Discrete Mathematics 340 (2017) 766–793

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Perspective

An introduction to the discharging method via graph coloring

Daniel W. Cranston [a], Douglas B. West [b,c,*]

[a] Virginia Commonwealth University, United States
[b] Zhejiang Normal University, China
[c] University of Illinois, United States

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form 15 November 2016
Accepted 15 November 2016
Available online 16 December 2016

Keywords:
Discharging method
Graph coloring
Planar graph
Edge-coloring
Maximum average degree
Light subgraph

ABSTRACT

We provide a "how-to" guide to the use and application of the Discharging Method. Our aim is not to exhaustively survey results proved by this technique, but rather to demystify the technique and facilitate its wider use, using applications in graph coloring as examples. Along the way, we present some new proofs and new problems.

1. Introduction

Arguments that can be phrased in the language of the Discharging Method have been used in graph theory for more than 100 years, though that name is much more recent. The most famous application of the method is the proof of the Four Color Theorem, stating that graphs embeddable in the plane have chromatic number at most 4. However, the method remains mysterious to many. Our aim is to explain its use and make the method more widely accessible. Although we mention many applications, including stronger versions of results proved here, cataloguing applications is not our goal. Borodin [22] presents a survey of applications of discharging to coloring of plane graphs.

Discharging is most commonly used as a tool in a two-pronged approach to inductive proofs, typically for sparse graphs. In this context, it is used to prove that a global sparseness hypothesis guarantees the existence of some desired local structure. The method has been applied to many types of problems (including graph embeddings and decompositions, spread of infections in networks, geometric problems, etc.). Nevertheless, we present only applications in graph coloring (where it has been used most often), in order to emphasize the discharging techniques.

In the simplest version, discharging just involves reallocation of vertex degrees in the context of a global bound on the average degree. We view each vertex as having an initial "charge" equal to its degree. To show that average degree less than b forces the presence of a desired local structure, we show that the absence of such a structure allows charge to be moved (via "discharging rules") so that the final charge at each vertex is at least b. This violates the hypothesis, and hence the desired structure must occur.

In an application of the resulting structure theorem, one shows that each such local structure is "reducible", meaning that it cannot occur in a minimal counterexample to the desired conclusion. This motivates the phrase "an unavoidable set of reducible configurations" to describe the overall process.

Definition 1.1. A *configuration* in a graph G can be any structure in G (often a specified sort of subgraph). A configuration is *reducible* for a graph property Q if it cannot occur in a minimal graph not having property Q. Let $d_G(v)$ or simply $d(v)$ denote the degree (number of neighbors) of vertex v in G, and let $\bar{d}(G)$ denote the average of the vertex degrees in G. *Degree charging* is the assignment to each vertex v of an "initial charge" equal to $d(v)$.

The notion of configuration is vague to permit use in various contexts. "Minimal" refers to some partial order on the graphs being considered; usually it is just minimality with respect to taking subgraphs, and the property Q is *monotone* (preserved by taking subgraphs).

Sparse local configurations aid in inductive proofs about coloring. For example, when $\bar{d}(G) < k$ with $k \in \mathbb{N}$, the pigeonhole principle guarantees a vertex with degree less than k in G. Also, when $d(v) < k$, a proper k-coloring of $G - v$ extends to a proper k-coloring of G. (A *k-coloring* is a function that assigns labels to vertices from a set of size k, a coloring of a graph G is *proper* if adjacent vertices have distinct colors, G is *k-colorable* if it admits a proper k-coloring, and the *chromatic number* $\chi(G)$ is the least k such that G is k-colorable.)

In other words, vertices of degree less than k are reducible for the property $\chi(G) \le k$. However, guaranteeing such a vertex from the global bound $\bar{d}(G) < k$ does not need discharging. To illustrate how discharging works and interacts with reducibility, we consider another elementary example after introducing notation convenient for discussing vertex degrees.

Definition 1.2. A *j-vertex*, *j^+-vertex*, or *j^--vertex* is a vertex with degree equal to j, at least j, or at most j, respectively. A *j-neighbor* of v is a j-vertex that is a neighbor of v. We write $\delta(G)$ for the minimum and $\Delta(G)$ for the maximum of the vertex degrees in G.

Lemma 1.3. *If $\overline{d}(G) < 3$, then G has a 1^--vertex or a 2-vertex with a 5^--neighbor.*

Proof. We use degree charging; each vertex v starts with charge $d(v)$. Suppose that G has no 1^--vertex and that no 2-vertex in G has a 5^--neighbor. We move charge so that each vertex ends with charge at least 3. The 2-vertices need charge; 4^+-vertices can give charge.

Let each 2-vertex take $\frac{1}{2}$ from each neighbor. Now each 2-vertex has charge 3, since no two 2-vertices are adjacent. Vertices of degrees 3, 4, 5 lose no charge, since we assumed that no 2-vertex has a 5^--neighbor. Every 6^+-vertex v loses charge at most $\frac{1}{2}$ to each neighbor, leaving it with charge at least $d(v)/2$, which is at least 3 when $d(v) \geq 6$. Thus $\overline{d}(G) \geq 3$ when no 2-vertex has a 5^--neighbor. □

A 2-vertex with a 5^--neighbor is a local sparseness condition, somehow more sparse than a 2-vertex with high-degree neighbors. We first consider its use for edge-coloring. (A *k-edge-coloring* of a graph G assigns labels to edges from a set of size k; it is *proper* if incident edges have distinct colors, G is *k-edge-colorable* if it has a proper k-edge-coloring, and the *edge-chromatic number* $\chi'(G)$ is the least k such that G is k-edge-colorable.)

Here we phrase the reducibility statement in more generality. The *weight* of a subgraph H of a graph G is $\sum_{v \in V(H)} d_G(v)$; we sum the degrees in the full graph G.

Lemma 1.4. *An edge with weight at most $k+1$ is a reducible configuration for the property of being k-edge-colorable.*

Proof. Let G be a graph having an edge e of weight at most $k+1$. If the graph $G - e$ is k-edge-colorable, then a color is available to extend the coloring to e, because e is incident to a total of at most $k-1$ other edges at its two endpoints. Thus a minimal graph G with $\chi'(G) > k$ cannot contain such a configuration. □

To complete an inductive proof of $\chi'(G) \leq 6$ from Lemmas 1.3 and 1.4, we also need average degree less than 3 in subgraphs of G.

Definition 1.5. The *maximum average degree* of a graph G, denoted $\mathrm{mad}(G)$, is the maximum of the average degree over all subgraphs of G.

The application is now easy. Note that always $\chi'(G) \geq \Delta(G)$. In fact, Vizing's Theorem [73,117] states that always $\chi'(G) \leq \Delta(G) + 1$, and distinguishing between $\chi'(G) = \Delta(G)$ and $\chi'(G) = \Delta(G) + 1$ is an important and difficult problem.

Theorem 1.6. *If* $\mathrm{mad}(G) < 3$ *and* $\Delta(G) \geq 6$, *then* $\chi'(G) = \Delta(G)$.

Proof. Fix an integer k at least 6. We prove more generally that if $\mathrm{mad}(G) < 3$ and $\Delta(G) \leq k$, then $\chi'(G) \leq k$. That is, among graphs with $\mathrm{mad}(G) < 3$ and $\Delta(G) \leq k$ there is no minimal graph satisfying $\chi' > k$. Note that the hypotheses also hold in subgraphs.

We may discard isolated vertices. By Lemma 1.3, G then has a 1-vertex or has a 2-vertex with a 5^--neighbor. The edge incident to a 1-vertex has weight at most $\Delta(G) + 1$; an edge joining a 2-vertex to a 5^--neighbor has weight at most 7. In either case, the weight of this edge e is at most $k + 1$, and Lemma 1.4 implies that G is not a minimal graph satisfying $\chi'(G) > k$. Hence there is no minimal counterexample. □

Before leaving Theorem 1.6, we note that many reducibility arguments for coloring problems involve deleting some parts of a graph (such as a 1-vertex or the edge e in the proof above) and then choosing colors for the missing pieces as they are replaced. Suitable choices can be made if there are enough available colors; it does not matter what the colors are. In this situation, the arguments yield stronger statements about coloring from lists.

Definition 1.7. A *list assignment* L on a graph G gives each $v \in V(G)$ a set $L(v)$ of colors, called its *list*. In a *k-uniform* list assignment, each list has size k. Given a list assignment L, an *L-coloring* of G is a proper coloring ϕ of G such that $\phi(v) \in L(v)$ for all $v \in V(G)$. A graph G is *k-choosable* if G is L-colorable whenever each list has size at least k (we may assume L is k-uniform). The *list chromatic number* or *choice number* of G, written $\chi_\ell(G)$, is the least k such that G is k-choosable. Analogous language is used for edge-colorings chosen from list assignments to edges.

Since the lists can be identical, always $\chi_\ell(G) \geq \chi(G)$. Thus $\chi_\ell(G) \leq b$ is stronger than $\chi(G) \leq b$. For example, $\mathrm{mad}(G) < k$ inductively yields $\chi_\ell(G) \leq k$. Similarly, an edge of weight at most $k+1$ is reducible for k-edge-choosability, and the proof of Theorem 1.6 yields $\chi_\ell(G) = \Delta(G)$ when $\mathrm{mad}(G) < 3$ and $\Delta(G) \geq 6$.

We present various classical applications, some with new proofs. We emphasize discharging arguments but include reducibility arguments to show how discharging is applied. For clarity and simplicity in illustrating the method, we often assume more restrictive hypotheses than used in the strongest known results. Often those results are proved similarly, but with more detail in the discharging arguments and more configurations to be proved reducible.

The basic idea of discharging proofs is simple, and the proofs are usually easy to follow, though they may have many details. The mystery arises in the choice of reducible configurations, the rules for moving charge, and how to find the best hypothesis. We will explain the interplay among these and suggest how the proofs are discovered, starting with the context of $\mathrm{mad}(G) < b$ in Section 2. We include related results as exercises to aid in self-study; most exercises have relatively short solutions (items labeled "Question" are unsolved).

As we have illustrated, structural results proved by discharging when $d(G) < b$ are applied inductively to obtain coloring conclusions under the hypothesis $\mathrm{mad}(G) < b$. The point is that every subgraph H satisfies $d(H) < b$. For natural hereditary families like planar graphs, bounds on $\mathrm{mad}(G)$ are easily obtained. The families satisfying $\mathrm{mad}(G) < b$ for various positive b provide a rich spectrum for study.

Discharging has been used to prove many results on coloring or structure of planar graphs (or planar graphs with large girth). Euler's Formula implies that (every subgraph of) a planar graph with girth at least g has average degree less than $\frac{2g}{g-2}$. Some results on such graphs in fact hold whenever $\mathrm{mad}(G) < \frac{2g}{g-2}$, regardless of planarity, often with the same proof by discharging. Others, as discussed in Section 3, truly need planarity and may assign charge to both the faces and the vertices (the dual graph is also sparse). This is the basic reason why discharging is so useful for planar graphs. Subsequent sections will discuss additional techniques of discharging, especially with examples from "list coloring".

Finally, we note that in addition to its usefulness as a proof technique, the discharging method also has algorithmic implications, often yielding fast constructive algorithms for good colorings or embeddings. Iterative application of the structure theorem yields reductions to smaller graphs. After a good coloring of a base graph is found, the intermediate graphs receive good colorings using the reducibility arguments, until the original graph is restored and its coloring obtained (see Section 6 of [54]).

2. Structure and coloring of sparse graphs

In studying discharging, the principle and the details are simple. The mystery is the source of the discharging rule and the hypothesis on $\overline{d}(G)$. The secret is that the discharging rule is found before knowing the hypothesis of the theorem and is used to discover it. To explain such aspects of discharging, we study the forcing of local configurations with small weight.

Remark 2.1. *Finding the best bound on* mad(G). Consider Lemma 1.3 more generally. When we want G with mad$(G) < b$ to have a 1^--vertex or have a 2-vertex with a j^--neighbor, what is the best choice of b? Actually, we start with the proof and let it produce the statement. We must make b at most 3, since otherwise G may be 3-regular with no 2-vertex. Given that, when we exclude 1^--vertices and use degree charging, only 2-vertices will need charge. The most natural way for them to obtain it is to take it from their neighbors.

If each 2-vertex takes ρ from each neighbor, then the final charge is at least b at each vertex if and only if 2-vertices obtain enough charge and vertices with degree larger than j do not lose too much. Such vertices can lose ρ to each neighbor, so we need $2 + 2\rho \geq b$ and $d - d\rho \geq b$ when $d \geq j + 1$. To find the largest b that works, set $2 + 2\rho = (j + 1)(1 - \rho)$, yielding $\rho = \frac{j-1}{j+3}$ and hence $b = 2 + 2\rho = 4\frac{j+1}{j+3}$. When $j = 5$, we obtain Lemma 1.3.

What we did was find the weakest hypothesis allowing the discharging proof to work. The proof also provides sharpness examples showing that the condition mad$(G) < b$ cannot be weakened. If every 2-vertex has only $(j + 1)$-neighbors, every $(j + 1)$-vertex has only 2-neighbors, and there are no other vertices, then all the equalities are tight, no 2-vertex has a j^--neighbor, all vertices end with charge exactly b, and the average degree is b. Hence we obtain a sharpness example by taking a $(j + 1)$-regular graph and subdividing every edge.

What the discharging argument does is count part of the degree of higher-degree vertices at their 2-neighbors. In this sense discharging is "amortized counting"; the counting of the degree of a vertex is allocated to (or "charged to") other vertices.

The discharging argument for a structure theorem guaranteeing local configurations is quite separate from the reducibility arguments used to give an inductive proof of the desired conclusion. Thus the unavoidable set resulting from a particular sparseness condition may be usable to prove other results. In practice, usually the configurations are those already known to be reducible for the desired property in the application. Nevertheless, Lemma 1.3 does apply to another coloring problem.

Definition 2.2. An *acyclic coloring* of a graph is a proper coloring such that the union of any two color classes induces an acyclic subgraph; equivalently, no cycle is 2-colored.

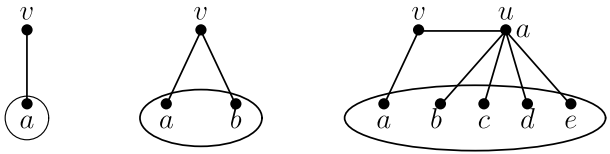

Fig. 1. Reducibility for acyclic 6-coloring.

Theorem 2.3. *If* $\mathrm{mad}(G) < 3$, *then* G *is acyclically* 6*-choosable.*

Proof. It suffices to show that the configurations forced by Lemma 1.3 when $\mathrm{mad}(G) < 3$ are reducible for the existence of an acyclic coloring chosen from a 6-uniform list assignment L. By definition, $\mathrm{mad}(G - v) < 3$. To show reducibility, we assume an acyclic L-coloring ϕ of $G - v$ and obtain such a coloring of G. The cases appear in Fig. 1.

If $d_G(v) \leq 1$, then we extend ϕ by letting $\phi(v)$ be a color in $L(v)$ not used on the neighbor of v. If $d_G(v) = 2$ and ϕ gives distinct colors on $N_G(v)$, then again we just avoid them on v. If $d_G(v) = 2$ and ϕ gives the same color to both vertices of $N_G(v)$, then there is danger of completing a 2-colored cycle. However, since v has a 5^--neighbor u, at most four other colors appear on the neighbors of u, so a color in $L(v)$ remains available for v. $\square$

If a structure theorem with hypothesis $\overline{d}(G) < b$ is sharp, then when $\overline{d}(G)$ exceeds b we must add other configurations to obtain a structure theorem. At $\overline{d}(G) = 3$ we may have no 2-vertices with 5^--neighbors and perhaps no 2-vertices at all. Nevertheless, when $\overline{d}(G) < 4$ the graph must have a 2^--vertex or have a 3-vertex with a 5^--neighbor (Exercise 2.1).

In the other direction, when we reduce the bound on $\overline{d}(G)$ we can impose more sparseness. By Remark 2.1, $\overline{d}(G) < \frac{12}{5}$ implies that G has two adjacent 2-vertices if it has no 1^--vertex. What sparser local configuration can we force when the average degree declines even further?

Definition 2.4. An *ℓ-thread* in a graph G is a trail of length $\ell + 1$ in G whose ℓ internal vertices have degree 2 in the full graph G.

Under this definition, an ℓ-thread contains two $(\ell - 1)$-threads, and the ends of a thread may be the same vertex.

Lemma 2.5. *If $\overline{d}(G) < 2 + \frac{2}{3\ell-1}$ and G has no 2-regular component, then G contains a 1^--vertex or an ℓ-thread.*

Proof. Let $\rho = \frac{1}{3\ell-1}$, so the hypothesis is $\overline{d}(G) < 2 + 2\rho$. Use degree charging. If neither stated configuration occurs, then we redistribute charge to leave each vertex with at least $2+2\rho$. Since G has no 1^--vertex, $\delta(G) \geq 2$. Since G has no 2-regular component, each 2-vertex lies in a unique maximal thread. Redistribute charge as follows:

(R1) Each 2-vertex v takes charge ρ from each end of its maximal thread.

Since each 2-vertex lies on a unique maximal thread, it ends with charge $2 + 2\rho$. Since ℓ-threads are forbidden, each j-vertex with $j \geq 3$ gives charge to at most $\ell - 1$ vertices along the thread started by each incident edge, losing at most $j(\ell - 1)\rho$. To show that its final charge is at least $2 + 2\rho$, we compute

$$j - j(\ell - 1)\rho \geq 3\left[1 - \frac{\ell - 1}{3\ell - 1}\right] = 2 + \frac{2}{3\ell - 1}.$$

Hence avoiding the specified configurations requires $\overline{d}(G) \geq 2 + \frac{2}{3\ell-1}$. $\square$

Remark 2.6. Once again, the hypothesis of Lemma 2.5 is discovered from the proof, and the structure theorem is sharp. When using degree charging with $\overline{d}(G) < 2 + 2\rho$, only 2-vertices need charge (once we restrict to $\delta(G) \geq 2$), and the natural (local) sources of charge are the nearest vertices of larger degree. This yields the discharging rule, taking ρ from each.

We choose ρ by finding the weakest hypothesis that avoids taking too much from 3^+-vertices. The inequality $j - j(\ell - 1)\rho \geq 2 + 2\rho$ implies that the proof guarantees an ℓ-thread when $\rho \leq \frac{j-2}{j(\ell-1)+2}$ for $j \geq 3$. Thus setting $\rho = \frac{1}{3\ell-1}$ both makes the proof work and gives the weakest hypothesis where it works.

Furthermore, to achieve sharpness in the proof, all vertices should have degree 2 or 3. Replace each edge of any 3-regular graph with an $(\ell - 1)$-thread. By the discharging computation, the average degree is $2 + \frac{2}{3\ell-1}$, and there are no ℓ-threads.

Discovering a discharging argument can be fun, but its value is in applications. To apply our result on threads inductively to a coloring problem, we replace the bound on $\bar{d}(G)$ by the same bound on $\mathrm{mad}(G)$. The condition $\mathrm{mad}(G) < 3$ already implies 3-colorability, and graphs having any odd cycle require at least three colors, so we need another coloring model in order to allow a stronger bound on $\mathrm{mad}(G)$ to have a chance to improve on 3-colorability.

Definition 2.7. A *t-fold coloring* of a graph G assigns each vertex a set of t colors so that adjacent vertices receive disjoint sets. The *t-fold chromatic number* $\chi_t(G)$ is the least k such that G has a t-fold coloring using subsets of $[k]$ (where $[n] = \{1, \ldots, n\}$). The *fractional chromatic number* $\chi^*(G)$ of G is $\inf_t \chi_t(G)/t$. The *odd girth* of G, written $g_o(G)$, is the length of a shortest odd cycle in G (infinite when G is bipartite).

An ordinary proper coloring is a 1-fold coloring, so always $\chi^*(G) \le \chi(G)$. The *independence number* $\alpha(G)$ of a graph G is the maximum size of an independent set of vertices. When G has n vertices, always a t-fold coloring of G needs at least $nt/\alpha(G)$ colors, since each color can only be used on an independent set. Hence $\chi^*(G) \ge n/\alpha(G)$; equality holds for vertex-transitive graphs using all automorphic images of a largest independent set. In particular, $\chi^*(C_{2t+1}) = 2 + \frac{1}{t}$, where C_n denotes the n-vertex cycle.

Theorem 2.8. *If $g_o(G) \ge 2t+1$ and $\mathrm{mad}(G) < 2+\frac{1}{3t-2}$, then G has a t-fold coloring with $2t+1$ colors, and hence $\chi^*(G) \le 2+\frac{1}{t}$.*

Proof. We have noted that $g_o(G) \ge 2t+1$ is needed. By Lemma 2.5, it suffices to show that 1^--vertices and $(2t-1)$-threads (which may be contained in longer threads) are reducible for t-fold $(2t+1)$-colorability. If $d(v) \le 1$, then a such a coloring ϕ of $G - v$ easily extends to v, choosing $\phi(v)$ from the complement of the set assigned to its neighbor when $d(v) = 1$.

When G contains a $(2t-1)$-thread with endpoints u and v, let G' be the graph obtained by deleting its internal vertices. The hypotheses hold for G', so G' admits a t-fold coloring ϕ using $2t+1$ colors. We want to extend ϕ along the thread. When two t-sets in $[2t+1]$ differ by one element, a unique t-set lies in the complement of both. Hence in two steps we can switch any color in a t-set to any missing color. In fact, this is the only change achievable in two steps (we can also return to the same t-set). Since there are $2t$ steps along the thread from u to v, we can thus extend ϕ along the thread to obtain the desired coloring of G. $\square$