N(x)|$, $|N^+(x)|$ and $|N^-(x)|$ over all vertices $x \in V(G)$ and call these the *minimum degree*, *minimum indegree* and *minimum outdegree*. The *minimum semi-degree* is defined as $\delta^0(G) := \min(\delta^+(G), \delta^-(G))$.

Given a set A of vertices of G, we write $N^+_G(A)$ for the set of all outneighbours of vertices in A. So $N^+_G(A)$ is the union of $N^+_G(a)$ over all $a \in A$. $N^-_G(A)$ is defined similarly. The directed subgraph $G[A]$ of G induced by $A \subseteq V(G)$ is the oriented graph whose edges are those edges of G with both vertices in A and we write $e(A)$ for the number of its edges. $G - A$ denotes the oriented graph obtained from G by deleting A and all edges incident to A. We say that A is *independent* if $G[A]$ contains no edges. Given disjoint vertex sets A and B in a graph G, an A-B edge is an edge ab where $a \in A$ and $b \in B$. We write $e(A, B)$ for the number of all these edges. We write $(A, B)_G$ for the induced bipartite subgraph of G whose vertex classes are A and B. We write (A, B) where this is unambiguous. (A bipartite graph has vertex set $V(G) = A \cup B$ with $e(A) = e(B) = 0$.)

A path is a sequence of distinct vertices $v_1 v_2 \dots v_\ell$ where each $v_i v_{i+1}$ $(1 \leq i \leq \ell - 1)$ is an edge. Such a path is said to go from v_1 to v_ℓ. An $x - y$ *path* is a path from x to y. We call a path with the standard orientation a *directed path*. A *cycle* is a closed path (i.e. a path where $v_\ell v_1$ is also an edge) and a *Hamilton cycle* is a cycle containing every vertex. An oriented graph is said to be *Hamiltonian* if and only if it contains a Hamilton cycle. A *walk* in G is a sequence $v_1 v_2 \dots v_\ell$ of (not necessarily distinct) vertices, where $v_i v_{i+1}$ (or $v_{i+1} v_i$ if the walk is not directed) is an edge for all $1 \leq i < \ell$. The length of a walk W is $\ell(W) := \ell - 1$. The walk is *closed* if $v_1 = v_\ell$. Given two vertices x, y of G, the *distance* *dist*(x, y) *from* x *to* y is the length of the shortest directed x-y path. The *diameter* of G is the maximum distance between any ordered pair of vertices. Outside of sections clearly pertaining to arbitrary orientations, when referring to paths, cycles and walks in oriented graphs we usually mean that they are directed without mentioning this explicitly.

Given two vertices x and y on a directed cycle C, we write xCy for the subpath of C from x to y. Similarly, given two vertices x and y on a directed path P such that x precedes y, we write xPy for the subpath of P from x to y.

The *underlying graph* of an oriented graph G is the graph obtained from G by ignoring the directions of its edges. We call an orientation of a complete graph a *tournament* and an orientation of a complete bipartite graph a *bipartite tournament*. An oriented graph G is *d-regular* if all vertices have indegree and outdegree d.[1] G is *regular* if it is d-regular for some d. It is easy to see (e.g. by induction) that for every odd n there exists a regular tournament on n vertices. A *1-factor* of an oriented graph is a 1-regular spanning oriented subgraph, i.e. a covering of the oriented graph by pairwise-disjoint cycles. Note that a Hamilton cycle is a connected 1-factor.

Almost all of these definitions apply equally well for digraphs and oriented graphs.

For two functions $f, g : \mathbb{N} \to \mathbb{R}$ we write $f(n) = o(g(n))$ to mean $f(x)/g(x) \to 0$ as $x \to \infty$. We write $0 < a_1 \ll a_2 \ll \ldots \ll a_k$ to mean that we can choose the constants $a_1, a_2, \ldots, a_k$ from right to left. More precisely, there are increasing functions $f_1, f_2, \ldots, f_{k-1}$ such that, given a_k, whenever we choose some $a_i \leq f_i(a_{i+1})$, all calculations needed using these constants are valid.

[1]We will later introduce the entirely separate property of being ε-regular. From the context and the use of Latin and Greek letters as the parameter we hope it is clear to the reader which of these are being used.

CHAPTER 3

SZEMERÉDI'S REGULARITY LEMMA

In this chapter we collect all the information we need about the Diregularity lemma and the Blow-up lemma. See [52] for a survey on the Regularity lemma, originally proved by Szemerédi [70], and [50] for a survey on the Blow-up lemma. We will use the Regularity lemma as a major tool twice. Once in Chapter 4 where we also need a powerful version of the Blow-up lemma due to Csaba (Lemma 3.4). The second time is in Chapter 6, where we use only a relatively weak 'path-embedding lemma'.

3.1 Further Notation

We start with some more notation. The density of a bipartite graph $G = (A, B)$ with vertex classes A and B is defined to be

$$d_G(A, B) := \frac{e_G(A, B)}{|A|\,|B|}.$$

We often write $d(A, B)$ if this is unambiguous. Given $\varepsilon > 0$, we say that G is *ε-regular* if for all subsets $X \subseteq A$ and $Y \subseteq B$ with $|X| > \varepsilon\,|A|$ and $|Y| > \varepsilon\,|B|$ we have that $|d(X, Y) - d(A, B)| < \varepsilon$. Given $d \in [0, 1]$ we say that G is *(ε, d)-super-regular* if it is ε-regular and furthermore $d_G(a) \geq (d - \varepsilon)\,|B|$ for all $a \in A$ and $d_G(b) \geq (d - \varepsilon)\,|A|$ for all $b \in B$. (This is a slight variation of the standard definition of (ε, d)-super-regularity where one requires $d_G(a) \geq d\,|B|$ and $d_G(b) \geq d\,|A|$.)

3.2 The Diregularity Lemma

The Diregularity lemma is a version of the Regularity lemma for digraphs due to Alon and Shapira [2]. Its proof is quite similar to the undirected version. We will use the degree form of the Diregularity lemma which can be easily derived (see e.g. [79] or [55]) from the standard version, in exactly the same manner as the undirected degree form.

Lemma 3.1 (Degree form of the Diregularity lemma)**.** *For every $\varepsilon \in (0,1)$ and every integer M' there are integers M and n_0 such that if G is a digraph on $n \geq n_0$ vertices and $d \in [0,1]$ is any real number, then there is a partition of the vertices of G into $V_0, V_1, \dots, V_k$ and a spanning subdigraph G' of G such that the following holds:*

- $M' \leq k \leq M$,
- $|V_0| \leq \varepsilon n$,
- $|V_1| = \cdots = |V_k| =: m$,
- $d^+_{G'}(x) > d^+_G(x) - (d+\varepsilon)n$ *for all vertices* $x \in G$,
- $d^-_{G'}(x) > d^-_G(x) - (d+\varepsilon)n$ *for all vertices* $x \in G$,
- *for all* $i = 1, \dots, k$ *the digraph* $G'[V_i]$ *is empty,*
- *for all* $1 \leq i, j \leq k$ *with* $i \neq j$ *the bipartite graph whose vertex classes are* V_i *and* V_j *and whose edges are all the* V_i*-*V_j *edges in* G' *is* ε*-regular and has density either* 0 *or density at least* d*.*

$V_1, \dots, V_k$ are called *clusters*, V_0 is called the *exceptional set* and the vertices in V_0 are called *exceptional vertices*. The last condition of the lemma says that all pairs of clusters are ε-regular in both directions (but possibly with different densities). We call the spanning digraph $G' \subseteq G$ given by the Diregularity lemma the *pure digraph*. Given clusters $V_1, \dots, V_k$ and the pure digraph G', the *reduced digraph* R' is the digraph whose vertices are $V_1, \dots, V_k$ and in which V_iV_j is an edge if and only if G' contains a V_i-V_j

edge. Note that the latter holds if and only if the bipartite graph whose vertex classes are V_i and V_j and whose edges are all the V_i-V_j edges in G' is ε-regular and has density at least d. It turns out that R' inherits many properties of G, a fact that is crucial in the proofs in this thesis using the Diregularity lemma. However, R' is not necessarily oriented even if the original digraph G is, but the next lemma shows that by discarding edges with appropriate probabilities one can go over to a reduced oriented graph $R \subseteq R'$ which still inherits many of the properties of G.

Lemma 3.2. *For every $\varepsilon \in (0,1)$ there exist integers $M' = M'(\varepsilon)$ and $n_0 = n_0(\varepsilon)$ such that the following holds. Let $d \in [0,1]$, let G be an oriented graph of order at least n_0 and let R' be the reduced digraph obtained by applying the Diregularity lemma to G with parameters ε, d and M'. Then R' has a spanning oriented subgraph R with*

(a) $\delta^+(R) \geq (\delta^+(G)/|G| - (3\varepsilon + d))\,|R|$,

(b) $\delta^-(R) \geq (\delta^-(G)/|G| - (3\varepsilon + d))\,|R|$,

(c) $\delta(R) \geq (\delta(G)/|G| - (3\varepsilon + 2d))\,|R|$.

Proof. Let us first show that every cluster V_i satisfies

$$|N_{R'}(V_i)|/|R'| \geq \delta(G)/|G| - (3\varepsilon + 2d). \tag{3.1}$$

To see this, consider any vertex $x \in V_i$. As G is an oriented graph, the Diregularity lemma implies that $|N_{G'}(x)| \geq \delta(G) - 2(d+\varepsilon)|G|$. On the other hand, $|N_{G'}(x)| \leq |N_{R'}(V_i)|m + |V_0| \leq |N_{R'}(V_i)||G|/|R'| + \varepsilon|G|$. Altogether this proves (3.1).

We first consider the case when

$$\delta^+(G)/|G| \geq 3\varepsilon + d \quad \text{and} \quad \delta^-(G)/|G| \geq 3\varepsilon + d. \tag{3.2}$$

Let R be the spanning oriented subgraph obtained from R' by deleting edges randomly as follows. For every unordered pair V_i, V_j of clusters we do nothing if either of the edges V_iV_j

and V_jV_i does not exist. Otherwise we delete the edge V_iV_j with probability

$$\frac{e_{G'}(V_j,V_i)}{e_{G'}(V_i,V_j)+e_{G'}(V_j,V_i)}, \tag{3.3}$$

deleting V_jV_i if not. So if R' contains at most one of the edges V_iV_j, V_jV_i then we do nothing. We do this for all unordered pairs of clusters independently and let X_i be the random variable which counts the number of outedges of the vertex $V_i \in R$. Then

$$\begin{aligned}
\mathbb{E}(X_i) &\geq \sum_{j\neq i} \frac{e_{G'}(V_i,V_j)}{e_{G'}(V_i,V_j)+e_{G'}(V_j,V_i)} \geq \sum_{j\neq i} \frac{e_{G'}(V_i,V_j)}{|V_i|\,|V_j|} \\
&\geq \frac{|R'|}{|G|\,|V_i|} \sum_{x\in V_i} (d^+_{G'}(x) - |V_0|) \geq (\delta^+(G')/|G| - \varepsilon)\,|R| \\
&\geq (\delta^+(G)/|G| - (2\varepsilon + d))\,|R| \overset{(3.2)}{\geq} \varepsilon |R|.
\end{aligned}$$

A Chernoff-type bound (see e.g. [3, Cor. A.14]) now implies that there exists a constant $c = c(\varepsilon)$ such that

$$\begin{aligned}
\mathbb{P}(X_i < (\delta^+(G)/|G| - (3\varepsilon + d))\,|R|) &\leq \mathbb{P}(|X_i - \mathbb{E}(X_i)| > \varepsilon \mathbb{E}(X_i)) \\
&\leq \mathrm{e}^{-c\mathbb{E}(X_i)} \leq \mathrm{e}^{-c\varepsilon|R|}.
\end{aligned}$$

Writing Y_i for the random variable which counts the number of inedges of the vertex V_i in R, it follows similarly that

$$\mathbb{P}(Y_i < (\delta^-(G)/|G| - (3\varepsilon + d))\,|R|) \leq \mathrm{e}^{-c\varepsilon|R|}.$$

As $2|R|\mathrm{e}^{-c\varepsilon|R|} < 1$ if M' is chosen to be sufficiently large compared to ε, this implies that there is some outcome R with $\delta^+(R) \geq (\delta^+(G)/|G| - (3\varepsilon + d))\,|R|$ and $\delta^-(R) \geq (\delta^-(G)/|G| - (3\varepsilon + d))\,|R|$. But $N_{R'}(V_i) = N_R(V_i)$ for every cluster V_i and so (3.1) implies that $\delta(R) \geq (\delta(G)/|G| - (3\varepsilon + 2d))|R|$. Altogether this shows that R is as required in the lemma.

If neither of the conditions in (3.2) hold, then (a) and (b) are trivial and one can obtain an oriented graph R which satisfies (c) from R' by arbitrarily deleting one edge from each double edge. If exactly one of the conditions in (3.2) holds, say the first, then (b) is trivial. To obtain an oriented graph R which satisfies (a) we consider the X_i as before, but ignore the Y_i. Again, $N_{R'}(V_i) = N_R(V_i)$ for every cluster V_i and so (c) is also satisfied. □

The oriented graph R given by Lemma 3.2 is called the *reduced oriented graph*. The spanning oriented subgraph G^* of the pure digraph G' obtained by deleting all the V_i-V_j edges whenever $V_iV_j \in E(R') \setminus E(R)$ is called the *pure oriented graph*. Given an oriented subgraph $S \subseteq R$, the *oriented subgraph of G^* corresponding to S* is the oriented subgraph obtained from G^* by deleting all those vertices that lie in clusters not belonging to S as well as deleting all the V_i-V_j edges for all pairs V_i, V_j with $V_iV_j \notin E(S)$.

3.3 The Blow-up Lemma

In the proof of Theorem 4.3 we need the Blow-up lemma, in both the original form of Komlós, Sárközy and Szemerédi [51] and a recent strengthening due to Csaba [26]. We will also use the Blow-up lemma when proving Theorem 6.2, but only in a weaker form discussed in Section 3.4. Roughly speaking, they say that an M-partite graph formed by M clusters such that all the pairs of these clusters are dense and ε-regular behaves like a complete M-partite graph with respect to containing graphs H of bounded maximum degree as subgraphs.

Lemma 3.3 (Blow-up Lemma, Komlós, Sárközy and Szemerédi [51]). *Given a graph R on $[M]$ and positive numbers d and Δ there exists a positive real $\varepsilon_0 = \varepsilon_0(d, \Delta, M)$ such that the following holds for all positive numbers m and all $0 < \varepsilon \leq \varepsilon_0$. Let F be the graph obtained from R by replacing each vertex $i \in R$ with a set V_i of M new vertices and joining all vertices in V_i to all vertices in V_j whenever ij is an edge of R. Let G be a spanning subgraph of F such that for every edge $ij \in R$ the graph $(V_i, V_j)_G$ is (ε, d)-super-regular. Then G contains a copy of every subgraph H of F with maximum degree $\Delta(H) \leq \Delta$.*

Moreover, this copy of H in G maps the vertices of H to the same sets V_i as the copy of H in F, i.e. if $h \in V(H)$ is mapped to V_i by the copy of H in F, then it is also mapped to V_i by the copy of H in G.

Furthermore, given $b > 0$ we can additionally require that for vertices $x \in H \subseteq F$ lying in V_i their images in the copy of H in G are contained in (arbitrary) given sets $C_x \subseteq V_i$ provided that $|C_x| \geq bM$ for each such x, in each V_i there are at most αM such vertices x and $\alpha < \alpha_0(d, \Delta, M, b)$.

The 'furthermore' section of this theorem, whilst not given in their original statement, is implicit in their proof. It also uses the standard definition of super-regularity, but this affects nothing since one can merely use a slightly larger d than would otherwise be necessary.

Observe that in this version the pairs of clusters have to be super-regular and the regularity constant ε_0 depends on the number M of clusters. It is also not explicitly formulated to allow for a set V_0 of exceptional vertices. So we also need the stronger (and more technical) version due to Csaba [26]. The case when $\Delta = 3$ of this is implicit in [27].

Lemma 3.4 (Blow-up Lemma, Csaba [26])**.** *For all integers Δ, K_1, K_2, K_3 and every positive constant c there exists an integer N such that whenever $\varepsilon, \varepsilon', \delta', d$ are positive constants with*

$$0 < \varepsilon \ll \varepsilon' \ll \delta' \ll d \ll 1/\Delta, 1/K_1, 1/K_2, 1/K_3, c$$

the following holds. Suppose that G^* is a graph of order $n \geq N$ and $V_0, \ldots, V_k$ is a partition of $V(G^*)$ such that the bipartite graph $(V_i, V_j)_{G^*}$ is ε-regular with density either 0 or d for all $1 \leq i < j \leq k$. Let H be a graph on n vertices with $\Delta(H) \leq \Delta$ and let $L_0 \cup L_1 \cup \cdots \cup L_k$ be a partition of $V(H)$ with $|L_i| = |V_i| =: m$ for every $i = 1, \ldots, k$. Furthermore, suppose that there exists a bijection $\phi : L_0 \to V_0$ and a set $I \subseteq V(H)$ of vertices at distance at least 4 from each other such that the following conditions hold:

(C1) $|L_0| = |V_0| \leq K_1 dn$.

(C2) $L_0 \subseteq I$.

(C3) L_i *is independent for every* $i = 1, \dots, k$.

(C4) $|N_H(L_0) \cap L_i| \leq K_2 dm$ *for every* $i = 1, \dots, k$.

(C5) *For each* $i = 1, \dots, k$ *there exists* $D_i \subseteq I \cap L_i$ *with* $|D_i| = \delta' m$ *and such that for* $D := \bigcup_{i=1}^{k} D_i$ *and all* $1 \leq i < j \leq k$

$$||N_H(D) \cap L_i| - |N_H(D) \cap L_j|| < \varepsilon m.$$

(C6) *If* $xy \in E(H)$ *and* $x \in L_i, y \in L_j$ *then* $(V_i, V_j)_{G^*}$ *is* ε*-regular with density* d.

(C7) *If* $xy \in E(H)$ *and* $x \in L_0, y \in L_j$ *then* $|N_{G^*}(\phi(x)) \cap V_j| \geq cm$.

(C8) *For each* $i = 1, \dots, k$*, given any* $E_i \subseteq V_i$ *with* $|E_i| \leq \varepsilon' m$ *there exists a set* $F_i \subseteq (L_i \cap (I \setminus D))$ *and a bijection* $\phi_i : E_i \to F_i$ *such that* $|N_{G^*}(v) \cap V_j| \geq (d - \varepsilon)m$ *whenever* $N_H(\phi_i(v)) \cap L_j \neq \emptyset$ *(for all* $v \in E_i$ *and all* $j = 1, \dots, k$*).*

(C9) *Writing* $F := \bigcup_{i=1}^{k} F_i$ *we have that* $|N_H(F) \cap L_i| \leq K_3 \varepsilon' m$.

Then G^* *contains a copy of* H *such that the image of* L_i *is* V_i *for all* $i = 1, \dots, k$ *and the image of each* $x \in L_0$ *is* $\phi(x) \in V_0$.

The additional properties of the copy of H in G^* are not included in the statement of the lemma in [26] but are stated explicitly in the proof.

Let us briefly motivate the conditions of the Blow-up lemma. The embedding of H into G guaranteed by the Blow-up lemma is found by a randomised algorithm which first embeds each vertex $x \in L_0$ to $\phi(x)$ and then successively embeds the remaining vertices of H. So the image of L_0 will be the exceptional set V_0. Condition (C1) requires that there are not too many exceptional vertices and (C2) ensures that we can embed the vertices in L_0 without affecting the neighbourhood of other such vertices. As L_i will be embedded into V_i we need to have (C3). Condition (C5) gives us a reasonably large set D of 'buffer vertices' which will be embedded last by the randomised algorithm. (C6) requires that edges between vertices of $H - L_0$ are embedded into ε-regular pairs of density d. (C7) ensures that the exceptional vertices have large degree in all 'neighbouring clusters'. (C8) and (C9) allow us to embed those vertices whose set of candidate images in G^* has grown very small at some point of the algorithm. Conditions (C6), (C8) and (C9) correspond to a substantial weakening of the super-regularity that the usual form of the Blow-up lemma requires, namely that whenever H contains an edge xy with and $x \in L_i, y \in L_j$ then $(V_i, V_j)_{G^*}$ is (ε, d)-super-regular.

The weakest commonly-used embedding lemma is the following result, often called the Key Lemma, which does not allow us to find spanning graphs but in recompense does not require that any pairs of clusters are super-regular. We use this lemma in Chapter 5.

Theorem 3.5 ([52], Theorem 2.1)**.** *For all $d \in (0, 1]$, $h_0 \in \mathbb{N}$ and $\Delta \geq 1$ there exists $\varepsilon_0 > 0$ and $m_0 \in \mathbb{N}$ such that the following holds. Suppose $\varepsilon \leq \varepsilon_0$ and R is any graph. Let F be the graph obtained from R by replacing each vertex $i \in R$ with a set V_i of $m \geq m_0$ new vertices and joining all vertices in V_i to all vertices in V_j whenever ij is an edge of R. Let G be a spanning subgraph of F such that for every edge $ij \in R$ the graph $(V_i, V_j)_G$ is ε-regular and has density at least d. Then G contains a copy of every subgraph H of F with maximum degree $\Delta(H) \leq \Delta$ and $h \leq h_0$ vertices.*

3.4 A Path-Embedding Lemma

In the proof of Theorem 6.2 we shall only use the following consequence of the Blow-up lemma, which uses similar ideas to those in recent work of Christofides, Keevash, Kühn and Osthus [19].

Lemma 3.6. *Suppose that all the following hold.*

- $0 < 1/m \ll \varepsilon \ll d \ll 1$.
- $U_1, \ldots, U_k$ *are pairwise disjoint sets of size m, for some $k \geq 6$, and G is a digraph*

on $U_1 \cup \ldots \cup U_k$ such that each $(U_i, U_{i+1})_G$ is (ε, d)-super-regular (where by convention we consider U_{k+1} to be U_1).

- *$A_1, \ldots, A_k$ are pairwise disjoint sets of vertices with $(1-\varepsilon)m \le |A_i| := m_i \le m$ and H is a digraph on $A_1 \cup \ldots \cup A_k$ which is a vertex-disjoint union of paths of length at least 3, where every edge going out of A_i end in A_{i+1} for all i.*
- *$S_1 \subseteq U_1, \ldots, S_k \subseteq U_k$ are sets of size $|S_i| = m_i$.*
- *For each path P of H we are given vertices $x_P, y_P \in V(G)$ such that if the initial vertex a_P of P belongs to A_i then $x_P \in S_i$ and if the final vertex b_P of P belongs to A_j then $y_P \in S_j$, and the vertices x_P, y_P are distinct as P ranges over the paths of H.*

Then there is an embedding of H into $G_S := G[\bigcup S_i]$ in which every path P of H is mapped to a path that starts at x_P and ends at y_P.

The following immediate consequence of the Blow-up lemma is needed in the proof of Lemma 3.6.

Lemma 3.7. *For every $0 < d < 1$ and $p \ge 4$ there exists $\varepsilon_0 > 0$ such that the following holds for $0 < \varepsilon < \varepsilon_0$. Let $U_1, \ldots, U_p$ be pairwise disjoint sets of size m, for some m, and suppose G is a graph on $U_1 \cup \ldots \cup U_p$ such that each pair (U_i, U_{i+1}), $1 \le i \le p-1$ is (ε, d)-super-regular. Let $f : U_1 \to U_p$ be any bijective map. Then there are m vertex-disjoint paths from U_1 to U_p so that for every $x \in U_1$ the path starting from x ends at $f(x) \in U_p$.*

We also need the following random partitioning property of super-regular pairs which says that with high probability (i.e. with probability tending to 1 as $m \to \infty$) all new pairs created by a random partition of a super-regular pair are themselves super-regular. (It can be deduced from, for example, Fact 1.5 in [52] and standard Chernoff-type bounds.)

Lemma 3.8. *Suppose that the following hold.*

- *$0 < \varepsilon < \theta < d < 1/2$, $k \ge 2$ and for $1 \le i \le k$ we have $a_i, b_i > \theta$ with $\sum_{i=1}^k a_i = \sum_{i=1}^k b_i = 1$.*

- *$G = (A, B)$ is an (ε, d)-super-regular pair with $|A| = |B| = m$ sufficiently large.*
- *Uniformly at random we choose partitions $A = A_1 \cup \ldots \cup A_k$ and $B = B_1 \cup \ldots \cup B_k$ with $|A_i| = a_i m$ and $|B_i| = b_i m$ for $1 \leq i \leq r$.*

Then with high probability (A_i, B_j) is $(\theta^{-1}\varepsilon, d/2)$-super-regular for every $1 \leq i, j \leq k$.

With these tools we can now prove Lemma 3.6.

Proof. [Of Lemma 3.6] Enumerate the paths of H as $P_1, \ldots, P_p$. We break P_i into consecutive paths $P_{i,1}, P_{i,2}, \ldots, P_{i,q_i}$ where the initial vertex $a_{i,j}$ of $P_{i,j}$ is the terminal vertex $b_{i,j-1}$ of $P_{i,j-1}$. We can take all these paths to have length 3, 4 or 5 as each path has length at least 3. Let E_s consist of all $a_{i,j}$ belonging to the cluster A_s and similarly let F_s consist of all $b_{i,j}$ belonging to the cluster A_s. For each $a_{i,j} \in E_s$ pick a distinct vertex $x_{i,j} \in S_s$ and for each $b_{i,j} \in F_s$ pick a distinct vertex $y_{i,j} \in S_s$ such that if $a_{i,j} = b_{i,j-1}$ then $x_{i,j} = y_{i,j-1}$, $x_{i,1} = x_{P_{i,j}}$ and $y_{i,m_i} = y_{P_{i,j}}$. It is sufficient to show that there is an embedding of H in which each path $P_{i,j}$ is mapped to a path in G_S starting at $x_{i,j}$ and ending at $y_{i,j}$.

For a path $P_{i,j}$ encode whether each edge in $P_{i,j}$ goes forwards or backwards. If $P_{i,j}$ has length 3 then, writing `f` for an edge going from some A_ℓ to $A_{\ell+1}$ and `b` for an edge going from A_ℓ to $A_{\ell-1}$, t encodes one of the following $2^3 = 8$ possibilities:

`fff ffb fbf fbb bff bfb bbf bbb.`

Similarly there are 2^4 possibilities for paths of length 4 and 2^5 for those of length 5. We divide the paths $P_{i,j}$ into $56k$ subcollections $\mathcal{P}_{i,t}$ with $1 \leq i \leq k$, $3 \leq \ell \leq 5$ and

$$t : \{0, 1, \ldots, \ell\} \to \{-\ell, -\ell + 1, \ldots, \ell\}$$

encoding one of the $2^3 + 2^4 + 2^5 = 56$ possibilities discussed above and the length ℓ of the paths. Note that we always have $t(0) = 0$. For example, a path oriented `ffb` would have $t : (0,1,2,3) \mapsto (0,1,2,1)$. $\mathcal{P}_{i,t}$ contains all paths $P_{i,j}$ of length $\ell = \ell(t)$ starting in A_i with each vertex in $\mathcal{P}_{i,j}$ going to the cluster relative to A_i given by t.

Observe that as $|U_i \setminus S_i| \leq \varepsilon m$, every pair (S_i, S_{i+1}) is $(2\varepsilon, d/2)$-super-regular. We first use a greedy algorithm to sequentially embed those collections $\mathcal{P}_{i,t}$ containing at most d^2m paths. That is, we pick any $|\mathcal{P}_{i,t}|$ vertices in S_i to be the start of these paths, and then construct these paths by selecting any (distinct) neighbours of these vertices in the S_j appropriate for each vertex in each path. Each set S_i is met by at most 11×56 of the collections so at any stage in this process we have used at most $6 \times 11 \times 56d^2m$ vertices from any cluster U_i. As we have $d \ll 1$ the restriction of any pair (S_i, S_{i+1}) to the remaining vertices is still $(4\varepsilon, d/4)$-super-regular and so we can indeed do this.

With all the $\mathcal{P}_{i,t}$ containing at most d^2m paths embedded we randomly split the remaining vertices so that for each large $\mathcal{P}_{i,t}$ we have sets $S^0_{i,t} \subseteq S_{t(0)=i}$, $S^1_{i,t} \subseteq S_{t(1)}$, $\ldots$, $S^\ell_{i,t} \subseteq S_{t(\ell)}$ each of size $|\mathcal{P}_{i,t}| > d^2m$. By Lemma 3.8 for each large collection $\mathcal{P}_{i,t}$ and for all $0 \leq r \leq \ell - 1$ the pair $(S^r_{i,t}, S^{r+1}_{i,t})$ if $t(r+1) > t(r)$ or the pair $(S^{r+1}_{i,t}, S^r_{i,t})$ if $t(r+1) < t(r)$ is $(4d^{-2}\varepsilon, d/8)$-super-regular with high probability. Thus for sufficiently large m we can choose a partition with this property and apply Lemma 3.7 to embed each large $\mathcal{P}_{i,t}$ within its allocated sets. □

3.5 Super-regular oriented subgraphs

At various stages in our proof we will need some pairs of clusters to be not just regular but super-regular. The following well-known result (for example, [20]) tells us that we can indeed do this whilst maintaining the regularity of all other pairs.

Lemma 3.9. *Let $\varepsilon \ll d, 1/\Delta$ and let R be a reduced oriented graph of G as given by Lemmas 3.1 and 3.2. Let S be an oriented subgraph of R of maximum degree Δ. Then we can move exactly $2\Delta\varepsilon|V_i|$ vertices from each cluster into V_0 such that each pair (V_i, V_j) corresponding to an edge of S becomes $(2\varepsilon, d/2)$-super-regular and every pair corresponding to an edge of $R \setminus S$ becomes 2ε-regular with density at least $d - \varepsilon$.*

In Chapter 4 we would like to apply the Csaba Blow-up lemma (Lemma 3.4) with G^* being obtained from the underlying graph of the pure oriented graph by adding the exceptional vertices. It will turn out that in order to satisfy (C8) it suffices to ensure that all the edges of a suitable 1-factor in the reduced oriented graph R correspond to (ε, d)-super-regular pairs of clusters. Lemma 3.9 states that this can be ensured by removing a small proportion of vertices from each cluster V_i, and so (C8) can be satisfied. However, (C6) requires all the edges of R to correspond to ε-regular pairs of density precisely d and not just at least d. (As remarked by Csaba [26], it actually suffices that the densities are close to d in terms of ε.) The following proposition shows that this does not pose a problem.

Proposition 3.10. *Let M', n_0, D be integers and let ε, d be positive constants such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \ll 1/D$. Let G be an oriented graph of order at least n_0. Let R be the reduced oriented graph and let G^* be the pure oriented graph obtained by successively applying first the Diregularity lemma with parameters ε, d and M' to G and then Lemma 3.2. Let S be an oriented subgraph of R with $\Delta(S) \leq D$. Let G' be the underlying graph of G^*. Then one can delete $2D\varepsilon|V_i|$ vertices from each cluster V_i to obtain subclusters $V_i' \subseteq V_i$ in such a way that G' contains a subgraph G'_S whose vertex set is the union of all the V_i' and such that*

- *$(V_i', V_j')_{G'_S}$ is $(\sqrt{\varepsilon}, d - 4D\varepsilon)$-super-regular whenever $V_iV_j \in E(S)$,*
- *$(V_i', V_j')_{G'_S}$ is $\sqrt{\varepsilon}$-regular and has density $d - 4D\varepsilon$ whenever $V_iV_j \in E(R)$.*

Proof. Consider any cluster $V_i \in V(S)$ and any neighbour V_j of V_i in S. Recall that $m = |V_i|$. Let d_{ij} denote the density of the bipartite subgraph $(V_i, V_j)_{G'}$ of G' induced by V_i and V_j. So $d_{ij} \geq d$ and this bipartite graph is ε-regular by the remarks before Lemma 3.2. Thus there are at most $2\varepsilon m$ vertices $v \in V_i$ such that $||N_{G'}(v) \cap V_j| - d_{ij}m| > \varepsilon m$. So in total there are at most $2D\varepsilon m$ vertices $v \in V_i$ such that $||N_{G'}(v) \cap V_j| - d_{ij}m| > \varepsilon m$ for some neighbour V_j of V_i in S. Delete all these vertices as well as some more vertices if necessary to obtain a subcluster $V_i' \subseteq V_i$ of size $(1 - 2D\varepsilon)m =: m'$. Delete any $2D\varepsilon m$

vertices from each cluster $V_i \in V(R) \setminus V(S)$ to obtain a subcluster V_i'. It is easy to check that for each edge $V_iV_j \in E(R)$ the graph $(V_i', V_j')_{G'}$ is still 2ε-regular and that its density d_{ij}' satisfies

$$d' := d - 4D\varepsilon < d_{ij} - \varepsilon \leq d_{ij}' \leq d_{ij} + \varepsilon.$$

Moreover, whenever $V_iV_j \in E(S)$ and $v \in V_i'$ we have that

$$(d_{ij} - 4D\varepsilon)m' \leq |N_{G'}(v) \cap V_j'| \leq (d_{ij} + 4D\varepsilon)m'.$$

For every pair V_i, V_j of clusters with $V_iV_j \in E(S)$ we now consider a spanning random subgraph G_{ij}' of $(V_i', V_j')_{G'}$ which is obtained by choosing each edge of $(V_i', V_j')_{G'}$ with probability d'/d_{ij}', independently of the other edges. Consider any vertex $v \in V_i'$. Then the expected number of neighbours of v in V_j' (in the graph G_{ij}') is at least $(d_{ij} - 4D\varepsilon)d'm'/d_{ij}' \geq (1 - \sqrt{\varepsilon})d'm'$ (for ε sufficiently small). So we can apply a Chernoff-type bound to see that there exists a constant $c = c(\varepsilon)$ such that

$$\mathbb{P}(|N_{G_{ij}'}(v) \cap V_j'| \leq (d' - \sqrt{\varepsilon})m') \leq \mathrm{e}^{-cd'm'}.$$

Similarly, whenever $X \subseteq V_i'$ and $Y \subseteq V_j'$ are sets of size at least $2\varepsilon m'$ the expected number of X-Y edges in G_{ij}' is $d_{G'}(X, Y)d'|X||Y|/d_{ij}'$. Since $(V_i', V_j')_{G'}$ is 2ε-regular this expected number lies between $(1 - \sqrt{\varepsilon})d'|X||Y|$ and $(1 + \sqrt{\varepsilon})d'|X||Y|$. So again we can use a Chernoff-type bound to see that

$$\mathbb{P}(|e_{G_{ij}'}(X, Y) - d'|X||Y|| > \sqrt{\varepsilon}|X||Y|) \leq \mathrm{e}^{-cd'|X||Y|} \leq \mathrm{e}^{-4cd'(\varepsilon m')^2}.$$

Moreover, with probability at least $1/(3m')$ the graph G_{ij}' has its expected density d' (see e.g. [10, p. 6]). Altogether this shows that with probability at least

$$1/(3m') - 2m'\mathrm{e}^{-cd'm'} - 2^{2m'}\mathrm{e}^{-4cd'(\varepsilon m')^2},$$

which is greater than 0 for sufficiently large m', we have that G'_{ij} is $(\sqrt{\varepsilon}, d')$-super-regular and has density d'. Proceed similarly for every pair of clusters forming an edge of S. An analogous argument applied to a pair V_i, V_j of clusters with $V_iV_j \in E(R) \setminus E(S)$ shows that with non-zero probability the random subgraph G'_{ij} is $\sqrt{\varepsilon}$-regular and has density d'. Altogether this gives us the desired subgraph G'_S of G'. □

CHAPTER 4

HAMILTON CYCLES IN ORIENTED GRAPHS

4.1 Introduction

When discussing cycles and paths in digraphs in this chapter we always mean that they are directed without mentioning this explicitly.

As discussed in the preface, there is an obvious analogue of a Hamilton cycle for digraphs. That is, an ordering $x_1, \dots, x_n$ of the vertices of a digraph D such that $x_i x_{i+1}$ is a directed edge for all i. A fundamental result of Dirac states that a minimum degree of $|G|/2$ guarantees a Hamilton cycle in an undirected graph G on at least 3 vertices. Ore in 1960 gave a stronger sufficient condition: if the sum of the degrees of every pair of non-adjacent vertices is at least $|G|$, then the graph is Hamiltonian [66]. There are corresponding versions of these famous theorems of Dirac and Ore for digraphs. Ghouila-Houri [34] proved in 1960 that every digraph D with minimum semi-degree at least $|D|/2$ contains a Hamilton cycle. Meyniel [60] showed that an analogue of Ore's theorem holds for digraphs; that is, a digraph on at least 4 vertices is either Hamiltonian or the sum of the degrees of a pair of non-adjacent vertices is less than $2|D| - 1$. All these bounds are best possible.

It is natural to ask for the (smallest) minimum semi-degree which guarantees a Hamilton cycle in an oriented graph G. This question was first raised by Thomassen [73], who [75] showed that a minimum semi-degree of $|G|/2 - \sqrt{|G|/1000}$ suffices (see also [74]).

Note that this degree requirement means that G is not far from being complete. Häggkvist [37] improved the bound further to $|G|/2 - 2^{-15}|G|$ and conjectured that the actual value lies close to $3|G|/8$. The best previously known bound is due to Häggkvist and Thomason [39], who showed that for each $\alpha > 0$ every sufficiently large oriented graph G with minimum semi-degree at least $(5/12 + \alpha)|G|$ has a Hamilton cycle. Our first result (Theorem 1.1 in the preface) implies that the actual value is indeed close to $3|G|/8$.

Theorem 4.1. *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that every oriented graph G of order $|G| \geq N$ with $\delta^0(G) \geq (3/8 + \alpha)|G|$ contains a Hamilton cycle.*

A construction of Häggkvist [37] shows that the bound in Theorem 1.1 is asymptotically best possible (see Proposition 4.6).

In fact, Häggkvist [37] formulated the following stronger conjecture. Given an oriented graph G, recall that $\delta(G)$ denotes the minimum degree of G (i.e. the minimum number of edges incident to a vertex) and set $\delta^*(G) := \delta(G) + \delta^+(G) + \delta^-(G)$.

Conjecture 4.2 (Häggkvist [37]). *Every oriented graph G with $\delta^*(G) > (3n - 3)/2$ has a Hamilton cycle.*

Our next result (stated in the preface as Theorem **??** provides an approximate confirmation of this conjecture for large oriented graphs.

Theorem 4.3. *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that every oriented graph G of order $|G| \geq N$ with $\delta^*(G) \geq (3/2 + \alpha)|G|$ contains a Hamilton cycle.*

Note that Theorem 4.1 is an immediate consequence of this. Once one has a Dirac-type result it is natural to ask if there is a corresponding Ore-type result and indeed in this case there is.

Theorem 4.4. *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that if G is an oriented graph on $n \geq N$ vertices with $d^+(u) + d^-(v) \geq 3n/4 + \alpha n$ for all non-adjacent vertices $u, v \in V(G)$ then G contains a Hamilton cycle.*

The proof for this is similar to that of Theorem 4.3 so we do not give the entire proof. We do though give a proof of the one important lemma which is different, along with a brief discussion, in Section 4.6.

Since this work was originally published, Keevash, Kühn and Osthus [46] have improved upon Theorem 4.1, proving that in any sufficiently large oriented graph G having a minimum semi-degree of at least $\delta^0(G) \geq (3|G| - 4)/8$ suffices. (See Theorem 1.9.)

Moreover, note that Theorem 4.1 immediately implies a partial result towards a classical conjecture of Kelly (see e.g. [5]), which states that every regular tournament on n vertices can be partitioned into $(n-1)/2$ edge-disjoint Hamilton cycles.

Corollary 4.5. *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that every regular tournament of order $n \geq N$ contains at least $(1/8 - \alpha)n$ edge-disjoint Hamilton cycles.*

Indeed, Corollary 4.5 follows from Theorem 4.1 by successively removing Hamilton cycles until the oriented graph G obtained from the tournament in this way has minimum semi-degree less than $(3/8 + \alpha)|G|$. The best previously known bound on the number of edge-disjoint Hamilton cycles in a regular tournament is the one which follows from the result of Häggkvist and Thomason [39] mentioned above. A related result of Frieze and Krivelevich [33] implies that almost every tournament contains a collection of edge-disjoint Hamilton cycles which covers almost all of its edges and that the same holds for almost all regular tournaments. Since this research was originally carried out, Kühn, Osthus and Treglown [57] have proved an approximate version of Kelly's conjecture for large tournaments.

4.2 Extremal Example

The following construction of Häggkvist [37] shows that Conjecture 4.2 is best possible for infinitely many values of $|G|$. We include it here for completeness.

Proposition 4.6. *There are infinitely many oriented graphs G with minimum semi-degree $(3|G| - 5)/8$ which do not contain a 1-factor and thus do not contain a Hamilton*

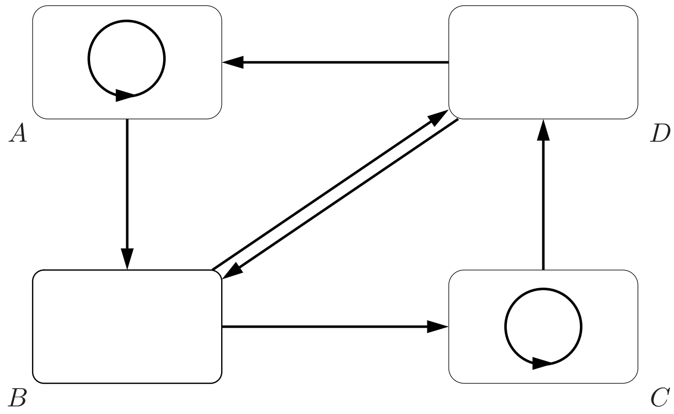


Figure 4.1: The oriented graph in the proof of Proposition 4.6.

cycle.

Proof. Let $n := 4m+3$ for some odd $m \in \mathbb{N}$. Let G be the oriented graph obtained from the disjoint union of two regular tournaments A and C on m vertices, a set B of $m+2$ vertices and a set D of $m+1$ vertices by adding all edges from A to B, all edges from B to C, all edges from C to D as well as all edges from D to A. Finally, between B and D we add edges to obtain a bipartite tournament which is as regular as possible, i.e. the indegree and outdegree of every vertex differ by at most 1. So in particular every vertex in B sends exactly $(m+1)/2$ edges to D (Figure 1).

It is easy to check that the minimum semi-degree of G is $(m-1)/2 + (m+1) = (3n-5)/8$, as required. Since every path which joins two vertices in B has to pass through D, it follows that every cycle contains at least as many vertices from D as it contains from B. As $|B| > |D|$ this means that one cannot cover all the vertices of G by disjoint cycles, i.e. G does not contain a 1-factor. □

4.3 Overview of the proof of Theorem 4.3

Let G be our given oriented graph. The rough idea of the proof is to apply the Diregularity lemma and Lemma 3.2 to obtain a reduced oriented graph R and a pure oriented graph G^*. The following result of Häggkvist implies that R contains a 1-factor.

Theorem 4.7 (Häggkvist [37])**.** *Let R be an oriented graph with $\delta^*(R) > (3|R| - 3)/2$. Then R has a directed path between every 2 vertices and contains a 1-factor.*

So one can apply the Blow-up lemma (together with Proposition 3.10) to find a 1-factor in $G^* - V_0 \subseteq G - V_0$. One now would like to glue the cycles of this 1-factor together and to incorporate the exceptional vertices to obtain a Hamilton cycle of G^* and thus of G. However, we were only able to find a method which incorporates a set of vertices whose size is small compared to the cluster size m. This is not necessarily the case for V_0. So we proceed as follows. We first choose a random partition of the vertex set of G into two sets A and $V(G) \setminus A$ having roughly equal size. We then apply the Diregularity lemma to $G - A$ in order to obtain clusters $V_1, \dots, V_k$ and an exceptional set V_0. We let m denote the size of these clusters and set $B := V_1 \cup \dots V_k$. By arguing as indicated above, we can find a Hamilton cycle C_B in $G[B]$. We then apply the Diregularity lemma to $G - B$, but with an ε which is small compared to $1/k$, to obtain clusters $V_1', \dots, V_\ell'$ and an exceptional set V_0'. Since the choice of our partition $A, V(G) \setminus A$ will imply that $\delta^*(G - B) \geq (3/2 + \alpha/2)|G - B|$ we can again argue as before to obtain a cycle C_A which covers precisely the vertices in $A' := V_1' \cup \dots \cup V_\ell'$. Since we have chosen ε to be small compared to $1/k$, the set V_0' of exceptional vertices is now small enough to be incorporated into our first cycle C_B. (Actually, C_B is only determined at this point and not yet earlier on.) Moreover, by choosing C_B and C_A suitably we can ensure that they can be joined together into the desired Hamilton cycle of G.

4.4 Shifted Walks

In this section we will introduce the tools we need in order to glue certain cycles together and to incorporate the exceptional vertices. Let R^* be a digraph and let $\mathcal{C}$ be a collection of disjoint cycles in R^*. We call a closed walk W in R^* *balanced w.r.t.* $\mathcal{C}$ if

- for each cycle $C \in \mathcal{C}$ the walk W visits all the edges on C an equal number of times,
- W visits every vertex of R^*,
- every vertex not in any cycle from $\mathcal{C}$ is visited exactly once.

Let us now explain why balanced walks are helpful in order to incorporate the exceptional vertices. Suppose that $\mathcal{C}$ is a 1-factor of the reduced oriented graph R and that R^* is obtained from R by adding all the exceptional vertices $v \in V_0$ and adding an edge vV_i (where V_i is a cluster and $v \in V_0$) whenever v sends edges to a significant proportion of the vertices in V_i, say we add vV_i whenever v sends at least cm edges to V_i. (Recall that m denotes the size of the clusters.) The edges in R^* of the form V_iv are defined in a similar way. Let G^c be the oriented graph obtained from the pure oriented graph G^* by making all the nonempty bipartite subgraphs between the clusters complete (and orienting all the edges between these clusters in the direction induced by R) and adding the vertices in V_0 as well as all the edges of G between V_0 and $V(G)\setminus V_0$. Suppose that W is a balanced closed walk in R^* which visits all the vertices lying on a cycle $C \in \mathcal{C}$ precisely $m_C \leq m$ times. Furthermore, suppose that $|V_0| \leq cm/2$ and that the vertices in V_0 have distance at least 3 from each other on W. Then by 'winding around' each cycle $C \in \mathcal{C}$ precisely $m - m_C$ times (at the point when W first visits C) we can obtain a Hamilton cycle in G^c. Indeed, the two conditions on V_0 ensure that the neighbours of each $v \in V_0$ on the Hamilton cycle can be chosen amongst the at least cm neighbours of v in the neighbouring clusters of v on W in such a way that they are distinct for different exceptional vertices. The idea then is to apply the Blow-up lemma to show that this Hamilton cycle corresponds to one in G. So our aim is to find such a balanced closed walk in R^*. However, as indicated in

Section 4.3, the difficulties arising when trying to ensure that the exceptional vertices lie on this walk will force us to apply the above argument to the subgraphs induced by a random partition of our given oriented graph G.

Let us now go back to the case when R^* is an arbitrary digraph and $\mathcal{C}$ is a collection of disjoint cycles in R^*. Given vertices $a, b \in R^*$, a *shifted a-b walk* is a walk of the form

$$W = aa_1C_1b_1a_2C_2b_2 \dots a_tC_tb_tb$$

where $C_1, \dots, C_t$ are (not necessarily distinct) cycles from $\mathcal{C}$ and a_i is the successor of b_i on C_i for all $i \leq t$. (We might have $t = 0$. So an edge ab is a shifted a-b walk.) We call $C_1, \dots, C_t$ the cycles which are *traversed* by W. So even if the cycles $C_1, \dots, C_t$ are not distinct, we say that W traverses t cycles. Note that for every cycle $C \in \mathcal{C}$ the walk $W - \{a, b\}$ visits the vertices on C an equal number of times. Thus it will turn out that by joining the cycles from $\mathcal{C}$ suitably via shifted walks and incorporating those vertices of R^* not covered by the cycles from $\mathcal{C}$ we can obtain a balanced closed walk on R^*.

Our next lemma will be used to show that if R^* is oriented and $\delta^*(R^*) \geq (3/2+\alpha)|R^*|$ then any two vertices of R^* can be joined by a shifted walk traversing only a small number of cycles from $\mathcal{C}$ (see Corollary 4.10). The lemma itself shows that the δ^* condition implies expansion, and this will give us the 'expansion with respect to shifted neighbourhoods' we need for the existence of shifted walks. The proof of Lemma 4.8 is similar to that of Theorem 4.7.

Lemma 4.8. *Let R^* be an oriented graph on N vertices with $\delta^*(R^*) \geq (3/2 + \alpha)N$ for some $\alpha > 0$. If $X \subseteq V(R^*)$ is nonempty and $|X| \leq (1-\alpha)N$ then $|N^+(X)| \geq |X| + \alpha N/2$.*

Proof. For simplicity, we write $\delta := \delta(R^*)$, $\delta^+ := \delta^+(R^*)$ and $\delta^- := \delta^-(R^*)$. Suppose the assertion is false, i.e. there exists $X \subseteq V(R^*)$ with $|X| \leq (1 - \alpha)N$ and

$$\left|N^+(X)\right| < |X| + \alpha N/2. \tag{4.1}$$

We consider the following partition of $V(R^*)$:

$$A := X \cap N^+(X), \quad B := N^+(X)\backslash X, \quad C := V(R^*)\backslash(X \cup N^+(X)), \quad D := X\backslash N^+(X).$$

(4.1) gives us

$$|D| + \alpha N/2 > |B|. \tag{4.2}$$

Suppose $A \neq \emptyset$. If $|N^+(x) \cap A| \geq |A|/2$ for every $x \in A$ then we would have $e(A) \geq |A|^2/2$, contradicting the fact that $e(A) \leq |A|(|A|-1)$. Thus there exists $x \in A$ with $|N^+(x) \cap A| < |A|/2$. From this we can that $\delta^+ \leq |N^+(x)| < |B| + |A|/2$. Combining this with (4.2) we get

$$|A| + |B| + |D| \geq 2\delta^+ - \alpha N/2. \tag{4.3}$$

If $A = \emptyset$ then $N^+(X) = B$ and so (4.2) implies $|D| + \alpha N/2 \geq |B| \geq \delta^+$. Thus (4.3) again holds. Similarly, if $C \neq \emptyset$ then considering the inneighbourhood of a suitable vertex $x \in C$ gives

$$|B| + |C| + |D| \geq 2\delta^- - \alpha N/2. \tag{4.4}$$

Suppose $C = \emptyset$: then if $D = \emptyset$ also we have $V(R^*) = X \cup N^+(X) = (X \setminus N^+(X)) \cup N^+(X) = N^+(X)$, but the right-hand side has order less than $(1-\alpha/2)N$, which is a contradiction. Thus $D \neq \emptyset$ and so $N^-(D) \subseteq B$ (as $C \neq \emptyset$), which gives $|B| \geq \delta^-$. Together with (4.2) this shows that (4.4) holds in this case too.

If $D = \emptyset$ then trivially $|A| + |B| + |C| = N \geq \delta$. If not, then for any $x \in D$ we have $N(x) \cap D = \emptyset$ and hence

$$2|A| + 2|B| + 2|C| \geq 2\,|N(x)| \geq 2\delta. \tag{4.5}$$

Combining (4.3), (4.4) and (4.5) gives

$$3|A| + 4|B| + 3|C| + 2|D| \geq 2\delta^- + 2\delta^+ + 2\delta - \alpha N = 2\delta^*(R^*) - \alpha N.$$

Finally, substituting (4.2) gives

$$3N + \alpha N/2 \geq 2\delta^*(R^*) - \alpha N \geq 3N + \alpha N,$$

which is a contradiction. □

As indicated before, we will now use Lemma 4.8 to prove the existence of shifted walks in R^* traversing only a small number of cycles from a given 1-factor of R^*. For this (and later on) the following fact will be useful.

Fact 4.9. *Let G be an oriented graph with $\delta^*(G) \geq (3/2+\alpha)|G|$ for some constant $\alpha > 0$. Then $\delta^0(G) > \alpha|G|$.*

Proof. Suppose that $\delta^-(G) \leq \alpha|G|$. As G is oriented we have that $\delta^+(G) < |G|/2$ and so $\delta^*(G) < 3n/2 + \alpha|G|$, a contradiction. The proof for $\delta^+(G)$ is similar. □

Corollary 4.10. *Let R^* be an oriented graph on N vertices with $\delta^*(R^*) \geq (3/2+\alpha)N$ for some $\alpha > 0$ and let $\mathcal{C}$ be a 1-factor in R^*. Then for any distinct $x, y \in V(R^*)$ there exists a shifted x-y walk traversing at most $2/\alpha$ cycles from $\mathcal{C}$.*

Proof. Let X_i be the set of vertices v for which there is a shifted x-v walk which traverses at most i cycles. So $X_0 = N^+(x) \neq \emptyset$ and $X_{i+1} = N^+(X_i^-) \cup X_i$, where X_i^- is the set of all predecessors of the vertices in X_i on the cycles from $\mathcal{C}$. Suppose that $|X_i| \leq (1-\alpha)N$. Then Lemma 4.8 implies that

$$|X_{i+1}| \geq |N^+(X_i^-)| \geq |X_i^-| + \alpha N/2 = |X_i| + \alpha N/2.$$

Thus $|X_i| \geq i\alpha N/2$ for all i, so taking $i^* := \lfloor 2/\alpha \rfloor - 1$ we have $|X_{i^*}^-| = |X_{i^*}| \geq (1-\alpha)N$. But $|N^-(y)| \geq \delta^-(R^*) > \alpha N$ by Fact 4.9 and so $N^-(y) \cap X_{i^*}^- \neq \emptyset$. In other words, $y \in N^+(X_{i^*}^-)$ and so there is a shifted x-y walk traversing at most $i^* + 1$ cycles. □

Corollary 4.11. *Let R^* be an oriented graph with $\delta^*(R^*) \geq (3/2 + \alpha)\,|R^*|$ for some $0 < \alpha \leq 1/6$ and let $\mathcal{C}$ be a 1-factor in R^*. Then R^* contains a closed walk which is balanced w.r.t. $\mathcal{C}$ and meets every vertex at most $|R^*|/\alpha$ times and traverses each edge lying on a cycle from $\mathcal{C}$ at least once.*

Proof. Let $C_1, \dots, C_s$ be an arbitrary ordering of the cycles in $\mathcal{C}$. For each cycle C_i pick a vertex $c_i \in C_i$. Denote by c_i^+ the successor of c_i on the cycle C_i. Corollary 4.10 implies that for all i there exists a shifted c_i-c_{i+1}^+ walk W_i traversing at most $2/\alpha$ cycles from $\mathcal{C}$, where $c_{s+1} := c_1$. Then the closed walk

$$W' := c_1^+ C_1 c_1 W_1 c_2^+ C_2 c_2 \dots W_{s-1} c_s^+ C_s c_s W_s c_1^+$$

is balanced w.r.t. $\mathcal{C}$ by the definition of shifted walks. Since each shifted walk W_i traverses at most $2/\alpha$ cycles of $\mathcal{C}$, the closed walk W' meets each vertex at most $(|R^*|/3)(2/\alpha) + 1$ times. Let W denote the walk obtained from W' by 'winding around' each cycle $C \in \mathcal{C}$ once more. (That is, for each $C \in \mathcal{C}$ pick a vertex v on C and replace one of the occurences of v on W' by vCv.) Then W is still balanced w.r.t. $\mathcal{C}$, traverses each edge lying on a cycle from $\mathcal{C}$ at least once and visits each vertex of R^* at most $(|R^*|/3)(2/\alpha) + 2 \leq |R^*|/\alpha$ times as required. □

4.5 Proof of Theorem 4.3

4.5.1 Partitioning G and applying the Diregularity lemma

Let G be an oriented graph on n vertices with $\delta^*(G) \geq (3/2+\alpha)n$ for some constant $\alpha > 0$. Clearly we may assume that $\alpha \ll 1$. Define positive constants ε, d and integers M'_A, M'_B such that

$$1/M'_A \ll 1/M'_B \ll \varepsilon \ll d \ll \alpha \ll 1.$$

Throughout this section, we will assume that n is sufficiently large compared to M'_A for our estimates to hold. Choose a subset $A \subseteq V(G)$ with $(1/2-\varepsilon)n \le |A| \le (1/2+\varepsilon)n$ and such that every vertex $x \in G$ satisfies

$$\frac{d^+(x)}{n} - \frac{\alpha}{10} \le \frac{|N^+(x)\cap A|}{|A|} \le \frac{d^+(x)}{n} + \frac{\alpha}{10}$$

and such that $N^-(x)\cap A$ satisfies a similar condition. (The existence of such a set A can be shown by considering a random partition of $V(G)$.) Apply the Diregularity lemma (Lemma 3.1) with parameters ε^2, $d+8\varepsilon^2$ and M'_B to $G-A$ to obtain a partition of the vertex set of $G-A$ into $k \ge M'_B$ clusters $V_1,\dots,V_k$ and an exceptional set V_0. Set $B := V_1 \cup \ldots \cup V_k$ and $m_B := |V_1| = \cdots = |V_k|$. Let R_B denote the reduced oriented graph obtained by an application of Lemma 3.2 and let G^*_B be the pure oriented graph. Since $\delta^+(G-A)/|G-A| \ge \delta^+(G)/n - \alpha/9$ by our choice of A, Lemma 3.2 implies that

$$\delta^+(R_B) \ge (\delta^+(G)/n - \alpha/8)|R_B|. \tag{4.6}$$

Similarly

$$\delta^-(R_B) \ge (\delta^-(G)/n - \alpha/8)|R_B| \tag{4.7}$$

and $\delta(R_B) \ge (\delta(G)/n - \alpha/4)|R_B|$. Altogether this implies that

$$\delta^*(R_B) \ge (3/2 + \alpha/2)|R_B|. \tag{4.8}$$

So Theorem 4.7 gives us a 1-factor $\mathcal{C}_B$ of R_B. We now apply Proposition 3.10 with $\mathcal{C}_B$ playing the role of S, ε^2 playing the role of ε and $d+8\varepsilon^2$ playing the role of d. This shows that by adding at most $4\varepsilon^2 n$ further vertices to the exceptional set V_0 we may assume that each edge of R_B corresponds to an ε-regular pair of density d (in the underlying graph of G^*_B) and that each edge in the union $\bigcup_{C\in\mathcal{C}_B} C \subseteq R_B$ of all the cycles from $\mathcal{C}_B$ corresponds to an (ε, d)-super-regular pair. (More formally, this means that we replace the clusters with the subclusters given by Proposition 3.10 and replace G^*_B with its oriented subgraph obtained by deleting all edges not corresponding to edges of the graph $G'_{\mathcal{C}_B}$ given by Proposition 3.10, i.e. the underlying graph of G^*_B will now be $G'_{\mathcal{C}_B}$.) Note that the new exceptional set now satisfies $|V_0| \leq \varepsilon n$.

Apply Corollary 4.11 with $R^* := R_B$ to find a closed walk W_B in R_B which is balanced w.r.t. $\mathcal{C}_B$, meets every cluster at most $2|R_B|/\alpha$ times and traverses all the edges lying on a cycle from $\mathcal{C}_B$ at least once.

Let G^c_B be the oriented graph obtained from G^*_B by adding all the V_i-V_j edges for all those pairs V_i, V_j of clusters with $V_iV_j \in E(R_B)$. Since $2|R_B|/\alpha \ll m_B$, we could make W_B into a Hamilton cycle of G^c_B by 'winding around' each cycle from $\mathcal{C}_B$ a suitable number of times. We could then apply the Blow-up lemma to show that this Hamilton cycle corresponds to one in G^*_B. However, as indicated in Section 4.3, we will argue slightly differently as it is not clear how to incorporate all the exceptional vertices by the above approach.

Set $\varepsilon_A := \varepsilon/|R_B|$. Apply the Diregularity lemma with parameters ε_A^2, $d+8\varepsilon_A^2$ and M'_A to $G[A \cup V_0]$ to obtain a partition of the vertex set of $G[A \cup V_0]$ into $\ell \geq M'_A$ clusters $V'_1, \dots, V'_\ell$ and an exceptional set V'_0. Let $A' := V'_1 \cup \dots \cup V'_\ell$, let R_A denote the reduced oriented graph obtained from Lemma 3.2 and let G^*_A be the pure oriented graph. Similarly as in (4.8), Lemma 3.2 implies that $\delta^*(R_A) \geq (3/2 + \alpha/2)|R_A|$ and so, as before, we can apply Theorem 4.7 to find a 1-factor $\mathcal{C}_A$ of R_A. Then as before, Proposition 3.10 implies that by adding at most $4\varepsilon_A^2 n$ further vertices to the exceptional set V'_0 we may assume that each edge of R_A corresponds to an ε_A-regular pair of density d and that each edge in the union $\bigcup_{C \in \mathcal{C}_A} C \subseteq R_A$ of all the cycles from $\mathcal{C}_A$ corresponds to an (ε_A, d)-super-regular pair. So we now have that

$$|V'_0| \leq \varepsilon_A n = \varepsilon n/|R_B|. \tag{4.9}$$

Similarly as before, Corollary 4.11 gives us a closed walk W_A in R_A which is balanced w.r.t. $\mathcal{C}_A$, meets every cluster at most $2|R_A|/\alpha$ times and traverses all the edges lying on a cycle from $\mathcal{C}_A$ at least once.

4.5.2 Incorporating V_0' into the walk W_B

Recall that the balanced closed walk W_B in R_B corresponds to a Hamilton cycle in G_B^c. Our next aim is to extend this walk to one which corresponds to a Hamilton cycle which also contains the vertices in V_0'. (The Blow-up lemma will imply that the latter Hamilton cycle corresponds to one in $G[B \cup V_0']$.) We do this by extending W_B into a walk on a suitably defined digraph $R_B^* \supseteq R_B$ with vertex set $V(R_B) \cup V_0'$ in such a way that the new walk is balanced w.r.t. $\mathcal{C}_B$. R_B^* is obtained from the union of R_B and the set V_0' by adding an edge vV_i between a vertex $v \in V_0'$ and a cluster $V_i \in V(R_B)$ whenever $\left|N_G^+(v) \cap V_i\right| > \alpha m_B/10$ and adding the edge V_iv whenever $\left|N_G^-(v) \cap V_i\right| > \alpha m_B/10$. Thus

$$|N_G^+(v) \cap B| \leq |N_{R_B^*}^+(v)|m_B + |R_B|\alpha m_B/10.$$

Hence

$$\begin{aligned}
|N_{R_B^*}^+(v)| &\geq |N_G^+(v) \cap B|/m_B - \alpha|R_B|/10 \geq |N_G^+(v) \cap B||R_B|/|B| - \alpha|R_B|/10 \\
&\geq (|N_{G-A}^+(v)| - |V_0|)|R_B|/|G - A| - \alpha|R_B|/10 \\
&\geq (\delta^+(G)/n - \alpha/2)|R_B| \geq \alpha|R_B|/2. \qquad (4.10)
\end{aligned}$$

(The penultimate inequality follows from the choice of A and the final one from Fact 4.9.) Similarly

$$|N_{R_B^*}^-(v)| \geq \alpha|R_B|/2.$$

Given a vertex $v \in V_0'$ pick $U_1 \in N_{R_B^*}^+(v)$, $U_2 \in N_{R_B^*}^-(v)\backslash\{U_1\}$. Let C_1 and C_2 denote the cycles from $\mathcal{C}_B$ containing U_1 and U_2 respectively. Let U_1^- be the predecessor of U_1 on C_1, and U_2^+ be the successor of U_2 on C_2. (4.10) implies that we can ensure $U_1^- \neq U_2^+$. (However, we may have $C_1 = C_2$.) Corollary 4.10 gives us a shifted walk W_v from U_1^- to U_2^+ traversing at most $2/(\alpha/2) = 4/\alpha$ cycles of $\mathcal{C}_B$. To incorporate v into the walk W_B, recall that W_B traverses all those edges of R_B which lie on cycles from $\mathcal{C}_B$ at least once.

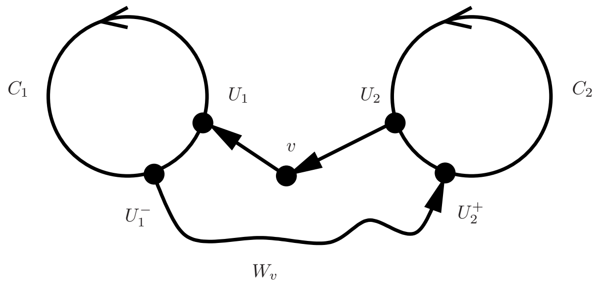


Figure 4.2: Incorporating the exceptional vertex v.

Replace one of the occurences of $U_1^- U_1$ on W_B with the walk

$$W_v' := U_1^- W_v U_2^+ C_2 U_2 v U_1 C_1 U_1,$$

i.e. the walk that goes from U_1^- to U_2^+ along the shifted walk W_v, it then winds once around C_2 but stops in U_2, then it goes to v and further to U_1, and finally it winds around C_1. The walk obtained from W_B by including v in this way is still balanced w.r.t. $\mathcal{C}_B$, i.e. each vertex in R_B is visited the same number of times as every other vertex lying on the same cycle from $\mathcal{C}_B$. We add the extra loop around C_1 because when applying the Blow-up lemma we will need the vertices in V_0' to be at a distance of at least 4 from each other. Using this loop, this can be ensured as follows. After we have incorporated v into W_B we 'ban' all the 6 edges of (the new walk) W_B whose endvertices both have distance at most 3 from v. The extra loop ensures that every edge in each cycle from $\mathcal{C}$ has at least one occurence in W_B which is not banned. (Note that we do not have to add an extra loop around C_2 since if $C_2 \neq C_1$ then all the banned edges of C_2 lie on W_v' but each edge of C_2 also occurs on the original walk W_B.) Thus when incorporating the next exceptional vertex we can always pick an occurence of an edge which is not banned to be replaced by a longer walk. (When incorporating v we picked $U_1^- U_1$.) Repeating this argument, we can incorporate all the exceptional vertices in V_0' into W_B in such a way that all the vertices of V_0' have distance at least 4 on the new walk W_B.

Recall that G_B^c denotes the oriented graph obtained from the pure oriented graph G_B^*

by adding all the V_i-V_j edges for all those pairs V_i, V_j of clusters with $V_iV_j \in E(R_B)$. Let $G^c_{B\cup V'_0}$ denote the graph obtained from G^c_B by adding all the V'_0-B edges of G as well as all the B-V'_0 edges of G. Moreover, recall that the vertices in V'_0 have distance at least 4 from each other on W_B and $|V'_0| \leq \varepsilon n/|R_B| \ll \alpha m_B/20$ by (4.9) and since $m_B|R_B| \approx n/2$ and $\alpha \ll 1$. As already observed at the beginning of Section 4.4, altogether this shows that by winding around each cycle from $\mathcal{C}_B$, one can obtain a Hamilton cycle $C^c_{B\cup V'_0}$ of $G^c_{B\cup V'_0}$ from the walk W_B, provided that W_B visits any cluster $V_i \in R_B$ at most m_B times. To see that the latter condition holds, recall that before we incorporated the exceptional vertices in V'_0 into W_B, each cluster was visited at most $2|R_B|/\alpha$ times. When incorporating an exceptional vertex we replaced an edge of W_B by a walk whose interior visits every cluster at most $4/\alpha + 2 \leq 5/\alpha$ times. Thus the final walk W_B visits each cluster $V_i \in R_B$ at most

$$2|R_B|/\alpha + 5|V'_0|/\alpha \overset{(4.9)}{\leq} 6\varepsilon n/(\alpha|R_B|) \leq \sqrt{\varepsilon} m_B \tag{4.11}$$

times. Hence we have the desired Hamilton cycle $C^c_{B\cup V'_0}$ of $G^c_{B\cup V'_0}$. Note that (4.11) implies that we can choose $C^c_{B\cup V'_0}$ in such a way that for each cycle $C \in \mathcal{C}_B$ there is a subpath P_C of $C^c_{B\cup V'_0}$ which winds around C at least

$$(1 - \sqrt{\varepsilon})m_B \tag{4.12}$$

times in succession.

4.5.3 Applying the Blow-up lemma to find a Hamilton cycle in $G[B \cup V'_0]$

Our next aim is to use the Blow-up lemma to show that $C^c_{B\cup V'_0}$ corresponds to a Hamilton cycle in $G[B \cup V'_0]$. Recall that $k = |R_B|$ and that for each exceptional vertex $v \in V'_0$ the outneighbour U_1 of v on W_B is distinct from its inneighbour U_2 on W_B. We will apply the Blow-up lemma with H being the underlying graph of $C^c_{B\cup V'_0}$ and G^* being the graph obtained from the underlying graph of G_B^* by adding all the vertices $v \in V_0'$ and joining each such v to all the vertices in $N_G^+(v) \cap U_1$ as well as to all the vertices in $N_G^-(v) \cap U_2$. Recall that after applying the Diregularity lemma to obtain the clusters $V_1, \dots, V_k$ we used Proposition 3.10 to ensure that each edge of R_B corresponds to an ε-regular pair of density d (in the underlying graph of G_B^* and thus also in G^*) and that each edge of the union $\bigcup_{C \in \mathcal{C}_B} C \subseteq R_B$ of all the cycles from $\mathcal{C}_B$ corresponds to an (ε, d)-super-regular pair.

V_0' will play the role of V_0 in the Blow-up lemma and we take $L_0, L_1, \dots, L_k$ to be the partition of H induced by $V_0', V_1, \dots, V_k$. $\phi : L_0 \to V_0'$ will be the obvious bijection (i.e. the identity). To define the set $I \subseteq V(H)$ of vertices of distance at least 4 from each other which is used in the Blow-up lemma, let P_C' be the subpath of H corresponding to P_C (for all $C \in \mathcal{C}_B$). For each $i = 1, \dots, k$, let $C_i \in \mathcal{C}_B$ denote the cycle containing V_i and let $J_i \subseteq L_i$ consist of all those vertices in $L_i \cap V(P_{C_i}')$ which have distance at least 4 from the endvertices of P_{C_i}'. Thus in the graph H each vertex $u \in J_i$ has one of its neighbours in the set L_i^- corresponding to the predecessor of V_i on C_i and its other neighbour in the set L_i^+ corresponding to the successor of V_i on C_i. Moreover, all the vertices in J_i have distance at least 4 from all the vertices in L_0 and (4.12) implies that $|J_i| \geq 9m_B/10$. It is easy to see that one can greedily choose a set $I_i \subseteq J_i$ of size $m_B/10$ such that the vertices in $\bigcup_{i=1}^k I_i$ have distance at least 4 from each other. We take $I := L_0 \cup \bigcup_{i=1}^k I_i$.

Let us now check conditions (C1)–(C9). (C1) holds with $K_1 := 1$ since $|L_0| = |V_0'| \leq \varepsilon_A n = \varepsilon n/k \leq d|H|$. (C2) holds by definition of I. (C3) holds since H is a Hamilton cycle in $G^c_{B \cup V_0'}$ (c.f. the definition of the graph $G^c_{B \cup V_0'}$). This also implies that for every edge $xy \in H$ with $x \in L_i, y \in L_j$ $(i, j \geq 1)$ we must have that $V_iV_j \in E(R_B)$. Thus (C6) holds as every edge of R_B corresponds to an ε-regular pair of clusters having density d. (C4) holds with $K_2 := 1$ because

$$|N_H(L_0) \cap L_i| \leq 2\,|L_0| = 2\,|V_0'| \overset{(4.9)}{\leq} 2\varepsilon n/|R_B| \leq 5\varepsilon m_B \leq d m_B.$$

For (C5) we need to find a set $D \subseteq I$ of buffer vertices. Pick any set $D_i \subseteq I_i$ with $|D_i| = \delta' m_B$ and let $D := \bigcup_{i=1}^k D_i$. Since $I_i \subseteq J_i$ we have that $|N_H(D) \cap L_j| = 2\delta' m_B$ for all $j = 1, \dots, k$. Hence

$$||N_H(D) \cap L_i| - |N_H(D) \cap L_j|| = 0$$

for all $1 \le i < j \le k$ and so (C5) holds. (C7) holds with $c := \alpha/10$ by our choice $U_1 \in N^+_{R^*_B}(v)$ and $U_2 \in N^-_{R^*_B}(v)$ of the neighbours of each vertex $v \in V'_0$ in the walk W_B (c.f. the definition of the graph R^*_B).

(C8) and (C9) are now the only conditions we need to check. Given a set $E_i \subseteq V_i$ of size at most $\varepsilon' m_B$, we wish to find $F_i \subseteq (L_i \cap (I \setminus D)) = I_i \setminus D$ and a bijection $\phi_i : E_i \to F_i$ such that every $v \in E_i$ has a large number of neighbours in every cluster V_j for which L_j contains a neighbour of $\phi_i(v)$. Pick any set $F_i \subseteq I_i \setminus D$ of size $|E_i|$. (This can be done since $|D \cap I_i| = \delta' m_B$ and so $|I_i \setminus D| \ge m_B/10 - \delta' m_B \gg \varepsilon' m_B$.) Let $\phi_i : E_i \to F_i$ be an arbitrary bijection. To see that (C8) holds with these choices, consider any vertex $v \in E_i \subseteq V_i$ and let j be such that L_j contains a neighbour of $\phi_i(v)$ in H. Since $\phi_i(v) \in F_i \subseteq I_i \subseteq J_i$, this means that V_j must be a neighbour of V_i on the cycle $C_i \in \mathcal{C}_B$ containing V_i. But this implies that $|N_{G^*}(v) \cap V_j| \ge (d - \varepsilon) m_B$ since each edge of the union $\bigcup_{C \in \mathcal{C}_B} C \subseteq R_B$ of all the cycles from $\mathcal{C}_B$ corresponds to an (ε, d)-super-regular pair in G^*.

Finally, writing $F := \bigcup_{i=1}^k F_i$ we have

$$|N_H(F) \cap L_i| \le 2\varepsilon' m_B.$$

This holds since $F_j \subseteq J_j$ implies that each element of F_j has its two neighbours in H in L_j^+ and L_j^-, so each of $\le \varepsilon' m_B$ elements in F_j contributes at most two to the intersection with a given L_i. Thus (C9) is satisfied with $K_3 := 2$. Hence (C1)–(C9) hold and so we can apply the Blow-up lemma to obtain a Hamilton cycle in G^* such that the image of L_i is V_i for all $i = 1, \dots, k$ and the image of each $x \in L_0$ is $\phi(x) \in V_0$. (Recall that G^* was obtained from the underlying graph of G^*_B by adding all the vertices $v \in V'_0$ and joining each such v to all the vertices in $N^+_G(v) \cap U_1$ as well as to all the vertices in $N^-_G(v) \cap U_2$, where U_1 and U_2 are the neighbours of v on the walk W_B.) Using the fact that H was obtained from the (directed) Hamilton cycle $C^c_{B\cup V'_0}$ and since $U_1 \neq U_2$ for each $v \in V'_0$, it is easy to see that our Hamilton cycle in G^* corresponds to a (directed) Hamilton cycle C_B in $G[B \cup V'_0]$.

4.5.4 Finding a Hamilton cycle in G

The last step of the proof is to find a Hamilton cycle in $G[A']$ which can be connected with C_B into a Hamilton cycle of G, recalling that $A = V'_1 \cup \ldots \cup V'_\ell$. Pick an arbitrary edge v_1v_2 on C_B and add an extra vertex v^* to $G[A']$ with outneighbourhood $N^+_G(v_1) \cap A'$ and inneighbourhood $N^-_G(v_2) \cap A'$. A Hamilton cycle C_A in the digraph thus obtained from $G[A']$ can be extended to a Hamilton cycle of G by replacing v^* with $v_2C_Bv_1$. To find such a Hamilton cycle C_A, we can argue as before. This time, there is only one exceptional vertex, namely v^*, which we incorporate into the walk W_A. Note that by our choice of A and B the analogue of (4.10) is satisfied and so this can be done as before. We then use the Blow-up lemma to obtain the desired Hamilton cycle C_A corresponding to this walk.

4.6 Ore-type Condition

The following observation guarantees that every oriented graph as in Theorem 4.4 has large minimum semidegree.

Fact 4.12. *Suppose that $0 < \alpha < 1$ and that G is an oriented graph such that $d^+(x) + d^-(y) \geq (3/4 + \alpha)|G|$ whenever $xy \notin E(G)$. Then $\delta^0(G) \geq |G|/8 + \alpha|G|/2$.*

Proof. Suppose not. We may assume that $\delta^+(G) \leq \delta^-(G)$. Pick a vertex x with $d^+(x) = \delta^+(G)$. Let Y be the set of all those vertices y with $xy \notin E(G)$. Thus $|Y| \geq$

$7|G|/8 - \alpha|G|/2$. Moreover, $d^-(y) \geq (3/4 + \alpha)|G| - d^+(x) \geq 5|G|/8 + \alpha|G|/2$. Hence $e(G) \geq |Y|(5|G|/8 + \alpha|G|/2) > 35|G|^2/64$, a contradiction. □

The proof of Theorem 4.4 is similar to that of Theorem 4.3. Fact 4.12 and Lemma 3.2 together imply that the reduced oriented graph R_A (and similarly R_B) has minimum semidegree at least $|R|/8$ and it inherits the Ore-type condition from G (i.e. it satisfies condition (d) of Lemma 3.2 with $c = 3/4 + \alpha$). Together with Lemma 4.13 below (which is an analogue of Lemma 4.8) this implies that R_A (and R_B as well) is an expander in the sense that $|N^+(X)| \geq |X| + \alpha|R_A|/2$ for all $X \subseteq V(R_A)$ with $|X| \leq (1-\alpha)|R_A|$. In particular, R_A (and similarly R_B) has a 1-factor: To see this, note that the above expansion property together with Fact 4.12 imply that for any $X \subseteq V(R_A)$, we have $|N^+_{R_A}(X)| \geq |X|$. Together with Hall's theorem, this means that the following bipartite graph H has a perfect matching: the vertex classes W_1, W_2 are 2 copies of $V(R_A)$ and we have an edge in H between $w_1 \in W_1$ and $w_2 \in W_2$ if there is an edge from w_1 to w_2 in R_A. But clearly a perfect matching in H corresponds to a 1-factor in R_A. Using these facts, one can now argue precisely as in the proof of Theorem 4.3.

Lemma 4.13. *Suppose that $0 < \varepsilon \ll \alpha \ll 1$. Let R^* be an oriented graph on N vertices and let U be a set of at most εN^2 ordered pairs of vertices of R^*. Suppose that $d^+(x) + d^-(y) \geq (3/4 + \alpha)N$ for all $xy \notin E(R^*) \cup U$. Then any $X \subseteq V(R^*)$ with $\alpha N \leq |X| \leq (1-\alpha)N$ satisfies $|N^+(X)| \geq |X| + \alpha N/2$.*

Proof. The proof is similar to that of Lemma 4.8. Suppose that Lemma 4.13 does not hold and let $X \subseteq V(R^*)$ with $\alpha N \leq |X| \leq (1-\alpha)N$ be such that

$$\left|N^+(X)\right| < |X| + \alpha N/2. \tag{4.13}$$

Call a vertex of R^* *good* if it lies in at most $\sqrt{\varepsilon}N$ pairs from U. Thus all but at most $2\sqrt{\varepsilon}N$ vertices of R^* are good. As in the proof of Lemma 4.8 we consider the following partition of $V(R^*)$:

$$A := X \cap N^+(X), \quad B := N^+(X)\backslash X, \quad C := V(R^*)\backslash(X \cup N^+(X)), \quad D := X\backslash N^+(X).$$

(4.13) implies

$$|D| + \alpha N/2 > |B|. \tag{4.14}$$

Suppose first that $|D| > 2\sqrt{\varepsilon}N$. It is easy to see that there are vertices $x \neq y$ in D such that $xy, yx \notin U$. Since no edge of R^* lies within D we have $xy, yx \notin E(R^*)$ and so $d(x) + d(y) \geq 3N/2 + 2\alpha N$. In particular, at least one of x, y has degree at least $3N/4 + \alpha N$. But then

$$|A| + |B| + |C| \geq 3N/4 + \alpha N. \tag{4.15}$$

If $|D| \leq 2\sqrt{\varepsilon}N$ then $|A|+|B|+|C| \geq N-|D|$ and so (4.15) still holds with room to spare. Note that (4.14) and (4.15) together imply that $2|A| + 2|C| \geq 3N/2 + 2\alpha N - 2|B| \geq 3N/2 - |B| - |D| \geq N/2$. Thus at least one of A, C must have size at least $N/8$. In particular, this implies that one of the following 3 cases holds.

Case 1. $|A|, |C| > 2\sqrt{\varepsilon}N$.

Let A' be the set of all good vertices in A. By an averaging argument there exists $x \in A'$ with $|N^+(x) \cap A'| < |A'|/2$. Since $N^+(A) \subseteq A \cup B$ this implies that $|N^+(x)| < |B| + |A \setminus A'| + |A'|/2$. Let $C' \subseteq C$ be the set of all those vertices $y \in C$ with $xy \notin U$. Thus $|C \setminus C'| \leq \sqrt{\varepsilon}N$ since x is good. By an averaging argument there exists $y \in C'$ with $|N^-(y) \cap C'| < |C'|/2$. But $N^-(C) \subseteq B \cup C$ and so $|N^-(y)| < |B| + |C \setminus C'| + |C'|/2$. Moreover, $d^+(x) + d^-(y) \geq 3N/4 + \alpha N$ since $xy \notin E(R^*) \cup U$. Altogether this shows that

$$|A'|/2 + |C'|/2 + 2|B| \geq d^+(x) + d^-(y) - |A \setminus A'| - |C \setminus C'| \geq 3N/4 + \alpha N/2.$$

Together with (4.15) this implies that $3|A| + 6|B| + 3|C| \geq 3N + 3\alpha N$, which in turn together with (4.14) yields $3|A| + 3|B| + 3|C| + 3|D| \geq 3N + 3\alpha N/2$, a contradiction.

Case 2. $|A| > 2\sqrt{\varepsilon}N$ and $|C| \le 2\sqrt{\varepsilon}N$.

As in Case 1 we let A' be the set of all good vertices in A and pick $x \in A'$ with $|N^+(x)| < |B| + |A \setminus A'| + |A'|/2$. Note that (4.14) implies that $|D| > N - |X| - |C| - \alpha N/2 \ge \sqrt{\varepsilon}N$. Pick any $y \in D$ such that $xy \notin U$. Then $xy \notin E(R^*)$ since R^* contains no edges from A to D. Thus $d^+(x) + d^-(y) \ge 3N/4 + \alpha N$. Moreover, $N^-(y) \subseteq B \cup C$. Altogether this gives

$$|A'|/2 + 2|B| \ge d^+(x) + d^-(y) - |A \setminus A'| - |C| \ge 3N/4 + \alpha N/2.$$

As in Case 1 one can combine this with (4.15) and (4.14) to get a contradiction.

Case 3. $|A| \le 2\sqrt{\varepsilon}N$ and $|C| > 2\sqrt{\varepsilon}N$.

This time we let C' be the set of all good vertices in C and pick $y \in C'$ with $|N^-(y) \cap C'| < |C'|/2$. Hence $|N^-(y)| < |B| + |C \setminus C'| + |C'|/2$. Moreover, we must have $|D| = |X| - |A| > \sqrt{\varepsilon}N$. Pick any $x \in D$ such that $xy \notin U$. Then $xy \notin E(R^*)$ since R^* contains no edges from D to C. Thus $d^+(x) + d^-(y) \ge 3N/4 + \alpha N$. Moreover, $N^+(x) \subseteq A \cup B$. Altogether this gives

$$|C'|/2 + 2|B| \ge d^+(x) + d^-(y) - |A| - |C \setminus C'| \ge 3N/4 + \alpha N/2,$$

which in turn yields a contradiction as before. □

CHAPTER 5

SHORT CYCLES

5.1 Introduction

5.1.1 Cycles of Given Length in Oriented Graphs

A central problem in digraph theory is the Caccetta-Häggkvist conjecture [18] (which generalised an earlier conjecture of Behzad, Chartrand and Wall [8]):

Conjecture 5.1. *An oriented graph on n vertices with minimum outdegree d contains a cycle of length at most $\lceil n/d \rceil$.*

Note that in Conjecture 5.1 it does not matter whether we consider oriented graphs or general digraphs. Chvátal and Szemerédi [22] showed that a minimum outdegree of at least d guarantees a cycle of length at most $\lceil 2n/(d+1) \rceil$. For most values of n and d, this is improved by a result of Shen [68], which guarantees a cycle of length at most $3\lceil 0.44n/d \rceil$. Chvátal and Szemerédi [22] also showed that Conjecture 5.1 holds if we increase the bound on the cycle length by adding a constant c. They showed that $c := 2500$ will do. Nishimura [65] refined their argument to show that one can take $c := 304$. The next result of Shen gives the best known constant.

Theorem 5.2 (Shen [67])**.** *An oriented graph on n vertices with minimum outdegree d contains a cycle of length at most $\lceil n/d \rceil + 73$.*

The special case of Conjecture 5.1 that has attracted most interest is when $d = \lceil n/3 \rceil$. Here the conjecture is that a minimum outdegree of $\lceil n/3 \rceil$ implies a cycle of length 3, that is, a directed triangle. The following bound towards this case improves an earlier one of Caccetta and Häggkvist [18].

Theorem 5.3 (Shen [67])**.** *If G is any oriented graph on n vertices with $\delta^+(G) \geq 0.355n$ then G contains a directed triangle.*

If one considers the *minimum semi-degree* $\delta^0(G) := \min\{\delta^+(G), \delta^-(G)\}$ instead of the minimum outdegree $\delta^+(G)$, then the constant can be improved slightly. The best known value for the constant in this case is currently 0.346 [43]. See the monograph [5] or the survey [64] for further partial results on Conjecture 5.1.

We consider the natural and related question of which minimum semi-degree forces cycles of length exactly $\ell \geq 4$ in an oriented graph. We will often refer to cycles of length ℓ as *ℓ-cycles*. Our main result answers this question completely when ℓ is not a multiple of 3.

Theorem 5.4. *Let $\ell \geq 4$. If G is an oriented graph on $n \geq 10^{10}\ell$ vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then G contains an ℓ-cycle. Moreover for any vertex $u \in V(G)$ there is an ℓ-cycle containing u.*

The extremal example showing this to be best possible for $\ell \geq 4$, $\ell \not\equiv 0 \mod 3$ is given by the blow-up of a 3-cycle. More precisely, let G be the oriented graph on n vertices formed by dividing $V(G)$ into 3 vertex classes V_1, V_2, V_3 of as equal size as possible and adding all possible edges from V_i to V_{i+1}, counting modulo 3. Then this oriented graph contains no ℓ-cycle and has minimum semi-degree $\lfloor n/3 \rfloor$.

Also, for all those $\ell \geq 4$ which are multiples of 3, the 'moreover' part is best possible for infinitely many n. To see this, consider the modification of the above example formed by deleting a vertex from the largest vertex class and adding an extra vertex u with $N^+(u) = V_2$ and $N^-(u) = V_1$. This gives an oriented graph with minimum semi-degree $\lfloor (n-1)/3 \rfloor$. For $\ell \equiv 0 \mod 3$ it contains no ℓ-cycle through u.

Perhaps surprisingly, we can do much better than Theorem 5.4 for some cycle lengths (if we do not ask for a cycle through a given vertex). Indeed, we conjecture that the correct bounds are those given by the obvious extremal example: when we seek an ℓ-cycle, the extremal example is probably the blow-up of a k-cycle, where $k \geq 3$ is the smallest integer which is not a divisor of ℓ.

Conjecture 5.5. *Let $\ell \geq 4$ be a positive integer and let $k \geq 3$ be minimal such that k does not divide ℓ. Then there exists an integer $n_0 = n_0(\ell)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lfloor n/k \rfloor + 1$ contains an ℓ-cycle.*

It is easy to see that the only values of k that can appear in Conjecture 5.5 are of the form $k = p^s$ with $k \geq 3$, where $p \geq 2$ is a prime and s a positive integer. Theorem 5.4 confirms this conjecture in the case when $k = 3$. The following result implies that Conjecture 5.5 is approximately true when $k = 4, 5$ and ℓ is sufficiently large. It also gives weaker bounds on the minimum semi-degree for large values of k.

Theorem 5.6. *Let $\ell \geq 4$ be a positive integer and let $k \geq 3$ be minimal such that k does not divide ℓ.*

(i) *There exists an integer $n_0 = n_0(\ell)$ such that whenever $k \geq 150$ and G is an oriented graph on $n \geq n_0$ vertices with $\delta^+(G) \geq n/k + 150n/k^2$ then G contains an ℓ-cycle.*

(ii) *If $k = 4$ and $\ell \geq 42$ then for every $\varepsilon > 0$ there exists an integer $n_0 = n_0(\ell, \varepsilon)$ such that every oriented graph G on $n \geq n_0$ vertices with $\delta^0(G) \geq n/k + \varepsilon n$ contains an ℓ-cycle.*

(iii) *The analogue of (ii) holds if $k = 5$ and $\ell \geq 2550$.*

Part (i) is obtained from Theorem 5.2 via a simple application of the Regularity lemma for digraphs (see Section 5.3). It would be interesting to find a proof which does not rely on the Regularity lemma. Moreover, part (i) suggests that one might be able to replace δ^0 by δ^+ in Conjecture 5.5. Even replacing it in Theorem 5.4 would be interesting.

In view of Theorem 5.4 and the Caccetta-Häggkvist Conjecture one might wonder whether a minimum semi-degree close to $n/3$ also forces a 3-cycle through any given vertex. However the next proposition (whose straightforward proof is given in Section 5.2) shows that the threshold in this case is much higher.

Proposition 5.7.

(i) If G is an oriented graph on n vertices with $\delta^0(G) \geq \lceil 2n/5 \rceil$ then for any vertex $u \in V(G)$ there exists a 3-cycle containing u.

(ii) For infinitely many n there exists an oriented graph G on n vertices with $\delta^0(G) = \lfloor 2n/5 \rfloor$ containing a vertex u which does not lie on a 3-cycle.

5.1.2 Arbitrary orientations of cycles

It is natural to ask whether these results still hold if we ask for arbitrary orientations of short cycles. It appears that the semi-degree required depends on the so-called cycle-type. Given an arbitrarily oriented ℓ-cycle C, the *cycle-type* $t(C)$ of C is the number of edges oriented forwards in C minus the number of edges oriented backwards in C. By traversing C in the opposite direction if necessary, we may assume that $t(C) \geq 0$. An oriented ℓ-cycle has cycle-type ℓ. Arbitrarily oriented cycles of cycle-type 0 are precisely those for which there is a digraph homomorphism into an oriented path. (A *digraph homomorphism* is a mapping between digraphs which sends edges to edges.) Moreover, if $t(C) \geq 3$ then $t(C)$ is the *maximum* length of an oriented cycle into which there is a digraph homomorphism of C.

Proposition 5.8.

- *Let $\ell \geq 4$ and let $\alpha > 0$. Then there exists $n_0 = n_0(\ell, \alpha)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (1/3 + \alpha)n$ contains every orientation of an ℓ-cycle.*

- *Let $\alpha > 0$ and let ℓ be some positive constant. Then there exists $n_0 = n_0(\alpha, \ell)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \alpha n$ contains every cycle of length at most ℓ and cycle-type 0.*

This result is proved in Section 5.4. Conjecture 5.5 has a natural strengthening to incorporate arbitrarily oriented cycles.

Conjecture 5.9. *Let C be an arbitrarily oriented cycle of length $\ell \geq 4$ and cycle-type $t(C) \geq 4$. Let k be the smallest integer which is greater than 2 and does not divide $t(C)$. Then there exists an integer $n_0 = n_0(\ell, k)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lfloor n/k \rfloor + 1$ contains C.*

As we shall see in Section 5.4, Conjecture 5.5 would imply an approximate version of Conjecture 5.9.

5.1.3 Cycles of Given Length in Digraphs

A straightforward application of the Regularity lemma shows that a solution to Conjecture 5.5 would also asymptotically give a result for general digraphs: Let $\delta_{\text{di}}(\ell, n)$ denote the smallest integer d such that every digraph with n vertices and minimum semi-degree at least d contains an ℓ-cycle and let $\delta_{\text{orient}}(\ell, n)$ denote the smallest integer d so that every oriented graph with n vertices and minimum semi-degree at least d contains an ℓ-cycle.

Proposition 5.10. *For any $\ell \geq 3$,*

$$\lim_{n\to\infty} \frac{\delta_{di}(\ell, n)}{n} = \begin{cases} 1/2 & \text{if } \ell \text{ is odd;} \\ \lim_{n\to\infty} \frac{\delta_{orient}(\ell,n)}{n} & \text{otherwise.} \end{cases}$$

It is easy to see that these limits exist.[1] We will prove Proposition 5.10 in Section 5.3. The corresponding density problem for digraphs was solved by Häggkvist and Thomassen.

[1]Suppose for example that $\lim_{n\to\infty} \delta_{\text{orient}}(\ell, n)/n$ does not exist. Then there is an $\varepsilon > 0$ such that for every $n' \in \mathbb{N}$ there exist $n_2 > n_1 \geq n'$ with $c_2 := \delta_{\text{orient}}(\ell, n_2)/n_2 \geq \delta_{\text{orient}}(\ell, n_1)/n_1 + \varepsilon =: c_1 + \varepsilon$. Let G_2 be any oriented graph on n_2 vertices with $\delta^0(G_2) \geq c_2 n_2 - 1$ (say) which does not contain an ℓ-cycle. Pick a random set $X \subseteq V(G_2)$ of size n_1. Then $G_2[X]$ has minimum semidegree at least $(c_2 - \varepsilon/2)n_1$, contradicting the fact that $\delta_{\text{orient}}(\ell, n_1)/n_1 = c_1$.

Let $\mathrm{ex}_{\mathrm{di}}(\ell, n)$ denote the largest number d so that there is digraph with n vertices and at least d edges which contains no ℓ-cycle. Häggkvist and Thomassen [40] proved that

$$\mathrm{ex}_{\mathrm{di}}(\ell, n) = \binom{n}{2} + \frac{(\ell-2)n}{2}. \tag{5.1}$$

The case $\ell = 3$ was proved earlier by Brown and Harary [15]. A transitive tournament (i.e. an acyclic orientation of a complete graph) shows that it does not make sense to consider this density problem for oriented graphs. More general extremal digraph problems are discussed in the surveys [16, 55].

5.2 Proofs of Theorem 5.4 and Proposition 5.7

We begin with two immediate facts about oriented graphs which will prove very useful.

Fact 5.11. *If G is an oriented graph and $X \subseteq V(G)$ is non-empty then $e(X) \leq |X|(|X|-1)/2$. In particular, there exists $x \in X$ with $|N^+(x) \cap X| \leq |X|/2 - 1/2$ and thus $|N^+(X) \setminus X| \geq |N^+(x) \setminus X| \geq \delta^0(G) - |X|/2 + 1/2$.* □

Fact 5.12. *If G is an oriented graph on n vertices then the maximum size of an independent set is at most $n - 2\delta^0(G)$.* □

Proof of Proposition 5.7. First we prove (i). By Fact 5.11 there exists a vertex $x \in N^+(u)$ with

$$|N^+(x) \setminus N^+(u)| \geq \delta^0(G) - |N^+(u)|/2 + 1/2.$$

Hence

$$|N^+(u)| + |N^-(u)| + |N^+(x) \setminus N^+(u)| \geq 5\delta^0(G)/2 + 1/2 > n$$

and so x must have an outneighbour in $N^-(u)$.

For (ii), pick $m \in \mathbb{N}$ and define an oriented graph G on $n := 5m-1$ vertices as follows. Let A, B, C be disjoint vertex sets of sizes $2m-1$, $2m-1$ and m respectively. Add all possible edges from A to B, B to C and C to A. Let $G[A]$ and $G[B]$ induce regular tournaments. So for example every vertex in A will have $m-1$ outneighbours and $m-1$ inneighbours in A. Add a single vertex u with $N^+(u) := B$ and $N^-(u) := A$. Then $\delta^0(G) = 2m-1 = \lfloor 2n/5 \rfloor$. By construction u is not contained in a 3-cycle. □

We now prove Theorem 5.4 in a series of lemmas. Lemmas 5.13, 5.14 and 5.16 deal with the special cases $\ell = 4, 5, 6$. Lemmas 5.17 and 5.18 deal with the general case $\ell \geq 7$.

Lemma 5.13. *If G is an oriented graph on $n \geq 4$ vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then for any vertex $x \in V(G)$, G contains a 4-cycle through x.*

Proof. Assume that there is a vertex $x \in V(G)$ for which no such cycle exists. Let X be a set of $\lfloor n/3 \rfloor + 1$ outneighbours of x and Y be a set of $\lfloor n/3 \rfloor + 1$ inneighbours. Suppose that both of the following hold.

(i) There exists $x' \in X$ with $|N^+(x') \setminus (X \cup Y)| \geq (\lfloor n/3 \rfloor + 1)/2$.

(ii) There exists $y' \in Y$ with $|N^-(y') \setminus (X \cup Y)| \geq (\lfloor n/3 \rfloor + 1)/2$.

Then

$$(N^+(x') \cap N^-(y')) \setminus (X \cup Y) \neq \emptyset$$

and hence the desired 4-cycle exists. So without loss of generality assume that (i) does not hold. (The case when (ii) does not hold is similar.) Let X' be the set of vertices $x' \in X$ with $d^-_X(x') > 0$. Note that Fact 5.12 implies that $X' \neq \emptyset$. Let $x' \in X'$ be such that $d^+_{X'}(x')$ is minimal. Since $N^+(x') \cap (X \setminus X') = \emptyset$, Fact 5.11 implies that

$$|N^+(x') \setminus X| = |N^+(x') \setminus X'| \geq \delta^0(G) - |X'|/2 \geq \delta^0(G) - |X|/2 \geq (\lfloor n/3 \rfloor + 1)/2.$$

Since we are assuming that (i) does not hold this means that x' has an outneighbour $y \in Y$. By definition of X' there exists an inneighbour $x'' \in X$ of x'. But then $xx''x'y$ is the required 4-cycle. □

Lemma 5.14. *If G is an oriented graph on $n \geq 5$ vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then for any vertex $x \in V(G)$, G contains a 5-cycle through x.*

Proof. As $N^-(x)$ is not independent by Fact 5.12 we can pick vertices $a, y \in N^-(x)$ such that $ya, ax, yx \in E(G)$. Let X be a set of $\lfloor n/3 \rfloor + 1$ outneighbours of x and Y be a set of $\lfloor n/3 \rfloor + 1$ inneighbours of y. Define $Z := X \cap Y$. Clearly, it suffices to prove the next claim.

Claim 1. *There exists at least one of the following:*

(i) an x-y path of length 4,

(ii) an x-y path of length 3 avoiding a.

Note that $x, y, a \notin X \cup Y$ since G is an oriented graph. So we may assume that $e(X, Y) = 0$, as otherwise (ii) is satisfied. In particular, Z is independent and $e(X, Z) = e(Z, Y) = 0$. The following claim immediately implies (i) (to see this, note that $x, y \notin N^+(x') \cap N^-(y')$).

Claim 2. *Both of the following hold.*

(a) There exists $x' \in X$ with $|N^+(x') \setminus (X \cup Y)| \geq (\lfloor n/3 \rfloor + 1 + |Z|)/2$.

(b) There exists $y' \in Y$ with $|N^-(y') \setminus (X \cup Y)| \geq (\lfloor n/3 \rfloor + 1 + |Z|)/2$.

We will only prove (a) (the argument for (b) is similar). If $X \setminus Z = \emptyset$ then $X = Z$ and so X is independent. But $|X| = \lfloor n/3 \rfloor + 1$ which contradicts Fact 5.12. So assume that $X \setminus Z \neq \emptyset$ and let $x' \in X \setminus Z$ be such that $d^+_{X \setminus Z}(x')$ is minimal. Fact 5.11 implies that

$$d^+_{\overline{X \setminus Z}}(x') > \delta^0(G) - (|X| - |Z|)/2 \geq (\lfloor n/3 \rfloor + 1 + |Z|)/2.$$

By assumption x' has no outneighbours in Y, so $d^+_{\overline{X \setminus Z}}(x') = d^+_{\overline{X \cup Y}}(x')$ and thus (a) holds. (Recall that for $X \subseteq V(G)$, $\overline{X} := V(G) \setminus X$.) □

In order to prove the cases $\ell = 6$ and $\ell \geq 7$ of Theorem 5.4 we need some more notation. An *xy-butterfly* is an oriented graph with vertices x, y, z, a, b such that xa, xz, az, zb, zy, by are all the edges (Figure 5.1). The crucial fact about a butterfly is that it contains x-y paths of lengths 2, 3 and 4, and is thus a useful tool in finding cycles

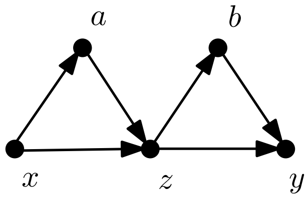


Figure 5.1: An xy-butterfly

of prescribed length: any y-x path of length $\ell - 2$, $\ell - 3$ or $\ell - 4$ whose interior avoids the xy-butterfly yields an ℓ-cycle containing x. The following fact tells us that a large minimum semi-degree guarantees the existence of a butterfly.

Fact 5.15. *If G is an oriented graph on n vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then for any vertex $x \in V(G)$ there exists a vertex y such that G contains an xy-butterfly.*

Proof. By Fact 5.12 the outneighbourhood of x is not independent, so pick an edge az in it. Reapply Fact 5.12 to find an edge by in the outneighbourhood of z. Note that as $x, a \in N^-(z)$ all the vertices are distinct. □

Lemma 5.16. *If G is an oriented graph on $n \geq 6$ vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then for any vertex $x \in V(G)$, G contains a 6-cycle through x.*

Proof. Fact 5.15 gives us an xy-butterfly for some vertex $y \in V(G)$, with vertices a, b, z as described in the definition of an xy-butterfly. To complete the proof we may assume that each of the following holds.

(i) There is no y-x path of length 2.

(ii) There is no y-x path of length 3 avoiding a.

(iii) There is no y-x path of length 4 avoiding z.

Indeed, it is easy to check that if one of these does not hold then this y-x path together with a suitable subpath of the xy-butterfly forms the required cycle.

Pick $Y \subseteq N^+(y) \setminus \{a, x\}$, $X \subseteq N^-(x) \setminus \{y\}$ such that $|Y| = \lfloor n/3 \rfloor - 1$ and $|X| = \lfloor n/3 \rfloor$. Observe that $b, z \notin Y$ and $a, z \notin X$. Moreover $X \cap Y = \emptyset$ by (i). Let $Y' := N^+(Y) \setminus Y$,

$X' := N^-(X) \setminus X$. Then $X \cap Y' = \emptyset$ and $Y \cap X' = \emptyset$ by (ii). Fact 5.11 implies that $|Y'| \geq \lfloor n/3 \rfloor/2 + 2$ and $|X'| \geq \lfloor n/3 \rfloor/2 + 3/2$. By (i) and the definitions of X and Y we have $x, y \notin X, Y, X', Y'$. Altogether this shows that

$$n + |X' \cap Y'| \geq |X| + |Y| + |X'| + |Y'| + 2 \geq 3\lfloor n/3 \rfloor + 9/2 \geq (n-2) + 9/2.$$

Hence $|X' \cap Y'| \geq 3$, and so $(X' \cap Y') \setminus \{z\} \neq \emptyset$. But this implies that there is a y-x path of length 4 avoiding z. □

The next two lemmas deal with the case $\ell \geq 7$.

Lemma 5.17. *Let C be some positive integer. If G is an oriented graph on $n \geq 8 \cdot 10^9 C$ vertices with $\delta^0(G) \geq n/3 - C + 1$ then for every pair $x \neq y$ of vertices there exists an x-y path of length 3, 4 or 5.*

Proof. Let $\varepsilon := 1/10^4$ and $C' := 10C/\varepsilon$. Let X be a set of $\lceil n/3 \rceil - 2C$ outneighbours of x in $G - y$ and let Y be a set of $\lceil n/3 \rceil - 2C$ inneighbours of y in $G - x$, chosen so that $|X \setminus Y|, |Y \setminus X| \geq C$. Let $Z := X \cap Y$. If there is an X-Y edge then we have an x-y path of length 3. So suppose there is no such edge. In particular this implies that Z is independent and there are no X-Z or Z-Y edges.

Let $X' := N^+(X \setminus Z) \setminus X$ and $Y' := N^-(Y \setminus Z) \setminus Y$. Note that $X' \cap Y = \emptyset$ and $Y' \cap X = \emptyset$, as otherwise we have an X-Y edge. Moreover, we may assume that $X' \cap Y' = \emptyset$, as otherwise we have an x-y path of length 4. As no vertex in $X \setminus Z$ has an outneighbour in Z we have $X' = N^+(X \setminus Z) \setminus (X \setminus Z)$. Hence by Fact 5.11

$$|X'| \geq \delta^0(G) - |X \setminus Z|/2 \geq \lceil n/3 \rceil/2 + |Z|/2.$$

Similarly, $|Y'| \geq \lceil n/3 \rceil/2 + |Z|/2$. Observe that this implies

$$|V(G) \setminus ((X \cup Y) \cup (X' \cup Y'))| = n - 2(\lceil n/3 \rceil - 2C) + |Z| - 2(\lceil n/3 \rceil/2 + |Z|/2) \leq 4C. \quad (5.2)$$

Note that

$$|X'| \leq n - |X \cup Y| - |Y'| \leq n - (2n/3 - |Z| - 4C) - (n/6 + |Z|/2) = n/6 + |Z|/2 + 4C. \quad (5.3)$$

We call a vertex $x' \in X \setminus Z$ *good* if $|N^+(x') \setminus X| \geq n/6 + |Z|/2 - C' \geq |X'| - 4C'/3$ (the last inequality follows from (5.3)). Suppose that at least $\varepsilon|X \setminus Z|$ vertices in $X \setminus Z$ are not good. Since $d^+_{X\setminus Z}(x') \geq \delta^0(G) - |N^+(x') \setminus (X \setminus Z)| = \delta^0(G) - |N^+(x') \setminus X|$ for every $x' \in X \setminus Z$ this implies that

$$\begin{aligned}
e(X \setminus Z) &\geq \varepsilon|X \setminus Z|(\delta^0(G) - (n/6 + |Z|/2 - C')) + (1 - \varepsilon)|X \setminus Z|(\delta^0(G) - |X'|) \\
&\overset{(5.3)}{\geq} \varepsilon|X \setminus Z|(n/6 - |Z|/2 + C'/2) + (1 - \varepsilon)|X \setminus Z|(n/6 - |Z|/2 - 5C) \\
&= |X \setminus Z|(n/6 - |Z|/2 + \varepsilon C'/2 - 5C(1 - \varepsilon)) \\
&\geq |X \setminus Z|(n/6 - |Z|/2) \geq |X \setminus Z|^2/2.
\end{aligned}$$

But this is a contradiction as G is an oriented graph. Thus we may assume that all but at most $\varepsilon|X \setminus Z|$ vertices in $X \setminus Z$ are good, and hence, since $|X'| \geq n/6 \geq 4C'/(3\varepsilon)$ we have

$$e(X \setminus Z, X') \geq (1 - \varepsilon)|X \setminus Z|(|X'| - 4C'/3) \geq (1 - 2\varepsilon)|X \setminus Z||X'|. \quad (5.4)$$

Call a vertex $x' \in X'$ *nice* if $|N^-(x') \cap (X \setminus Z)| \geq (1 - 2\sqrt{\varepsilon})|X \setminus Z|$. Then at least $(1 - 2\sqrt{\varepsilon})|X'|$ vertices in X' are nice, as otherwise

$$e(X \setminus Z, X') \leq 2\sqrt{\varepsilon}|X'|(1 - 2\sqrt{\varepsilon})|X \setminus Z| + (1 - 2\sqrt{\varepsilon})|X'||X \setminus Z| < (1 - 2\varepsilon)|X'||X \setminus Z|,$$

which contradicts (5.4). Consider a nice vertex $x' \in X' \setminus \{y\}$. Note that $N^+(x') \cap (Y \cup Y')$ is either empty or equal to $\{x\}$ (as otherwise we get an x-y path of length 4 or 5). Since x'

is nice it has at most $2\sqrt{\varepsilon}|X \setminus Z|$ outneighbours in $X \setminus Z$ and so

$$|N^+(x') \cap X'| \overset{(5.2)}{\geq} \delta^0(G) - 2\sqrt{\varepsilon}|X \setminus Z| - 1 - 4C \geq n/3 - \sqrt{\varepsilon}n. \tag{5.5}$$

In particular, $|X'| \geq n/3 - \sqrt{\varepsilon}n$. Similarly, $|Y'| \geq n/3 - \sqrt{\varepsilon}n$. But $|X \cup Y| \geq n/3 - C$ and so

$$|X'| \leq n - |X \cup Y| - |Y'| \leq n/3 + 2\sqrt{\varepsilon}n. \tag{5.6}$$

Now we combine this with the fact that at least $|X'| - 1 - 2\sqrt{\varepsilon}|X|' \geq (1 - 3\sqrt{\varepsilon})|X'|$ vertices in $X' \setminus \{y\}$ are nice to obtain

$$|X'|^2/2 \geq e(X') \overset{(5.5)}{\geq} (1-3\sqrt{\varepsilon})|X'|(n/3-\sqrt{\varepsilon}n) \overset{(5.6)}{\geq} (1-3\sqrt{\varepsilon})|X'|(|X'|-3\sqrt{\varepsilon}n) > 2|X'|^2/3.$$

This contradiction completes the proof. □

Lemma 5.18. *Suppose $\ell \geq 7$ and $n \geq 10^{10}\ell$. If G is an oriented graph on n vertices with $\delta^0(G) \geq \lfloor n/3 \rfloor + 1$ then for every vertex $x \in V(G)$, G contains an ℓ-cycle through x.*

Proof. Fact 5.15 gives us an xy-butterfly for some vertex $y \in V(G)$, with a, b and z as in the definition of an xy-butterfly. Greedily pick a path P of length $\ell - 7$ from y to some vertex v such that P avoids a, b, x, z (the minimum semi-degree condition implies the existence of such a path).

Now apply Lemma 5.17 to $G - (\{a, b, z\} \cup (V(P) \setminus \{v\}))$ with $C := \ell$ (say) to find a v-x path of length 3, 4 or 5. Pick a path from x to y in the xy-butterfly of appropriate length to obtain the desired ℓ-cycle through x. □

5.3 Proofs of Theorem 5.6 and Proposition 5.10

The following lemma implies that if we allow ourselves a linear 'error term' in the degree conditions then instead of finding an ℓ-cycle, it suffices to look for a closed walk of length ℓ. We will use (i) and (ii) in the proof of Theorem 5.6 and (iii) in the proof of Proposition 5.10.

The proof of this lemma is a standard application of the Regularity lemma. As mentioned in the introduction to this chapter, it would be interesting to find a proof which avoids the Regularity lemma. This would probably yield a much better bound on n_1.

Lemma 5.19. *Let $\ell \geq 2$ be an integer.*

(i) *Suppose that $c > 0$ and there exists an integer n_0 such that every oriented graph H on $n \geq n_0$ vertices with $\delta^0(H) \geq cn$ contains a closed walk of length ℓ. Then for each $\varepsilon > 0$ there exists $n_1 = n_1(\varepsilon, \ell, n_0)$ such that if G is an oriented graph on $n \geq n_1$ vertices with $\delta^0(G) \geq (c + \varepsilon)n$ then G contains an ℓ-cycle.*

(ii) *The analogue holds if we replace $\delta^0(H)$ by $\delta^+(H)$ and $\delta^0(G)$ by $\delta^+(G)$.*

(iii) *The analogue of (i) also holds if we consider directed graphs instead of oriented graphs.*

Proof. We only consider (i). (The arguments for the remaining parts are similar.) Apply the degree form of the Diregularity lemma (Lemma 3.1) and Lemma 3.2 to G to obtain a partition of $V(G)$ into clusters and a reduced oriented graph R. By Lemma 3.2 R almost inherits the minimum semi-degree of G, i.e. $\delta^0(R) \geq (c + \varepsilon/2)|R|$. Applying our assumption with $H := R$ gives a closed walk of length ℓ in R. Since n_1 is large compared to ℓ, this also holds for the size of the clusters. So we can apply Theorem 3.5 to find an ℓ-cycle in G. □

Proof of Theorem 5.6(i). Note that Lemma 5.19(ii) implies that in order to prove part (i) it suffices to show that every oriented graph H with $\delta^+(H) \geq |H|/k + 149|H|/k^2$ contains a closed walk of length ℓ. Theorem 5.2 implies that H contains an a-cycle C for some $a \leq 1/(1/k + 149/k^2) + 74 < k$. But $a > 2$ since H is oriented and thus a divides ℓ by our definition of k. By traversing C precisely ℓ/a times we obtain the required closed walk of length ℓ in H. □

Note that the proof actually shows the following: Let c be such that every oriented graph G with $\delta^+(G) \geq d$ has a cycle of length at most $\lceil cn/d \rceil$. Then for each $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon, \ell)$ such that every oriented graph G on $n \geq n_0$ vertices with $\delta^+(G) \geq cn/(k-1) + \varepsilon n$ contains an ℓ-cycle (where ℓ and k are as in Theorem 5.6). In particular, if we assume the Caccetta-Häggkvist conjecture, then this implies that Conjecture 5.5 is approximately true if we replace k by $k-1$. Similarly, the result in [22] which gives a cycle of length at most $\lceil 2n/(d+1) \rceil$ in an oriented graph of minimum outdegree at least d implies that we may take $c := 2$. It would be interesting to find improved approximate versions of Conjecture 5.5.

To prove parts (ii) and (iii) of Theorem 5.6, we will use the following lemma.

Lemma 5.20. *Let G be an oriented graph on n vertices.*

(i) *If $\delta^0(G) \geq n/4$ then either the diameter of G is at most 6 or G contains a 3-cycle.*

(ii) *If $\delta^0(G) > n/5$ then either the diameter of G is at most 50 or G contains a 3-cycle.*

Proof. We first prove (i). Consider $x \in V(G)$ and define $X_1 := N^+(x)$ and $X_{i+1} := N^+(X_i) \cup X_i$ for $i \geq 1$. If there exists an i with $\delta^+(G[X_i]) > 3|X_i|/8$ then $G[X_i]$ contains a 3-cycle by Theorem 5.3. So assume not. Then there exists a vertex $x_i \in X_i$ with $|N^+(x_i) \cap X_i| \leq 3|X_i|/8$. Hence

$$|X_{i+1}| \geq |X_i| + (\delta^0(G) - 3|X_i|/8) \geq 5|X_i|/8 + n/4.$$

In particular $|X_2| \geq 13n/32$ and $|X_3| \geq 65n/256 + n/4 = 129n/256 > n/2$. Similarly, for any vertex $y \neq x$ we have that $|\{v \in V(G) : dist(v, y) \leq 3\}| > n/2$, and thus there exists an x-y path of length at most 6, which completes the proof of (i).

To prove (ii), define sets X_i as before. Consider any i for which $|X_i| \leq n/2$. Similarly as before

$$|X_{i+1}| \geq |X_i| + (\delta^0(G) - 3|X_i|/8) > |X_i| + (n/5 - 3|X_i|/8) \geq |X_i| + (n/5 - 3n/16)$$
$$= |X_i| + n/80.$$

Thus $|X_{25}| > n/2$. Similarly, for any vertex $y \neq x$ we have that $|\{v \in V(G) : dist(v, y) \leq 25\}| > n/2$. Thus there exists an x-y path of length at most 50. □

Proof of Theorem 5.6(ii). As in the proof of (i), by Lemma 5.19(i) it suffices to show that every sufficiently large oriented graph H with $\delta^0(H) \geq |H|/4 + 1$ contains a closed walk of length ℓ. If H has a 3-cycle then it contains such a walk since 3 divides ℓ by definition of k. Thus we may assume that H has no 3-cycle. Fact 5.12 implies that the maximum size of an independent set is smaller than the neighbourhood $N_H(v)$ of any vertex v. Thus H contains some orientation of a triangle. By assumption this is not a 3-cycle, and so it must be transitive, i.e. the triangle consists of vertices x, y, z and edges xz, xy, zy.

Since $H - z$ has no 3-cycle, Lemma 5.20(i) implies that $H - z$ contains a y-x path P of length $t \leq 6$. This gives us 2 cycles $C_1 := yPxy$ and $C_2 := yPxzy$ of lengths $t + 1$ and $t+2$ respectively. Write ℓ as $\ell = a(t+1)+r$ with $0 \leq r \leq t \leq 6$. We can wind r times around C_2 and $(a - r)$ times around C_1 to find a closed walk of length ℓ in H provided that $r \leq a$. But the latter holds as $a = \lfloor \ell/(t + 1) \rfloor \geq 6$. □

In the proof of Theorem 5.6(iii), we will use the following result (on undirected graphs) of Andrásfai, Erdős and Sós [4]:

Theorem 5.21. *Every triangle-free graph F on n vertices with minimum degree $\delta(F) > 2n/5$ is bipartite.*

Proof of Theorem 5.6(iii). Again, by Lemma 5.19(i) it suffices to show that every sufficiently large oriented graph H on n vertices with $\delta^0(H) > n/5 + 1$ contains a closed walk of length ℓ.

Let F be the underlying undirected graph of H. Since H has no double edges, we have $\delta(F) > 2n/5$. Suppose first that F contains a triangle. This cannot correspond to a 3-cycle in H, as this in turn immediately yields a closed walk of length ℓ in H. So H must contain a transitive triangle, i.e. vertices x, y, z with $xz, xy, zy \in E(H)$. We can now proceed similarly as in the proof of Theorem 5.6(ii): by Lemma 5.20(ii) we can find a y-x path P of length $t \leq 50$ in $H - z$. This gives us 2 cycles $C_1 := yPxy$ and $C_2 := yPxzy$ of lengths $t+1$ and $t+2$ respectively. To obtain a closed walk of length ℓ, write ℓ as $\ell = a(t+1) + r$ with $0 \leq r \leq t \leq 50$. We can wind r times around C_2 and $(a-r)$ times around C_1 to find a closed walk of length ℓ in H provided that $r \leq a$. But the latter holds as $a = \lfloor \ell/(t+1) \rfloor \geq 50$.

So now suppose that F does not contain a triangle. Then Theorem 5.21 implies that F (and thus H) is bipartite. We will now use this to find a 4-cycle in H. (This immediately yields a closed walk of length ℓ in H.) So suppose that H has no 4-cycle. Write $\delta_0 := \lceil n/5 \rceil + 1$. Denote the vertex classes of H by A and B. Let $a := |A|$ and $b := |B|$, where without loss of generality we have $b \leq n/2$. On the other hand $b \geq \delta(F) \geq 2n/5$ and so $a \leq 3n/5$. Now consider any $v \in A$. Choose a set $X_1 \subseteq N^+(v)$ and $Y_1 \subseteq N^-(v)$ with $|X_1| = |Y_1| = \delta_0$. Let $X_2 := N^+(X_1)$ and $Y_2 := N^-(Y_1)$. Note that X_2 and Y_2 are disjoint, as otherwise we would have a 4-cycle (through v) in H. The number of edges from X_1 to X_2 is at least $|X_1|\delta_0$, so by averaging there is a vertex $x \in X_2$ which receives at least $|X_1|\delta_0/|X_2|$ edges from X_1. This in turn means that x sends at most $|X_1|(1 - \delta_0/|X_2|)$ edges to X_1. Recall that x does not send an edge to Y_1 since otherwise $x \in X_2 \cap Y_2 = \emptyset$. So if we let $Z := B \setminus (X_1 \cup Y_1)$, then x sends at least $\delta_0 - |X_1|(1 - \delta_0/|X_2|) = \delta_0^2/|X_2|$ edges to Z. In particular, $|Z| \geq \delta_0^2/|X_2|$. On the other hand, $|Z| = b - 2\delta_0 \leq n/10$. So $|X_2| \geq \delta_0^2/(n/10) \geq 2\delta_0$. Since X_2 and Y_2 are disjoint, this implies that $|Y_2| \leq a - |X_2| \leq 3n/5 - 2\delta_0 < n/5$. On the other hand, the definition of Y_2 implies that $|Y_2| \geq \delta^0(H)$, a contradiction. □

Proof of Proposition 5.10. First suppose that ℓ is even. The inequality $\delta_{di}(\ell, n) \geq \delta_{orient}(\ell, n)$ is trivial. For the upper bound on $\delta_{di}(\ell, n)$, suppose we are given a digraph H

on n vertices with $\delta^0(H) \geq \delta_{orient}(\ell, n)$. If H has a double edge, it has a closed walk of length ℓ. If it has no double edge, then H has an ℓ-cycle by definition of $\delta_{orient}(\ell, n)$. So in both cases, H has a closed walk of length ℓ. So part (iii) of Lemma 5.19 implies that for each $\varepsilon > 0$ there is an n_0 so that for all $n \geq n_0$ we have $\delta_{di}(\ell, n) \leq \delta_{orient}(\ell, n) + \varepsilon n$, as required.

If ℓ is odd, we obtain the lower bound by considering the complete bipartite digraph with vertex class sizes as equal as possible. The upper bound follows e.g. from (5.1). □

5.4 Proof of Proposition 5.8

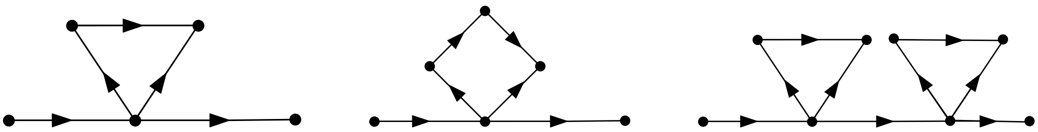
For both parts of Proposition 5.8, the proof divides into three steps.

1. For a given ℓ-cycle C with cycle-type k find an appropriate walk W with prescribed orientation (which will be a cycle for $k \geq 3$) into which there is a digraph homomorphism of C.

2. Prove that the minimum semi-degree condition in Proposition 5.8 guarantees a copy of W in any sufficiently large oriented graph G.

3. Apply Lemma 5.19(i) to 'lift' this walk to one on the cycle C itself.

Let us start with the first step. For $k = 0$ it is clear that there is a digraph homomorphism of C into a directed path of length ℓ. For $k \geq 3$ we can let W be a directed k-cycle and then construct our digraph homomorphism greedily. Suppose that $k = 1$. Then the number of edges of C oriented forwards is one larger than the number of its edges oriented backwards. So C must contain a subpath of the form `ffb`, where we write `f` for an edge oriented forwards and `b` for an edge oriented backwards. But this means that there exist constants $0 \leq k_1, k_2 < \ell$ (depending on C) such that there is a digraph homomorphism of C into the oriented walk W obtained by adding a transitive triangle to the k_1th vertex of a directed path of length k_2 (see Figure 5.2(a)).



(a) $k=1$, `ffb` in cycle (b) $k=2$, `fffb` in cycle (c) $k=2$, `ffb` in cycle twice

Figure 5.2: The walks needed in the cases $k=1$ and $k=2$.

Finally suppose that $k=2$. So C contains two more edges oriented forwards than backwards. Hence C contains two subpaths of the form `ffb` or one subpath of the form `fffb`. In the first case we take W to be a suitable directed path of length less than ℓ with two transitive triangles attached, possibly to the same vertex (see Figure 5.2(c)). In the second case we let W be a suitable directed path with a 4-cycle oriented `fffb` attached (see Figure 5.2(b)).

For the second step we have to show that the relevant minimum semi-degree condition implies the existence of W in G. If W is a path then we only need the minimum semi-degree to be at least ℓ. If W is a k-cycle then we just apply Theorem 5.4. So suppose that $k=1,2$ and consider any vertex x of G. The minimum semi-degree condition $\delta^0(G) \geq (1/3+\alpha)n$ implies each vertex $y \in N^+(x)$ has at least αn neighbours in $N^+(x)$. So G contains the transitive triangles needed in Figures 5.2(a) and (c). To see that we can also find the 4-cycle oriented `fffb`, suppose that $N := G[N^+(x)]$ does not contain a directed path of length 2 (otherwise we are done). Then N must contain two distinct vertices y and y' such that y has no outneighbours in N and y' has no inneighbours in N. But this means that there is some $z \in N^+(y) \cap N^-(y')$ and then $xyzy'$ has the required orientation `fffb`. Hence we can find any of the walks in Figure 5.2 greedily. An application of Lemma 5.19(i) now completes the proof of Proposition 5.8. The argument for the case $k \geq 3$ also shows that Conjecture 5.5 would imply an approximate version of Conjecture 5.9.

CHAPTER 6

ARBITRARY ORIENTATIONS OF HAMILTON CYCLES

6.1 Introduction

It is natural to ask whether the bounds giving Hamilton cycles in Chapter 4 give only directed Hamilton cycles or whether they give every possible orientation of a Hamilton cycle. Indeed this question was answered for digraphs, asymptotically at least, by Häggkvist and Thomason in 1995.

Theorem 6.1 (Häggkvist and Thomason [38]). *There exists n_0 such that every digraph D on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(D) \geq n/2 + n^{5/6}$ contains every orientation of a Hamilton cycle.*

For oriented graphs this question was asked originally by Häggkvist and Thomason [39] who proved that for all $\alpha > 0$ and all sufficiently large oriented graphs G a minimum semi-degree of $(5/12 + \alpha)|G|$ suffices to give *any* orientation of a Hamilton cycle. They conjectured that $(3/8 + \alpha)|G|$ suffices, the same bound as for the directed Hamilton cycle up to the error term $\alpha|G|$. Whilst not asked explicitly before Häggkvist and Thomason's paper, there is some previous work of Thomason and Grant relevant to this area. Grant [36] proved in 1980 that any digraph D with minimum semi-degree $\delta^0(D) \geq 2|D|/3 + \sqrt{|D| \log |D|}$ contains an anti-directed Hamilton cycle, provided that n is even. (An anti-directed cycle is one in which the edge orientations alternate.) Thomason [71] showed in 1986 that every sufficiently large tournament contains every possible orientation of a Hamilton cycle (except possibly the directed Hamilton cycle if the tournament is not strong). The following theorem confirms the conjecture of Häggkvist and Thomason.

Theorem 6.2. *For every $\alpha > 0$ there exists an integer $n_0 = n_0(\alpha)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3/8 + \alpha)n$ contains every orientation of a Hamilton cycle.*

6.1.1 Robust Expansion

The property underlying the proofs of many of the recent results on Hamilton cycle in oriented graphs is *robust expansion*. This is a notion which was introduced by Kühn, Osthus and Treglown in [56] and has proved to be the correct notion of expansion in a digraph when dealing with this kind of question or when using the Diregularity lemma. Informally speaking, a digraph G is a robust outexpander if all subsets of $V(G)$ have outneighbourhoods larger than themselves unless they are very large or very small and, moreover, this still holds after the removal of a small number of edges.

Having a minimum semi-degree $\delta^0(G) \geq (3/8 + \alpha)|G|$ for some $\alpha > 0$, satisfying an approximate Ore-type condition or satisfying an approximate Chvátal condition imply robust outexpansion (see Lemma 11 in [56]). Hence an extension of Theorem 6.2 to robust outexpanders would imply approximate Ore-type and Chvátal-type results for arbitrary orientations of Hamilton cycles. The author believes it is likely that the argument given in this chapter could be straightforwardly extended to prove this.

6.1.2 Extremal Example

As discussed in Section 4.2, Häggkvist [37] constructed an example in 1993 giving a graph on $n = 8k - 1$ vertices with minimum semi-degree $(3n - 5)/8$ containing no Hamilton

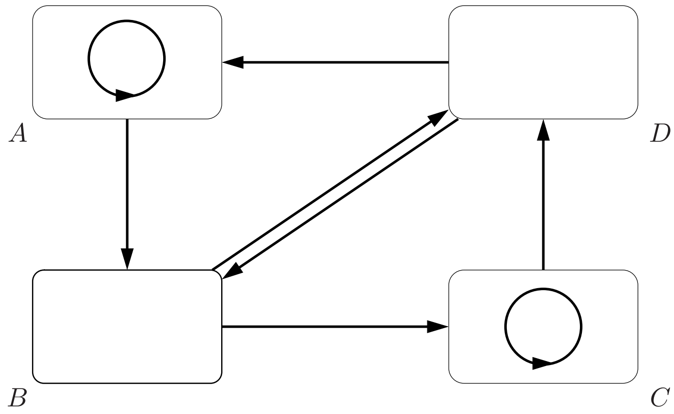


Figure 6.1: The oriented graph constructed in Proposition 6.3

cycle. In 2009 Keevash, Kühn and Osthus [46] extended this to all n. This means that Theorem 1.9 is best possible and that Theorem 6.2 is best possible up to the linear error term. Interestingly, this example can be improved upon when considering arbitrary orientations. Hence the additive constant in Theorem 1.9 is not the correct bound when seeking any orientation of a Hamilton cycle. It is an open question as to what the correct additional term should be.

Proposition 6.3. *There are infinitely many oriented graphs G with minimum semi-degree exactly $(3|G|-4)/8$ which do not contain an anti-directed Hamilton cycle.*

Proof. Let $n := 8m + 4$ for some integer $m \in \mathbb{N}$. Let G be the oriented graph obtained from the disjoint union of two regular tournaments A and C on $2m+1$ vertices and sets B and D of $2m + 1$ vertices by adding all edges from A to B, all edges from B to C, all edges from C to D and all edges from D to A. Finally, between B and D we add edges to obtain a bipartite tournament which is as regular as possible, i.e. the indegree and the outdegree of every vertex differs by at most 1. So in particular every vertex in B sends at least m edges to D. It is easy to check that the minimum semi-degree of G is $3m + 1 = (3n - 4)/8$, as required.

Let us try to construct an anti-directed Hamilton cycle in G and let us start in B with an edge going forwards. This edge can go either to C or to D. (Starting with an edge oriented backwards produces an identical argument and result.) The next edge must go backwards. It can go from C to either B or C. It can go from D to either B or C. So after two steps we can be in either B or C. Our next edge must go forwards. If we are in B our possible locations after the next two steps are B and C as before. From C we can go forwards either to C or to D. Both options repeat situations we have already met. In no case do we have a means to reach A whilst respecting the orientation of our anti-directed Hamilton cycle. Hence the longest anti-directed cycle in G has length at most $3n/4$ and we have no anti-directed Hamilton cycle as claimed. □

6.2 Overview of the Proof

The proof of Theorem 6.2 splits into two parts, both relying on the expansion properties that our minimum semi-degree condition implies. The cases are distinguished by the similarity of the Hamilton cycle C we are trying to embed to the standard orientation of a Hamilton cycle. It turns out that the correct measure, at least for this problem, of whether a cycle is close to a directed cycle is the number of pairs of consecutive edges with different orientations. Given an oriented graph C we call the subgraph induced by three vertices $x, y, z \in V(C)$ a *neutral pair* if $xy, zy \in E(C)$. Given an arbitrarily oriented cycle C on n vertices let $n(C)$ be the number of neutral pairs in C. Write C_n^* for the standard orientation of a cycle on n vertices. When there is no ambiguity we will merely write C^*.

The essential idea is to split the cycle up into alternating short and long paths and use the probabilistic method to find an approximate embedding of the long paths into a Hamilton cycle of the reduced oriented graph created by applying the Diregularity lemma. We connect these paths up greedily using the short paths and then adjust the embedding to obtain something which, after the Blow-up lemma has been applied, gives us the desired orientation of a Hamilton cycle in our graph.

The case distinction comes in the manner in which we alter our embedding. In Section 6.6 we give the argument for cycles far from C^*, where we use the neutral pairs for our adjustments. In Section 6.7 we assume that we have few neutral pairs, and thus many long sections of C containing no changes in direction, and use these to adjust our embedding.

Our need to have more control over the number of exceptional vertices than provided directly by the Diregularity lemma means that some technical difficulties are introduced. So we control the number of exceptional vertices by randomly splitting our oriented graph G. In still vague, but slightly more precise terms, the Diregularity lemma will for any $\varepsilon > 0$ give us a partition with the property of ε-regularity. It will also give us a set of 'exceptional vertices' which are in some sense badly behaved, but tells us that these make up at most an ε proportion of our vertices. Our method can only cope with $\eta n \ll \varepsilon n$ such vertices. Hence we split the vertices of our given graph G into two sets A and B of roughly equal size (satisfying some 'nice' properties). We apply the Diregularity lemma to $G[B]$, giving us at most $\varepsilon|G|$ exceptional vertices V_0. We then apply the Diregularity lemma to $G[A \cup V_0]$ only this time not with ε but with η. This gives us at most $\eta|G|$ exceptional vertices V_0'. We then consider $G_B := G[(B \setminus V_0) \cup V_0']$, which is ε-regular and has $\eta|G| \ll \varepsilon|G_B|$ exceptional vertices and $G_A := G - G_B$, which is η-regular and has $0 \ll \eta|G_A|$ exceptional vertices. Hence, at the cost of some technical work and having to stitch everything back together we are able to control the number of exceptional vertices.

6.3 Skewed Traverses and Shifted Walks

In this section we introduce some tools needed to tweak a random embedding of an arbitrarily oriented Hamilton cycle into a directed Hamilton cycle of the reduced oriented graph to make it correspond (in some sense) to the desired orientation of a Hamilton cycle in our original graph.

First we must recall from Section 4.4 the following crucial result which says that our minimum semi-degree condition implies outexpansion.

Lemma 6.4 (K., Kühn, Osthus [48])**.** *Let R be an oriented graph with $\delta^0(R) \geq (3/8 + \alpha)|R|$ for some $\alpha > 0$. If $X \subset V(R)$ with $0 < |X| \leq (1 - \alpha)|R|$ then $|N^+(X)| \geq |X| + \alpha|R|/2$.*

(The condition on $\delta^0(G)$ implies the necessary $\delta^*(G)$ bound to allow us to apply Lemma 4.8.) Suppose that F is a Hamilton cycle (with the standard orientation) of the reduced oriented graph R and relabel the vertices of R such that $F = V_1V_2 \dots V_M$, where we let $M := |R|$. Create a new digraph R^* from R by adding all the exceptional vertices $v \in V_0$ and adding an edge vV_i (where V_i is a cluster containing m vertices) whenever v sends edges to a significant proportion of the vertices in V_i, say we add vV_i whenever v sends at least cm edges to V_i for some constant $c > 0$. (Recall that m denotes the size of the clusters.) The edges in R^* of the form V_iv are defined in a similar way. Let G^c be the digraph obtained from the pure oriented graph G^* by making all the non-empty bipartite subgraphs between the clusters complete (and orienting all the edges between these clusters in the direction induced by R) and adding the vertices in V_0 as well as all the edges of G between V_0 and $V(G - V_0)$.

Let W be an assignment of the vertices of an arbitrarily oriented cycle C on n vertices to the vertices of R^* which respects edges (i.e. is a digraph homomorphism from C to R^*). We denote by $a(i)$ the number of vertices of C assigned to the cluster V_i. Observe that we can think of W either as a (possibly degenerate) embedding into G^c or as a closed walk in R^*. It will be useful to the reader to keep both interpretations in mind when reading the rest of the proof. We say that an assignment W of C to R^* is *γ-balanced* if $\max_i |a(i) - m| \leq \gamma n$ and *balanced* if $a(i) = m$ for all i. Furthermore, we say that an embedding *(γ, μ)-corresponds to C* if the following conditions hold.

- W is γ-balanced.
- Each exceptional vertex $v \in V_0$ has exactly one vertex of C assigned to it.

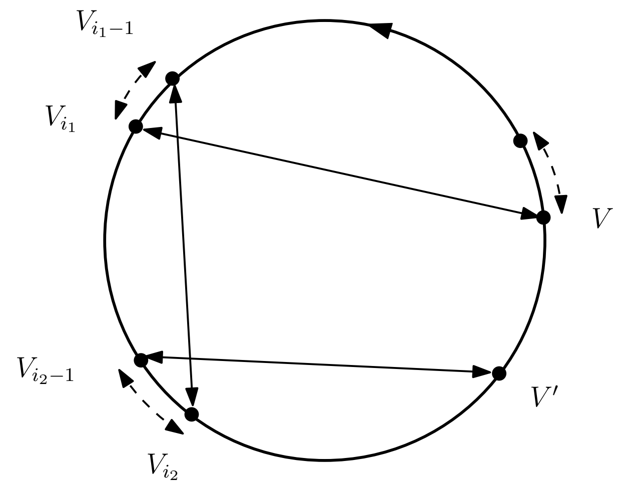


Figure 6.2: A skewed V-V' traverse

- In every $V_i \in V(R)$ at least $m - \mu n$ of the vertices of C assigned to V_i have both of their neighbours assigned to $V_{i-1} \cup V_{i+1}$.

We say that the assignment *μ-corresponds to C* if it $(0, \mu)$-corresponds to C.

Once we have found such an assignment we can use the Blow-up lemma (Lemma 3.3) to show that it corresponds to a copy of C in G. Our immediate aim then is to find such a closed walk corresponding to C.

Given clusters V and V', a *skewed V-V' traverse* $T(V, V')$ is a collection of edges of the form

$$T(V, V') := VV_{i_1}, V_{i_1-1}V_{i_2}, V_{i_2-1}V_{i_3}, \dots, V_{i_t-1}V'.$$

The *length* of a skewed traverse in the number of its edges minus one; so the length of the above skewed traverse is t. Suppose that we have a γ-balanced assignment W of C to R^* and that many neutral pairs of C are assigned to each vertex of R. We would like to make this a balanced embedding by modifying W. Let V_i, V_j be clusters with $a(V_i) > m$ and $a(V_j) < m$. If $V_{i-1}V_j \in E(R)$ then we could replace one neutral pair assigned to $V_{i-1}V_iV_{i-1}$ in the embedding with $V_{i-1}V_jV_{i-1}$. This would reduce $a(V_i)$ by one and increase $a(V_j)$ by one. Repeating this process would give the desired balanced embedding. We can not though guarantee that $V_{i-1}V_j \in E(R)$ so we are forced to use skewed traverses to achieve the same affect, which we are able to show always exist under certain conditions. Let

$$V_{i-1}V_{i_1}, V_{i_1-1}V_{i_2}, V_{i_2-1}V_{i_3}, \dots, V_{i_t-1}V_j$$

be a skewed V_{i-1}-V_j traverse. Then replacing neutral pairs starting at $V_{i-1}, V_{i_1-1}, \dots, V_{i_t-1}$ with the edges in the skewed V_{i-1}-V_j traverse we reduce $a(V_i)$ by one, increase $a(V_j)$ by one and crucially do not alter $a(V_k)$ for any $V_k \in V(R) \setminus \{V_i, V_j\}$. See Figure 6.2 for an illustration of this, where the dashed edges represent the neutral pairs which will be replaced by the solid edges representing the edges of the skewed traverse. We always assume that a skewed traverse has minimal length and thus that each vertex $V_i \in V(R)$ appears at most once as the first vertex of an edge in a skewed traverse.

Given vertices $V, V' \in V(R)$, a *shifted* V-V' *walk* $S(V, V')$ is a walk of the form

$$S(V, V') := VV_{i_1}FV_{i_1-1}V_{i_2}FV_{i_2-1} \dots V_{i_t}FV_{i_t-1}V',$$

where we write V_iFV_j for the path

$$V_iFV_j := V_iV_{i+1}V_{i+2} \dots V_j,$$

counting modulo $|F| = M$. (The case $t = 0$, and thus a walk VV', is allowed.) We say that W traverses F t times and always assume that a shifted walk $S(V, V')$ traverses F as few times as possible. Its length is the length of the corresponding walk in R. Note that if we can find a skewed V-V' traverse then we can find a shifted V-V' walk.

The most important property of shifted walks is that the walk $W - \{V, V'\}$ visits every vertex in R an equal number of times. Observe also that by our minimality assumption each vertex V_i is visited at most one time from a vertex other than V_{i-1}. I.e. of the t times that V_i is visited at most one does not come from winding around F. This fact will be useful later when we try and bound the number of edges of an embedding not lying on the edges of F.

As with skewed traverses, we can use shifted walks to go from an approximate assignment W of a cycle C to R^* to a balanced assignment. Let V_i, V_j be clusters with $a(V_i) > m$ and $a(V_j) < m$. If $V_{i-1}V_j, V_jV_{i+1} \in E(R)$ then we could replace one section of W isomorphic to F by $V_{i-1}V_jV_{i+1}FV_{i-1}$, that is, replace $V_{i-1}V_iV_{i+1}$ by $V_{i-1}V_jV_{i+1}$. This new section has the same length as before and so would not alter the rest of W. Clearly we can not ensure that such edges always exist. Instead we use shifted walks and replace a section of the embedding that looks like $FF\dots F$ with

$$S(V_{i-1}, V_j)S(V_j, V_{i+1})FV_{i-1}F\dots FV_{i-1};$$

where the $F\dots F$ in the new embedding contains the appropriate number of F to ensure that it is of exactly the same length as the section of the assignment it replaced. This is a shifted walk from V_{i-1} to V_j, then a shifted walk from V_j to V_{i+1} and then wind around F. By our definition of shifted walks each cluster will have the same number of vertices assigned to it (except V_{i-1}, V_i and V_j) the total number of vertices assigned will not be altered. Clearly this method needs the cycle we're trying to embed to contain many long sections with no changes of orientation (and oriented in the same direction as F). In the case where the cycle we are trying to embed is close to C^*, the standard orientation of a cycle, we are indeed able to ensure this.

Corollary 6.5. *Let R be an oriented graph on M vertices with $\delta^0(R) \geq (3/8+\alpha)M$ for some $\alpha > 0$ and let $F = V_1V_2\dots V_M$ be a Hamilton cycle of R. Define $r := \lceil 2/\alpha \rceil$. Then for any distinct $V, V' \in V(R)$ there exists the following.*

(i) A skewed V-V' traverse of length at most r.

(ii) A shifted V-V' walk traversing at most r cycles.

Proof. Let A_i be the set of vertices which can be reached from V by a skewed traverse of length at most i and let $A_i^- := \{V_i \in V(R) : V_{i+1} \in A_i\}$. If $|A_{r-2}| \geq (1-\alpha)M$ then $N^-(V') \cap A_{r-2}^- \neq \emptyset$ and so we have a skewed V-V' traverse of length at most $r-1$.

Otherwise $|A_{r-2}| \leq (1-\alpha)M$, so $|A_i| \leq (1-\alpha)M$ for all $i \leq r-2$: then, applying Lemma 4.8, we have that for each $i \leq r-2$ we have $|A_{i+1}| \geq |A_i| + \alpha M/2$. Thus $|A_{r-2}| \geq (r-2)\alpha M/2 \geq (2/\alpha - 2)\alpha M/2 \geq (1-\alpha)M > M - |N^-(V')|$ and so again $N^-(V') \cap A^-_{r-2} \neq \emptyset$ and thus we get a skewed traverse

This skewed traverse also gives the desired shifted walk, merely 'wind around' F after each edge. □

When linking together sections of our cycle we will sometimes need to find a path between two vertices which is not just short but is isomorphic to a given path. To do this we use the following lemma of Häggkvist and Thomason.

Lemma 6.6 (Häggkvist and Thomason [39], Lemma 2)**.** *Let R be an oriented graph on M vertices with $\delta^0(R) \geq (3/8+\alpha)M$ for some $\alpha > 0$. Let $4\lceil \log_2(1/\alpha)\rceil \leq k \leq \alpha M/4$ and let P be an arbitrarily oriented path of length k. Then, if M is large enough and $V, V' \in V(R)$ are distinct vertices, there exists a path from V to V' isomorphic to P.*

6.4 An approximate embedding lemma

Our main tool in our proof of Theorem 6.2 is the following probabilistic result which says that we can assign a series of paths P_i to the vertices of a small graph R such that each vertex of R is assigned approximately the same number of vertices. Furthermore, we show that if we have a large collection of subpaths of the P_i we can assure that every vertex of R is assigned a reasonable number of these. When we talk about 'greedily embedding an oriented path P_i around a cycle F given a starting point $V \in V(F)$' we mean the following. Assign the first vertex of P_i to V. Given an embedding of some initial segment of P_i which ends at $V' \in V(F)$ assign the next vertex of P_i to either the successor or the predecessor of V' in F according to the orientation of the edge in P_i.

Lemma 6.7. *Let R be an oriented graph on M vertices and let F be a Hamilton cycle in R. Let $\mathcal{P} = \{P_1, \ldots, P_s\}$ be a collection of arbitrarily oriented paths on t vertices and*

$\mathcal{Q}$ be a collection of pairwise disjoint oriented subpaths of the P_i. Then for any $\gamma > 0$ and sufficiently large s there exists a map $\phi : [s] \to V(R)$ such that if the paths are greedily embedded around F with the embedding of each $P(i)$ starting at $\phi(i)$ then the following holds. Define $a(i)$ to be the number of vertices in $\bigcup_{j=1}^{s} P_j$ assigned to V_i by this embedding and define $n(i, \mathcal{Q})$ to be the number of oriented subpaths in $\mathcal{Q}$ starting at V_i. Then for all $V_i \in V(R)$

$$\left| a(i) - \frac{st}{M} \right| \leq \gamma st, \tag{6.1}$$

$$\left| n(i, \mathcal{Q}) - \frac{|\mathcal{Q}|}{M} \right| \leq \gamma st. \tag{6.2}$$

To prove it we need the following well-known probabilistic bound (see [61] for example).

Theorem 6.8. *Let X be a random variable determined by s independent trials $X_1, \ldots, X_s$ such that changing the outcome of any one trial can affect X by at most c. Then for any $\lambda > 0$,*

$$\Pr(|X - \mathbb{E}(X)| > \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2c^2 s} \right).$$

Proof. [of Lemma 6.7] We construct ϕ by picking each $\phi(i)$ independently and uniformly at random. Observe that the assignment of any one path P_i can change the number of vertices assigned to any vertex of R by at most t. Clearly $\mathbb{E}(a(i)) = st/M$. By Theorem 6.8 we have

$$\Pr(|a(i) - st/M| > \gamma st) \leq 2\exp(-\frac{\gamma^2 s^2 t^2}{2t^2 s}) = 2\exp(-\frac{\gamma^2 s}{2}) < 1/(2M)$$

for $s \gg M$.

A similar argument gives that the probability that $n(i, \mathcal{Q})$ differs too much from the expected value is at most $1/(2M)$. Thus the probability that there exists V_i which does not have almost the expected number of vertices or almost the expected number of starting points of paths in $\mathcal{Q}$ assigned to it by ϕ is less than 1. Hence a map satisfying the conclusion of the lemma exists. □

6.5 Preparations for the Proof of Theorem 6.2

6.5.1 The Two Cases

We split into two cases depending on the number of neutral pairs. Let G be an oriented graph on n vertices with $\delta^0(G) \geq (3/8 + \alpha)n$ for some constant $0 < \alpha \ll 1$. Let C be an orientation of a cycle on n vertices with $n(C) = \lambda n$ neutral pairs. Define the following hierarchy of constants.

$$0 < \varepsilon_1 \ll \varepsilon_2 \ll \varepsilon_3 \ll \varepsilon_4 \ll \varepsilon_5 \ll \varepsilon_6 \ll \alpha < 1.$$

Let $\mathcal{Q}$ be a maximal collection of neutral pairs all at a distance of at least 3 from each other, where the distance between two neutral pairs is understood to be minimum of the distances between the ends.

If $\lambda \ll \varepsilon_4$ then let $\varepsilon := \varepsilon_6$, $\varepsilon_A := \varepsilon_5$ and $\varepsilon^* := \varepsilon_4$. The proof of this case is given in Section 6.7.

Otherwise we have $\lambda \gg \varepsilon_3$ (more strictly, we construct our hierarchy of constants such that either this or $\lambda \ll \varepsilon_4$ is true) and we set $\varepsilon := \varepsilon_3$, $\varepsilon_A := \varepsilon_2$ and $\varepsilon^* := \varepsilon_1$. The proof of this case is in Section 6.6.

The following two sections, where we partition G and C in preparation for our embedding, are common to both cases.

6.5.2 Preparing G for the Proof of Theorem 6.2

Define a positive constant d and integers M'_A, M'_B (all functions of α) such that

$$0 < \varepsilon^* \ll 1/M'_A \ll \varepsilon_A \ll 1/M'_B \ll \varepsilon \ll d \ll \alpha \ll 1.$$

Chernoff-type bounds applied to a random partition of $V(G)$ show the existence of a subset $A \subset V(G)$ with $(1/2 - \varepsilon)n \leq |A| \leq (1/2 - \varepsilon)n$ such that every vertex $x \in V(G)$

satisfies

$$\frac{d^+(x)}{n} - \frac{\alpha}{10} \leq \frac{|N_A^+(x)|}{|A|} \leq \frac{d^+(x)}{n} + \frac{\alpha}{10}$$

and similarly for $d^-(x)$. Apply the Diregularity lemma (Lemma 3.1) with parameters ε^2, $d+8\varepsilon^2$ and M'_B to $G-A$ to obtain a partition of the vertex set of $G-A$ into $M_B := k \geq M'_B$ clusters $V_1, \dots, V_k$ and an exceptional set V_0. Set $B := V_1 \cup \ldots \cup V_k$ and $m_B := |V_1| = \cdots = |V_k|$. Let $G'_B := G[B]$, let R_B denote the reduced oriented graph obtained by an application of Lemma 3.2 and let G^*_B be the pure oriented graph. Since $\delta^+(G-A)/|G-A| \geq \delta^+(G)/n - \alpha/9$ by our choice of A, Lemma 3.2 implies that

$$\delta^0(R_B) \geq \left(\frac{\delta^0(G)}{n} - \frac{\alpha}{4}\right)|R_B| \geq \left(\frac{3}{8} + \frac{3\alpha}{4}\right)|R_B|. \tag{6.3}$$

So Theorem 4.3 gives us a Hamilton cycle F_B of R_B. Relabel the clusters of R_B so that $V_iV_{i+1} \in E(F_B)$ for all i. We now apply Lemma 3.9 with F_B playing the role of S, ε^2 playing the role of ε and $d+8\varepsilon^2$ playing the role of d. This shows that by adding at most $4\varepsilon^2 n$ further vertices to the exceptional set V_0 we may assume that each edge of R_B corresponds to an ε-regular pair of density at least d (in the underlying graph of G^*_B) and that each edge in F_B corresponds to an (ε, d)-super-regular pair. Note that the new exceptional set now satisfies $|V_0| \leq \varepsilon n$.

Now apply the Diregularity Lemma with parameters $\varepsilon_A^2/4$, $d+2\varepsilon_A^2$ and M'_A to $G[A \cup V_0]$ to obtain a partition of the vertex set of $G[A \cup V_0]$ into $M_A := \ell \geq M'_A$ clusters $V'_1, \dots, V'_\ell$ and an exceptional set V'_0. Let $A' := V'_1 \cup \cdots \cup V'_\ell$, let R_A denote the reduced oriented graph obtained from Lemma 3.2 and let G^*_A be the pure oriented graph. As before Lemma 3.2 implies that $\delta^0(R_A) \geq (3/8 + 3\alpha/4)|R_A|$ and so, as before, we can apply Theorem 4.3 to find a Hamilton cycle F_A of R_A. Then as before, Lemma 3.9 implies that by adding at most $\varepsilon_A^2|A \cup V_0|$ further vertices to the exceptional set V'_0 we may assume that each edge of R_A corresponds to an ε_A-regular pair of density at least d and that each edge in F_A corresponds to an (ε_A, d)-super-regular pair. Finally define $G_B := G[B \cup V'_0]$

and $n_B := |G_B|$ and observe that we now have

$$|V_0'| \leq \varepsilon_A |A \cup V_0|/2 < \varepsilon_A n_B. \tag{6.4}$$

In both cases of our proof we now have

$$0 < \varepsilon^* \ll 1/M_A \ll \varepsilon_A \ll 1/M_B \ll \varepsilon \ll d \ll \alpha \ll 1.$$

6.5.3 Preparing C

We would like to divide C into a number of paths and use Lemma 6.7 to obtain a ε-balanced assignment of C to R. Since we have split our graph G into two parts, we have to split C into two paths P_A and P_B and embed these into (an oriented graph similar to) $G[A']$ and G_B respectively.

Define $r := 4\lceil \log_2(4/\alpha) \rceil$. Lemma 6.6 tells us that if P is an orientation of a path of length r then between any two distinct vertices in $V(R_B)$ or $V(R_A)$ there exists a path between them isomorphic to P.

Define

$$s := \lfloor (\log n)^2 \rfloor, \quad t := \left\lfloor \frac{n-(s+1)(r-1)}{s+2} \right\rfloor - 1 \approx \frac{n}{(\log n)^2}.$$

Recall that $\mathcal{Q}$ is a maximal collection of neutral pairs in C all at a distance of at least 3 from each other. If $\mathcal{Q}$ is large, i.e. we are in the case where C is far from C^*, let v^* be a vertex in C such that the subpath of C of length $n/2$ following v^* and the subpath of C preceding v^* both contain at least $2|\mathcal{Q}|/5$ elements of $\mathcal{Q}$. Divide C into (overlapping) paths

$$C := Q_1P_1Q_2P_2 \ldots Q_{s-1}P_{s-1}Q_sP_sQ^*P^*$$

where their lengths satisfy $\ell(P_i) = t$, $\ell(Q_i) = \ell(Q^*) = r$ and $2t \leq \ell(P^*) < 3t$ and Q_1 starts at v^*. Let s_B be such that

$$1 < n_B - s_B(t+r) < \ell(P^*)$$

and let

$$P_B := P_B^* Q_1 P_1 \dots Q_{s_B} P_{s_B}$$

where P_B^* is an initial segment of P^* of such length as to ensure $\ell(P_B) + 1 = n_B$. Let

$$P_A := Q_1' P_1' \dots Q_{s_A}' P_{s_A}' Q^* P_A^*$$

where $Q_i' := Q_{s_B+i}$, $P_i' := P_{s_B+i}$, $s_A := s - s_B$ and P_A^* is the terminal segment of P^* which overlaps P_B^* in exactly one vertex. Observe that we now have

$$n_B = s_B t + s_B r + \ell(P_B^*) = |V(P_B)| \tag{6.5}$$

and define

$$n_A := n - n_B = s_A t + (s_A + 1)r + \ell(P_A^*) - 1 = |V(P_A)| - 2.$$

6.6 Cycle is Far From C^*

6.6.1 Approximate Embedding

First we assign the paths P_i to the clusters of R_B in such a way as to ensure that all the clusters are assigned approximately the same number of vertices and the neutral pairs are relatively evenly distributed. Recall that $\mathcal{Q}$ is a maximal collection of neutral pairs in C all at a distance of at least 3 from each other. Let $\mathcal{Q}_B$ contain all neutral pairs in P_B from $\mathcal{Q}$ which are contained in and at a distance of at least three from the ends of the P_i. Apply Lemma 6.7 to R_B, $\mathcal{P}_B := \{P_1, P_2 \dots, P_{s_B}\}$ and $\mathcal{Q}_B$ with ε^* as γ to obtain an embedding of the P_i into $V(R_B)$ with

$$\left| a(i) - \frac{s_B t}{M_B} \right| \leq \varepsilon^* s_B t, \quad \left| n(i, \mathcal{Q}_B) - \frac{|\mathcal{Q}_B|}{M_B} \right| \leq \varepsilon^* s_B t.$$

for all $V_i \in V(R_B)$. (Recall that $a(i)$ is defined to be the number of vertices assigned to the cluster V_i by the embedding.) In a slight abuse of notation let $n(i)$ be the number of neutral pairs in $\mathcal{Q}_B$ starting at V_i. Note that

$$\begin{aligned} |a(i) - m_B| &\overset{(6.5)}{\leq} \left|a(i) - \frac{s_B t}{M_B}\right| + \left|\frac{s_B t}{M_B} - m_B\right| \\ &\leq \left|a(i) - \frac{s_B t}{M_B}\right| + \left|\frac{s_B r + 2t}{M_B}\right| \\ &\leq \left|a(i) - \frac{s_B t}{M_B}\right| + \varepsilon^* m_B. \end{aligned} \tag{6.6}$$

The last term here is $\leq \varepsilon^* m_B$ if and only if $|a_B t - n_B| \leq \varepsilon^* n_B$, which follows from the definition of s_B. The requirement that the neutral pairs in $\mathcal{Q}$ are at a distance of at most three from each other means that $|\mathcal{Q}| \geq n(C)/4$. By the observation in Section 6.5.3 we know that P_B contains at least $2|\mathcal{Q}|/5 \geq \lambda n/10$ neutral pairs. The paths Q_i and P_B^* together contain fewer than $s_B r + 3t$ neutral pairs and at most $4s_B$ neutral pairs can be in the P_i but within a distance of at most three from a Q_i. Thus for all i

$$n(i) \geq \frac{\lambda n}{10 M_B} - \varepsilon^* s_B t - (s_B r + 3t + 4s_B) \geq \frac{\lambda n_B}{6 M_B} - 2\varepsilon^* n_B \geq \frac{\lambda m_B}{7},$$

where $m_B := |V_i| = (n_B - |V_0'|)/M_B$ is the size of a cluster. For all $2 \leq i \leq s_B$ we can join P_{i-1} and P_i by a path in R_B isomorphic to Q_i using Lemma 6.6. Furthermore we can greedily extend P_1 backwards by a path isomorphic to $P_B^* Q_1$. This will increase $a(i)$ by at most $s_B r + 3t < \varepsilon^* m_B$ for n sufficiently large. We now have an assignment of P_B to the clusters of R_B which we can think of as a walk W_B in R_B.

6.6.2 Incorporating the Exceptional Vertices

Let G_B^c be the digraph obtained from the pure oriented graph G_B^* by making all the non-empty bipartite subgraphs between the clusters complete (and orienting all the edges between these clusters in the direction induced by R_B) and adding the vertices in V_0' as well as all the edges of G between V_0' and $V(G_B - V_0')$. Our next aim is to incorporate

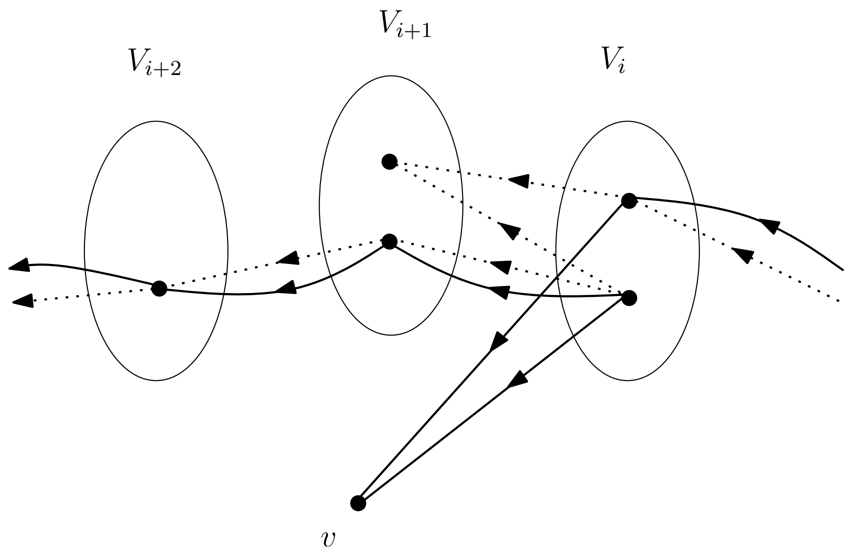


Figure 6.3: Incorporating an exceptional vertex when C is far from C^*.

the exceptional vertices V_0' into the walk W_B. We do this by considering the following extension of R_B. Define $R_B^* \supseteq R_B$ to be the digraph formed by adding to R_B the vertices in V_0' and, for $v \in V_0'$ and $V_i \in V(R_B)$, the edge vV_i if $|N_G^+(v) \cap V_i| > \alpha m_B/10$ and $V_i v$ if $|N_G^-(v) \cap V_i| > \alpha m_B/10$.

Then for each $v \in V_0'$ pick an inneighbour $V_i \in V(R_B)$ and replace one neutral pair $V_iV_{i+1}V_i$ starting at V_i with V_ivV_i. This reduces $a(i+1)$ and $n(i)$ by one. Figure 6.3 contains an illustration of this, where we consider W_B as being in G_B^c and the dotted lines as the section of the embedding to be replaced by the solid lines. After doing this for every exceptional vertex we will have that for all $V_i \in V(R_B)$

$$|a(i) - m_B| \overset{(6.6)}{\leq} \left| a(i) - \frac{s_B t}{M_B} \right| + \varepsilon^* m_B \tag{6.7}$$

$$\leq (\varepsilon^* s_B t + \varepsilon_A m_B + |V_0'|) + \varepsilon^* m_B \overset{(6.4)}{<} 4\varepsilon_A n_B, \tag{6.8}$$

where the second term in the second line comes from greedily embedding the Q_i. We also still have a reasonable number of neutral pairs starting at each cluster of R_B for all $V_i \in V(R_B)$:

$$n(i) \geq \frac{\lambda m_B}{7} - |V_0'| > \frac{\lambda m_B}{7} - \varepsilon_A n_B > \frac{\lambda m_B}{8}.$$

Note that of the $a(i)$ vertices of P_B assigned to any $V_i \in V(R)$, at most $\varepsilon_A n_B + 2|V_0'| \leq 3\varepsilon_A n_B$ do not have their neighbours assigned to $V_{i-1} \cup V_{i+1}$, where the first term came from the Q_i and the second came from incorporating the exceptional vertices. Thus we currently have a $(4\varepsilon_A, 3\varepsilon_A)$-corresponding embedding of P_B into R_B^*.

6.6.3 Adjusting the Embedding

We now adjust W_B to obtain a $5\varepsilon_A M_B$-corresponding assignment of P_B to R_B^*; i.e. we adjust W_B to ensure that $a(i) = m_B$ for all $V_i \in V(R_B)$. Recall from Corollary 6.5 that between any two vertices in R_B there exists a skewed traverse of length at most $\lceil 8/3\alpha \rceil < r' := \lceil 4/\alpha \rceil$. Then for each cluster $V_i \in V(R_B)$ with $a(i+1) > m_B$ pick $V_j \in V(R_B)$ with $a(j) < m_B$ and find a skewed V_i-V_j traverse of length $q \leq r'$:

$$V_i V_{k_1}, V_{k_1-1} V_{k_2}, V_{k_2-1} V_{k_3}, \ldots, V_{k_q} V_{k_q-1}, V_{k_q-1} V_j.$$

As discussed in Section 6.3 we can use this skewed traverse to modify W_B to reduce $a(i+1)$ by one, increase $a(j)$ by one and leave the number of vertices assigned to every other cluster of R_B the same. We do this by, for every $0 \leq p \leq q$, replacing a neutral pair $V_{k_p-1} V_{k_p} V_{k_p-1}$ in W_B by $V_{k_p-1} V_{k_{p+1}} V_{k_p-1}$ where we define $V_{k_0-1} := V_i$ and $V_{k_{q+1}} := V_j$.

Since $\sum_{i=1}^{M_B} |a(i) - m_B| \leq 4\varepsilon_A M_B n_B$, doing this will consume at most $4\varepsilon_A M_B n_B$ neutral pairs starting at any vertex of R_B. This is fine though as for all $V_i \in V(R_B)$ we have $n(i) \geq \lambda m_B/8 \gg 4\varepsilon_A M_B n_B$. Each cluster V_i now has at most $3\varepsilon_A n_B + 4\varepsilon_A M_B n_B < 5\varepsilon_A M_B n_B$ vertices of P_B assigned to it that do not have both their neighbours assigned to $V_{i-1} \cup V_{i+1}$. Hence we have constructed a $5\varepsilon_A M_B$-corresponding embedding W_B of P_B into R_B^*.

6.6.4 Finding a copy of P_B in G_B

We will now use Lemma 3.6 to find a copy of P_B in G_B. To do this we use W_B to find an embedding W_B' of P_B into G_B such that

- Every vertex of W_B in V_0' is unchanged in W_B'.
- Each appearance of a cluster of R_B in W_B is replaced by a unique vertex in the corresponding cluster in G_B.
- Every edge of W_B which does not lie upon an edge of F_B is mapped to an edge of G_B.

First we split W_B into two digraphs W_B^1 and W_B^2. Let W_B^1 consist of all maximal walks

$$u_{i,1}u_{i,2}\dots u_{i,\ell_i}$$

in W_B of length at least three whose edges all lie on F_B. Let W_B^2 consist of all edges not in W_B^1. Then W_B^2 is a union of walks $v_{i,1}v_{i,2}\dots v_{i,\ell_i}$, where we relabel if necessary to ensure that $u_{i,1} = v_{i-1,\ell_{i-1}}$ and $u_{i,\ell_i} = v_{i,1}$. We now greedily find an embedding of W_B^2 into G_B which will satisfy the third requirement above.

The walks in W_B^2 are of one of three types. The first type comes from the incorporation of an exceptional vertex, in which we have an exceptional vertex $x \in V_0'$ and a cluster $V_i \in V(R_B)$ with $|N_G^-(x) \cap V_i| > \alpha m_B/10$. In this case we choose any two distinct vertices $u, v \in N_G^-(x) \cap V_i$, which we can do as there are at most $|V_0'| \ll \varepsilon m_B \ll \alpha m_B/10$ exceptional vertices. The second type comes from the paths Q_i and the path P_B^*. These we find in G_B^* (and hence in $G_B \supseteq G_B^*$) greedily. We can do so as the total length of the Q_i is at most $s_B r + 2t \ll \varepsilon m_B$ and all their edges are assigned to edges in R_B corresponding to ε-regular pairs of density at least d in G_B^*. The final type are pairs of edges ij, ji with $i, j \in V(R_B)$ which come from the skewed traverses used to ensure that the correct number of vertices of P_B were assigned to each vertex of R_B. There are at most $5\varepsilon_A M_B n \ll \varepsilon m_B$ of these and so we can again find these greedily. Note that our requirement that all the neutral pairs in $\mathcal{Q}$ are at a distance of at least three from each other and the ends of the P_i implies that we have now considered all possible walks in W_B^2. To satisfy the second condition above we simply assign each vertex of W_B not already assigned to a vertex in the corresponding cluster in G_B. As W_B is balanced (i.e. W_B assigns exactly m_B vertices to each cluster) we can do this.

For all i let S_i consist of the vertices of $G_B - V_0'$ to which the vertices of W_B^1 that are not at the end of a path have been assigned. We can now apply Lemma 3.6 to $G_B - V_0'$ with W_B^1 as H, the $u_{i,1}$ and u_{i,ℓ_i} as the x_P and y_P respectively and the S_i as just defined. Combining this with the embedding of W_B^2 into G gives us a copy of P_B in G_B.

6.6.5 Finding a copy of C in G

Recalling how we 'chopped up' C at the start of this section, let $u, v \in V(G_B)$ be the vertices to which the endpoints of P_B were assigned. To complete the proof of this case we find a copy of P_A in $G_A := G[A' \cup \{u, v\}]$ starting at v and ending at u. We find a copy of P_A exactly as we found the copy of P_B with three differences. Firstly there are no exceptional vertices, except u and v and these are handled in a different manner. Secondly, recalling that

$$P_A := Q_1'P_1' \dots Q_{s_A}'P_{s_A}'Q^*P_A^*,$$

we require that the embeddings of Q_1' and P_A^* start and end at v and u respectively. Since Q_1' is long enough for Lemma 6.6 we can specify the cluster to which its initial vertex is assigned and use Lemma 6.6 to join it to P_1'. We embed P_A^* greedily and use Q^* and Lemma 6.6 to connect it with the rest of the embedding. Hence we can indeed start and end at the required vertices. This doesn't affect the constants in the rest of the proof. Since the number of exceptional vertices and the imbalances created by the approximate embedding are both small (and small as functions of M_A) we can proceed exactly as before and find the desired cycle C in G. The calculations work as before as a result of us only having two exceptional vertices. The equation (6.7) becomes

$$\begin{aligned} |a(i) - m_A| &\leq \left| a(i) - \frac{s_A t}{M_A} \right| + \varepsilon^* m_A \\ &\leq (\varepsilon^* s_A t + \varepsilon_A m_A + |\{u, v\}|) + \varepsilon^* m_A \leq 4\varepsilon_A m_A. \end{aligned}$$

Hence in Section 6.6.4 we now have

$$\sum_{i=1}^{M_A} |a(i) - m_A| \le 4\varepsilon_A M_A m_A,$$

which is fine as we will have that $n(i) \ge \lambda m_A/8 \ll 4\varepsilon_A M_A m_A$ for all clusters $V_i' \in V(R_A)$.

6.7 Cycle is Close to C^*

Our argument closely follows that in the previous section, the difference being in the means of correcting imbalances. To correct imbalances we will need long sections of P_B with no changes in orientation. Define $\ell_B := \lceil \frac{4}{\alpha} \rceil M_B$, which is at least the maximum length of a shifted walk between two vertices in R_B. As before we split up C into P_A and P_B, the only difference being that we do not need a special vertex v^* this time. Let $\mathcal{Q}_B'$ consists of a maximal collection of paths in P_B of length $3\ell_B$ all at a distance of at least 3 from each other, oriented in the same direction and containing no changes in orientation. We will call these *long runs*. There are at least

$$m(P_B, \mathcal{Q}_B') \ge \frac{n_B}{3\ell_B + 6} - 2\lambda n \ge \frac{\alpha n_B}{14 M_B}$$

of these in P_B. (We subtract $2\lambda n$ not λn as both neutral pairs $V_i V_{i+1} V_i$ and their inverse $V_i V_{i-1} V_i$ kill possible long runs.)

Let $\mathcal{Q}_B$ be the subset of $\mathcal{Q}_B'$ containing those long runs contained in the P_i, at a distance of at least 4 from the ends of all the P_i and all oriented in the same direction. We assume that these all are oriented in the same direction as F_B. Keeping only long runs oriented in one direction loses us at most half of them. The paths Q_i and the path $Q^* P_B^*$ (and the 3 vertices neighbouring them in the P_i in each direction) can intersect at most $2s + 2$ of the long runs and so, abusing notation slightly,

$$m(\mathcal{P}_B) \ge \frac{\alpha n_B}{28 M_B} - 2s - 2 \ge \frac{\alpha n_B}{30 M_B}$$

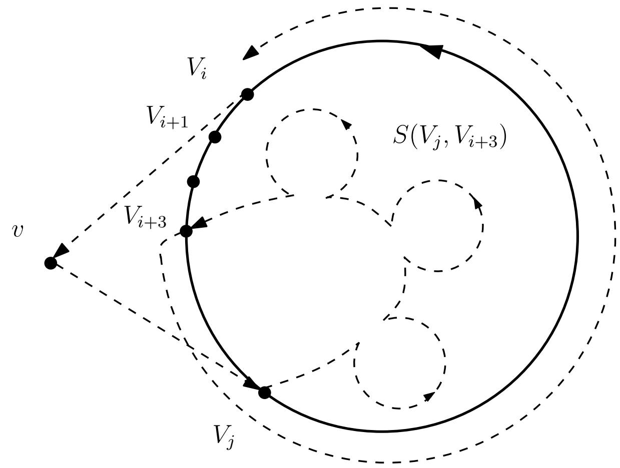


Figure 6.4: Incorporating an exceptional vertex when C is close to C^*.

for sufficiently large n, where we recall that $\mathcal{P}_B := \{P_1, P_2 \dots, P_{s_B}\}$. Similarly defining $\ell_A := \lceil \frac{4}{\alpha} \rceil M_A$ and $\mathcal{Q}'_A$ and $\mathcal{Q}_A$ in the obvious way we have $m(P_A) \geq \alpha(n_A)/30M_A$.

Apply Lemma 6.7 to R_B, $\mathcal{Q}_B$ and $\mathcal{P}_B$ with ε^* as γ to obtain an embedding of the P_i into $V(R_B)$ with

$$\left| a(i) - \frac{s_B t}{M_B} \right| \leq \varepsilon^* s_B t, \quad m(i) \geq \frac{\alpha n_B}{30 M_B^2} - \varepsilon^* s_B t \geq \frac{\alpha n_B}{32 M_B^2} \tag{6.9}$$

for all $V_i \in V(R_B)$, where we write $m(i)$ for the number of elements of $\mathcal{Q}_B$ whose initial vertex is assigned to $V_i \in V(R)$.

For all $2 \leq i \leq s_B$ we can join P_{i-1} and P_i by a path in R_B isomorphic to Q_i using Lemma 6.6. Furthermore we can greedily extend P_1 backwards by a path isomorphic to $P_B^* Q_1$. This will increase $a(i)$ by at most $s_B r + 3t \leq \varepsilon_A m_B$ for n sufficiently large. We now have an embedding of P_B into R_B which we can think of as a walk W_B in R_B.

Let G_B^*, G_B^c and R_B^* be defined exactly as in Section 6.6.2. Let $v \in V_0'$ be an exceptional vertex and let $V_i v, vV_j \in E(R_B^*)$. Take a long run in $\mathcal{Q}_B$ whose initial vertex is currently assigned to V_i. Since M_B divides ℓ_B it also ends at V_i. We cannot replace the long run simply by $V_i v V_j F_B \dots F_B$ because this would not end at V_i. Thus it would require us to alter the rest of our approximate embedding, possibly causing (6.9) to no longer hold. Instead we use shifted walks and a 'jump' to ensure that our modification incorporates v

into our walk and does not alter $a(i)$ or $m(i)$ significantly for any cluster of R_B. We replace the long run starting at V_i with the following walk

$$V_i v V_j S(V_j, V_{i+3}) F_B F_B \dots F_B V_i,$$

where $S(V_j, V_{i+3})$ is a shifted walk from V_j to V_{i+3}. The number of F_B is chosen so that the new section has exactly the same length as the long run it replaces. This is illustrated in Figure 6.4. This is a walk that goes out to v, back to V_j, follows a shifted walk to V_{i+3} and then winds around F until we have a walk of length $3\ell_B$ ending at V_i. This new walk visits V_{i+1} and V_{i+2} one time fewer than previously and V_j one time more. Observe that the shifted walk by definition visits every cluster in R_B the same number of times, which allows us to observe that we still end at V_i. Repeating this for each exceptional vertex creates a new assignment now satisfying

$$\begin{aligned} |a(i) - m_B| &\overset{(6.5)}{\leq} \left| a(i) - \frac{s_B t}{M_B} \right| + \left| \frac{s_B r + 2t}{M_B} \right| \\ &\leq (\varepsilon_A s_B t + \varepsilon_A m_B + |V_0'|) + \varepsilon_A m_B \leq 3\varepsilon_A n_B. \end{aligned}$$

for all i. We also still have a reasonable number of long runs starting at each cluster.

$$m(i) \geq \frac{\alpha n_B}{32M_B^2} - |V_0'| \geq \frac{\alpha n_B}{40M_B^2}.$$

Note that of the $a(i)$ vertices of P_B assigned to $V_i \in V(R)$, at most

$$\varepsilon_A m_B + 4|V_0'| \leq 5\varepsilon_A n_B$$

do not have their neighbours assigned to $V_{i-1} \cup V_{i+1}$. The first term here comes from connecting the P_i and the second term from incorporating the exceptional vertices: each exceptional vertex has one direct edge from a given cluster in R_B and the shifted walk can add at most two edges outside F_B to each cluster. Thus we currently have a $(3\varepsilon_A, 5\varepsilon_A)$-

corresponding assignment of P_B into R_B^*.

6.7.1 Correcting the imbalances

We now adjust our current assignment of P_B to R_B^* to obtain a $15\varepsilon_A$-corresponding assignment, i.e. we adjust W_B to ensure that $a(i) = m_B$ for all $V_i \in V(R_B)$. To do this we find a pair $V_i, V_j \in V(R_B)$ such that $a(i) > m_B$ and $a(j) < m_B$ and replace a long run starting at V_{i-1} with the following walk:

$$S(V_{i-1}, V_j)S(V_j, V_{i+1})F_B \dots F_B V_{i-1},$$

where the number of F_B is chosen to ensure that the new section has length $3\ell_B$. This walk removes the assignment of one vertex to V_i, assigns one extra vertex to V_j and does not change the number of vertices assigned to all other clusters in R_B. Since $\sum_{i=1}^{M_B} a(i) = m_B M_B$ we can always find such a pair unless we have corrected all the imbalances. Each pair requires a long run and we still have at least $\alpha n_B/40M_B^2 \gg 3\varepsilon_A n_B$ of these starting at each cluster and so can indeed correct all the imbalances. This leaves us with a balanced assignment with at most

$$3\varepsilon_A n_B + 4 \times 3\varepsilon_A n_B = 15\varepsilon_A n_B$$

edges outside F_B from each vertex. Hence there are at most $15\varepsilon_A M_B n_B \ll \varepsilon m_B$ edges in total not in a path of length at least 3 all of whose edges lie on F_B or not lying entirely on F_B. This is exactly the same position as in Section 6.6.4. We can now proceed as before to first find a copy of P_B in G_B and then repeat the procedure with P_A (using $\mathcal{Q}_A$ not $\mathcal{Q}_B$) to find the desired cycle C in G.

CHAPTER 7

PANCYCLICITY

7.1 Introduction

7.1.1 An Exact Pancyclicity Result

Building on the proof of Theorem 4.3, Keevash, Kühn and Osthus [46] recently gave an exact minimum semi-degree bound which forces a Hamilton cycle in an oriented graph. More precisely, they showed (Theorem 1.9) that every sufficiently large oriented graph G with $\delta^0(G) \geq (3n-4)/8$ contains a Hamilton cycle. This is best possible for all n and settles a problem of Thomassen. The arguments in [46] can straight-forwardly be modified to show that G even contains an ℓ-cycle through any given vertex for every $\ell \geq n/10^{10}$ and we do so in Section 7.2. Together with Theorems 5.3 and 5.4 this implies that G is *pancyclic*, i.e. it contains cycles of all possible lengths.

Theorem 7.1. *There exists an integer n_0 such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3n-4)/8$ contains an ℓ-cycle for all $3 \leq \ell \leq n$. Moreover, if $4 \leq \ell \leq n$ and if u is any vertex of G then G contains an ℓ-cycle through u.*

This improves a bound of Darbinyan [28], who proved that a minimum semi-degree of $\lfloor n/2 \rfloor - 1 \geq 4$ implies pancyclicity. Another degree condition which implies pancyclicity in oriented graphs which are close to being tournaments is given by Song [69]. Proposition 5.7

shows that we cannot have $\ell = 3$ in the 'moreover' part of Theorem 7.1.

For (general) digraphs, Thomassen [72] as well as Häggkvist and Thomassen [40] gave degree conditions which imply that every digraph with minimum semi-degree $> n/2$ is pancyclic. (The complete bipartite digraph whose vertex class sizes are as equal as possible shows that the latter bound is best possible.) Alon and Gutin [1] observed that one can use Ghouila-Houri's theorem [34] (which states that a minimum semi-degree of at least $n/2$ guarantees a Hamilton cycle in a digraph) to show that every digraph G with minimum semi-degree $> n/2$ is even vertex-pancyclic, i.e. for every $\ell = 2, \ldots, n$ each vertex of G lies on an ℓ-cycle.

7.1.2 Universal Pancyclicity

In Chapter 6 we discussed the following result on arbitrary orientations of Hamilton cycles in oriented graphs.

Theorem 7.2. *For every $\alpha > 0$ there exists an integer $n_0 = n_0(\alpha)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3/8 + \alpha)n$ contains every orientation of a Hamilton cycle.*

We also proved in Chapter 5 a result on arbitrary orientations of short cycles (Proposition 5.8).

In this section we extend Theorem 7.2 to a pancyclicity result for arbitrary orientations: If an oriented graph G on n vertices contains every possible orientation of an ℓ-cycle for all $3 \leq \ell \leq n$ we say that G is *universally pancyclic*. The following result says that asymptotically universal pancyclicity requires the same minimum semi-degree as pancyclicity.

Theorem 7.3. *For all $\alpha > 0$ there exists an integer $n_0 = n_0(\alpha)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq (3/8 + \alpha)n$ is universally pancyclic.*

In Section 7.3 we derive this universal pancyclicity result (Theorem 7.3) by combining the short-cycle result (Proposition 5.8) with a probabilistic argument applied to Theorem 7.2 giving all long cycles.

7.2 An Exact Result

With the results of Section 5.2 in mind, we are in a position to prove Theorem 7.1. The proof that this result holds for 'long' cycles uses somewhat similar methods to those in [46], and we will use some results from that paper. Using the 'stability method' we will distinguish between a non-extremal case where our oriented graph has some form of expansion property, and an extremal case where the oriented graph is shown to be similar to that in Figure 4.1.

We have already proved the result for $4 \leq \ell \leq n/10^{10}$ in Theorem 5.4 and the case $\ell = 3$ is dealt with by Theorem 1.4. Thus we can assume that $n/10^{10} \leq \ell < n$.

We will need the following slight extension of Lemma 3.2, due to Keevash, Kühn and Osthus [46].

Lemma 7.4. *For every $\varepsilon \in (0,1)$ there exists numbers $M' = M'(\varepsilon)$ and $n_0 = n_0(\varepsilon)$ such that the following holds. Let $d \in [0,1]$ with $\varepsilon \leq d/2$, let G be an oriented graph of order $n \geq n_0$ and let R' be the reduced digraph with parameters (ε, d) obtained by applying the Diregularity Lemma to G with M' as the lower bound on the number of clusters. Then R' has a spanning oriented subgraph R such that*

(a) $\delta^+(R) \geq (\delta^+(G)/|G| - (3\varepsilon + d))|R|$,

(b) $\delta^-(R) \geq (\delta^-(G)/|G| - (3\varepsilon + d))|R|$,

(c) *for all disjoint sets $S, T \subset V(R)$ with $e_G(S^*, T^*) \geq 3dn^2$ we have $e_R(S,T) > d|R|^2$, where $S^* := \bigcup_{i \in S} V_i$ and $T^* := \bigcup_{i \in T} V_i$.*

(d) *for every set $S \subset V(R)$ with $e_G(S^*) \geq 3dn^2$ we have $e_R(S) > d|R|^2$, where $S^* := \bigcup_{i \in S} V_i$.*

Define a hierarchy of constants so that

$$1/n_0 \ll \varepsilon \ll d \ll c \ll \eta \ll 1.$$

Let G be an oriented graph on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lceil (3n-4)/8 \rceil$ and let $u \in V(G)$. Suppose that G contains no cycle of length ℓ containing u. Apply the Diregularity Lemma (Theorem 3.1) and Lemma 7.4 to G with parameters ($\varepsilon^2/3$, d). This gives us a partition of $V(G)$ into $V_0, V_1, \dots, V_k$ with $m := |V_1| = \dots = |V_k|$ and a reduced oriented graph R on vertex set $\{1, 2, \dots, k\}$. Lemma 7.4 gives us that

$$\delta^0(R) > (3/8 - 1/(2n) - d - \varepsilon^2)k > (3/8 - 2d)k. \tag{7.1}$$

Case 1. *$|N_R^+(S)| \geq |S| + 2ck$ for every $S \subset [k]$ with $k/3 < |S| < 2k/3$.*

In this case we use probabilistic methods to find a subdigraph G' of G with ℓ vertices and a new reduced oriented graph which still satisfies the conditions of Case 1, possibly with modified constants. Also, we can ensure that $u \in V(G')$. We can then use the following result from [46], which says that all such graphs contain a Hamilton cycle.

Lemma 7.5. *Let M', n_0 be positive numbers and let $\varepsilon, d, \eta, \nu, \tau$ be positive constants such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \ll \nu \leq \tau \ll 1$. Let G be an oriented graph on $n \geq n_0$ vertices such that $\delta^0(G) \geq 2\eta n$. Let R' be the reduced digraph of G with parameters (ε, d) and such that $|R'| \geq M'$. Suppose that there exists a spanning oriented subgraph R of R' with $\delta^0(G) \geq \eta|R|$ and such that $|N_R^+(S)| \geq |S| + \nu|S|$ for all sets $S \subseteq V(R)$ with $|S| < (1-\tau)|R|$. Then G contains a Hamilton cycle.*

The argument we use to find an appropriate subdigraph G' is similar to that in [46], and uses standard probabilistic techniques. Recall that there are k (non-exceptional) clusters, each with size m.

Claim 1.1. *Let m' satisfy $10^{-11}n/k < m' < m$ and $p := m'/m$. Then there exists a partition of $V(G) \setminus V_0$ into sets A and B which has the following properties:*

(a) $|A_i| = m'$*, where we write* $A_i := V_i \cap A$ *for every* $i \in [k]$*;*

(b) $|N^+_G(v) \cap A_i| = p|N^+_G(v) \cap V_i| \pm n^{2/3}$ *for every vertex* $v \in V(G)$*; and similarly for* $N^-_G(v)$*;*

(c) R *is the oriented reduced graph with parameters* $(\varepsilon^2/10^{11}, 3d/4)$ *corresponding to the partition* $A_1, \dots, A_k$ *of the vertex set of* $G[A]$*;*

(d) $\delta^0(G[A]) \geq (3/8 - \varepsilon)|A|$*.*

Proof. For each cluster V_i define a partition into A_i and B_i as follows. Let $\eta := n^{2/3}/(4|V_i|)$ and put $x \in V_i$ in A_i with probability $p + \eta$, independently of all other vertices. Then standard Chernoff-type bounds give that the probability that $p|V_i| < |A_i| < p|V_i| + n^{2/3}/2$ does not occur is exponentially small in $|V_i|$. Further, they also give that the probability that any vertex $v \in A_i$ has outneighbourhood varying from $p|N^+_G(v) \cap V_i|$ by more than $n^{2/3}/2$ is exponentially small. Thus for sufficiently large n a partition exists satisfying both these conditions, and we can discard up to $n^{2/3}/2$ vertices from each A_i to obtain a partition satisfying (a) and (b).

To see (c) note that the definition of regularity implies that the pair (A_i, A_j) consisting of all the A_i-A_j edges in the pure oriented graph G^* is $\varepsilon^2/10^{11}$-regular and has density at least $3d/4$ whenever $ij \in E(R)$. On the other hand, (A_i, A_j) is empty whenever $ij \notin E(R)$ since $(V_i, V_j) \supset (A_i, A_j)$ is empty in this case. Property (d) follows immediately from (b).

□

If $\ell \geq n - |V_0|$ then form G' by discarding $n - \ell$ arbitrary vertices from $V(G) \setminus \{u\}$. Otherwise apply the previous claim to G with $m' := \lfloor \ell/k \rfloor - 1$. Let $G' := G[A \cup V_0']$, where $V_0' \subseteq V(G) \setminus A$ is an arbitrary set of vertices containing u (if $u \notin A$) of size $\ell - |A|$. Then G' has exactly ℓ vertices and satisfies the conditions of Lemma 7.5 with $\tau = 1/3$, $\eta = 1/6$ and $\nu = 2c$. Apply that result to obtain a Hamilton cycle in G' and thus a cycle of length ℓ through u in G.

Case 2. *There is a set* $S \subset [k]$ *with* $k/3 < |S| < 2k/3$ *and* $|N^+_R(S)| < |S| + 2ck$*.*

In this case we exploit the minimum semi-degree condition to demonstrate that G has roughly the same structure as the extremal graph. The proof proceeds in three steps.

(i) Show that the G has roughly the same structure as the extremal graph.

(ii) Show that if the cluster sizes and vertices satisfy certain conditions then using the Blow-up Lemma (Lemma 3.3) we have the desired cycle (Claim 2.3).

(iii) Use (ii) to obtain further structural refinements, eventually showing that G either contains a Hamilton cycle or contradicts the minimum semi-degree condition.

The difference between the proof here and the proof of the exact Hamiltonicity result in [46] is primarily in Step (ii), Claim 2.3. We have similar conditions here, but the stronger conclusion that we get a cycle of any length, not just a Hamilton cycle. Their proofs of the results needed for (iii) in the Hamiltonicity case implicitly require only that the conditions of (ii) are not satisfied, and so the proof of Step (iii) for us is implicit in their paper. Hence we will not give their proofs for either Step (i) or (iii). Instead we give a complete proof of the result in Step (ii) and refer the reader to [46] for all remaining details.

Let

$$A_R := S \cap N_R^+(S),\ B_R := N_R^+(S) \setminus S,\ C_R := [k] \setminus (S \cup N_R^+(S)),\ D_R := S \setminus N_R^+(S).$$

These sets will have similar properties as the sets A, B, C and D in the extremal example. Let $A := \bigcup_{i \in A_R} V_i$ and define B, C, D similarly. The following notation will prove useful. Let $P(1) := A$, $P(2) := B$, $P(3) := C$ and $P(4) := D$. When we refer to $P(i+1)$ or $P(i-1)$ we will always mean modulo 4. Define $P(i \oplus 1)$ by $P(1 \oplus 1) := P(1)$, $P(2 \oplus 1) := P(4)$, $P(3 \oplus 1) := P(3)$ and $P(4 \oplus 1) := P(2)$. This operation should be viewed with reference to the extremal graph as being the 'other' out-class of $P(i)$ (so C so A, D for B, A for C and B for D), and has the obvious inverse $P(1 \ominus 1) := P(1)$, $P(2 \ominus 1) := P(4)$, $P(3 \ominus 1) := P(3)$ and $P(4 \ominus 1) := P(2)$. Since we will show that G

has a somewhat similar structure to the extremal graph it will be useful to define the following graph on $V(G)$. Let $F[(P(i), P(i+1)]$ contain all edges from $P(i)$ to $P(i+1)$, let $F[A]$ and $F[C]$ be tournaments which are as regular as possible. Finally let $F[B, D]$ be a bipartite tournament which is as regular as possible. We will show that G roughly looks like F, and hence contains a cycle of length ℓ. From now on we will not calculate explicit constants multiplying c, and just write $O(c)$. The constants implicit in the $O(*)$ notation will always be absolute.

We call a vertex $x \in P(i)$ *cyclic* if it has almost the same number of neighbours in $P(i-1)$ and $P(i+1)$ as a vertex in the corresponding vertex class in F. More precisely, call a vertex $x \in P(i)$ *cyclic* if $|N_G^+(x) \cap P(i+1)| \geq (1 - O(\sqrt{c}))|P(i+1)|$ and $|N_G^-(x) \cap P(i-1)| \geq (1 - O(\sqrt{c})|)P(i-1)|$, counting modulo 4. A vertex is *acceptable* if it has a significant outneighbourhood in one of its two 'out-classes' and one of its two 'in-classes', where these are understood with reference to F. More precisely, $x \in P(i)$ is *acceptable* if both the following hold.

- $|N_G^+(x) \cap P(i+1)| \geq (1/100 - O(\sqrt{c}))n$ or $|N_G^+(x) \cap P(i \oplus 1)| \geq (1/100 - O(\sqrt{c}))n$,
- $|N_G^-(x) \cap P(i-1)| \geq (1/100 - O(\sqrt{c}))n$ or $|N_G^-(x) \cap P(i \ominus 1)| \geq (1/100 - O(\sqrt{c}))n$.

An edge from $P(i)$ to $P(j)$ in G is acceptable if $P(j) = P(i+1)$ or $P(j) = P(i \oplus 1)$.

The next claim combines several results from [46] and shows that these sets have roughly the same structure as in F.

Claim 2.2 (Keevash, Kühn and Osthus, [46])**.** *The following hold for all i.*

(a) $|P(i)| = (1/4 \pm O(c))n$,

(b) $e(P(i), P(i+1)) > (1 - O(c))n^2/16$,

(c) $e(P(i), P(i \oplus 1)) > (1/2 - O(c))n^2/16$.

Furthermore, by reassigning vertices that are not cyclic to A, B, C or D we can arrange that every vertex of G is acceptable. We can also arrange that there are no vertices that are not cyclic but would become so if they were reassigned.

Note that these properties of A, B, C and D are invariant under the relabelling $A \leftrightarrow C$, $B \leftrightarrow D$. Thus we may assume that $|B| \geq |D|$.

Given a path $P := v_1 \dots v_k$ in G with $v_1, v_k \in P(i)$ we say we *contract* P to refer to the following process, which yields a new digraph H. Remove $v_1, \dots, v_k$ from G and add an extra vertex v^* to $P(i)$ with outneighbourhood $N^+(v_k)$ and inneighbourhood $N^-(v_1)$. The 'moreover' part of the next claim is not in the statement of the corresponding claim in [46]. That we are not seeking a Hamilton cycle allows us this modified condition and a simpler proof than would otherwise be the case.

Claim 2.3. *If $|B| = |D|$ and every vertex is acceptable then G has an ℓ-cycle containing u. Moreover, the assertion also holds if we allow one non-acceptable vertex $x \in A \cup C$.*

Proof. The idea is as follows. First we contract suitable paths to leave us with a digraph G_1 containing only cyclic vertices. Then we find suitable paths to contract to give a digraph G_2 with $|A| = |B| = |C| = |D|$. We can then apply the Blow-up Lemma to the underlying graph to find a cycle in G_2 which 'winds around' A, B, C, D. By our choice of the vertices in this cycle and the definition of our contractions this will correspond to the desired cycle in G. We will say that a 4-partite graph with vertex classes $(P(1), P(2), P(3), P(4))$ has *type* (p_1, p_2, p_3, p_4) if $|P(i)| = p_i + q$ and $p_i \in \mathbb{N}$ for all i and some q. Our initial condition on the sizes means that G has type $(p_1, 0, p_3, 0)$. The *type sum* is $p_1 + p_2 + p_3 + p_4$.

Firstly, move the non-acceptable vertex x (if it exists) to a vertex class in which it is acceptable, and readjust the $O(c)$ notation if necessary. This gives us type $(p_1, 0 \leq p_2 \leq 1, p_3, 0)$, possibly with new values for the p_i. Let $v_1, \dots, v_t$ be vertices which are acceptable but not cyclic. Claim 2.2 (a) and (b) give us that $t = O(\sqrt{c})n$ (easily shown by counting edges), so we can pick cyclic neighbours v_i^+ and v_i^- of each v_i such that the edges $v_i v_i^+$ and $v_i^- v_i$ are acceptable and all these vertices are distinct. We want to contract $v_i^- v_i v_i^+$ so that we form a new graph in which all vertices are cyclic. We need to ensure that after contracting we are still of type $(p_1, 0 \leq p_2 \leq 1, p_3, 0)$ (although possibly with different p_i to above) and $p_1, p_3 = O(\sqrt{c})n$. For each v_i find a path P_i' of length at most 3 starting at v_i^+, ending at some cyclic vertex in the same cluster as v_i^- and 'winding around the clusters,' i.e. following the order $P(i)$, $P(i+1)$ etc. If v_i^+ and v_i^- are in the same cluster then the path P_i' is the empty path. Let $P_i := v_i^- v_i v_i^+ P_i'$ and note that we can choose the P_i to be disjoint.

Contract the paths P_i to form a new digraph G_1. Note that G_1 is not necessarily oriented. Every vertex in G_1 is cyclic by construction, possibly with a new constant in the $O(\sqrt{c})$ notation in the definition of a cyclic vertex. G_1 also has type $(p_1, 0 \leq p_2 \leq 1, p_3, 0)$ and $p_1, p_3 = O(\sqrt{c})n$.

Now suppose that $|A| < |C|$ and let $s := |C| - |A| = p_3 - p_1$. Greedily find a path P_C in G_1 which follows the pattern $CCDAB$ s times and then ends in C. I.e. find an edge between 2 cyclic vertices in C, extend around the clusters back to C and repeat until we have a path from C to C with s (cyclic) vertices from A, B and D and $2s+1$ vertices from C. We can do this as Claim 2.2 (a) and (c) imply that almost all unordered pairs of vertices in C are connected by an edge and $s = O(\sqrt{c})n$. Let G_2 be the digraph obtained by contracting P_C. Then in G_2 has type $(p_1, 0 \leq p_2 \leq 1, p_1, 0)$. If $|A| > |C|$ we can achieve type $(p_1, 0 \leq p_2 \leq 1, p_1, 0)$ in a similar way by contracting a path P_A from A to A following the pattern $AABCD$. Note that since $s = O(\sqrt{c})n$, all vertices of G_2 are still cyclic. Now suppose that in G_2 we have $|D| > |A|$. Let $s := |D| - |A| = -p_1$. This time we find a path P_D from D to D following the pattern $DBCDABDABC$ which contains $s+1$ more vertices from D than it contains from A, and similarly for C. Note that contracting P_D does not change $|B| - |D|$. Contracting P_D gives us a digraph (which we still call G_2) with type $(0, 0 \leq p_2 \leq 1, 0, 0)$ and all of whose vertices are still cyclic. The last case to consider is when we have $|D| < |A|$. In this case we can equalize the sets by contracting two paths P_A and P_C of appropriate length as above.

We now find and contract a short path in G_2 to form a new oriented graph G_3 with $|G_3| - n + \ell \equiv 0 \pmod 4$. Let $p := n - \ell \pmod 4$. This is (congruent to) the number of vertices we do not want in the cycle we will find in G_3. We now contract paths to ensure that G_3 has type sum p, and thus $|G_3| - n + \ell \equiv 0 \pmod 4$. Suppose G_3 has type $(0,0,0,0)$. If $p = 0$ we are done. If $p = 1$ use one path P_C and one path P_D as above to obtain type $(1,0,0,0)$. If $p = 2$ then a path P_D gives us type $(1,0,1,0)$ and finally if $p = 3$ a path P_C gives us type $(1,1,0,1)$. Now suppose G_3 has type $(0,1,0,0)$. If $p = 1$ we are done already. If $p = 2$ contract one path P_D and one path P_C to get type $(1,1,0,0)$. If $p = 3$ a path P_D gives us type $(1,1,1,0)$. Finally if $p = 4$ two paths P_D and one path P_C gives type $(2,1,1,0)$.

At most $O(\sqrt{c})n$ vertices in G_3 correspond to paths in G. Call these vertices and u *special vertices*. We now contract the special vertices. Let S_1 consist of the special vertices in A. Find a path from A to A that 'winds around' the 4 clusters of the oriented graph G_3 $|S_1|$ times and contains all vertices in S_1. As $|S_1| \leq O(\sqrt{c})n$ we can find such a path easily with a greedy algorithm. Contract this path and repeat for B, C and D to reduce the number of special vertices to at most 4 without otherwise affecting the structure of G_3. Let S consist of these remaining special vertices.

Let G_3' be the underlying graph corresponding to the set of edges oriented from $P(i)$ to $P(i+1)$, for $1 \leq i \leq 4$. Since all vertices of G_3 are cyclic and we chose $c \ll \eta \ll 1$, each pair $(P(i), P(i+1))$ is $(\eta, 1)$-super-regular in G_3'. Furthermore, G_3' contains no multiple edges. Let F' be the 4-partite graph with vertex classes $P(i)$ where the 4 bipartite graphs induced by $(P(i), P(i+1))$ are all complete. Define $\ell' := |F'| - n + \ell$ and note that it satisfies $\ell' \equiv 0 \pmod 4$ and $\ell'/4 \leq |D|$. Thus 'winding around' the 4 clusters $\ell'/4$ times we can find a cycle of length ℓ' in F' including all the special vertices. Note that we need $\ell < n$ here, since the one non-acceptable vertex means that we cannot ensure that G_3 has type $(0,0,0,0)$. Remove each special vertex $v_j \in S$ from this cycle to split the cycle into a series of disjoint paths $P_1 := v_1^+ P_1' v_2^-$, $P_2 := v_2^+ P_2' v_3^-$ etc. For each $v_j \in S \cap P(i)$ and every i pick sets $C_j^+ \subset N^+(v_j) \cap P(i+1)$ and $C_j^- \subset N^-(v_j) \cap P(i-1)$ of size $10^{-8}|G_3|$. We now apply Lemma 3.3 with $M = 4$, $\Delta = 2$, $b = 10^{-8}$ and the C_j^+ and C_j^- as the sets C_x (for $x \in \{v_1^+, v_2^-, v_2^+, \dots, v_1^-\}$) to embed the paths $P_1, \dots, P_{|S|}$. This gives us disjoint paths in $G_3' - S$ with endpoints in the C_j^+ and C_j^- and the sum of whose lengths is $|G_3| - n + \ell - 2|S|$. The 'moreover' part of Lemma 3.3 implies that we can assume that these paths continually 'wind around' A, B, C, D. The condition on the endpoints of the paths ensures that we can add in the special vertices to obtain a cycle C in G_3 of length $|G_3| - n + \ell$. As every vertex outside C in G_3 corresponds to a single vertex in G, the cycle C_ℓ in G corresponding to C has length ℓ and contains u. □

Since we are done if we satisfy the conditions of Claim 2.3, assume that $|B| > |D|$. The argument in [46] reaches a similar point to us here, and proceeds by showing that either G contains a Hamilton cycle, or is even more like the extremal graph. More precisely, they show that G either satisfies certain structural conditions, which we state below, or the conditions of Claim 2.3 are satisfied. They do this by moving vertices between clusters to obtain $|B| = |D|$ whilst ensuring that all vertices are acceptable. The situation can arise though that $|B| = |D| + 1$ and the only vertex class that it is possible to move vertices in B to without stopping them being acceptable is D. In this case we can shift an arbitrary vertex in B to $A \cup C$ to satisfy the conditions of Claim 2.3.

Claim 2.4. *For each of the following properties, there are fewer than $|B| - |D|$ vertices with that property or the conditions of Claim 2.3 are satisfied.*

- $x \in A$ *and* $|N^-(x) \cap C| \geq (1/100 - O(\sqrt{c}))n$.
- $x \in A$ *and* $|N^-(x) \cap B| \geq (1/100 - O(\sqrt{c}))n$.
- $x \in C$ *and* $|N^+(x) \cap A| \geq (1/100 - O(\sqrt{c}))n$.
- $x \in C$ *and* $|N^+(x) \cap B| \geq (1/100 - O(\sqrt{c}))n$.

We now define a new class of vertices. We say that a vertex is *good* if it is acceptable and satisfies one of the following.

- $x \in A$ and $|N^-(x) \cap C|$, $|N^-(x) \cap B| \leq (1/100 + O(\sqrt{c}))n$.
- $x \in B$ and $|N^+(x) \cap A|$, $|N^+(x) \cap B| \leq (1/100 + O(\sqrt{c}))n$ and $|N^-(x) \cap B|$, $|N^-(x) \cap C| \leq (1/100 + O(\sqrt{c}))n$.
- $x \in C$ and $|N^+(x) \cap A|$, $|N^+(x) \cap B| \leq (1/100 + O(\sqrt{c}))n$.

- $x \in D$.

Note that cyclic vertices are not necessarily good.

Claim 2.5. *By reassigning at most $O(\sqrt{c})n$ vertices we can arrange that every vertex is good or the conditions of Claim* 2.3 *are satisfied.*

Let M be a maximal matching in $e(B,A) \cup e(B) \cup e(C,A) \cup e(C,B)$.

Claim 2.6. $e(M) = 0$ *and* $|B| - |D| = 1$ *or the conditions of Claim* 2.3 *are satisfied.*

If the conditions of Claim 2.3 are satisfied we are done, so assume not. Since $e(M) = 0$ we now have $e(B \cup C, A) = 0$. Since $e(A) < |A|^2/2$ there exists a vertex $a \in A$ with $d^-(a) \leq (|A|-1)/2 + |D|$. Furthermore, we also now have that $e(C,B) = 0$ and $e(B) = 0$, and so there exist vertices $c \in C$ and $b \in B$ with $d^+(c) \leq (|C|-1)/2 + |D|$ and $d(b) \leq |A| + |C| + |D|$. Since $|D| = |B| - 1$ we see that

$$(3n-4)/2 \leq d^-(a) + d^+(c) + d(b) \leq \frac{3}{2}(|A| + |C| + 2|D|) - 1 = \frac{3}{2}(n-1) - 1.$$

This contradiction completes the proof. □

7.3 Proof of universal pancyclicity result

To deduce Theorem 7.3 from Theorem 7.2 and Proposition 5.8 we will use the following observation which is similar to one in [54].

Lemma 7.6. *There exists an integer n_1 such that the following holds for all $0 < \alpha < 1$. Suppose we are given an oriented graph G on $n \geq n_1$ vertices with minimum semi-degree $\delta^0(G) \geq (3/8 + \alpha - n^{-3/8})n$ where $n/2 \in \mathbb{N}$. Then there is a subset $U \subseteq V(G)$ of size $|U| = n/2 := u$ such that $\delta^0(G[U]) \geq (3/8 + \alpha - u^{-3/8})u$.*

To prove it we need a large deviation bound for the hypergeometric distribution (see e.g. [45, Theorem 2.10]).

Lemma 7.7. *Given $q \in \mathbb{N}$ and sets $A \subseteq T$ with $|T| \geq q$, let Q be a subset of size q of T chosen uniformly at random. Let $X := |A \cap Q|$. Then for all $0 < \varepsilon < 1$ we have*

$$\mathbb{P}[|X - \mathbb{E}(X)| \geq \varepsilon \mathbb{E}(X)] \leq 2 \exp\left(-\frac{\varepsilon^2}{3}\mathbb{E}(X)\right).$$

Proof of Lemma 7.6. Consider a subset U of vertices of G chosen uniformly at random from all subsets of $V(G)$ of size u. Let $\varepsilon := (1 - 2^{-3/8})u^{-3/8}$. Consider any vertex x of G and define a random variable $X^+ := |N^+(x) \cap U|$. Observe that $\varepsilon \mathbb{E}(X^+) \leq \varepsilon u = (1 - 2^{-3/8})u^{5/8}$ and hence

$$\mathbb{E}(X^+) \geq (3/8 + \alpha - n^{-3/8})u = (3/8 + \alpha - u^{-3/8})u + \varepsilon \mathbb{E}(X^+).$$

Then by Lemma 7.7 we have

$$\mathbb{P}[X^+ \leq (3/8 + \alpha - u^{-3/8})u] \leq \mathbb{P}[X^+ \leq (1 - \varepsilon)\mathbb{E}(X^+)] \leq 2 \exp\left(-\frac{(1 - 2^{-3/8})^2}{3u^{3/4}}\frac{u}{4}\right) \leq n^{-2}.$$

The final inequality holds since we assume n, and hence u, to be sufficiently large. The same bound holds when we consider inneighbourhoods of vertices. Hence with positive probability there exists a set $U \subseteq V(G)$ with the desired minimum semi-degree property. □

We are now in a position to derive Theorem 7.3.

Proof of Theorem 7.3. Given $\alpha > 0$, set $\ell_0 := \max\{n_0(\alpha/3), n_1, (6/\alpha)^{8/3}\}$, where n_0 is the function defined in Theorem 7.2 and n_1 is as in Lemma 7.6. Let $n \gg \ell_0, 1/\alpha$ and consider an oriented graph G on n vertices with minimum semi-degree $\delta^0(G) \geq (3/8+\alpha)n$. Choose any $3 \leq \ell \leq n$ and any orientation C of an ℓ-cycle. We have to show that G contains a copy of C. This is clear if $4 \leq \ell \leq \ell_0$, since $n \gg \ell_0, 1/\alpha$ and thus an application of Proposition 5.8 gives us C immediately, at least for $\ell \geq 4$. For $\ell = 3$ we can use Fact 5.12 to get that $N^+(v)$ is not independent for any $v \in V(G)$ and hence to find a transitive triangle. Theorem 5.3 gives us the directed 3-cycle.

So we may assume that $\ell > \ell_0$. Let k be an integer such that $2^k\ell \leq n < 2^{k+1}\ell$. A straightforward application of Lemma 7.7 implies the existence of a subgraph G' of G on $n' := 2^k\ell$ vertices with $\delta^0(G') \geq (3/8 + \alpha/2)n'$. Apply Lemma 7.6 k times to obtain a subgraph G'' of G' on ℓ vertices with $\delta^0(G'') \geq (3/8 + \alpha/2 - \ell^{-3/8})\ell \geq (3/8 + \alpha/3)\ell$. Since $\ell > n_0(\alpha/3)$ we can now apply Theorem 7.2 to obtain a Hamilton cycle oriented as C in G'' and hence the desired orientation of an ℓ-cycle in G. $\square$

CHAPTER 8

OPEN PROBLEMS

8.1 Long Cycles

There are two natural directions in which to extend our work on Hamilton cycles in oriented graphs. Firstly, we can seek stronger sufficient conditions for the existence of such a Hamilton cycle in an oriented graph. The best-known open problem here is the Nash-Williams conjecture, which would provide a digraph analogue of Chvátal's theorem for digraphs. If the *degree sequence* $d_1 \leq d_2 \leq \ldots \leq d_n$ of a graph satisfies $d_k \geq k+1$ or $d_{n-k} \geq n-k$ whenever $k < n/2$ then Chvátal's theorem tells us that it is Hamiltonian. For digraphs we need two sequences: $d_1^+ \leq d_2^+ \leq \ldots \leq d_n^+$ for the out-degree sequence and $d_1^- \leq d_2^- \leq \ldots \leq d_n^-$ for the in-degree sequence.

Conjecture 8.1. *Let G be a strongly connected digraph of order n and suppose that for all $k < n/2$*

(i) $d_k^+ \geq k+1$ *or* $d_{n-k}^- \geq n-k$ *and*

(ii) $d_k^- \geq k+1$ *or* $d_{n-k}^+ \geq n-k$.

Then G contains a Hamilton cycle.

An approximate version of Conjecture 8.1 for large digraphs was recently proved by Kühn, Osthus and Treglown [56]:

Theorem 8.2. *For every $\eta > 0$ there exists an integer $n_0 = n_0(\eta)$ such that the following holds. Suppose G is a digraph on $n < n_0$ vertices such that for all $k < n/2$*

(i) $d_k^+ \geq k + \eta n$ *or* $d_{n-k-\eta n}^- \geq n - k$ *and*

(ii) $d_k^- \geq k + \eta n$ *or* $d_{n-k-\eta n}^+ \geq n - k$.

Then G contains a Hamilton cycle.

It is natural to ask whether this could be made exact and the error terms removed.

The other direction in which the question can be extended is to expand the class of cycles we are seeking. In Chapter 6 we have done so, looking at arbitrary orientations of Hamilton cycles. The obvious open problem here is whether Theorem 6.2, our result on arbitrary orientations of Hamilton cycles, can this be made exact and the error term removed. The first step to doing so, and an interesting question in its own right, is likely to be obtaining an understanding of the extremal oriented graphs. That is, what do those oriented graphs who almost (in some appropriately defined sense) satisfy the minimum semi-degree condition of Theorem 6.2 but do not contain some orientation of a Hamilton cycle look like? It is not clear that this is a simple family, as is the case with the standard orientation where we have the example of Häggkvist (Figure 4.1).

8.2 Short Cycles

With short cycles the first open problem is obvious: we have not solved our own conjecture.

Conjecture 8.3. *Let $\ell \geq 4$ be a positive integer and let $k \geq 3$ be minimal such that k does not divide ℓ. Then there exists an integer $n_0 = n_0(\ell)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lfloor n/k \rfloor + 1$ contains an ℓ-cycle.*

As discussed in Section 5.1.2, there is a natural strengthening of this conjecture to arbitrary orientations of cycles.

Conjecture 8.4. *Let C be an arbitrarily oriented cycle of length $\ell \geq 4$ and cycle-type $t(C) \geq 4$. Let k be the smallest integer which is greater than 2 and does not divide $t(C)$. Then there exists an integer $n_0 = n_0(\ell, k)$ such that every oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lfloor n/k \rfloor + 1$ contains C.*

Some of our partial results towards these conjectures require use of Lemma 5.19, which says that if we allow a linear 'error term' in the degree conditions then instead of finding an ℓ-cycle, it suffices to look for a closed walk of length ℓ. The proof of this lemma is a standard application of the Regularity lemma. It would be interesting to find a proof which avoids the Regularity lemma. This would probably allow some of our partial results to be applied to much smaller graphs than is the case at present, as well as being an interesting result itself.

BIBLIOGRAPHY

[1] N. Alon and G. Gutin, Properly colored Hamilton cycles in edge colored complete graphs, *Random Structures and Algorithms* **11** (1997), 179–186.

[2] N. Alon and A. Shapira, Testing subgraphs in directed graphs, *Journal of Computer and System Sciences* **69** (2004), 354–382.

[3] N. Alon and J. Spencer, *The Probabilistic Method* (2nd edition), Wiley-Interscience 2000.

[4] B. Andrásfai, P. Erdős and V.T. Sós, On the connection between chromatic number, maximal clique and minimal degree of a graph, *Discrete Math.* **8** (1974), 205–218.

[5] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer 2000.

[6] J. Bang-Jensen and G. Gutin, Paths, trees and cycles in tournaments, *Surveys in Graph Theory* **115** (1996), 131–170.

[7] J. Bang-Jensen, G. Gutin and H. Li, Sufficient conditions for a digraph to be Hamiltonian, *Journal of Graph Theory* **22(2)** (1996), 181–187.

[8] M. Behzad, G. Chartrand and C. Wall, On minimal regular digraphs with given girth, *Fundamenta Mathematicae* **69** (1970), 227–231.

[9] B. Bollobás, *Modern Graph Theory*, Springer-Verlag 1998.

[10] B. Bollobás, *Random Graphs* (2nd edition), Cambridge University Press 2001.

[11] B. Bollobás and A. Frieze, On matchings and Hamilton cycles in random graphs, *Annals of Discrete Mathematics* **28** (1985), 23-46.

[12] B. Bollobás and R. Häggkvist, Powers of Hamilton cycles in tournaments, *Journal of Combinatorial Theory, Series B* **50(2)** (1990), 309–318.

[13] J.A. Bondy and C. Thomassen, A short proof of Meyniel's theorem, *Discrete Mathematics* **19(2)** (1977), 195–197.

[14] L.M. Bregman, Some properties of nonnegative matrices and their permanents, *Soviet Mathematics Doklady* **14** (1973), 945–949.

[15] W.G. Brown and F. Harary, Extremal digraphs, in *Combinatorial Theory and its Applications*, János Bolyai Mathematical Society 1970, 135–198.

[16] W.G. Brown and M. Simonovits, in *Extremal Multigraph and Digraph Problems*, János Bolyai Mathematical Society 2002, 157–203.

[17] A. Busch, A note on the number of Hamiltonian paths in strong tournaments, *The Electronic Journal Of Combinatorics* **13(1)** (2006).

[18] L. Caccetta and R. Häggkvist, On minimal graphs with given girth, in *Proceedings of the Ninth Southeastern Conference on Combinatorics, Graph Theory and Computing*, Utilitas Mathematica 1978, 181–187.

[19] D. Christofides, P. Keevash, D. Kühn and D. Osthus, Finding Hamilton cycles in robustly expanding digraphs, preprint.

[20] D. Christofides, P. Keevash, D. Kühn and D. Osthus, A semi-exact degree condition for Hamilton cycles in digraphs, preprint.

[21] M. Chudnovsky, P. Seymour and B. Sullivan, Cycles in dense graphs, *Combinatorica* **28(1)** (2008), 1–18.

[22] V. Chvátal and E. Szemerédi, Short cycles in directed graphs, *Journal of Combinatorial Theory Series B* **35(3)** (1983), 323–327.

[23] C. Cooper and A. Frieze, Hamilton cycles in a class of random directed graphs, *Journal of Combinatorial Theory Series B* **62(1)** (1991), 151–163.

[24] C. Cooper and A. Frieze, Hamilton cycles in random graphs and directed graphs, *Random Structures and Algorithms* **16(4)** (2000), 369–401.

[25] C. Cooper, A. Frieze and M. Molloy, Hamilton cycles in random regular digraphs, *Combinatorics, Probability and Computing* **3** (1994), 39–49 .

[26] B. Csaba, On the Bollobás-Eldridge conjecture for bipartite graphs, *Combinatorics, Probability and Computing* **16** (2007), 661–691.

[27] B. Csaba, A. Shokoufandeh, and E. Szemerédi, Proof of a conjecture of Bollobás and Eldridge for graphs of maximum degree three, *Combinatorica* **23** (2003), 35–72.

[28] S. Darbinyan, Pancyclicity of digraphs with large semidegrees, *Akademiya Nauk Armyanskoĭ SSR. Doklady* **80** (1985), 51–54.

[29] G. Fan, New sufficient conditions for cycles in graphs, *Journal of Combinatorial Theory, Series B* **37(3)** (1984), 221–227.

[30] E. Friedgut and J. Kahn, On the number of Hamilton cycles in a tournament, *Combinatorics, Probability and Computing* **14** (2005), 769–781.

[31] P. Fraisse and C. Thomassen, Hamiltonian dicycles avoiding prescribed arcs in tournaments, *Graphs and Combinatorics* **3** (1986), 239–250.

[32] A. Frieze, An algorithm for finding Hamilton cycles in random directed graphs, *Journal of Algorithms* **(9)** (1988), 181–204.

[33] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, *Journal of Combinatorial Theory, Series B* **94** (2005), 159–172.

[34] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, *Comptes Rendus Mathématique de l'Académie des Sciences de Paris* **25** (1960), 495–497.

[35] M. de Graaf, A. Schrijver, and P. D. Seymour, Directed triangles in directed graphs, *Discrete Mathematics* **110** (1992), 279–282.

[36] D. Grant, Antidirected Hamiltonian cycles in digraphs, *Ars Combinatoria* **10** (1980), 205–209.

[37] R. Häggkvist, Hamilton cycles in oriented graphs, *Combinatorics, Probability and Computing* **2** (1993), 25–32.

[38] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in digraphs, *Journal of Graph Theory* **20** (1995), 471–479.

[39] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in oriented graphs, in *Combinatorics, Geometry and Probability*, Cambridge University Press 1997, 339–353.

[40] R. Häggkvist and C. Thomassen, On pancyclic digraphs, *Journal of Combinatorial Theory, Series B* **20(1)** (1976), 20–40.

[41] P. Hamburger, P. Haxell and A. Kostochka, On directed triangles in digraphs, *The Electronic Journal of Combinatorics* **14** (2007), Note 19, 9 pp. (electronic).

[42] Y. Hamidoune, An application of connectivity theory in graphs to factorizations of elements in groups, *European Journal of Cominatorics* **2(4)** (1981), 349–355.

[43] J. Hladký, D. Král and S. Norine, Counting Flags in Triangle-Free Digraphs, *ArXiv Mathematics e-prints* (2009).

[44] C. Hoang and B. Reed, A note on short cycles in digraphs, *Discrete Mathematics* **66(1-2)** (1987), 103–107.

[45] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience 2000.

[46] P. Keevash, D. Kühn and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, *Journal of the London Mathematical Society* **79** (2009), 144–166.

[47] L. Kelly, Arbitrary orientations of Hamilton cycles in oriented graphs, preprint.

[48] L. Kelly, D. Kühn and D. Osthus, A Dirac type result on Hamilton cycles in oriented graphs, *Combinatorics, Probability and Computing* **17** (2008), 689–709.

[49] L. Kelly, D. Kühn and D. Osthus, Cycles of given length in oriented graphs, *Journal of Combinatorial Theory, Series B*, to appear.

[50] J. Komlós, The Blow-up lemma, *Combinatorics, Probability and Computing* **8** (1999), 161–176.

[51] J. Komlós, G. Sárközy, and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.

[52] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, in *Paul Erdős is Eighty (Vol. 2)*, János Bolyai Mathematical Society 1996, 295–352.

[53] A. Kotzig, The decomposition of a directed graph into quadratic factors, *Acta Facultatis, Rerum Naturalium, Universitatis Comenianae, Mathematica* **22** (1969), 27–29.

[54] D. Kühn and D. Osthus, Loose Hamilton cycles in 3-uniform hypergraphs of large minimum degree, *Journal of Combinatorial Theory, Series B* **96** (2006), 767–821.

[55] D. Kühn and D. Osthus, Embedding large subgraphs into dense graphs, in *Surveys in Combinatorics 2009*, Cambridge University Press 2009, 137–167.

[56] D. Kühn, D. Osthus and A. Treglown, Hamiltonian degree sequences in digraphs, preprint.

[57] D. Kühn, D. Osthus and A. Treglown, Hamilton decompositions of regular tournaments, preprint.

[58] D. Kühn, D. Osthus and A. Young, k-ordered Hamilton cycles in digraphs, *Journal of Combinatorial Theory, Series B* **98(6)** (2008), 1165–1180.

[59] C. McDiarmid, General percolation and random graphs, *Advances In Applied Probability* **13** (1981), 40–60.

[60] M. Meyniel, Une condition suffisante d'existence d'un circuit hamiltonien dans un graphe orient, *Journal of Combinatorial Theory, Series B* **14** (1973), 137–147.

[61] M. Molloy and B. Reed, *Graph Colouring and the Probabilistic Method*, Springer 2002.

[62] J. Moon, On subtournaments of a tournament, *Canadian Mathematical Bulletin* **9** (1966), 297–301.

[63] J. Moon, *Topics on tournaments*, Holt, Rinehart and Winston 1968.

[64] M. Nathanson, The Caccetta-Häggkvist Conjecture and additive number theory, *ArXiv Mathematics e-prints* (2006).

[65] T. Nishimura, Short cycles in digraphs, *Proceedings of the First Japanese Conference on Graph Theory and Applications* **72** (1988), 295–298.

[66] Ø. Ore, Note on Hamiltonian circuits, *American Mathematics Monthly* **67** (1960), 55.

[67] J. Shen, Directed triangles in digraphs, *Journal of Combinatorial Theory, Series B* **74(2)** (1998), 405–407.

[68] J. Shen, On the girth of digraphs, *Discrete Mathematics* **211** (2000), 167–181.

[69] Z. Song, Pancyclic oriented graphs, *Journal of Graph Theory* **18(5)** (1994), 461–468.

[70] E. Szemerédi, Regular partitions of graphs, in *Problèmes combinatoires et théorie des graphes*, CNRS 1978, 399–401.

[71] A. Thomason, Paths and cycles in tournaments, *Transactions of the American Mathematical Society* **296** (1986), 167–180.

[72] C. Thomassen, An Ore-type condition implying a digraph to be pancyclic, *Discrete Mathematics* **19(1)** (1977), 85–92.

[73] C. Thomassen, Long cycles in digraphs with constraints on the degrees, in *Surveys in Combinatorics*, Cambridge University Press 1979, 211–228.

[74] C. Thomassen, Long cycles in digraphs. *Proceedings of the London Mathematical Society* **42** (1981), 231–251.

[75] C. Thomassen, Edge-disjoint Hamiltonian paths and cycles in tournaments, *Proceedings of the London Mathematicsl Society* **45(1)** (1982), 151–168.

[76] C. Thomassen, Hamiltonian-connected tournaments, *Journal of Combinatorial Theory, Series B* **45(1)** (1982), 151–168.

[77] C. Thomassen, Connectivity in tournaments, in *Graph Theory and Combinatorics*, Cambridge University Press 1983, 305–313.

[78] T. Tillson, A Hamiltonian decomposition of K^*_{2m}, $2m \geq 8$, *Journal of Combinatorial Theory, Series B* **29(1)** (1980), 68–74.

[79] A. Young, Extremal problems for dense graphs and digraphs, PhD thesis, School of Mathematics, University of Birmingham 2007.