Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARTIN GRÖTSCHEL

My Favorite Theorem: Characterizations of Perfect Graphs

Preprint SC 99-17 (June 1999)

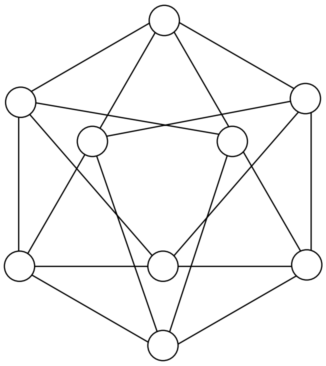
My Favorite Theorem: Characterizations of Perfect Graphs

Martin Grötschel

The favorite topics and results of a researcher change over time, of course. One area that I have always kept an eye on is that of perfect graphs. These graphs, introduced in the late 50s and early 60s by Claude Berge, link various mathematical disciplines in a truly unexpected way: graph theory, combinatorial optimization, semidefinite programming, polyhedral and convexity theory, and even information theory.

This is not a survey of perfect graphs. It's just an appetizer. To learn about the origins of perfect graphs, I recommend to read the historical papers [1] and [2]. The book [3] is a collection of important articles on perfect graphs. Algorithmic aspects of perfect graphs are treated in [13]. A comprehensive survey of graph classes, including perfect graphs, can be found in [5]. Hundreds of classes of perfect graphs are known, 96 important classes and the inclusion relations among them are described in [16].

So, what is a perfect graph? Complete graphs are perfect, bipartite, interval, comparability, triangulated, parity, and unimodular graphs are perfect as well. The following beautiful perfect graph is the line graph of the complete bipartite graph $K_{3,3}$.



Due to the evolution of the theory, definitions of perfection (and versions thereof) have changed over time. To keep this article short, I do not follow the historical development of the notation. I use definitions that streamline the presentation. Berge defined

G is a *perfect* graph,

if and only if

(1) $\omega(G') = \chi\,(G')$ for all node-induced subgraphs $G' \subseteq G$,

where $\omega(G)$ denotes the *clique number* of G (= largest cardinality of a *clique* of G, i.e., a set of mutually adjacent nodes) and $\chi(G)$ is the *chromatic number* of G (= smallest number of colors needed to color the nodes of G). Berge discovered that all classes of perfect graphs he found also have the property that

(2) $\alpha(G') = \overline{\chi}(G')$ for all node-induced subgraphs $G' \subseteq G$,

where $\alpha(G)$ is the *stability number* of G (= largest cardinality of a *stable set* of G, i.e., a set of mutually nonadjacent nodes) and $\overline{\chi}(G)$ denotes the *clique covering number* of G (= smallest number of cliques needed to cover all nodes of G exactly once).

Note that complementation (two nodes are adjacent in the *complement* $\overline{G}$ of a graph G iff they are nonadjacent in G) transforms a clique into a stable set and a coloring into a clique covering, and vice versa. Hence, the complement of a perfect graph satisfies (2). This observation and his discovery mentioned above led Berge to conjecture that G is a perfect graph if and only if

(3) $\overline{G}$ is a perfect graph.

Developing the antiblocking theory of polyhedra, Fulkerson launched a massive attack on this conjecture, see [10], [11], and [12]. The conjecture was solved in 1972 by Lovász [17], who gave two short and elegant proofs. Lovász [18], in addition, characterized perfect graphs as those graphs $G = (V, E)$ for which the following holds:

(4) $\omega(G') \cdot \alpha(G') \;\geq\; |V(G')|$ for all node-induced subgraphs $G' \subseteq G$.

A link to geometry can be established as follows. Given a graph $G = (V, E)$, we associate with G the vector space $\mathbb{R}^V$ where each component of a vector of $\mathbb{R}^V$ is indexed by a node of G. With every subset $S \subseteq V$, we can associate its incidence vector $\chi^S = (\chi^S_v)_{v \in V} \in \mathbb{R}^V$ defined by

$$\chi^S_v := 1 \text{ if } v \in S, \;\; \chi^S_v := 0 \text{ if } v \notin S.$$

The convex hull of all the incidence vectors of stable sets in G is denoted by $\mathrm{STAB}(G)$, i.e.,

$$\mathrm{STAB}(G) = \text{ conv } \{\chi^S \in \mathbb{R}^V \mid S \subseteq V \;\; \text{stable}\}$$

and is called the *stable set polytope* of G. Clearly, a clique and a stable set of G can have at most one node in common. This observation yields that, for every clique $Q \subseteq V$, the so-called *clique inequality*

$$x(Q) := \sum_{v \in Q} x_v \leq 1$$

is satisfied by every incidence vector of a stable set. Thus, all clique inequalities are valid for STAB(G). The polytope

$$\text{QSTAB}(G) := \{x \in \mathbb{R}^V \mid 0 \leq x_v \; \forall \; v \in V, \;\; x(Q) \leq 1 \; \forall \; \text{ cliques } Q \subseteq V\},$$

called *fractional stable set polytope* of G, is therefore a polyhedron containing STAB(G), and trivially,

$$\text{STAB}(G) = \text{conv}\; \{x \in \{0,1\}^V \mid x \in \text{QSTAB}(G)\}.$$

Knowing that computing $\alpha(G)$ (and its weighted version) is $\mathcal{NP}$-hard one is tempted to look at the LP relaxation

$$\max c^T x, \;\; x \in \; \text{QSTAB}(G),$$

where $c \in \mathbb{R}^V_+$ is a vector of node weights. However, solving LPs of this type is also $\mathcal{NP}$-hard for general graphs G, see [14]. For the class of perfect graphs G, though, these LPs can be solved in polynomial time — albeit via an involved detour, see below.

Let us now look at the following chain of inequalities and equations, typical for IP/LP approches to combinatorial problems. Let $G = (V, E)$ be some graph and $c \geq 0$ a vector of node weights:

$$\begin{aligned}
& \max \;\; \{\sum_{v \in S} c_v \mid S \subseteq V \text{ stable set of } G\} \\
= \; & \max \;\; \{c^T x \mid x \in \text{STAB}(G)\} \\
= \; & \max \;\; \{c^T x \mid x \geq 0, \; x(Q) \leq 1 \; \forall \; \text{ cliques } Q \subseteq V, x \in \{0,1\}^V\} \\
\leq \; & \max \;\; \{c^T x \mid x \geq 0, \; x(Q) \leq 1 \; \forall \; \text{ cliques } Q \subseteq V\} \\
= \; & \min \;\; \{\sum_{Q \text{ clique}} y_Q \mid \sum_{Q \ni v} y_Q \geq c_v \; \forall \; v \in V, y_Q \geq 0 \; \forall \; \text{ cliques } Q \subseteq V\} \\
\leq \; & \min \;\; \{\sum_{Q \text{ clique}} y_Q \mid \sum_{Q \ni v} y_Q \geq c_v \; \forall \; v \in V, y_Q \in \mathbb{Z}_+ \; \forall \text{ cliques } Q \subseteq V\}
\end{aligned}$$

The inequalities come from dropping or adding integrality constraints, the last equation is implied by LP duality. The last program can be interpreted as an IP formulation of the weighted clique covering problem. It follows from (2) that equality holds throughout the whole chain for all 0/1 - vectors c iff G is a perfect graph. This, in turn, is equivalent to

(5) The value $\max \{c^T x \mid x \in \text{ QSTAB}(G)\}$ is integral for all $c \in \{0,1\}^V$.

Results of Fulkerson [10] and Lovász [17] imply that (5) is in fact equivalent to

(6) The value $\max \{c^T x \mid x \in \text{ QSTAB}(G)\}$ is integral for all $c \in \mathbb{Z}_+^V$.

and that, for perfect graphs, equality holds throughout the above chain for all $c \in \mathbb{Z}_+^V$. This, as a side remark, proves that the constraint system defining QSTAB(G) in totally dual integral for perfect graphs G. Chvátal [6] observed that (6) holds iff

(7) $$\text{STAB}(G) = \text{QSTAB}(G)$$

These three characterizations of perfect graphs provide the link to polyhedral theory (a graph is perfect iff certain polyhedra are identical) and integer programming (a graph is perfect iff certain LPs have integral solution values).

Another, quite surprising road towards understanding properties of perfect graphs was paved by Lovász [19]. He introduced a new geometric representation of graphs linking perfectness to convexity and semidefinite programming.

An *orthonormal representation* of a graph $G = (V, E)$ is a sequence $(u_i \mid i \in V)$ of $|V|$ vectors $u_i \in \mathbb{R}^V$ such that $||u_i|| = 1$ for all $i \in V$ and $u_i^T u_j = 0$ for all pairs i, j of nonadjacent nodes. For any orthonormal representation $(u_i \mid i \in V)$ of G and any additional vector c of unit length, the so-called *orthonormal representation constraint*

$$\sum_{i \in V} (c^T u_i)^2 x_i \leq 1$$

is valid for $\text{STAB}(G)$. Taking an orthonormal basis $B = \{e_1, ..., e_{|V|}\}$ of $\mathbb{R}^V$ and a clique Q of G, setting c:= $u_i := e_1$ for all $i \in Q$, and assigning different vectors of $B \backslash \{e_1\}$ to the remaining nodes $i \in V \backslash Q$, one observes that every clique constraint is a special case of this class of infinitely many inequalities. The set

$$\text{TH}(G) := \{x \in \mathbb{R}_+^V \mid x \text{ satisfies all orthonormal representation constraints}\}$$

is thus a convex set with

$$\text{STAB}(G) \subseteq \text{TH}(G) \subseteq \text{ QSTAB}(G).$$

It turns out, see [14], that a graph G is perfect if and only if any of the following conditions is satisfied:

$$\text{TH}(G) = \text{STAB}(G). \tag{8}$$

$$\text{TH}(G) = \text{QSTAB}(G). \tag{9}$$

$$\text{TH}(G) \text{ is a polytope.} \tag{10}$$

The last result is particularly remarkable. It states that a graph is perfect if and only if a certain convex set is a polytope.

If $c \in \mathbb{R}^V_+$ is a vector of node weights, the optimization problem (with infinitely many linear constraints)

$$\max c^T x, x \in \text{TH}(G)$$

can be solved in polynomial time for any graph G. This implies, by (8), that the weighted stable set problem for perfect graphs can be solved in polynomial time, and by LP duality, that the weighted clique covering problem, and by complementation, that the weighted clique and coloring problem can be solved in polynomial time. These results rest on the fact that the value

$$\vartheta(G, c) := \max \{c^T x \mid x \in \text{TH}(G)\}$$

can be characterized in many equivalent ways, e.g., as the optimum value of a semidefinite program, the largest eigenvalue of a certain set of symmetric matrices, or the maximum value of some function involving orthornormal representations.

Details of this theory can be found, e.g., in Chapter 9 of [14]. The algorithmic results involve the ellipsoid method. It would be nice to have "more combinatorial" algorithms that solve the four optimization problems for perfect graphs in polynomial time.

Let us now move into information theory. Given a graph $G = (V, E)$, we call a vector $p \in \mathbb{R}^V_+$ a *probability distribution* on V if its components sum to 1. Let $G^{(n)} = (V^n, E^{(n)})$ denote the so-called *n-th conormal power* of G, i.e., V^n is the set of all n-vectors $x = (x_1, \ldots, x_n)$ with components $x_i \in V$, and

$$E^{(n)} := \{xy \mid x, y \in V^n \text{ and } \exists\, i \text{ with } x_i y_i \in E\}.$$

Each probability distribution p on V induces a probability distribution p^n on V^n as follows: $p^n(x) := p(x_i) \cdot p(x_2) \cdot \ldots \cdot p(x_n)$. For any node set $U \subseteq V^n$, let

$G^{(n)}[U]$ denote the subgraph of $G^{(n)}$ induced by U and $\chi(G^{(n)}[U])$ its chromatic number. Then one can show that, for every $0 < \epsilon < 1$, the limit

$$H(G,p) := \lim_{n\to\infty} \frac{1}{n} \min_{p^n(U)\geq 1-\epsilon} \log \chi(G^{(n)}\,[U])$$

exists and is independent of ϵ (the logs are taken to base 2). $H(G,p)$ is called the *graph entropy* of the graph G with respect to the probability distribution p. If $G = (V,E)$ is the complete graph, we get the well known *Shannon entropy*

$$H(p) = -\sum_{i\in V} p_i \log p_i.$$

Let us call a graph $G = (V,E)$ *strongly splitting* if for every probability distribution p on V

$$H(p) = H(G,p) + H(\overline{G},p)$$

holds. Csiszár et al [9]. have shown that a graph is perfect if and only if

(11) G is strongly splitting.

I.e., G is perfect iff, for every probabiltity distribution, the entropies of G and of its complement $\overline{G}$ add up to the entropy of the complete graph (the Shannon entropy). I recommend [9] for the study of graph entropy and related topics.

Given all these beautiful characterizations of perfect graphs and polynomial time algorithms for many otherwise hard combinatorial optimization problems, it is really astonishing that nobody knows to date whether perfectness of a graph can be recognized in polynomial time. There are many ways to prove that, deciding whether a graph is not perfect, is in $\mathcal{NP}$. But that's all we know!

Many researchers hope that a proof of the most famous open problem in perfect graph theory, the *strong perfect graph conjecture:*

> *A graph G is perfect if and only if G neither contains an odd hole nor an odd antihole as an induced subgraph.*

results in structural insights that lead to a polynomial time algorithm for recognizing perfect graphs. It is trivial that every odd hole (= chordless cycle of length at least five) and every odd antihole (= complement of an odd hole) are not perfect. Whenever Claude Berge encountered an imperfect graph G he discovered that G contains an odd hole or an odd antihole and, thus, came to the strong perfect graph conjecture. In his honor, it is customary to call graphs without odd holes and odd antiholes *Berge graphs*. Hence, the strong perfect graph conjecture essentially reads: every Berge graph is perfect.

This conjecture stimulated a lot of research resulting in fascinating insights into the structure of graphs that are in some sense nearly perfect or imperfect. E.g., Padberg [20], [21] (introducing perfect matrices and using proof techniques from linear algebra) showed that, for an imperfect graph $G = (V, E)$ with the property that the deletion of any node results in a perfect graph, satifies the following:

- $|V| = \alpha(G) \cdot \omega(G) + 1$,
- G has exactly $|V|$ maximum cliques, and every node is contained in exactly $\omega(G)$ such cliques.
- G has exactly $|V|$ maximum stable sets, and every node is contained in exactly $\alpha(G)$ such stable sets.
- QSTAB(G) has exactly one fractional vertex, namely the point $x_v = 1/\omega(G)$ $\forall\ v \in V$, which is contained in exactly $|V|$ facets and adjacent to exactly $|V|$ vertices, the incidence vectors of the maximum stable sets.

Similar investigations (but not resulting in such strong structural results) have recently been made by Annegret Wagler [24] on graphs which are perfect and have the property that deletion (or addition) of any edge results in an imperfect graph. The graph of Figure 1 is from Wagler's Ph.D. thesis. It is the smallest perfect graph G such that whenever any edge is added to G or any edge is deleted from G the resulting graph is imperfect.

Particular efforts have been made to characterize perfect graphs "constructively" in the following sense. One first establishes that a certain class $\mathcal{C}_\infty$ of graphs is perfect and considers, in addition, a finite list $\mathcal{C}_\in$ of special perfect graphs. Then one defines a set of "operations" (e.g., replacing a node by a stable set or a perfect graph) and "compositions" (e.g., take two graphs G and H and two nodes $u \in V(G)$ and $v \in V(H)$, define $V(G \circ H) := (V(G) \cup V(H)) \backslash \{u, v\}$ and $E(G \circ H) := E(G-u) \cup E(G-v) \cup \{x, y \mid xu \in E(G), yv \in E(H)\}$ and shows that every perfect graph can be constructed from the basic classes $\mathcal{C}_\infty$ and $\mathcal{C}_\in$ by a sequence of operations and compositions. Despite ingenious constructions (that were very helpful in proving some of the results mentioned above) and lots of efforts, this route of research has not led to success, yet. A paper describing many compositions that construct perfect graphs from perfect graphs is, e.g., [8].

Chvátal [7] initiated research into another "secondary structure" related to perfect graphs in order to come up with a (polynomial time recognizable) certificate of perfection. For a given graph $G = (V, E)$, its P_4-*structure* is the 4-uniform hypergraph on V whose hyperedges are all the 4-element node sets of V that induce a P_4 (path on four nodes) of G. Chvátal observed that any graph whose P_4-structure is that of an odd hole is an odd hole or its complement and, thus, conjectured that perfection of a graph depends solely on its P_4-structure. Reed [23] solved Chvátal's *semi-strong perfect graph conjecture* by showing that a graph G is perfect iff

(12) $\qquad$ G has the P_4-structure of a perfect graph.

There are other such concepts, e.g., the *partner-structure*, that have resulted in further characterizations of perfect graphs through secondary structures. We recommend [15] for a thorough investigation of this topic. But the polynomial-time-recognition problem for perfect graphs is still open.

A relatively recent line of research in the area of structural perfect graph theory is the use of the probability theory. I would like to mention just one nice result of Prömel und Steger [22]. Let us denote the number of perfect graphs on n nodes by *Perfect* (n) and the number of Berge graphs on n nodes by *Berge* (n), then

$$\lim_{n\to\infty} \frac{\text{Perfect } (n)}{\text{Berge } (n)} = 1.$$

In other words, almost all Berge graphs are perfect which means that, if there are counterexamples to the strong perfect graph conjecture, they are "very rare".

The theory of random graphs provides deep insights into the probabilistic behavior of graph parameters, see [4] for instance. To take a simple example, consider a random graph $G = (V, E)$ on n nodes where each edge is chosen with probality $\frac{1}{2}$. It is well known that the expected values of $\alpha(G)$ and $\omega(G)$ are of order $\log n$ while $\chi(G)$ and $\overline{\chi}(G)$ both have expected values of order $n/\log n$. This implies that such random graphs are almost surely not perfect. An interesting question is to see whether the "LP-relaxation of $\alpha(G)$", the so-called fractional stability number $\alpha^*(G) := \max\{\mathbb{1}^T x \mid x \in \text{ QSTAB}(G)\}$, is a good approximation of $\alpha(G)$. Observing that the point $x = (x_v)_{v\in V}$ with $x_v := 1/\omega(G), v \in V$, satisfies all clique constraints and is thus in QSTAB(G) and knowing that $\omega(G)$ is of order $\log n$ one can deduce that the expected value of $\alpha^*(G)$ is of order $n/\log n$, i.e., it is much closer to $\overline{\chi}(G)$ than to $\alpha(G)$. Hence, somewhat surprisingly, $\alpha^*(G)$ is a pretty bad approximation of $\alpha(G)$ in general – not so for perfect graphs, though.

To summarize this quick tour through perfect graph theory (omitting quite a number of the other interesting developments and important results), here is my favorite theorem:

Theorem *Let G be a graph. The following twelve conditions are equivalent and characterize G as a perfect graph:*

(1) $\omega(G') = \chi\,(G')$ *for all node-induced subgraphs* $G' \subseteq G$.

(2) $\alpha(G') = \overline{\chi}(G')$ *for all node-induced subgraphs* $G' \subseteq G$.

(3) $\overline{G}$ *is a perfect graph.*

(4) $\omega(G') \cdot \alpha(G') \;\geq\; |V(G')|$ *for all node-induced subgraphs* $G' \subseteq G$.

(5) *The value* $\max\,\{c^T x \mid x \in \textit{QSTAB}(G)\}$ *is integral for all* $c \in \{0,1\}^V$.

(6) *The value* $\max\,\{c^T x \mid x \in \textit{QSTAB}(G)\}$ *is integral for all* $c \in \mathbb{Z}_+^V$.

(7) $\textit{STAB}(G) = \textit{QSTAB}(G)$.

(8) $\textit{TH}\,(G) = \textit{STAB}(G)$.

(9) $\textit{TH}(G) = \textit{QSTAB}(G)$.

(10) $\textit{TH}(G)$ *is a polytope.*

(11) G *is strongly splitting.*

(12) G *has the P_4-structure of a perfect graph.*

References

[1] C. Berge, The history of the perfect graphs, *Southeast Asian Bull. Math. 20*, No.1 (1996) 5-10.

[2] C. Berge, Motivations and history of some of my conjectures, *Discrete Mathematics* 165-166 (1997) 61-70.

[3] C. Berge, V. Chvátal (eds.), *Topics on perfect graphs,* Annals of Discrete Mathematics 21, North-Holland, Amsterdam, 1984.

[4] B. Bollobás, *Random Graphs*, Academic Press, 1985.

[5] A. Brandstädt, V. B. Le, J. P. Spinrad, *Graph Classes: A Survey*, SIAM, 1999.

[6] V. Chvátal, On certain polytopes associated with graphs, *Journal of Combinatorial Theory B 18* (1975) 138-154.

[7] V. Chvátal, A semi-strong perfect graph conjecture, *Annals of Discrete Mathematics 21* (1984) 279-280.

[8] G. Cornuéjols, W. H. Cunningham, Compositions for perfect graphs, *Discrete Mathematics 55* (1985) 245-254.

[9] I. Csiszár, J. Körner, L. Lovász, K. Marton, G. Simony, Entropy splitting for anti-blocking corners and perfect graphs, *Combinatorica 10 (1)* (1990) 27-40.

[10] D. R. Fulkerson, The perfect graph conjecture and pluperfect graph theorem *Proc. 2nd Chapel Hill Conf. Combin. Math. Appl.*, Univ. North Carolina (1970), 171-175.

[11] D. R. Fulkerson, Blocking and anti-blocking pairs of polyhedra, *Mathematical Programming 1* (1971) 168-194.

[12] D. R. Fulkerson, On the perfect graph theorem, *Math. Programming, Proc. Advanced Seminar*, Univ. Wisconsin, Madison (1972) 69-76.

[13] M. C. Golumbic, *Algorithmic graph theory and perfect graphs,* Academic Press, New York, 1980.

[14] M. Grötschel, L. Lovász, A. Schrijver, *Geometric algorithms and combinatorial optimization*, Springer, Berlin-Heidelberg, 1988.

[15] S. Hougardy, *On the P_4-structure of perfect graphs,* Shaker-Verlag, Aachen 1996.

[16] S. Hougardy, Inclusions Between Classes of Perfect Graphs, Preprint, Institut für Informatik, Humboldt-Universität zu Berlin, 1998, see URL: http://www.informatik.hu-berlin.de/ hougardy/paper/classes.html

[17] L. Lovász, Normal hypergraphs and the perfect graph conjecture, *Discrete Mathematics 2* (1972) 253-267.

[18] L. Lovász, A characterization of perfect graphs, *Journal of Combinatorial Theory (B) 13* (1972) 95-98.

[19] L. Lovász, On the Shannon capacity of a graph, *IEEE Transf. Inform. Theory 25* (1979) 1-7.

[20] M. W. Padberg, Perfect zero-one matrices, *Math. Programming 6* (1974) 180-196.

[21] M. W. Padberg, Almost integral polyhedra related to certain combinatorial optimization problems, *Lin. Alg. Appl. 15* (1976) 69-88.

[22] H. J. Prömel, A. Steger, Almost all Berge graphs are perfect, *Combin. Probab. Comput. 1* (1992) 53-79.

[23] B. Reed, *A semi-strong perfect graph theorem*, Journal of Comb. Theory, Ser. B 43 (1987) 223-240.

[24] A. Wagler, *Critical edges in perfect graphs,* Ph.D. thesis, TU Berlin, 1999.