Transforming i.d. random variables into each other with conditional expectations.

Georgy Gaitsgori

Based on joint work with Stanislav A. Molchanov and Isaac M. Sonin

Stochastic Analysis and its Application in Economics RS at HSE

October 2, 2020

Story begins in 2007, when A.Cherny and P.Grigoriev published a noteworthy result for Risk Theory.

Theorem 1

Let (Ω, F, P) be a nonatomic probability space, X ,Y two bounded functions with the same distribution. Then for any $\epsilon > 0$ there is a sequence of σ-subalgebras $F_1, F_2, \dots, F_n \subseteq F$ such that for a sequence of random variables $X_0 = X, X_1 = E(X_0|F_1)$, $X_2 = E(X_1|F_2), \dots, X_n = E(X_{n-1}|F_n)$, the following inequality holds $||X_n - Y||_\infty < \epsilon$.

Story begins in 2007, when A.Cherny and P.Grigoriev published a noteworthy result for Risk Theory.

Theorem 1

Let (Ω, F, P) be a nonatomic probability space, X ,Y two bounded functions with the same distribution. Then for any $\epsilon > 0$ there is a sequence of σ-subalgebras $F_1, F_2, \ldots, F_n \subseteq F$ such that for a sequence of random variables $X_0 = X, X_1 = E(X_0|F_1)$, $X_2 = E(X_1|F_2), \ldots, X_n = E(X_{n-1}|F_n)$, the following inequality holds $||X_n - Y||_\infty < \epsilon$.

This result is highly connected to **Risk Theory**.

- One of the important objects studied in Financial Mathematics and Risk Theory is **risk measures**.
- For example, risk measure is used to determine the amount of an asset (currency) to be kept in reserve.
- A risk measure is defined as a mapping $\rho : \mathcal{L} \to \mathbb{R}$ from a set of random variables to the real numbers:
  1. **Normalized** $\rho(0) = 0$;
  2. **Translative** If $a \in \mathbb{R}, X \in \mathcal{L}$, then $\rho(X + a) = \rho(X) - a$;
  3. **Monotone** If $X, Y \in \mathcal{L}$ and $X \leq Y$, then $\rho(Y) \leq \rho(X)$.

- One of the important objects studied in Financial Mathematics and Risk Theory is **risk measures**.
- For example, risk measure is used to determine the amount of an asset (currency) to be kept in reserve.
- A risk measure is defined as a mapping $\rho : \mathcal{L} \to \mathbb{R}$ from a set of random variables to the real numbers:
  1. **Normalized** $\rho(0) = 0$;
  2. **Translative** If $a \in \mathbb{R}, X \in \mathcal{L}$, then $\rho(X + a) = \rho(X) - a$;
  3. **Monotone** If $X, Y \in \mathcal{L}$ and $X \leq Y$, then $\rho(Y) \leq \rho(X)$.
- For example: $\rho(X) = \mathbb{E}[-X]$.
- Counterexample: $\rho(X) = Var[X]$

- One of the important objects studied in Financial Mathematics and Risk Theory is **risk measures**.
- For example, risk measure is used to determine the amount of an asset (currency) to be kept in reserve.
- A risk measure is defined as a mapping $\rho : \mathcal{L} \to \mathbb{R}$ from a set of random variables to the real numbers:
  1. **Normalized** $\rho(0) = 0$;
  2. **Translative** If $a \in \mathbb{R}, X \in \mathcal{L}$, then $\rho(X + a) = \rho(X) - a$;
  3. **Monotone** If $X, Y \in \mathcal{L}$ and $X \leq Y$, then $\rho(Y) \leq \rho(X)$.
- For example: $\rho(X) = \mathbb{E}[-X]$.
- Counterexample: $\rho(X) = Var[X]$
- Important family: **Law invariant risk measures**.

- **Coherent risk measures**: $\rho : L^\infty \to \mathbb{R}$ with
  1. **Normalized + Monotonicity + Translation invariance**
  2. **Subadditivity** $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
  3. **Positive homogenity** If $\lambda \geq 0$ then $\rho(\lambda X) = \lambda \rho(X)$.

- **Coherent risk measures**: $\rho : L^\infty \to \mathbb{R}$ with
  1. **Normalized + Monotonicity + Translation invariance**
  2. **Subadditivity** $\rho(X + Y) \leq \rho(X) + \rho(Y)$.
  3. **Positive homogenity** If $\lambda \geq 0$ then $\rho(\lambda X) = \lambda\rho(X)$.
- Föllmer and Schied showed that it is too restrictive.
- **Convex risk measures**: $\rho : L^\infty \to \mathbb{R}$ with
  1. **Normalized + Monotonicity + Translation invariance**
  2. **Convexity** $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$, where $\lambda \in [0,1]$.

- **Coherent risk measures**: $\rho : L^\infty \to \mathbb{R}$ with
  1. **Normalized + Monotonicity + Translation invariance**
  2. **Subadditivity** $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
  3. **Positive homogenity** If $\lambda \geq 0$ then $\rho(\lambda X) = \lambda\rho(X)$.
- Föllmer and Schied showed that it is too restrictive.
- **Convex risk measures**: $\rho : L^\infty \to \mathbb{R}$ with
  1. **Normalized + Monotonicity + Translation invariance**
  2. **Convexity** $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$, where $\lambda \in [0,1]$.

Theorem (Föllmer, Schied in '04)

Any law invariant convex risk measure on an atomless probability space has to satisfy

$$\rho(\mathbb{E}(X|G)) \leq \rho(X)$$

for any $X \in L^\infty$ and $\mathcal{G} \subset \mathcal{F}$.

- Property $\rho(\mathbb{E}(X|G)) \leq \rho(X)$ for any $X \in \mathcal{L}^\infty$ and $\mathcal{G} \subset \mathcal{F}$ was introduced by Leitner in '04 and is called **dilatation monotoninicty**.
- Are dilatation monotone risk measures are law invariant?

- Property $\rho(\mathbb{E}(X|G)) \leq \rho(X)$ for any $X \in \mathcal{L}^\infty$ and $\mathcal{G} \subset \mathcal{F}$ was introduced by Leitner in '04 and is called **dilatation monotoninicty**.
- Are dilatation monotone risk measures are law invariant?

Theorem (Cherny, Grigoriev in '07)

On an atomless probability space any L^∞-continuous dilatation monotone map $R : L^\infty \to \mathbb{R}$ is law invariant.

- Property $\rho(\mathbb{E}(X|G)) \leq \rho(X)$ for any $X \in \mathcal{L}^\infty$ and $\mathcal{G} \subset \mathcal{F}$ was introduced by Leitner in '04 and is called **dilatation monotoninicty**.
- Are dilatation monotone risk measures are law invariant?

Theorem (Cherny, Grigoriev in '07)

On an atomless probability space any L^∞-continuous dilatation monotone map $R : L^\infty \to \mathbb{R}$ is law invariant.

Theorem

On an atomless probability space a convex risk measure is law invariant iff it is dilatation monotone.

Theorem 1

Let (Ω, F, P) be a nonatomic probability space, X ,Y two bounded functions with the same distribution. Then for any $\epsilon > 0$ there is a sequence of σ-subalgebras $F_1, F_2, \ldots, F_n \subseteq F$ such that for a sequence of random variables $X_0 = X, X_1 = E(X_0|F_1)$, $X_2 = E(X_1|F_2), \ldots, X_n = E(X_{n-1}|F_n)$, the following inequality holds $||X_n - Y||_\infty < \epsilon$.

Theorem 1
Let (Ω, F, P) be a nonatomic probability space, X ,Y two bounded functions with the same distribution. Then for any $\epsilon > 0$ there is a sequence of σ-subalgebras $F_1, F_2, \ldots, F_n \subseteq F$ such that for a sequence of random variables $X_0 = X, X_1 = E(X_0|F_1)$, $X_2 = E(X_1|F_2), \ldots, X_n = E(X_{n-1}|F_n)$, the following inequality holds $||X_n - Y||_\infty < \epsilon$.

Note that $\mathbb{F} := (F_i)$ is a not a filtration, which is unusual for financial mathematics.

Theorem 1
Let (Ω, F, P) be a nonatomic probability space, X ,Y two bounded functions with the same distribution. Then for any $\epsilon > 0$ there is a sequence of σ-subalgebras $F_1, F_2, \ldots, F_n \subseteq F$ such that for a sequence of random variables $X_0 = X, X_1 = E(X_0|F_1)$, $X_2 = E(X_1|F_2), \ldots, X_n = E(X_{n-1}|F_n)$, the following inequality holds $||X_n - Y||_\infty < \epsilon$.

Note that $\mathbb{F} := (F_i)$ is a not a filtration, which is unusual for financial mathematics.

In other words, we can «fully» transform r.v. X into Y with the sequence of «averaging» operations.

Points to improve:

- Highly formal concept of conditional expectations;
- No explicit sequence or way how to build it;
- No optimality or asymptotic behavior.

Points to improve:

- Highly formal concept of conditional expectations;
- No explicit sequence or way how to build it;
- No optimality or asymptotic behavior.

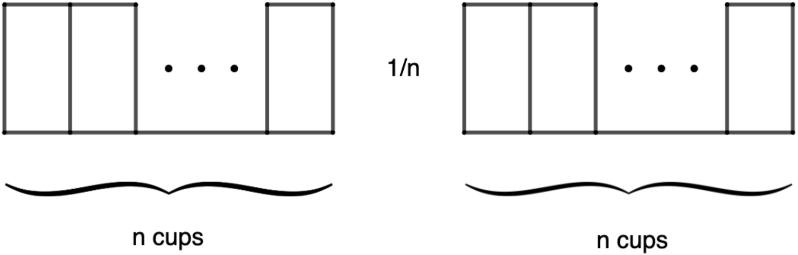
Main results:

- Transparent interpretation using «hydrostatic» setup;
- Explicit and optimal algorithm of transformations;
- Exact first term of the asymptotic behavior.

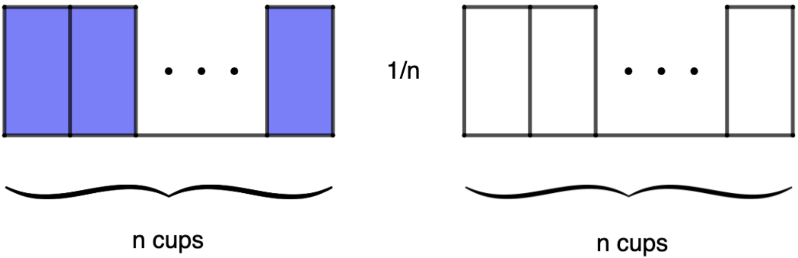
The structure:

- Transparent interpretation;
- Optimal algorithm for a simple case;
- Asymptotic behaviour of the optimal algorithm (Basic Lemma);
- Implication of the Basic Lemma to the Theorem 1.
- Optimality of the algorithm.

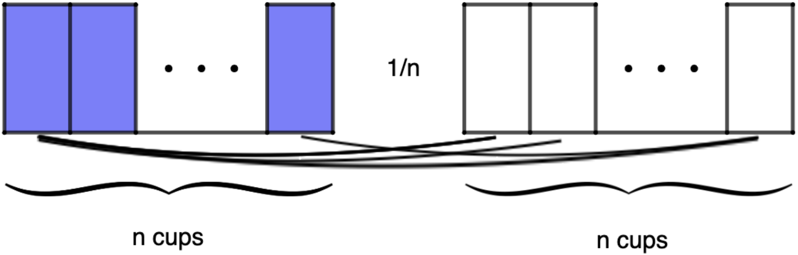
Transparent interpretation



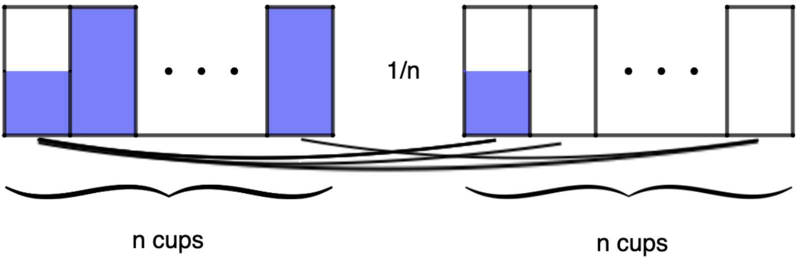
Transparent interpretation



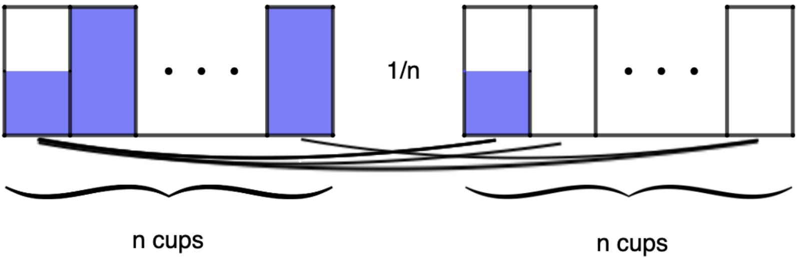
Transparent interpretation



Transparent interpretation

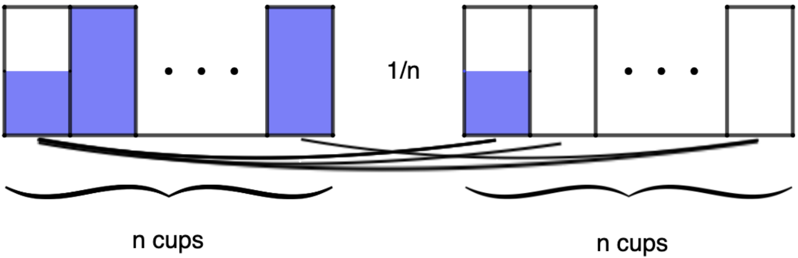


Transparent interpretation



Problem

How much water can be transferred from the left to the right cups with a series of such «mixing operations»?



Problem

How much water can be transferred from the left to the right cups with a series of such «mixing operations»?

Basic Lemma

In the setup with $2n$ cups the first term of the asymptotic behaviour of water amount left under the optimal transfer is $\frac{2}{\sqrt{n\pi}}$.

- Let the cups be numbered as $1, 2, \ldots, n$ from the center to the right, and as $-1, -2, \ldots, -n$ from the center to the left;
- On the first stage, we start with connecting the full cup -1 sequentially with all cups $1, 2, 3, \ldots, n$;

Optimal algorithm

- Let the cups be numbered as $1, 2, \dots, n$ from the center to the right, and as $-1, -2, \dots, -n$ from the center to the left;
- On the first stage, we start with connecting the full cup -1 sequentially with all cups $1, 2, 3, \dots, n$;
- As a result, these n cups receive levels $1/2, 1/4, 1/8, \dots, 1/2^n$, the cup -1 gets the same level as the cup n, i.e. $1/2^n$;

- Let the cups be numbered as $1, 2, \ldots, n$ from the center to the right, and as $-1, -2, \ldots, -n$ from the center to the left;
- On the first stage, we start with connecting the full cup -1 sequentially with all cups $1, 2, 3, \ldots, n$;
- As a result, these n cups receive levels $1/2, 1/4, 1/8, \ldots, 1/2^n$, the cup -1 gets the same level as the cup n, i.e. $1/2^n$;
- On our second stage, we connect the full cup -2 sequentially again with all cups $1, 2, 3, \ldots, n$, or again with all cups $1, 2, \ldots;$
- As a result, these n cups involved receive levels $(1 + 1/2)/2 = 3/4, (3/4 + 1/4)/2 = 1/2, (1/2 + 1/8) = 5/16, \ldots$

- Let $x_i(j)$ be the relative level in the cup $j, j = 1, 2, \dots, n$ after i stages of our procedure, $i = 1, 2, \dots, n$;
- Denote by d_n (the «deficit») the total amount of water untransferred to the right after n stages of finite transfer;
- The deficit equals

$$d_n = \frac{1}{n} \sum_{i=1}^{n} x_i(n). \tag{1}$$

- Let $x_i(j)$ be the relative level in the cup $j, j = 1, 2, \ldots, n$ after i stages of our procedure, $i = 1, 2, \ldots, n$;
- Denote by d_n (the «deficit») the total amount of water untransferred to the right after n stages of finite transfer;
- The deficit equals

$$d_n = \frac{1}{n}\sum_{i=1}^{n} x_i(n). \tag{1}$$

Objective

We want to prove that

$$d_n = \frac{1}{n}\sum_{i=1}^{n} x_i(n) \sim \frac{2}{\sqrt{n\pi}}.$$

Lemma 1

Denote S_k as a random variable with Negative Binomial distribution with parameters $k, p = 1/2$ (a number of failures before the k-th success occurs). Then

$$x_k(j) = P(S_1 = j-1) + \cdots + P(S_k = j-1), j = 1, 2, \ldots \qquad (2)$$

Lemma 1

Denote S_k as a random variable with Negative Binomial distribution with parameters $k, p = 1/2$ (a number of failures before the k-th success occurs). Then

$$x_k(j) = P(S_1 = j-1) + \cdots + P(S_k = j-1), j = 1, 2, \ldots \quad (2)$$

- Introduce generating function $F_n(z) = \sum_{j=1}^{\infty} x_n(j) z^j, n = 1, 2, \ldots$
- Use recursive relation $x_i(j) = \frac{x_{i-1}(j)}{2} + \frac{x_{i-1}(j-1)}{2^2} + \cdots + \frac{x_{i-1}(1)}{2^j} + \frac{1}{2^j}, j = 1, 2, \ldots$

Lemma 1

Denote S_k as a random variable with Negative Binomial distribution with parameters $k, p = 1/2$ (a number of failures before the k-th success occurs). Then

$$x_k(j) = P(S_1 = j-1) + \cdots + P(S_k = j-1), j = 1, 2, \ldots \quad (2)$$

- Introduce generating function $F_n(z) = \sum_{j=1}^{\infty} x_n(j) z^j, n = 1, 2, \ldots$
- Use recursive relation $x_i(j) = \frac{x_{i-1}(j)}{2} + \frac{x_{i-1}(j-1)}{2^2} + \cdots + \frac{x_{i-1}(1)}{2^j} + \frac{1}{2^j}, j = 1, 2, \ldots$
- Get $F_n(z) = \sum_{k=1}^{n} \frac{z}{(2-z)^k}, n = 1, 2, \ldots$

Lemma 1

Denote S_k as a random variable with Negative Binomial distribution with parameters $k, p = 1/2$ (a number of failures before the k-th success occurs). Then

$$x_k(j) = P(S_1 = j-1) + \cdots + P(S_k = j-1), j = 1, 2, \ldots \qquad (2)$$

- Introduce generating function $F_n(z) = \sum_{j=1}^{\infty} x_n(j) z^j, n = 1, 2, \ldots$
- Use recursive relation $x_i(j) = \frac{x_{i-1}(j)}{2} + \frac{x_{i-1}(j-1)}{2^2} + \cdots + \frac{x_{i-1}(1)}{2^j} + \frac{1}{2^j}, j = 1, 2, \ldots$
- Get $F_n(z) = \sum_{k=1}^{n} \frac{z}{(2-z)^k}, n = 1, 2, \ldots$

Proposition 2

Using (1) and (2) we obtain $nd_n = \sum_{k=1}^{n} P\left(S_k \geq n\right)$.

Using notation $m_n = n^{3/5}$ we can rewrite (3) as

$$\begin{aligned} nd_n = \sum_{k=1}^{n-\lfloor m_n \rfloor} P\left(S_k \geq n\right) + \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(S_k \geq n + m_n\right) + \\ + \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(n \leq S_k < n + m_n\right) = A_1 + A_2 + A_3 \end{aligned} \quad (3)$$

Using notation $m_n = n^{3/5}$ we can rewrite (3) as

$$nd_n = \sum_{k=1}^{n-\lfloor m_n \rfloor} P\left(S_k \geq n\right) + \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(S_k \geq n + m_n\right) + \\ + \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(n \leq S_k < n + m_n\right) = A_1 + A_2 + A_3 \tag{3}$$

Proposition 3

If $k \leq n - m_n$, then $P\left(S_k \geq n\right) \leq \exp\left(-\frac{n^{\frac{1}{5}}}{4} + O\left(n^{-\frac{1}{5}}\right)\right)$.

Proposition 4

For any $k \leq n$ $P\left(S_k \geq n + m_n\right) \leq \exp\left(-\frac{n^{\frac{1}{5}}}{4} + O\left(n^{\frac{1}{6}}\right)\right)$.

- If $1 \leq k \leq n - m_n$ and $1 < z < 2$, Markov inequality implies

$$P\left(S_k \geq n\right) = P\left(z^{S_k} \geq z^n\right) \leq \frac{Ez^{S_k}}{z^n} = \frac{1}{(2-z)^k z^n}.$$

- If $1 \leq k \leq n - m_n$ and $1 < z < 2$, Markov inequality implies
$$P(S_k \geq n) = P\left(z^{S_k} \geq z^n\right) \leq \frac{Ez^{S_k}}{z^n} = \frac{1}{(2-z)^k z^n}.$$
- Put $z = 1 + \varepsilon$. Then for large n and small ε we have
$$P(S_k \geq n) = P\left((1+\varepsilon)^{S_k} \geq (1+\varepsilon)^n\right) \leq (1-\varepsilon)^{-k}(1+\varepsilon)^{-n} \leq$$
$$\leq (1-\varepsilon)^{-n+m_n}(1+\varepsilon)^{-n} =$$
$$= \exp\left[-m_n\varepsilon + n\varepsilon^2 + O\left(m_n\varepsilon^2\right) + O\left(n\varepsilon^3\right)\right].$$

Proof of Propositions 3

- If $1 \leq k \leq n - m_n$ and $1 < z < 2$, Markov inequality implies

$$P\left(S_k \geq n\right) = P\left(z^{S_k} \geq z^n\right) \leq \frac{Ez^{S_k}}{z^n} = \frac{1}{(2-z)^k z^n}.$$

- Put $z = 1 + \varepsilon$. Then for large n and small ε we have

$$P\left(S_k \geq n\right) = P\left((1+\varepsilon)^{S_k} \geq (1+\varepsilon)^n\right) \leq (1-\varepsilon)^{-k}(1+\varepsilon)^{-n} \leq$$
$$\leq (1-\varepsilon)^{-n+m_n}(1+\varepsilon)^{-n} =$$
$$= \exp\left[-m_n\varepsilon + n\varepsilon^2 + O\left(m_n\varepsilon^2\right) + O\left(n\varepsilon^3\right)\right].$$

- Put $\varepsilon = \varepsilon_n = \frac{1}{2n^{2/5}}$, as it approximately minimizes of the exponent. It gives for $k \leq n - m_n$

$$P\left(S_k \geq n\right) \leq \exp\left(-\frac{1}{4}n^{\frac{1}{5}} + O\left(n^{-\frac{1}{5}}\right)\right),$$

Therefore, $A_1(n) \to 0, A_2(n) \to 0$ and $nd_n = A_3(n) + o(1)$.

Therefore, $A_1(n) \to 0, A_2(n) \to 0$ and $nd_n = A_3(n) + o(1)$.

- Recall $A_3(n) = \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(n \leq S_k < n + m_n\right)$;
- Note that $P\left(S_{k+1} = m\right) = b(k|k+m)$, where $b(k|n)$ is a Binomial distribution with parameters $n, 1/2$;

Therefore, $A_1(n) \to 0, A_2(n) \to 0$ and $nd_n = A_3(n) + o(1)$.

- Recall $A_3(n) = \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(n \le S_k < n + m_n\right)$;
- Note that $P\left(S_{k+1} = m\right) = b(k|k+m)$, where $b(k|n)$ is a Binomial distribution with parameters $n, 1/2$;
- By the De Moivre-Laplace Theorem $b(k|n,p) = \frac{1}{\sqrt{2\pi npq}} \exp\left[-\frac{(k-np)^2}{2npq} + o(1)\right]$;
- Therefore, we obtain

$$A_3(n) = \sum_{k=n-\lfloor m_n \rfloor}^{n-1} \sum_{m=n}^{n+\lfloor m_n \rfloor} \sqrt{\frac{2}{\pi(k+m)}} \exp\left(-\frac{(m-k)^2}{2(k+m)} + o(1)\right);$$

Therefore, $A_1(n) \to 0, A_2(n) \to 0$ and $nd_n = A_3(n) + o(1)$.

- Recall $A_3(n) = \sum_{k=n-\lfloor m_n \rfloor+1}^{n} P\left(n \leq S_k < n + m_n\right)$;
- Note that $P\left(S_{k+1} = m\right) = b(k|k+m)$, where $b(k|n)$ is a Binomial distribution with parameters $n, 1/2$;
- By the De Moivre-Laplace Theorem $b(k|n,p) = \frac{1}{\sqrt{2\pi npq}} \exp\left[-\frac{(k-np)^2}{2npq} + o(1)\right]$;
- Therefore, we obtain

$$A_3(n) = \sum_{k=n-\lfloor m_n \rfloor}^{n-1} \sum_{m=n}^{n+\lfloor m_n \rfloor} \sqrt{\frac{2}{\pi(k+m)}} \exp\left(-\frac{(m-k)^2}{2(k+m)} + o(1)\right);$$

- Or $A_3(n) = \frac{1}{\sqrt{\pi n}} \sum_{l=1}^{\lfloor n^{3/5} \rfloor} (l+1) \exp\left(-\frac{l^2}{4n}\right) + O\left(n^{\frac{7}{10}} \exp\left(-\frac{n^{\frac{1}{5}}}{4}\right)\right)$

Euler-Maclaurin formula

Suppose that L is a positive integer, $f(x)$ is a non-negative continuous function with the absolute maximum at the point $x_0 \in [0, L]$, and monotonously increasing on $[0, x_0]$, and monotonously decreasing on $[x_0, L]$. Then for $S_L = f(0) + f(1) + \cdots + f(L)$

$$\left| S_L - \int_0^L f(x)dx \right| \leq 3f\left(x_0\right) = 3 \max_{x \in [0,L]} f(x).$$

Euler-Maclaurin formula

Suppose that L is a positive integer, $f(x)$ is a non-negative continuous function with the absolute maximum at the point $x_0 \in [0, L]$, and monotonously increasing on $[0, x_0]$, and monotonously decreasing on $[x_0, L]$. Then for $S_L = f(0) + f(1) + \cdots + f(L)$

$$\left| S_L - \int_0^L f(x)dx \right| \leq 3f(x_0) = 3 \max_{x \in [0,L]} f(x).$$

- Apply to function $(x+1)\exp\left(-\frac{x^2}{4n}\right)$ with maximum of an order $\sqrt{n}$.
- Get

$$\frac{1}{\sqrt{\pi n}} \sum_{l=1}^{\lfloor n^{3/5} \rfloor} (l+1)\exp\left(-\frac{l^2}{4n}\right) =$$

$$= \frac{1}{\sqrt{\pi n}} \int_1^{\lfloor n^{3/5} \rfloor} (x+1)\exp\left(-\frac{x^2}{4n}\right) dx + O(1) \sim \frac{2}{\sqrt{n\pi}}.$$

- We devided nd_n into sum $nd_n = A_1(n) + A_2(n) + A_3(n)$;
- We showed that $A_1(n) \to 0, A_2(n) \to 0$;
- We showed that $A_3(n) \sim 2\sqrt{\frac{n}{\pi}} + o(1)$;
- Therefore, we proved Basic Lemma.

Basic Lemma

In the setup with $2n$ cups the first term of the asymptotic behaviour of water amount left under the optimal transfer is $d_n = \frac{2}{\sqrt{n\pi}}$.

- We devided nd_n into sum $nd_n = A_1(n) + A_2(n) + A_3(n)$;
- We showed that $A_1(n) \to 0, A_2(n) \to 0$;
- We showed that $A_3(n) \sim 2\sqrt{\frac{n}{\pi}} + o(1)$;
- Therefore, we proved Basic Lemma.

Basic Lemma

In the setup with $2n$ cups the first term of the asymptotic behaviour of water amount left under the optimal transfer is $d_n = \frac{2}{\sqrt{n\pi}}$.

Remark

Basic Lemma proves Theorem 1 for the case when $X = c_1 1_A + c_2 1_B, Y = c_2 1_A + c_1 1_B$ for some disjoint subsets $A, B \subseteq \Omega$ such that $\mu(A) = \mu(B)$.

We should show the derivation of the Cherny-Grigoriev Theorem 1 from this basic case.

- It is sufficient to consider only simple bounded function X, Y (functions with finitely many values).
- Indeed, for an arbitrary $\delta > 0$ and equidistributed bounded functions X, Y there exist equidistributed bounded **simple** functions $\tilde{X}, \tilde{Y}$ such that $||X - \tilde{X}||_\infty < \delta, ||Y - \tilde{Y}||_\infty < \delta$.

- It is sufficient to consider only simple bounded function X, Y (functions with finitely many values).
- Indeed, for an arbitrary $\delta > 0$ and equidistributed bounded functions X, Y there exist equidistributed bounded **simple** functions $\tilde{X}, \tilde{Y}$ such that $||X - \tilde{X}||_\infty < \delta, ||Y - \tilde{Y}||_\infty < \delta$.
- Say that X takes values $x_1 < x_2 < \cdots < x_N$ on subsets $A_1, \dots, A_N \subset \Omega$, Y takes the same values $x_1 < x_2 < \cdots < x_N$ on subsets $B_1, \dots, B_N \subset \Omega$, and $\forall i = 1, \dots, N : \mu(A_i) = \mu(B_i)$.
- We now will *transform* X into Y step by step.

- It is sufficient to consider only simple bounded function X, Y (functions with finitely many values).
- Indeed, for an arbitrary $\delta > 0$ and equidistributed bounded functions X, Y there exist equidistributed bounded **simple** functions $\tilde{X}, \tilde{Y}$ such that $||X - \tilde{X}||_\infty < \delta, ||Y - \tilde{Y}||_\infty < \delta$.
- Say that X takes values $x_1 < x_2 < \cdots < x_N$ on subsets $A_1, \ldots, A_N \subset \Omega$, Y takes the same values $x_1 < x_2 < \cdots < x_N$ on subsets $B_1, \ldots, B_N \subset \Omega$, and $\forall i = 1, \ldots, N : \mu(A_i) = \mu(B_i)$.
- We now will *transform* X into Y step by step.
- At the first step we will make X equal (up to fixed ϵ) to x_N on B_N.
- Then without touching previous results we will make our new X_1 equal (up to fixed ϵ) to x_{N-1} on B_{N-1} and so on.
- As the result, we will obtain X_N such that on every B_N it is equal to Y up to fixed ϵ, i.e. $||X_N - Y||_\infty < \epsilon$.

- Suppose that $A_N \cap B_N = \varnothing$, if not we will not touch their intersection.
- Suppose X takes values $x_1 < x_2 < \cdots < x_N$ on subsets $C_1, \ldots, C_N \subset B_N$ respectively, $\cup_{i=1}^{N} C_i = B_N$.
- Divide A_N into disjoint sets $D_1, \ldots, D_N \subset A_N$ such that $\cup_{i=1}^{N} D_i = A_N$ and $\forall i = 1, \ldots, N : \mu(D_i) = \mu(C_i)$.
- Basic Lemma implies that we can *swap* the values of X on all these subsets C_i and D_i up to a fixed ϵ: first, swap C_1 and D_1, then C_2 and D_2, etc.

- After that, we will obtain new function X_1 such that $\forall i = 1, \dots, N : ||X_1|_{C_i} - Y|_{C_i}||_\infty < \epsilon$ or $||X_1|_{B_N} - Y|_{B_N}||_\infty < \epsilon$.
- Moreover, on $\Omega \setminus B_N$ up to an arbitrary small ϵ X_1 and Y again equidistributed, but take only $N - 1$ values.
- Therefore, we can similarly obtain next X_2 such that $||X_2|_{B_{N-1}} - Y|_{B_{N-1}}||_\infty < \epsilon$ or $||X_2|_{B_N \cup B_{N-1}} - Y|_{B_N \cup B_{N-1}}||_\infty < \epsilon$.
- After N such steps we will obtain $||X_N - Y||_\infty < \epsilon$.

- After that, we will obtain new function X_1 such that $\forall i = 1, \dots, N : ||X_1|_{C_i} - Y|_{C_i}||_\infty < \epsilon$ or $||X_1|_{B_N} - Y|_{B_N}||_\infty < \epsilon$.
- Moreover, on $\Omega \setminus B_N$ up to an arbitrary small ϵ X_1 and Y again equidistributed, but take only $N-1$ values.
- Therefore, we can similarly obtain next X_2 such that $||X_2|_{B_{N-1}} - Y|_{B_{N-1}}||_\infty < \epsilon$ or $||X_2|_{B_N \cup B_{N-1}} - Y|_{B_N \cup B_{N-1}}||_\infty < \epsilon$.
- After N such steps we will obtain $||X_N - Y||_\infty < \epsilon$.

- What is about **optimality**?

- Consider the problem of transferring water from n left cups to n right cups.

- Consider the problem of transferring water from n left cups to n right cups.
- **Triple discretization problem:**

- Consider the problem of transferring water from n left cups to n right cups.
- **Triple discretization problem:**
  1. A space (for example $(0, 1)$) is divided into $2n$ equal parts;
  2. Possible values of functions are in a fixed grid;
  3. Discretizing discrete time even more, allowing only one unit of a transfer at a time.

- Consider the problem of transferring water from n left cups to n right cups.
- **Triple discretization problem:**
  1. A space (for example $(0,1)$) is divided into $2n$ equal parts;
  2. Possible values of functions are in a fixed grid;
  3. Discretizing discrete time even more, allowing only one unit of a transfer at a time.
- Suppose that levels of the cups are integers $s = 0, 1, \ldots$ or fractions s/M, where s and M are integers (for continuous case let n and M tend independently to infinity).

- Suppose that the i-th cup on the left has level $a(i)$ and the j-th cup on can connect if $a(i) > b(j)$).
- After the connection levels become $a(i)' = a(i) - 1$ and $b(j)' = b(j) + 1$.

- Suppose that the i-th cup on the left has level $a(i)$ and the j-th cup on can connect if $a(i) > b(j)$).
- After the connection levels become $a(i)' = a(i) - 1$ and $b(j)' = b(j) + 1$.
- Say that a cup i on the left is **open** at moment m, if there is at least one cup on the right with the lower the level.
- Denote $A^0_m(i)$, the «*inner set*» of a cup i at the moment m, i.e. all cups available to it, which are not available to any cup with a lower level.

- Suppose that the i-th cup on the left has level $a(i)$ and the j-th cup on can connect if $a(i) > b(j)$).
- After the connection levels become $a(i)' = a(i) - 1$ and $b(j)' = b(j) + 1$.
- Say that a cup i on the left is **open** at moment m, if there is at least one cup on the right with the lower the level.
- Denote $A^0_m(i)$, the «*inner set*» of a cup i at the moment m, i.e. all cups available to it, which are not available to any cup with a lower level.

Optimal strategy

Let π_* be the following strategy: *At each moment a cup with lowest level from the left among all open cups is connected with an available for it cup with the greatest level on the right*. And if at any moment there are some draws, connections can be ordered in some arbitrary way.

- At each moment the state of a system is described by a $2n$ dimensional vector
  $c = (a, b), a = (a(i), i = 1, 2, \ldots, n), b = (b(j), j = 1, 2, \ldots, n)$.
- Define $k(c|\pi)$ as the total number of connections obtained from state c using a strategy π and let $\pi(c)$ be the next state obtained from c, when π is applied.
- Define $k_0(c) = \sup_\pi k(c|\pi)$, i.e. maximum possible number of connections from state c. Let $k_*(c)$ be the number of connections from state c using the strategy π_*.

- At each moment the state of a system is described by a $2n$ dimensional vector $c = (a, b), a = (a(i), i = 1, 2, \ldots, n), b = (b(j), j = 1, 2, \ldots, n)$.
- Define $k(c|\pi)$ as the total number of connections obtained from state c using a strategy π and let $\pi(c)$ be the next state obtained from c, when π is applied.
- Define $k_0(c) = \sup_\pi k(c|\pi)$, i.e. maximum possible number of connections from state c. Let $k_*(c)$ be the number of connections from state c using the strategy π_*.

Objective

We want to show that for any given state c we have $k_*(c) = k_0(c)$.

- We shall prove Theorem 2 by induction on the maximal number of remaining steps n.
- The induction statement $P_n, n = 0, 1, \ldots$ is: For any state c, if $k_0(c) = n$, then $k_*(c) = n$.

- We shall prove Theorem 2 by induction on the maximal number of remaining steps n.
- The induction statement $P_n, n = 0, 1, \ldots$ is: For any state c, if $k_0(c) = n$, then $k_*(c) = n$.
- Base is trivial as if for a particular state $c : n = 0$, then it means that there is no open cups and P_0 is trivially true.

- We shall prove Theorem 2 by induction on the maximal number of remaining steps n.
- The induction statement $P_n, n = 0, 1, \ldots$ is: For any state c, if $k_0(c) = n$, then $k_*(c) = n$.
- Base is trivial as if for a particular state $c : n = 0$, then it means that there is no open cups and P_0 is trivially true.
- Now suppose that P_k holds for all $1 \leq k \leq n$, and $k_0(c) = n + 1$. Let an optimal strategy π_0 asks (s, t) connection and the strategy π_* asks (i, j) connection.
- Let us denote states $c_0 = \pi_0(c)$ and $c_* = \pi_*(c)$. According to the induction assumption the strategy π_0 after the first step can be continued by the strategy π_*, i.e. $k_0(c_0) = k_*(c_0) = n$.

- We shall prove Theorem 2 by induction on the maximal number of remaining steps n.
- The induction statement $P_n, n = 0, 1, \dots$ is: For any state c, if $k_0(c) = n$, then $k_*(c) = n$.
- Base is trivial as if for a particular state $c : n = 0$, then it means that there is no open cups and P_0 is trivially true.
- Now suppose that P_k holds for all $1 \leq k \leq n$, and $k_0(c) = n + 1$. Let an optimal strategy π_0 asks (s, t) connection and the strategy π_* asks (i, j) connection.
- Let us denote states $c_0 = \pi_0(c)$ and $c_* = \pi_*(c)$. According to the induction assumption the strategy π_0 after the first step can be continued by the strategy π_*, i.e. $k_0(c_0) = k_*(c_0) = n$.
- And we want to show that $k_*(c_*) = k_0(c_0) = n$ either, or that there exists a strategy π' such that $k(c_*|\pi') = n$.

- If an optimal strategy π_0 is different from strategy π_*, then the definition of π_* implies that $a(s) \geq a(i)$ and if $b(t) < a(i)$, then $b(j) \leq b(j)$.
- Only the following four situations are possible:
  1. $a(s) > a(i)$ and $b(t) \geq a(i)$;
  2. $a(s) > a(i)$ and $b(t) < b(j)$;
  3. $a(s) > a(i)$ and $b(t) = b(j)$;
  4. $a(s) = a(i)$ and $b(t) < b(j)$.

- If an optimal strategy π_0 is different from strategy π_*, then the definition of π_* implies that $a(s) \geq a(i)$ and if $b(t) < a(i)$, then $b(j) \leq b(j)$.
- Only the following four situations are possible:
  1. $a(s) > a(i)$ and $b(t) \geq a(i)$;
  2. $a(s) > a(i)$ and $b(t) < b(j)$;
  3. $a(s) > a(i)$ and $b(t) = b(j)$;
  4. $a(s) = a(i)$ and $b(t) < b(j)$.
- In all cases we shall show that the state c_* is «at least good» as state c_0, i.e. there is a matching strategy π' applied to c_* with at least n steps.

- In case 1) the strategy π_* applied to c_0 asks for (i,j) connection, because in this state, as in state c, the left cup i is still the lowest open and the right cup j is still its greatest available.

- In case 1) the strategy π_* applied to c_0 asks for (i,j) connection, because in this state, as in state c, the left cup i is still the lowest open and the right cup j is still its greatest available.
- Hence $c' = \pi_*(c_0) = (a(s) - 1, a(i) - 1, \dots; b(t) + 1, b(j) + 1, \dots)$ and $k_*(c') = k_0(c') = n - 1$.

- In case 1) the strategy π_* applied to c_0 asks for (i,j) connection, because in this state, as in state c, the left cup i is still the lowest open and the right cup j is still its greatest available.
- Hence $c' = \pi_*(c_0) = (a(s) - 1, a(i) - 1, \dots; b(t) + 1, b(j) + 1, \dots)$ and $k_*(c') = k_0(c') = n - 1$.
- But state c' can be obtained also from state c_* using strategy π' with the first connection (s, t).

- In case 1) the strategy π_* applied to c_0 asks for (i,j) connection, because in this state, as in state c, the left cup i is still the lowest open and the right cup j is still its greatest available.
- Hence $c' = \pi_*(c_0) = (a(s) - 1, a(i) - 1, \dots; b(t) + 1, b(j) + 1, \dots)$ and $k_*(c') = k_0(c') = n - 1$.
- But state c' can be obtained also from state c_* using strategy π' with the first connection (s, t).
- If this connection is continued by π_0 then $k_0(c_*) \geq k(c_*|\pi') = 1 + k_0(c') = 1 + (n - 1) = n$.
- Therefore, we obtain P_{n+1}.

Main results:

- We showed transparent interpretation of Theorem 1 using «hydrostatic» setup;
- We provided explicit and optimal algorithm of transformations for the simplest case;
- We found an exact first term of the asymptotic behavior (Basic Lemma);
- We proved that our simplest case implies Theorem 1.

- V.I. Arnold (2006), Mathematical Understanding of Nature: Essays on Amazing Physical Phenomena and Their Understanding by Mathematicians. *MCCME*.
- A.S. Cherny, P.G. Grigoriev (2007), Dilatation monotone risk measures are law invariant. *Finance and Stochastic, Vol. 2*, 291-298.
- W. Feller (1967), An Introduction to Probability Theory and Its Applications. Vol. 1. *Wiley*.
- H. Föllmer, A. Schied (2002), Convex measures of risk and trading constraints. *Finance and Stochastics, 6, No. 4*, 429-447.
- J. Leitner (2004), Belayage monotonous risk measures. *International Journal of Theoretical and Applied Finance, 7, No. 7*, 887-900.