Zero-Sum Games and Linear Programming Duality

Bernhard von Stengel

Department of Mathematics, London School of Economics, b.von-stengel@lse.ac.uk

July 26, 2023

Abstract

The minimax theorem for zero-sum games is easily proved from the strong duality theorem of linear programming. For the converse direction, the standard proof by Dantzig (1951) is known to be incomplete. We explain and combine classical theorems about solving linear equations with nonnegative variables to give a correct alternative proof, more directly than Adler (2013). We also extend Dantzig's game so that any max-min strategy gives either an optimal LP solution or shows that none exists.

1 Introduction and summary

LP duality (the strong duality theorem of linear programming) is a central result in optimization. It helps proving many results with ease, such as the minimax theorem for zero-sum games, first proved by von Neumann in 1928 [27]. In October 1947, George Dantzig explained his nascent ideas on linear programming to John von Neumann [9, p. 45]. In response, he got an "eye-popping" lecture on LP duality, which von Neumann conjectured to be equivalent to his minimax theorem. This "equivalence" is commonly assumed (for example, Schrijver [22, p. 218]), but on closer inspection does not hold at all.

"Equivalence" is actually not a good term – all theorems, as logical statements without free variables, are equivalent, to "true". We therefore say that theorem A *proves* (rather than "implies") theorem B, typically by a suitable but different use of the variables in theorem A, and state straightforward proof relations of this kind as propositions (see Proposition 1 for an example).

The classic proof by Dantzig [7] of LP duality from the minimax theorem needs an additional assumption about the game solution, namely strict complementarity in the last column of the game matrix that corresponds to the right-hand side of the LPs. (We state Dantzig's game in (35) below; it differs from the original in a trivial change of signs so that the primal LP is a maximization problem subject to upper bounds, in line with the row player in a zero-sum game as the maximizer.) This complementarity assumption, acknowledged by Dantzig [7][8, p. 291], applies only to non-generic LPs and seems technical. Adler [1] fixed this "hole" in Dantzig's proof, and showed how an algorithm that solves a zero-sum game can be used to either solve an LP or certify that it has no optimal solution. Recently, Brooks and Reny [3] gave a zero-sum game whose solution directly provides such a solution or certificate.

The aim of this article is to clarify the underlying problem, with two new main results (explained later). Our narrative is self-contained, not least because LP duality is so familiar that it can be overlooked as a silent assumption. For example, reducing optimality of maximizing $c^\top x$ subject to $Ax \leq b$, $x \geq \mathbf{0}$ to feasibility of $Ax \leq b$, $x \geq \mathbf{0}$, $A^\top y \geq c$, $y \geq \mathbf{0}$, $b^\top y \leq c^\top x$ assumes that there cannot be a positive "duality gap" $b^\top y - c^\top x$, which is the strong duality theorem. Our presentation shows how one could prove, in full, LP duality via the minimax theorem, if one were to take that route. Some of the presented less-known elegant proofs from the literature are also of historical interest.

Dantzig's assumption holds if a pure strategy that is a best response in every solution of the zero-sum game has positive probability in some solution. As noted by Adler [1, p. 167], this can be shown (e.g., [21, p. 742]) using a version of the Lemma of Farkas [10]. However, the Lemma of Farkas proves LP duality directly. Our first, easy observation is that Dantzig's assumption amounts to the Lemma of Tucker [25]. This, in turn, directly proves the Lemma of Farkas [25, p. 7], even for the special case of Dantzig's game (Proposition 6 below). The assumption is therefore extremely strong and in a sense useless for proving LP duality from the minimax theorem. Curiously, Tucker did not consider the converse that in nearly the same way the Lemma of Farkas proves his Lemma (see Proposition 8 below). This suggests that Tucker thought he had proved a more general statement. Tucker's proof of his Lemma is indeed short and novel, but in this light we agree with Adler's view of Tucker's Lemma as a "variant of Farkas's Lemma" [1, p. 174].

LP duality and the minimax theorem are closely related to solving, respectively, inhomogeneous and homogeneous linear equations in nonnegative variables. The Lemma of Farkas characterizes when the inhomogeneous linear equations $Ax = b$ have no solution vector x such that $x \geq \mathbf{0}$. The Theorem of *Gordan* [14] characterizes when the homogeneous equations $Ax = \mathbf{0}$ have no solution $x \geq \mathbf{0}$ other than the trivial one $x = \mathbf{0}$. Gordan's Theorem and its "inequality version" due to Ville [26] prove the minimax theorem and vice versa.

Our first main result, Theorem 6 in Section 7, is a proper proof of LP duality from the minimax theorem. Inspired by Adler [1, section 4], we use Gordan's Theorem to prove the *Theorem* of Tucker [25], an easy but powerful generalization of his Lemma (like Broyden [4] we think that it deserves more recognition). Tucker's

Theorem shows that any system of homogeneous equations $Ax = \mathbf{0}$ such that $x \geq \mathbf{0}$ has a natural partition of its solution vector x into a set of variables that can take positive values and the others that are zero in any nonnegative solution. It is easy to see that one can drop the nonnegativity requirement for the variables that can be positive. By *eliminating* these unconstrained variables from the system $Ax = \mathbf{0}$ with a bit of linear algebra, applying Gordan's Theorem to the variables that are always zero in any nonnegative solution then gives Tucker's Theorem. Compared to the detailed computations of this variable elimination by Adler [1], our proof is self-contained and more direct. Using Dantzig's game (35), Tucker's theorem proves LP duality in a stronger version, namely the existence of a "strictly complementary" solution to the LPs if they are feasible (Proposition 10 below).

Our second main result, Theorem 7 in Section 8, extends Dantzig's elegant game (35) with an extra row in (51) that "enforces" the desired complementarity in the last column. *Every* max-min strategy of this game either gives an optimal pair of solutions to the primal and dual LPs, or represents an unbounded ray for at least one of the LPs if it is feasible, so that the other LP is therefore infeasible. This result is similar to Adler's "Karp-type" reduction of an LP to a zero-sum game [1, section 3.1], but with the extra certificate of infeasibility. It is also similar to, and inspired by, the main result of Brooks and Reny [3]. The proof of Theorem 7 (in a separate Theorem 8) does *not* rely on LP duality and was surprisingly hard to find. Compared to either [1] or [3], our game (51) more naturally extends Dantzig's original game. Similar to both, it imposes an upper bound on the LP variables that does not affect whether the LPs are feasible. This bound follows from Carathéodory's theorem [5] that nonnegative solutions x to $Ax = b$ can be found using only linearly independent columns of A (of which there are only finitely many sets). That bound is determined apriori and of polynomial encoding size from the sizes of the entries of A and b if these are integer or algebraic numbers, otherwise abstractly from all "basic solutions" x to $Ax = b$.

We give a self-contained introduction to linear programming duality (for LPs in inequality form) and to the minimax theorem in Section 2. Section 3 recalls how LP duality is proved from the Lemma of Farkas. The theorems of Gordan [14] and Ville [26] are the topic of Section 4. Stiemke [23] gave a two-page proof of the Theorem of Gordan (without referencing it, even though published in the same journal, presumably with no editor around to remember it). His proof uses implicitly that the null space and row space of a matrix are orthogonal complements. But there are no matrices in these papers – people manipulated linear equations with their unknowns instead. For historical interest, and because of its structural similarity to Tucker's proof of his Lemma [25, p. 5–7], we reproduce Stiemke's proof in Section 5. We also present a most elegant half-page proof of the minimax theorem due to Loomis [17], which then leads to Gordan's Theorem as an easy additional step. As we explain at the end of Section 5, it seems difficult to extend the proof by Loomis to proving LP duality directly, which was the original aim of this research.

Section 6 presents the classic derivation of LP duality from the minimax theorem due to Dantzig [7]. Even though its additional assumption looks minor, we show that it amounts to the Lemma of Tucker [25], which, as noted by Tucker [25, p. 7], proves the Lemma of Farkas. This shows that the assumption is way too strong to make Dantzig's derivation useful.

Section 7 proves Tucker's Theorem and thus LP duality from the minimax theorem using Gordan's theorem. As mentioned, this is distilled from Adler [1, section 4]. In Section 8, we add another row to Dantzig's game to obtain a new game where every max-min strategy either gives a solution to the LP or a certificate that no optimal solution exists. Theorems 6 and 7 in Sections 7 and 8 are the main results of this paper.

Section 9 gives a detailed comparison of our work with the closely related papers by Adler [1] and Brooks and Reny [3]

In the final Section 10 we present a little-known gem of a proof of the Lemma of Farkas due to Conforti, Di Summa, and Zambelli [6]. Their theorem states that a system of inequalities $Ax \le b$ is *minimally* infeasible if and only if the corresponding *equalities* $Ax = b$ are minimally infeasible. Because the linear equations are infeasible, a suitable linear combination of them states $0 = -1$, which proves the Lemma of Farkas in this context.

2 LP duality and the minimax theorem

Throughout, m and n are positive integers, and $[n] = \{1, \ldots, n\}$. All vectors are column vectors. The jth component of a vector x is written x_j. All matrices have real entries. The transpose of a matrix A is written $A^\top$. Vectors and scalars are treated as matrices of appropriate dimension, so that a vector x times a scalar α is written as $x\alpha$, and a row vector $x^\top$ times a scalar α as $\alpha x^\top$. The matrix A with all entries multiplied by the scalar α is written as αA. We usually transpose vectors rather than the matrix, to emphasize that Ax is a linear combination of the columns of A and $y^\top A$ is a linear combination of the rows of A. The all-zero and the all-one vector are written as $\mathbf{0} = (0, \ldots, 0)^\top$ and $\mathbf{1} = (1, \ldots, 1)^\top$, their dimension depending on the context, and the all-zero matrix just as 0. Inequalities between vectors or matrices such as $x \ge \mathbf{0}$ hold between all components.

A linear program (LP) in *inequality form* is given by an $m \times n$ matrix A and vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ and states, with a vector $x \in \mathbb{R}^n$ of variables:

$$\underset{x}{\text{maximize}}\ c^\top x \quad \text{subject to} \quad Ax \le b, \quad x \ge \mathbf{0}. \tag{1}$$

This LP is called *feasible* if there is some $x \in \mathbb{R}^n$ that fulfills the constraints $Ax \le b$ and $x \ge \mathbf{0}$, otherwise *infeasible*. If there are arbitrarily large values of $c^\top x$ with $Ax \le b$ and $x \ge \mathbf{0}$, then the LP is called *unbounded*.

With (1) considered as the *primal* LP, its *dual LP* states, with a vector $y \in \mathbb{R}^m$ of variables:

$$\underset{y}{\text{minimize}} \;\; y^\top b \quad \text{subject to} \quad y^\top A \ge c^\top, \quad y \ge \mathbf{0}, \tag{2}$$

with feasibility and unboundedness defined accordingly. An equivalent way of writing the dual constraints in (2) is $A^\top y \ge c$, which transposes only the matrix and can be more readable.

The *weak duality* theorem states that if both primal and dual LP have feasible solutions x and y, respectively, then their objective function values are mutual bounds, that is,

$$c^\top x \le y^\top b, \tag{3}$$

which holds because feasibility implies $c^\top x \le y^\top Ax \le y^\top b$. Hence, if there are feasible solutions x and y so that the two objective functions are *equal*, $c^\top x = y^\top b$, then both are optimal. The (strong) *LP duality* theorem states that this is always the case if the two LPs are feasible:

Theorem 1 (LP duality)**.** *If the primal LP* (1) *and the dual LP* (2) *are feasible, then there exist feasible x and y with $c^\top x = y^\top b$, which are therefore optimal solutions.*

A *zero-sum game* is given by an $m \times n$ matrix A and is played between a *row player*, who chooses a row i of the matrix, simultaneously with the *column player*, who chooses a column j of the matrix, after which the row player receives the matrix entry a_{ij} from the column player as a *payoff* (which is a *cost* to the column player). That is, the row player is the maximizer and the column player the minimizer. The rows and columns are called the players' *pure strategies*.

The players can *randomize* their actions by choosing them according to a probability distribution, called a *mixed strategy*. The other player may know the probability distribution but not the chosen pure strategy. The row player is then assumed to maximize his *expected payoff* and the column player to minimize her *expected cost*. We denote the set of mixed strategies of the row player by

$$Y = \{y \in \mathbb{R}^m \mid y \ge \mathbf{0},\ \mathbf{1}^\top y = 1\}, \tag{4}$$

and of the column player by

$$X = \{x \in \mathbb{R}^n \mid x \ge \mathbf{0},\ \mathbf{1}^\top x = 1\}, \tag{5}$$

in order to stay close to the LP notation (normally row and column player are considered as first and second player, respectively, so that the letters for their mixed strategies should be in alphabetical order, but this is already violated with the very common naming of the LP variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$).

With mixed strategies y and x of row and column player, the expected payoff to the maximizing row player and expected cost to the minimizing column player is $y^\top Ax$.

The minimizing column player who chooses a mixed strategy x should expect that the row player responds with a mixed strategy y (called a *best response*) that maximizes her payoff $y^\top Ax$. That best-response payoff $\max_{y\in Y} y^\top Ax$ is the weighted sum $\sum_{i\in[m]} y_i(Ax)_i$ of the expected payoffs $(Ax)_i$ for the rows i and therefore equal to their maximum, which in turn is the least upper bound v of these row payoffs. That is,

$$\max_{y\in Y} y^\top Ax \;=\; \max_{i\in[m]} (Ax)_i \;=\; \min\{v\in\mathbb{R} \mid Ax\le \mathbf{1}v\}\,. \tag{6}$$

A *min-max* strategy x of the column player minimizes this worst-case cost v that he has to pay, that is, it is an optimal solution to

$$\underset{x,v}{\text{minimize}}\; v \quad \text{subject to} \quad Ax\le\mathbf{1}v, \quad x\in X \tag{7}$$

and then v is called the *min-max value* of the game.

Similarly, a *max-min* strategy y and the *max-min value* u is an optimal solution to

$$\underset{y,u}{\text{maximize}}\; u \quad \text{subject to} \quad y^\top A\ge u\mathbf{1}^\top, \quad y\in Y. \tag{8}$$

The minimax theorem of von Neumann [27] states

$$\max_{y\in Y}\,\min_{x\in X}\, y^\top Ax \;=\; v \;=\; \min_{x\in X}\,\max_{y\in Y}\, y^\top Ax \tag{9}$$

where the unique real number v is called the *value* of the game. Via (6) and the corresponding expression for $\min_{x\in X} y^\top Ax$ (the best-response cost to $y\in Y$), we state this as follows.

Theorem 2 (The minimax theorem)**.** *Consider optimal x,v for (7) and y,u for (8). Then $u=v$ (the value of the game), x is a min-max strategy, and y is a max-min strategy.*

The LP (7) is in *general form* with an equation $\mathbf{1}^\top x=1$ and an unconstrained variable v (with $-v$ to be maximized), and so is (8), which is the dual LP to (7) with u as the unconstrained variable (with $-u$ to be minimized) that corresponds to the equation for X written as $-\mathbf{1}^\top x=-1$. Since both LPs are feasible, the strong duality theorem (which also holds for LPs in general form) implies that their optimal values are equal ($-v=-u$), which proves Theorem 2.

One can avoid stating LPs in general form by ensuring that the min-max value is positive, by adding a constant α to the payoffs a_{ij}, which defines a new payoff matrix $A+\mathbf{1}\alpha\mathbf{1}^\top$. Then for $y\in Y$ and $x\in X$

$$y^\top(A+\mathbf{1}\alpha\mathbf{1}^\top)x \;=\; y^\top Ax+y^\top\mathbf{1}\alpha\mathbf{1}^\top x \;=\; y^\top Ax+\alpha\,, \tag{10}$$

which shows that best responses and min-max and max-min strategies are unaffected and the corresponding values just shifted by α. If all entries of A are positive, then $v > 0$ for any feasible v in (7). Division of each row in (7) by v (where we now maximize $1/v$) then gives the LP

$$\underset{x}{\text{maximize}}\ \mathbf{1}^\top x \quad \text{subject to} \quad Ax \le \mathbf{1}, \quad x \ge \mathbf{0} \tag{11}$$

with its dual

$$\underset{y}{\text{minimize}}\ y^\top \mathbf{1} \quad \text{subject to} \quad y^\top A \ge \mathbf{1}^\top, \quad y \ge \mathbf{0}\,. \tag{12}$$

Both LPs are feasible with nonzero optimal solutions x and y, which give the min-max and max-min strategies xv and yv with $v = 1/\mathbf{1}^\top x = 1/\mathbf{1}^\top y$ and game value v.

These are the standard ways to derive the minimax theorem from LP duality [8, section 13-2]. Section 6 describes the classical converse approach, which we show to be incomplete.

3 The Lemma of Farkas and LP duality

The standard way to prove the LP duality theorem uses the Lemma of Farkas [10], stated in (13) below, which characterizes when an inhomogeneous system $Ax = b$ of linear equations has no solution $x \ge \mathbf{0}$ in nonnegative variables. Two related theorems are (14) and (15). The following proposition asserts how close they are, by using the respective matrix in different ways (we say "proves" rather than "implies" because it is not the same matrix).

Proposition 1. *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then each of the following three assertions proves the others: The Lemma of Farkas with equalities and nonnegative variables*

$$\nexists x \in \mathbb{R}^n \ :\ Ax = b,\ \ x \ge \mathbf{0} \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^m \ :\ y^\top A \ge \mathbf{0}^\top,\ \ y^\top b < 0\,, \tag{13}$$

the Lemma of Farkas with inequalities and nonnegative variables

$$\nexists x \in \mathbb{R}^n \ :\ Ax \le b,\ \ x \ge \mathbf{0} \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^m \ :\ y^\top A \ge \mathbf{0}^\top,\ \ y \ge \mathbf{0},\ \ y^\top b < 0\,, \tag{14}$$

and the Lemma of Farkas with inequalities and unconstrained variables

$$\nexists x \in \mathbb{R}^n \ :\ Ax \le b \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^m \ :\ y^\top A = \mathbf{0}^\top,\ \ y \ge \mathbf{0},\ \ y^\top b < 0\,. \tag{15}$$

Proof. In each of (13), (14), (15) the direction "$\Leftarrow$" is immediate, for example in (13) because $y^\top A \ge \mathbf{0}^\top$ and $Ax = b$, $x \ge \mathbf{0}$ imply $y^\top b = y^\top Ax \ge 0$ which contradicts $y^\top b < 0$. We therefore only consider "$\Rightarrow$". Condition (13) proves (14) by writing $Ax \le b$ as $Ax + s = b$, $s \ge \mathbf{0}$ for a vector of *slack variables* $s \in \mathbb{R}^m$, and then applying (13) to the matrix $[A\ \ I\,]$ instead of A, where I is the $m \times m$ identity matrix.

Conversely, if there is no solution $x \geq \mathbf{0}$ to $Ax = b$, that is, to $Ax \leq b$ and $-Ax \leq -b$, then by (14) there are nonnegative $y^+, y^- \in \mathbb{R}^m$ with $(y^+)^\top A - (y^-)^\top A \geq \mathbf{0}^\top$ and $(y^+)^\top b - (y^-)^\top b < 0$. This shows (13) with $y = y^+ - y^-$.

Condition (15) follows from (14) by writing $Ax \leq b$ in (15) as $Ax^+ - Ax^- \leq b$ with nonnegative x^+ and x^-. The converse holds by writing $Ax \leq b$, $x \geq \mathbf{0}$ in (14) as $Ax \leq b$, $-x \leq \mathbf{0}$ in (15). □

These versions of the Lemma of Farkas are "theorems of the alternative" in that exactly one of two conditions is true, as in (13): Either there is a solution x to $Ax = b$, $x \geq \mathbf{0}$, or a solution y to $y^\top A \geq \mathbf{0}^\top$, $y^\top b < 0$, but not to both. We always state such theorems so that "⇒" is the nontrivial direction.

The following is standard (e.g., Gale [11, p. 79]), and similar arguments as used in the proof will be used repeatedly.

Proposition 2. *The inequality version* (14) *of the Lemma of Farkas proves LP duality.*

Proof. Suppose that the primal LP (1) has a feasible solution $\bar{x}$ and the dual LP (2) has a feasible solution $\bar{y}$ and that, contrary to the claim of the LP duality theorem, there are no feasible x and y so that $c^\top x = y^\top b$. That is, the system of inequalities

$$\begin{array}{rcrcl} & & Ax & \leq & b \\ -A^\top y & & & \leq & -c \\ b^\top y & - & c^\top x & \leq & 0 \end{array} \tag{16}$$

has no solution $(y, x) \in \mathbb{R}^m \times \mathbb{R}^n$ with $y \geq \mathbf{0}$ and $x \geq \mathbf{0}$. Hence, by (14) (written transposed), there are nonnegative $(\hat{y}, \hat{x}, t) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ such that

$$\begin{array}{rcrcrcl} & - & A\hat{x} & + & bt & \geq & \mathbf{0} \\ A^\top \hat{y} & & & - & ct & \geq & \mathbf{0} \\ b^\top \hat{y} & - & c^\top \hat{x} & & & < & 0\,. \end{array} \tag{17}$$

If $t > 0$ then $\hat{x}\frac{1}{t}$ and $\hat{y}\frac{1}{t}$ are feasible solutions to the primal (1) and dual (2) with $\frac{1}{t}\hat{y}^\top b < c^\top \hat{x}\frac{1}{t}$ in violation of weak duality (3). If $t = 0$ then $A\hat{x} \leq \mathbf{0}$ and $\hat{y}^\top A \geq \mathbf{0}^\top$. The last inequality in (17) implies that at least one of the inequalities $\hat{y}^\top b < 0$ or $0 < c^\top \hat{x}$ holds. Suppose the latter. For $\alpha \in \mathbb{R}$ we have $A(\bar{x} + \hat{x}\alpha) \leq b$ and $\bar{x} + \hat{x}\alpha \geq \mathbf{0}$, but $c^\top(\bar{x} + \hat{x}\alpha) \to \infty$ as $\alpha \to \infty$, that is, the objective function of the primal LP is unbounded, contradicting its upper bound $\bar{y}^\top b$ from the dual LP. Similarly, $\hat{y}^\top b < 0$ implies that the dual LP is unbounded and thus the primal LP infeasible, again a contradiction. This shows that (16) has a nonnegative solution (y, x) with $y^\top b \leq c^\top x$ and thus $y^\top b = c^\top x$ by weak duality, as claimed. □

The converse also holds, as well as a useful extension of LP duality.

Proposition 3. *The LP duality Theorem 2 proves* (14)*. Moreover, if the primal LP* (1) *is infeasible and the dual LP* (2) *is feasible, then the dual LP is unbounded.*

Proof. Suppose there is no $x \geq \mathbf{0}$ with $Ax \leq b$. Then the LP (with a new scalar variable t)

$$\underset{x,t}{\text{maximize}} \ -t \quad \text{subject to} \quad Ax - \mathbf{1}t \leq b, \quad x \geq \mathbf{0},\ t \geq 0 \tag{18}$$

(which is feasible by choosing $t \geq -b_i$ for all $i \in [m]$ and $x = \mathbf{0}$) has an optimum solution with $t > 0$. The dual LP to (18) states

$$\underset{y}{\text{minimize}} \ y^\top b \quad \text{subject to} \quad y^\top A \geq \mathbf{0}^\top, \quad -y^\top \mathbf{1} \geq -1, \ \ y \geq \mathbf{0}, \tag{19}$$

is feasible with $y = \mathbf{0}$, and therefore has an optimal solution $y \geq \mathbf{0}$ with equal objective function value to the primal, that is, $y^\top b = -t < 0$. This shows (14).

To prove the second part, suppose $\bar{y}^\top A^\top \geq c^\top$ for some $\bar{y} \geq \mathbf{0}$. Then with the preceding $y \geq \mathbf{0}$ such that $y^\top b < 0$ we have $(\bar{y}^\top + \alpha y^\top)A \geq c^\top$ and $\bar{y} + y\alpha \geq \mathbf{0}$ and $(\bar{y}^\top + \alpha y^\top)b \to -\infty$ as $\alpha \to \infty$. □

4 The theorems of Gordan and Ville

The Lemma of Farkas with equalities (13) characterizes when the inhomogeneous linear equations $Ax = b$ have no solution $x \geq \mathbf{0}$ in nonnegative variables. The following Theorem (20) of Gordan [14] for homogeneous equations characterizes when the system $Ax = \mathbf{0}$ has no nontrivial solution $x \geq \mathbf{0}$. Its "inequality version" (21) is known as the Theorem of Ville [26]. Ville's Theorem essentially states the minimax theorem for a game with positive value. To prove the minimax theorem from Ville's Theorem, the game should have its value normalized to zero. A common way to achieve this is to symmetrize the game [12]. Instead, we shift the payoffs as in (10) so that the max-min value is zero. Note that the min-max and max-min values in (7) and (8) exist without having to assume LP duality.

Proposition 4. *Let $A \in \mathbb{R}^{m\times n}$. Then the following Theorem (20) of Gordan proves the Theorem (21) of Ville and vice versa, and (21) proves the minimax theorem and vice versa:*

$$\nexists x \in \mathbb{R}^n \ : \ Ax = \mathbf{0}, \ \ x \geq \mathbf{0}, \ \ x \neq \mathbf{0} \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^m \ : \ y^\top A > \mathbf{0}^\top, \tag{20}$$

$$\nexists x \in \mathbb{R}^n \ : \ Ax \leq \mathbf{0}, \ \ x \geq \mathbf{0}, \ \ x \neq \mathbf{0} \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^m \ : \ y^\top A > \mathbf{0}^\top, \ \ y \geq \mathbf{0}. \tag{21}$$

Proof. Assume (20) holds. We prove (21). Suppose there is no $x \in \mathbb{R}^n$ with $Ax \leq \mathbf{0}$, $x \geq \mathbf{0}$, $x \neq \mathbf{0}$. Then there is no $x \in \mathbb{R}^n$ and $s \in \mathbb{R}^m$ with $Ax + s = \mathbf{0}$ and $x \geq \mathbf{0}$, $s \geq \mathbf{0}$, and $(x, s) \neq (\mathbf{0}, \mathbf{0})$ (this clearly holds if $x \neq \mathbf{0}$, and if $x = \mathbf{0}$ then $s = \mathbf{0}$). Hence, by (20), there is some $y \in \mathbb{R}^m$ with $y^\top A > \mathbf{0}^\top$ and $y > \mathbf{0}$ and thus $y \geq \mathbf{0}$. This shows the nontrivial direction "⇒" in (21).

Conversely, suppose there is no $x \geq \mathbf{0}$, $x \neq \mathbf{0}$ with $Ax = \mathbf{0}$ and hence no $x \geq \mathbf{0}$, $x \neq \mathbf{0}$ with $Ax \leq \mathbf{0}$ and $-Ax \leq \mathbf{0}$. Then by (21) there exist $y^+ \geq \mathbf{0}$ and $y^- \geq \mathbf{0}$

with $(y^+)^\top A + (y^-)^\top(-A) > \mathbf{0}^\top$, that is, $(y^+ - y^-)^\top A > \mathbf{0}^\top$, which shows (20) with $y = y^+ - y^-$.

Assume the minimax Theorem 2 holds for the game matrix A. The left-hand side of (21) states that the value v of the game is positive, because otherwise there would be a mixed strategy $x \in X$ with nonpositive min-max value v in (7). With the optimal $y \in Y$ and $u > 0$ in (8) we have $y^\top A \geq u\mathbf{1}^\top > \mathbf{0}^\top$ as asserted in (21).

Conversely, assume (21) and consider a game matrix A. Let u be its max-min value and $y \in Y$ be a max-min strategy as in (8). Let $A' = A - \mathbf{1}u\mathbf{1}^\top$. Then $y^\top A' = y^\top A - u\mathbf{1}^\top \geq \mathbf{0}^\top$. We claim that $A'x \leq \mathbf{0}$ for some $x \in X$. If not then there is no $x \geq \mathbf{0}$, $x \neq \mathbf{0}$ with $A'x \leq \mathbf{0}$ (otherwise scale x so that $x \in X$), and therefore by (21) we have $y^\top A' > \mathbf{0}^\top$ for some $y \geq \mathbf{0}$. Because $y \neq \mathbf{0}$, we can scale y such that $y \in Y$ and choose $\varepsilon > 0$ such that $y^\top A' \geq \varepsilon\mathbf{1}^\top$ and hence $y^\top A \geq (u + \varepsilon)\mathbf{1}^\top$, which contradicts the maximality of u in (8). Hence, there is $x \in X$ with $A'x \leq \mathbf{0}$, so A' has min-max value zero and therefore A has min-max value u, which proves the minimax theorem. □

5 The theorems of Stiemke and Loomis

This section is about two proofs of the minimax theorem, for example in order to use it for proving LP duality. For historical interest, we first reproduce a short proof of Gordan's Theorem (20) by Stiemke [23]. In modern language, it uses the property that the null space and row space of a matrix are orthogonal complements, as stated in (25) below. We state this property as the following "theorem of the alternative" about the solvability of linear equations without nonnegativity constraints, which is well known (e.g., [15]). We also use this lemma in Section 10 for a short proof of the Lemma of Farkas.

Lemma 1. *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then*

$$\nexists x \in \mathbb{R}^n \ : \ Ax = b \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^m \ : \ y^\top A = \mathbf{0}^\top, \ \ y^\top b \neq 0\,. \tag{22}$$

Proof. We show the nontrivial direction "⇒". Assume that b is not a linear combination of the columns $A_1, \ldots, A_n$ of A. Let k be the column rank of A and $\{A_j\}_{j\in K}$ be a basis of the column space of A, with $|K| = k \geq 0$, and let A_K be the matrix of these columns. By assumption, the $m \times (k+1)$ matrix $[A_K\ b]$ has rank $k+1$, which is also its row rank. Its rows span therefore all of $\mathbb{R}^{1\times(k+1)}$, in particular the vector $(\mathbf{0}^\top, 1)$, that is, $y^\top A_K = \mathbf{0}^\top$ and $y^\top b = 1$ for some $y \in \mathbb{R}^m$. Any other column A_j of A for $j \notin K$ is a linear combination of the basis columns, $A_j = A_K z^{(j)}$ for some $z^{(j)} \in \mathbb{R}^k$, which implies $y^\top A_j = y^\top A_K z^{(j)} = 0$. This shows that overall $y^\top A = \mathbf{0}^\top$ and $y^\top b \neq 0$, as required. □

Theorem 3 (Stiemke [23])**.** *Let $A \in \mathbb{R}^{m\times n}$. Then*

$$\nexists y \in \mathbb{R}^m \ : \ y^\top A \geq \mathbf{0}^\top, \ \ y^\top A \neq \mathbf{0} \quad \Leftrightarrow \quad \exists x \in \mathbb{R}^n \ : \ Ax = \mathbf{0}\,, \ \ x > \mathbf{0}\,. \tag{23}$$

Proof. Define

$$\begin{aligned} \mathsf{rowspace}(A) &= \{y^\top A \mid y \in \mathbb{R}^m\}\,, \\ \mathsf{nullspace}(A) &= \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\}\,. \end{aligned} \tag{24}$$

We have for $c \in \mathbb{R}^n$

$$c^\top \in \mathsf{rowspace}(A) \quad \Leftrightarrow \quad \forall x \in \mathsf{nullspace}(A) \;:\; c^\top x = 0 \tag{25}$$

because this is equivalent to

$$\exists y \;:\; y^\top A = c^\top \quad \Leftrightarrow \quad \nexists x \;:\; Ax = \mathbf{0}\,,\;\; c^\top x \neq 0\,, \tag{26}$$

which (with both sides negated) is the transposed version of (22).

The nontrivial direction in (23) is "$\Rightarrow$". It states: Suppose $\mathbf{0}^\top$ is the only nonnegative vector in $\mathsf{rowspace}(A)$. Then there is some $x \in \mathsf{nullspace}(A)$ with $x > \mathbf{0}$. We show this by induction on n. If $n = 1$ then the single column of A is $\mathbf{0}$, and we can choose $x = 1$. Let $n > 1$ and suppose the claim is true for $n - 1$.

Case 1. There is some $a \in \mathbb{R}^{n-1}$, $a \geq \mathbf{0}$, $a \neq \mathbf{0}$ so that $(1, -a^\top) \in \mathsf{rowspace}(A)$. Consider a set of row vectors $(1, -a^\top)$, $(0, a_2^\top)$, $\ldots$, $(0, a_m^\top)$ that span $\mathsf{rowspace}(A)$ (easily obtained from the rows of A). There is no $w \in \mathbb{R}^{m-1}$ such that $c^\top = \sum_{i=2}^m w_{i-1}\, a_i^\top$ is nonnegative and nonzero, because otherwise $(0, c^\top)$ is in $\mathsf{rowspace}(A)$ and nonnegative and nonzero. Hence, by inductive hypothesis, there is some $z \in \mathbb{R}^{m-1}$, $z > \mathbf{0}$, such that $a_i^\top z = 0$ for $2 \leq i \leq m$. Then $x_1 = a^\top z > 0$, and $x = \binom{x_1}{z} \in \mathsf{nullspace}(A)$ by (25) because $(1, -a^\top)x = 0$ and $(0, a_i^\top)x = 0$ for $2 \leq i \leq m$, and $x > \mathbf{0}$.

Case 2. Otherwise, consider any $y \in \mathbb{R}^m$ and let $(c_1, c^\top) = y^\top A$ with $c \in \mathbb{R}^{m-1}$. Then $c \geq \mathbf{0}$ implies $c = \mathbf{0}$, which holds by assumption if $c_1 \geq 0$, and if $c_1 < 0$ and $c \geq \mathbf{0}$, $c \neq \mathbf{0}$ then $(1, \frac{1}{c_1} c^\top) \in \mathsf{rowspace}(A)$ and Case 1 applies. By inductive hypothesis, there is some $z \in \mathbb{R}^{m-1}$, $z > \mathbf{0}$, such that $A\binom{0}{z} = \mathbf{0}$. If $x_1 = 0$ for all $x \in \mathsf{nullspace}(A)$ then by (25) we have $(1, 0, \ldots, 0) \in \mathsf{rowspace}(A)$ contrary to assumption. So there is some $x' \in \mathsf{nullspace}(A)$ with $x'_1 > 0$, and therefore $x = x'\varepsilon + \binom{0}{z} > \mathbf{0}$ for sufficiently small $\varepsilon > 0$, where $Ax = \mathbf{0}$. This completes the induction. □

The preceding theorem is statement I of Stiemke [23], and Gordan's Theorem (20) is statement II.

Proposition 5. *Stiemke's Theorem 3 proves Gordan's Theorem (20).*

Proof. Let $A \in \mathbb{R}^{m \times n}$. Let $\{b_1, \ldots, b_k\}$ with $k \geq 1$ be a spanning set of $\mathsf{nullspace}(A)$ and $B = [b_1 \cdots b_k]$. Then for b and c in $\mathbb{R}^n$

$$b \in \mathsf{nullspace}(A) \quad \Leftrightarrow \quad b^\top \in \mathsf{rowspace}(B^\top) \tag{27}$$

and, using (25),

$$\begin{aligned} & c^\top \in \mathsf{rowspace}(A) \\ \Leftrightarrow\quad & \forall x \in \mathsf{nullspace}(A) \ : \ c^\top x = 0 \\ \Leftrightarrow\quad & c^\top b_i = 0 \qquad (1 \le i \le k) \\ \Leftrightarrow\quad & c^\top B = \mathbf{0}^\top \\ \Leftrightarrow\quad & c \in \mathsf{nullspace}(B^\top)\,. \end{aligned} \tag{28}$$

Stiemke's Theorem (23) applied to $B^\top$ instead of A states

$$\nexists b^\top \in \mathsf{rowspace}(B^\top)\,,\ \ b \ge \mathbf{0}\,,\ \ b \ne \mathbf{0} \quad \Leftrightarrow \quad \exists c \in \mathsf{nullspace}(B^\top) \ : \ c > \mathbf{0} \tag{29}$$

which by (27) and (28) is equivalent to

$$\nexists b \in \mathsf{nullspace}(A)\,,\ \ b \ge \mathbf{0}\,,\ \ b \ne \mathbf{0} \quad \Leftrightarrow \quad \exists c^\top \in \mathsf{rowspace}(A) \ : \ c > \mathbf{0} \tag{30}$$

which is Gordan's Theorem (20). □

Via Propositions 4 and 5, Stiemke's Theorem 3 therefore proves the minimax theorem. Using symmetric games, this was also shown by Gale, Kuhn, and Tucker [12].

Our favorite proof of the minimax theorem is based on the following theorem.

Theorem 4 (Loomis [17])**.** *Let A and B be two $m \times n$ matrices with $B > 0$. Then there exist $x \in X$, $y \in Y$, and $v \in \mathbb{R}$ such that $Ax \le Bxv$ and $y^\top A \ge vy^\top B$.*

The case $B = \mathbf{1}\mathbf{1}^\top$ gives the minimax theorem. Conversely, the minimax theorem proves Theorem 4 [16, p. 19]: Because $B > 0$, the value of the game $A - \alpha B$ is negative for sufficiently large α, positive for sufficiently negative α, is a continuous function of α, and therefore zero for some α, which then gives Theorem 4 with $v = \alpha$.

The following is the proof by Loomis [17] of Theorem 4 specialized to the minimax theorem. It is an induction proof about the min-max value v and max-min value u (which exist, irrespective of LP duality). It is easy to remember: If the players have optimal strategies that equalize v and u for all rows and columns, then $u = v$. Otherwise (if needed by exchanging the players), there is at least one row with lower payoff than v, which will *anyhow* not be chosen by the row player. By omitting this row from the game, the minimax theorem holds (using a bit of convexity and continuity) by the inductive hypothesis.

Proof of Theorem 2. Consider optimal solutions v, x to (7) and u, y to (8), where

$$u = u\mathbf{1}^\top x \le y^\top Ax \le y^\top \mathbf{1}v = v. \tag{31}$$

We prove $u = v$ by induction on $m + n$. It holds trivially for $m + n = 2$. If all inequalities in (31) hold as equalities, then $u = v$. Hence, assume that at least one inequality is strict, say $(Ax)_k < v$ for some row $k \in [m]$ (the case for a column is similar). Let $\bar{A}$ be the matrix A with the kth row deleted. By induction hypothesis, $\bar{A}$ has game value $\bar{v}$ with $\bar{A}\bar{x} \le \mathbf{1}\bar{v}$ for some $\bar{x} \in X$, where it is easy to see that

$$\bar{v} \le v, \qquad \bar{v} \le u \tag{32}$$

because compared to A the game $\bar{A}$ strengthens the minimizing column player.

We claim that $\bar{v} = v$. Namely, if $\bar{v} < v$, let $0 < \varepsilon \le 1$ and consider the strategy $x(\varepsilon) = x(1-\varepsilon) + \bar{x}\varepsilon$ where $x \in X$ because X is convex. Then

$$\bar{A}x(\varepsilon) \;=\; \bar{A}(x(1-\varepsilon)+\bar{x}\varepsilon) \le \mathbf{1}v(1-\varepsilon) + \mathbf{1}\bar{v}\varepsilon = \mathbf{1}(v - \varepsilon(v-\bar{v})) < \mathbf{1}v\,. \tag{33}$$

For the missing row k of A where $(Ax)_k < v$ we have for sufficiently small ε

$$(Ax(\varepsilon))_k = (Ax)_k(1-\varepsilon) + (A\bar{x})_k\varepsilon < v\,. \tag{34}$$

Hence, $Ax(\varepsilon) < \mathbf{1}v$ for some $x(\varepsilon) \in X$, in contradiction to the minimality of v in (7). This shows $v = \bar{v}$, and, by (32), $\bar{v} \le u \le v = \bar{v}$ and therefore $u = v$. This completes the induction. □

The proof by Loomis [17] has been noted (in particular by von Neumann and Morgenstern [28, p. vi]) but is not widely known, and should be a standard textbook proof (as in [29, p. 216]). (A better title of Loomis's paper would have been "An elementary proof of the minimax theorem", given that Theorem 4 is not substantially more general.) It was, in essence, re-discovered by Owen [19]. However, Owen needlessly manipulates the max-min strategy y; the proof by Loomis is more transparent. Owen's proof is discussed further by Binmore [2].

The research in this paper originated with an attempt to extend the induction proof by Loomis to a direct proof of LP duality, via the existence of a strictly complementary pair of optimal strategies in a zero-sum game, applied to Dantzig's game in (35) below. This existence seems to be difficult to prove within this induction. For example, the game $\begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ has a max-min and min-max strategy where every pure best response is played with positive probability (such as both players mixing uniformly), but also the left column as a pure min-max strategy. However, omitting the unplayed second or third column in an induction would alter the game substantially, because then a strictly complementary pair has the first column as a unique min-max strategy, with a positive slack in the column that was not omitted.

6 The minimax theorem and LP duality

The following theorem assumes the minimax theorem.

Theorem 5 (Dantzig [7])**.** *Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Consider the zero-sum game with the payoff matrix B (with $k = m + n + 1$ rows and columns) defined by*

$$B = \begin{bmatrix} 0 & A & -b \\ -A^\top & 0 & c \\ b^\top & -c^\top & 0 \end{bmatrix}. \tag{35}$$

Then B has value zero, with a min-max strategy $z = (y, x, t) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ that is also a max-min strategy, with $Bz \le \mathbf{0}$. If $z_k = t > 0$ then $x\frac{1}{t}$ is an optimal solution to the primal LP (1) and $y\frac{1}{t}$ is an optimal solution to dual LP (2). If $(Bz)_k < 0$ then $t = 0$ and at least one of the LPs (1) or (2) is infeasible.

Proof. Because $B = -B^\top$, this game is symmetric and its game value v is zero. Let $z = (y, x, t)$. Then $Bz \le \mathbf{0}$ states $Ax - bt \le \mathbf{0}$, $-A^\top y + ct \le \mathbf{0}$, and $b^\top y - c^\top x \le 0$. If $t > 0$ then $x\frac{1}{t}$ and $y\frac{1}{t}$ are primal and dual feasible with $b^\top y\frac{1}{t} \le c^\top x\frac{1}{t}$ and therefore optimal.

If $(Bz)_k < 0$, that is, $b^\top y - c^\top x < 0$, then $t > 0$ would violate weak duality, so $t = 0$. Moreover, $Ax \le \mathbf{0}$ and $y^\top A \ge \mathbf{0}^\top$, and $y^\top b < 0$ or $0 < c^\top x$. As shown following (17), this implies infeasibility of at least one of the LPs (1) or (2). □

Hence, Theorem 5 seems to show that the minimax theorem proves LP duality. The known "hole" in this argument is that it is does not cover the case of a min-max strategy z where $z_k = 0$ and $(Bz)_k = 0$, which is therefore uninformative, as noted by Dantzig [8, p. 291]. Luce and Raiffa [18, p. 421] claim without proof (or forgot a reference, e.g. to corollary 3A in their cited work [13]) that if $(Bz)_k = 0$ for all min-max strategies z, then $\bar{z}_k > 0$ for some max-min strategy $\bar{z}$. Because B is skew-symmetric ($B = -B^\top$), this would solve the problem with $\bar{z}$ as a min-max strategy. We will show that this assumption is essentially the Lemma of Tucker [25, p. 5] for the case of a skew-symmetric matrix. Already for the special case of B in (35), this proves the Lemma of Farkas (14) (see also [4, theorem 1.1]), and this defeats the purpose of proving LP duality from the minimax theorem.

Proposition 6. *Consider B in (35) with $c = \mathbf{0}$, and suppose that there is always some $z \ge \mathbf{0}$ with $Bz \le \mathbf{0}$ and $z_k - (Bz)_k > 0$. Then this proves (14).*

Proof. Let $z = (y, x, t)$ as described, where $Ax - bt \le \mathbf{0}$ and $-A^\top y \le \mathbf{0}$ and $b^\top y \le 0$ because $Bz \le \mathbf{0}$, and $z_k - (Bz)_k = t - b^\top y > 0$. Then if $t > 0$ we have $Ax\frac{1}{t} \le b$, and if $t = 0$ then $y^\top A \ge \mathbf{0}^\top$ and $y^\top b < 0$, which proves (14). □

The Lemma of Tucker comes in several variants.

Proposition 7. *Let $A \in \mathbb{R}^{m\times n}$. Then the following Lemma of Tucker*

$$\exists y \in \mathbb{R}^m,\ x \in \mathbb{R}^n \ :\ y^\top A \ge \mathbf{0}^\top, \quad x \ge \mathbf{0}, \quad Ax = \mathbf{0}, \quad x_n + (y^\top A)_n > 0 \tag{36}$$

proves the following inequality version and vice versa:

$$\exists y \in \mathbb{R}^m,\ x \in \mathbb{R}^n \ :\ y \ge \mathbf{0},\ \ y^\top A \ge \mathbf{0}^\top,\ \ x \ge \mathbf{0},\ \ Ax \le \mathbf{0},\ \ x_n + (y^\top A)_n > 0\,, \tag{37}$$

and similarly its version for a skew-symmetric matrix $B \in \mathbb{R}^{k\times k}$, that is, $B = -B^\top$:

$$\exists z \in \mathbb{R}^k \;:\; z \geq \mathbf{0}, \quad Bz \leq \mathbf{0}, \quad z_k - (Bz)_k > 0\,. \tag{38}$$

Proof. Applying (36) to the matrix $[\,I \;\; A]$ with the identity matrix I gives (37). For the converse, write $Ax = \mathbf{0}$ as $Ax \leq \mathbf{0}\,, -Ax \leq \mathbf{0}\,.$

Condition (38) follows from (37) with $A = B$ and $z = x + y$ because $-Bz = z^\top B$ and $y_n \geq 0$ and $(x^\top B)_n \geq 0$. For the converse, use $B = \begin{bmatrix} 0 & A \\ -A^\top & 0 \end{bmatrix}$ and $z = \binom{y}{x}$. □

Tucker [25, p. 7] used (36) to prove the Lemma of Farkas in its version (13). Less known, but similarly easy, is that the converse holds as well.

Proposition 8. *The Lemma of Farkas (13) proves Tucker's Lemma (36).*

Proof. Let $A = [A_1 \cdots A_n] \in \mathbb{R}^{m\times n}$. By (13), either $\sum_{j=1}^{n-1} A_j z_j = -A_n$ for some $z \in \mathbb{R}^{n-1}$ with $z \geq \mathbf{0}$, in which case let $x = \binom{z}{1}$ and $y = \mathbf{0}$, or otherwise $y^\top A_j \geq 0$ for $1 \leq j < n$ and $y^\top(-A_n) < 0$ for some $y \in \mathbb{R}^m$, in which case let $x = \mathbf{0}$. In both cases we have $Ax = \mathbf{0}$ and $x_n + y^\top A_n > 0$, and (36) holds. □

In the next section, we show a proper way of proving LP duality from the minimax theorem.

7 Proving Tucker's Theorem from Gordan's Theorem

In Tucker's Lemma (36), the last (nth) column of the matrix A plays a special role, which can be taken by any other column. This proves the following stronger version (39) known as the *Theorem* of Tucker [25, p. 8].

Proposition 9. *Let $A \in \mathbb{R}^{m\times n}$. Tucker's Lemma (36) proves Tucker's Theorem*

$$\exists y \in \mathbb{R}^m,\, x \in \mathbb{R}^n \;:\; y^\top A \geq \mathbf{0}^\top, \quad x \geq \mathbf{0}, \quad Ax = \mathbf{0}, \quad x^\top + y^\top A > \mathbf{0}^\top. \tag{39}$$

Proof. Let $j \in [n]$. By applying (36) to the jth column of A with j instead of n, choose $y^{(j)} \in \mathbb{R}^m$ and $x^{(j)} \in \mathbb{R}^n$ such that

$$(y^{(j)})^\top A \geq \mathbf{0}^\top, \quad x^{(j)} \geq \mathbf{0}, \quad Ax^{(j)} = \mathbf{0}, \quad x_j^{(j)} + ((y^{(j)})^\top A)_j > 0\,. \tag{40}$$

Then $y = \sum_{j\in[n]} y^{(j)}$ and $x = \sum_{j\in[n]} x^{(j)}$ fulfill (39). □

Tucker's Theorem (39) is a very versatile theorem that proves a number of theorems of the alternative (see [25]), for example immediately Gordan's Theorem (20) or Stiemke's Theorem 3.

The main Theorem 6 of this section shows that Gordan's Theorem (20) proves Tucker's Theorem (39). It is based on the following observation. If $Ax = \mathbf{0}$ and

$x \geq \mathbf{0}$, then any y with $y^\top A \geq \mathbf{0}^\top$ has the property that if $x_j > 0$ then $(y^\top A)_j = 0$ because otherwise $0 = y^\top Ax = \sum_{j\in[n]}(y^\top A)_j x_j > 0$. Hence, (39) implies that the *support*

$$S = \mathsf{supp}(x) = \{j \in [n] \mid x_j > 0\} \tag{41}$$

of x is unique. The main idea is that the nonnegativity constraints for the variables x_j for $j \in S$ can be dropped and these variables therefore be eliminated, which allows applying Gordan's Theorem to the remaining variables. The following proof is distilled from the more complicated computational approach of Adler [1, section 4].

Theorem 6. *Gordan's Theorem (20) proves Tucker's Theorem (39).*

Proof. Let $A = [A_1 \cdots A_n]$. For any $S \subseteq [n]$ and $J = [n] - S$ write $A = [A_J\ A_S]$ and $x = (x_J, x_S)$ for $x \in \mathbb{R}^n$. If $Ax = \mathbf{0}$, $x \geq \mathbf{0}$, $Ax' = \mathbf{0}$, $x' \geq \mathbf{0}$, then $A(x + x') = \mathbf{0}$, $x + x' \geq \mathbf{0}$, and $\mathsf{supp}(x + x') = \mathsf{supp}(x) \cup \mathsf{supp}(x')$. Choose S as the inclusion-maximal support of any $x \geq \mathbf{0}$ such that $Ax = \mathbf{0}$. Then any y with $y^\top A \geq \mathbf{0}^\top$ fulfills $y^\top A_S = \mathbf{0}^\top$ (because otherwise $y^\top Ax = y^\top A_S x_S > 0$).

On the other hand, (39) states $x_j + y^\top A_j > 0$ for all $j \in [n]$, which requires $y^\top A_j > 0$ for $j \in J = [n] - S$. We now show that there indeed exist $y \in \mathbb{R}^m$ and $x = (\mathbf{0}, x_S)$ such that

$$y^\top A_J > \mathbf{0}^\top,\ \ y^\top A_S = \mathbf{0}^\top,\ \ Ax = A_S x_S = \mathbf{0},\ \ x_S > \mathbf{0}\,, \tag{42}$$

which implies (39). Consider some $\tilde{x} \geq \mathbf{0}$ with maximum support $S = \mathsf{supp}(\tilde{x})$ such that $A\tilde{x} = \mathbf{0}$, that is, $\tilde{x}_S > \mathbf{0}$. If $S = [n]$ we are done. Let k be the rank of A_S. Suppose $k = m$. We claim that then $S = [n]$, which implies (39) with $y = \mathbf{0}$. Namely, if $j \in [n] - S$, then $A_j = A_S \hat{x}_S$ for some $\hat{x}_S$ because A_S has full rank, and therefore $A_j + A_S(\tilde{x}_S \alpha - \hat{x}_S) = \mathbf{0}$ where $\tilde{x}_S \alpha - \hat{x}_S > \mathbf{0}$ for sufficiently large α, which gives a solution $x \geq \mathbf{0}$ to $Ax = \mathbf{0}$ with $\mathsf{supp}(x) = \{j\} \cup S$ in contradiction to the maximality of S.

Hence, let $k < m$. In order to apply Gordan's Theorem (20), we eliminate the variables x_S from the system $Ax = A_J x_J + A_S x_S = \mathbf{0}$ by replacing it with an equivalent system $CAx = 0$ with a suitable invertible $m \times m$ matrix C. Let a_{iS} be the ith row of A_S for $i \in [m]$. Suppose for simplicity that the last k rows of A_S are linearly independent and define the matrix F, and that for $i = 1, \ldots, m - k$ we have $a_{iS} = z^{(i)} F$ for some row vector $z^{(i)}$ in $\mathbb{R}^{1\times k}$. Then the $m \times m$ matrix

$$C = \begin{bmatrix} 1 & \cdots & 0 & -z^{(1)} & & \\ & \ddots & & \vdots & & \\ 0 & \cdots & 1 & -z^{(m-k)} & & \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ & \vdots & & & \ddots & \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{43}$$

is clearly invertible, and any solution (x_J, x_S) to $A_J x_J + A_S x_S = \mathbf{0}$ is a solution to

$$CA_J x_J + CA_S x_S = \mathbf{0} \tag{44}$$

and vice versa, with

$$CA_J = \begin{bmatrix} D \\ E \end{bmatrix}, \quad CA_S = \begin{bmatrix} 0 \\ F \end{bmatrix} \tag{45}$$

where $D \in \mathbb{R}^{(m-k)\times|J|}$, $E \in \mathbb{R}^{k\times|J|}$, and $F \in \mathbb{R}^{k\times|S|}$.

Suppose there is some $x_J \in \mathbb{R}^{|J|}$ with

$$Dx_J = \mathbf{0}\,, \quad x_J \geq \mathbf{0}, \;\; x_J \neq \mathbf{0}\,. \tag{46}$$

Because F has rank k there exists x_S so that $Fx_S = -Ex_J$. Then $Ex_J + Fx_S = \mathbf{0}$ and hence $CA_J x_J + CA_S x_S = \mathbf{0}$ and thus $A_J x_J + A_S x_S = \mathbf{0}$. With $x(\alpha) = (x_J, x_S + \tilde{x}_S \alpha)$ we have $Ax(\alpha) = \mathbf{0}$ (because $A_S \tilde{x}_S = \mathbf{0}$) and $x(\alpha) \geq \mathbf{0}$ for $\alpha \to \infty$, where $x(\alpha)$ has larger support that S, but S was maximal. Hence, there is no x_J so that (46) holds. By Gordan's Theorem (20), there is some $w \in \mathbb{R}^{m-k}$ with $w^\top D > \mathbf{0}^\top$, that is,

$$(w^\top, \mathbf{0}^\top) \begin{bmatrix} D \\ E \end{bmatrix} > \mathbf{0}^\top, \quad (w^\top, \mathbf{0}^\top) \begin{bmatrix} 0 \\ F \end{bmatrix} = \mathbf{0}^\top.$$

With $y^\top = (w^\top, \mathbf{0}^\top)C$ and (45), this implies (42) with $x = \tilde{x}$, as claimed. □

Because the minimax theorem proves Gordan's Theorem (see Proposition 4), it proves Tucker's Theorem (39) and Tucker's Lemma (36) and the Lemma of Farkas and therefore LP duality.

Instead of the minimax theorem we can by Proposition 5 use Stiemke's Theorem 3 to prove Gordan's Theorem (20). The short proof by Tucker [25, p. 5–7] of his Lemma (36) has some structural similarities to Stiemke's proof but uses more explicit computations.

We conclude this section to show how Tucker's Theorem proves, as one of its main applications [25, theorem 6], the condition of *strict complementarity* in linear programming. For the LP (1) and its dual LP (2), a feasible pair x, y of solutions is optimal if and only if we have equality in (3), that is, $c^\top x = y^\top Ax = y^\top b$, which means

$$y^\top(b - Ax) = 0\,, \qquad (y^\top A - c^\top)x = 0\,. \tag{47}$$

This orthogonality of the nonnegative vectors y and $b - Ax$, and of $y^\top A - c^\top$ and x, means that they are complementary in the sense that in each component *at least one* of them is zero:

$$y_i(b - Ax)_i = 0 \quad (i \in [m]), \qquad (y^\top A - c^\top)_j\, x_j = 0 \quad (j \in [n]), \tag{48}$$

also called "complementary slackness". The following theorem asserts *strict* complementarity, namely that if (1) and (2) are feasible, then they have feasible solutions x and y where *exactly one* of each component in (48) is zero.

Proposition 10. *If the LPs* (1) *and* (2) *are feasible, then they have optimal solutions* x *and* y *such that* (47) *holds and*

$$y + (b - Ax) > \mathbf{0}, \qquad x^\top + (y^\top A - c^\top) > \mathbf{0}^\top. \tag{49}$$

Proof. Optimality of x and y means $c^\top x = y^\top b$ and therefore (47). Similar to Proposition 7 and (38), Tucker's Theorem (39) proves that for a skew-symmetric matrix B there is some z such that

$$z \geq \mathbf{0}, \quad Bz \leq \mathbf{0}, \quad z - Bz > \mathbf{0}. \tag{50}$$

Applied to the game matrix B in (35), because the LPs are feasible, this gives a solution $z = (y', x', t')$ with $t' > 0$, where $y = y'\frac{1}{t}$ and $x = x'\frac{1}{t}$ fulfill (49). □

The proof of Proposition 10 demonstrates a very good use of Dantzig's game B in (35). Geometrically, the LP solutions x and y are then in the relative interior of the set of optimal solutions. Unless this set is a singleton, x and y are not unique, but their supports $\mathsf{supp}(x)$ and $\mathsf{supp}(y)$ are unique, shown similarly to the initial argument in the proof of Theorem 6.

8 Extending Dantzig's game

In this section, we give a longer but more constructive proof of LP duality from the minimax theorem. We present a natural extension of Dantzig's game B in (35) by adding an extra row to B, giving the game B_M in (51) below. The aim is to "enforce" the last column of B to be played with positive probability t if that is possible. Any max-min strategy for B_M gives not only information about solutions to the LPs (1) and (2) if both are feasible, but also a certificate in (52) if not.

Theorem 7. *There is some* $M \in \mathbb{R}$ *with the following properties: If both the primal LP* (1) *and its dual* (2) *are feasible, then they also have respective feasible solutions* x *and* y *with* $\mathbf{1}^\top x + \mathbf{1}^\top y + 1 \leq M$. *Moreover, consider the zero-sum game*

$$B_M = \begin{bmatrix} 0 & A & -b \\ -A^\top & 0 & c \\ b^\top & -c^\top & 0 \\ \mathbf{1}^\top & \mathbf{1}^\top & -M \end{bmatrix} \tag{51}$$

with value v. *Then* $v \geq 0$, *and*

(a) $v = 0$ *with min-max strategy* (y, x, t) *and max-min strategy* $(y, x, t, 0)$ *for* B_M *if and only if* (1) *and* (2) *are feasible, in which case* $x\frac{1}{t}$ *is optimal for* (1) *and* $y\frac{1}{t}$ *is optimal for* (2).

(b) *If* $v > 0$ *with max-min strategy* (y, x, r, s) *for* B_M, *then* $r = 0$, $s = v$, *and*

$$Ax \leq \mathbf{0}, \quad x \geq \mathbf{0}, \quad A^\top y \geq \mathbf{0}, \quad y \geq \mathbf{0}, \quad b^\top y - c^\top x < 0, \tag{52}$$

which proves that (1) or (2) is infeasible. Moreover, $v < 1$, and the smallest number w such that

$$A\bar{x} \le b + \mathbf{1}w, \quad \bar{x} \ge \mathbf{0}, \quad -A^\top \bar{y} \le -c + \mathbf{1}w, \quad \bar{y} \ge \mathbf{0} \tag{53}$$

has feasible solutions $\bar{x}$ and $\bar{y}$ is given by

$$w = \frac{M+1}{1/v - 1}\,. \tag{54}$$

(c) *If the entries of A, b, c are rational numbers, let α be the maximum of the absolute value of the numerators of these numbers, let β be the maximum denominator, and $\ell = m + n + 1$. Then a suitable choice of M is*

$$M = \ell!\,\ell\alpha^{\ell}\beta^{\ell^2+\ell} + 1, \tag{55}$$

which in bit-size is polynomial in the bit-size of A, b, c.

We first discuss Theorem 7. We will prove it (in Theorem 8 below) without using LP duality, which will therefore be an alternative proof of LP duality from the minimax theorem. Although this proof is longer than that of Theorem 6, it provides a reduction of the problem of solving an LP (in the sense of providing an optimal solution or a certificate that the LP is unbounded or infeasible) to the problem of solving a zero-sum game. This reduction is new, as discussed further in Section 9.

Some observations in Theorem 7 are immediate: The value v of B_M is non-negative because the row player can ignore the last row and play as in Dantzig's game B in (35). Furthermore, if $v = 0$, then the second-to-last row in B_M states $\mathbf{1}^\top y + \mathbf{1}^\top x - Mt \le 0$ for any min-max strategy (y, x, t), which means $t > 0$. That strategy can be used as a max-min strategy (with the last row of B_M unplayed), with optimal solutions $x\frac{1}{t}$ and $y\frac{1}{t}$ to (1) and (2). For the converse, however, (1) and (2) may have feasible solutions x and y, respectively, but none of them fulfill $c^\top x \ge y^\top b$ unless we assume the LP duality theorem (which then proves (a)). In order to avoid using strong LP duality, we have to argue more carefully, as done in Theorem 8 below. Also, the optimal strategies $x\frac{1}{t}$ and $y\frac{1}{t}$ fulfill $\mathbf{1}^\top y\frac{1}{t} + \mathbf{1}^\top x\frac{1}{t} \le M$, so this constraint does not (and must not) affect feasibility of (1) and (2).

Theorem 7(b) gives a certificate that at least one of the LPs (1) and (2) is infeasible, if that is the case, via any max-min strategy (y, x, r, s). Then (52) holds (which follows from $r = 0$ and $s = v$), which implies $c^\top x > 0$ or $b^\top y < 0$ (or both) and thus unbounded solutions to (1) or (2), respectively, if either LP is feasible (and then the other LP is not). Furthermore, the value v of B_M defines, in a strictly monotonic relation (54), the minimal constant w in (53) added as extra slack to the right-hand sides that makes both LPs feasible. Given A, b, c, the value of w in (53) is clearly unique (and finite and independent of M), whereas the game value v of B_M depends on M.

Theorem 7(c) shows that a suitable constant M can be found by identifying the largest numerator (in absolute value) α and denominator β of the entries in A, b, c if these are given as rational numbers. (A similar a priori bound is known if these entries are algebraic numbers [1, p. 172], but not if they are general real numbers.) Although M in (55) is large, its description as a binary number is of polynomial size in the description of A, b, c. The conversion of the LPs (1) and (2) to the game matrix B_M is therefore a polynomial "Karp-type" reduction, where any minimax solution of B_M either solves the LPs or proves the infeasibility of at least one of them.

Finding M as in Theorem 7 uses the following well-known concept. A *basic* solution x to $Ax = b$ is given by a solution x where the columns A_j of A with $x_j \neq 0$ are linearly independent, which then determine uniquely the solution x. These columns are then easily extended to a basis of the column space of A and define a *basis matrix*. If A has full row rank m, then a basis has size m, and the basis matrix is an invertible $m \times m$ matrix. A basic *feasible* solution x also fulfills $x \geq \mathbf{0}$. A basic feasible solution to inequalities $Ax \leq b$ (and $x \geq \mathbf{0}$) is meant to be a basic feasible solution to the system $Ax + p = b$ (and $x, p \geq \mathbf{0}$), which has full row rank.

Part (b) in the following lemma and its proof are due to Ilan Adler (personal communication, 2022).

Lemma 2. *Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.*

(a) *If $Ax = b$, $x \geq \mathbf{0}$ has a feasible solution x, then it also has a basic feasible solution.*

Furthermore, suppose the LP: minimize $c^\top x$ subject to $Ax = b$, $x \geq \mathbf{0}$ is feasible and has a known lower bound λ, that is, $c^\top x \geq \lambda$ for all feasible x. Then

(b) *for every feasible solution x to $Ax = b$, $x \geq \mathbf{0}$ there is a basic feasible solution x^* with $c^\top x^* \leq c^\top x$,*

(c) *and* $\min\{c^\top x^* \mid Ax^* = b,\ x^* \geq \mathbf{0},\ x^* \text{ is basic}\} = \min\{c^\top x \mid Ax = b,\ x \geq \mathbf{0}\}$.

Proof. Choose a feasible x to $Ax = b$, $x \geq \mathbf{0}$ with minimal support. Then the columns A_j of A for $x_j > 0$ are linearly independent: Namely, if $Az = \mathbf{0}$ for some $z \neq \mathbf{0}$ where $z_j \neq 0$ implies $x_j > 0$, let $P = \{j \mid z_j > 0\}$ where $P \neq \emptyset$ (otherwise replace z by $-z$). Then with

$$\alpha = \min\{x_j/z_j \mid j \in P\}, \qquad x' = x - z\alpha \tag{56}$$

we have $Ax' = b$, $x' \geq \mathbf{0}$, and x' of smaller support than x. Hence, no such z exists, which proves the claimed linear independence. This shows (a).

To show (b), suppose $Ax = b$ and $x \geq \mathbf{0}$ and x is not basic, with $Az = \mathbf{0}$ for some $z \neq \mathbf{0}$ where $z_j \neq 0$ implies $x_j > 0$ as before. If $c^\top z < 0$, or if $c^\top z = 0$ and $z \leq \mathbf{0}$, replace z by $-z$. Let $P = \{j \mid z_j > 0\}$. Then $P \neq \emptyset$, which holds if $c^\top z = 0$ because $z \neq \mathbf{0}$, and if $P = \emptyset$ and $c^\top z > 0$ then $z \leq \mathbf{0}$, and $x - z\alpha$ is feasible but $c^\top(x - z\alpha)$ is arbitrarily negative as $\alpha \to \infty$, which contradicts boundedness. Then with α and x' as in (56), x' has smaller support than x and $c^\top x' \le c^\top x$. If A has n columns, then this process terminates after at most n steps with a basic feasible solution x^* with $c^\top x^* \le c^\top x$, as claimed.

Part (c) follows from (b) because there are finitely many basic feasible solutions, so the minimum on the left exists, and the minimum on the right also exists and equals its infimum. □

With the added equation $\mathbf{1}^\top x = 1$, Lemma 2(a) is *Carathéodory's theorem*: Any convex combination b of points in $\mathbb{R}^m$ is already the convex combination of a suitable set of at most $m+1$ of these points [5, p. 200].

We prove Theorem 7 using the following Theorem 8 (mostly to simplify notation) applied to

$$C = \begin{bmatrix} 0 & A \\ -A^\top & 0 \end{bmatrix}, \qquad d = \begin{bmatrix} b \\ -c \end{bmatrix}. \tag{57}$$

The proof of Theorem 8 does *not* use strong LP duality.

Theorem 8. *Let $C \in \mathbb{R}^{k\times k}$ such that $C = -C^\top$, and $d \in \mathbb{R}^k$. Let $(z,w) = (z^*,w^*) \in \mathbb{R}^k \times \mathbb{R}$ be a basic feasible solution that minimizes w subject to*

$$Cz - \mathbf{1}w \le d, \quad d^\top z - w \le 0, \quad z \ge \mathbf{0}, \quad w \ge 0, \tag{58}$$

and let $M \in \mathbb{R}$ with

$$\mathbf{1}^\top z^* + 1 \le M\,. \tag{59}$$

Consider the zero-sum game

$$D_M = \begin{bmatrix} C & -d \\ d^\top & 0 \\ \mathbf{1}^\top & -M \end{bmatrix} \tag{60}$$

with game value v. Then $v \ge 0$ and

(a) *$v = 0$ if and only if $w^* = 0$. If $w^* = 0$, let $t = \frac{1}{\mathbf{1}^\top z^* + 1}$ and $z = z^* t$. Then (z,t) is a min-max strategy and $(z,t,0)$ is a max-min strategy for D_M.*

(b) *Suppose $v > 0$. Then every max-min strategy (q,r,s) of D_M fulfills $r = 0$, $s = v$, and*

$$Cq \le \mathbf{0}, \quad d^\top q < 0, \tag{61}$$

which proves that there is no $z \ge \mathbf{0}$ with $Cz \le d$.

Proof. In the following, letters (and their decorated versions) q and z denote vectors in $\mathbb{R}^k$, and r,s,t,u,v,w denote scalars in $\mathbb{R}$.

The system (58) is feasible, for example with $z = \mathbf{0}$ and large enough w, and w is bounded from below, so that (58) has an optimal basic feasible solution (z^*,w^*) by Lemma 2(c).

We have $v \geq 0$, because the game matrix

$$D = \begin{bmatrix} C & -d \\ d^\top & 0 \end{bmatrix} \tag{62}$$

is skew-symmetric and has game value 0, so by adopting any max-min strategy for D and not playing the last row in D_M the row player will get at least 0.

For the "if" part of case (a), if $w^* = 0$ then with $t = \frac{1}{\mathbf{1}^\top z^* + 1}$ and $z = z^* t$ we have $\mathbf{1}^\top z - Mt \leq -t < 0$ by (59). This shows that (z, t) is a min-max strategy and $(z, t, 0)$ a max-min strategy for D_M, and $v = 0$. For the "only if" part, if $v = 0$ then a min-max strategy (z', t) for D_M requires $t > 0$ to get a nonpositive cost in the last row, and then $z = z' \frac{1}{t}$ solves (58) with $w = 0$.

To show (b), let $v > 0$. The following properties hold for any optimal strategies of D_M. The min-max value of D_M with min-max strategy (z, t) is the smallest real number v such that

$$\begin{array}{rcrcl} Cz & - & dt & \leq & \mathbf{1}v \\ d^\top z & & & \leq & v \\ \mathbf{1}^\top z & - & Mt & \leq & v \\ \mathbf{1}^\top z & + & t & = & 1 \\ z & , & t & \geq & \mathbf{0}\,. \end{array} \tag{63}$$

The max-min value of D_M with max-min strategy (q, r, s) is the largest v such that

$$\begin{array}{lcrcrcl} q^\top C & + & rd^\top & + & s\mathbf{1}^\top & \geq & v\mathbf{1}^\top \\ q^\top(-d) & & & - & sM & \geq & v \\ q^\top \mathbf{1} & + & r & + & s & = & 1 \\ q & , & r & , & s & \geq & \mathbf{0}\,. \end{array} \tag{64}$$

Then $0 < s < 1$ because if $s = 0$ then $(q, r, 0)$ would be a max-min strategy for the symmetric game D in (62) with max-min value $v > 0$ which is not possible, and if $s = 1$ then the last row of D_M alone would be a max-min strategy for D_M, but that row has the negative entry $-M$.

Because $s > 0$, we have $\mathbf{1}^\top z - Mt = v$ in (63), and, using $\mathbf{1}^\top z = 1 - t$,

$$v = 1 - (M + 1)t\,. \tag{65}$$

We show that $v \leq s$. If $v > s$, then by (64),

$$\begin{array}{lcrcl} q^\top C & + & rd^\top & \geq & (v - s)\mathbf{1}^\top \\ q^\top(-d) & & & \geq & v + Ms \end{array} \tag{66}$$

which would define a max-min strategy $(q\frac{1}{1-s}, \frac{r}{1-s})$ with positive max-min value for the symmetric game D, a contradiction.

Hence, $0 < v \leq s < 1$ and by (65),

$$t = \frac{1 - v}{M + 1} > 0\,. \tag{67}$$

Then (63) implies

$$Cz\tfrac{1}{t} \le d + \mathbf{1}\tfrac{v}{t}, \quad d^\top z\tfrac{1}{t} \le \tfrac{v}{t}, \tag{68}$$

and therefore

$$w^* \le \frac{v}{t} = \frac{v(M+1)}{1-v}. \tag{69}$$

In order to show that every max-min strategy (q, r, s) for D_M is of the form $(q, 0, v)$, we will in essence use weak duality. We write $s = u + v$ with $u \ge 0$ (we know $s \ge v$) and let v in (64) be *fixed* where we now in essence maximize u. That is, we consider the constraints

$$\begin{array}{llllll} q^\top C & + rd^\top & + u\mathbf{1}^\top & \ge & \mathbf{0}^\top \\ q^\top(-d) & & - uM & \ge & v(M+1) \\ q^\top\mathbf{1} & + r & + u & = & 1-v \\ q & , r & , u & \ge & \mathbf{0}. \end{array} \tag{70}$$

They have solutions with the current max-min strategy (q, r, s) and $u = s - v$. We use that $Cz^* - d - \mathbf{1}w^* \le \mathbf{0}$ and $d^\top z^* - w^* \le 0$ in (58), and $-1 \ge \mathbf{1}^\top z^* - M - w^*$ by (59), and $v(M+1) - (1-v)w^* \ge 0$ by (69) in the following chain of inequalities, obtained by multiplying the first inequality in (70) by z^*, the second by 1, and the equation by $-w^*$ and summing up:

$$\begin{aligned} 0 &\ge -u \\ &\ge q^\top(Cz^* - d - \mathbf{1}w^*) + r(d^\top z^* - w^*) + u(\mathbf{1}^\top z^* - M - w^*) \\ &\ge v(M+1) - (1-v)w^* \ge 0. \end{aligned} \tag{71}$$

Hence, all inequalities hold as equalities, in particular

$$w^* = \frac{M+1}{1/v - 1} \tag{72}$$

and $u = 0$. This shows $s = v$ in any solution (q, r, s) to (64). In addition, $q^\top C + rd^\top \ge \mathbf{0}^\top$, that is, $Cq - dr \le \mathbf{0}$, and $q^\top d \le -v(M+1) < 0$. The skew-symmetry of C implies $q^\top Cq = (q^\top Cq)^\top = q^\top C^\top q = -q^\top Cq$ and therefore $q^\top Cq = 0$, for any q. If we had $r > 0$ then $Cq\frac{1}{r} \le d$ and $0 = q^\top Cq\frac{1}{r} \le q^\top d < 0$, a contradiction, which shows $r = 0$. This shows $Cq \le \mathbf{0}$ and $d^\top q < 0$ as claimed in (61). In turn, this shows that there is no $z \ge \mathbf{0}$ with $Cz \le d$, because this would imply $0 \le z^\top(-Cq) = z^\top C^\top q = q^\top Cz \le q^\top d < 0$. □

Proof of Theorem 7. We apply Theorem 8 to C and d in (57). Let v be the value of the game D_M. Then by Theorem 8(a), $v = 0$ implies feasibility and optimality of the LPs (1) and (2). Conversely, suppose that (1) and (2) are feasible. Then $v = 0$, because if $v > 0$ then (61) contradicts feasibility. This shows part (a) in Theorem 7, and also part (b) via (72).

To show Theorem 7(c), suppose first that $\beta = 1$, that is, all entries of A, b, c are integers. The system (58) has ℓ rows, and written as equations with slack variables has entries from A, b, c or $0, 1, -1$. Any basic solution is uniquely determined by the basis matrix, where each variable is the quotient of two determinants where the denominator is at least 1 and the numerator bounded in absolute value by $\ell!\,\alpha^\ell$. Only the ℓ basic variables can be nonzero, so that we can choose $M = \ell!\,\ell\alpha^\ell + 1$ by (59). See also [20, p. 30] or [1, p. 172]; I did not find the next description, clearly standard, if $\beta > 1$.

If $\beta > 1$, multiply each column of $\binom{C}{d^\top}$ and $\binom{d}{0}$ in (58) with the least common multiple of the denominators in that column, called the *scale factor* σ_j for that column j (with $j = 0$ if the column is d). This gives an integral system where each basic solution has to be changed by multiplying each variable in column j with its scale factor σ_j and dividing it by σ_0 to give the solution to the original system. Each entry of the integral system has been multiplied by at most β^ℓ (this is an overestimate because each column of C in (57) has m or n zeros), so we have to replace α^ℓ by $\alpha^\ell(\beta^\ell)^\ell$, with the extra factor β^ℓ for the re-scaling of the variables, which shows (55). The number of bits to represent M is its binary logarithm, which is polynomial in ℓ and in the bit-sizes of α and β, and hence in the bit-size of A, b, c. □

9 Discussion and related work

Because Dantzig's proof in Theorem 5 works for generic LPs, a first question is if genericity can be achieved by perturbing a given LP. However, this may alter its feasibility. For example, consider the LP of maximizing x_2 subject to $x_2 \le 1$, $x \ge \mathbf{0}$, $x \in \mathbb{R}^2$. The corresponding game B in (35) has an all-zero row and column, which when played as an optimal pure-strategy pair does not play the last column ($t = 0$). The LP has optimal solutions $(x_1, 1)$ for any $x_1 \ge 0$. However, maximizing the perturbed objective function $\varepsilon x_1 + x_2$ (for some small $\varepsilon > 0$) with the same constraints gives an unbounded LP. Hence, there is no obvious way of perturbing the LP to make Dantzig's proof generally applicable.

The closest related works to ours are Adler [1] and Brooks and Reny [3]. We continue here our discussion from the introduction.

A main goal of [1] is to reduce the computational problem of solving an LP (in the sense of finding an optimal solution or proving there is none) to the problem of solving a zero-sum game by means of a strongly polynomial-time reduction. Adler considers the feasibility problem with equalities, that is, to find $x \in \mathbb{R}^n$ such that

$$Ax = b, \qquad x \ge \mathbf{0}\,, \tag{73}$$

for an $m \times n$ matrix A, or to show that no such x exists. He constructs a symmetric game with $m+n+3$ rows and columns. An optimal strategy to that game produces either a solution to (73), or a vector $y \in \mathbb{R}^m$ such that $y^\top A \ge \mathbf{0}^\top$ and $y^\top b < 0$ (which by (13) shows that (73) is infeasible), or some $\tilde{x} \ne \mathbf{0}$ such that $A\tilde{x} = \mathbf{0}$ and

$\tilde{x} \geq \mathbf{0}$. The first two cases answer whether (73) is feasible or not. In the third case, $Ax = b$ is replaced by an equivalent system where the variables x_S in the support S (written J^+ in [1]) of $\tilde{x}$ are eliminated. In a solution to that equivalent system, the variables x_S can be substituted back, and irrespective of their sign can be replaced by $x_S + \tilde{x}_S\alpha$ for sufficiently large α to find a solution to (73). (The latter step is implicit in the claim (10b) of [1, p. 173] and attributed to [8] but without a page number; I could not find it and found these computations the hardest to follow.) Repeating this at most n times, with corresponding calls to solving a zero-sum game, then answers the feasibility problem. This is known as a "Cook-type" reduction. It also leads to a proof of Tucker's Theorem from Gordan's Theorem in [1, section 4], which we have given in a more direct way in Theorem 6.

A different "Karp-type" reduction uses only a single step from the feasibility problem (73) to solving a zero-sum game, by adding a constraint $\mathbf{1}^\top x \leq M$ where M is large enough to not affect feasibility. If the entries of A and b are algebraic numbers (in particular, integers), they determine an explicit bound on M of polynomial encoding size [1, p. 172].

We have done the same in Theorem 7 above. However, our game B_M is directly derived from the original LPs (1) and (2) defined by inequalities (also first considered by Adler) with a single extra row added to Dantzig's original game B in (35), rather than converting them to equalities as in (73) (with a new, larger matrix A) and then back to inequalities to construct an even larger symmetric game. As an additional, new property, Theorem 7(b) shows that a max-min strategy of B_M provides a certificate that the LPs are infeasible if that is the case.

Brooks and Reny [3] prove the following theorem. For any matrix D, let $\|D\|$ be the maximum absolute value of its entries.

Theorem 9 (Brooks and Reny [3])**.** *Consider the LPs* (1) *and* (2)*. Let r be the rank of the matrix*

$$\hat{A} = \begin{bmatrix} 0 & -A^\top \\ A & 0 \\ -c^\top & b^\top \end{bmatrix} \tag{74}$$

and let

$$\alpha = 2r^2 \max\{\|b\|, \|c\|\} \max_W \|W^{-1}\| + 1, \tag{75}$$

where the second maximum is taken over all invertible sub-matrices W of $\hat{A}$. Then for the game P with $n + m + 1$ rows and columns

$$P = \begin{bmatrix} 0 & -\alpha A^\top & \mathbf{0} \\ \alpha A & 0 & \mathbf{0} \\ -\alpha c^\top & \alpha b^\top & 0 \end{bmatrix} + \begin{bmatrix} c \\ -b \\ 0 \end{bmatrix} \mathbf{1}^\top \tag{76}$$

either

(a) *the value of P is zero, and then for a min-max strategy (x^*, y^*, t^*) of P, a pair of optimal solutions to the LPs* (1) *and* (2) *is $(x^*\alpha, y^*\alpha)$, or*

(b) *the value of P is positive, and then any max-min strategy (x, y, t) of P fulfills $Ax \leq \mathbf{0}$, $x \geq \mathbf{0}$, $A^\top y \geq \mathbf{0}$, $y \geq \mathbf{0}$, and $c^\top x > b^\top y$, which shows that at least one LP is infeasible.*

The main effect of the definition of P is that for any min-max strategy (x^*, y^*, t^*) with min-max value v, we have

$$\begin{array}{rcl} -\ \alpha A^\top y^* & \leq & -c + \mathbf{1}v \\ \alpha A x^* \qquad\qquad & \leq & b + \mathbf{1}v \\ -\alpha c^\top x^* + \ \alpha b^\top y^* & \leq & v \end{array} \tag{77}$$

with *constant* right-hand sides $-c$ and b rather than these being scaled by t^*. The number α is similar to the bound M in Theorem 8 and (59), because if the LPs (1) and (2) have feasible solutions, then with x^* and y^* as in Theorem 9(a), they have feasible solutions $x^*\alpha, y^*\alpha$ with $\mathbf{1}^\top x^*\alpha + \mathbf{1}^\top y^*\alpha \leq \alpha$, as noted by Brooks and Reny [3, Remark 7]. If the value of P in (76) is positive, then any max-min strategy (x, y, t) in Theorem 9(b) proves the infeasibility of at least one of the LPs just as in (52) in Theorem 7.

Given the constraints (77), the definition of P can be seen as "canonical" as claimed by Brooks and Reny, although one could also call it "proof-induced". From the viewpoint of using this game, it has the disadvantage that all entries of A are multiplied by the large number α, and P is a full matrix and no longer half-empty, with zero entries replaced by the rows of $-c$ and b. In contrast, in our matrix B_M in Theorem 7 the large number M appears in a single place, and the zero entries remain. The game B_M also naturally extends Dantzig's original game.

In summary, it seems that proving LP duality from the minimax theorem requires quite a bit of linear algebra, most concisely in our relatively short proof of Theorem 6. We show an elegant use of linear algebra in the next, final section.

10 Minimally infeasible sets of inequalities

We conclude this article with a short elementary proof of the Lemma of Farkas in its inequality-only version (15) due to Conforti, Di Summa, and Zambelli [6]. The main trick is to state the *minimal* infeasibility of these inequalities in terms of infeasibility of the corresponding equalities, which is canonically proved by induction. The second step is to apply the linear algebra Lemma 1 to the infeasible equalities to obtain the required vector y in (15).

A set of linear equations and inequalities is called infeasible if it has no solution, and *minimally* infeasible if omitting any one equation or inequality makes it feasible. The following proofs of theorem 2.1 and lemma 2.1 of [6], in simplified notation, show (15) based on minimally infeasible sets of inequalities.

Theorem 10 (Conforti, Di Summa, and Zambelli [6])**.** *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$ and let $a_1, \ldots, a_m$ be the rows of A. Suppose the system $Ax \leq b$ is minimally infeasible.*

(i) *Then the system $Ax = b$ is minimally infeasible.*

(ii) *Reversing any inequality $a_i x \le b_i$ in $Ax \le b$ creates a feasible system:*

$$\forall\, i \in [m]\ \exists\, x^{(i)} \in \mathbb{R}^n\ :\ a_i x^{(i)} > b_i\,,\quad \forall\, k \in [m]-\{i\}\ :\ a_k x^{(i)} = b_k\,. \tag{78}$$

Proof. We prove that for any $R \subseteq [m]$ the constraints

$$a_i x = b_i \quad (i \in R), \qquad a_i x \le b_i \quad (i \in [m]-R) \tag{79}$$

are minimally infeasible. The proof is by induction on $|R|$. For $|R| = 0$ condition (79) holds by assumption. Suppose it holds for all R up to a certain size $|R|$. If $R = [m]$ then the proof of (i) is complete, so let $h \notin R$, where we want to show that

$$a_i x = b_i \quad (i \in R), \qquad a_h x = b_h\,, \qquad a_i x \le b_i \quad (i \in [m]-R-\{h\}) \tag{80}$$

is minimally infeasible. The system (80) is infeasible (because $Ax \le b$ is infeasible), so we have to prove that omitting any constraint $a_j x = b_j$ or $a_j x \le b_j$ for $j \in [m]$ produces a feasible system. This is clearly the case if $j = h$, or if $j \in R$ by applying the inductive hypothesis to $R \cup \{h\} - \{j\}$, so let $j \notin R$. The constraints (79) for $i \ne h$ and $i \ne j$ have solutions $x^{(h)}$ and $x^{(j)}$, respectively, with

$$\begin{array}{llll} a_i x^{(h)} = b_i & (i \in R), & a_i x^{(h)} \le b_i \quad (i \in [m]-R-\{h\}), & a_h x^{(h)} > b_h \\ a_i x^{(j)} = b_i & (i \in R), & a_i x^{(j)} \le b_i \quad (i \in [m]-R-\{j\}). & \end{array} \tag{81}$$

If $a_h x^{(j)} = b_h$ then $x^{(j)}$ is a feasible solution to (80) with row $a_j x \le b_j$ omitted. Otherwise $a_h x^{(j)} < b_h$, and a suitable convex combination of $x^{(j)}$ and $x^{(h)}$ is such a solution because $a_h x^{(h)} > b_h$. This completes the induction.

Condition (ii) is an immediate consequence of (i): Let $i \in [m]$. Because $Ax = b$ is minimally infeasible, there is some $x^{(i)} \in \mathbb{R}^n$ such that $a_k x^{(i)} = b_k$ for all $k \ne i$ and $a_i x^{(i)} \ne b_i$, where $a_i x^{(i)} < b_i$ would imply that $Ax \le b$ is feasible, hence $a_i x^{(i)} > b_i$. □

Proof of (15) using Theorem 10. The direction "⇐" in (15) is immediate (and will be used below). To prove "⇒", assume that $Ax \le b$ is infeasible, and (by dropping sufficiently many rows from these inequalities, whose components of y will be set to zero) that $Ax \le b$ is minimally infeasible. Denote the number of rows of this minimally infeasible system again by m. By Theorem 10, $Ax = b$ is minimally infeasible. By Lemma 1, there is some $y \in \mathbb{R}^m$ so that $y^\top A = \mathbf{0}^\top$ and $y^\top b = -1$. It remains to show that $y \ge \mathbf{0}$. If not, suppose that $I = \{i \in [m] \mid y_i < 0\} \ne \emptyset$. Define the system $A'x \le b'$ as $Ax \le b$ with the rows in I reversed, that is, each of its rows $a'_i x \le b'_i$ means $-a_i x \le -b_i$ if $i \in I$ and $a_i x \le b_i$ otherwise. Take some $i \in I$, and $x^{(i)}$ as in (78) in Theorem 10(ii). Then $A'x^{(i)} \le b'$. On the other hand, define $w \in \mathbb{R}^m$ by $w_k = |y_k|$ for $k \in [m]$. Then $w \ge \mathbf{0}$ and $w^\top A' = y^\top A = \mathbf{0}^\top$ and $w^\top b' = y^\top b = -1$. But this contradicts $0 = w^\top A' x^{(i)} \le w^\top b'$. Hence, $I = \emptyset$ and therefore $y \ge \mathbf{0}$ as required. □

The proof of Theorem 10 is canonical and easy to reconstruct. As for proving the Lemma of Farkas, in the same version (15), perhaps the most natural and elementary proof is "projection" or Fourier-Motzkin elimination (see Schrijver [22, p. 155f] and references). It expresses the constraints in $Ax \leq b$ in terms of x_1 by dividing each row by the coefficient of x_1 when it is nonzero, which reverses the inequality when the coefficient is negative. This induces mutual bounds among the other linear terms in $x_2, \ldots, x_n$ and eliminates x_1. This elimination is then iterated (and may lead to an exponential increase in the number of constraints). See Kuhn [15] and Tao [24, p. 180] for deriving (15) in this way.

Acknowledgments

I thank Ahmad Abdi, Ben Brooks, Phil Reny, Giacomo Zambelli, and anonymous referees for helpful comments and discussions, and Sylvain Sorin for alerting me to the works of Loomis [17] and Ville [26]. I thank Ilan Adler for great help in proving Theorem 8 without assuming strong LP duality.

References

[1] Adler, I. (2013). The equivalence of linear programs and zero-sum games. *International Journal of Game Theory*, 42(1), 165–177.

[2] Binmore, K. (2004). Guillermo Owen's proof of the minimax theorem. In: G. Gambarelli (ed.), *Essays in Cooperative Games*, 19–23. Kluwer, Boston.

[3] Brooks, B. and Reny, P. J. (2022). A canonical game – 75 years in the making – showing the equivalence of matrix games and linear programming. Version of 5 August 2022; first version 26 August 2021. Working Paper, Department of Economics, University of Chicago.

[4] Broyden, C. G. (2001). On theorems of the alternative. *Optimization Methods and Software*, 16(1-4), 101–111.

[5] Carathéodory, C. (1911). Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen. *Rendiconti del Circolo Matematico di Palermo (1884–1940)*, 32(1), 193–217.

[6] Conforti, M., Di Summa, M., and Zambelli, G. (2007). Minimally infeasible set-partitioning problems with balanced constraints. *Mathematics of Operations Research*, 32(3), 497–507.

[7] Dantzig, G. B. (1951). A proof of the equivalence of the programming problem and the game problem. In: T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation* (Cowles Commission Monograph No. 13), 330–335. Wiley, New York.

[8] Dantzig, G. B. (1963). *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ.

[9] Dantzig, G. B. (1982). Reminiscences about the origins of linear programming. *Operations Research Letters*, 1(2), 43–48.

[10] Farkas, J. (1902). Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik*, 124(1), 1–27.

[11] Gale, D. (1960). *The Theory of Linear Economic Models*. McGraw-Hill, New York.

[12] Gale, D., Kuhn, H. W., and Tucker, A. W. (1950). On symmetric games. In: H. W. Kuhn and A. W. Tucker (eds.), *Contributions to the Theory of Games, Vol. I*, volume 24 of *Annals of Mathematics Studies*, 81–87. Princeton University Press, Princeton, NJ.

[13] Goldman, A. J. and Tucker, A. W. (1956). Theory of linear programming. In: H. W. Kuhn and A. W. Tucker (eds.), *Linear Inequalities and Related Systems*, volume 38 of *Annals of Mathematics Studies*, 53–97. Princeton University Press, Princeton, NJ.

[14] Gordan, P. (1873). Ueber die Auflösung linearer Gleichungen mit reellen Coefficienten. *Mathematische Annalen*, 6(1), 23–28.

[15] Kuhn, H. W. (1956). Solvability and consistency for linear equations and inequalities. *The American Mathematical Monthly*, 63(4), 217–232.

[16] Laraki, R., Renault, J., and Sorin, S. (2019). *Mathematical Foundations of Game Theory*. Springer.

[17] Loomis, L. H. (1946). On a theorem of von Neumann. *Proceedings of the National Academy of Sciences of the United States of America*, 32(8), 213–215.

[18] Luce, R. D. and Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*. John Wiley & Sons, New York.

[19] Owen, G. (1967). Communications to the editor: An elementary proof of the minimax theorem. *Management Science*, 13(9), 765–765.

[20] Papadimitriou, C. H. and Steiglitz, K. (1982). *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Englewood Cliffs, NJ.

[21] Raghavan, T. E. S. (1994). Zero-sum two-person games. In: R. Aumann and S. Hart (eds.), *Handbook of Game Theory*, volume 2, 735–768. North-Holland, Amsterdam.

[22] Schrijver, A. (1986). *Theory of Linear and Integer Programming*. John Wiley & Sons, Chichester, UK.

[23] Stiemke, E. (1915). Über positive Lösungen homogener linearer Gleichungen. *Mathematische Annalen*, 76(2), 340–342.

[24] Tao, T. (2008). *Structure and Randomness: Pages from Year One of a Mathematical Blog*. American Mathematical Society, Providence, RI.

[25] Tucker, A. W. (1956). Dual systems of homogeneous linear relations. In: H. W. Kuhn and A. W. Tucker (eds.), *Linear Inequalities and Related Systems*, volume 38 of *Annals of Mathematics Studies*, 1–18. Princeton University Press, Princeton, NJ.

[26] Ville, J. (1938). Sur la théorie générale des jeux ou intervient l'habileté des joueurs. In: E. Borel et al. (ed.), *Traité du Calcul des Probabilités et de ses Applications, Tome IV*, 105–113. Gauthier-Villars, Paris.

[27] von Neumann, J. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100(1), 295–320.

[28] von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior, 2nd ed*. Princeton University Press, Princeton, NJ.

[29] von Stengel, B. (2022). *Game Theory Basics*. Cambridge University Press, Cambridge, UK.