Math 330 Final Report

Peter Prastakos

December 10, 2020

1 Introduction

The paper "Mechanism Design via Optimal Transport" by Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos is broadly concerned with the problem of designing a revenue-optimal auction for selling n items to m bidders whose valuations are drawn from known prior distributions. More specifically, their goal in the paper is to develop a general optimization framework for obtaining closed-form descriptions of the revenue-optimal mechanisms for selling n goods to a single additive whose values for the goods are drawn independently from probability distributions with given probability density functions.

As a core element of many of the results in the paper, the authors rely on Strassen's theorem regarding the stochastic dominance of measures, discussed in section 2. However, to avoid the cumbersome process of checking stochastic dominance, they develop an alternate condition which implies stochastic dominance, which they use for all of their 2-item auction applications (e.g. obtaining a closed-form description of the optimal mechanism for two independent exponentially distributed items). The proof of the sufficiency of this alternate condition is of independent interest to measure theory and is the primary focus of this report. It is mostly self-contained and requires no prior exposure to mechanism design.

2 Preliminaries and Strassen's Theorem

Throughout the report, whenever we use the term *measurable* or *measure*, we will use them with respect to the Borel σ-algebra. Additionally, we will use $\Gamma(\mu^{\mathcal{X}}, \nu^{\mathcal{Y}})$ to refer to the set of all measures on $\mathcal{X} \times \mathcal{Y}$ with marginal measures $\mu^{\mathcal{X}}$ and $\nu^{\mathcal{Y}}$, respectively. That is, for any function $\gamma \in \Gamma(\mu^{\mathcal{X}}, \nu^{\mathcal{Y}})$, we have that $\gamma(A, \mathcal{Y}) = \mu^{\mathcal{X}}(A)\nu^{\mathcal{Y}}(\mathcal{Y})$ for all measurable $A \subseteq \mathcal{X}$ and $\gamma(\mathcal{X}, B) = \mu^{\mathcal{X}}(\mathcal{X})\nu^{\mathcal{Y}}(B)$ for all measurable $B \subseteq \mathcal{Y}$.

We start with some necessary definitions.

Definition 1 (Partial ordering $\preceq$)**.** *For any two vectors $a, b \in \mathbb{R}^n_{\geq 0}$, we denote that $a \preceq b$ iff $a_i \leq b_i \ \forall i$. Using this ordering, we state the following additional definitions:*

A function $f : \mathbb{R}^n_{\geq 0} \mapsto \mathbb{R}$ is increasing *if $a \preceq b \Rightarrow f(a) \leq f(b)$.*

A set $S \subset \mathbb{R}^n$ is increasing *if $a \in S$ and $a \preceq b$ implies $b \in S$, and* decreasing *if $a \in S$ and $b \preceq a$ implies $b \in S$.*

Definition 2 (Stochastic dominance of measures)**.** *For two measures α, β on $\mathbb{R}^n_{\geq 0}$, we say that α* stochastically dominates *β (with respect to the partial order $\preceq$), denoted $\beta \preceq \alpha$, if $\int_{\mathbb{R}^n_{\geq 0}} f d\alpha \geq \int_{\mathbb{R}^n_{\geq 0}} f d\beta$ for all increasing bounded measurable functions f. Similarly, if g, h are density functions, $h \preceq g$ if $\int_{\vec{x} \in \mathbb{R}^n_{\geq 0}} f(\vec{x}) h(\vec{x}) d\vec{x} \leq \int_{\vec{x} \in \mathbb{R}^n_{\geq 0}} f(\vec{x}) g(\vec{x}) d\vec{x}$ for all increasing bounded measurable functions f.*

We can now state Strassen's theorem for the partial ordering $\preceq$ just defined.

Theorem 1 (Strassen)**.** *If α and β are probability measures on $\mathbb{R}^n_{\geq 0}$ and α stochastically dominates β with respect to the partial order $\preceq$ (i.e. $\beta \preceq \alpha$), then there exists a probability measure $\hat{\gamma} \in \Gamma(\alpha, \beta)$ on $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}$ with marginals α and β respectively such that $\hat{\gamma}(\{(a, b) : b \preceq a\}) = 1$, where the set $\{(a, b) : b \preceq a\}$ is closed.*

3 An Equivalent Condition for Stochastic Dominance

In order to use Strassen's theorem, stochastic dominance needs to hold. However, since such a condition can be hard to verify, the authors use an equivalent condition, which we will prove in Lemma 2. Before the statement and proof of the lemma, let us start with a few claims and definitions.

Claim 1. *Let α, β be finite measures on $\mathbb{R}^n_{\geq 0}$. Then $\beta \preceq \alpha$ if and only if $\beta(A) \leq \alpha(A)$ for all increasing measurable sets A.*

Proof. Without loss of generality, assume that $\alpha(\mathbb{R}^n_{\geq 0}) = 1$.

($\Rightarrow$) If α stochastically dominates β then by definition, $\int_{\mathbb{R}^n_{\geq 0}} f d\alpha \geq \int_{\mathbb{R}^n_{\geq 0}} f d\beta$ for all increasing bounded measurable functions f. Then let $f = \mathbb{1}\{a \in A\}$. Clearly, this is a bounded measurable function, and it is increasing because if $a \in A$ (i.e. $f(a) = 1$) and $a \preceq b$, then since A is an increasing set, $b \in A$ so $f(b) = 1 \geq f(a)$. Then we note simply that $\int_{\mathbb{R}^n_{\geq 0}} \mathbb{1}\{a \in A\} d\alpha = \alpha(A)$ and $\int_{\mathbb{R}^n_{\geq 0}} \mathbb{1}\{a \in A\} d\beta = \beta(A)$ to conclude that $\beta(A) \leq \alpha(A)$ for all increasing measurable sets A.

($\Leftarrow$) Suppose by way of contradiction that $\beta(A) \leq \alpha(A)$ for all increasing measurable sets A but α does *not* stochastically dominate β. Then, there exists an increasing, bounded, measurable function f such that

$$\int f d\beta - \int f d\alpha > 2^{-k+1} \text{ for some } k \in \mathbb{N}.$$

Without loss of generality, we can assume that f is nonnegative by adding the constant of $f(0)$ to all values. We now define the function $\tilde{f}$ by point-wise rounding f upwards to the nearest multiple of 2^{-k}. Clearly, $\tilde{f}$ is also an increasing, measurable, and bounded. Additionally, we have that

$$\int \tilde{f} d\beta - \int \tilde{f} d\alpha \geq \int f d\beta - \int f d\alpha - 2^{-k} > 2^{-k+1} - 2^{-k} > 0.$$

Thus, we have that

$$\int \tilde{f} d\beta > \int \tilde{f} d\alpha.$$

We now design the increasing sets such that we arrive at a contradiction. This can be done by decomposing $\tilde{f}$ into the weighted sum of indicator functions of increasing sets. Let $\{r_1, ..., r_m\}$ be the set of all values taken by $\tilde{f}$, where $r_1 > r_2 > ... > r_m$. We notice that, for any $s \in \{1, ..., m\}$, the set $A_s = \{z : \tilde{f}(z) \geq r_s\}$ is measurable, since $\tilde{f}$ is measurable, and increasing. The latter is because if $a \in A_s$ and $a \preceq b$, then since $\tilde{f}$ is increasing, $\tilde{f}(a) \leq \tilde{f}(b)$, so $\tilde{f}(b) \geq r_s$ and hence $b \in A_s$.

Therefore, we can write

$$\tilde{f} = \sum_{s=1}^{m} (r_s - r_{s-1}) \mathbb{1}\{s \in A_s\},$$

setting $r_0 = 0$. We now apply the condition that, since A_s is an increasing measurable set $\forall s \in [m]$,

$$\beta(A_s) \leq \alpha(A_s).$$

We now compute

$$\int \tilde{f} d\beta = \sum_{s=1}^{m} (r_s - r_{s-1}) \beta(A_s) \leq \sum_{s=1}^{m} (r_s - r_{s-1}) \alpha(A_s) = \int \tilde{f} d\alpha.$$

But this is a contradiction to the fact that we had $\int \tilde{f} d\beta > \int \tilde{f} d\alpha$ above.

■

We now proceed to state some definitions.

Definition 3. *For any $z \in \mathbb{R}^n_{\geq 0}$, we define the* base rooted at z *to be*

$$B_z \triangleq \{z' : z \preceq z'\},$$

the minimal increasing set containing z.

We denote Q_k as the set of points in $\mathbb{R}^n_{\geq 0}$ with all coordinates multiples of 2^{-k}.

Definition 4. *An increasing set S is k-discretized if $S = \bigcup_{z \in S \cap Q_k} B_z$. A corner c of a k-discretized set S is a point $c \in S \cap Q_k$ such that there does not exist $z \in S \setminus \{c\}$ with $z \preceq c$.*

Lemma 1. *Every k-discretized set S has only finitely many corners. Furthermore, $S = \cup_{c \in \mathcal{C}} B_c$, where $\mathcal{C}$ is the collection of corners of S.*

Proof. We prove that there are finitely many corners by induction on the dimension, n.

For the base case of $n = 1$, note that if S is nonempty it has exactly one corner.

Now suppose S has dimension n. Pick some corner $\hat{c} = (c_1, \ldots, c_n) \in S$. We know that any other corner $c \in \mathcal{C}$ must be strictly less than $\hat{c}$ in some coordinate. Therefore,

$$|\mathcal{C}| \leq \underbrace{1}_{\hat{c}} + \sum_{i=1}^{n} |\{c \in \mathcal{C} \text{ s.t. } c_i < \hat{c}_i\}| = 1 + \sum_{i=1}^{n} \sum_{j=1}^{2^k \hat{c}_i} |c \in \mathcal{C} \text{ s.t. } c_i = \hat{c}_i - 2^{-k} j|.$$

By the inductive hypothesis, we know that each set $\{c \in \mathcal{C} \text{ s.t. } c_i = \hat{c}_i - 2^{-k}j\}$ is finite, since it is contained in the set of corners of the $(n-1)$-dimensional subset of S whose points have i^{th} coordinate $\hat{c}_i - 2^{-k}j$. Therefore $|\mathcal{C}|$ is finite.

To show that $S = \bigcup_{c \in \mathcal{C}} B_c$, we will show $S \subseteq \bigcup_{c \in \mathcal{C}} B_c$ and $\bigcup_{c \in \mathcal{C}} B_c \subseteq S$. The latter holds trivially since we know S is a k-discretized set and $\mathcal{C} \subseteq S \cap Q_k$. To prove $S \subseteq \bigcup_{c \in \mathcal{C}} B_c$, pick any $z \in S$. Since S is k-discretized, there exists a $b \in S \cap Q_k$ such that $z \in B_b$. If b is a corner, then z is clearly contained in $\bigcup_{c \in \mathcal{C}} B_c$. If b is not a corner, then there is some other point $b' \in S \cap Q_k$ with $b' \preceq b$. If b' is a corner, we're done. Otherwise, we repeat this process at most $2^k \sum_j b_j$ times to arrive at a point c such that there will not be another point $e \in S$ where $e_i \leq c_i \; \forall i \in [n]$ (i.e. e will be strictly greater in at least one coordinate). Thus, $c \in \mathcal{C}$. Since we already have that $z \in B_c$ by construction, this implies that $z \in \bigcup_{c \in \mathcal{C}} B_c$ and we are done. ■

We now show that, to verify that one measure stochastically dominates another on all increasing sets, it suffices to verify that this holds for all sets that are the union of finitely many bases.

Lemma 2. *Let $g, h : \mathbb{R}^n_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$ be bounded density functions such that $\int_{\mathbb{R}^n_{\geq 0}} g(\vec{x}) d\vec{x} < \infty$ and $\int_{\mathbb{R}^n_{\geq 0}} h(\vec{x}) d\vec{x} < \infty$. Suppose that, for all finite collections Z of points in $\mathbb{R}^n_{\geq 0}$, we have*

$$\int_{\bigcup_{z \in Z} B_z} g(\vec{x}) d\vec{x} \geq \int_{\bigcup_{z \in Z} B_z} h(\vec{x}) d\vec{x}.$$

Then, for all increasing sets A,

$$\int_A g(\vec{x}) d\vec{x} \geq \int_A h(\vec{x}) d\vec{x}.$$

Proof. Let A be an increasing set. We clearly have $A = \bigcup_{z \in A} B_z$. For any point $z \in \mathbb{R}^n_{\geq 0}$, we denote $z^{n,k} \in \mathbb{R}^n_{\geq 0}$ as follows:

$$z^{n,k}_i = \max\{0, z_i - 2^{-k}\} \quad \forall i \in [n].$$

Now, we define

$$A^l_k \triangleq \bigcup_{z \in A \cap Q_k} B_z; A^u_k \triangleq \bigcup_{z \in A \cap Q_k} B_{z^{n,k}}.$$

Clearly, both A^l_k and A^u_k are k-discretized, and we have that $A^l_k \subseteq A$ since $A \cap Q_k \subseteq A$. Also, for any $z \in A$, we have that there exists $z' \in A \cap Q_k$ such that each component of z' is at most 2^{-k} more than the corresponding component of z. So $A \subseteq A^u_k$. In summary, we have

$$A^l_k \subseteq A \subseteq A^u_k.$$

The key now is to bound

$$\int_{A^u_k} g d\vec{x} - \int_{A^l_k} g d\vec{x}.$$

Define

$$W_k = \{z : z_i > k \text{ for some } i\}; \quad W^c_k = \{z : z_i \leq k \text{ for all } i\}.$$

First, note that

$$\int_{A_k^u \cap W_k} g d\vec{x} - \int_{A_k^l \cap W_k} g d\vec{x} \leq \int_{W_k} g d\vec{x}.$$

Furthermore, since $\lim_{k\to\infty} W_k^c = \mathbb{R}^n_{\geq 0}$, we have that

$$\lim_{k\to\infty} \int_{W_k^c} g d\vec{x} = \int_{\mathbb{R}^n_{\geq 0}} g d\vec{x}.$$

Thus, we have that

$$\lim_{k\to\infty} \int_{W_k} g d\vec{x} = 0.$$

Using that

$$\int_{A_k^u \cap W_k} g d\vec{x} - \int_{A_k^l \cap W_k} g d\vec{x} \geq 0,$$

we get that

$$\lim_{k\to\infty} \left(\int_{A_k^u \cap W_k} g d\vec{x} - \int_{A_k^l \cap W_k} g d\vec{x} \right) = 0.$$

Next, we bound

$$\int_{A_k^u \cap W_k^c} g d\vec{x} - \int_{A_k^l \cap W_k^c} g d\vec{x} \leq |g|_{sup}(\lambda(A_k^u \cap W_k^c) - \lambda(A_k^l \cap W_k^c))$$

where $|g|_{sup} < \infty$ is the supremum of g (remember g is bounded), and $\lambda(\cdot)$ denotes the Lebesgue measure.

For each $m \in \{1, \ldots, n+1\}$ and $z \in \mathbb{R}^n_{\geq 0}$, we define the point $z^{m,k}$ by:

$$z_i^{m,k} = \begin{cases} \max\{0, z_i - 2^{-k}\} & \text{if } i < m \\ z_i & \text{otherwise} \end{cases}$$

and set

$$A_k^m \triangleq \bigcup_{z \in A \cap Q_k} B_{z^{m,k}}.$$

Note then that, $A_k^l = A_k^1$ (if $m = 1$ then $z_i^{m,k} = z_i \; \forall i$) and $A_k^u = A_k^{n+1}$ (if $m = n+1$ then $z_i^{m,k} = \max\{0, z_i - 2^{-k}\} \; \forall i$). Therefore,

$$\lambda(A_k^u \cap W_k^c) - \lambda(A_k^l \cap W_k^c) = \lambda(A_k^{n+1} \cap W_k^c) - \lambda(A_k^1 \cap W_k^c) = \sum_{m=1}^{n} (\lambda(A_k^{m+1} \cap W_k^c) - \lambda(A_k^m \cap W_k^c)).$$

Now, we notice that, for any point $(z_1, z_2, \ldots, z_{m-1}, z_{m+1}, \ldots, z_n) \in [0, k]^{n-1}$, there is an interval I of length at most 2^{-k} such that the point

$$(z_1, z_2, \ldots, z_{m-1}, w, z_{m-2}, \ldots, z_n) \in (A_k^{m+1} \setminus A_k^m) \cap W_k^c$$

if and only if $w \in I$. Therefore,

$$\lambda(A_k^{m+1} \cap W_k^c) - \lambda(A_k^m \cap W_k^c) \leq \int_0^k \cdots \int_0^k \int_0^k \cdots \int_0^k 2^{-k} dz_1 \cdots dz_{m-1} dz_{m+1} \cdots dz_n = 2^{-k} k^{n-1}.$$

Therefore, we have the bound

$$|g|_{sup}(\lambda(A_k^u \cap W_k^c) - \lambda(A_k^l \cap W_k^c)) \leq |g|_{sup} \sum_{m=1}^{n} 2^{-k} k^{n-1} = n|g|_{sup} 2^{-k} k^{n-1}$$

and thus

$$\begin{aligned} \int_{A_k^u} g d\vec{x} - \int_{A_k^l} g d\vec{x} &= \int_{A_k^u \cap W_k} g d\vec{x} - \int_{A_k^l \cap W_k} g d\vec{x} + \int_{A_k^u \cap W_k^c} g d\vec{x} - \int_{A_k^l \cap W_k^c} g d\vec{x} \\ &\leq \left(\int_{A_k^u \cap W_k} g d\vec{x} - \int_{A_k^l \cap W_k} g d\vec{x} \right) + n|g|_{sup} 2^{-k} k^{n-1}. \end{aligned}$$

Since

$$\lim_{k \to \infty} n|g|_{sup} 2^{-k} k^{n-1} = 0,$$

we have

$$\lim_{k \to \infty} \left(\int_{A_k^u} g d\vec{x} - \int_{A_k^l} g d\vec{x} \right) = 0.$$

Now, using that

$$\int_{A_k^u} g d\vec{x} \geq \int_A g d\vec{x} \geq \int_{A_k^l} g d\vec{x},$$

we have

$$\lim_{k \to \infty} \int_{A_k^u} g d\vec{x} = \int_A g d\vec{x} = \lim_{k \to \infty} \int_{A_k^l} g d\vec{x}.$$

Similarly, we have

$$\int_A h d\vec{x} = \lim_{k \to \infty} \int_{A_k^l} h d\vec{x}$$

and thus

$$\int_A (g - h) d\vec{x} = \lim_{k \to \infty} \left(\int_{A_k^l} g d\vec{x} - \int_{A_k^l} h d\vec{x} \right).$$

Since A_k^l is k-discretized, by Lemma 1, it has finitely many corners. Letting Z_k denote the corners of A_k^l, we have $A_k^l = \bigcup_{z \in Z_k} B_z$, and thus by our assumption $\int_{A_k^l} g d\vec{x} - \int_{A_k^l} h d\vec{x} \geq 0$ for all k. Therefore $\int_A g d\vec{x} \geq \int_A h d\vec{x}$, as desired. ■

Using Claim 1 and Lemma 2, we can see as a corollary that to verify $\beta \preceq \alpha$ it suffices to check that $\beta(B) \leq \alpha(B)$ for all sets B that are unions of finitely many bases.

4 Sufficient Condition for Stochastic Dominance in Two Dimensions

We will now apply Lemma 2 to prove a sufficient condition for stochastic dominance in two dimensions that is used for all the 2-item optimal mechanism arguments in the original paper.

Informally, Theorem 2 below deals with the scenario where two density functions, g and h, are both nonzero only on some set $\mathcal{C} \setminus R$, where R is a decreasing subset of $\mathcal{C}$. To prove that $g \succeq h$, it suffices to verify that (1) $g - h$ has an appropriate form (2) the integral of $g - h$ on $\mathcal{C}$ is *positive* and (3) if we integrate $g - h$ along either a vertical or horizontal line outwards starting from any point in R, the result is *negative*.

Theorem 2. *Let $\mathcal{C} = [c_1, d_1^+) \times [c_2, d_2^+)$, R be a decreasing nonempty subset of $\mathcal{C}$, and $g, h : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be bounded density functions which are 0 on R, have finite total mass, and satisfy*

1. $\int_{\mathcal{C}} (g - h) dx dy \geq 0$.
2. *For any basis vector $e_i \in \{(0, 1), (1, 0)\}$ and any point $z^* \in R$:*

$$\int_0^{d_i^+ - z_i^*} g(z^* + \tau e_i) - h(z^* + \tau e_i) d\tau \leq 0.$$

3. *There exist non-negative functions $\alpha : [c_1, d_1^+) \to \mathbb{R}_{\geq 0}$, $\beta : [c_2, d_2^+) \to \mathbb{R}_{\geq 0}$ and an increasing function $\eta : \mathcal{C} \to \mathbb{R}$ such that*

$$g(z_1, z_2) - h(z_1, z_2) = \alpha(z_1) \cdot \beta(z_2) \cdot \eta(z_1, z_2)$$

for all $(z_1, z_2) \in \mathcal{C} \setminus R$.

Then $g \succeq h$.

Proof. We begin by defining, for any $c_1 \leq a \leq b \leq d_1^+$, the function $\zeta_a^b : [c_2, d_2^+) \to \mathbb{R}$ by

$$\zeta_a^b(w) \triangleq \int_a^b (g(z_1, w) - h(z_1, w)) dz_1.$$

This function represents, for each w, the integral of $g - h$ along the line from (a, w) to (b, w).

Claim 2. *If $(a, w) \in R$, then $\zeta_a^b(w) \leq 0$.*

Proof. If $g(z_1, w) \leq h(z_1, w) \, \forall z_1 \in [a, b]$, then we have that $\int_a^b g(z_1, w) dz_1 \leq \int_a^b h(z_1, w) dz_1$ so $\zeta_a^b(w) \leq 0$.

Now suppose there exists a $z_1 \in [a, b]$ such that $g(z_1, w) > h(z_1, w)$. It must be that $(z_1, w) \notin R$ as both g and h are 0 in R. Moreover, because R is a decreasing set it is also true that $(\tilde{z}_1, w) \notin R$ for all $\tilde{z}_1 \geq z_1$ (contrapositive of the definition of a decreasing set). This implies by assumption 3 that

$$g(\tilde{z}_1, w) - h(\tilde{z}_1, w) = \alpha(\tilde{z}_1) \cdot \beta(w) \cdot \eta(\tilde{z}_1, w),$$

for all $\tilde{z}_1 \geq z_1$. Now given that $g(z_1, w) > h(z_1, w)$ and α, β are non-negative functions, we have that $\eta(z_1, w) > 0$, and since $\eta(\cdot, w)$ is an increasing function, $\eta(\tilde{z}_1, w) > 0$ for all $\tilde{z}_1 \geq z_1$ so we have that $g(\tilde{z}_1, w) \geq h(\tilde{z}_1, w)$ for all $\tilde{z}_1 \geq z_1$. Therefore, we have

$$\zeta_a^{z_1}(w) \leq \zeta_a^b(w) \leq \zeta_a^{d_1^+}(w).$$

By assumption 2, we have that $\zeta_a^{d_1^+}(w) \leq 0$ and thus we are done. ∎

We now claim the following:

Claim 3. *Suppose that $\zeta_a^b(w^*) > 0$ for some $w^* \in [c_2, d_2^+)$. Then $\zeta_a^b(w) \geq 0$ for all $w \in [w^*, d_2^+)$.*

Proof. Given that $\zeta_a^b(w^*) > 0$, by the contrapositive of the previous claim we have that $(a, w^*) \notin R$. Furthermore, since R is a decreasing set and $w \geq w^*$, it follows by the contrapositive of the definition of a decreasing set that $(a, w) \notin R$, and furthermore that $(c, w) \notin R$ for any $c \in [a, d_1^+)$. Therefore, by assumption 3 we may write

$$\zeta_a^b(w) = \int_a^b (g(z_1, w) - h(z_1, w))dz_1 = \int_a^b (\alpha(z_1) \cdot \beta(w) \cdot \eta(z_1, w))dz_1$$

where $w \geq w^*$. Specifically, we have that

$$\zeta_a^b(w^*) = \int_a^b (\alpha(z_1) \cdot \beta(w^*) \cdot \eta(z_1, w^*))dz_1.$$

Since η is increasing,

$$\zeta_a^b(w) = \int_a^b (\alpha(z_1) \cdot \beta(w) \cdot \eta(z_1, w))dz_1 \geq \int_a^b (\alpha(z_1) \cdot \beta(w) \cdot \eta(z_1, w^*))dz_1$$

Rewriting in terms of $\zeta_a^b(w^*)$, we have that

$$\int_a^b (\alpha(z_1) \cdot \beta(w) \cdot \eta(z_1, w^*))dz_1 = \frac{\beta(w)}{\beta(w^*)}\zeta_a^b(w^*).$$

Now, note that, since $\zeta_a^b(w^*) > 0$, we have $\beta(w^*) > 0$ (otherwise we are integrating the zero function). Therefore,

$$\zeta_a^b(w) \geq \frac{\overbrace{\beta(w)}^{\geq 0}}{\underbrace{\beta(w^*)}_{>0}} \underbrace{\zeta_a^b(w^*)}_{>0} \geq 0$$

as desired. ∎

We extend g and h to all of $\mathbb{R}^2_{\geq 0}$ by setting them to be 0 outside of $\mathcal{C}$. By Lemma 2, to prove that $g \succeq h$ it suffices to prove that $\int_A g dx dy \geq \int_A h dx dy$ for all sets A which are the union of finitely many bases. Since g and h are 0 outside of $\mathcal{C}$, it suffices to consider only bases $B_{z'}$ where $z' \in \mathcal{C}$, since otherwise we can either remove the base (if it is disjoint from $\mathcal{C}$) or can increase the coordinates of z' moving it to $\mathcal{C}$ without affecting the value of either integral.

We now finish the proof of Theorem 2 by induction on the number of bases in the union.

- **Base Case.**

  We aim to show $\int_{B_r}(g-h)dxdy \geq 0$ for any $r=(r_1,r_2)\in\mathcal{C}$. We have

$$\int_{B_r}(g-h)dxdy = \int_{r_2}^{d_2^+}\int_{r_1}^{d_1^+}(g-h)dz_1dz_2 = \int_{r_2}^{d_2^+}\zeta_{r_1}^{d_1^+}(z_2)dz_2.$$

  Now we know that either there exists some $z_2^* \in [c_2, r_2]$ such that $\zeta_{r_1}^{d_1^+}(z_2^*) > 0$, in which case, by Claim 3 we have that $\zeta_{r_1}^{d_1^+}(z_2) \geq 0$ for all $z_2 \geq r_2$, or $\zeta_{r_1}^{d_1^+}(z_2) \leq 0$ for all $z_2 \in [c_2, r_2]$. In the first case, the integral is clearly nonnegative, so we may assume that we are in the second case. We then have

$$\int_{r_2}^{d_2^+}\zeta_{r_1}^{d_1^+}(z_2)dz_2 \geq \int_{c_2}^{d_2^+}\zeta_{r_1}^{d_1^+}(z_2)dz_2 = \int_{c_2}^{d_2^+}\int_{r_1}^{d_1^+}(g-h)dz_1dz_2 = \int_{r_1}^{d_1^+}\int_{c_2}^{d_2^+}(g-h)dz_2dz_1.$$

  By an analogous argument to that above, we know that either $\int_{c_2}^{d_2^+}(g-h)(z_1,z_2)dz_2$ is nonnegative for all $z_1 \geq r_1$ (in which case the desired inequality holds trivially) or is nonpositive for all z_1 between c_1 and r_1. We assume therefore that we are in the second case, and thus

$$\int_{r_1}^{d_1^+}\int_{c_2}^{d_2^+}(g-h)dz_2dz_1 \geq \int_{c_1}^{d_1^+}\int_{c_2}^{d_2^+}(g-h)dz_2dz_1 = \int_{\mathcal{C}}(g-h)dxdy,$$

  which is nonnegative by assumption 1.

- **Inductive Step.** Suppose that we have proven the result for all finite unions of at most k bases. Consider now a set

$$A = \bigcup_{i=1}^{k+1} B_{z^{(i)}}.$$

  We may assume that all $z^{(i)}$ are distinct, and that there do not exist distinct $z^{(i)}$, $z^{(j)}$ with $z^{(i)} \preceq z^{(j)}$, since otherwise we could remove one such $B_{z^{(i)}}$ from the union without affecting the set A, and we would be left with k bases, allowing us to apply the inductive hypothesis to get the desired inequality.

  Therefore, we may order the $z^{(i)}$ such that

$$c_1 \leq z_1^{(k+1)} < z_1^{(k)} < z_1^{(k-1)} < \cdots < z_1^{(1)}$$

  and

$$c_2 \leq z_2^{(1)} < z_2^{(2)} < z_2^{(3)} < \cdots < z_2^{(k+1)}.$$

  Similarly to the base case, by Claim 3, we know that one of the two following cases must hold:

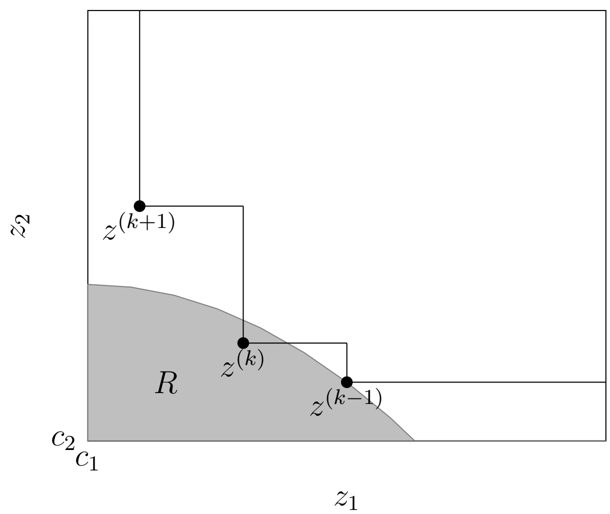

Figure 1: We show that either decreasing $z_2^{(k+1)}$ to $z_2^{(k)}$ or removing $z^{(k+1)}$ entirely decreases the value of $\int_A (g-h)$. In either case, we can apply our inductive hypothesis.

- **Case 1:** $\zeta_{z_1^{(k+1)}}^{z_1^{(k)}}(w) \leq 0$ for all $c_2 \leq w \leq z_2^{(k+1)}$.

  In this case, we see that

$$\int_{z_2^{(k)}}^{z_2^{(k+1)}} \int_{z_1^{(k+1)}}^{z_1^{(k)}} (g-h) dz_1 dz_2 = \int_{z_2^{(k)}}^{z_2^{(k+1)}} \zeta_{z_1^{(k+1)}}^{z_1^{(k)}}(w) dw \leq 0.$$

  Define the set

$$S \triangleq \left\{ z : z_1 \in [z_1^{(k+1)}, z_1^{(k)}] \wedge z_2 \in [z_2^{(k)}, z_2^{(k+1)}] \right\}.$$

  We now have

$$\int_A (g-h) dz_1 dz_2 \geq \int_A (g-h) dz_1 dz_2 + \int_S (g-h) dz_1 dz_2.$$

  Now, note that

$$\int_A (g-h) dz_1 dz_2 + \int_S (g-h) dz_1 dz_2 = \int_{S'} (g-h) dz_1 dz_2$$

  where

$$S' = \bigcup_{i=1}^{k} B_{z^{(i)}} \cup B_{(z_1^{(k+1)}, z_2^{(k)})}.$$

  Finally, since $(z_1^{(k)}, z_2^{(k)}) \succeq (z_1^{(k+1)}, z_2^{(k)})$, we have that

$$\bigcup_{i=1}^{k} B_{z^{(i)}} \cup B_{(z_1^{(k+1)}, z_2^{(k)})} = \bigcup_{i=1}^{k-1} B_{z^{(i)}} \cup B_{(z_1^{(k+1)}, z_2^{(k)})}.$$

Now, we can apply the inductive hypothesis to conclude that

$$\int_{S''} (g-h) dz_1 dz_2 \geq 0$$

where $S'' = \bigcup_{i=1}^{k-1} B_{z^{(i)}} \cup B_{(z_1^{(k+1)}, z_2^{(k)})}$, and we have that $\int_A (g-h) dz_1 dz_2 \geq 0$.

- **Case 2:** $\zeta_{z_1^{(k+1)}}^{z_1^{(k)}}(w) \geq 0$ for all $w \geq z_2^{(k+1)}$.
  In this case, we have

$$\int_{z_2^{(k+1)}}^{d_2^+} \int_{z_1^{(k+1)}}^{z_1^{(k)}} (g-h) dz_1 dz_2 = \int_{z_2^{(k+1)}}^{d_2^+} \zeta_{z_1^{(k+1)}}^{z_1^{(k)}}(w) dw \geq 0.$$

  After defining the set

$$S \triangleq \left\{ z : z_1 \in [z_1^{(k+1)}, z_1^{(k)}] \wedge z_2 \in [z_2^{(k+1)}, d_2^+) \right\},$$

  we have that

$$\int_A (g-h) dz_1 dz_2 = \int_{\bigcup_{i=1}^{k} B_{z^{(i)}}} (g-h) dz_1 dz_2 + \underbrace{\int_S (g-h) dz_1 dz_2}_{\geq 0} \geq \int_{\bigcup_{i=1}^{k} B_{z^{(i)}}} (g-h) dz_1 dz_2.$$

  We now apply the inductive hypothesis to conclude that

$$\int_{\bigcup_{i=1}^{k} B_{z^{(i)}}} (g-h) dz_1 dz_2 \geq 0$$

  and we are done.

■

References

[1] C. Daskalakis, A. Deckelbaum, C. Tzamos. Mechanism Design via Optimal Transport. In *the 14th ACM conference on Electronic Commerce (EC)*, 2013.

[2] T. Lindvall. On Strassen's theorem on stochastic domination. *Electronic Communications in Probability*, 4:51–59, 1999.