Independence and the Havel–Hakimi residue

Jerrold R. Griggs*

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

Daniel J. Kleitman**

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 12 December 1990
 Revised 24 June 1991

Abstract

Favaron et al. (1991) have obtained a proof of a conjecture of Fajtlowicz’ computer program Graffiti that for every graph \( G \) the number of zeroes left after fully reducing the degree sequence as in the Havel–Hakimi Theorem is at most the independence number of \( G \). In this paper we present a simplified version of the proof of Graffiti’s conjecture, and we find how the residue relates to a natural greedy algorithm for constructing large independent sets in \( G \).

The Havel–Hakimi Theorem [5, 6] recursively characterizes the sorted integer sequences \( (d) = (d_1 \geq d_2 \geq \cdots \geq d_n \geq 0) \) which arise as the degree sequences of simple graphs. It states that \( (d) \) is graphically realizable if and only if the derived sequence \( L_1(d) \) is graphically realizable, where \( L_1(d) \) denotes the sorted sequence obtained after dropping the term \( d_1 \) and decreasing the \( d_i \) largest other terms by one each. Hence, \( (d) \) is graphical if and only if the repeated application of operation \( L_1 \) leads to a sequence of zeroes. For example,

\[
(633333331) \rightarrow (32222221) \rightarrow (2221111) \rightarrow \cdots \rightarrow (000),
\]

so \( (633333331) \) is graphically realizable.

Fajtlowicz [1] proposed considering the number of zeroes left when a graphically realizable sequence is reduced. This is called the residue, denoted \( R = R(d) \), e.g., \( R(633333331) = 3 \). His computer program ‘Graffiti’ checked many examples and found that in every case \( \alpha \geq R \), where \( \alpha = \alpha(G) \) is the size of the largest set of independent (i.e.,

*Supported in part by NSA/MSP Grant MDA90-H-4028.
**Supported in part by grants AF89-0271 and NSF DMS-8606225.
nonadjacent) vertices and $R$ is the residue for the degree sequence of $G$. Favaron, Mahéo and Saclé [2] discovered a rather complicated proof of this surprising inequality, which Fajtlowicz had announced as a conjecture. They proved other interesting facts about the residue.

A simplified version of the proof of Graffiti's conjecture is presented here. It is also the purpose of this note to learn how the residue is related to the result of a natural heuristic algorithm for producing large independent sets in $G$. This greedy algorithm, called Maxine by Fajtlowicz, repeatedly removes some vertex $v$ which has maximum degree in the remaining graph $H$, and then continues on the induced subgraph $H/v$, until no edges remain. We denote by $M = M(G)$ the size of the resulting independent set of vertices. Of course, $M$ depends on the choice of vertex $v$ when at some stage more than one vertex has maximum degree, e.g. for a path on 5 vertices, $M(P_5)$ is either 2 or 3. This algorithm was introduced by Johnson [7] and again by Griggs [3]. Obviously, $x \geq M$. We prove here that $M \geq R$, which was conjectured by Shearer [1].

The Havel–Hakimi method of realizing a sequence $(d)$ is to run the reduction procedure backwards: Beginning with an independent set of size $R(d)$, one successive-ly adds vertices so that, when added, they are adjacent to the correct number of vertices of largest degrees. The realization is not unique since there may be a choice of vertices of largest degree, e.g. if $(d) = (22211111)$, we begin with $R(d) = 4$ independent vertices and build up either to a path $P_5$ and two disjoint edges ($x = 5$) or to a triangle and three disjoint edges ($x = 4$). The graph consisting of three disjoint paths $P_3$ also realizes $(d)$, but is not Havel–Hakimi. It has $M = x = 6$.

We see that the sequence of reductions on a sequence $(d)$ leading to its residue $R(d)$ corresponds to a sequence of vertex deletions by Maxine on a Havel–Hakimi realization $G$ of $(d)$. So our result can be restated as: The minimum of the independent set size $M$ over all applications of Maxine on graphs realizing $(d)$ is the residue $R(d)$.

We say for degree sequences $(a) = (a_1 \geq \cdots \geq a_n)$ and $(b) = (b_1 \geq \cdots \geq b_n)$ that $(a)$ dominates $(b)$ providing that $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$ and for all $k$, $\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i$.

**Theorem 1.** For any graph $G$, $x \geq M \geq R$.

**Proof.** We deduce our theorem from the following fact, which is also central in the proof in [2].

**Lemma 2.** If sequences $(a)$ and $(b)$ are graphically realizable and $(a)$ dominates $(b)$, then $R(a) \geq R(b)$.

We may prove the theorem by observing that the degree sequence obtained when a vertex $v_1$ of largest degree in $G$ is deleted dominates the sequence $L_1(d)$, where $d = d(G)$. By the lemma, $R(d(G_v_1)) \geq R(L_1(d))$. By induction, $M(G_v_1) \geq R(d(G_v_1))$. Since by definition, $M(G) = M(G_v_1)$ and $R(d) = R(L_1(d))$, the theorem follows.

We prove the lemma itself by induction on the sequence length $n$ and on the sequence sum for any given value $n$. The result is trivial if $n$ is one or if the sum is zero.
It is well-known that if \((a)\) dominates \((b)\) and \((a)\) is graphically realizable, then so is \((b)\). (This can be deduced from the Erdös-Gallai conditions [2], or it can be proven as part of the induction here.) By transitivity, it suffices to prove the lemma when \((a)\) covers \((b)\) in the dominance order. This can happen in two ways. Either way, the sequences coincide in all but two entries, where they differ by just one. In one case, the two sequences disagree only in consecutive entries, say \((a)\) has \(y + 1, x\) and \((b)\) has \(y, x + 1\). In the other case, \((a)\) has consecutive entries \(x + 1, x\) \((j > 0\) times\), \(x - 1\), while \((b)\) has \(x\) \((j + 2\) times\).

If the first entries of \((a)\) and \((b)\) are identical, then one can compare the orders in which successive entries of \((a)\) and \((b)\) are decreased to form \(L_1(a)\) and \(L_1(b)\) to see that \(L_1(a)\) dominates \(L_1(b)\) (however, it can happen that \(L_1(a) = L_1(b)\) or that some sequence is between the two). Our claim then follows by induction:

\[
R(a) = R(L_1(a)) \geq R(L_1(b)) = R(b).
\]

For the remainder assume that \((a)\) and \((b)\) disagree in their first positions. (We cannot repeat the induction argument above since the sums for \(L_1(a)\) and \(L_1(b)\) no longer agree.)

Consider the second case above, where \((a)\) and \((b)\) also disagree in position \(j + 2 > 2\). First suppose that \(0 < j \leq x + 1\). Let \(z = a_{x + 2}\). We can check that \(L_1(a)\) and \(L_2(a)\) are equal except that \(L_1(a)\) begins with \(x - 1\) and \(L_2(a)\) has \(z\), and later \(L_1(a)\) has \(z - 1\) and \(L_2(a)\) has \(x\). Applying \(L_1\) again to each, one examines the stretch of terms in positions descended from \(z\) terms in \((a)\) and verifies that \(L_1(L_2(a)) = L_1(z(a))\). It is easily checked that \(L_2(a)\) dominates \(L_1(b)\), so we can apply induction to them and obtain

\[
R(a) = R(L_1(a)) = R(L_1(L_2(a))) = R(L_2(a)) \geq R(L_1(b)) = R(b).
\]

(Note that the sequence \(L_2(a)\) is graphical, since a suitable vertex can be added to a graph which realizes \(L_1(L_2(a)) = L_1(z(a))\). A similar remark applies to the special sequences \(L(a), (a'),\) and \((b')\) below.)

Suppose instead that \(j > x + 1\) in the second case above. Let \(L(b)\) be obtained from \((a)\) by removing one \(x\) and reducing \(x\) other copies of \(x\) to \(x - 1\). We get easily that \(L(a)\) covers \(L_1(b)\), so arguing as before, by induction, it suffices to observe that \(L_1(L(a)) = L_1^2(a)\). To do this, one defines \(\delta = j - 2x - 2\) and checks that terms \(\leq x - 2\) in \((a)\) are not affected in either sequence and that the initial terms \(\geq x - 1\) give rise in both sequences to \(\delta x\) terms when \(\delta \geq 0\) or to \(-\delta(x - 2)\) terms when \(\delta \leq 0\), and to \((x - 1)\) terms otherwise.

We finally come to the first type of covering above in which \((a)\) and \((b)\) disagree in just their first two positions. Let \(j \geq 0\) be the number of additional \(x\) terms after the first two terms.

If \(j \leq y\), then \(L_1\) reduces all \(x\) terms in \((a)\) to \(x - 1\). Let \(z = a_{y + 2} = b_{y + 2}\). Then \(L_1(a)\) begins with \(x - 1\) compared to \(x\) for \(L_1(b)\), and later \(L_1(a)\) has \(z - 1\) while \(L_1(b)\) has \(z\). One then checks that \(L_1^2(a)\) dominates \(L_1^2(b)\), so by induction we obtain

\[
R(a) = R(L_1^2(a)) \geq R(L_1^2(b)) = R(b).
\]
This leaves the case that \( j \geq y + 1 \). Now \( L_1^2(a) \) has smaller sum than \( L_1^2(b) \). Form a new sequence \( (a') \) from \( (a) \) by reducing \( y + 1 \) to \( y \) (at \( a_j \)) and the last \( x \) to \( x - 1 \) (at \( a_{j+2} \)). Form another new sequence \( (b') \) from \( (b) \) by reducing \( y \) to \( y - 1 \) and \( x + 1 \) to \( x \). One can then check that \( L_1(a) = L_1(a') \), \( L_1(b) = L_1(b') \), and \( (a') \) dominates \( (b') \). Since \( (a') \) has smaller sum than \( (a) \), on the same number of terms, our conclusion again follows by induction, which completes the proof of the lemma and the theorem. \( \square \)

References