Here I give a proof of the "Fundamental Theorem of Linear Programming," as well as a discussion of its significance. Note that a statement of the theorem and a discussion of its geometric meaning can be found in Chapter 3 of the Vanderbei text.

Theorem. *Consider the linear program:*

$$\begin{aligned} \text{maximize:} \quad & \xi = c^t x \\ \text{subject to:} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{1}$$

where c is a nonzero vector of length n and A is an $m \times n$ matrix, $m \leq n$, with rank m. If the problem (1) admits a solution x — that is, if there exists an optimal feasible vector x — then there is an optimal feasible vector $\tilde{x} \in \mathbb{R}^n$ such that $\tilde{x}$ has exactly m nonzero entries.

Before we launch into a proof, let's discuss the significance of the theorem. Suppose that the L.P. (1) has a solution. Then we know that it has a solution with exactly m nonzero entries. The reason this is so useful, is that if we pick m entries $i_1, \ldots, i_m$ to be nonzero, then the constraints $Ax = b$ determine the rest of the entries.

That is, there is at most one solution $\tilde{x}$ to the $m \times n$ system of equations

$$A\tilde{x} = b \tag{2}$$

such that the entries $\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_m}$ are the only nonzero entries (prove this). Note, that its quite possible that there are NO solutions. Moreover, its entirely possible that there is a solution $\tilde{x}$, but $\tilde{x}$ does not satisfy the other constraints; i.e., such that

$$\tilde{x} \geq 0 \tag{3}$$

does not hold. The important point is that there is AT MOST ONE such solution $\tilde{x}$.

This means that we already have an algorithm for solving the L.P. problem (1). Namely, we consider each of the

$$\begin{pmatrix} n \\ m \end{pmatrix} = \frac{n!}{(n-m)!m!} \tag{4}$$

ways of choosing m of the n indices to be nonzero. Each such choice gives us at most one possible feasible vector, which we obtain by trying to solve the constraint equation $Ax = b$. We simply compute $c^t x$ for each of those vectors. The largest such value is the maximum of the objective function.

Proof of the Theorem: We will start by showing that if x is an feasible vector with $k > m$ nonzero entries, then we can produce a feasible vector $\tilde{x}$ with $k-1$ nonzero entries. To do this, we first note that given a set of $k > m$ indices $I = \{i_1, \ldots, i_k\}$, we can find a solution y to the equation

$$Ay = 0 \tag{5}$$

such that $y_i = 0$ for $i \notin I$. This follows from the linear dependence of the columns $i_1, \ldots, i_k$ of the matrix A. Now set

$$x^\epsilon = x - \epsilon y. \tag{6}$$

Note that by our choice of Y, $Ax^\epsilon = Ax = b$, so any such vector satisfies the first constraint equation. Our goal is to find a feasible vector with one fewer nonzero entry than x. To do that, we will look at each coordinate of x^ϵ separately. Let ϵ_i satisfy

$$x_i^\epsilon - \epsilon_i y_i = 0 \tag{7}$$

for $i \in I$. We can, without loss of generality, assume that at least one ϵ_i is positive (why?). Let ϵ be the smallest positive ϵ_i. Then it is easy to see that the vector x^ϵ satisfies the constraints $x^\epsilon > 0$ AND it has at most $k-1$ nonzero entries.

Now it is clear that if there is a feasible vector x, then there exists one with at most m nonzero entries (we just repeatedly apply the argument from the preceding paragraph until we get down to at most m nonzero entries). We need to make one additional observation to prove the theorem.

Let x be an optimal feasible vector — that is, a solution to the problem. Consider the vector

$$x^\epsilon = x - \epsilon y \tag{8}$$

where y is a solution to $Ay = 0$ and $y_i \neq 0$ only if $x_i \neq 0$. By a simple continuity argument, we see that for any ϵ of small enough magnitude — either positive or negative — the vector x^ϵ satisfies the constraints

$$\begin{aligned} Ax^\epsilon &= 0 \\ x^\epsilon &\geq 0. \end{aligned} \tag{9}$$

Since we assumed that x is an optimal feasible vector — that is, that x maximizes $c^t x$ — it follows that

$$c^t y = 0. \tag{10}$$

Otherwise, we could make $c^t(x - \epsilon y)$ larger while still satisfying the constraints by choosing a epsilon (either positive or negative) of small enough magnitude.

This means that if we start with an optimal feasible vector x with $k > m$ nonzero entries and apply the argument of the first paragraph, the result will be a feasible vector $\tilde{x}$ with at most $k-1$ nonzero entries such that

$$c^t \tilde{x} = c^t x. \tag{11}$$

In other words, $\tilde{x}$ will again be an optimal feasible vector. The theorem now follows by repeated application of the procedure of the first paragraph to an optimal feasible vector.