London Mathematical Society Monographs No.11
Series Editors. P.M. Cohn and G.E.H. Reuter

Extremal Graph Theory

Béla Bollobás

Academic Press

London New York San Francisco

EXTREMAL GRAPH THEORY

BÉLA BOLLOBÁS

Department of Pure Mathematics
and Mathematical Statistics
University of Cambridge
Cambridge, England

1978

ACADEMIC PRESS

LONDON NEW YORK SAN FRANCISCO

A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

Library of Congress Catalog Card Number: 77–76680
ISBN: 0–12–111750–2

PRINTED IN GREAT BRITAIN BY
PAGE BROS (NORWICH) LTD, MILE CROSS LANE, NORWICH

To My Parents

Preface

There is Pleasure sure,
In being Mad, which none but Madmen know!
John Dryden's "The Spanish Friar"

Extremal graph theory, in its strictest sense, is a branch of graph theory developed and loved by Hungarians. Its study, as a subject in its own right, was initiated by Turán in 1940, although a special case of his theorem and several other extremal results had been proved many years earlier. The main exponent has been Paul Erdös who, through his many papers and lectures, as well as uncountably many problems, has virtually created the subject. (In retrospect, it seems inevitable that it was Erdös who, when I was fourteen, introduced me to graph theory.)

In extremal graph theory one is interested in the relations between the various graph invariants, such as order, size, connectivity, minimum degree, maximum degree, chromatic number and diameter, and also in the values of these invariants which ensure that the graph has certain properties. Often, given a property $\mathscr{P}$ and an invariant μ for a class $\mathscr{H}$ of graphs, we wish to determine the least value m for which every graph G in $\mathscr{H}$ with $\mu(G) > m$ has property $\mathscr{P}$. Those graphs G in $\mathscr{H}$ without the property $\mathscr{P}$ and with $\mu(G) = m$ are called the *extremal graphs* for the problem. For instance, every graph of order n and size at least n contains a cycle, and the extremal graphs are the trees of order n. At a slightly less frivolous level, a graph of order $2u$ contains a triangle if the minimum degree is at least $u + 1$, and the only extremal graph is $K^{u,u}$, the complete bipartite graph. The prime example of an extremal problem is the following: given a graph F, determine $ex(n; F)$, the maximum number of edges in a graph of order n not containing F as a subgraph.

Having said this, I hasten to emphasize that in this book extremal graph theory is interpreted in a much broader sense, including in its scope various structural results and any relations among the invariants of a graph, especially those concerned with best possible inequalities. The chapter titles give a broad outline of the content of the text and, although most of the material appears here for the first time in a book, the topics covered in most standard treatises on graph theory are also dealt with in depth. The most notable omissions are algebraic graph theory, matroids and the problems of enumeration and reconstruction. The relative importance of the topics covered in the different chapters is not reflected in their lengths; otherwise the results concerning Hamiltonian cycles, colouring graphs on surfaces, and graphs without certain subgraphs would take up most of the space. Inevitably, the selection of material and its presentation have been greatly influenced by my personal preferences.

The readers are expected to have some familiarity with graph theory, though the book is self-contained. It has grown out of two Part III courses given at the University of Cambridge (1970/71 and 1975/76) and is intended for research students and professional mathematicians. It seemed desirable to expand the lecture notes into a book, since even expert graph theorists seem to be unaware of quite a few of the results which were proved years ago. I hope the book will help a little to stem the present tide of duplications. Although it is exciting to introduce new concepts and to find new problems, there is also merit in the continuity and development of a theory. The main aim of this book is to bring readers up to date with the results in a number of areas and to entice at least a few of them to continue the work. There is a false myth that extremal results are rapidly superseded. I hope that this book will help to make the myth a reality.

I would like to emphasize that the proofs of the results are important; though it is easy to flip through the book and pick out some results, in many cases it is more advantageous to be familiar with the methods than to know the results. The exercises at the end of each section vary a great deal in importance and difficulty. They contain many results and a few, marked with the sign '+', are really too difficult to be called exercises. In many cases hints are given to bring the problem within reach. Very easy problems are marked with the sign '−'. Unresolved questions are called Problems.

The end of a proof or the absence of a proof is indicated by the symbol ■; the greatest integer less than or equal to x is denoted by $\lfloor x \rfloor$ and $-\lfloor -x \rfloor$ is denoted by $\lceil x \rceil$.

It is a great pleasure to acknowledge the generous help of Professors P. Erdös, G. A. Dirac, R. K. Guy, R. Halin, N. Sauer and C. Thomassen. My research students, Stephen Eldridge and Andrew Thomason, made many helpful suggestions. In addition Keith Carne, Andrew Cornford, Michael

Davies, Donald Duval, David Goto, Ian Holyer, David Jackson, Terry Lyons, Richard Mason, Richard Nowakowski, Jonathan Partington, Richard Pinch, John Rickard, Geoff Thorpe and Antony Wassermann helped me find some of the mistakes; for the many which undoubtedly remain I apologise. For six months while writing this book I enjoyed the hospitality of mathematicians at the University of Calgary to whom I am very grateful. I acknowledge with thanks the excellent typing of Mrs Joan Scutt and Miss Karen McDermid.

I am especially grateful to my wife, without whose patience and understanding this book would never have been written. My research students in analysis also had to put up with a lot during the later stages of producing the book.

Finally, I would like to thank the editor of the series, Professor P. M. Cohn, for his speed and efficiency in handling the manuscript and for his help with the proofs, often beyond the call of duty.

Cambridge B.B.
January 1978

Basic Definitions

The brass band stirred themselves, took a deep breath and played through the "International" three times without a break.

I. A. Ilf and E. P. Petrov; "The Twelve Chairs".

Some of the concepts occurring in this book have a set theoretical or topological flavour. However, most of the structures we investigate are *finite* and every problem we discuss is free of set theoretical and topological difficulties. In view of this we try to avoid pretentious notations and keep the definitions as pedestrian as possible. Sometimes we carry this to the extent of abusing the notation slightly. It is unlikely that many of the readers are unfamiliar with the basic concepts of graph theory but to make sure that we speak the same language we run through the definitions needed in the sequel. In order to help the reader familiarize himself with the definitions we shall mention a few results as well. These results are hardly more than simple observations. For the convenience of the reader some of the definitions will be repeated in the chapter they are most used.

Unless otherwise stated every set is finite. The number of elements of a set X is denoted by $|X|$. If $|Y| = r$ then we say that Y is an r-set. If furthermore $Y \subset X$ then Y is an r-subset of X. The set of r-subsets of a set X is denoted by $X^{(r)}$, i.e. $X^{(r)} = \{Y: Y \subset X, |X| = r\}$. A *graph* G is an ordered pair of disjoint sets (V, E) such that $E \subset V^{(2)}$ and $V \neq \emptyset$. The set V is the set of *vertices* of G and E is the set of *edges*. An edge $\{x, y\}$ is said *to join* the vertex x to the vertex y and is denoted by xy. We also say that x and y are *adjacent* vertices and the vertex x is *incident* with the edge xy. Two distinct edges with a common endvertex are *adjacent*. Two graphs are *isomorphic* if there exists a 1–1 correspondence between their vertex sets that preserves adjacency. Usually we do not distinguish between isomorphic graphs, unless we want to specify the vertices and edges. This is reflected in the convention that if G and H are isomorphic graphs then we write $G \cong H$ or simply $G = H$.

The vertex set of a graph G is denoted by $V(G)$ and the edge set by $E(G)$; if there is no danger of ambiguity, these are abbreviated to V and E. Even more, if the letter G occurs without any explanation then it stands for an *arbitrary graph.* Instead of $x \in V(G)$ we usually write $x \in G$ to denote that x is a vertex of G. In the same spirit the number of vertices of G, called the *order* of G, is denoted by $|G|$. A graph of order 1 is said to be *trivial.* The number of edges of G, called the *size* of G, is denoted by $e(G)$. We use the notation G^n to denote an *arbitrary* graph of order n. Similarly $G(n, m)$ denotes an *arbitrary* graph of order n and size m. The class of graphs of order n is $\mathcal{G}^n$.

A graph $G' = (V', E')$ is a *subgraph* of a graph $G = (V, E)$ if $V' \subset V$ and $E' \subset E$. In this case we write $G' \subset G$. If $V' = V$ then G' is a *factor* of G. If $W \subset V$ then the graph $(W, E \cap W^{(2)})$ is said to be the subgraph *induced* or *spanned* by W, and is denoted by $G[W]$. We say that H is an *induced subgraph* of G if $H \subset G$ and $H = G[V(H)]$.

The set of vertices adjacent to a vertex $x \in G$ is denoted by $\Gamma(x)$ and $d(x) = |\Gamma(x)|$ said to be the *degree* of x. If it is not clear which is the underlying graph, we put the symbol of the graph into the *suffix* of the appropriate symbol. Thus, if H is an induced subgraph of G and $x \in H$ then

$$\Gamma_H(x) = \Gamma_G(x) \cap V(H) = \Gamma(x) \cap V(H) \quad \text{and} \quad d_H(x) = |\Gamma_H(x)|.$$

For $W \subset V(G)$ we put $\Gamma(W) = \cup\{\Gamma(x): x \in W\}$. The *minimum degree* of the vertices of G is denoted by $\delta(G)$ and the *maximum degree* is denoted by $\Delta(G)$. If $\delta(G) = \Delta(G) = k$, i.e. every vertex of G has degree k, then G is said to be *regular of degree k* or *k-regular.* A 3-regular graph is said to be *cubic.*

If $E' \subset E(G)$ then $G - E'$ denotes the graph resulting from G if we omit the edges belonging to E', i.e. $G - E' = (V(G), E(G) - E')$. Similarly, if $W \subset V(G)$ then $G - W$ is the graph obtained from G by the removal of the vertices belonging to W. Of course, if a vertex $x \in W$ was to be omitted then every edge incident with x was to be omitted as well, i.e. if $G = (V, E)$ then $G - W = (V - W, E \cap (V - W)^{(2)})$. If $W = \{x\}$ we usually write $G - x$ instead of $G - \{x\}$, analogously we may write $G - xy$ instead of $G - \{xy\}$, $xy \in E$. If $H \subset G$ then we may write $G - H$ instead of $G - V(H)$. If $xy \in V^{(2)} - E$ then the graph obtained from G by addition of the edge xy is $G + xy = (V, E \cup \{xy\})$. We might use similar notation for the addition of vertices.

If $V(G) = \{x_1, x_2, \dots, x_n\}$ then $(d(x_i))_1^n$ is said to be a *degree sequence* of G. Usually we order the vertices in such a way that the degree sequence is monotone increasing or monotone decreasing. Clearly

$$\sum_1^n d(x_i) = 2e(G)$$

so if $(d_i)_1^n$ is a degree sequence of a graph then

$$\sum_1^n d_i \equiv 0 \pmod 2. \tag{0.1}$$

Let x and y be two not necessarily different vertices of G. By an *x–y walk W* we mean an alternating sequence of vertices and edges, say $x_1, \alpha_1, x_2, \alpha_2, \ldots, x_l, \alpha_l, x_{l+1}$, such that $x_1 = x$, $x_{l+1} = y$ and $\alpha_i = x_i x_{i+1} \in E(G)$, $1 \leqslant i \leqslant l$. We usually put $W = x_1 x_2 \ldots x_{l+1}$ since from this form it is clear which are the edges in the sequence. The *length* of this walk W is l. The vertex set of W is $V(W) = \{x_i : 1 \leqslant i \leqslant l+1\}$ and the edge set of W is $E(W) = \{\alpha_i : 1 \leqslant i \leqslant l\}$. The walk above is a *trail* if all its edges are distinct and it is a *path*-or $x_1 - x_{l+1}$ *path* if all its vertices are distinct. A trail whose endvertices coincide is a *circuit*. If $l \geqslant 3$, $x_1 = x_{l+1}$ but the other vertices are distinct from each other and x_1 then we call the walk a *cycle*.

This cycle is usually denoted by $x_1 x_2 \ldots x_l$ (instead of $x_1 x_2 \ldots x_l x_1$). A path P and a cycle C are identified with the graphs $(V(P), E(P))$ and $(V(C), E(C))$. In particular, $x_1 x_2 \ldots x_{l+1}$ and $x_{l+1} x_l \ldots x_1$ denote the same path, so an x–y path is also a y–x path. Similarly $x_1 x_2 \ldots x_l$ and $x_2 x_3 \ldots x_l x_1$ denote the same cycle. An edge of the form $x_1 x_j$ $(3 \leqslant j \leqslant l-1)$ is a *diagonal* of this cycle. We denote by P^l a *path of length l* and by C^l a *cycle of length l*. We call C^3 a *triangle*, C^4 a *quadrilateral*, C^5 a *pentagon*, etc. A cycle is *odd* (*even*) if its length is odd (even).

If $P = x_1 x_2 \ldots x_{l+1}$ is a path, $u = x_i$, $v = x_j$ and $1 \leqslant i < j \leqslant l+1$ then the *$u - v$ segment of P* is the $u - v$ path $x_i x_{i+1} \ldots x_{j-1} x_j$. We denote it by uPv. If P is an x–y path and Q is a y–z path then $xPyQz$ is the x–z walk obtained by stringing the two paths together. Similarly we may string together *segments* of paths to obtain a walk or a path with the self-explanatory notation $x_1 P_1 x_2 P_2 x_3 \ldots P_l x_{l+1}$, where $x_i x_{i+1} \in V(P_i)$, $i = 1, 2, \ldots, l$.

If $x \in X$ and $y \in Y$ then an x–y path is also said to *X–Y path*. Similarly $\alpha \in E(G)$ is a X–Y edge if $\alpha = xy$ and $x \in X$, $y \in Y$. The number of X–Y edges is denoted by $e(X, Y)$. If $X = \{x\}$ we usually write $e(x, Y)$.

A graph is *connected* if every pair of vertices are joined by a path. A *maximal connected subgraph* is said to be a *component* of the graph. A connected graph not containing cycles is a *tree*, and a graph without cycles (an *acyclic* graph) is a *forest*. Clearly a forest is a graph whose every component is a tree. A tree of order n has n–1 edges and a forest or order n with c components has n–c edges.

The *distance* between two vertices x and y, denoted by $d(x, y)$ is the *minimum length* of an *x–y path*. If there is no x–y path, i.e, x and y belong to different components, then we put $d(x, y) = \infty$. The *diameter* of a graph G is defined as

$$\operatorname{diam} G = \max\{d(x, y):\quad x, y \in G\}.$$

A related concept is the *radius* of G, $\operatorname{rad} G = \min_x \max_y d(x, y)$. The *girth* of G, $g(G)$, is the minimum length of a cycle in G and the *circumference* of G, $c(G)$, is the maximum length of a cycle. If G does not contain a cycle then the girth and circumference are usually not defined though one might put $g(G) = c(G) = \infty$.

It might be appropriate to remark here that, following recent custom, we use the words "maximum" and "maximal" with different meanings. "Maximal" refers to a maximal element of an ordered set, in which, unless otherwise stated, the ordering is given by *inclusion.* "Maximum" refers to an element of maximal size. Thus P is a *maximal* path of a graph G if it is not properly contained in any other path and Q is a *maximum* path of G if G does not contain a path R *longer* than Q (i.e. $e(R) \leqslant e(Q)$ for every path R in G).

A graph G is an *r-partite graph with vertex classes* $V_1, V_2, \ldots, V_r$ if $V = V(G)$ is the disjoint union $V_1, V_2, \ldots, V_r$ and every edge joins vertices belonging to different vertex classes. Instead of 2-partite we say *bipartite.* We denote by $G_r(n_1, n_2, \ldots, n_r)$ an arbitrary r-partite graph whose ith class contains exactly n_i vertices.

A *k-colouring* or simply *colouring* c of a set X with colours $c_1, c_2, \ldots, c_k$ is a function $c: X \to \{c_1, \ldots, c_k\}$. We usually consider what one might call a *proper* colouring of the vertices or edges of a graph. This is a colouring in which adjacent elements (i.e. vertices in the vertex colouring and edges in the edge colouring) are assigned *different* colours. If G has a (proper) k-colouring of the vertices then G is said to be *k-colourable*. The *chromatic number* of G is $\chi(G) = \min\{k: G \text{ is } k\text{-colourable}\}$. If $\chi(G) = r$ we say that G is *r-chromatic.* It is easily seen that if G is a minimal r-chromatic graph then

$$\delta(G) \geqslant r - 1. \qquad (0.2)$$

For if $x \in G$, $d(x) \leqslant r - 2$, then a (proper) $(r - 1)$-colouring of $G - x$ can be extended to a (proper) $(r - 1)$-colouring of G. In particular,

$$\text{if } \chi(G) \geqslant r \text{ then } \delta(H) \geqslant r - 1 \text{ for some } H \subset G. \qquad (0.3)$$

Note that a (proper) r-colouring of the vertices of G is exactly a way of considering G as an r-partite graph: the ith vertex class is the set of vertices coloured with the ith colour. This is the reason why a vertex class is often referred to as a *colour class.* In spite of the equivalence of the terms r-partite and r-colourable we use both since when speaking about an r-partite graph we usually fix the vertex classes but the colour classes of an r-colourable graph are almost never supposed to be given *a priori*.

It is easily seen (e.g. [E16]) that every graph G contains a bipartite graph $B = G_2(n_1, n_2)$ whose size is at least half the size of G:

$$e(B) \geqslant \tfrac{1}{2}e(G). \tag{0.4}$$

For let B be a *maximum* bipartite subgraph of G. We may assume that B is the *bipartite subgraph of G spanned by the classes V_1 and V_2, where $V_1 \cup V_2 = V$.* If $x \in V_1$ then x is joined to at least as many vertices in V_2 as in V_1 since otherwise $V_1 - \{x\}$ and $V_2 \cup \{x\}$ could be chosen for the vertex classes, giving a bipartite subgraph of larger size. Consequently $d_B(x) \geqslant \frac{1}{2}d(x)$, implying (0.4).

The reader may find it amusing to prove that if $e(G) > 0$ then there is a subgraph $B = G_2(n_1, n_2) \subset G$ such that $n_1 + n_2 = n$, $|n_1 - n_2| \leqslant 1$ and $e(B) > \frac{1}{2}e(G)$. In particular, we may require strict inequality in (0.4). It is obvious that similar results can be proved for r-partite graphs. The weakest of these states that $e(G_r(n_1, \ldots, n_r)) \geqslant (1 - 1/r)e(G)$ for some $G_r(n_1, \ldots, n_r) \subset G$.

In a number of cases we shall find it convenient to consider a class of graphs defined as follows. Let $d \geqslant 1$ and put

$$\mathscr{D}_d = \left\{G: |G| \geqslant d, e(G) \geqslant d|G| - \binom{d+1}{2} + 1\right\}.$$

Note that if $G \in \mathscr{D}_d$ then $|G| > d$ since $|G| = d$ would imply

$$\binom{d}{2} \geqslant e(G) \geqslant d^2 - \binom{d+1}{2} + 1 = \binom{d}{2} + 1.$$

Furthermore, if $G \in \mathscr{D}_d$ then

$$G \text{ contains a subgraph } H \text{ with } \delta(H) \geqslant d + 1. \tag{0.5}$$

To see this note that if $\delta(G) \leqslant d$, say $d(x) \leqslant d$, then $G - x \in \mathscr{D}_d$ since $|G| > d$. By repeated application of this reduction we must arrive at a subgraph with minimal degree at least $d + 1$ since otherwise we would arrive at a graph $G_0 \in \mathscr{D}_d$ with $|G_0| = d$.

There are a number of structures related to graphs. A *hypergraph* or *set system H* is a set V together with a family Σ of subsets of V. Naturally $x \in V$

is a vertex and $S \in \Sigma$ is an edge of the hypergraph H. If $\Sigma \subset V^{(r)}$ then H is said to be an *r-graph* or *r-uniform hypergraph.*

By definition a graph does not contain a *loop*, i.e. an edge joining a vertex to itself, and two distinct vertices are joined by at most one edge, i.e. the graph does not contain *multiple edges.* If we allow *multiple edges* then instead of a graph we obtain a *multigraph.* The number of edges joining a vertex x to y is the *multiplicity* of the edge xy. Sometimes a multigraph is allowed to have loops (of course multiple loops) but it is more customary to call such an object a *pseudograph.* If $G = (V, E^*)$ is a multigraph, the underlying graph H of G has vertex set V and two vertices are joined in H iff they are joined in G. One might describe G as a graph H in which certain specified edges are multiple edges.

A *directed graph D* is a set $V = V(D)$ together with a collection $\vec{E} = \vec{E}(D)$ of *ordered pairs of distinct elements of V.* Of course, V is the set of vertices and $\vec{E}$ is the set of directed edges. A directed edge $(x, y) \in \vec{E}$ is denoted by $\overrightarrow{xy}$. An *oriented graph* is a directed graph containing no symmetric pair of directed edges, i.e. in which at most one of $\overrightarrow{xy}$ and $\overrightarrow{yx}$ is an edge. In other words an oriented graph G is obtained from a graph G by ordering each edge of G. Then we say that $\vec{G}$ is obtained from G by *giving G an orientation* or simply that $\vec{G}$ is an *orientation* of G. Finally we remark that the definition of an infinite graph is the obvious one: $G = (V, E)$ is an *infinite graph* if V is an infinite set, $E \subset V^{(2)}$ and $V \cap E = \emptyset$. In order to emphasize that an object in question is a graph we might call it a *simple* graph. Most of the concepts mentioned above can be carried over immediately to directed graphs. Note however that an $x_0 - x_k$ path corresponds to a *directed* $x_0 - x_k$ path, i.e. to a path $x_0 x_1 \dots x_k$ such that $x_i x_{i+1}$ is directed from x_i to x_{i+1}. Accordingly $d(x, y)$ is the minimum length of a directed $x - y$ path.

Let $G = (V, E)$, $G' = (V', E')$ be graphs. A map $\phi: V \to V'$ is said to be a *homomorphism* of G into G' if $xy \in E$ implies that $\phi(x)\,\phi(y) \in E'$. If ϕ is also 1–1 then it is an *embedding* of G into G': clearly ϕ gives an isomorphism between G and a subgraph (denoted by $\phi(G)$) of G'.

Let G be a pseudograph, i.e. multigraph in which loops are permitted. We say that a multigraph G' is an *elementary subdivision* of G if there is an edge of G joining $x \in G$ to $y \in G$ ($x = y$ may hold) such that G' is obtained from $G - xy$ by adding a new vertex and joining it to x and y. (Thus to obtain G' we replace an edge by a path of length 2.) We say that H is a *subdivision of G* or that H is a *topological G*, in notation $H = TG$, if H can be obtained from G by a sequence of elementary subdivisions. Note that the notation TG is analogous to G^n and $G(n, m)$, since it denotes an *arbitrary* subdivision of G.

In fact throughout the book we use the notation TG only in the case when G does not have a vertex of degree 2 joined to distinct vertices. Two multigraphs are said to be *homeomorphic* if they have isomorphic subdivisions. It is trivial to see that two multigraphs are homeomorphic iff the topological spaces naturally associated with them (cf. Ch V, §3) are homeomorphic.

Let H be a connected subgraph of a graph G. Add a new vertex x_H to the graph $G–H$ and join it to every vertex $y \in G–H$ for which G contains a $y–H$ edge. The resulting graph is denoted by G/H and it is said to be the graph obtained from G by *contracting* H (to a vertex). If $E(H) = xy$ then $G/xy = G/H$ is an *elementary contraction* of G. We say that L is a *contraction* of G, in notation $G > L$, if L can be obtained from G by a sequence of contractions (of connected subgraphs). L is a *subcontraction* of G, in notation $G \succ L$ if L is a contraction of a subgraph of G.

The *complement* of a graph $G = (V, E)$ is the graph $\bar{G} = (V, V^{(2)} - E)$. The *complete graph of order n*, K^n, has every pair of its n vertices adjacent. In other words K^n is the graph of order n and size $\binom{n}{2}$, that is $K^n = G(n, \binom{n}{2})$. Note that $K^3 = C^3$ is the triangle. We call K^4 a complete quadrilateral, etc. We have chosen the notation K^n instead of the more common K_n since we shall use capital letters with subscripts (G_k, H_l, K_p, etc.) to denote *specific* graphs. Thus K_p might denote a *given* complete subgraph. The complement of K^n is the *empty* or *null* graph of order n: $E^n = \bar{K}^n = G(n, 0)$. The unique maximal graph $G_r(n_1, n_2, \ldots, n_r)$ is denoted by $K_r(n_1, n_2, \ldots, n_r)$. It has r vertex classes, the ith class has n_i vertices and every pair of vertices belong to distinct classes are joined by an edge. Clearly

$$e(K_r(n_1, \ldots, n_r)) = \sum_{1 \leq i < j \leq r} n_i n_j.$$

If $r = 2$ then the index r might be omitted and n_1, n_2 might become upper indices, e.g. $K_2(3, 4) = K(3, 4) = K^{3,4}$. The tree $K^{1,p}$ is the *star* of order $p + 1$. The maximum order of a complete subgraph of G is the *clique number* of G. We denote it by $cl(G)$.

The *union* of the graphs G_1 and G_2 is denoted by $G_1 \cup G_2$. If

$$(V(G_1) \cup E(G_1)) \cap (V(G_2) \cup E(G_2)) = \emptyset \qquad (0.6)$$

then

$$V(G_1 \cup G_2) = V(G_1) \cup V(G_2)$$

and

$$E(G_1 \cup G_2) = E(G_1) \cup E(G_2).$$

If (6) does not hold then we add an index to the elements of $V(G_i) \cup E(G_i)$ to make (6) hold and define the union as before. Sometimes we emphasize this by saying that $G_1 \cup G_2$ is the *disjoint* union of G_1 and G_2; e.g. $K^3 \cup K^4$ is the complement of $K(3, 4)$. The only exception occurs when G_1 and G_2 are subgraphs of a given graph. Then, naturally, $G_1 \cup G_2$ is defined by

$$V(G_2 \cup G_2) = V(G_1) \cup V(G_2) \quad \text{and} \quad E(G_1 \cup G_2) = E(G_1) \cup E(G_2).$$

It will be clear from the text which of these cases is at hand.

The union of several graphs is defined analogously. The disjoint union of k copies of the same graph G is denoted by kG. Thus $kK^1 = kE^1 = E^k$.

The *join* $G_1 + G_2$ of G_1 and G_2 is obtained from $G_1 \cup G_2$ by joining *each vertex of* G_1 *to each vertex of* G_2. Thus $E^3 + E^4 = K(3, 4)$. The join of several graphs is defined analogously:

$$E^{n_1} + E^{n_2} + \ldots + E^{n_r} = K_r(n_1, n_2, \ldots, n_r).$$

In a number of graph constructions it is convenient to choose a prime for one of the parameters. In order to extend the construction to every possible value of the parameters one then uses a shallow or deep result about the distribution of primes. *Bertrand's postulate* claims that for every natural number $n > 3$ there is a prime between n and $2n - 2$. This was verified by Bertrand for $n < 3\,000\,000$ and proved by Tchebychev in 1850 (cf. [HW1; p. 373]). Furthermore, the quotient of consecutive primes tends to 1. In fact there are $0 < \eta < 1$ and $C_\eta > 0$ such that for every $n \geqslant 2$ there is a prime between n and $n + C_\eta n^\eta$. This was proved by Hoheisel [H27] ($\eta = 1 - 3300^{-1} + \varepsilon$), Ingham [I1] ($\eta = \frac{5}{8} + \varepsilon$), Montgomery [M32] ($\eta = \frac{3}{5} + \varepsilon$) and Huxley [H27] ($\eta = \frac{7}{12} + \varepsilon$) where $\varepsilon > 0$ is arbitrary. We shall make use of the fact that if n is sufficiently large then

$$\text{there is a prime between } n - \tfrac{1}{10}n^{2/3} \text{ and } n. \qquad (0.7)$$

If q is a prime power then there is a finite projective plane $PG(2, q)$ over the field of order q. We represent the points and lines of this plane by triples (a, b, c) and $[a, b, c]$ of elements of the ground field such that each triple has at least one non-zero element. If $\lambda \neq 0$ then (a, b, c) and $(\lambda a, \lambda b, \lambda c)$ represent the same point; similarly $[a, b, c]$ and $[\lambda a, \lambda b, \lambda c]$ represent the same line. A point (x, y, z) is on a line $[a, b, c]$ if $ax + by + cz = 0$.

Contents