MAT1855 Mathematical Problems in Economics

1 Stable Matching

A paring of *e.g.* students to colleges is called *unstable*, if there is an unmatched pair who both prefer each other to their partners, called a blocking pair.

- Transferable Utility (TU): preferences can be changed with *e.g.* cash transfer (optimal transport)
- Non-Transferable Utility (NTU): preferences cannot be changed with external factors
- Partially Transferable Utility (PTU)

Example (Roommate Problem): Four students assigned to two double rooms. Does any set of preferences admit a stable match? Each student has a rank list of roommate preferences with no ties.
The answer is no. Consider the case where three students all rank the same person the least preferred.

1.1 Two-sided (Bipartite) Matching

Example:

- College admission: many-to-one matching
- Marriage problems: One-to-one. Further simplification
  - Same number on each side
  - Preferences: ordered list of potential opposite partners
  - Assume everyone prefers marriage to non-married state

***Definition:* 1.1: Assignment**

Denote $[n] = \{1, ..., n\}$. In the 1-1 context, an assignment or match refers to a 1-1 map $\sigma : [n] \to [n]$. *i.e.* a permutation on n letters.
Equivalently, $\mu : [2n] \to [2n]$, a 1-1 map s.t. $\mu(i) > n$ if $i \le n$, $\mu(j) \le n$ if $j > n$, $\mu^2(i) = i$ $\forall i$.

***Theorem:* 1.1: Gale & Shapley (1962)**

In 1-1 bipartite settings for preferences as above, a stable match always exist.

Proof. The proof is an algorithm which identifies a stable match. This is the deferred acceptance (dating algorithm)

- 1st round: each man proposes to favorite woman. Each woman keeps favorite suitor, and rejects all others.
- ith round: each rejected man now proposes to favorite woman among those who have not yet rejected him.

Claim:

1. This process terminates in finite time (after at most $(n-1)^2$ rounds)
2. At termination, stability has been achieved.

Facts (two monotonicites):

1. At each round, each woman is at least as satisfied with her partner as the previous round
2. At each round, each man is no more enthusiastic about the person he is proposing to in previous rounds.

Each of n men has n preferences, creating a $n \times n$ matrix. In each round, we iterating through the matrix and advances the rows or eliminates elements. It always terminates.

Stability: Suppose (M, w) is a blocking pair at termination.
Let (M, m) and (W, w) be partners. But M prefers w to m ($w >_M m$), and w prefers M to W ($M >_w W$).
Since $w >_M m$, M must have been rejected by w before he proposed to m.
But by the first monotonicity, w must like W better than M ($W >_w M$). Contradiction. □

Definition: 1.2: Achievable Matching

A pair (M, w) is achievable if it forms part of some stable matching.

Theorem: 1.2: Gale & Shapley (1962-2)

If (M, w) is achievable, then w will never reject a proposal by M during the algorithm.

Proof. Induciton on time T of the first rejection by an achievable partner to draw a contradiction.
Let T be the first round in which such a rejection takes place.
i.e. say a rejects A for B, $B >_a A$, yet $\sigma = (Aa, Bb, ...)$ is part of some stable match, because (A, a) is achievable.
Also, $b >_B a$, since otherwise Ba would be a blocking pair, contradicting stability.
From the two monotonicities, B must have proposed to b before a, but by minimality of time T and stability of σ, she cannot have rejected yet. Contradiction. □

Corollary 1. *Proposer optimality of deferred acceptance. i.e. male proposing deferred acceptance proposes a stable match that all men weakly prefer to any other stable match.*

Proof. No man is rejected by an achievable mate during the deferred acceptance algorithm. □

Theorem: 1.3: Knuth - Battle of the Seres (1981)

If two stable matchings σ and σ' have the property that all men weakly prefer σ to σ', then all women weakly prefer σ' to σ.

Proof. To derive a contradiction, suppose all men weakly prefer μ to μ', yet some woman a also strictly prefer μ to μ'. *i.e.* $B = \mu'(a) <_a \mu(a) = A$, $\mu(A) = a >_A \mu'(A) \neq a$. But now Aa blocks μ'. Contradiction. □

Theorem: 1.4: Conway's Distributive Lattice

If matchings $\sigma = (Aa, Bb, ...)$ and $\sigma' = (Aa', Bb', ...)$ are both stable, then so are $\sigma \vee \sigma' = (A \max_{<A} \{a, a'\}, B \max_{<B} \{b, b'\}, ...)$ (men's preferred wifes) and $\sigma \wedge \sigma' = (A \min_{<A} \{a, a'\}, B \min_{<B} \{b, b'\}, ...)$ (women's preferred husbands).

Proof. The two matchings are symmetric, so we prove $\sigma \vee \sigma'$.

1) Join produces a 1-1 matching:
Suppose WLOG, $\exists A \neq B$, $a = \max_{<A} \{a, a'\} = \max_{<B} \{b, b'\} = b'$
Then $\sigma' = (Aa', Bb', ...)$ and $a >_A a'$, so stability of σ' implies $A <_a B$.
Similarly, $\sigma = (A(a = b'), Bb)$ and $a = b' >_B b$, so stability of σ implies $A >_a B$. Contradiction.

2) Join is stable.
Suppose WLOG some pair Ab'' blocks, $\sigma \vee \sigma' = (Aa'', Bb'', ...)$
i.e. $b'' >_A a''$ and $A >_{b''} B$. Also, $a'' = \max_{<A} \{a, a'\}$, $b'' = \max_{<B} \{b, b'\}$.
Then $a, a' \leq_A a''$, $b, b' \leq_B b''$.
Either Aa'', Bb'' both occurred in the same stable match σ (or σ') or one occurred in σ, the other in σ'/
By symmetry, either $\sigma = (Aa'', Bb'', ...)$ or $\sigma = (Aa'', ...)$ and $\sigma' = (..., Bb'', ...)$.
In case 1 ($\sigma = (Aa'', Bb'', ...)$), Ab'' blocks stability of σ. Contradiction.
In case 2, $a = a'' >_A a'$, $b <_B b' = b''$, then $b'' >_A a'' = a >_A a'$ and $A >_{b''} B$, Ab'' blocks σ'. Contradiction. □

The lattice is distributive, meaning that $\sigma \wedge (\sigma' \vee \sigma'') = (\sigma \wedge \sigma') \vee (\sigma \wedge \sigma'')$ and similar distribution rules are satisfied.

1.2 Game Theory

Definition: 1.3: Economic Games

Economic games have: players, feasible outcomes, and rules. Player share preferences over feasible outcomes.
Player i has strategy $S_i = \left\{e_i^1, ..., e_i^{k(i)}\right\}$. Outcome: $S_1 \times \cdots \times S_n \to \Omega$. For each i, define a relation $\leq_i \in \Omega^2$ to show the preference of outcomes.
An outcome ω *dominates* ω' if there is a coalition S of player s.t. each player in S strictly prefers ω to ω' and the rules give S the power to enforce ω than ω'.
The *core* of the game refers to the set of all undominated outcomes.

Theorem: 1.5:

The set of stable matching forms the core of the marriage game.

Proof. If a matching is unstable, it's because a blocking pair strictly prefers to marry each other rather than their assigned partner, σ is not in the core.
Conversely, if any matching σ is not in the core, there is some coalition S of players willing and able to prevent it. At least one man $A \in S$ prefers to marry some woman $b \in S$, who is willing to marry him. Ab is a pair which blocks the stability of σ. Hence, σ is unstable. □

Definition: 1.4: Nash Equilibrium

A Nash equilibrium $(s_1, ..., s_n) \in S_1 \times \cdots \times S_n$ is a strategy s.t. no player can strictly improve his outcome, acting unilaterally.
$(s_1, ..., s_n)$ fails to be a Nash equilibrium $\Leftrightarrow \exists i \in [n]$ s.t. $f(s_1, .., s_n) <_i f(s_1, ...s_{i-1}, \tilde{s}, s_{i+1}, ..., s_n)$.

Two-Player Zero Sum Game $n = 2$, payoff is $P : S_1 \times S_2 \to \mathbb{R}$. Player 2 wins $P(s,t)$, player 1 loses $P(s,t)$ (or wins $-P(s,t)$).

Example: The penalty kick game with payoff in Table 1 does not admit Nash equilibrium. It's an advantage to know opponents strategies. If 2 is changed to 1, a random strategy will work.

Table 1: Penalty Kick Game

	kick left	kick right
dive left	1	-1
dive right	-1	2

Definition: 1.5: Randomized Strategy (von Neumann - Morgenstern)

Replace $S_1 = \left\{s^1, ..., s^m\right\} = \left\{e_1^1, ..., e_1^m\right\} \subset \mathbb{R}^m$ by simplex
$\Delta^{m-1} = \left\{(x_1, ..., x_m) \in [0,1]^m : \sum_{i=1}^{m} x_i = 1\right\}$, and $S_2 = \left\{t^1, ..., t^n\right\}$ by
$\Delta^{m-1} = \left\{(y_1, ..., y_n) \in [0,1]^n : \sum_{j=1}^{n} y_j = 1\right\}$. Replace the payoff matrix $(P_{ij})_{i\in[m], j\in[n]}$ by expected payoffs $P(\bar{x}, \bar{y}) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j P_{ij}$.

Remark 1. It never hurts and sometimes helps to know your opponents strategy. *i.e.* If P_1 knows P_2's strategy j, the payoff if P_2 forced to declare their strategy first is $\sup_{j\in S_2} \inf_{i\in S_1} P_{ij}$. The payoff if P_1 forced to declare strategy first is $\inf_{i\in S_1} \sup_{j\in S_2} P_{ij}$. We have $\sup_{j\in S_2} \inf_{i\in S_1} P_{ij} \leq \inf_{i\in S_1} \sup_{j\in S_2} P_{ij}$. (i_0, j_0) is a Nash equilibrium (saddle point) if and only if $P(i_0, j) \leq P(i_0, j_0) \leq P(i, j_0), \forall (i,j) \in S_1 \times S_2$.

Theorem: 1.6: Minmax Theorem

If a Nash equilibrium exists, then $\sup_{j\in S_2} \inf_{i\in S_1} P(i,j) = \inf_{i\in S_1} \sup_{j\in S_2} P(i,j)$ (No duality gap)

Proof. Let (i_0, j_0) be a Nash equilibrium,

$$\inf_{i\in S_1} \sup_{j\in S_2} P(i,j) \leq \sup_j P(i_0, j) \leq P(i_0, j_0) \leq \inf_i P(i, j_0) \leq \sup_j \inf_i P(i,j)$$

∎

Definition: 1.6: Coercive

$f : M \to (-\infty, \infty]$ is coercive $\Leftrightarrow \forall c \in \mathbb{R}$, $f^{-1}((-\infty, c])$ is compact in M.

> ***Theorem:* 1.7: Existence of Nash Equilibrium (von Neumann)**
>
> If $X \subset \mathbb{R}^m, Y \subset \mathbb{R}^n$ are compact convex sets and $P \in C(X \times Y)$ (continuous function from X to Y) and $\forall (x_0, y_0) \in (X, Y)$. $P(x_0, y), -P(x, y_0)$ are convex or at least have convex sublevel sets (coercive and uniquely minimized), then a Nash equilibrium exists

Proof. If both functions are strictly convex, the best response function $x_b(y_0) = \arg\min_x P(x, y_0)$ is unique and continuous. Similarly, P_2's best response $y_b(x_0) = \arg\max_y P(x_0, y)$ is also unique and continous. Therefore, $x_b \circ y_0 : X \to X$ is continous.
Brouwer's fixed point theorem implies that $\exists \tilde{x} \in X$ s.t. $\tilde{x} = x_b(y_b(\tilde{x}))$.
Because $P(\tilde{x}, y) \leq P(\tilde{x}, \tilde{y}) \leq P(x, \tilde{y})$, $\forall (x, y) \in X \times Y$, $\tilde{y} = y_b(\tilde{x})$ makes $(\tilde{x}, \tilde{y})$ a Nash equilibrium

If convexity is not strict, apply perturbation $P_\epsilon(x, y) = P(x, y) + \frac{\epsilon}{2}(\|x\|^2 - \|y\|^2)$.
Then $\exists (x_\epsilon, y_\epsilon)$ s.t. $P_\epsilon(x_\epsilon, y) \leq P_\epsilon(x_\epsilon, y_\epsilon) \leq P_\epsilon(x, y_\epsilon)$, $\forall (x, y) \in X \times Y$.
Because the sets are compact, there exists a subsequence $\epsilon(k) \to 0$ s.t. $\lim_{k\to\infty} (x_{\epsilon(k)}, y_{\epsilon(k)}) = (x_\infty, y_\infty)$. As $k \to \infty$, $P(x_\infty, y) \leq P(x_\infty, y_\infty) \leq P(x, y_\infty)$.
Claim: If $\arg\max_y P(x_0, y) = \{y_b(x_b)\}$, then $y_b : X \to Y$ is continuous.
Let $x_k \to x_\infty$ in X. Set $y_k = y_b(x_k), \forall k \in \mathbb{N}$. *i.e.* $P(x_k, y_k) \geq P(x_k, y), \forall y \in Y$.
Consider a subsequence $y_k(j) \to y_\infty$, $P(x_\infty, y_\infty) \geq P(x_\infty, y)$. Then $y_\infty \in \arg\max_y P(x_\infty, y)$. Therefore $y_\infty = y_b(x_\infty)$ since $\arg\max$ is unique.
The arbitrariness of subsequence gives that $y_b(x_\infty) = \lim_{k\to\infty} y_k = \lim_{k\to\infty} y_b(x_k)$ on the full sequence. □

Example: In a football game, the offense decide to pass or run, defense decide to defend the pass or run.

Table 2: Football Game

	P	R
DP	5	6
DR	7	1

P_1 (Defense) chooses probability $s \in [0, 1]$ to defend the pass (defends the runs with probability $1 - s$).
P_2 (Offense) chooses probability $t \in [0, 1]$ to pass (runs with probability $1 - t$).
Expected yards when offense pass: $y_{OP}(s) = 5s + 7(1 - s)$.
Expected yards when offense runs: $y_{OR}(s) = 6s + (1 - s)$.
The optimal strategy $\tilde{s}$ is the solution $y_{OP}(\tilde{s}) = y_{OR}(\tilde{s})$, $\tilde{s} = \frac{6}{7}$.
Similarly for offense, $5t + 6(1 - t) = 7t + (1 - t)$, $\tilde{t} = \frac{5}{7}$.

1.3 Transferrable Utility Matching (Shapley & Shubik)

Match between i and j produces benefit b_{ij}. i tries to maximize share u_i of the benefit b_{ij}, j tries to maximize the share v_j of b_{ij}.

1. **Stability**: $u_i + v_j \geq b_{ij}$, $\forall (i, j) \in I \times J$. Otherwise if $u_i + v_j < b_{ij}$ for some $(i, j) \in I \times J$, i and j will leave partner to marry each other to make b_{ij} better off.
2. **Market Clearing Condition**: If $|I| = |J| = n$ and matching is 1-1, an assignment is a permutation $\sigma \in \Sigma(n)$ on n letters. Alternatively, allow randomized assignments. *e.g.* i has probability $\gamma_{ij} \geq 0$ of

matching with i. Define doubly stochastic matrices: $DS(n) = \left\{ (\gamma_{ij})_{i,j}^n : \sum_{j=1}^n \gamma_{ij} = \sum_{i=1}^n \gamma_{ij} = 1 \right\}$. *e.g.* $\gamma_{ij}^\sigma = \begin{cases} 1, j = \sigma(i) \\ 0, \text{ else} \end{cases}$

3. **Budget Constraint**: $u_i + v_j = b_{ij}$ if $\gamma_{ij} > 0$.

Question: Given direct utility $(b_{ij})_{i,j}$, do there always exist vectors $\bar{u}, \bar{v}$ (indirect utility) and matrix $(\gamma_{ij})_{i,j}$ satisfying the above 3 conditions (stability, market clearing, budget constraint). If yes, then it is a stable matching.

More formally, let b_{ij} =benefit produced if i matches with j, does there exists $(\gamma_{ij})_{i,j=1}^n \in DS(n)$ and $u, v : [n] \to \mathbb{R}$ s.t.

1. Stability: $(u, v) \in L_b = \{u, v \in \mathbb{R}^n : u_i + v_j \geq b_{ij}\}$
2. Market clears: $\gamma \in DS(n) = \left\{ (\gamma_{ij} \geq 0)_{i,j=1}^n : \sum_{i=1}^n \gamma_{ij} = \sum_{j=1}^n \gamma_{ij} = 1 \right\}$
3. Budget: $u_i + v_j = b_{ij}$, $\forall i, j$ if $\gamma_{i,j} > 0$.

Proof. Let $X = \{u, v \in \mathbb{R}^n\} = \mathbb{R}^{2n}$, $Y = \left\{ \gamma \in \mathbb{R}^{n^2} : \gamma_{ij} \geq 0 \right\}$. Define

$$P((u, v); \gamma) = \sum_{i=1}^n \sum_{j=1}^n (b_{ij} - u_i - v_j)\gamma_{ij} + \sum_{i=1}^n u_i + \sum_{j=1}^n v_j.$$

$P((u, v); \gamma)$ is bi-affine in (u, v) and γ.

$$\sup_{\gamma \in Y} P((u, v); \gamma) = \begin{cases} \sum_{i=1}^n u_i + \sum_{j=1}^n v_j, \text{ if } (b_{ij} - u_i - v_j) \leq 0, \forall i, j \leq n, (u, v) \in L_b \\ \infty, \text{ otherwise} \end{cases}$$

Therefore, $\inf_{(u,v)\in X} \sup_{\gamma \in Y} P = \inf_{(u,v)\in L_b} \sum_{i=1}^n u_i + \sum_{j=1}^n V_j$.

Now, rewrite $P = \sum_{i=1}^n \sum_{j=1}^n b_{ij}\gamma_{ij} + \sum_{i=1}^n u_i \left(1 - \sum_{j=1}^n \gamma_{ij}\right) + \sum_{j=1}^n v_j \left(1 - \sum_{i=1}^n \gamma_{ij}\right)$.

$$\inf_{(u,v)\in X} P = \begin{cases} \sum_{i=1}^n \sum_{j=1}^n b_{ij}\gamma_{ij}, \text{ if } \gamma \in DS(n) \\ -\infty, \text{ otherwise} \end{cases}$$

Therefore, $\sup_{\gamma \in Y} \inf_{(u,v)\in X} P = \sup_{\gamma \in Y} \sum_{i=1}^n \sum_{j=1}^n b_{ij}\gamma_{ij}$.

For Nash equilibrium, we need $\inf \sup P \leq P((u^0, v^0), \gamma^0) \leq \sup \inf P$.
This only happens when $(u^0, v^0) \in L_b$ and $\gamma^0 \in DS(n)$. Otherwise we have $\pm\infty$ for $\inf \sup P$ and $\sup \inf P$. It automatically satisfies the stability and market clearing conditions. When they are exactly equal,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\gamma_{ij}^0 = \sum_{i=1}^{n} u_i^0 + \sum_{j=1}^{n} v_j^0 = \sum_{i=1}^{n} u_i^0 \left(\sum_{j=1}^{n} \gamma_{ij}\right) + \sum_{j=1}^{n} v_j^0 \left(\sum_{i=1}^{n} \gamma_{ij}\right)$$

$$\Rightarrow \sum_{i=1}^{n}\sum_{j=1}^{n} (u_i^0 + v_j^0 - b_{ij})\gamma_{ij}^0 = 0$$

$\Rightarrow$ either $\gamma_{ij}^0 = 0$ or $u_i^0 + v_j^0 = b_{ij}$ if $\gamma_{ij}^0 > 0$.

This is the budget constraint (complementary slackness) $\square$

This induces two variational problems:

1. The **social planners problem**: $\sup_{\gamma \in DS(n)} \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} b_{ij}$ (matchmaker)
2. Minimize total surplus subject to stability, hoping to achieve budget constraint (**Affordability**): $\inf_{(u,v)\in L_b} \sum_{i=1}^{n} u_i + \sum_{j=1}^{n} v_j$

In the affordability problem, stability $\Rightarrow u_i + v_j \geq b_{ij}, \forall i, j \in [n] \Rightarrow u_i \geq \max_j b_{ij} - v_j$ and $v_j \geq \max_i b_{ij} - u_i$. When b_{ij} has unique maximum for each i and each j, then we get perfect matching.

This is an instance of the **2nd welfare theorem**:
Supply and demand b_{ij} determines equilibrium prices (shadow prices/ Lagrange multipliers for market clearing) $(\bar{u}, \bar{v})$, which then decentralize the market. *i.e.* for a.e. b_{ij}, the corresponding $(\bar{u}^0, \bar{v}^0)$ lead each man and woman to have a unique preferred partner, so no matchmaker/social planner is needed. γ_{ij} is the Lagrange multiplier for stability constraint.

2 Optimal Transport

Now we generalize to continuous types/heterogeneity and consider Monge-Kantorovich/Optimal Transport Problems.

Definition: 2.1: Polish Space

A space X is Polish if its topology is metrizable by a complete separable metric. Let $\mathcal{P}(x) = \{\mu \geq 0 \text{ on } X : \mu \text{ is a Borel probability measure}, \mu(X) = 1\}$. Topologize $\mathcal{P}(X)$ using the narrow topology, *i.e.* $\mu_n \to \mu \Leftrightarrow \lim_{n\to\infty} \int_X f d\mu_n \to \int_X f d\mu$ for every $f \in C_b(X) = \{f : X \to \mathbb{R} : f \text{ continuous and bounded}\}$.

Definition: 2.2: Tight Measures

$C \subset \mathcal{P}(X)$ is tight $\Leftrightarrow \forall \epsilon > 0$, $\exists X_\epsilon \subset X$ compact s.t. $\sup_{\mu \in C} \mu(X - X_\epsilon) < \epsilon$.

Theorem: 2.1: Prokhorov

$C \subset \mathcal{P}(X)$ is narrowly pre-compact $\Leftrightarrow$ C is tight.

Corollary 2. *$\mu \in \mathcal{P}(X)$ is inner regular. i.e. $\forall \epsilon > 0$, $\exists X_\epsilon$ compact, $\mu(X - X_\epsilon) < \epsilon$.*

Definition: 2.3: Monge's Optimal Transport (1781)

Given X, Y Polish spaces, $\mu^+ \in \mathcal{P}(X)$, $\mu^- \in \mathcal{P}(Y)$, a cost function $c = -b \in C_b(X \times Y)$. We seek a Borel map $G : X \to Y$, μ^+-measurable. Define the push-forward $G_\# \mu^+$ of μ^+ by G s.t. $G_\# \mu^+(V) = \mu^+(G^{-1}(V))$ for $V \subset Y$ measurable. If $\mu^+ \in \mathcal{P}(X)$, then $G_\# \mu^+ \in \mathcal{P}(Y)$, given G is defined μ^+-a.e. The optimal solution is

$$\inf_{G_\# \mu^+(V)} \int_X c(x, G(x)) d\mu(x)$$

If $\frac{d\mu^\pm}{dvol} = f^\pm$ is smooth enough probability density and $G : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism, then

$$\int_{G^{-1}(V)} f^+(x) dvol(x) = \int_V f^-(y) dvol(y) = \int_{G^{-1}(V)} f^{-1}(G(x)) \left|\det DG(x)\right| dvol(x)$$

Since V is arbitrary, $f^+(x) = f^-(G(x)) \left|\det DG(x)\right|$ vol-a.e. for x. It is difficult to find the solution for the differential equation.

Definition: 2.4: Kantorovich's Optimal Transport (1942)

Seek a joint distribution $\gamma \in \mathcal{P}(X, Y)$, $\gamma \in \Gamma(\mu^+, \mu^-) = \{\gamma \geq 0 : \mu^+(U) = \gamma(U \times Y), \mu^-(V) = \gamma(X \times V), \forall (U, V) \subset X \times Y, \text{ Borel measurable}\}$. Or equivalently, $\mu^+ = \Pi^X_\# \gamma$ is the projection of γ onto X, $\mu^- = \Pi^Y_\# \gamma$ is the projection of γ onto Y.

$$\sup_{\gamma \in \Gamma(\mu^+, \mu^-)} \int_{X \times Y} b(x, y) d\gamma(x, y)$$

In optimization and calculus of variations, we need a topology for which

1. Objective is semi-continuous

2. Feasible competitors form a compact set

Note that in Kantorovich's definition, $\Gamma(\mu^+,\mu^-)$ is convex, and the integral functional is linear in γ, so the supremum (maximum) is always attained. F.O.C. (first order conditions, e.g. KKT/Euler-Lagrange equations for any concave maximization problems) become sufficient as well as necessary for optimality. Steepest ascent finds the global optimum.

If $G_{\#}\mu^+ = \mu^-$, then $\gamma_G = (id_X \times G)_{\#\mu^+} \in \Gamma(\mu^+,\mu^-)$.

$$\sup \int b(x,y)d\gamma \geq \sup \int b(x,G(x))d\mu^+$$

We want to find a topology making $\Gamma(\mu^+,\mu^-)$ compact and $\gamma \in \Gamma(\mu^+,\mu^-)$ s.t. $\int bd\gamma$ is continuous. A simple one would be both X, Y are compact. A generalization is X, Y both Polish or Radon.

Fact: If X is Polish, then the narrow topology in Definition 2.1 is metrizable on $\mathcal{P}(X)$.

Claim: If (X,d_X) and (Y,d_Y) are Polish (or Radon) and $\mu^+ \in \mathcal{P}(X)$, then $\Gamma(\mu^+,\mu^-)$ is tight and narrowly closed, hence narrowly compact.

Proof. (1) tightness: $\{\mu^+\}$ is narrowly compact as a set of one element.
Fix $\epsilon > 0$, $\exists X_\epsilon \subset X$ compact s.t. $\mu^+(X - X_\epsilon) < \frac{\epsilon}{2}$. Similarly, $\exists Y_\epsilon \subset Y$ compact s.t. $\mu^-(Y - Y_\epsilon) < \frac{\epsilon}{2}$. Therefore,

$$\begin{aligned}\gamma(X_\epsilon \times Y_\epsilon) &= 1 - \gamma(X_\epsilon^C \times Y) - \gamma(X \times Y_\epsilon^C) + \gamma(X_\epsilon^C \times Y_\epsilon^C) \\ &\geq 1 - \gamma(X_\epsilon^C \times Y) - \gamma(X \times Y_\epsilon^C) \\ &> 1 - \frac{\epsilon}{2} - \frac{\epsilon}{2} = 1 - \epsilon\end{aligned}$$

Therefore $\gamma((X_\epsilon \times Y_\epsilon)^C) < \epsilon$, and $\Gamma(\mu^+,\mu^-)$ is tight.

(2) Narror closedness:
Let $\{\gamma_k\}_{k\in\mathbb{N}} \subset \Gamma(\mu^+,\mu^-)$ converging narrowly to $\gamma_\infty \in \mathcal{P}(X \times Y)$.
Let $f(x,y) = \tilde{f}(x), \forall x \in X$, where $f \in C_b(X,Y)$.
By Definition of narrow topology,

$$\int f d\gamma_\infty = \lim_{k\to\infty} \int_{X\times Y} f(x,y)d\gamma_k = \lim_{k\to\infty} \int_{X\times Y} \tilde{f}(x)d\gamma_k = \int_X \tilde{f}d\mu^+$$

Since $\tilde{f} \in C_b(X)$ is arbitrary, $\Pi^X_{\#}\gamma_\infty = \mu^+$. Similarly, $\Pi^Y_{\#}\gamma_\infty = \mu^-$. Therefore, $\gamma_\infty \in \Gamma(\mu^+,\mu^-)$. □

Remark 2. More generally, if $\mu_k^+ \to \mu_\infty^+$ and $\mu_k^- \to \mu_\infty^-$ narrowly, then any $\gamma_k \in \Gamma(\mu_k^, \mu_k^-)$ admits a subsequential narrow limit $\gamma_\infty \in \Gamma(\mu_\infty^+, \gamma_\infty^-)$. (Narrow stability).

This shows that $\sup \int bd\gamma = \max \int bd\gamma$ can be attained.

2.1 Linear Programming Approach

In this section, we consider a heuristic linear programming dual problem, inspired by the discrete case.

$$
\begin{aligned}
\max_{\gamma\in\Gamma(\mu^+,\mu^-)} \int bd\gamma &= \sup_{\gamma\geq 0} \inf_{u:X\to\mathbb{R},v:Y\to\mathbb{R}} \int bd\gamma - \int ud\gamma + \int ud\mu^+ - \int vd\gamma + \int vd\mu^- \\
&\leq \inf_{u,v} \sup_{\gamma} \int ud\mu^+ + \int vd\mu^- + \int (b-u-v)d\gamma \\
&= \inf_{u,v\in L_b} \int ud\mu^+ + \int vd\mu^-,
\end{aligned}
$$

where $L_b = \{u : X \to \mathbb{R}, v : Y \to \mathbb{R} : u(x) + v(y) \geq b(x,y), \forall x, y \in X \times Y\}$

Claim: If $\gamma \in \Gamma(\mu^+,\mu^-)$ and $(u,v) \in L_b$, then

1. $\displaystyle\int_{X\times Y} (u(x) + v(y) - b(x,y))d\gamma \geq 0$, since $u + v - b \geq 0$ for $(u,v) \in L_b$ and $\gamma \geq 0$
2. If the integral is 0, and $u \in L^1(d\mu^+)$, $v \in L^1(d\mu^-)$, then γ maximizes and (u,v) minimizes.
3. If the infimum is attained and equality holds, then γ maximizes iff $\exists (u,v) \in L_b$ s.t. the integral is 0, whereas $(u,v) \in L_b$ minimizes iff $\exists \gamma \in \Gamma(\mu^+,\mu^-)$ s.t. integral is 0.

To show that the infimum is attained with no gap:

1. characterize maximizers $\gamma \in \Gamma(\mu^+,\mu^-)$ by a property of their support spt$\gamma = S$, where $S \subset X \times Y$ is the smallest closed subset with $\gamma(S) = 1$.
2. use the properties to construct minimizers u and its partner v in L_b.

Motivating Example: $X = Y = \mathbb{R}^n$, $b(x,y) = x \cdot y$.

***Definition:* 2.5: b-cyclically Monotone**

A set $S = X \times Y$ is b-cyclically monotone if and only if $\forall k \in \mathbb{N}$ and $(x_1,y_1), ..., (x_k,y_k) \in S$,

$$
\sum_{i=1}^{k} b(x_i,y_i) \geq \sum_{i=1}^{k} b(x_i,y_{i-1}),
$$

where $y_0 = y_k$.

***Theorem:* 2.2:**

If $\gamma \in \Gamma(\mu^+,\mu^-)$ maximizes $b \in \Gamma(X \times Y)$, then sptγ is b-cyclically monotone.

Proof. Suppose not, *i.e.* $\exists k \in \mathbb{N}$, $(x_1,y_1), ..., (x_k,y_k) \in \text{spt}(\gamma)$ s.t. $\displaystyle\sum_{i=1}^{k} b(x_i,y_i) < \sum_{i=1}^{k} b(x_i,y_{i-1})$.

Since $b \in C(X \times Y)$, the same inequality holds for all x_i' near x_i and y_i' near y_i. *i.e.* $x_i' \in I_i$ and $y_i' \in J_i$ for some open neighborhood $I_i \times J_i$ of (x_i,y_i).
Let $\epsilon = \min_{i\leq k} \gamma(I_i \times J_i) > 0$. Set $\gamma^i(z) = \frac{\gamma(z\cap I_i\times J_i)}{\gamma(I_i\times J_i)}$, $\gamma - \epsilon\gamma^i \geq 0$ and $\gamma - \frac{\epsilon}{k}\sum_{i=1}^{k} \gamma^i \geq 0$.

Suppose $\gamma_i = z_{i\#}\omega = (x_i,y_i)_{\#}\omega$ on (Ω,ω) where $\omega \in \mathcal{P}(\Omega)$ is a probability measure.

$$
\gamma_\epsilon = \gamma - \frac{\epsilon}{k} \sum_{i=1}^{k} ((x_i,y_i)_{\#}\omega - (x_i,y_{i-1})_{\#}\omega) \in \Gamma(\mu^+,\mu^-),
$$

but $\int bd(\gamma - \gamma_\epsilon) = \frac{\epsilon}{k}\sum_{i=1}^{k}\int bd(x_i, y_i)_{\#}\omega - bd(x_i, y_{i-1})_{\#}\omega < 0$. Contradicting the b-maximiality of γ. ☐

Definition: 2.6: Proper Function and b-subdifferential

$u : X \to [-\infty, \infty]$ is proper unless $u^{-1}(\infty) = X$.
The b-subdifferential $\partial_b u = \{(x, y) \in X \times Y : u(\cdot) \geq u(x) + b(\cdot, y) - b(x, y), \forall \cdot \in X\}$.

Example: $b(x, y) = x \cdot y$ on $X = Y = \mathbb{R}^n$, $(x, y) \in \partial_b u \Leftrightarrow u(\cdot) \geq u(x) + \langle \cdot - x, y\rangle$. $\partial_b u$ is a set of (point, slope) pairs that is affine supporting hyperplane for Graph(u) at (x, y).

Theorem: 2.3: Rockafellar (1966) and Rochet (1986)

$S \subset X \times Y$ is b-cyclically monotone $\Leftrightarrow S \subset \partial_b u$ for some proper $u : X \to [-\infty, \infty]$.

Proof. ($\Rightarrow$) Fix $(x_0, y_0) \in S$. For $x \in X$, define

$$u(x) = \sup_{k\in\mathbb{N}} \sup_{(x_1,y_1),\dots,(x_k,y_k)\in S} b(x, y_k) - b(x_k, y_k) + \sum_{i=0}^{k-1}[b(x_{i+1}, y_i) - b(x_i, y_i)]$$

Claim: $S \subset \partial_b u$, because if $(x', y') \in S$, then $\forall\epsilon > 0$, $\exists k \in \mathbb{N}$, $(x_1, y_1), \dots, (x_k, y_k)$ s.t.

$$u(x') \leq \epsilon + b(x', y_k) - b(x_k, y_k) + \sum_{i=0}^{k-1}[b(x_{i+1}, y_i) - b(x_i, y_i)]$$

Also, $\cdot \in X$ means

$$u(\cdot) \geq b(\cdot, y_{k+1}) - b(x_{k+1}, y_{k+1}) + \sum_{i=0}^{k}[b(x_{i+1}, y_i) - b(x_i, y_i)]$$

$u(\cdot) - u(x') \geq -\epsilon + b(\cdot, y') - b(x', y')$, but ϵ is arbitrary, so $(x', y') \in \partial_b u$.
S is b-cyclically monotone ensures $u(x_0) \leq 0$ and u is proper and bounded above. ☐

Definition: 2.7: b-transform and Legendre-Fenchel Transform

The Legendre-Fenchel transform is

$$u^*(y) = \sup_{y\in\mathbb{R}^n} \langle x, y\rangle - u(x).$$

This is a convex and lower semi continuous (l.s.c.) function. $u \geq u^{**}$ is the convex hull of u, with equality if and only if u is convex l.s.c., $u^{***} = u^*$.
The b-transforms are

$$u^b(y) = \sup_{x\in X} b(x, y) - u(x)$$

$$v^{\tilde{b}}(x) = \sup_{y\in Y} b(x, y) - v(y)$$

Lemma: 2.1: Properties of u^b

$L_b = \{u : X \to [-\infty, \infty], v : Y \to [-\infty, \infty] \text{ proper } : u(x) + v(y) \geq b(x, y), \forall x, y \in X \times Y\}$

1. If $(u, v) \in L_b$, then $(u, u^b) \in L^b$ and $u^b \leq v$
2. $(u^b)^{\tilde{b}} \leq u$
3. $((u^b)^{\tilde{b}})^b = u^b$

Proof. 1. $u^b(y) = \sup_{x \in X} b(x, y) - u(x) \geq b(x, y) - u(x), \forall x \in X$, so $(u, u^b) \in L_b$

If $v(y) \geq b(x, y) - u(x), \forall x \in X$, then $v(y) \geq \sup_{x \in X} b(x, y) - u(x) = u^b(y)$.

2. Using symmetry of x, y, if $(u, u^b) \in L_b$, then $((u^b)^{\tilde{b}}, u^b) \in L_b$ and $(u^b)^{\tilde{b}} \leq u$.

3. Because of the minus sign, $\tilde{u} \geq u \Rightarrow \tilde{u}^b \leq u^b$. Therefore $u^b \leq ((u^b)^{\tilde{b}})^b \underset{\text{By 2}}{\leq} u^b$. □

Theorem: 2.4: Kantorovich Duality

Let X, Y be Polish, $b \in C_{bdd}(X \times Y)$, $\mu^+ \in \mathcal{P}(X)$, $\mu^- \in \mathcal{P}(Y)$. Then

$$\max_{\gamma \in \Gamma(\mu^+, \mu^-)} \int b d\gamma \leq \inf_{(u,v) \in L_b} \int_X u d\mu^+ + \int_Y v d\mu^-$$

And infimum is attained.

Proof. Let $\gamma \in \Gamma(\mu^+, \mu^-)$ optimize b, then $\text{spt}(\gamma)$ is b-cyclically monotone. By Theorem 2.3, there exists some u proper s.t. $\text{spt}(\gamma) \subset \partial_b u$.

Claim: if $v(y) = u^b(y) = \sup_{x \in X} b(x, y) - u(x)$, then $(u, v) \in L_b$ and $\int (u + v - b) d\gamma = 0$, also $u \in L^1(d\mu^+), v \in L^1(d\mu^-)$, where u^b is a b-transform on u.

Claim: $(x, y) \in \partial_b u \Leftrightarrow u^b(y) = b(x, y) - u(x)$.
By Definition 2.6, $b(x, y) - u(x) \geq b(\cdot, y) - u(\cdot), \forall \cdot \in X$.
This is equivalent to $u^b(y) = b(x, y) - u(x)$. Equivalently, $x \in \arg\max b(\cdot, y) - u(\cdot)$.

Therefore, $\text{spt}(\gamma) = \big\{(x, y) \in X \times Y : u(x) + u^b(y) - b(x, y) = 0\big\}$.

$u + v \geq b \geq \inf b = B > -\infty$, so $\int (u + v) d\gamma \geq \int b d\gamma \geq B > -\infty$.

Then $\exists x_0 \in X$ s.t. $u(x_0) \in \mathbb{R}$, so $v(y) = u^b(y) = \sup b - u \geq b(x_0, y) - u(x_0) \geq B - u(x_0)$ for any $y \in Y$.

Then $\int u d\mu^+ + \int v d\mu^-$ must be finite. □

Consider $b(x, y) = x \cdot y$ on $X = Y = \mathbb{R}^n$, $\mu^\pm \in \mathcal{P}_C(\mathbb{R}^n)$ compact metric on $\mathbb{R}^n$. Then

$$\max_{\gamma \in \Gamma(\mu^+, \mu^-)} \int \langle x, y \rangle \, d\gamma = \min_{(u,v) \in L_b} \int_X u d\mu^+ \int_Y v d\mu^-,$$

and the optimizers satisfy $\text{spt}(\gamma) \subset \partial_b u = \text{Graph}(Du)$ if $u \in C^1(\mathbb{R}^n)$.

Let $u : \mathbb{R}^n \to (-\infty, \infty]$ be convex and lower semi-continuous (l.s.c.), then $\text{Dom}(u) = \{x \in \mathbb{R}^n : u(x) < \infty\}$ is convex and $\partial(\text{Dom}(u))$ is $n - 1$ dimensional.

Theorem: 2.5: Rademacher's

If u is Lipschitz on $\mathbb{R}^n$, then u is differentiable a.e.

Lemma: 2.2:

Let X, Y be Polish/Radon. If $\gamma \in \mathcal{P}(X \times Y)$ and $G : X \to Y$ s.t. $X \times Y - \text{Graph}(G)$ has zero γ outer measure, where $\text{Graph}(G) = \{(x, y) \in X \times Y : y \in G(x)\}$, then G is μ-measurable and $\gamma = (\text{id} \times G)_{\#}\mu$ where $\mu = \Pi^X_{\#}\gamma$.

Proof. The Radon (inner measurability) property of Polish spaces implies there exist compact $K_i \subset K_{i+1} \subset \text{Graph}(G) \subset X \times Y$ s.t. $\gamma(K_\infty) = 1$, where $K_\infty = \cup_i K_i$.

Claim: $G_i = G|_{X_i}$ is continuous where $X_i = \Pi^X(K_i)$.
Fix $i \in \mathbb{N}$. Let $x^j \in X_i$ be arbitrary, convergent to $x^\infty = \lim j \to \infty x^j$. Therefore, there exists $(x^j, y^j) \in K_i$, compactness implies that every subsequence admits a sub-subsequence with some limit $y^\infty = \lim\limits_{j\to\infty} y^j$ and $(x^\infty, y^\infty) \in K_i$. Since $K_i \subset \text{Graph}(G)$, $y^\infty = G(x^\infty)$ and $y^j = G(x^j)$. Also, $x^j \to x^\infty$, then $G(x^j) \to G(x^\infty)$ and arbitrariness of the subsequence. Hence $G|_{X_i}$ is continuous.

G_i admits a continuous extension to (X, d) by Dugundji variant of Tietze's extension theorem (If A is a closed subset of X and $f : A \to Y$ is continuous, then f can be extended to X).

$$\mu(X_\infty) = \gamma(X_\infty \times Y) = \gamma((X_\infty \times Y) \cap \text{Graph}(G)) = \gamma(K_\infty) = 1$$

So X_∞ is μ-measurable.

Given $U \subset X$ and $V \subset Y$ Borel. Then

$$\begin{aligned}
\gamma(U \times V) &= \gamma((U \times V) \cap \text{Graph}(G_\infty)) \\
&= \gamma((U \cap G^{-1}_\infty(V)) \times (Y \cap \text{Graph}(G_\infty))) \\
&= \gamma((U \cap G^{-1}_\infty(V)) \times Y) \\
&= \mu(U \cap G^{-1}_\infty(V)) = (\text{id} \times G)_{\#}\mu(U \times V)
\end{aligned}$$

$G_\infty = \lim\limits_{i\to\infty} \tilde{G}_i$ on X_∞. Hence μ-a.e. agrees with the limit of continuous functions. □

Consider $L^1_b = \{(u, v) \in L_b : u \in L^1(d\mu), v \in L^1(d\nu) l.s.c.\}$. Stability implies that $u(x) + v(y) - b(x, y) \geq 0$, $S(u, v) = \{(x, y) \in X \times Y : u(x) + v(y) - b(x, y) \leq 0\}$ is closed if $(u, v) \in L^1_b$ and $b \in C(X \times Y)$. (Upper semi-continuity of b is enough).

Proposition: 2.1:

$v^{\tilde{b}}(x) = \sup\limits_{y\in Y} b(x, y) - v(y)$ inherits the x-modulus of continuity of $x \in X \mapsto b(x, y)$.

Proof. Let $x_0, x_1 \in X$ and $v^{\tilde{b}}(x_0) = b(x_0, y_0) - v(y_0)$.
Then $v^{\tilde{b}}(x_0) - v^{\tilde{b}}(x_1) \leq b(x_0, y_0) - v(y_0) - b(x_1, y_0) + v(y_0) \leq w^b_{y_0}(d(x_0, x_1))$.
Similarly, $v^{\tilde{b}}(x_1) - v^{\tilde{b}}(x_0) \leq w^b_{y_1}(d(x_0, x_1))$ □

Definition: 2.8: Semi-Convexity

$u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ has semiconvexity constant C if and only if $\forall x_0 \in \Omega, h \in \mathbb{R}^n$, $x_0 \pm h \in \Omega$, $\frac{u(x_0+h)+u(x_0-h)-2u(x_0)}{h^2} \geq -C$. $C = 0$ if and only if u is convex.

Corollary 3. *1. If $v \in C(X \times Y)$ and X, Y compact, then $v^{\tilde{b}} \in C(X)$*

2. If $x \in X \mapsto b(x, y)$ has a Lipschitz constant independent of y, then $v^{\tilde{b}}$ has the same Lipschitz constant

3. If $x \in X \mapsto b(x, y)$ has semiconvexity constant independent of y, then $v^{\tilde{b}}$ has the same semiconvexity constant

Proof. 1. Compactness of $X \times Y$ means b is uniformly continuous. Hence $v^{\tilde{b}}(x)$ inherits the modulus of continuity. □

Remark 3. u has semiconvexity constant if and only if $u(x) + \frac{C}{2}x^2$ is convex.

Theorem: 2.6: Bremer 1987 & McCann 1995

Let $\mu, \nu \in \mathcal{P}_C(\mathbb{R}^n)$ (compactly supported probability measures), with $\mu \ll H^n$ (asbolutely continuous w.r.t. Lebesgue measure). Then $\exists u : \mathbb{R}^n \to \mathbb{R}$ convex $Du : \mathbb{R}^n \to \mathbb{R}^n$ satisfies $(Du)_\# \mu = \nu$. This map is unique μ-a.e. Moreover, $\gamma = (\text{id} \times Du)_\# \mu$ uniquely maximizes Kantorovich's problem for $b(x, y) = x \cdot y$.
Generalization (McCann): $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$. μ vanishes on all Lipschitz hypersurfaces or all sets of Hausdoff dimension $n - 1$.

Proof. There exists a Kantorovich optimizer $\gamma \in \Gamma(\mu, \nu)$. spt$(\gamma)$ is b-cyclically monotone, so $\exists u = (u^b)^{\tilde{b}}$ in spt$(\gamma) \subset \partial u$. u is convex and $u = \sup_{y \in Y} x \cdot y - u^b(y)$ is Lipschitz by compactness of spt(γ). Therefore u is differentiable a.e. by Theorem 2.5. $\partial u = \{(x, y) \in (domu) \times \mathbb{R}^n : u(\cdot) \geq u(x) + y(\cdot - x) \forall \cdot \in \mathbb{R}^n\}$, $\partial u \cap (domDu \times \mathbb{R}^n) = \text{Graph}(Du)$. Therefore, $1 = \gamma(\partial U \cap (domDu \times \mathbb{R}^n)) = \gamma(\text{Graph}(Du))$.

By Lemma 2.2, $\gamma = (\text{id} \times Du)_\# \mu$, $\nu = Du_\# \mu$.

Uniqueness:
Suppose $\tilde{u} : \mathbb{R}^n \to \mathbb{R}$ convex with $D\tilde{u}_\# \mu = \nu$, but $Du = D\tilde{u}$, μ-a.e. fails.
i.e. in some coordinate system, let $U = \left\{x : \frac{\partial u}{\partial x_1} > \frac{\partial \tilde{u}}{\partial x_1}\right\}$, $\mu\left(\left\{x : \frac{\partial u}{\partial x_1} > \frac{\partial \tilde{u}}{\partial x_1}\right\}\right) > 0$.
$\tilde{\gamma} = \gamma|_{U \times \mathbb{R}^n} \in \Gamma(\tilde{\mu}, \tilde{nu})$, so $\tilde{\gamma}(\text{Graph}(Du)) = 1$. $\tilde{\gamma} = (\text{id} \times Du)_\# \tilde{\mu}$, and $\tilde{\nu} = Du_\# \tilde{\mu}$.

Suppose $D\tilde{u}_\# \tilde{\mu} = \tilde{\nu}$, but $\int y_1 d(Du_\# \tilde{\mu} - D\tilde{u}_\# \tilde{\mu}) = \int \left(\frac{\partial u}{\partial x_1} - \frac{\partial \tilde{u}}{\partial x_1}\right) d\tilde{\mu} > 0$. Contradicting $(D\tilde{\mu})_\# \tilde{\mu} = \tilde{\nu}$.

When $\gamma = (\text{id} \times D\tilde{u})_\# \mu$, $\tilde{u}(x) + \tilde{u}^b(y) - x \cdot y \geq 0$. Equality $\Leftrightarrow (x, y) \in \partial \tilde{u} \Leftrightarrow y = D\tilde{u}^b$. Integrate w.r.t. γ. □

Theorem: 2.7: Isoperimetric Inequality

Let $\Omega \subset \mathbb{R}^n$ with $|\Omega| = H^n(\Omega) = |B_1^n(0)|$ (unit ball). Then $H^{n-1}(\partial \Omega) = |\partial \Omega| \geq |\partial B_1^n(0)|$.

Proof. Let $\mu = \frac{1_\Omega}{|\Omega|}$, $\nu = \frac{1_{B_1^n(0)}}{|B_1^n|}$. By Theorem 2.6, there exists a convex $u : \mathbb{R}^n \to \mathbb{R}$ s.t. $(Du)_\# \mu = \nu$. Then $D^2 u(x_0)$ exists μ-a.e.

$u(x) = u(x_0) + \langle x - x_0, p\rangle + \frac{1}{2}\langle x - x_0, Q(x-x_0)\rangle + o(|x-x_0|^2)$ where $p = Du(x_0), Q = D^2u(x_0)$ and $|Du| \leq 1$, $\det D^2u(x_0) = 1$ a.e.

$$1 = \det\left[D^2u(x_0)\right]^{1/n} \leq \frac{1}{n}\mathrm{tr}D^2u(x_0) = \frac{1}{n}\triangle u$$
$$1 = |\Omega| = \int_\Omega dH^n \leq \frac{1}{n}\int_\Omega \triangle u = \frac{1}{n}\int_{\partial\Omega} Du(x)\cdot \hat{n}_\Omega(x) dH^{n-1} \leq \frac{1}{n}H^{n-1}(\partial\Omega) = \frac{1}{n}|\partial\Omega|$$

Also $1 = |B_1^n(0)| = \frac{1}{n}|\partial B_1^n(0)|$. □

***Theorem:* 2.8:**

Let $b \in C_b(X\times Y)$, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$. $\gamma \in \Gamma(\mu,\nu)$ is b-optimal (b-maximality) $\Leftrightarrow \exists (u,v) \in L_b^1$, $\gamma(S_{(u,v)}) = 1$ where $S_{(u,v)} = \{(x,y) \in X\times Y : u(x) + v(y) - b(x,y) = 0\}$.

Proof. ($\Leftarrow$) Suppose $\gamma \in \Gamma(\mu,\nu)$ and $\exists (u,v) \in L_b^1$ s.t. $\gamma(S_{(u,v)}) = 1$, $u(x)+v(y) = b(x,y)$, γ-a.e. $\int u d\mu + \int v d\nu = \int b d\gamma$, but $\inf\limits_{(\tilde{u},\tilde{v})\in L_b^1} \int \tilde{u}d\mu + \tilde{v}d\nu \geq \sup\limits_{\tilde{\gamma}\in\Gamma(\mu,\nu)}\int bd\tilde{\gamma}$, (u,v) minimizes, γ maximizes.

($\Rightarrow$) If $\gamma \in \Gamma(\mu,\nu)$ is b-optimal. Theorem 2.3 implies that $\mathrm{spt}(\gamma) \subset \partial_b u$ for some proper u s.t. $S_{(u,u^b)} \subset S_{(u^{b\tilde{b}},u^b)}$, where $(u^{b\tilde{b}}, u^b) \in L_b^1$. □

***Lemma:* 2.3: Restriction Property**

If $\gamma \in \Gamma(\mu,\nu)$ is b-optimal and $0 \leq \tilde{\gamma} \leq \gamma$, then so is $\hat{\gamma} = \frac{\tilde{\gamma}}{\tilde{\gamma}[X\times Y]}$

Proof. b-optimality of $\gamma \Rightarrow \exists (u,v) \in L_b^1$, $\mathrm{spt}(\hat{\gamma}) \subset \mathrm{spt}(\gamma) \subset S_{(u,v)}$. Therefore, $\hat{\gamma}$ is b-optimal. □

***Definition:* 2.9: Twist**

Let X be a manifold (or $\mathbb{R}^n$), $b \in C^1(X\times Y)$ satisfies twist if and only if $\forall y, y' \in Y, x \in X$, $b(x,y) - b(x,y')$ has no critical points. Equivalently, $\forall x_0 \in X$, the map $y \in Y \mapsto D_{x_0}b(x,y)$ is 1-1.

Remark 4. Twist implies that X is not compact.

Example: $b(x,y) = -h(x-y)$ where h is strictly convex/concave. Then b is a twist

Proof. $D_{x_0}b(x,y) = -Dh(x_0 - y) = p$, then $Dh^{-1}(-D_{x_0}b(x,y)) = x_0 - y$, so $y = x_0 - Dh^{-1}(p)$. □

Example: $b(x,y) = x\cdot y$, and $b(x,y) = -\frac{1}{2}|x-y|^2$ or any distance metrics are twists.

Example: Fix a Lagrangian $L \in C^1(TM)$ where TM is the tangent bundle of a manifold. Then $\forall x \in M$, $v \in T_xM \mapsto L(x,v)$ is strictly convex. Consider the cost of action along a smooth path. $\forall \sigma : [t_0,t_1] \to M$, its action is $A[\sigma] = \displaystyle\int_{t_0}^{t_1} L(\sigma(t),\dot{\sigma}(t))dt$. The cost $C(x_0,x_1) = \inf A[\sigma]$ on $\sigma \in C^2([t_0,t_1],M)$, $\sigma(t_0) = x_0$, $\sigma(t_1) = x_1$ (C^2 or preferrably Lipschitz curve from x_0 to x_1). $b = -c$ is twisted if inf is attained. This problem is following the idea: given the tangent vector at final time t_1 at position x_1, can we find the initial position x_0? The answer is yes.

Lemma: 2.4:

If $b \in C^1(X \times Y)$ is twisted, then $S_{(u,v)} \cap (\text{dom}Du \times Y) \subset \text{Graph}(G)$, where $G(x) = (D_x b)(x,\cdot)^{-1}(Du(x))$. If $u = u^{b\tilde{b}}$, Y is compact and $\partial X \cap \text{dom}Du = \emptyset$, then $S_{(u,v)} \cap (\text{dom}Du \times Y) = \text{Graph}(G)$.

Proof. Let $(\bar{x}, \bar{y}) \in S_{(u,v)} \cap (\text{dom}Du \times Y)$.
Since $u(x) + v(y) - b(x,y) \geq 0$, $\forall (x,y) \in X \times Y$ at $(\bar{x}, \bar{y})$, $Du(x) = D_x b(x,y)$. Therefore, it is a twist and $y = G(x)$.

If $x \in \text{dom}D^2u$, then $D^2u(\bar{x}) \geq D^2_{xx}b(\bar{x}, \bar{y})$, and $(\bar{x}, \bar{y}) \in \text{dom}D^2 b$. □

Theorem: 2.9: Gangbo 1996, Levin 1999

If $\mu, \nu \in \mathcal{P}_C(\mathbb{R}^n)$, $\mu \ll H^n$, $\|b\|_{C^1} < \infty$ and twisted, then $\exists u = u^{b\tilde{b}}$ s.t. $G(x) = D_x b(x,\cdot)^{-1} Du(x)$ satisfies $G_{\#}\mu = \nu$. This map is unique. Also, $\gamma = (\text{id} \times G)_{\#}\mu$ uniquely solves Kantorovich. Finally G is invertible if also $\nu \ll H^n$ and both b and $\tilde{b}$ are twisted.

Proof. By Theorem 2.8. There exist opimizers $\gamma, \tilde{\gamma}$ for Kantorovich problem and $\text{spt}(\gamma) \subset \partial_b u = S_{(u,u^b)}$ with $u = u^{b\tilde{b}}$.
$\|b\|_{C^1} < \infty$ means that u is Lipschitz, hence $H^n(X - \text{dom}Du) = 0$ by Theorem 2.5.
Lemma 2.4 implies that $S_{(u,u^b)} \cap (\text{dom}Du \times Y) = \text{Graph}(G)$, where $G(x) = (D_x b)(x,\cdot)^{-1}(Du(x))$.
$\gamma[\text{dom}Du \times Y] = \mu[\text{dom}Du] = 1$ because $\mu \ll H^n$. Therefore, $\gamma = (\text{id} \times G)_{\#}\mu$ from Lemma 2.2, $G_{\#}\mu = \nu$.

$$\int b d\tilde{\gamma} = \int b d\gamma = \int u d\mu + \int v d\nu = \int \tilde{u} d\mu + \int \tilde{u}^b d\nu$$

Therefore, $\tilde{\gamma} = (\text{id} \times G)_{\#}\mu$.

Also, $\gamma = (\text{id} \times \tilde{G})_{\#}\mu = (\text{id} \times G)_{\#}\mu$, where $\tilde{G}(x) = D_x b(x,\cdot)^{-1}(D\tilde{u}(x))$ on $D\tilde{u}$.

Claim: $G = \tilde{G}$ on $\text{dom}Du \cap \text{dom}D\tilde{u}$ μ-a.e.
If $b(x,y) = x \cdot y$, then $G = Du, \tilde{G} = D\tilde{u}$, $V = \left\{x \in X : \frac{\partial u}{\partial x^1} < \frac{\partial \tilde{u}}{\partial x^1}\right\}$. Aassume $\mu(V) > 0$. Let $\hat{\mu} = \frac{\mu|_V}{\mu(V)}$ and $\hat{\gamma} = \frac{\gamma|_{V \times Y}}{\mu(V)} \in \Gamma(\hat{\mu}, \hat{\nu})$.

Lemma 2.3 means that $(\text{id} \times \tilde{G})_{\#}\hat{\mu} = (\text{id} \times G)_{\#}\hat{\mu}$, so $G_{\#}\hat{\mu} = \tilde{G}_{\#}\hat{\mu}$.

$$\int y^1 dG_{\#}\hat{\mu} = \int y^1 d\tilde{G}_{\#}\hat{\mu} = \int \frac{\partial u}{\partial x^1} d\hat{\mu} = \int \frac{\partial \tilde{u}}{\partial x^1} d\hat{\mu}$$

But $\frac{\partial u}{\partial x^1} < \frac{\partial \tilde{u}}{\partial x^1}$ μ-a.e. Contradiction.

For general $b \in C^1(X \times Y)$ twisted, define $d_Y(G(x), \tilde{G}(x)) = \sup_{\|\phi\|_{Lip(y)} \leq 1} \phi(G(x)) - \phi(\tilde{G}(x)) = \sup \phi_i(G(x)) - \phi_i(\tilde{G}(x))$ for a countable collection of Lipschitz function.

Consider $U_i^{\pm} = \Big\{x \in X : \pm(\phi_i(G(x)) - \phi_i(\tilde{G}(x)) > 0\Big\}$. If $\mu(U_i^{\pm}) = 0, \forall i$, then $G = \tilde{G}$ μ-a.e. If not, $\exists i$ s.t. $\mu(U_i^{\pm}) > 0$, but it reaches the same contradition. □

2.2 Regularity

$b(x,y) = x \cdot y \Rightarrow F(x) = Du(x)$ is optimal between $d\mu^+(x) = f(x)d^n x$ and $d\mu^-(y) = g(y)d^n y$.
Claim: u convex $\Rightarrow F$ is approximately differentiable a.e. and $\det(DF(x)) = \frac{f(x)}{g(F(x))}$.

***Theorem:* 2.10:**

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ be probability measures, $\gamma \in \Gamma(\mu, \nu)$ is b-optimal. Then $\text{spt}(\gamma) \subset S \subset (\mathbb{R}^n)^2$ with S a 1-Lipschitz graph over diagonal.

Proof. By Theorem 2.2, b-optimality means $\text{spt}(\gamma)$ is b-cyclically monotone. Then $(x_0, y_0), (x_1, y_1) \in \text{spt}(\gamma) \Rightarrow \delta_x \delta y \geq 0$, where $\delta_x = x_1 - x_0$, $\delta_y = y_1 - y_0$.
Change bais, $\delta_x = \frac{\delta_z - \delta_w}{\sqrt{2}}$, $\delta_y = \frac{\delta_z + \delta_w}{\sqrt{2}}$. Then

$$0 \leq \delta_x \delta_y = \frac{1}{2}(\delta_z - \delta_w)(\delta_z + \delta_w) = \frac{1}{2}(|\delta_z|^2 - |\delta_w|^2)$$

$\Rightarrow |\delta_w| \leq |\delta_z|$, *i.e.* $|w_1 - w_0| \leq |z_1 - z_0|$ which is 1-Lipschitz.
Kirszbraun's extension theorem implies that $\exists$ 1-Lipschitz $W : \mathbb{R}^n \to \mathbb{R}^n$ s.t. $(z, w) \in \text{Graph}(W) \subset \text{spt}(\gamma)$. $\forall (z, w) = \left(\frac{y+x}{\sqrt{2}}, \frac{y-x}{\sqrt{2}}\right)$ with $(x, y) \in \text{spt}(\gamma)$. □

***Definition:* 2.10: Area and Coarea**

For Lipschitz change of variables $F : \mathbb{R}^n \to \mathbb{R}^m$, $DF : \mathbb{R}^n \to \mathbb{R}^m$ (all derivatives), define the Jacobian: $JF(x) = \begin{cases} \sqrt{\det(DFDF^\dagger)}, n \leq m \\ \sqrt{\det(DF^\dagger DF)}, n > 1m \end{cases}$, $\forall A \subset \mathbb{R}^n$ measurable, the following holds:

$$\int_A JF(x) dH^n(x) = \int_{\mathbb{R}^m} H^{(n-m)_+}(A \cap F^{-1}(y)) dH^{\min(n,m)}(y)$$

$$\int_{\mathbb{R}^n} \chi_A JF dH^n = \int_{\mathbb{R}^m} \int (\chi_A \circ F^{-1})(y) dH^{(n-m)_+}(y)$$

If we approximate with simple functions, we get:

$$\int_{\mathbb{R}^n} \phi JF dH^n = \int_{\mathbb{R}^m} \int (\phi \circ F^{-1})(y) dH^{(n-m)_+}(y)$$

Corollary 4. *If $A \subset \mathbb{R}^n$ has $H^n(A) = 0$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, then $H^m(F(A)) = 0$*

Q: How does the tagent space a.e. to $S \subset (\mathbb{R}^n)^2$ relate to differentiability of $G = Du : \mathbb{R}^n \to \mathbb{R}^n$?
Recall $\text{Graph}(G) \subset S = \text{Graph}(woverz)$. Define $X(z) = \frac{z - W(z)}{\sqrt{2}}, Y(z) = \frac{z + W(z)}{\sqrt{2}}$. Since W is 1-Lipschitz, then X, Y are Lipschitz.
Define $Z_{bad} = \{z \in \mathbb{R}^n : DW(z) \text{ does not exist}\}$. $H^n(Z_{bad}) = 0$, so by the corollary, $H^n(X(Z_{bad})) = 0$.

Claim: $G = Du$ is countably Lipschitz, *i.e.* $\cup_i X_i \subset \text{dom}(G)$ s.t. $G|_{X_i}$ has Lipschitz constant i and $H^n(\text{dom}(G) \setminus \cup_i X_i) = 0$. Then by Theorem 2.5 and extension, G is approximately differentiable H^n-a.e.

Proof. Heuristically $y = G(x)$ μ-a.e. $Y(z) = G(X(z))$, $\frac{z+W(z)}{\sqrt{2}} = G\left(\frac{z-W(z)}{\sqrt{2}}\right)$.
Differentiate both sides, $\frac{1}{\sqrt{2}}(I + D_zW(z)) = D_xG|_{X(z)}\frac{1}{\sqrt{2}}(I - D_zW(z))$.
Then $D_xG|_{X(z)} = (I + D_zW(z))(I - D_zW(z))^{-1}$ provided that $I - D_zW(z)$ is invertible.

Let $Z_1 = \{z : I - D_zW(z) \text{ is not invertible}\}$. We show that $H^n(X(Z_1)) = 0$.
Since $X(z)$ is Lipschitz, $z \in Z_1 \Rightarrow JX(z) = \sqrt{\det(DFDF^\dagger)} = 0$.

$$0 = \int_{Z_1} JX(z)dH^n = \int_{\mathbb{R}^n}\int_{Z\cap X^{-1}(x)} dH^0(Z)dH^n(X) \geq H^n(X(Z_1))$$

□

Note also that $Y(z) = G(X(z))$ implies $Y \circ X^{-1} = G$. X has a Lipschitz inverse except on Z_1. This applies implicit function theorem for Lipschitz maps.

Example: If $u : \mathbb{R}^n \to \mathbb{R}$ is convex and $Du(0) \neq 0$ exists. Assume $\frac{\partial u}{\partial x_n} \neq 0$, then there exists difference of convex functions $W : \mathbb{R}^{n-1} \to \mathbb{R}$ s.t. $u(x_1, ..., x_{n-1}, x_n) = 0$ near 0, then $x_n = W(x_1, ..., x_{n-1})$.

However, this does not work for Lipschitz functions due to potential bad sets.

***Theorem:* 2.11: Clarke 1976**

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be Lipschitz. Define $\partial F(0)$ to be the convex hull of $\left\{y \in \mathbb{R}^n : x_i \in \text{dom}(DF) \subset \mathbb{R}^n, y = \lim_{i\to\infty} DF(x_i), \lim_{i\to\infty} x_i = 0\right\}$. If $0 \notin \partial F(0)$, then F is invertible and has Lipschitz inverse in a neighborhood of 0.

***Theorem:* 2.12:**

Let X, Y be compact subset of $\mathbb{R}^n$. If γ is b-optimal and $b \in C^2(X \times Y)$ is non-degenerate at (x_0, y_0), *i.e.* $\det \frac{\partial^2 b}{\partial x^i \partial y^j} \neq 0$, then $\exists \epsilon > 0$ s.t. $\text{spt}(\gamma) \cap B_\epsilon(x_0, y_0) \subset S$, where S is an n-dim Lipschitz submanifold of $\mathbb{R}^{2n}$.

Proof. Since $\det \frac{\partial^2 b}{\partial x^i \partial x^j} \neq 0$, there exists new coordinate $\tilde{y}(y)$ s.t. $\tilde{b}(x, \tilde{y}) = b(x, y)$ with $\tilde{y}(y_0) = 0$ and $\frac{\partial^2 b}{\partial x^i \partial y^j}(x_0, 0) = \delta_{ij}$, $\frac{\partial^2 \tilde{b}}{\partial x^i \partial \tilde{y}^j}(x, \tilde{y}) = \delta_{ij} + o((x - x_0)^2 + \tilde{y}^2)$.

WLOG, set $x_0 = 0$, $z = \frac{\tilde{y}+x}{\sqrt{2}}$, $w = \frac{\tilde{y}-x}{\sqrt{2}}$.
$\forall (x_0, y_0), (x_1, y_1) \in \text{spt}(\gamma)$, define $\Delta(x_1, y_1, x_0, y_0) = b(x_1, y_1) - b(x_0, y_0) - b(x_1, y_0) - b(x_0, y_1)$. $\Delta \geq 0$ on $(\text{spt}(\gamma))^2$. Let $\Delta_0(x, y) = \Delta(x, y, x_0, y_0)$. Apply Taylor expansion:

$$\begin{aligned}
\Delta_0(x_0 + \delta_x, y_0 + \delta_y) &= (\delta_x, \delta_y)D^2_{xy}b\begin{pmatrix}\delta_x \\ \delta_y\end{pmatrix} + o(|\delta_x|^2 + |\delta_y|^2) \\
0 \leq \tilde{\Delta}_0(x_0 + \delta_x, 0 + \delta_{\tilde{y}}) &= \langle \delta_x, \delta_{\tilde{y}}\rangle + o(|\delta_x|^2 + |\delta_y|^2) \\
&\leq \langle \delta_x, \delta_{\tilde{y}}\rangle + \frac{\eta}{2}o(|\delta_x|^2 + |\delta_y|^2),
\end{aligned}$$

for all $\eta > 0$. Then $\exists \delta > 0$, $(x_1, \tilde{y}_1) \in B_\epsilon(x_0, y_0)$ s.t.

$$0 \leq \frac{1}{2}(|\delta_z|^2 - |\delta_w|^2) + \frac{\eta}{2}(|\delta_w|^2 + |\delta_z|^2) \leq (1 + \eta)|\delta_z|^2$$

Then $|\delta_w| \leq \sqrt{\frac{1+\eta}{1-\eta}}|\delta_z|$. The transformation $W(z)$ is Lipschitz. □

Summary of Conditions:
Let $\Delta_0(x,y) = b(x,y) + b(x_0,y_0) - b(x,y_0) - b(x_0,y) \geq 0$ on $\text{spt}(\gamma) \subset \text{Graph}(G)$ γ-a.e.

(B0) $b \in C(\overline{X \times Y})$

(B1) Twist: $\forall x_0 \in X \subset \mathbb{R}^m, y \in Y \subset \mathbb{R}^n$, $D_{x_0} b(x,y)$ is 1-1. b is twisted $\Leftrightarrow$ b and $\tilde{b}$ are twisted $\Leftrightarrow$ $\Delta_0(x,y)$ has no critical points except $(x,y) = (x_0,y_0)$. $m = n$

(B2) Non-degeneracy: $\det \frac{\partial^2 b}{\partial x^i \partial y^j}(x_0,y_0) \neq 0$ for $(x_0,y_0) \in X \times Y$, $m = n$. $\Delta_0(x_0+\delta_x, y_0+\delta_y) = \frac{1}{2}(\delta_x,\delta_y) D^2_{xy}\Delta_0 \binom{\delta_x}{\delta_y} + o(|\delta_x|^2 + |\delta_y|^2)$ as $(\delta_x,\delta_y) \to 0$ if $b \in C^3$. $D^2_{xy}\Delta_0 = \begin{bmatrix} 0 & D^2_{xy}b \\ (D^2_{xy}b)^\dagger & 0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$.
$\det D^2_{xy}\Delta_0 = (-1)^n \det D^2_{xy} b$.
Then $\Delta_0(x_0+\delta_x, y_0+\delta_y) = \delta_x D^2_{xy} b(x_0,y_0)\delta_y + o(|\delta_x|^2 + |\delta_y|^2)$ if $b \in C^2$.
Also $\Delta_0(x_0+\delta_x, y_0-\delta_y) = -\Delta_0(x_0+\delta_x, y_0+\delta_y) + o(|\delta_x|^2 + |\delta_y|^2)$.
Let $H_0 = D^2_{xy}\Delta_0$. If $H_0\binom{\delta_x}{\delta_y} = \lambda\binom{\delta_x}{\delta_y}$, then $H_0\binom{\delta_x}{-\delta_y} = -\lambda\binom{\delta_x}{\delta_y}$. So dianoalizing the symmetric matrix H_0 produces $\pm$ eigenvalue pairs.

If $\gamma \in \Gamma(\mu,\nu)$ is b-optimal, then B2 holds at $(x_0,y_0) \in \text{spt}(\gamma) \Rightarrow \text{spt}(\gamma) \cap B_\epsilon(x_0,y_0) \subset S$ of dim n in $\mathbb{R}^{2n}$ a Lipschitz submanifold for $\epsilon \ll 1$. $S : \mathbb{R} \to (x(s),y(s))$ is a smooth curve in $\text{spt}(\gamma)$ with $(x(0),y(0)) = (x_0,y_0)$. Then

$$0 \leq \Delta_0(x(s),y(s)) = \frac{st}{2}(\dot{x}(0),\dot{y}(0))H_0\begin{pmatrix}\dot{x}(0)\\ \dot{y}0\end{pmatrix} + o(s^2+t^2) \text{ as } (s,t) \to 0$$

$$(\dot{x}(0),\dot{y}(0))H_0\begin{pmatrix}\dot{x}(0)\\ \dot{y}0\end{pmatrix} \geq 0$$

Assume $\text{spt}(\gamma) \subset \text{Graph}(G)$, $G : X \to Y$ smooth, $u(x) + v(y) - b(x,y) \geq 0$. Equality on $y = G(x)$.
F.O.C. gives $Du(x) = D_x b(x,G(x))$. S.O.C. gives $D^2u(x) - D^2_{xx}b(x,G(x)) \geq 0$. Since functions are smooth, differentiating F.O.C. gives $D^2u(x) - D^2_{xx}b(x,G(x)) = D^2_{xy}b(x,G(x))DG(x)$. Take det over both sides,

$$\det(D^2u(x) - D^2_{xx}b(x,G(x))) = \det D^2_{xy}b(x,G(x)) \det DG(x)$$

Let $d\mu(x) = f(x)d^n x, d\nu(y) = g(y)d^n y$. Since G is a smooth diffeomorphism, $|\det DG| = \frac{f(x)}{g(G(x))}$.
Twist (B1) $\Rightarrow G(x) = D_x b(x,\cdot)^{-1}Du(x) = Y_b(x,Du(x))$. Then

$$g(Y_b(x,Du(x)))\det(D^2u(x) - D^2_{xx}b(Y_b(x,Du(x)))) = |\det D^2_{xy}b| f(x)$$

This is the *Monge-Ampere equation* (prescribed Jacobian equation)

If $b(x,y) = x \cdot y$, $Y(x,y) = y$, $g(Du)\det D^2 u = f(x)$.

We need convexity of Y and $\log g, \log f \in C^\infty \cap L^\infty$ to have smoothness of u and G. (Caffarelli' 1990)

3 Asymmetric Information (Principal Agent Framework)

Monopolists: Single entity (seller) on one side of the market, parametrized by $y \in Y \subset \mathbb{R}^n$.
Agents: Large population of heterogenous agents (Each agent has different preferences) parametrized by $x \in X \subset \mathbb{R}^m$.

x is private information. Public knowledge is $\mu \in \mathcal{P}(X)$, $d\mu(x)$ =relative frequency of agent type $x \in X$.

Assume monopolists are selling cars. The public knowlege are: $\phi(x,y,z)$ =value of car y to agent x at price z. Assume $\frac{\partial \phi}{\partial z} < 0$.
$\pi(x,y,z)$ =monopolists' profit from selling y to x at price z

Assume there is an outside option $y_\phi \in Y$.

Agent problem:

$$u(x) = \sup_{y \in Y} \phi(x,y,v(y))$$

Agent x buys a car $y_{\phi,v}(x) \in \arg\max_{y \in Y} \phi(x,y,v(y))$. Hopefully uniquely attained μ-a.e.

Monopolists problem: choose a price menu $v : Y \to \mathbb{R}$ s.t. $v(y_\phi) = 0$ to maximize expected profits:

$$\sup_{v, v(y_\phi)=0} \int \pi(x, y_{\phi,v}(x), v(y_{\phi,v}(x))) d\mu(x)$$

If v is l.s.c, or Y is compact, then the supremum can be attained.

This is the *Monge* type formulation.

In *Kantorovich* setting:
Seek $\gamma \geq 0$ on $X \times Y$, $\Pi^X_{\#}\gamma = \mu$, $u(x) = \phi(x, y_{\phi,v}(x), v(y_{\phi,v}(x)))$. Let $v^\phi = u$, $u(x) + v(y) - b(x,y) = v^\phi(x) - \phi(x,y,v(y)) \geq 0$ and equality is achieved γ-a.e.

$$\sup_{v:Y \to \mathbb{R} l.s.c., v(y_\phi)=0} \sup_\gamma \int \pi(x,y,v(y)) d\gamma(x,y).$$

If Y compact, π, ϕ continous, v l.s.c. s.t. $v(y_\phi) = 0$, then supremum can be attained. Maximum problem on γ is an infinite-dimensional linear programming problem, given γ is convex and compact. However, supremum on v is more complicated.

Example: Consider the quasilinear case $\phi(x,y,z) = b(x,y) - z$, $\pi(x,y,z) = z - a(y)$, where a, b can be non-linear functions.
Monopolists: $\int \pi d\gamma = \int v(y) - a(y) d\gamma$
Agent x: $u(x) = \sup_y b(x,y) - v(y)$.
$u(x) + v(y) - b(x,y) \geq 0$ on $X \times Y$ and equality holds a.e.
Let $U = \left\{u : X \to \overline{R} : u = v^\phi \text{ for some } v : Y \to \overline{R} \text{ l.s.c.}\right\}$. If $b \in C^1$, then $u \in U$ are equi-Lipschitz. If $b \in C^2$, then $u \in U$ are equi-semiconvex. In both cases, U is convex via Ascoli-Arzela. Then

$$\int v(y) - a(y) d\gamma = \int b(x,y) - u(x) - a(y) d\gamma$$

Equality holds γ-a.e. $\Leftrightarrow$ spt$(\gamma) \subset \partial_b u$. When U is compact ($b \in C^1$), sup is attained.
spt$(\gamma) \subset \partial_b u \Rightarrow \gamma$ is b-optimal hence optimal transport theorey applies.

Rochet-Chone (1998) Formulation: Suppose $n = m$, $y_\phi = 0 \in \mathbb{R}^n$, $Y = [0,\infty)^n$, $d\mu(x) = f(x)dH^n$, $y_b(x) = Du(x)$. $u = v^*$ is convex. Solve for

$$\max_{u\geq 0, u \text{ convex}} \int_X (xDu(x) - u - a(Du(x)))f(x)dH^n x$$

The function is concave in u if $y \mapsto a(y)$ is convex. Choose $a(y) = \frac{1}{2}|y|^2$, we get

$$\begin{aligned}
&\max_{u,u',u''\geq 0} \int \left(xDu - u - \frac{1}{2}|Du|^2\right) \\
=&\max \int \left(-\frac{1}{2}(Du-x)^2 - \left(u - \frac{1}{2}|x|^2\right)\right) f \\
=&\min \int \left(\frac{1}{2}(Du-x)^2 + \left(u - \frac{1}{2}|x|^2\right)\right) f
\end{aligned}$$

Consider $m = n = 1$, $f(x) = \chi_{[a,a+1]}(x)$, then the optimization problem becomes:

$$\min \int_a^{a+1} \left(\frac{1}{2}(u'-x)^2 + \left(u - \frac{1}{2}x^2\right)\right)$$

Simpler Version (Obstacle problem): $\min\limits_{w\geq h \text{ on } \Omega} \displaystyle\int_\Omega \frac{1}{2}|Dw|^2 dx$ s.t. w vanishes on $\partial\Omega$ and $w \in W^{1,2}(\Omega)$. Then $\triangle w = 0$ on $\{w > h\}$, and $w = h$ on $\Omega \setminus \{w > h\}$. There are fewer constraints on w.

Let $L(u) = \displaystyle\int (c(Du(x)) + u(x) - xDu(x))f(x)dH^n x$.
The Rochet-Chone is solving for $-\min\limits_{u\geq 0 \text{convex}, \ Du(x)\subset \text{conv}(Y)} L(u)$. If c is convex, then $L(u)$ is convex, *i.e.* if both u_0, u_1 minimize, so does $(1-t)u_0 + tu_1, \forall t \in [0,1]$. If c is strictly convex, then $L(u_{1/2}) < L(u_0) + L(u_1)$ unless $Du_0(x) = Du_1(x)$ f-a.e.
Uniqueness: If $H^n \ll \mu \ll H^n$ and $X \subset \mathbb{R}^n$ convex (or open + connected), then $u_0 = u_1 + c_1$ inside X. $L(u_0) = L(u_1) + c_1$, $c_1 = 0$

Corollary 5. *If the solution is unique, it inherits any symmetries of the problem.*

Example: $X \subset Y = [0,\infty)^n$. Let $\hat{X} = \{(x_1,...,x_n) \in \mathbb{R}^n : (|x_1|,...,|x_n|) \in X\}$, $\hat{Y} = \mathbb{R}^n$.
Define $\hat{f}(x_1,...,x_n) = f(|x_1|,...,|x_n|)$, $\hat{L}(u) = \int_{\hat{X}}(c(Du(x)) + u(x) - xDu(x))\hat{f}(x)dH^n x$.
$\hat{u} \in \arg\min_{u\geq 0 \text{ convex}} \hat{L}(u)$. By uniqueness, $\hat{u}(x_1,...,x_n) = \hat{u}(|x_1|,...,|x_n|)$. This is an *unconditionally symmetric minimizer.*

Claim: $\hat{u}$ unconditional $\Rightarrow$ if $x \in X \subset Y$, then $Du(x) \in Y = [0,\infty)^n$.

Proof. Assume not, *i.e.* there exists $x \in X$ with $y = Du(x) \notin Y$, say $y_i < 0$.
Define $\hat{x} = (\pm x_1,...,\pm x_n)$. Choose $-$ if $y_i < 0$. $\hat{y} = D\hat{u}(\hat{x}) \in Y$.
$0 \leq \langle D\hat{u}(\hat{x}) - Du(x), \hat{x} - x\rangle < 0$. (from convexity of $\hat{u}$) Contradiction. ☐

> ***Theorem:* 3.1: Rochet-Chone**
>
> Let $X \subset Y = [0,\infty)^n$, c strictly convex, $c \in C^{1,1}(\mathbb{R}^n)$. $u_0 \in \arg\min\limits_{u\geq 0 \text{ convex}} L(u) \Leftrightarrow L(u_0) \leq L(u_0 + w)$ $\forall w \geq 0$ convex or w convex and $\text{spt}(w) \subset \{u > 0\}$.

Proof. $(\Rightarrow)$ is obvious. Consider $\Rightarrow$.
Recall that to minimize $E(x)$ on $\Omega \subset \mathbb{R}^n$ a compact convex set:

1. If $E \in C^1(\mathrm{int}\Omega)$, then $DE(x_0) = 0 \Leftrightarrow x_0 \in \mathrm{int}\Omega$ is a minimizer
2. If $E \in C^1(\overline{\Omega})$ and $\partial\Omega \in C^1$, then $DE(x_0) = \lambda \hat{n}_\Omega(x_0)$, where $\lambda \geq 0$, $\lambda = 0$ unless $x_0 \in \partial\Omega \Leftrightarrow x_0 \in \partial\Omega$ is a minimizer.
3. Non-smooth version: ELKKT (Euler-Lagrange-Karush-Kuhn-Tucker) E, Ω convex, $\partial E(x_0) \cap N_\Omega(x_0) \neq \emptyset$, where $N_\Omega(x_0) = \{v \in \mathbb{R}^n : v(x - x_0) \leq 0, \forall x \in \Omega\}$ (cone of generalized normals) $\Leftrightarrow x_0$ minimizes E.

Let $U = \{u \geq 0 \text{ convex on } \mathbb{R}^n\}$, U is a convex cone. By ∞-dim version of ELKKT, $DL(u_0) \cap N_U(x_0) \neq \emptyset$ $\Leftrightarrow u_0 \in \arg\min_{u \in U} L(u)$ and $v \in N_u(x_0)$ has $v(u - u_0) \leq 0, \forall u \in U$. $\square$

Consider a small perturbation on $L(u)$:

$$\begin{aligned} \frac{d}{d\epsilon}\bigg|_{\epsilon=0} L(u + \epsilon w) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_X (c(Du + \epsilon Dw) + u + \epsilon w - xDu - \epsilon x Dw) f dx \\ &= \int_X (Dc|_{Du} Dw + w - xDw) f dx \\ &= \int_{\partial X} (Dc|_{Du} - x)\hat{n}_X(x) f(x) w(x) dH^{n-1}x + \int_X (f + \nabla \cdot ((x - Dc|_D u) f)) w dx \\ &= \int_{\mathbb{R}^n} \frac{\delta L}{\delta u} w dx, \text{ where} \\ \sigma := \frac{\delta L}{\delta u} &= (Dc|_{Du} - x)\hat{n}_x f dH^{n-1}|_{\partial X} + (f + \nabla \cdot ((x - Dc|_D u) f)) dH^n|_X \end{aligned}$$

If $u > 0$ and $D^2 u \geq \lambda I > 0$ on $\Omega \subset X$ open and smooth, then $u + \epsilon w \in U$. If $w \in C_C^2(\Omega)$, then $\frac{d}{d\epsilon}\big|_{\epsilon=0} L(u_0 + \epsilon w) = 0$. This gives the E-L equation:

$$\begin{cases} 1 = f^{-1} \nabla \cdot ((x - Dc|_{Du}) f) = \nabla \cdot (Dc|_{Du} - x) + \langle Dc|_{Du}, D \log f \rangle \text{ on } \Omega \\ (Dc|_{Du} - x) \cdot \hat{n}_X f = 0 \text{ on } \partial X. \end{cases}$$

Suppose $c(y) = \frac{1}{2}|y|^2$, $Dc(y) = y$, $f = \chi_X$. These give a Poisson equation + Neumann BC:

$$\begin{cases} \Delta u = n + 1 \\ (Du - x) \cdot \hat{n} = 0 \end{cases}$$

Assume $n = 1$.

In the homogenous case, $\mu = \delta_{x_0}$. $L(u) = c(y) + u(x_0) - x_0 \cdot y$, minimum is attained when $Dc(y_0) = x_0$. $u(x) = \max(0, y_0(x - x_0))$, *i.e.* $y_0 = (Dc)^{-1}(x_0)$. It is the undistorted choice of consumer x_0.

In heterogenous case, $X = [a, a + 1] \subset Y = [0, \infty)$.

$$\begin{cases} u'' = 2 \text{ on } X \\ (u' - x) \cdot \hat{n} = 0 \end{cases}$$

When $a \leq 1$, $u(x) = \begin{cases} \left(x - \frac{a+1}{2}\right)^2, x \geq \frac{a+1}{2} \\ 0, \text{ else} \end{cases}$. A positive fraction of consumers don't buy (buyer's market).

When $a > 1$, $u(x) = \left(x - \frac{a+1}{2}\right)^2 - \left(\frac{a-1}{2}\right)^2$ (seller's market)

Distortion: Always downwards. $x - u'(x) = a + 1 - x \geq 0$. As a increases, more seller's market, more distortion.

$\sigma = \frac{\delta L}{\delta u} = (n + 1 - \Delta u) dH^n|_X + (Du - x)\hat{n}_X dH^{n-1} x|_{\partial X}$ is a measure of finite total variation.

For a given u_0, we define the equivalence relation and class by $x \sim x' \Leftrightarrow Du_0(x) = Du_0(x')$, $\tilde{x} = \{x' \in X : Du_0(x) = Du_0(x')\} = \partial u_0^*(Du_0(x)) \subset \mathbb{R}^n$, where $u_0^*(y) = \sup_{x \in X} x \cdot y - u_0(x)$.

Let $X_i = \Omega_i = \{x \in X : \dim \tilde{x} = n - 1\}$.

***Definition:* 3.1: Convex Order**

Let $\mu, \nu \in m_+(X)$, $X \subset \mathbb{R}^n$, U any convex cone, $\mu \leq_U \nu \Leftrightarrow \int u d\mu \leq \int u d\nu$ for all $u \in u \subset C(X)$.

Example: If $U = \{u : X \to \mathbb{R} \text{ convex}\}$, $\leq_{cx} = \leq_U$ is called convex order or second order stochastic dominance.

***Definition:* 3.2: Sweeping Operator**

A sweeping operator is an operator $T : x \in X \to Tx \in \mathcal{P}(X)$ s.t.

1. $\forall E \subset X$ Borel, $x \to Tx(E)$ is also Borel
2. $x = \int_X z dTx(z)$, $\forall x \in X$

Given $\omega \in \mathcal{P}(X)$, define $T\omega \in \mathcal{P}(X)$ by

$$\int_X \phi dT\omega = \int_X \left(\int_X \phi(z) dTx(z) \right) d\omega(x)$$

***Theorem:* 3.2: Strassen**

Let $\mu, \nu \in \mathcal{P}(X)$ s.t. $\mu \leq_{cx} \nu$ if and only if there exists a sweeping operator T s.t. $T\mu = \nu$.

***Lemma:* 3.1: Restoring Neutrality**

Let $U_0 = \{u : U + u \geq 0\}$. There exists a Lagrange multiplier $\lambda \in m_+(X)$ s.t. $u \in \arg\min_{U_0} L(u) \Leftrightarrow u \in \arg\min_{U} L(u) - \lambda u$. In fact, $\lambda \in m_+(\{u_0 = 0\})$.

Corollary 6. *For* $\frac{\delta L}{\delta u} = \sigma = \sigma^+ - \sigma^-$, $\exists \lambda \in \mathcal{P}(\{u_0 = 0\})$, $\sigma - \lambda = (\sigma - \lambda)^+ - (\sigma - \lambda)^- = \omega^+ - \omega^-$ *s.t.* $\omega^- \leq_U \omega^+$, $\sigma^- \leq_U \sigma^+$.

Corollary 7. $(Du_0)_\# \sigma = \delta_0$ *and* $(Du_0)_\# \omega^- = (Du_0)_\# (T\omega^-) = (Du_0)_\# \omega^+$.

***Lemma:* 3.2:**

If $\omega^+ = T\omega^-$ for a sweeping operator T, then $Tx(X - \tilde{x}) = 0$ ω^--a.e. x. *i.e.* sweeping only occurs within equivalent classes.

Theorem: 3.3: Disintegration of Measures

Given X, Y Polish, $F : X \to Y$ Borel, $\mu \in \mathcal{P}(X)$. Then $\exists \{\mu_y\}_{y \in Y} \subset \mathcal{P}(X)$, $\nu = F_{\#}\mu$-a.e. y satisfies $\mu_y(F^{-1}(y)) = 1$ and ν-a.e. $\mu_y \in \mathcal{P}(X)$ is unique s.t. $\forall$ Borel test functions $\phi \geq 0$ on X.

$$\int_X \phi d\mu = \int_Y \left(\int_X \phi(z) d\mu_y(z) \right) d(F_{\#}\mu)(y)$$

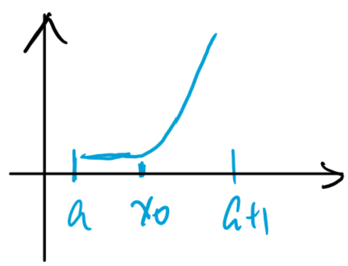
This is similar to Fubini's and Bayes' theorem. All conditional probability measures are unique. For us $F = Du_0, \mu = \omega^+, \nu = \omega^-$.

Corollary 8. *Not only $\omega^- \leq_{cx} \omega^+$, but the conditional measures $\omega_y^- \leq_{cx} \omega_y^+$ for $\nu = (Du_0)_{\#}\omega^-$-a.e. y.*

Example: For $n = 1$, $f = \chi_{[a,a+1]}$, $\sigma = \frac{\delta L}{\delta u} = (2 - u'')dH|_X + (u' - x)\hat{n}dH^0x$. $\Omega_0 = [a, x_0]$, $\Omega_1 = (x_0, a+1]$. Figure 1.

$$\sigma|_{\Omega_0} = 2dH|_{[a,x_0]} - x \cdot \hat{n}_a = 2\chi_{[a,x_0]} + adH^0$$
$$1 = \int \sigma|_{\Omega_0} = 2(x_0 - a) + a = 2x_0 - a$$
$$x_0 = \frac{a+1}{2}$$

Figure 1: 1D Example



Example:For $n = 2$, $f = \chi_X$, where $X = [a, a+1]^2$. Figure 2.
Assume $a > \frac{7}{2}, \Omega_0 = \{u_0 = 0\}$. Then

$$\sigma|_{\Omega_0} = 3dH^2|_{\Omega_0} - x \cdot \hat{n}dH^1|_{\Omega_0 \cap \partial X}$$
$$1 = \int \sigma|_{\Omega_0} = 3h^2 + 2ah$$
$$h = \frac{2a}{3}\left(-1 + \sqrt{1 + \frac{3}{2a^2}}\right) \sim \frac{1}{2a} \text{ as } a \to \infty$$

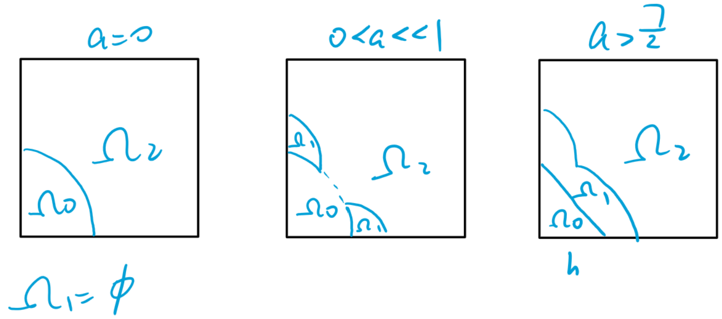
Summarization With $b(x, y) = x \cdot y$, $a(y) = \frac{1}{2}|y|^2$, $f = 1_\Omega$, where $\Omega \subset \mathbb{R}^n$ is open and convex.

$$E(u) = \int_\Omega \left(\frac{1}{2}|Du - x|^2 + \left(u - \frac{1}{2}|x|^2\right)\right) dx$$

We want to minimzie $E(u)$ on U_0, where $U = \{u : \Omega \to \mathbb{R}^n \text{ convex}\}$, $U_0 = \{u : u + U \geq 0\}$. Define

$$E'_u(w) = \frac{d}{d\lambda}\bigg|_{\lambda=0} E(u + \lambda w) = \int_{\overline{\Omega}} w d\sigma + \int_\Omega (n + 1 - \Delta u) w dH^n(x) + \int_{\partial\Omega} (Du - x)\hat{n} w dH^{n-1}x$$

Figure 2: 2D Example



Lemma: 3.3: Variational Lemma

If $u \in U_0$, $E(u) \leq E(\overline{u})$ forall $\overline{u} \in U_0, w = \overline{u} - u$, then

1. $E'_u(w) \geq 0$
2. $E'_u(\overline{u}) \geq 0$
3. $E'_{\overline{u}}(-w) < 0$ strictly unless $\overline{u} = u$+const

Corollary 9. *Normal distortion is never inward, i.e.* $(Du(x) - x) \cdot \hat{n}_\Omega(x) \geq 0, \forall x \in \partial\Omega$ *s.t.* $\hat{n}$ *is unique*

Proof. WLOG, assume $\partial\Omega \in C^1$, strictly convex (otherwise approximate)
Let $x_0 \in \partial\Omega$. Define $v(x) = u(x) - u(x_0) - Du(x_0)(x - x_0) \geq 0$, $v(x_0) = 0$
WLOG, assume $x_0 = 0$, $\hat{n}(x_0) = (1, 0, 0, \dots)$. If not, perform translation.
Set $p_0(x) = u(x_0) + Du(x_0)(x - x_0)$ the Taylor expansion. $\hat{u}_t(x) = \frac{n+1}{2}(x_1 + t)^2 + p_0(x)$, and $\overline{u}_t = \max\{u, u_t\}$.
$U_t = \{x : \overline{u}_t > u\} \subset \{x : -t \leq x_1 \leq 0\}$. Hence, $\lim_{t \to 0} U_t = \{x_0 = 0\}$.
By Lemma 3.3,

$$0 < E'_{\overline{u}_t}(w) = \int_{U_t \cap \partial\Omega} (D\hat{u}_t - x) \cdot \hat{n}_\Omega dH^{n-1}(x),$$

if $w = \overline{u}_t - u = \begin{cases} \hat{u}_t, & \text{in } U_t \\ 0, & \text{else} \end{cases}$ □

By Theorem 3.1 and Theorem 3.3,

$$u \in \arg\min_{U_0} E(u) \Leftrightarrow E'_u(w) \geq 0, \forall u + w \in U_0 \Leftrightarrow \int w d\sigma_{\tilde{x}} \geq 0 \text{ a.e.-}x$$

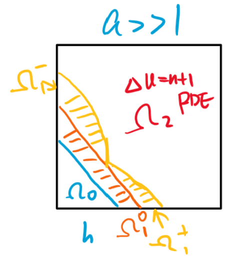
The last inequality is also equivalent to $\int w d\sigma^+_{\tilde{x}} \geq \int w d\sigma^-_{\tilde{x}}$.

In the equations, $\tilde{x} = \left\{x' \in \overline{\Omega} : Du(x) = Du(x')\right\}$ and

$$\int_{\overline{\Omega}} \phi d\sigma^\pm = \int_{Du(\Omega)} \left[\int_{\tilde{x}} \phi(z) d\sigma^\pm_{\tilde{x}}(z)\right] d(Du_\# \sigma^\pm)(x)$$

Consider the 2D example with $a \gg 1$ (Figure 3). Split Ω_1 into Ω_1^0 (where everything is well-behaved) and $\Omega_1^\pm$ (symmetric regions).

Figure 3: 2D Example $a \gg 1$



In Ω_1^0: Perform a transformation of basis, let $z = x_1 + x_2, w = \frac{x_1 - x_2}{2}$, $dx_1 dx_2 = 2dzdw$.

$$u(x_1, x_2) = g(x_1 + x_2) = g(z), Du = (g', g'), \Delta u = 2g''$$

$$\begin{aligned} d\sigma(z, w) &= (3 - \Delta u) dx_1 dx_2|_{\Omega_1^0} + (Du - x) \cdot \hat{n} dH^1|_{\Omega^0 \cap \partial\Omega} \\ &= 2(3 - 2g'')dzdw + (a - g') \left(\left| \frac{1}{2}dz - dw \right|_{x_2 = a} + \left| \frac{1}{2}dz + dw \right|_{x_1 = a} \right) \end{aligned}$$

Fix z, integrate over $\tilde{x}$, *i.e.* over $w \in \left(a - \frac{z}{2}, a + \frac{z}{2}\right)$.

$$0 = \int_{\tilde{x}} d\sigma = \int_{a - \frac{z}{2}}^{a + \frac{z}{2}} \sigma(z, dw) = 2(3 - 2g'')(z - 2a) + 2(a - g')$$

This is an ODE for g, with BCs: $g(z_0) = g'(z_0) = 0$. $\Omega \cap \partial\Omega_0 = \{z = z_0\}$.
Homogenous part solved by power law, ansatz with power -1. This leads to

$$g'(z) = \frac{3}{4}z - \frac{a}{2} + \frac{\text{const}}{2(z - 2a)}$$

In Ω_1^-: Let $\overline{x}(r, \theta(t)) = (a, h(\theta)) + r(\cos\theta, \sin\theta)$ and $u(\overline{x}(r, \theta)) = b(\theta) + m(\theta)r = \overline{u}(r, \theta)$.
Assume $(r, \theta) \mapsto \overline{x}(r, \theta)$ is locally bi-Lipschitz.

$$dH^2|_{\Omega_1^-} = \left| h' \cos\theta + r \right| dr d\theta$$

$$dH^1|_{\Omega_1 \cap \partial\Omega} = \left| h'(\theta) \right| d\theta$$

$$Du(\overline{x}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} m \\ m' \end{bmatrix}$$

$$D^2 u(\overline{x}) = \frac{m'' + m}{h' \cos\theta + r} \begin{bmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{bmatrix}$$

$$b'(\theta) = h'(\theta) \frac{\partial u}{\partial x_2}(\overline{x})$$

This is because $\frac{\partial(x_1, x_2)}{\partial(r, \theta)} = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & h' + r\cos\theta \end{bmatrix}$ and $\det = h' \cos\theta + r$, and we apply chain rule and matrix inversion to get all the equalities.

Therefore,

$$\begin{aligned} d\sigma(r, \theta) &= (n + 1 - \Delta u) dH^2|_\Omega + (Du - x) \cdot \hat{n} dH^1|_{\partial\Omega} \\ \pm d\sigma(r, \theta) &= \left(3 - \frac{m'' + m}{h' \cos\theta + r}\right)(h' \cos\theta + r) dr d\theta + (Du - \overline{x}) \cdot \hat{n}_\Omega \delta_0(r) h'(\theta) d\theta dr \\ &= (3(h' \cos\theta + r) - m'' - m) dr d\theta + (Du - \overline{x}) \cdot \hat{n}_\Omega \delta_0(r) h'(\theta) d\theta dr \end{aligned}$$

For Ω_1^- in $a \gg 1$, $0 < r < R(\theta)$, $\theta \in [\theta_0, 0]$. Since $\sigma_{\bar{x}}^+ \geq \sigma_{\bar{x}}^-$, the singular term is in $\sigma_{\bar{x}}^+$ and not in $\sigma_{\bar{x}}^-$. Either $h' > 0$ and choose $d\sigma(r,\theta)$, or $h' < 0$ and choose $-d\sigma(r,\theta)$, but as we will see, the second choice can be ruled out.

$$0 = \pm \int_0^{R(\theta)} \sigma(dr, \theta) = (3h' \cos\theta - m'' - m)R + \frac{3}{2}R^2 + (Du - \overline{x}) \cdot \hat{n}_\Omega h'(\theta)$$
$$0 = \pm \frac{1}{R^2} \int_0^{R(\theta)} r\sigma(dr, \theta) = (3h' \cos\theta - m'' - m)\frac{1}{2} + R$$

Given $R : [\theta_0, 0] \to [0, \infty)$ Lipschitz, and $\theta_0, h(\theta_0)$, we can solve for $m(\theta), h(\theta)$ subject to ICs:

1. $a \gg 1$, $m(\theta_0) = 0$, $m'(\theta_0)$ depends on the initial slope of the line
2. $a \ll 1$, $m(\theta_0) = m'(\theta_0) = 0$

Then $(Du - \overline{x}) \cdot \hat{n}_\Omega h'(\theta) = \frac{R^2}{2} > 0$, so $-d\sigma(r,\theta)$ cannot happen.

For Ω_2, $\begin{cases} \Delta u = 3 \\ (Du - x) \cdot \hat{n}_\Omega = 0 \text{ on } \partial\Omega \end{cases}$, $u_1 = u_2$, $\frac{\partial u_1}{\partial u} = \frac{\partial u_2}{\partial u}$ on $\partial\Omega_1 \cap \partial\Omega_2$, $u \in C^{1,1}_{loc} \cap C^1(\Omega)$.

$b' = h' \frac{\partial u}{\partial x_2} = h'(m \sin\theta + m' \cos\theta)$. Once h', m, m' are solved,

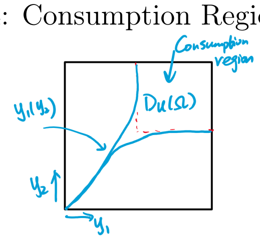
$$b(\hat{\theta}) = b(\theta_0) + \int_{\theta_0}^{\hat{\theta}} h'(\theta)(m \sin\theta + m' \cos\theta) d\theta$$

On a Lipschitz domain Ω_2 and u_1 on Ω_1, there exists a unique $u_2 + const$ s.t. $\Delta u_2 = 3$, $(Du - x) \cdot \hat{n}_\Omega = 0$ on $\partial\Omega \cap \partial\Omega_2$ (**Fixed Boundary**), and $\frac{\partial u_1}{\partial u} = \frac{\partial u_2}{\partial u}$ on $\partial\Omega_1 \cap \partial\Omega_2$ (**Free Boundary**)

Claim: At most one such choice $\{R(\theta)\}_{\theta_0 < \theta < 0}$, $h(\theta_0)$ and θ_0 exist, such that $u_2 = u_1 + const$ on free boundary and u is convex.
Conjecture: There is at least one such choice.

Corollary 10. *Monotonicity and concavity of stingray's tail. (Figure 4)*

Figure 4: Consumption Region $a \gg 1$



Proof.

$$e(\theta) = y_2 = \frac{\partial u}{\partial x_2} = m' \cos\theta + m \sin\theta \geq 0$$
$$f(\theta) = a - y_1 = (Du - \overline{x}) \cdot \hat{n} = m' \sin\theta - m \cos\theta + a \geq 0$$
$$-\frac{dy_1}{dy_2} = \frac{df}{de} = \frac{f'(\theta)}{e'(\theta)} = \frac{(m'' + m)\sin\theta}{(m'' + m)\cos\theta} = \tan\theta < 0$$
$$-\frac{d^2 y_1}{dy_2^2} = \frac{d^2 f}{de^2} = \frac{1}{e'(\theta)}\frac{d}{d\theta}\tan\theta = \frac{1}{(m'' + m)\cos^3\theta} > 0$$

□

Lemma: 3.4:

If $u : \mathbb{R}^n \to \mathbb{R}$ is convex ($u = u^{**}$) and $u^*(y) = \sup_{x \in \mathbb{R}^n} \langle x, y \rangle - u(x)$ is its dual. If $(x, y) \in \partial u = \left\{(x_0, y_0) \in \mathbb{R}^{2n} : u(z) \geq u(x_0) + \langle y_0, z - x_0 \rangle, \forall z \in \mathbb{R}^n\right\}$ and $Q = Q^\dagger > 0$, then for $\delta_x = z - x$, the following two statments are equivalent

$$u(x + \delta_x) \geq u(x) + \langle y, \delta_x \rangle + \frac{1}{2} \langle \delta_x, Q\delta_x \rangle + o(\delta_x^2) \text{ as } \delta_x \to 0$$
$$u^*(y + \delta_y) \leq u(y) + \langle x, \delta_y \rangle + \frac{1}{2} \left\langle \delta_y, Q^{-1}\delta_y \right\rangle + o(\delta_y^2) \text{ as } \delta_y \to 0$$

Proof. We only need to prove the $\Rightarrow$ direction, and the other direction follows from duality.

WLOG, take $x = y = 0, u(0) = 0$. Assume $u(\delta_x) \geq (1 - s)\frac{1}{2}\delta_x Q \delta_x$ holds on neighborhood U_s of 0.
For $\delta_u \in \partial u(U_s) = \bigcup_{x \in U_s} \partial u(x)$, where $\partial u(x) = \{y \in \mathbb{R}^n : (x, y) \in \partial u\}$,

$$u^*(\delta_y) = \sup_{\delta_x \in U_s} \langle \delta_x, \delta_y \rangle - u(\delta_x) \leq \sup_{\delta_x \in U_s} \langle \delta_x, \delta_y \rangle - \frac{1 - s}{2}\delta_x Q \delta_x$$

F.O.C. gives $\delta_y = (1 - s)Q\delta_x$ if $\delta_y \in ((1 - s)QU_s) \cap \partial u(U_s)$, and

$$u^*(\delta_y) \leq \frac{1}{2} \left\langle \delta_y, ((1 - s)Q)^{-1}\delta_y \right\rangle$$

□

Lemma: 3.5: Interior Consumption

Fix u optimal. Let $x_0 \in \Omega \cap \mathrm{dom} D^2 u$. If $Du(x_0) \in \mathrm{int}(D(\Omega))$, then $\Delta u \geq n + 1$. If in addition, $\tilde{x_0} = \{x_0\}$, then $\Delta u(x_0) = n + 1$.

Proof. Assume $x_0 \in \Omega \cap \mathrm{dom} D^2 u$ with $Du(x_0) \in \mathrm{int}(D(\Omega))$, but $\Delta u + sn < n + 1$ for some $s > 0$.
WLOG, take $x_0 = u(x_0) = Du(x_0) = 0$.
Then $u(x) < \frac{1}{2} \left\langle x, (D^2 u(x_0) + sI)x \right\rangle$ near $x_0 = 0$.
Let $\overline{v}(y) = \frac{1}{2} \left\langle y, (D^2 u(x_0) + sI)^{-1} y \right\rangle$, then $u^*(y) > \overline{v}(y)$ in a punctured neighborhood of 0.

The connected component y^h of $\{u^* < \overline{v} + h\}$ containing $y_0 = 0$ shrinks to $\{0\}$ as $h \to 0$.
Let $v_h(y) = \begin{cases} \overline{v}(y) + h \text{ on } y^h \\ u^*(y) \text{ otherwise} \end{cases}$, the max of u^* and $\overline{v} + h$, and $u_h = v_h^* \leq u^*$ strictly at $x_0 = 0$.
However, $D^2 v_h = (D^2 u(0) + sI)^{-1}$ throughout y^h, so $\Delta u_h \leq \Delta u(x_0) + ns < n + 1$ on $\{u_h < u\}$.
But $0 < E'_{u_h}(u_h - u) = \int (n + 1 - \Delta u_h)(u_h - u) + 0$ which gives a contradiction, since $n + 1 - \Delta u_h > 0$, but $u_h - u < 0$. □

Theorem: 3.4: PDE Laplacian

$u \in C^2(\Omega) \cap C(\overline{\Omega})$ satisfies $\Delta u = 0$ a.e. if and only if $\forall x_0 \in \Omega$ and $B_r(x_0) \subset \Omega$, $u(x_0) = \int_{\partial B_r(x_0)} u(x) dH^{n-1}(x)$.

Corollary 11. *u as above is $C^\infty(\Omega)$ (smooth).*

Corollary 12. *If u as above satisfies $u \geq 0$ on Ω, but $u(x_0) = 0$, then $u = 0$ throughout Ω.*

Theorem: 3.5: All Non-trivial Bunches Reaches $\partial\Omega$

Suppose $\Omega \subset \mathbb{R}^n$ is a convex, open, bounded subset.

1. $\Omega_0 = \left\{x \in \overline{\Omega} : H^n(\tilde{x}) > 0\right\} \subset \{x : u = 0\}$, where $\tilde{x}$ is the equivalence class.
2. If $x \in \Omega_1 \cup \cdots \cup \Omega_{n-1}$, then $\tilde{x} \cap \partial\Omega \neq \emptyset$
3. $\Omega_n = \left\{x \in \overline{\Omega} : \tilde{x} = \{x\}\right\}$ is relatively open in $\overline{\Omega}$ and $u \in C^\infty(\Omega_n \cap \Omega)$

Proof. 1. Let $x_0 \in \Omega_0$. Choose $w = \pm u$, if $-u$ also works, then the equality holds

$$0 \leq E'_u(w) = \int_\Omega (n + 1 - \Delta u) w dH^n + \int_{\partial\Omega} (Du - x) \cdot \hat{n}_\Omega w dH^{n-1}$$

$$0 = E'_u(u)|_{\tilde{x}_0} = \int_{\tilde{x}_0} (n + 1 - \Delta u) u + \int_{\tilde{x}_0 \cap \partial\Omega} (Du - x) \cdot \hat{n}_\Omega u$$

Note that $\Delta u = 0$ on the equivalence class, so all terms are nonnegativem and hence $u = 0$ a.e. $\tilde{x}_0$. $-u$ is affine on the equivalence class, so the equality holds.

2. For a contradiction, suppose $\exists \{x_0\} \neq \tilde{x}_0 \subset \Omega$ and $\tilde{x}_0 \cap \partial\Omega = \emptyset$. Then $y_0 = Du(x_0) \in \text{int}(Du(\Omega))$.
But $u \in C^{1,1}$ and $X^h = \left\{x \in \overline{\Omega} : u(x) \leq u_h(x) = u(x_0) + Du(x_0)(x - x_0) + h\right\}$
Then $\tilde{x}_0 \subset \Omega$ as $h \to 0$. $Du(X^h) \subset Du(\Omega)$ for $0 < h \ll 1$ compactly.
By Lemma 3.5, $\Delta u \geq n + 1$ a.e. X^h. Setting $\overline{u} = \max(u, u_h)$ yields.:

$$0 \leq E'_u(\overline{u} - u) = \int_{X^h} (n + 1 - \Delta u)(\overline{u} - u) dH^n + \int_{X^n \cap \partial\Omega} (Du - x) \cdot \hat{n}_\Omega (\overline{u} - u) dH^{n-1}$$

Note $\overline{u} > u$, $\Delta u \geq n + 1$ and the boundary term vanishes, so RHS≤ 0.
To make the equality true, we need $\Delta u = n + 1$ a.e. on X^h.
Let $j \in \{1, ..., n\}$, $\Delta u_{jj} = 0$, because $u_{jj} = \partial^2_{jj} u \geq 0$. Either $u_{jj} > 0$ on X^h or $X^h \cap \partial\Omega \neq \emptyset$. Both are contradictions.

3. Recall $\partial u = \{(x, y) : u(z) \geq u(x) + y(z - x), \forall z \in \mathbb{R}^n\} \subset \mathbb{R}^n \times \mathbb{R}^n$ is closed.
Claim: $x_0 \in \Omega_n \cap \Omega \Rightarrow Du(x_0) \in \text{int}(Du(\Omega))$.
Define $\overline{u}(x) = \begin{cases} u(x), x \in \overline{\Omega} \subset \mathbb{R}^n \\ \infty, \text{ else} \end{cases}$, $\overline{u}^*(y)$ is Lipschitz.
$\forall y_k \to Du(x_0)$, $\exists x_k \in \Omega$ s.t. $(y_k, x_k) \in \partial\overline{u}^*$ *i.e.* $(x_k, y_k) \in \partial\overline{u}$.
If $x_k \to x_\infty$, then $(x_\infty, Du(x_0)) \in \partial\overline{u} \Rightarrow x_\infty = x_0$.
If $x_0 \in \text{dom}D^2 u$, then $\Delta u(x_0) = n + 1$. Therefore, $\Delta u = n + 1$ H^n-a.e. on $\Omega_n \cap \Omega$.
Openness of $\Omega_n \cap \Omega$ follows from 2. Hence $u \in C^\infty(\Omega_n \cap \Omega)$. ☐

Lemma: 3.6:

$R(\theta)$ which is the diameter of equivalence class of $\overline{x}(\theta)$ is upper semi-continous.

Proof. If u convex is affine on $[a_k, b_k] \subset \mathbb{R}^n$, and $(a_k, b_k) \to (a_\infty, b_\infty)$, then u is affine on $[a_\infty, b_\infty]$. ☐

Lemma: 3.7:

Two disjoint segments of length $\geq 2\delta$ in the plane whose midpoints are distance 2ϵ aprt make an angle 2θ s.t. $\arctan\theta \leq \frac{\epsilon}{\delta}$. *i.e.* as $\epsilon \to 0$, θ(mid point) has Lipschitz constant $\frac{1}{\delta}$.

Lemma: 3.8:

$0 \le (Du(x_0) - x_0) \cdot \hat{n} \le C\text{diam}(\tilde{x}_0)$, where $C = \sup_{B_\epsilon(x_0)} \Delta u$.

Let $w = u - u_1 \ge 0$. If $\Delta w = f(x)\chi_{\{u>0\}} \ge 0$, $f \in C(x_0 = 0)$, then $\tilde{u}(x) = \lim_{r\to 0} \frac{u(rx)}{r^2}$ exists subsequentially, if $u \in C^{1,1}$, $u \ge 0$, then $\tilde{u} = f(0)\chi_{\{\tilde{u}>0\}}$ on $\mathbb{R}^n$.

Theorem: 3.6: Caffarelli (1972)

Either $\tilde{u}$ is convex and quadratic or half parabola. *i.e.* after rotation and translation, $\tilde{u}(x_1, ..., x_n) = \begin{cases} \frac{f(0)}{2} x_1^2, x_1 > 0 \\ 0, \text{ else} \end{cases}$

If $x_0 = 0 \in \partial\{u > 0\}$ free boundary, then it blows up.

Define Lebesgue upper density $\overline{\theta}(x_0, z) = \limsup_{r\to 0} \frac{H^n(Z \cap B_r(x_0))}{H^n(B_r(x_0))}$, and Lebesgue lower density $\underline{\theta}(x_0, z) = \liminf_{r\to 0} \frac{H^n(Z \cap B_r(x_0))}{H^n(B_r(x_0))}$. Then, at regular free boundary points, $\overline{\theta}(x_0, \{u > 0\}) = \frac{1}{2}$, and free boundary may be Lipschitz; at singular free boundary points $\underline{\theta}(x_0, \{u > 0\}) = 0$, and free boundary is non-Lipschitz at x_0.

Remark 5. If $R(\theta)$ is monotone on an interval $J \subset S'$, then $R \in C(\overline{J})$.

Theory:

1. w detaches quadratically from $\{w = 0\} \subset \mathbb{R}^n$ provided $f(x) \ge c_0 > 0$
2. The Hausdorff dimension of $\partial\{w = 0\} < n - \epsilon(c_0) < n \Rightarrow R(\theta)$ is continous H^1-a.e.
3. Caffarelli, Kinderlehrer, and Nirenberg (1977): $\forall k \in \{0, 1, 2, ...\}$, if $f \in C^{k+2,\alpha}$ and $\partial\{u = 0\}$ is C^1, then $\partial\{u = 0\}$ is $C^{k+1,\alpha}$. Specifically, Cafferalli shows $f \in C^{2,\alpha} \Rightarrow \partial\{u = 0\} \in C^1$.

Theorem: 3.7: Bootstrapping

If $R \in C^{0,1}$, then R is smooth.

Proof. Suppose $R \in C^{0,1}$ is a neighborhood of a tame ray, then $\overline{x}(0, r)$ is bi-Lipschitz or a smaller neighborhood.

$$\frac{R^2}{2} = h'(Du - x) \cdot \hat{n} \quad 3 - \Delta u = \frac{3r - 2R}{h' \cos\theta + r}$$

$\Rightarrow R \in C^{0,1}$, $u \in C^{1,1} \Rightarrow h \in C^{1,1} \Rightarrow$ coordinates improve to bi-$C^{1,1}$.

Let $r = R$, $3 - \Delta u = \frac{R}{h' \cos\theta + R} \in C^{0,1}$, $u_1 \in C^{2,\alpha}(B_\epsilon(\theta_0, R(\theta_0)))$ for some ϵ and $\alpha \in (0, 1)$.
$\Rightarrow R \in C^{1,\alpha}$, provided the ray is transverse to free boundary by CKN-1977. Then $h \in C^{2,\alpha} \Rightarrow u \in C^{3,\alpha} \Rightarrow R \in C^{2,\alpha}$. □

Lemma: 3.9:

If $H^{\dim(\tilde{x})}(\tilde{x} - \text{dom}D^2 u) = 0$, then $\forall \xi \in \mathbb{R}^n$, $\partial^2_{\xi\xi} u = \xi D^2 u \xi$ agrees a.e. on $\tilde{x}$ with convex function. In fact, the relative interior $\text{relint}(\tilde{x}) \subset \text{dom}D^2 u$.

Corollary 13. *Δu is conves on relint$(\tilde{x})$.*

Proof. Fix $x_0, x_1, x_t \in \tilde{x} \cap \mathrm{dom} D^2 u$, $\xi \in \mathbb{R}^2, r > 0$ with $x_t = (1-t)x_0 + tx_1, t \in (0,1)$.

$$u(x_t + r\xi) = u(x_t) + rDu(x_t)\xi + \frac{r^2}{2}\partial^2_{\xi\xi}(x_t) + o(r^2)$$
$$u(x_t + r\xi) \leq (1-t)u(x_0 + r\xi) + tu(x_1 + r\xi)$$

Substituting the first equation into the second inequality, we get $\partial^2_{\xi\xi}(x_t) \leq (1-t)\partial^2_{\xi\xi}u(x_0) + t\partial^2_{\xi\xi}u(x_1)$. □

***Proposition:* 3.1:**

Let $x_0 \in \partial\Omega$ and $\epsilon > 0$, $(Du - x)\cdot \hat{n}_\Omega = 0$ throughout $B_\epsilon(x_0) \cap \partial\Omega$. Then $\tilde{x}_0 = \{x_0\}$, *i.e.* $x_0 \in \Omega_2$.

Proof. Take $x_0 \in \partial\Omega$ and $\epsilon > 0$ as above, $w = \Delta u - 3$ is convex on $\tilde{x}$ by Lemma 3.9.

$$0 \leq E'_w(u)|_{\tilde{x}} = \int_{\tilde{x}} (3 - \Delta u)w + \int_{\tilde{x}\cap\partial\Omega} (Du - x)\cdot \hat{n} w = -\int_{\tilde{x}} w^2$$

$\Rightarrow w = 0$ a.e. Let $\tilde{N} = \bigcup_{x \in B_\epsilon(x_0)\cap\partial\Omega} \tilde{x}$, $\Delta u = 3$ on $\tilde{N}$ and Ω_2. □

Recall the ODE on tame part of Ω_1:

$$m'' + m - 3h'\cos\theta - \frac{3}{2}R^2 = h'(Du - x)\cdot\hat{n}$$
$$m'' + m - 3h'\cos\theta = 2R$$

This gives $(m'' + m - 2R)(m'\sin\theta - m\cos\theta + a) = \frac{3}{2}R^2\cos\theta$.

***Theorem:* 3.8:**

Let u optimize for $\Omega = (a, a+1)^2$, $a \geq 0$. Then

1. Ω_0 is a convex set including a neighborhood of (a,a) in $\overline{\Omega}$, *i.e.* $(a,a) \in \mathrm{int}\Omega_0 = \mathrm{int}\{u = 0\}$
2. The set Ω_1^0 of two-ended rays is connected and if $a \geq \frac{7}{2} - \sqrt{2}$, Ω_1^0 is non-empty, but if $a \ll 1$, $\Omega_1^0 = \emptyset$.
3. For $a > 0$, there are exactly two connected components of $\Omega_1 - \Omega_1^0$: one consisting of one-ended rays intersecting west boundary Ω_W. The other intersecting south boundary Ω_S. These one-ended rays are all tame and satisfy the above ODE. For $a = 0$, $\Omega_1 = \emptyset$
4. The north and east boundaries $\Omega_N \cup \Omega_E \subset \Omega_2$ and $(Du - x)\cdot\hat{n} = 0$ for all $x \in \Omega_2 \cap \partial\Omega$ and $\Delta u = 3$ on Ω_2.
5. $x \in \Omega_1 \cap \partial\Omega$ is tame if and only if $(Du(x) - x)\cdot\hat{n} > 0$, stray otherwise, happens only at $\partial\Omega_1 \cap \partial\Omega_2 \cap \partial\Omega$.

Proof. 1. Recall that $\Omega_0 = \{u = 0\}$ if $\Omega_0 \neq \emptyset$. The convexity of $\{u = 0\}$ and $\sigma(\{u = 0\}) = \frac{\delta E}{\delta u}(\{u = 0\}) = 1$, then $\mathrm{int}\{u = 0\} \neq \emptyset$, hence $\{u = 0\} = \Omega_0$. $\frac{\partial u}{\partial x_i} \geq 0$ for $i = 1, 2$, by symmtry, $(a,a) \in \Omega_0$.

4. Claim: $(Du - x)\cdot\hat{n} = 0$ on $\Omega_N \cup \Omega_E$. For a contradiction, suppose a tame ray intersects Ω_E with negative slope, so $\theta \geq 0$ measured clockwise from $(-1, 0)$.

$\sigma^- \le \sigma^+ \Rightarrow$ sign of Jacobian $dH^1(x^2) = h'(\theta)d\theta$, so nearby rays are spreading as we move away from $\partial\Omega$.
2-Monotonicity of $Du \Rightarrow$ if $x_2 \in \tilde{x}_0$ and $x_3 = x_2 + \lambda e_1 \in \partial\Omega$ for $\lambda > 0$. Then

$$\begin{aligned}(Du(x_3) - Du(x_2)) \cdot e_1 &> 0\\ (Du(x_3) - x_3)e_1 &> 0\\ (Du(x_2) - x_2)e_1 &> 0\\ \Rightarrow \frac{\partial}{\partial\theta}(Du(\overline{x}(0,\theta)) - \overline{x}(0,\theta))e_1 &> 0\end{aligned}$$

Rays continue above x_0 all the way to $(a+1, a+1)$.
One of the rays $\tilde{x}_4$ above x_0 is two-sided, and cuts off an isoceles triangle T with right-angle at $\Omega_N \cap \Omega_E$.
Taylor expansion gives

$$\overline{u}(x) = \begin{cases} u(x), x \notin T\\ u(x_4) + Du(x_4)(x - x_4), x \in T\end{cases}$$

Claim: $E(\overline{u}) < E(u) = \int_\Omega \frac{1}{2}|Du - x|^2 + \left(u - \frac{1}{2}x^2\right)$.
$\overline{u} \le u$ on Ω. For $x = (x_1, x_2) \in T$,

$$\begin{aligned} x_1 \le a+1 < \frac{\partial u}{\partial x_1}(x_4) = \frac{\partial \overline{u}}{\partial x_1}(x_4) \le \frac{\partial u}{\partial x_1}(x)\\ \frac{\partial u}{\partial x_1}(x_4) - (a+1) > 0\\ x_2 \le a+1 < \frac{\partial u}{\partial x_2}(x_5) = \frac{\partial \overline{u}}{\partial x_1}(x_5) < \frac{\partial u}{\partial x_2}(x).\end{aligned}$$

Then $|D\overline{u} - x|^2 \le |Du - x|^2, \forall x \in T$. Contradiction.

Suppose there are no two-ended rays and instead have single-ended tame rays all the way to the corner. In this case, we have uniform control on Δu for any point $x \in \mathrm{dom} D^2 u$ on these rays of

$$\infty > C > \Delta u - 3 = \frac{2R - 3r}{h' \cos\theta + r} \overset{r=0}{\to} \frac{2R}{h'\cos\theta} = \frac{4(Du - x)\cdot \hat{n}}{R} \to \infty \text{ as } R \to 0$$

Contradiction.

3. Similar arguments show no tame ray intersecting either Ω_W or Ω_S can have nonnegative slope. But we can have rays of negative slope.
Similarly, arguements show such rays must limit to either Ω_0 or Ω_1^0.
Strong maximum principle $\Rightarrow$ lower limit of Ω_1^0, if $\Omega_1^0 \neq \emptyset$, it must lie on $\partial\Omega_0$.
Lower limit of any connected component of Ω_1^- either has to be Ω_0, Ω_1^0 or have zero length.

2. We want to show that $a \ge \frac{7}{2} - \sqrt{2} \Rightarrow \Omega_1^0 \neq \emptyset$.

Assume $\Omega_1^0 = \emptyset$. Let $[(a, \underline{x_2}), (a, \overline{x_2})]$ be maximal in $\overline{\Omega}_1 \cap \partial\Omega$.

$$\begin{aligned}
a - 0 &= \partial_1 u(a, \overline{x_2}) - \partial_1 u(a, \underline{x_2}) \\
&= \int_{\underline{x_2}}^{\overline{x_2}} \frac{\partial^2 u}{\partial x_1 \partial x_2}(a, x_2) dx_2 \\
&= -\int_{\underline{\theta}}^{\overline{\theta}} \frac{m'' + m}{h' \cos\theta} \sin\theta \cos\theta h'(\theta) d\theta \\
&= -\int_{\underline{\theta}}^{\overline{\theta}} (2R + 3h' \cos\theta) \sin\theta d\theta \text{ because } 3 - \Delta u = \frac{-2R}{h' \cos\theta} \text{ when } r \to 0 \\
&< 2\left\|R\right\|_\infty \left[\cos 0 - \cos\left(-\frac{\pi}{4}\right)\right] + \frac{3}{2}\int_{\underline{x_2}}^{\overline{x_2}} dx_2 \\
&= 2\left(1 - \frac{1}{\sqrt{2}}\right) + \frac{3}{2} = \frac{7}{2} - \sqrt{2}
\end{aligned}$$

Then we want to show that $a \ll 1 \Rightarrow \Omega_1^0 = \emptyset$.
Let $u^{(0)} \in \arg\min \int_{(a,a+1)^2} \frac{1}{2}|Du - x|^2 + u - \frac{1}{2}x^2$, $u^{(0)} \neq 0$.
$\Omega_1^{(0)} = \emptyset$, $\Omega_1^{(0)} \supset \lim_{a \to 0} \Omega_1^{(a)}$ with area $\frac{1}{3}$. ☐

3.1 Regularity

Let $u \geq 0$ be convex, Ω be a compact convex subset of $[0, \infty)^n$, consider

$$\inf E(u) = \inf \int (c(Du) - xDu + u) f(x) dH^n(x),$$

where $c(y) = \frac{1}{2}|y|^2$ or $D^2 c \geq \epsilon I > 0$.

***Theorem:* 3.9: Caffarelli-Lions (2006+)**

u optimal $\Rightarrow u \in C^{1,1}_{loc}(\Omega)$ with norms $\|u\|_{C^{1,1}(X')}$ depending only on X' a convex compact subset of Ω, $\|\log f\|_{C^{0,1}(X')}$, $d(X', \partial\Omega)$, and $\epsilon > 0$.

Idea: Estimate energies of locally affine replacement.

Proof. Assuming Lemma 3.10, comparing energy of u to $\max\{u, A\}$ yields

$$\begin{aligned}
0 &\leq E(\max\{u, A\}) - E(u) \\
&= \int_S [c(\overline{y}) - x \cdot \overline{y} + u]_{\overline{y} = Du(x_0)}^{\overline{y} = y} f(x) dH^n(x) \\
&\leq \left(c_1 h - c_2 \left(\frac{h}{r}\right)^2 + h\right) |S| \text{ by 2,3 of lemma}
\end{aligned}$$

Therefore, $\left(\frac{h}{r}\right)^2 \leq \frac{c_1+1}{c_2} h$, or $h \leq \frac{c_1+1}{c_2} r^2$. ☐

Lemma: 3.10:

Assume $u \in C^1(\overline{\Omega})$, $\exists c_1, c_2$ constant and $r_0 < d(X', \partial\Omega)$ depending only on the same data $(\epsilon, \Omega, \|\log f\|_{C^{0,1}(X')}, \|u\|_{C^{1,1}(X')})$, $\forall (x_0, y_0) \in \partial\Omega$ with $x \in X'$ and $0 < r < r_0$, $\exists A(x) = x \cdot y + \beta$ s.t.

1. $x_0 \in S = \{x_0 \in \Omega : u(x) < A(x)\}$
2. $0 \leq (A-u)(x) \leq h = \sup_{x \in B_r(x_0)} u(x) - u(x_0) - Du(x_0)(x - x_0)$ on S
3. $\frac{1}{|S|}\int_S (c(y) - x \cdot y - c(Du) + xDu) f(x) dH^n(x) \leq c_1 h - c_2 \left(\frac{h}{r}\right)^2$.

Proof. Choose A s.t. 1, 2 holds. Taylor expansion gives

$$c(\overline{y}) \geq c(y) + Dc(y)(\overline{y} - y) + \frac{\epsilon}{2}|y - \overline{y}|^2$$

$$[c(\overline{y}) - x \cdot y]_{\overline{y}=Du(x)}^{\overline{y}=DA(x)=y} \leq -(Dc(DA) - x)(Du - DA) - \frac{\epsilon}{2}|Du - DA|^2$$

Integrate over S to get

$$\begin{aligned}
\int_S [c(\overline{y}) - x \cdot y]_{\overline{y}=Du(x)}^{\overline{y}=DA(x)=y} f dH^n \leq &- \int_{\partial S} (Dc(DA) - x) \cdot \hat{n}_S (u - A) f dH^{n-1} \\
&+ \int_S \nabla \cdot f (Dc(DA) - x)(u - A) dH^n \\
&- \frac{\epsilon}{2}\int_S |Du - DA|^2 f dH^n \\
&\leq c_1'' h|S| + c_1 h|S| - \frac{\epsilon}{2}\int |Du - DA|^2 f dH^n
\end{aligned}$$

The last inequality comes from the following: f is constant. Convexity of $S, \Omega \Rightarrow \nabla \cdot x = n$.

$$\begin{aligned}
n|S| &= \int_S \nabla \cdot (x - x_0) dH^n = \int_{\partial S} (x - x_0) \cdot \hat{n}_S dH^{n-1} \\
&\geq \int_{\partial S \cap \partial\Omega} (x - x_0) \cdot \hat{n}_S dH^{n-1} \\
&\geq \int_{\partial S \cap \partial\Omega} r_0 dH^{n-1}
\end{aligned}$$

Similarly,

$$\int_{\partial S} (Dc(DA) - x) \cdot \hat{n}_S (u - A) f dH^{n-1} \leq f_2 h \int_{\partial S \cap \partial\Omega} (x - Dc(DA)) \cdot \hat{n}$$

To choose A, WLOG, assume $x_0 = 0 = Du(x_0)$, $h = u(re_1)$, $Du(re_1) = \lambda e_1$ for some λ.
Let $A(x) = \frac{hx_1}{2r} + \frac{h}{2}$, $A(re_1) = h$, $A(-re_1) = 0$, $DA(re_1) = \frac{h}{2r}e_1$. $(u - A)|_{(r, x_2, \ldots, x_n)} \geq 0$, with equality at re_1. $S = \{x \in \Omega : u < A\} \subset \{x \in \mathbb{R}^n : -r \leq x_1 \leq r\}$.

Now we want to show that $\int_S |Du - DA|^2 dH^n \geq c_2''|S|\left(\frac{h}{r}\right)^2$.
Choose $0 < k < \frac{\operatorname{diam}\Omega}{r_0}$, then $kS \subset B_{r_0}(x_0 = 0) \subset \Omega$.

Let $\tilde{x} = P_1^{\perp}(x_1, x_2, ..., x_n) = (x_2, ..., x_n)$ be the projection, $\Delta L_{\tilde{x}} = L_{\tilde{x}}^+ - L_{\tilde{x}}^-$.

$$\begin{aligned}
\int_S |Du - DA|^2 dH^n &= \int_{P_1^{\perp}(kS)} \Delta L_{\tilde{x}} \left(\int_{L_{\tilde{x}}^-}^{L_{\tilde{x}}^+} \left(\frac{\partial u}{\partial x_1} - \frac{\partial A}{\partial x_1} \right)^2 \frac{\partial x_1}{\Delta L_{\tilde{x}}} \right) dH^{n-1}(\tilde{x}) \\
&\geq \int_{P_1^{\perp}(kS)} \Delta L_{\tilde{x}} (u - A)|_{L_{\tilde{x}}^-}^{L_{\tilde{x}}^+} \frac{1}{\Delta L_{\tilde{x}}} dH^{n-1} \\
&\geq \int_{P_1^{\perp}(kS)} \frac{1}{2r} \left(\frac{h(1-k)}{2} \right)^2 dH^{n-1} \\
&\geq \frac{k^{n-1}}{2r} \left(\frac{h(1-k)}{2} \right)^2 H^{n-1}(P_1^{\perp}(S)) \\
&\geq \frac{k^{n-1}}{(2r)^2} \left(\frac{h(1-k)}{2} \right)^2 |S| \quad \text{(Because } H^n(S) \leq 2rH^{n-1}(P_1^{\perp}(S))) \\
&= c_2'''|S| \left(\frac{h}{r} \right)^2 \geq c_1 h|S|
\end{aligned}$$

□