

Testing Consumer Rationality using Perfect Graphs and Oriented Discs

Shant Boodaghians¹ and Adrian Vetta²

¹ Department of Mathematics and Statistics, McGill University
shant.boodaghians@mail.mcgill.ca

² Department of Mathematics and Statistics, and School of Computer Science, McGill University
vetta@math.mcgill.ca

Abstract. Given a consumer data-set, the axioms of revealed preference proffer a binary test for rational behaviour. A natural (non-binary) measure of the degree of rationality exhibited by the consumer is the minimum number of data points whose removal induces a rationalisable data-set. We study the computational complexity of the resultant CONSUMER RATIONALITY problem in this paper. We explain how to formulate this problem in terms of a directed revealed preference graph and show, for markets with a large number of commodities, that it is equivalent (in terms of approximation) to the *directed feedback vertex set* problem. Our main result is to obtain an exact threshold on the number of commodities that separates easy cases and hard cases. Specifically, for two-commodity markets the CONSUMER RATIONALITY problem is polynomial time solvable; we prove this via a reduction to the *vertex cover* problem on perfect graphs. For three-commodity markets, however, the problem is NP-complete; we prove this using a reduction from PLANAR 3-SAT that is based upon oriented-disc drawings.

1 Introduction

The theory of revealed preference, introduced by Samuelson [27, 28], has long been used in economics to test for rational behaviour. Specifically, given a set of m commodities with price vector \mathbf{p} , we wish to determine whether the consumer always demands an affordable bundle \mathbf{x} of maximum utility. To test this question, assume we are given a collection of *consumer data* $\{(\mathbf{p}_1, \mathbf{x}_1), (\mathbf{p}_2, \mathbf{x}_2), \dots, (\mathbf{p}_n, \mathbf{x}_n)\}$. Each pair $(\mathbf{p}_i, \mathbf{x}_i)$ denotes the fact that the consumer purchased the bundle of goods $\mathbf{x}_i \in \mathbf{R}^n$ when the prices were $\mathbf{p}_i \in \mathbf{R}^n$. (Here $\mathbf{R} = \mathbb{R}_{\geq 0}$ denotes the set of non-negative real numbers.) Now, assuming the consumer is rational, the selection of \mathbf{x}_i reveals information about the consumer's preferences; in particular, suppose that $\mathbf{p}_i \cdot \mathbf{x}_i \geq \mathbf{p}_i \cdot \mathbf{x}_j$ for some $j \neq i$. This means that the bundle \mathbf{x}_j was affordable, and available for selection, when \mathbf{x}_i was chosen. In this case, we say \mathbf{x}_i is *directly revealed preferred* to \mathbf{x}_j and denote this $\mathbf{x}_i \succeq \mathbf{x}_j$. Furthermore, suppose we observe that $\mathbf{x}_i \succeq \mathbf{x}_j$ and that $\mathbf{x}_j \succeq \mathbf{x}_k$. Then, by transitivity of preference, we say \mathbf{x}_i is *indirectly revealed preferred* to \mathbf{x}_k .

For clarity of presentation, we will assume that all the chosen bundles are distinct and that all revealed preferences are strict (no ties). For a rational consumer, the data-set should then have the following property:

The Generalized Axiom of Revealed Preference.¹

If $\mathbf{x}_1 \succeq \mathbf{x}_2$, $\mathbf{x}_2 \succeq \mathbf{x}_3$, \dots , $\mathbf{x}_{k-2} \succeq \mathbf{x}_{k-1}$ and $\mathbf{x}_{k-1} \succeq \mathbf{x}_k$ then $\mathbf{x}_k \not\succeq \mathbf{x}_1$.

Moreover, Afriat [1] showed that the Generalized Axiom of Revealed Preference (GARP) is also sufficient for the construction of a utility function which *rationalises* the data-set. That is, Afriat showed that if the consumer data satisfies GARP then one can construct a utility function $v : \mathbf{R}^n \rightarrow \mathbf{R}$ such that v is maximised at \mathbf{x}_i among the set of affordable bundles at prices \mathbf{p}_i . Hence, GARP is a necessary and sufficient condition for consumer rationality.

We can represent the preferences revealed by the consumer data via a directed graph, $D_{\succeq} = (V, A)$. This directed *revealed preference graph* contains a vertex $\mathbf{x}_i \in V$ for each data-pair $(\mathbf{p}_i, \mathbf{x}_i)$, and an arc from \mathbf{x}_i to \mathbf{x}_j if and only if $\mathbf{x}_i \succeq \mathbf{x}_j$. Observe that GARP holds *if and only if* the revealed preference graph is acyclic. Consequently, Afriat's theorem implies that the consumer is rational if and only if D_{\succeq} contains no directed cycles.

For example, Figure 1 displays visually two sets of consumer data. Each bundle \mathbf{x}_i is paired with its price vector \mathbf{p}_i , and a dotted line is drawn through \mathbf{x}_i perpendicular to \mathbf{p}_i . Note that $\mathbf{p}_i \mathbf{x}_i \geq \mathbf{p}_i \mathbf{y}$ if and only if \mathbf{y} lies on the opposite side of the dotted line to the drawing of \mathbf{p}_i . Hence, for the first consumer (left), we have $\mathbf{x}_3 \succeq \mathbf{x}_2$, $\mathbf{x}_3 \succeq \mathbf{x}_1$ and $\mathbf{x}_2 \succeq \mathbf{x}_1$. This produces an acyclic revealed preference graph D_{\succeq} and, therefore, her behaviour can be rationalized. On the otherhand, the second consumer (right) reveals $\mathbf{x}_3 \succeq \mathbf{x}_2 \succeq \mathbf{x}_3$. This produces a directed 2-cycle in D_{\succeq} and, so, her behaviour cannot be rationalised.

1.1 A Measure of Consumer Rationality

We have seen that graph acyclicity can be used to provide a test for consumer rationality. However such a test is binary and, in practice, leads to the immediate conclusion of irrationality, as observed data sets typically induce cycles in the revealed preference graph. Consequently, there has been a large body of experimental and theoretical work designed to measure how close to rational the behaviour of a consumer is. Examples include measurements based upon best-fit perturbation errors (*e.g.* Afriat [2] and Varian [34]), measurements based upon counting the number of rationality violations present

¹ When ties are possible, this formulation is called the *strong axiom of revealed preference*; see Houthakker [19]. We refer the reader to the survey by Varian [33] for details concerning the assorted axioms of revealed preference.

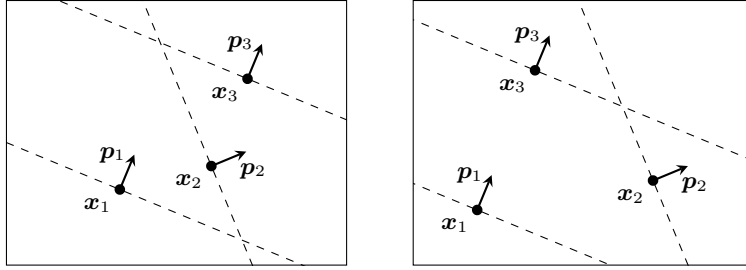


Fig. 1: A rational consumer and an irrational consumer.

in the data (*e.g.* Swofford and Whitney [32] and Famulari [17]), and measurements based upon the maximum size of a *rational* subset of the data (*e.g.* Koo [23] and Houtman and Maks [20]). Gross [13] provides a review and analysis of some of these measures. Recently new measures have been designed by Echenique et al. [12], Apesteguia and Ballester [3], and Dean and Martin [8].

Combinatorially, perhaps the most natural measure is simply to count the number of “irrational” purchases. That is, what is the minimum number of data-points whose removal induces a rational set of data? The associated decision problem is called the CONSUMER RATIONALITY problem.

CONSUMER RATIONALITY

Instance: Consumer data $(\mathbf{p}_1, \mathbf{x}_1), \dots, (\mathbf{p}_m, \mathbf{x}_m) \in \mathbf{R}^n \times \mathbf{R}^n$, and an integer k .

Problem: Is there a sub-collection of at most k data points whose removal produces a data set satisfying GARP?

We note that this CONSUMER RATIONALITY problem is dual to the measure of Houtman and Maks [20]. Using the graphical representation, it can be seen that the consumer rationality problem is a special case of the DIRECTED FEEDBACK VERTEX SET problem. In fact, as we explain in Section 2, when there are many goods, the two problems are equivalent. However, the consumer rationality problem becomes easier to approximate as the number of commodities falls. Indeed, the main contribution of this paper is to obtain an exact threshold on the number of commodities that separates easy cases (polynomial) and hard cases (NP-complete). In particular, we prove the problem is polytime solvable for a two-commodity market (Section 3), but that it is NP-complete for a three-commodity market (Section 4).

2 The General Case: Many Commodities

In this section we show that the CONSUMER RATIONALITY problem in full generality is computationally equivalent to the DIRECTED FEEDBACK VERTEX SET (DFVS) problem.

DIRECTED FEEDBACK VERTEX SET

Instance: A directed graph $D = (V, A)$, and an integer k .

Problem: Is there a set S of at most k vertices such that the induced subgraph $D[V \setminus S]$ is acyclic? (Such a set S is called a *feedback vertex set*.)

First, observe that the CONSUMER RATIONALITY PROBLEM is a special case of the DIRECTED FEEDBACK VERTEX SET PROBLEM: we have seen that the dataset is rationalizable if and only if the preference graph is acyclic. Thus, the minimum feedback vertex set in the preference graph D_{\succeq} clearly corresponds to the minimum number of data points that must be removed to create a rationalizable data-set.

On the other hand, provided the number of commodities is large, DFVS is a special case of the CONSUMER RATIONALITY PROBLEM. Specifically, Deb and Pai [9] show that for any directed graph D there is a data-set on $m = n$ commodities whose preference graph is $D_{\succeq} = D$; for completeness, we include the short proof of this result.

Lemma 2.1. [9] *Given sufficiently many commodities, we may construct any digraph as a preference graph.*

Proof. Let D be any digraph on n nodes. We will construct n pairs in $\mathbf{R}^n \times \mathbf{R}^n$ such that $D_{\succeq} \cong D$. Denote $\mathbf{p}^i = (p_1^i, \dots, p_n^i)$, and set $p_i^i = 1$, $p_j^i = 0$ for $j \neq i$. Similarly, denote $\mathbf{x}^i = (x_1^i, \dots, x_n^i)$, and set

$$x_j^i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } (i, j) \in D \\ 2 & \text{if } (i, j) \notin D \end{cases}$$

We then have, $\mathbf{p}_i \cdot \mathbf{x}_i = 1$, $\mathbf{p}_i \cdot \mathbf{x}_j = 0$ if we want an arc from i to j , and $\mathbf{p}_i \cdot \mathbf{x}_j = 2$ if we do not want an arc, as desired. \square

It follows that any lower and upper bounds on approximation for (the optimization version of) DFVS immediately apply to (the optimization version of) the CONSUMER RATIONALITY problem.² The exact hardness of ap-

² Interestingly, suppose we apply the construction in the proof of Lemma 2.1 to a symmetric directed graph D , where an arc (i, j) is in D if and only if (j, i) is in D . Then

proximation for DFVS is not known. The best upper bound is due to Seymour [29] who gave an $O(\log n \log \log n)$ approximation algorithm. With respect to lower bounds, the DIRECTED FEEDBACK VERTEX SET problem is NP-complete [21]. Furthermore, as we will see in Section 3, the CONSUMER RATIONALITY problem is at least as hard to approximate as VERTEX COVER. It follows that DFVS problem cannot be approximated to within a factor 1.36 [10] unless $P = NP$. A stronger hardness bound is achievable if we strengthen our underlying complexity assumption. Specifically, assuming the Unique Games Conjecture [22], the minimum directed feedback vertex set cannot be approximated to within any constant factor [16, 31].

Lemma 2.1 shows the equivalence with DIRECTED FEEDBACK VERTEX SET applies when the number of commodities is at least the size of the dataset. However, Deb and Pai [9] also show that for an m -commodity market, there exists a directed graph on $O(2^m)$ vertices that cannot be realised as a preference graph. This suggests that the hardness of the consumer rationality problem may vary with the quantity of goods. Indeed, we now prove that this is the case.

3 The Case of Two Commodities

We begin by outlining the basic approach to proving polynomial solvability for two goods. As described, the CONSUMER RATIONALITY problem is a special case of DVFS. For two goods, however, rather than considering all directed cycles, it is sufficient to find a vertex hitting set for the set of *digons* (directed cycles consisting of two arcs). The resulting problem can be solved by finding a minimum vertex cover in a corresponding auxiliary undirected graph. The VERTEX COVER problem is, of course, itself hard [10]. But we prove that the auxiliary undirected graph is perfect, and VERTEX COVER is polytime solvable in perfect graphs.

3.1 Two-Commodity Markets and the Vertex Cover Problem

So, our first step is to show that it suffices to hit only digons. Specifically, we prove that every vertex-minimal cycle in the revealed preference graph D_{\succeq} is a digon. This fact corresponds to the result that for two goods the *Weak Axiom of Revealed Preference* is equivalent to the *Generalised Axiom of Revealed Preference*. This equivalence was noted by Samuelson [28] and

this must also produce 2-cycles in the preference graph D_{\succeq} . Thus a subset S of vertices of D_{\succeq} is rational if and only if the induced graph $D[S]$ contains no arcs. As observed by Smeulders et al. [30], this immediately implies, that the Houtman-Maks measure [20] is as hard to approximate as the INDEPENDENT SET problem.

formally proven by Rose [26] in 1958; for a recent structurally motivated proof see [18]. For completeness, and to illustrate some of the notation and techniques required in this paper, we present a short geometric proof here.

We begin with the required notation. Let $\mathbf{x} = (x_1, x_2) \in \mathbf{R}^2$, and define

$$\mathbf{x}^{\searrow} := \{(y_1, y_2) \in \mathbf{R}^2 : y_1 \geq x_1, y_2 \leq x_2\} ,$$

i.e. the points which lie “below and to the right” of \mathbf{x} in the plane. Define \mathbf{x}^{\swarrow} , \mathbf{x}^{\nearrow} and \mathbf{x}^{\nwarrow} similarly. In addition, define $\mathbf{x}^{\searrow\searrow}$, $\mathbf{x}^{\swarrow\swarrow}$ and $\mathbf{x}^{\nwarrow\nwarrow}$ by replacing the inequalities with strict inequalities. Furthermore, if ℓ is a line in the plane of non-positive slope which intersects the positive quadrant, we say a point *lies below* ℓ if it lies in the same closed half-plane as the origin. For each data pair $(\mathbf{p}_i, \mathbf{x}_i)$, we define ℓ_i to be the line through \mathbf{x}_i perpendicular to \mathbf{p}_i . Hence, in our setting $\mathbf{x}_i \succeq \mathbf{x}_j$ if and only if \mathbf{x}_j lies below ℓ_i . Note that, if $\mathbf{x}_i \succeq \mathbf{x}_j$, then we may not have $\mathbf{x}_j \in \mathbf{x}_i^{\nearrow\nearrow}$ since \mathbf{p}_i is non-negative.

Lemma 3.1. [26] *For two commodities, every minimal cycle is a digon.*

Proof. Let $C_k = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k-1}\}$, listed in order, be a vertex-minimal directed cycle in D_{\succeq} . Suppose, for a contradiction, that $k \geq 3$. By minimality, the cycle C_k is chordless, therefore, $\mathbf{x}_i \succeq \mathbf{x}_j$ if and only if $j = i + 1 \pmod{k}$. (Henceforth, we will often assume without statement that indices are taken modulo k . Furthermore, “left” will stand for the negative x direction.) Without loss of generality, suppose \mathbf{x}_i is the leftmost bundle – or one of them. Since $\mathbf{x}_i \succeq \mathbf{x}_{i+1}$, we have that \mathbf{x}_{i+1} must fall in \mathbf{x}_i^{\searrow} . We claim that ℓ_i must be steeper than ℓ_{i+1} . To see this, suppose this is not true. Then, as shown in Figure 2(a), ℓ_{i+1} must intersect the line ℓ_i strictly to the left of \mathbf{x}_i . If not, $\mathbf{x}_{i+1} \succeq \mathbf{x}_i$. Now \mathbf{x}_{i+2} lies under ℓ_{i+1} but not under ℓ_i , but this implies that \mathbf{x}_{i+2} lies strictly to the left of \mathbf{x}_i as illustrated. This gives the desired contradiction. Hence, ℓ_i must be steeper than ℓ_{i+1} . This situation is illustrated

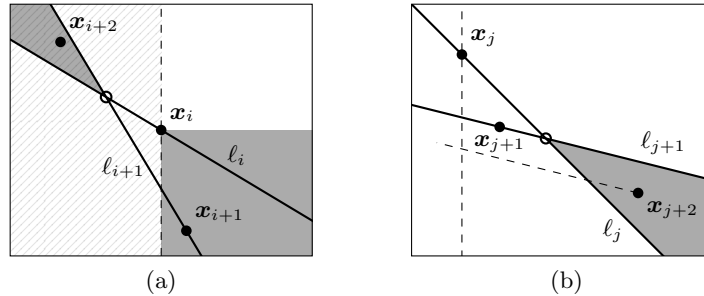


Fig. 2: Leftmost 2-commodity bundles on a cycle.

in Figure 2(b) where we set $j = i$. We claim the following:

Claim. Suppose $\mathbf{x}_{j+1} \in \mathbf{x}_j^{\searrow}$ and ℓ_j is steeper than ℓ_{j+1} , then we must have that $\mathbf{x}_{j+2} \in \mathbf{x}_{j+1}^{\searrow}$ and that ℓ_{j+1} is steeper than ℓ_{j+2} .

As shown in Figure 2(b), because ℓ_j is steeper than ℓ_{j+1} , we must have $\mathbf{x}_{j+2} \in \mathbf{x}_{j+1}^{\searrow}$. It remains to show that ℓ_{j+1} is steeper than ℓ_{j+2} . Suppose not, then, since \mathbf{x}_{j+1} must fall above ℓ_{j+2} , the (highlighted) point where ℓ_j and ℓ_{j+1} meet must also fall above ℓ_{j+2} . Thus, the region which falls above both ℓ_j and ℓ_{j+1} cannot intersect the region below ℓ_{j+2} . Therefore, there is no valid position for \mathbf{x}_{j+3} . Consequently, ℓ_{j+1} must be steeper than ℓ_{j+2} , as desired.

Hence, by induction, for every $0 \leq j \leq k - 1$, we have that ℓ_j is steeper than ℓ_{j+1} and that $\mathbf{x}_{j+1} \in \mathbf{x}_j^{\searrow}$, where our base case is $j = i$. However, this cannot hold for $j = i - 1$, since \mathbf{x}_i is the leftmost point in the cycle, amounting to a contradiction, and refuting the assumption that there existed a minimal cycle on at least 3 vertices. \square

Lemma 3.1 implies that a vertex set that intersects every digon will also intersect each directed cycle of any length. Hence, to solve the CONSUMER RATIONALITY problem for two goods, it suffices to find a minimum cardinality hitting vertex set for the digons of D_{\succeq} . We can do this by transforming the problem into one of finding a minimum vertex cover in an undirected graph. Recall the VERTEX COVER problem is:

VERTEX COVER

INSTANCE: Given an undirected graph $G = (V, E)$ and an integer k .

PROBLEM: Is there a set S of at most k vertices such that every edge has an endpoint in S ?

The transformation is then as follows: given the directed revealed preference graph D_{\succeq} we create an *auxiliary undirected graph* G_{\succeq} . The vertex set $V(G_{\succeq}) = V(D_{\succeq})$ so the undirected graph also has a vertex for each bundle \mathbf{x}_i . There is an edge $(\mathbf{x}_i, \mathbf{x}_j)$ in G_{\succeq} if and only if \mathbf{x}_i and \mathbf{x}_j induce a digon in D_{\succeq} . It is easy to verify that a vertex cover in G_{\succeq} corresponds to a hitting set for digons of D_{\succeq} .

Let's see some simple examples for the auxiliary graph G_{\succeq} . First consider Figure 3(a), where bundles are placed on a concave curve. Now every pair of vertices \mathbf{x}_i and \mathbf{x}_j induce a digon in D_{\succeq} . Thus G_{\succeq} is an undirected clique. Now consider Figure 3(b). The vertices on the left induce a directed path in D_{\succeq} ; the vertices along the bottom also induce a directed path in D_{\succeq} . However each pair consisting of one vertex on the left and one vertex on the bottom induce a digon in D_{\succeq} . Thus G_{\succeq} is a complete bipartite graph.

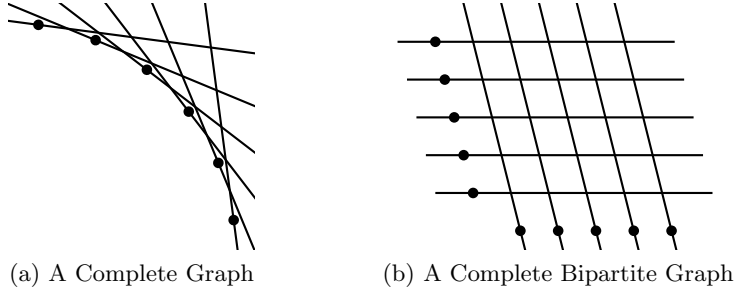


Fig. 3: Examples of the Auxiliary Undirected Graph.

3.2 Perfect Graphs

An undirected graph G is perfect if the chromatic number of any induced subgraph is equal to the cardinality of the maximum clique in the subgraph. In 1961, Berge [5] made the famous conjecture that an undirected graph is perfect if and only if it contains neither an odd length hole nor an odd length antihole as induced subgraphs. Here a *hole* is a chordless cycle with at least four vertices. An *antihole* is the complement of a chordless cycle with at least four vertices. Berge’s conjecture was finally proven by Chudnovsky, Robertson, Seymour and Thomas [7] in 2006.

Theorem 3.1 (The Strong Perfect Graph Theorem [7]). *An undirected graph is perfect if and only if it contains no odd holes and no odd antiholes.*

There are many important classes of perfect graphs, for example, cliques, bipartite graphs, chordal graphs, line graphs of bipartite graphs, and comparability graphs.³ Interestingly, we now show that the class of 2D auxiliary revealed preference graphs are also perfect. To prove this, we will need the following geometric lemma.

Lemma 3.2. *Let $\{x_i, x_j, x_k\}$, listed in order, be an induced path in the 2D auxiliary revealed preference graph G_{\succeq} . If $x_i \in x_j^{\nwarrow}$ then $x_k \in x_j^{\nwarrow}$. (Similarly, if $x_i \in x_j^{\searrow}$ then $x_k \in x_j^{\searrow}$.)*

Proof. Recall the assumption that the bundles distinct, that is, $x_i \neq x_j$ for all $i \neq j$. Because $\{x_i, x_j\}$ is an edge in the auxiliary undirected graph G_{\succeq} , we know that $x_i \succeq x_j$ and $x_j \succeq x_i$. Therefore it cannot be the case that

³ By the (Weak) Perfect Graph Theorem [24], the complements of these classes of graphs are also perfect.

$\mathbf{x}_i \in \mathbf{x}_j^{\nearrow\searrow}$ or $\mathbf{x}_j \in \mathbf{x}_i^{\nearrow\searrow}$. Thus, either $\mathbf{x}_j \in \mathbf{x}_i^{\nwarrow}$ or $\mathbf{x}_j \in \mathbf{x}_i^{\swarrow}$, but not both. Similarly, because $\{x_j, x_k\}$ is an edge in G_{\succeq} , either $\mathbf{x}_k \in \mathbf{x}_j^{\nwarrow}$ or $\mathbf{x}_k \in \mathbf{x}_j^{\swarrow}$.

Now, without loss of generality, let $\mathbf{x}_i \in \mathbf{x}_j^{\nwarrow}$. For a contradiction, assume that $\mathbf{x}_k \in \mathbf{x}_j^{\swarrow}$. Hence, we have $\mathbf{x}_j \in \mathbf{x}_i^{\swarrow} \cap \mathbf{x}_k^{\nwarrow}$. Suppose \mathbf{x}_j lies strictly below the line $\ell_{i,k}$ through \mathbf{x}_i and \mathbf{x}_k . But then we cannot have both $\mathbf{x}_j \succeq \mathbf{x}_i$ and $\mathbf{x}_j \succeq \mathbf{x}_k$. This is because the line ℓ_j must cross the segment of $\ell_{i,k}$ between \mathbf{x}_i and \mathbf{x}_k if it is to induce either of the two preferences. Thus, the line ℓ_j separates \mathbf{x}_i and \mathbf{x}_k and, so, at most one of bundles can lie below the line. This is illustrated in Figure 4(a).

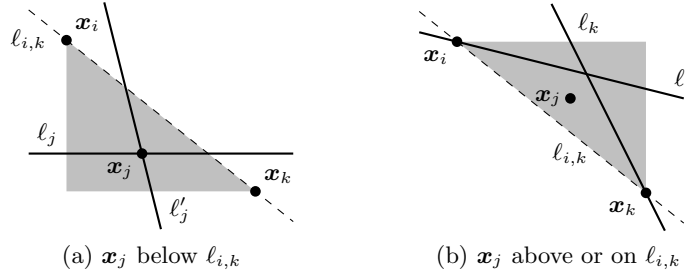


Fig. 4: Induced path on three vertices.

On the other hand, suppose \mathbf{x}_j lies on or above the line $\ell_{i,k}$ through \mathbf{x}_i and \mathbf{x}_k . Now we know that $\mathbf{x}_i \succeq \mathbf{x}_j$. This implies that $\mathbf{x}_i \succeq \mathbf{x}_k$, as illustrated in Figure 4(b). Furthermore, we know that $\mathbf{x}_k \succeq \mathbf{x}_j$ which implies that $\mathbf{x}_k \succeq \mathbf{x}_i$. Thus $\{x_i, x_j\}$ is an edge in G_{\succeq} . This contradicts the fact that $\{x_i, x_j, x_k\}$ is an induced path. \square

Lemma 3.3. *The 2D auxiliary revealed preference graph G_{\succeq} contains no odd holes on at least 5 vertices.*

Proof. Take a hole $C_k = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k-1}\}$, listed in order, where $k \geq 5$ is odd. For any $0 \leq i \leq k-1$, the three vertices $\{\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}\}$ induce a path in G_{\succeq} . Consequently, by Lemma 3.2, either both \mathbf{x}_{i-1} and \mathbf{x}_{i+1} are in \mathbf{x}_i^{\nwarrow} or both \mathbf{x}_{i-1} and \mathbf{x}_{i+1} are in \mathbf{x}_i^{\swarrow} . In the former case, colour \mathbf{x}_i yellow. In the latter case, colour \mathbf{x}_i red. Thus we obtain a 2-coloring of C_k . Since k is odd, there must be two adjacent vertices, \mathbf{x}_i and \mathbf{x}_{i+1} , with the same colour. Without loss of generality, let both vertices be yellow. Thus, \mathbf{x}_{i+1} is \mathbf{x}_i^{\nwarrow} and \mathbf{x}_i is in $\mathbf{x}_{i+1}^{\nwarrow}$. This contradicts the distinctness of \mathbf{x}_i and \mathbf{x}_{i+1} . \square

We remark that the parity condition in Lemma 3.3 is necessary. To see this consider the example in Figure 5 which produces an even hole on six ver-

tices. Specifically, the only mutually adjacent pairs are the $(\mathbf{x}_i, \mathbf{x}_{i+1})$ pairs, with indices taken modulo 6.

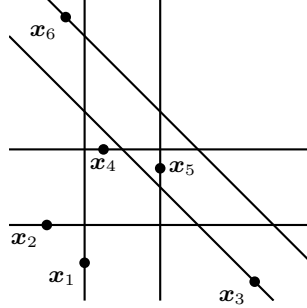


Fig. 5: Construction for C_6

Lemma 3.4. *The 2D auxiliary revealed preference graph G_{\succeq} contains no antiholes on at least 5 vertices.*

Proof. Note that the complement of an odd hole on five vertices is also an odd hole. Thus, by Lemma 3.3, the graph G_{\succeq} may not contain an antihole on five vertices.

Next consider an antihole $\bar{C}_k = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{k-1}\}$, listed in order, with $k \geq 6$. The neighbours in \bar{C}_k of \mathbf{x}_i , for any $0 \leq i \leq k-1$, are $\Gamma_i = \{\mathbf{x}_{i+2}, \mathbf{x}_{i+3}, \dots, \mathbf{x}_{i-2}\}$. We claim that either every vertex of Γ_i is in \mathbf{x}_i^{\nwarrow} or every vertex of Γ_i is in \mathbf{x}_i^{\searrow} . To see this note that $(\mathbf{x}_{i+2}, \mathbf{x}_{i+3})$ is not an edge, and therefore $\{\mathbf{x}_{i+2}, \mathbf{x}_i, \mathbf{x}_{i+3}\}$ is an induced path in G_{\succeq} . By Lemma 3.2, without loss of generality, both \mathbf{x}_{i+2} and \mathbf{x}_{i+3} are in \mathbf{x}_i^{\nwarrow} . But $\{\mathbf{x}_{i+3}, \mathbf{x}_i, \mathbf{x}_{i+4}\}$ is also an induced path in G_{\succeq} . Consequently, as \mathbf{x}_{i+3} is in \mathbf{x}_i^{\nwarrow} , Lemma 3.2 implies that \mathbf{x}_{i+4} is in \mathbf{x}_i^{\nwarrow} . Repeating this argument through to the induced path $\{\mathbf{x}_{i-3}, \mathbf{x}_i, \mathbf{x}_{i-2}\}$ gives the claim.

Now consider the three vertices $\mathbf{x}_0, \mathbf{x}_2$ and \mathbf{x}_4 . Since $k \geq 6$ these vertices are pairwise adjacent in \bar{C}_k . Without loss of generality, by the claim, Γ_0 is in \mathbf{x}_0^{\nwarrow} . Thus, \mathbf{x}_2 and \mathbf{x}_4 are in \mathbf{x}_0^{\nwarrow} . However \mathbf{x}_0 is in $\Gamma_2 \cap \Gamma_4$. Thus every vertex in Γ_2 is in \mathbf{x}_2^{\searrow} and every vertex in Γ_4 is in \mathbf{x}_4^{\searrow} . Hence, \mathbf{x}_4 is in \mathbf{x}_2^{\searrow} and \mathbf{x}_2 is in \mathbf{x}_4^{\searrow} , a contradiction. \square

Lemmas 3.3 and 3.4 together show, by applying the Strong Perfect Graph Theorem, that the auxiliary undirected graph is perfect.

Theorem 3.2. *The 2D auxiliary revealed preference graph G_{\succeq} is perfect.*

\square

3.3 A Polynomial Time Algorithm

In classical work, Grötschel, Lovász and Schrijver [14, 15] show that the VERTEX COVER problem in a perfect graph can be solved in polynomial time via the ellipsoid method.

Theorem 3.3. [14] *The VERTEX COVER problem is solvable in polynomial time on a perfect graph.* \square

But by Theorem 3.2, the auxiliary undirected graph is perfect. Since the consumer rationality problem for two commodities corresponds to a vertex cover problem on this auxiliary undirected graph, we have:

Theorem 3.4. *In a two-commodity market, the CONSUMER RATIONALITY problem is solvable in polynomial time.* \square

4 The Case of Three Commodities

We have shown that for two commodities, the consumer rationality problem can be solved in polynomial time. We now prove the problem is NP-complete if there are three (or more) commodities by presenting a reduction from PLANAR 3-SAT. The proof has three parts: first we transform an instance of PLANAR 3-SAT to an instance of VERTEX COVER in an associated undirected *gadget graph*. Second, we show that a vertex cover in the gadget graph corresponds to a directed feedback vertex set in a directed *oriented disc graph*. Finally, we prove that every oriented disc graph corresponds to a preference graph in a three-commodity market. Consequently, we can solve PLANAR 3-SAT using an algorithm for the three-commodity case of the CONSUMER RATIONALITY problem.

We begin by defining the class of oriented-disc graphs. Let $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be points in the plane and let $\{B_1, \dots, B_n\}$ be closed discs of varying radii such that B_i contains \mathbf{x}_i on its boundary. We call this collection of points and discs an *oriented-disc drawing*. Given a drawing, we construct a directed graph $D = (V, A)$ on the vertex set $V = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. There is an arc from \mathbf{x}_i to \mathbf{x}_j in D if $\mathbf{x}_j, j \neq i$, is contained in the disc B_i . A directed graph that can be built in this manner is called an *oriented-disc graph*.

An example is given in Figure 6. The oriented-disc drawing is shown on the left and the the resulting oriented disc graph, a directed cycle on 3 vertices, is shown on the right. (We remark that, for enhanced clarity in the larger figures that follow, the boundary circles are drawn half-dotted.) Note that, even if the discs have uniform radii, the resulting oriented-disc graphs need not be symmetric – that is, $(\mathbf{x}_i, \mathbf{x}_j)$ can be an arc even if $(\mathbf{x}_j, \mathbf{x}_i)$ is not. This is due to the fact that \mathbf{x}_i lies on the boundary, not at the centre,

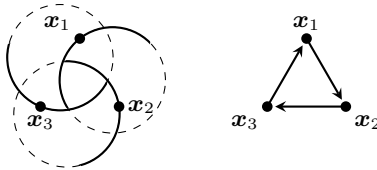


Fig. 6: An oriented disc drawing and its corresponding oriented disc graph.

of its disc B_i . We now start by proving the third part of the reduction: every oriented disc graph corresponds to a preference graph in a three-commodity market.

Lemma 4.1. *Every oriented-disc graph corresponds to a preference graph induced by consumer data in a three-commodity market.*

Proof. Let D be any oriented-disc graph. We wish to build a three-commodity data set whose preference graph is D . Recall that the plane is homomorphic to the 3-dimensional sphere minus a point. Moreover, the inverse of the *stereographic projection* is a map from the plane to a sphere which preserves the shape of circles; see, for example, [11]. This motivates us to attempt to draw the points and discs on the unit sphere centered at $(1, 1, 1) \in \mathbf{R}^3$. To do this, we scale the oriented-disc drawing appropriately and embed it in a small region on the “underside” of the sphere, that is, around the point where the inwards normal vector is $(1, 1, 1)$. An example of this, where the oriented-disc graph is the directed 3-cycle, is shown in Figure 7(a).

We now need to create the corresponding collection of consumer data. Let $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be the n points of some oriented-disc drawing of D embedded onto the underside of the sphere. Note that the intersection of a sphere and a plane is a circle. Furthermore, a plane through a point on the sphere will create a circle containing that point. Thus we may select the \mathbf{x}_i to be the bundles chosen by the market and we may choose \mathbf{p}_i such that the plane with normal \mathbf{p}_i that passes through \mathbf{x}_i intersects the sphere exactly along the boundary of the embedding of the disc B_i . An example is shown in Figure 7(b). Because \mathbf{p}_i is non-negative it points into the sphere. Therefore, \mathbf{x}_i is revealed preferred to every point on the inside of the embedding of B_i ; it is not revealed preferred to any other point on the sphere. Hence, the preference graph D_{\succeq} is isomorphic to the original oriented-disc graph, as desired. \square

Now, recall the first part of the reduction: we wish to transform an instance of PLANAR 3-SAT to an instance of VERTEX COVER in an associated

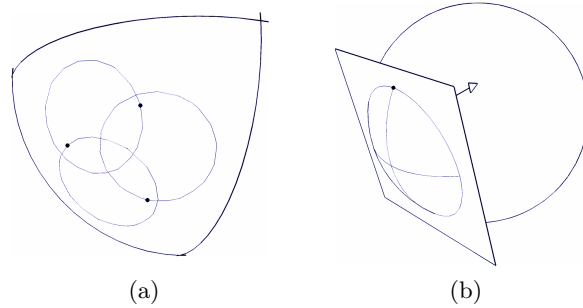


Fig. 7: A 3-cycle embedded on a sphere section, and a disc on a sphere.

undirected *gadget graph*. Our gadget graph is based upon a network used by Wang and Kuo [35] to prove the hardness of MAXIMUM INDEPENDENT SET in undirected unit-disc graphs. However, we are able to simplify their non-planar network by using an instance of PLANAR 3-SAT rather than the general 3-SAT. This simplification will be useful when implementing the second part of the reduction.

Let φ be an instance of PLANAR 3-SAT with variables u_1, \dots, u_n and clauses C_1, \dots, C_m . Recall that φ is *planar* if the bipartite graph H_φ – consisting of a vertex for each variable, a vertex for each clause, and edges connecting each clause to its three variables – is planar. The associated, undirected, *gadget graph* G_φ is constructed as shown in Figure 8. For each clause $C = (u_i \vee u_j \vee u_k)$, add to the graph a 3-cycle whose vertices are labelled by the appropriate literals for the variables u_i, u_j and u_k . We call these the *clause gadgets*. For each variable u_i , add a large cycle of even length whose vertices are alternately labelled as the literals u_i and \bar{u}_i . We call these the *variable gadgets*. Finally, add an edge from each variable in the clause gadgets to some vertex on the corresponding variable gadget with the opposite label – we choose a different variable vertex for each clause it is contained in.

The next lemma then follows by applying the arguments of Wang and Kuo [35].

Lemma 4.2. [35] *The PLANAR 3-SAT instance φ is satisfiable if and only if G_φ has vertex cover set of size at most $2m + \frac{1}{2} \sum_{i=1}^n r_i$, where r_i is the number of vertices in the variable gadget’s cycle for u_i .*

Proof. Suppose φ is satisfiable. Take any satisfying assignment, and let U be the set of literals which take TRUE values in the assignment, *i.e.* the literal “ u_i ” if the variable u_i was assigned TRUE, and the literal “ \bar{u}_i ” if the variable u_i was assigned FALSE. Let \bar{U} be the remaining literals. Now, every vertex in

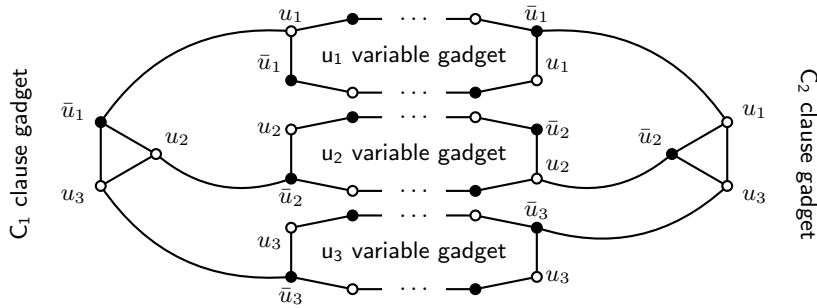


Fig. 8: The gadget graph G_φ for $\varphi = (\bar{u}_1 \vee u_2 \vee u_3) \wedge (u_1 \vee \bar{u}_2 \vee u_3)$.

a variable gadget of G_φ whose label is in \bar{U} will be selected to be in the vertex cover. In total this amounts to $\frac{1}{2} \sum_{i=1}^n r_i$ vertices, and these *cover* every edge in the variable gadgets of G_φ . Next consider the clause gadgets of G_φ . We must select two vertices of each clause gadget to cover the edges of the 3-cycle. This amounts to $2m$ vertices. Since we have a satisfying assignment and we chose the nodes corresponding to \bar{U} in the variable gadgets, each clause gadget must have at least one incident edge covered by the variable gadgets' selected vertices. Hence, selecting the other two vertices will cover all incident edges to the clause gadgets, and all edges of the gadget's cycle. Thus we have a vertex cover with $2m + \frac{1}{2} \sum_{i=1}^n r_i$ nodes. For example, in Figure 8, if we set all variables to FALSE, one possible vertex cover is the set of vertices labelled with non-negated literals, *i.e.* those coloured in white. (Note, clearly, the set of white vertices will not typically form a vertex cover in the gadget graph.)

Conversely, suppose we have a vertex cover \mathcal{C} containing at most $2m + \sum_{i=1}^n \frac{r_i}{2}$ vertices. Each variable gadget must contribute at least $\frac{r_i}{2}$ vertices, otherwise we cannot cover every edge in its cycle. Each clause gadget must contribute at least two vertices, or one edge in the 3-cycle will be uncovered. Hence, \mathcal{C} contains exactly $2m + \sum_{i=1}^n \frac{r_i}{2}$ vertices. The $\frac{1}{2}r_i$ vertices from the variable gadget for u_i corresponds either to the set of all vertices with negated labels or to the set with non-negated labels, otherwise there is an uncovered edge in the cycle. This induces a truth assignment; set u_i to TRUE if all the " \bar{u}_i "-labelled vertices are selected, and FALSE if the " u_i "-labelled vertices are selected. Furthermore this is a satisfying assignment. To see this note that as \mathcal{C} covers all edges, the unselected vertex in each clause is a literal which evaluates to TRUE by the selected assignment. \square

Hence, to solve for the satisfiability of φ , it suffices to test whether G_φ admits a vertex cover with at most $2m + \frac{1}{2} \sum_{i=1}^n r_i$ vertices. It remains to show

the second of the three parts of the reduction. That is, we need to show that this VERTEX COVER problem in the undirected gadget graph can be solved by finding a minimum directed feedback vertex set in an oriented-disc graph D_φ . The basic idea is straightforward (albeit that the implementation is intricate). The oriented-disc graph D_φ will contain a digon for each edge in some G_φ . However, it will also contain a collection of additional arcs. The key fact will be that these additional arcs form an acyclic subgraph of D_φ . Thus every cycle in D_φ must induce a digon. Consequently, a minimum directed feedback vertex set need only intersect each digon to ensure that every cycle is hit. As argued previously, hitting the underlying graph formed by the digons of D_φ corresponds to selecting a vertex cover in G_φ , as desired. We now formalise this argument.

Lemma 4.3. *For every instance φ of PLANAR 3-SAT, there exists an oriented-disc graph D_φ on which the DIRECTED FEEDBACK VERTEX SET problem is equivalent to the VERTEX COVER problem on G_φ .*

Proof. We prove this by explicitly constructing the oriented-disc drawing. Recall the disc graph D_φ should contain a digon for each edge in G_φ . To do this, we begin with a sufficiently large planar drawing of H_φ , the planar bipartite network associated with φ . At each clause vertex, we place an oriented-disc construction for the clause gadget. This construction, along with its resulting graph, is shown in Figure 9. The figure shows a clause gadget and a segment of each of the three neighbouring variable gadgets to which it is attached. Observe from the figure that, as claimed, the set of arcs created in D_φ which are not in a digon, form an acyclic subgraph of D_φ .

It remains to construct the large cycles for the variable gadgets, and connect them to the clause gadgets. However, segments of these cycles are already included in the clause gadgets. Thus, it suffices to connect these cycle segments via paths of digons. This can be done using the oriented disc constructions shown in Figure 10. To draw the cycle for some variable, say u_i , we note that u_i 's vertex in the planar network H_φ shares an edge with every clause gadget which connects to u_i 's gadget. Hence, as illustrated in Figure 11, we may follow along the edges of H_φ to construct the cycle. For example, in the figure, the variable cycle for u_1 (highlighted) follows the topology of the edges incident to u_1 's vertex, and joins the clause gadgets (circled) to one another.

Observe that constructions in Figure 10 produce paths of digons in D_φ , where every arc produced is contained in a digon. It follows that the only arcs in D_φ that are not in digons are in the neighbourhoods of the clause gadgets and, as we have seen, these are acyclic. But then, to hit all the cycles

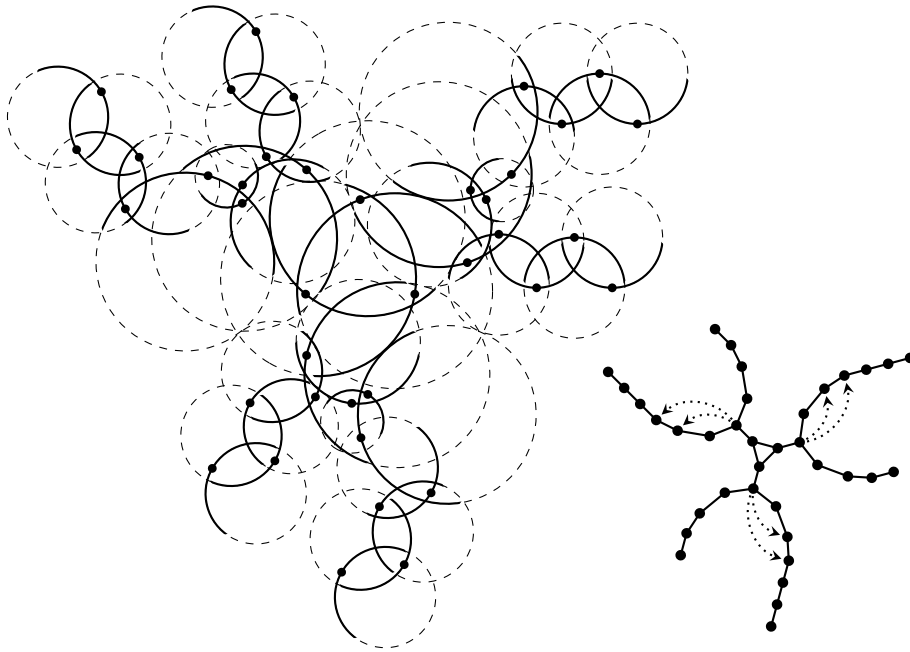


Fig. 9: Oriented-disc graph construction of the clause gadget, and its resulting graph. Solid edges represent digons.

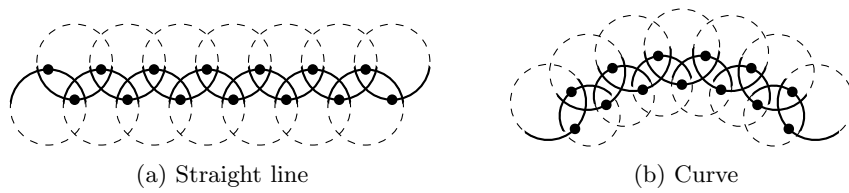


Fig. 10: Paths of bidirected edges as oriented-disc drawings.

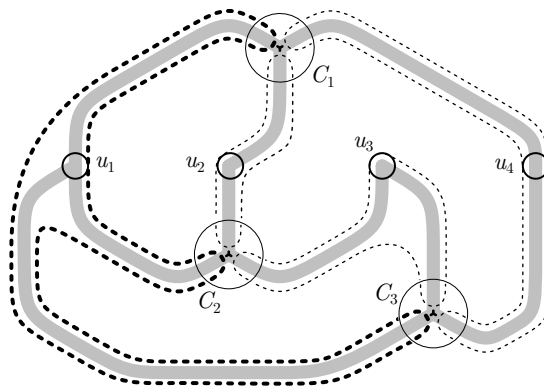


Fig. 11: G_φ as an oriented-disc graph.

in D_φ , it suffices to hit all the digons, which, in turn, corresponds to a vertex cover in G_φ , completing the proof. \square

This completes all the steps in the reduction and we obtain:

Theorem 4.1. *The CONSUMER RATIONALITY problem is NP-complete for a market with at least 3 commodities.* \square

References

1. S. Afriat, “The construction of a utility function from expenditure data”, *International Economic Review*, **8**, pp67–77, 1967.
2. S. Afriat, “On a system of inequalities in demand analysis: an extension of the classical method”, *International Economic Review*, **14**, pp460-472, 1967.
3. J. Apestequia and M. Ballester, “A measure of rationality and welfare”, to appear in *Journal of Political Economy*, 2015.
4. L. Ausubel, P. Cramton and P. Milgrom, “The clock-proxy auction: a practical combinatorial auction design”, in P. Cramton, Y. Shoham and R. Steinberg (eds.), *Combinatorial Auctions*, pp115–138, MIT Press, 2006.
5. C. Berge, “Färbung von Graphen deren sämtliche beziehungsweise deren ungerade Kreise starr sind (Zusammenfassung)”, *Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe*, **10**, pp114-115, 1961.
6. S. Choi, R. Fisman, D. Gales and S. Kariv, “Consistency and heterogeneity of individual behaviour under uncertainty”, *American Economic Review*, **97(5)**, pp1921–1938, 2007.
7. M. Chudnovsky, N. Robertson, P. Seymour and R. Thomas. “The strong perfect graph theorem”, *Annals of Mathematics*, pp51–229, 2006.
8. M. Dean and D. Martin, “Measuring rationality with the minimum cost of revealed preference violations”, to appear in *Review of Economics and Statistics*, 2015.
9. R. Deb and M. Pai, “The geometry of revealed preference”, *Journal of Mathematical Economics*, **50**, pp203–207, 2014.
10. I. Dinur and S. Safra, “On the hardness of approximating minimum vertex cover”, *Annals of Mathematics*, **162(1)**, pp439-485, 2005.
11. Earl, R. (2007). Geometry II: 3.1 Stereographic Projection and the Riemann Sphere. Retrieved from <https://people.maths.ox.ac.uk/earl/G2-lecture5.pdf>
12. F. Echenique, S. Lee and M. Shum, “The money pump as a measure of revealed preference violations”, *Journal of Political Economy*, **119(6)**, pp1201–1223, 2011.
13. J. Gross, “Testing data for consistency with revealed preference”, *The Review of Economics and Statistics*, **77(4)**, pp701–710, 1995.
14. M. Grötschel, L. Lovász and A. Schrijver, “Polynomial algorithms for perfect graphs”, *Annals of Discrete Mathematics*, **21**, pp325-356, 1984.

15. M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimisation*, Springer-Verlag, 1988.
16. V. Guruswami, J. Hastad, R. Manokaran, P. Raghavendra, and M. Charikar, “Beating the random ordering is hard: Every ordering CSP is approximation resistant”, *SIAM J. Computing*, **40(3)**, pp878–914, 2011.
17. M. Famulari, “A household-based, nonparametric test of demand theory”, *Review of Economics and Statistics*, **77**, pp372-383, 1995.
18. J. Heufer, “A geometric approach to revealed preference via Hamiltonian cycles”, *Theory and Decision*, **76(3)**, pp329–341, 2014.
19. H. Houthakker, “Revealed preference and the utility function”, *Economica, New Series*, **17(66)**, pp159–174, 1950.
20. M. Houtman and J. Maks, “Determining all maximal data subsets consistent with revealed preference”, *Kwantitatieve Methoden*, **19**, pp89–104, 1950.
21. R. Karp, “Reducibility among combinatorial problems”, *Complexity of Computer Computations*, New York: Plenum, pp85–103, 1972.
22. S. Khot, “On the power of unique 2-prover 1-round games”, Proceedings of STOC, pp767–775, 2002.
23. A. Koo, “An empirical test of revealed preference theory”, *Econometrica*, **31(4)**, pp646–664, 1963.
24. L. Lovász, “Normal hypergraphs and the perfect graph conjecture”, *discrete Mathematics*, **2(3)**, pp253–267, 1972.
25. D. Lichtenstein, “Planar formulae and their uses”, *SIAM Journal on Computing*, **11(2)**, pp329–343, 1982.
26. H. Rose. “Consistency of preference: the two-commodity case”, *Review of Economic Studies*, **25**, pp124–125, 1958.
27. P. Samuelson, “A note on the pure theory of consumer’s behavior”, *Economica*, **5(17)**, pp61–71, 1938.
28. P. Samuelson, “Consumption theory in terms of revealed preference”, *Economica*, **15(60)**, pp243–253, 1948.
29. P. Seymour, “Packing directed circuits fractionally”, *Combinatorica*, **15(2)**, pp281–288, 1995.
30. B. Smeulders, F. Spieksma, L. Cherchye and B. de Rock, “Goodness of fit measures for revealed preference tests: complexity results and algorithms”, *Transactions on Economics and Computation*, **2(1)**, Article #3, 2014.
31. O. Svensson, “Hardness of vertex deletion and project scheduling”, In Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, volume 7408 of Lecture Notes in Computer Science, pages 301–312. 2012.
32. J. Swafford and G. Whitney, “Nonparametric test of utility maximization and weak separability for consumption, leisure and money”, *Review of Economics and Statistics*, **69**, pp458-464, 1987.
33. H. Varian, “Revealed preference”, in M. Szenberg, L. Ramrattand and A. Gottesman (eds.), *Samulesonian Economics and the 21st Century*, pp99–115, Oxford University Press, 2005.
34. H. Varian, “Goodness-of-fit in optimizing models”, *Journal of Econometrics*, **46**, pp125-140, 1990.
35. D. Wang and Y. Kuo, “A study on two geometric location problems”, *Information Processing Letters*, **28(6)**, pp281–286, 1988.