Introduction

by Alan D. Taylor

Notions of obvious breadth and importance should, when possible, be examined under a number of different disciplinary lenses. This is the spirit in which the present offering by Julius Barbanel (a mathematician) joins recent books by Hervé Moulin (an economist) [32] and Nicholas Rescher (a philosopher) [35]. But fairness – or, more explicitly, fair division – comes in a number of different flavors, and we should begin by setting forth something of a general framework in which we can place the present book.

One of the more important dichotomies in the treatment of fairness is the extent to which the treatment is normative. Is the author trying to argue that certain methods of allocation are superior to others? The treatment of fair division by economists, philosophers, and political scientists tends to lie in the normative camp. Mathematicians, on the other hand, focus on what is possible and what is not, and often leave subjective judgments to others, as Barbanel does here.

Yet there is a normative aspect of the present work that sets it apart from the great majority of mathematical treatments, and it is revealed in Barbanel's choice of title. The work is not called "The Geometry of Fair Division" but "The Geometry of Efficient Fair Division." Efficiency – also called Pareto optimality, after the nineteenth-century Italian scholar Vilfredo Pareto – is, according to Hervé Moulin, "the single most important tool of normative economics" [32, pg. 8].

Economists also tend to focus (although not exclusively) on issues of fairness in the context of a finite collection of divisible homogeneous goods. Mathematicians, however, far more often work with a single divisible heterogeneous good and typically phrase the discussion in terms of the cake-cutting metaphor that dates back to the seventeen-century English political theorist James Harrington [26].

Discussions of cake cutting almost always begin with the procedure known as divide-and-choose. Historically, this two-person scheme traces its origins back 5000 years to the Bible's account of land division between Abram (later to be called Abraham) and Lot, and it resurfaces more explicitly two-and-a-half millennia ago as Hesiod, in his *Theogony*, describes the division of meat into two piles by Prometheus, with Zeus then choosing the pile that he preferred.

But even narrowed to the context of a mathematician's non-normative treatment of cake cutting, there is an important second dichotomy that sets the present work apart from earlier efforts such as those by Steven Brams and myself [16] and by Jack Robertson and William Webb [36]. This dichotomy is, in a sense, one of process versus product. Is one searching for a constructive procedure – a process – that will, in a step-by-step fashion, lead to desirable allocations, or is one trying to establish, by any mathematical means possible, the very existence of the desired allocation itself – the product?

The Brams–Taylor book and the Roberston–Webb book both focus on constructive procedures. The present work, on the other hand, is the first one on fair division to sit squarely in the existence camp. Yet economists will find it remarkably accessible – and an absolute gem in terms of illustrating how much insight the hands of an expert can wring from a couple of abstract results.

This distinction between constructive procedures and existence results is also reflected in the assumptions made in formalizing the preferences of the various participants in a fair-division situation. But in order to illustrate these differences, we need a few procedures on which to hang such a discussion. A quick historical tour will provide what is required.

Mathematical investigations of fair division date from the early 1940s. The constructive vein was first opened by the Polish mathematician Hugo Steinhaus (see [40]) and his colleagues Stefan Banach and Bronislaw Knaster. Steinhaus appears to have been the first to ask if there is an obvious extension of divide-and-choose to the case wherein there are three participants instead of two, and he derived the scheme referred to in a number of mathematical texts for non-majors (see [18] and [42]) as "the lone-divider method." But extending this procedure to four or more participants is somewhat complicated, and was not actually achieved until Harold Kuhn [30] did so in 1967. Banach and Knaster, however, took an entirely different tack and devised a fair-division scheme for any number of participants that is known today as the "last-diminisher method."

Each of these schemes generalizes divide-and-choose in the sense of providing a finite constructive procedure by which a group of people can allocate a "cake" among themselves in such a way that each has a strategy that ensures his or her own "satisfaction" even in the face of a conspiracy by all of the others. The word "protocol" is often used to capture both the algorithmic and the strategic aspects of such procedures, and this game-theoretic view results in the use of "player" in place of "participant."

Yet it turns out that the devil is in the details. "Satisfied" in what sense? For the procedures of Steinhaus, Banach, and Knaster, the answer is something called "proportionality" – each of n players is assured of receiving a piece that he or she thinks is at least $\frac{1}{n}$th of the total in size or value. Divide-and-choose is obviously proportional: if the divider makes it a 50–50 division, he or she will get exactly one-half; the chooser can't go wrong. Proportionality, however, is only the easy answer.

In 1959, the physicist George Gamow and the mathematician Marvin Stern published a book [24] in which they pointed out that with divide-and-choose, each of the two players is assured of getting a piece that he or she thinks is at least tied for largest (or tied for most valuable). They asked if there were procedures that would do the same for three or more players. The name attached to such allocations today is "envy-free" or "no-envy," a notion that economists trace back to Duncan Foley [22] in 1967. Envy-freeness is harder to come by than proportionality, although the existence results we turn to momentarily show that much more is, in some sense, possible.

Within a year of the Gamow–Stern question, John Conway of Princeton and John Selfridge of Northern Iowa University independently constructed an elegant envy-free protocol for three parties (see [16]), although the general question for four or more parties remained open until it was settled in the affirmative in 1992 [15]. There is, however, an important issue that still awaits attention: The three-person scheme never requires more than five cuts, whereas the general procedure, even if there are only four players, has no upper bound on the number of cuts needed that is independent of the preferences of the people involved.

But how do we formalize these "preferences" of the players, and what kind of an object do we take this "cake" to be? If we begin with the most general context that suggests itself, the "cake" C would be an arbitrary set and each player's preferences over (certain) subsets of C would be given by a binary relation R that is reflexive, transitive, and complete (with XRY intuitively meaning that this player finds the subset X to be at least as desirable as the subset Y). And, as first pointed out by David Gale [23], discrete cake-cutting protocols implicitly assume only three additional postulates: (i) a partitioning postulate, asserting that a player can divide a piece of cake into any number of smaller pieces that he or she considers equivalent to each other, (ii) a trimming postulate asserting that if a player prefers one piece of cake to another, then there is a subset of the former that he or she considers equivalent to the latter, and (iii) a weak-additivity postulate asserting that if a player prefers piece 1 to piece 2, and piece 3 to piece 4, and pieces 1 and 3 are disjoint, then that player will prefer the union of pieces 1 and 3 to the union of pieces 2 and 4.

The easiest way to obtain such a relation is to let Player i's preferences be given by a finitely additive, non-atomic probability measure over some algebra of subsets of the arbitrary set C. That is, one starts with a collection of subsets of C that is closed under complementation, finite unions, and finite intersections – this is what an algebra is – and a function μ that assigns a real number in the interval $[0, 1]$ to each set in the algebra so that if $A_1, \ldots, A_n$ is a finite collection of pairwise disjoint sets in the algebra, then $\mu(A_1 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n)$ – this is finite additivity – and such that, if $\mu(A) > 0$, then there is some $B \subseteq A$ such that $0 < \mu(B) < \mu(A)$ – this is what it means to be non-atomic.

In point of fact, there is only one difference between working in the general context of a preference relation satisfying Gale's three postulates and working with a finitely additive, non-atomic probability measure: the latter ensures that the players' preferences satisfy an Archimedean property asserting that, if a subset of C is strictly preferred to the empty set, then the entire cake C can be partitioned into finitely many pieces, all of which are less desirable than the given piece. This is the only difference in the sense that one can prove [10] that any Archimedean preference relation satisfying Gale's three postulates is induced by a finitely additive, non-atomic probability measure.

Protocols – or, more generally, all cake-division schemes with a legitimate claim to being finite and constructive – work in the context of finitely additive, non-atomic probability measures. Existence results, on the other hand, both assume more and deliver more.

Historically, the first existence result to explicitly address fair division may have been Jerzy Neyman's 1946 result [34] asserting that a cake can be divided among n players in such a way that every player thinks every piece is $\frac{1}{n}$th of the total. This theorem assumes, as do virtually all of what are called "existence results" in this context, that the players' preferences are given by non-atomic probability measures that are not only finitely additive, but countably additive: If $A_1, A_2, \ldots$ is a collection of pairwise disjoint sets in the algebra indexed by the set of natural numbers, then $\mu(A_1 \cup A_2 \cup \cdots) = \mu(A_1) + \mu(A_2) + \cdots$.

There are stepping stones between protocols and existence results that deserve mention. These are the so-called "moving-knife schemes" that date back to the observation of Lester Dubins and E. H. Spanier [20] that the Banach–Knaster scheme can be envisioned as one in which a knife is slowly moved across the cake, with each player having the option to call "cut" at any time and to exit the game with the resulting piece. A moving-knife alternative to the three-player envy-free Selfridge–Conway procedure was found by Walter Stromquist [41] in 1980, and, in 1982, A. K. Austin [3] introduced a moving-knife version of the $n = 2$ case of Neyman's theorem. A number of questions in the context of moving-knife schemes remain open (see [17] and [9]). The reader seeking an additional challenge can try to extend to the moving-knife arena the myriad of results set forth by Barbanel in what follows.

So now we have the context: Barbanel is giving a non-normative, mathematical treatment of existence results that deal with efficiency as well as fairness, in the context of a single heterogeneous good with the preferences of players given by countably additive, non-atomic, probability measures. All that remains is to address the question of how geometry enters the picture.

Geometry is the study of size and shape. Thus, one might expect Barbanel to study the size and shape of, well, the cake (or at least pieces thereof). But that's not at all what he does. His study of the geometry of fair division is much more in the spirit of Donald Saari's study of the geometry of voting [39]. For Saari, a ballot in an election corresponds to a point in n-space. For Barbanel, an allocation of the cake corresponds to a point of n-space in one of the two main geometric objects considered. In the other, each point of the cake corresponds to a point in n-space, but in a non-obvious manner. Either way, once he has a set of points in n-space, he is geometrically off and running.

The book is laid out as follows. After introducing some basic notation, terminology, and background in Chapter 1, Barbanel defines the first geometric object on which he focuses: the Individual Pieces Set (IPS). He introduces the IPS for two players in Chapter 2 and then exploits it in the context of fairness and efficiency in Chapter 3.

In Chapter 4, Barbanel moves on to the general case of n players, where he generalizes the IPS to the FIPS, the Full Individual Pieces Set, and he proves an important result concerning the possible shapes of the FIPS. In Chapter 5, he considers what the IPS and FIPS reveal about fairness and efficiency in the general n-player context.

Barbanel next focuses exclusively on efficiency, and he presents three quite different characterizations of Pareto optimality. After some introductory notions in Chapter 6, he characterizes Pareto optimality using the optimization of convex combinations of measures (Chapter 7) and partition ratios (Chapter 8). In Chapter 9, Barbanel introduces the second of his two main geometric objects: the Radon–Nikodym Set (RNS), and he uses it, together with an idea of Dietrich Weller, to present a third characterization of Pareto optimality in Chapter 10.

In Chapter 11, Barbanel considers the possible shapes of the IPS, and he provides a complete characterization in the case of two players and a partial result in the general n-player context. In Chapters 12 and 13, he studies the relationship between the IPS and the RNS, and he provides a new presentation of the fundamental result that ensures the existence of a partition that is both Pareto optimal and envy-free.

In Chapter 14, Barbanel introduces a strengthening of Pareto optimality that he calls “strong Pareto optimality,” and he presents both characterization theorems and existence results. He also discusses the relationships between the number of strongly Pareto optimal partitions and the number of Pareto optimal partitions that are not strongly Pareto optimal.

Barbanel’s characterizations of Pareto optimality in Chapters 7 and 10 involve what is essentially an iterative procedure. In Chapter 15, he shows that these ideas can be greatly simplified by the use of hyperreal numbers and non-standard analysis.

Finally, in Chapter 16, Barbanel shows that the IPS can be viewed as a piece of a larger structure that he calls the Multicake Individual Pieces Set (MIPS). Earlier chapters reveal certain peculiarly non-symmetric possibilities for the IPS; symmetry reasserts itself in the MIPS.

1

Notation and Preliminaries

Our "cake" C is some set. We wish to partition C among n players, whom we shall refer to as Player 1, Player 2, ..., Player n. For each $i = 1, 2, \ldots, n$, Player i uses a measure m_i to evaluate the size of pieces of cake (i.e., subsets of C). Unless otherwise noted, we shall always assume that C is non-empty.

Definition 1.1 A σ-*algebra* on C is a collection of subsets W of C satisfying that

a. $C \in W$,
b. if $A \in W$ then $C \backslash A \in W$, and
c. if $A_i \in W$ for every $i \in \mathbf{N}$, then $(\bigcup_{i \in \mathbf{N}} A_i) \in W$ (where $\mathbf{N}$ denotes the set of natural numbers).

Definition 1.2 Assume that some σ-algebra W has been defined on C. A *countably additive measure* on W is a function $\mu : W \to \mathbf{R}$ (where $\mathbf{R}$ denotes the set of real numbers) satisfying that

a. $\mu(A) \geq 0$ for every $A \in W$,
b. $\mu(\emptyset) = 0$, and
c. if $A_1, A_2, \ldots$ is a countable collection of elements of W and this collection is pairwise disjoint, then $\mu(\bigcup_{i \in \mathbf{N}} A_i) = \sum_{i \in \mathbf{N}} \mu(A_i)$.

In addition, μ is

d. *non-atomic* if and only if, for any $A \in W$, if $\mu(A) > 0$ then for some $B \subseteq A$, $B \in W$ and $0 < \mu(B) < \mu(A)$ and
e. a *probability measure* if and only if $\mu(C) = 1$.

Unless otherwise noted, all measures that we shall consider will be countably additive, non-atomic probability measures, and we shall simply use the term "measure" to refer to them. Notice that for any measure μ and any $a \in C$, the non-atomic nature of μ implies that $\mu(a) = 0$.

Also, unless otherwise specified, C shall denote an arbitrary cake. We assume that there are n players, Player 1, Player 2, ..., Player n, with corresponding measures $m_1, m_2, \ldots, m_n$, respectively. At times, we shall work with specific examples and shall give specific definitions of C and $m_1, m_2, \ldots, m_n$.

Whenever a subset of C is mentioned, we assume it is a member of some common σ-algebra on which all of the measures are defined. We shall never explicitly define a specific σ-algebra.

We will be concerned with partitions of the cake C among the players. When we consider an ordered partition $\langle P_1, P_2, \ldots, P_n\rangle$ of C, our intention is that P_1 goes to Player 1, P_2 goes to Player 2, etc. The term "partition" always means "ordered partition." *Part* denotes the set of all partitions of the appropriated size, which will always be clear by context.

Consider the set $\{(m_1(A), m_2(A), \ldots, m_n(A)) : A \subseteq C\}$, which is a subset of $\mathbf{R}^n$. This set will be important for us. A central tool concerning this set is Lyapounov's theorem.

Theorem 1.3 (Lyapounov's Theorem [31]) $\{(m_1(A), m_2(A), \ldots, m_n(A)) : A \subseteq C\}$ *is a closed and convex subset of* $\mathbf{R}^n$.

Another important set is $\{[m_i(P_j)]_{i,j\le n} : \langle P_1, P_2, \ldots, P_n\rangle$ is a partition of $C\}$. This is a subset of the set of all $n \times n$ matrices and can be viewed as a subset of $\mathbf{R}^{(\mathbf{n}^2)}$. An element of this set gives each player's evaluation of the size of each piece of cake in a given partition. A central tool concerning this set is Dvoretsky, Wald, and Wolfovitz's theorem.

Theorem 1.4 (Dvoretsky, Wald, and Wolfovitz's Theorem [21]) $\{[m_i(P_j)]_{i,j\le n} : \langle P_1, P_2, \ldots, P_n\rangle$ *is a partition of* $C\}$ *is a closed and convex subset of the set of all* $n \times n$ *matrices.*

(Dvoretsky, Wald, and Wolfovitz's theorem actually is more general than the preceding statement. The number of players need not equal the number of pieces of the partition, and thus the set under consideration is $\{[m_i(P_j)]_{i\le m; j\le n} : \langle P_1, P_2, \ldots, P_n\rangle$ is a partition of $C\}$. The theorem says that this set is a closed and compact subset of the set of all m × n matrices. We shall always have the number of players equal to the number of pieces of partitions, and so we have stated the theorem in this more restricted form.)

Notice that $\{(m_1(A), m_2(A), \ldots, m_n(A)) : A \subseteq C\}$ is the set of all first (or second, or third, etc.) columns of $\{[m_i(P_j)]_{i,j\le n} : \langle P_1, P_2, \ldots, P_n\rangle$ is a partition of $C\}$. This tells us that Lyapounov's theorem follows immediately from Dvoretsky, Wald, and Wolfovitz's theorem.

We shall frequently need to find subsets of C having certain sizes on which all players agree. The following corollary to Lyapounov's theorem will often provide exactly what we need.

Corollary 1.5 *Fix non-negative real numbers $p_1, p_2, \ldots, p_n$ such that $p_1 + p_2 + \cdots + p_n = 1$. There is a partition $P = \langle P_1, P_2, \ldots, P_n \rangle$ of C such that for all $i, j = 1, 2, \ldots, n,\ m_i(P_j) = p_j$.*

Proof: Fix $p_1, p_2, \ldots, p_n$ as in the statement of the corollary and let $G = \{[m_i(P_j)]_{i,j \leq n} : \langle P_1, P_2, \ldots, P_n \rangle$ is a partition of $C\}$. For each $i = 1, 2, \ldots, n$, let M_i be the matrix with all ones in column i and zeros everywhere else. Then, by considering the partitions $\langle C, \emptyset, \emptyset, \ldots, \emptyset, \emptyset \rangle$, $\langle \emptyset, C, \emptyset, \ldots, \emptyset, \emptyset \rangle, \ldots, \langle \emptyset, \emptyset, \emptyset, \ldots, \emptyset, C \rangle$, we see that each M_i is in G. By Dvoretsky, Wald, and Wolfovitz's theorem, G is convex and hence $p_1 M_1 + p_2 M_2 + \cdots + p_n M_n \in G$. But $p_1 M_1 + p_2 M_2 + \cdots + p_n M_n$ is the matrix with every entry in the first column equal to p_1, every entry in the second column equal to p_2, etc. This implies that there is a partition $P = \langle P_1, P_2, \ldots, P_n \rangle$ of C such that for all $i, j = 1, 2, \ldots, n,\ m_i(P_j) = p_j$, as desired. □

Corollary 1.5 has many simple applications. Two are given by the following two corollaries.

Corollary 1.6 *For any $A \subseteq C$ and non-negative real numbers $q_1, q_2, \ldots, q_n$ with $q_1 + q_2 + \cdots + q_n = 1$, there is a partition $Q = \langle Q_1, Q_2, \ldots, Q_n \rangle$ of A such that for all $i, j = 1, 2, \ldots, n$, $m_i(Q_j) = q_j m_i(A)$.*

Proof: Fix A and $q_1, q_2, \ldots, q_n$ as in the statement of the corollary and let $\delta = \{i \leq n : m_i(A) > 0\}$. For each $i \in \delta$, we define m'_i on A as follows:

$$\text{for each } B \subseteq A,\ m'_i(B) = \frac{m_i(B)}{m_i(A)}$$

Each such m'_i is a measure on A. For each $i \notin \delta$, let m'_i be any measure on A.

It follows from Corollary 1.5, with A playing the role of C, that there is a partition $Q = \langle Q_1, Q_2, \ldots, Q_n \rangle$ of A satisfying that $m'_i(Q_j) = q_j$ for all $i, j = 1, 2, \ldots, n$. We claim that for all $i, j = 1, 2, \ldots, n$, $m_i(Q_j) = q_j m_i(A)$. Fix such an i and j. We consider two cases.

Case 1: $i \in \delta$. Then $m_i(Q_j) = m'_i(Q_j) m_i(A) = q_j m_i(A)$.
Case 2: $i \notin \delta$. Then $m_i(A) = 0$ and hence, since $Q_j \subseteq A$, $m_i(Q_j) = 0$.

Therefore, $m_i(Q_j) = 0 = (q_j)(0) = q_j m_i(A)$.

This establishes that for all $i, j = 1, 2, \ldots, n,\ m_i(Q_j) = q_j m_i(A)$, as desired. □

Corollary 1.7 *Fix some* $A \subseteq C$ *and* $k = 1, 2, \ldots, n$. *If* $m_k(A) > 0$, *then for any* r *with* $0 \leq r \leq m_k(A)$, *there is a* $B \subseteq A$ *with* $m_k(B) = r$.

Proof: Let A and k be as in the statement of the corollary, assume that $m_k(A) > 0$, and fix some r with $0 \leq r \leq m_k(A)$. Set $q_k = \frac{r}{m_k(A)}$. Then $0 \leq q_k \leq 1$. For each $i = 1, 2, \ldots, n$ with $i \neq k$, let q_i be an arbitrary non-negative real number, subject to the condition that $q_1 + q_2 + \cdots + q_n = 1$. By Corollary 1.6, there is a partition $Q = \langle Q_1, Q_2, \ldots, Q_n \rangle$ of A such that for all $i, j = 1, 2, \ldots, n$, $m_i(Q_j) = q_j m_i(A)$. Set $B = Q_k$. Then $m_k(B) = m_k(Q_k) = q_k m_k(A) = r$, as desired. □

We will be interested in what it means for a partition of C to be a "good" partition. Various notions of what good means in this context have been considered. These notions are of two types. One is concerned with fairness and the other with efficiency. Before we can define fairness and efficiency properties, we must first consider a more basic question: Do players want as much cake as possible or do they want as little cake as possible? For example, if the cake represents money to be distributed among the players, then it is reasonable to assume that each player wants as much of the cake as possible. On the other hand, if the cake represents some task that all players view as unpleasant, then each player wants as little of the cake as possible. We shall refer to the first setting, in which "bigger is better," as the *standard setting*, and shall refer to the latter setting, in which "smaller is better," as the *chores setting* (since, in this case, pieces of cake may be viewed as "chores"). Unless otherwise noted, we shall assume that we are working in the standard setting. Our approach for most sections is to first concentrate on the standard setting and then on the chores setting. (However, there will be some sections where we find it most convenient to consider the standard setting and the chores setting at the same time.) Most of the time, results about the chores setting will simply be symmetric adjustments of results about the standard setting. However, there will be important exceptions.

What does it mean to say that a partition of the cake is fair? We shall say that a partition is fair if and only if every player thinks that it is fair, and so the question becomes: When does a player think that a partition is fair? Consider the following five answers for the standard setting. A player thinks that a partition is fair if and only if that player thinks that his or her piece of cake is

a. at least of average size.
b. of bigger-than-average size.
c. at least as big as every other piece.