Regularity and optimization

Bruno Gaujal

INRIA

SDA2, 2007

Majorization

$x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, $x^{\downarrow}$ is a permutation of x with decreasing coordinates.

$$x \preccurlyeq y \text{ si } \begin{cases} \sum_{i=1}^{k} x_i^{\downarrow} \leqslant \sum_{i=1}^{k} y_i^{\downarrow} & \forall k, 1 \leqslant k \leqslant n-1, \\ \sum_{i=1}^{n} x_i^{\downarrow} = \sum_{i=1}^{n} y_i^{\downarrow}. \end{cases}$$

Marshall Olkin, 1979

Majorization

$x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, $x^{\downarrow}$ is a permutation of x with decreasing coordinates.

$$x \preccurlyeq y \text{ si } \begin{cases} \sum_{i=1}^{k} x_i^{\downarrow} \leqslant \sum_{i=1}^{k} y_i^{\downarrow} & \forall k, 1 \leqslant k \leqslant n-1, \\ \sum_{i=1}^{n} x_i^{\downarrow} = \sum_{i=1}^{n} y_i^{\downarrow}. \end{cases}$$

Marshall Olkin, 1979

If $x = (0, -1, -3, 1)$ and $y = (3, 1, -8, 1)$,
$x^{\downarrow} = (1, 0, -1, -3), y^{\downarrow} = (3, 1, 1, -8)$, partial sums verify :

$$\begin{aligned} 1 &\leqslant 3 \\ 1+0 &\leqslant 3+1 \\ 1+0-1 &\leqslant 3+1+1 \\ 1+0-1-3 &= 3+1+1-8, \end{aligned}$$

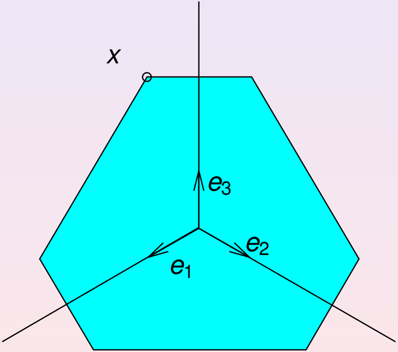
which implies $x \preccurlyeq y$.

Majorization (II)

If P is a doubly stochastic matrix, $y^{\downarrow} = x^{\downarrow}P \Leftrightarrow y \preccurlyeq x$.
Using Birkhoff Theorem (doubly stochastic matrices are convex combination of permutation matrices).

Majorization (II)

If P is a doubly stochastic matrix, $y^{\downarrow} = x^{\downarrow} P \Leftrightarrow y \preccurlyeq x$.
Using Birkhoff Theorem (doubly stochastic matrices are convex combination of permutation matrices).



Schur convexity

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ is *Schur-convex* if $x \preccurlyeq y \Rightarrow f(x) \leqslant f(y)$.

Schur convexity

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ is *Schur-convex* if $x \preccurlyeq y \Rightarrow f(x) \leqslant f(y)$.

Theorem

If $f(x) \leqslant f(y)$ for all Schur convex functions f, then $x \preccurlyeq y$.

Schur convexity

$f : \mathbb{R}^n \to \mathbb{R}$ is *Schur-convex* if $x \preccurlyeq y \Rightarrow f(x) \leqslant f(y)$.

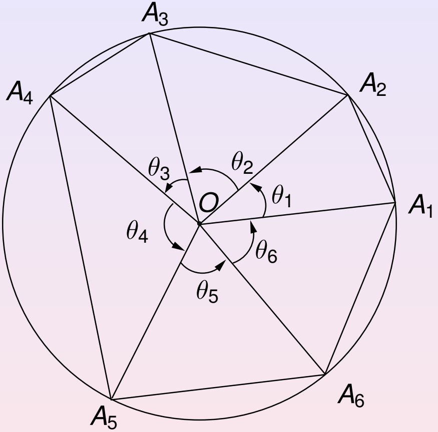
Theorem

If $f(x) \leqslant f(y)$ for all Schur convex functions f, then $x \preccurlyeq y$.

Let $\phi : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ such that $\psi : \mathbb{R}^n \to \mathbb{R}$ is such that $\psi(x_1, \ldots, x_n) = \phi(g(x_1), \ldots, g(x_n))$. If ϕ is increasing and Schur convex and g is convex, then ψ est Schur-decreasing.

Some applications

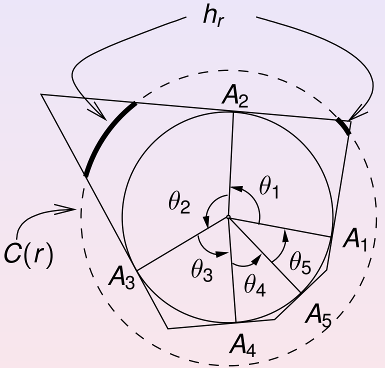
Polygons with n vertices inside a disk.



The surface is $1/2 \sum_{i=1}^{n} \sin(\theta_i)$, which is Schur-concave.
$(\theta_1, \ldots \theta_n) \preccurlyeq (\alpha_1, \ldots \alpha_n)$ implies that the surfaces $S_\theta \geqslant S_\alpha$. The polygon with the largest surface S is regular.

Some applications

Polygon with n vertices outside a disk, with the smallest k-th moment, for all k.



$h_r = r\sum_{i=1}^{n} \max(0, \theta_i - 2\cos^{-1}(1/r))$. is Schur convex.
The polygon with the smallest k-th moment is regular, for all k.

Application in Networks

Bandwidth allocation in Networks Eitan Altman
Using Shannon SINR Formula, The total throughput is a sum of a convex function of the emited power, so that it is maximized by small (w.r.t.majorization)power allocation.

Application in Networks

Bandwidth allocation in Networks Eitan Altman
Using Shannon SINR Formula, The total throughput is a sum of a convex function of the emited power, so that it is maximized by small (w.r.t.majorization)power allocation.

Structure of optimal control Ger Koole

Routing applications

Majorization and Schur convexity cannot be used when the problem depends on the actual sequence and is not invariant up to permutations.
Actually, the following applications will be invariant up to cyclic permutations.

Mechanical words

This talk will now focus on *finite or periodic mechanical words*.

1. Extremal properties of rational mechanical words for scheduling problems
2. Factorizations of mechanical words using Continued Fractions and their consequences on computational problems
3. Compute the extremal points in the scheduling problems.

Mechanical words

This talk will now focus on *finite or periodic mechanical words*.

1. Extremal properties of rational mechanical words for scheduling problems
2. Factorizations of mechanical words using Continued Fractions and their consequences on computational problems
3. Compute the extremal points in the scheduling problems.

There exist many ways to link continued fraction decompositions with mechanical words. Here, we will exhibit a link which has some algorithmic and computational interest.

Mechanical words

Definition (Mechanical words)

A mechanical word with slope (density) $0 \leqslant \alpha \leqslant 1$ and intercept ρ is the infinite word $w_{\alpha,\rho}$ whose letters are

$$\forall n \geqslant 0,\ w_{\alpha,\rho}(n) = \lceil (n+1)\alpha + \rho \rceil - \lceil n\alpha + \rho \rceil, or$$

$$\forall n \geqslant 0,\ w_{\alpha,\rho}(n) = \lfloor (n+1)\alpha + \rho \rfloor - \lfloor n\alpha + \rho \rfloor$$

Mechanical words

Definition (Mechanical words)

A mechanical word with slope (density) $0 \leqslant \alpha \leqslant 1$ and intercept ρ is the infinite word $w_{\alpha,\rho}$ whose letters are

$$\forall n \geqslant 0,\ w_{\alpha,\rho}(n) = \lceil (n+1)\alpha + \rho \rceil - \lceil n\alpha + \rho \rceil, or$$

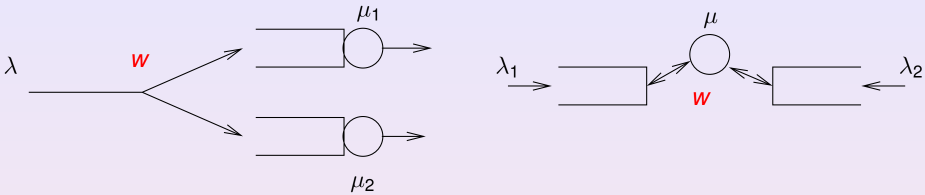
$$\forall n \geqslant 0,\ w_{\alpha,\rho}(n) = \lfloor (n+1)\alpha + \rho \rfloor - \lfloor n\alpha + \rho \rfloor$$

When α is a rational number ($\alpha = p/q$), then $w_{\alpha,\rho}$ is periodic of period q (*i.e.* $\forall n \geqslant 0$, $w_{\alpha,\rho}(n) = w_{\alpha,\rho}(n+q)$)

Definition (Finite mechanical words)

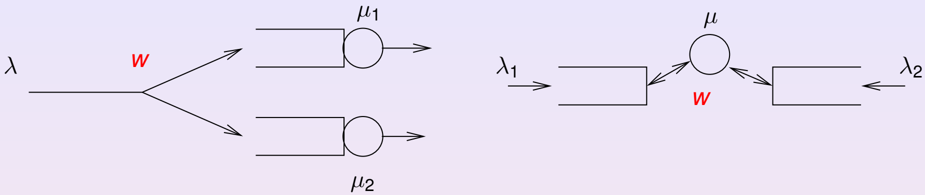
A finite word m is mechanical if $w = m^{\infty}$ is a mechanical word .

Optimal properties of rational mechanical words



The arrival rate(s) λ (λ_1 and λ_2) are arbitrary positive real numbers, the service rates μ_1, μ_2 (μ) are also arbitrary. They verify the stability property : $\mu_1 + \mu_2 \geqslant \lambda$ ($\mu > \lambda_1 + \lambda_2$).

Optimal properties of rational mechanical words



The arrival rate(s) λ (λ_1 and λ_2) are arbitrary positive real numbers, the service rates μ_1, μ_2 (μ) are also arbitrary. They verify the stability property : $\mu_1 + \mu_2 \geqslant \lambda$ ($\mu > \lambda_1 + \lambda_2$).

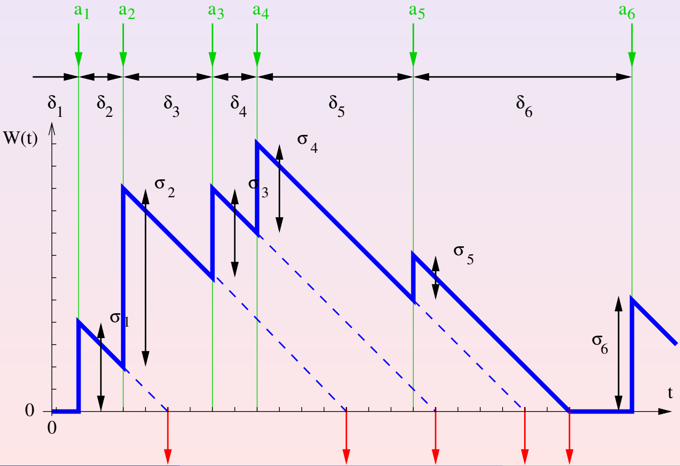
Theorem

G.-Hyon 03, G.-Hordijk-Van der Laan 06 The routing (allocation) sequence minimizing the workload is a mechanical word with a rational density α.

Workload

If σ_n is the size of packet n, δ_n the time between the nth arrival and the $n-1$th arrival, and $a_n \in \{0,1\}$, the routing decision to Queue 1, then the workload verifies :

$$W_n = \max\left(W_{n-1} + a_{n-1}\sigma_{n-1} - \delta_n, 0\right) .$$



Optimality (II)

Under routing sequence w, let $W_n^1(w)$ be the workload in queue 1 at the n-th decision and

$$W^1(w) = \limsup \frac{1}{N} \sum_{n=1}^{N} W_n^1(w).$$

Theorem (AGH)

Altman, G., Hordijk, 2000 Over all sequences w of density at least α, $W^1(w)$ is minimal when $w = m_\alpha$, the mechanical sequence of density α.

Based on multimodularity properties.

Convexity of $\alpha \mapsto \alpha W^1(m_\alpha)$

Let $\alpha_1 = \frac{p_1}{q_1}$ and $\alpha_2 = \frac{p_2}{q_2}$ in the stability interval, $J = [1 - \mu_2, \mu_1]$.
Let $m = m_{\alpha_1}^{q_2} m_{\alpha_2}^{q_1}$.
Let $N(m)$ be the average number of packets in queue 1 under the routing sequence m.

Convexity of $\alpha \mapsto \alpha W^1(m_\alpha)$

Let $\alpha_1 = \frac{p_1}{q_1}$ and $\alpha_2 = \frac{p_2}{q_2}$ in the stability interval,$J = [1 - \mu_2, \mu_1]$.
Let $m = m_{\alpha_1}^{q_2} m_{\alpha_2}^{q_1}$.
Let $N(m)$ be the average number of packets in queue 1 under the routing sequence m.

$N(m) = \frac{q_1 q_2 N(m_1) + q_2 q_1 N(m_2)}{q_1 q_2 + q_2 q_1}$.

Convexity of $\alpha \mapsto \alpha W^1(m_\alpha)$

Let $\alpha_1 = \frac{p_1}{q_1}$ and $\alpha_2 = \frac{p_2}{q_2}$ in the stability interval,$J = [1 - \mu_2, \mu_1]$.
Let $m = m_{\alpha_1}^{q_2} m_{\alpha_2}^{q_1}$.
Let $N(m)$ be the average number of packets in queue 1 under the routing sequence m.

$$N(m) = \frac{q_1 q_2 N(m_1) + q_2 q_1 N(m_2)}{q_1 q_2 + q_2 q_1}.$$

Let $\alpha = \frac{p_1 q_2 + p_2 q_1}{q_1 q_2 + q_2 q_1}$ (note that $\alpha = \alpha_1/2 + \alpha_2/2$ is the density of m).

Convexity of $\alpha \mapsto \alpha W^1(m_\alpha)$

Let $\alpha_1 = \frac{p_1}{q_1}$ and $\alpha_2 = \frac{p_2}{q_2}$ in the stability interval,$J = [1 - \mu_2, \mu_1]$.
Let $m = m_{\alpha_1}^{q_2} m_{\alpha_2}^{q_1}$.
Let $N(m)$ be the average number of packets in queue 1 under the routing sequence m.

$$N(m) = \frac{q_1 q_2 N(m_1) + q_2 q_1 N(m_2)}{q_1 q_2 + q_2 q_1}.$$

Let $\alpha = \frac{p_1 q_2 + p_2 q_1}{q_1 q_2 + q_2 q_1}$ (note that $\alpha = \alpha_1/2 + \alpha_2/2$ is the density of m).

Using Theorem [AGH], $N(m_\alpha) \leqslant N(m) = 1/2 N(m_{\alpha_1}) + 1/2 N(m_{\alpha_2})$, and $\forall \lambda \in [0, 1]$

$$N(m_{\lambda\alpha_1 + (1-\lambda)\alpha_2}) \leqslant \lambda N(m_{\alpha_1}) + (1 - \lambda) N(m_{\alpha_2}).$$

Convexity of $\alpha \mapsto \alpha W^1(m_\alpha)$

Let $\alpha_1 = \frac{p_1}{q_1}$ and $\alpha_2 = \frac{p_2}{q_2}$ in the stability interval,$J = [1 - \mu_2, \mu_1]$.
Let $m = m_{\alpha_1}^{q_2} m_{\alpha_2}^{q_1}$.
Let $N(m)$ be the average number of packets in queue 1 under the routing sequence m.

$N(m) = \frac{q_1 q_2 N(m_1) + q_2 q_1 N(m_2)}{q_1 q_2 + q_2 q_1}$.

Let $\alpha = \frac{p_1 q_2 + p_2 q_1}{q_1 q_2 + q_2 q_1}$ (note that $\alpha = \alpha_1/2 + \alpha_2/2$ is the density of m).

Using Theorem [AGH], $N(m_\alpha) \leqslant N(m) = 1/2 N(m_{\alpha_1}) + 1/2 N(m_{\alpha_2})$, and $\forall \lambda \in [0, 1]$

$$N(m_{\lambda\alpha_1 + (1-\lambda)\alpha_2}) \leqslant \lambda N(m_{\alpha_1}) + (1 - \lambda) N(m_{\alpha_2}).$$

Using Little's Formula, $N(m_\alpha) = \alpha(W^1(m_\alpha) + 1/\mu)$.

Linearity over Farey intervals

The interval $[d_1, d_2]$ ($d_i = \frac{p_i}{q_i}$) is called a Farey interval if $q_1 \cdot p_2 - p_1 \cdot q_2 = 1$.

Linearity over Farey intervals

The interval $[d_1, d_2]$ ($d_i = \frac{p_i}{q_i}$) is called a Farey interval if $q_1 \cdot p_2 - p_1 \cdot q_2 = 1$.

Proposition

Let $I = [\alpha_1, \alpha_2]$ be a Farey interval, then for all $\alpha \in I$, m_α can be factorized into m_{α_1} and m_{α_2}.

Linearity over Farey intervals

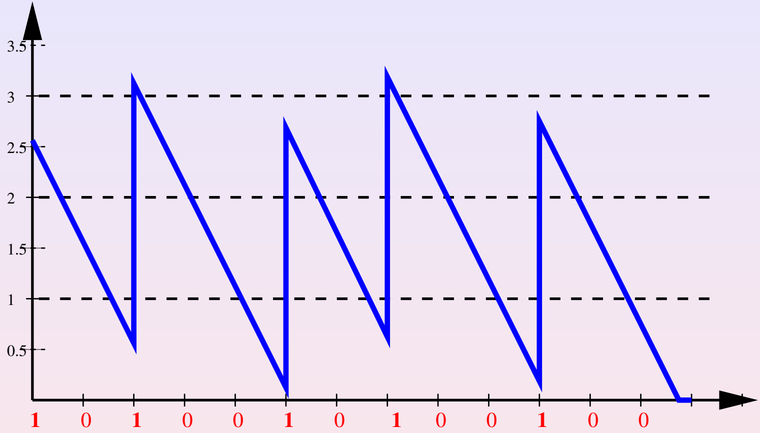
The interval $[d_1, d_2]$ ($d_i = \frac{p_i}{q_i}$) is called a Farey interval if $q_1 \cdot p_2 - p_1 \cdot q_2 = 1$.

Proposition

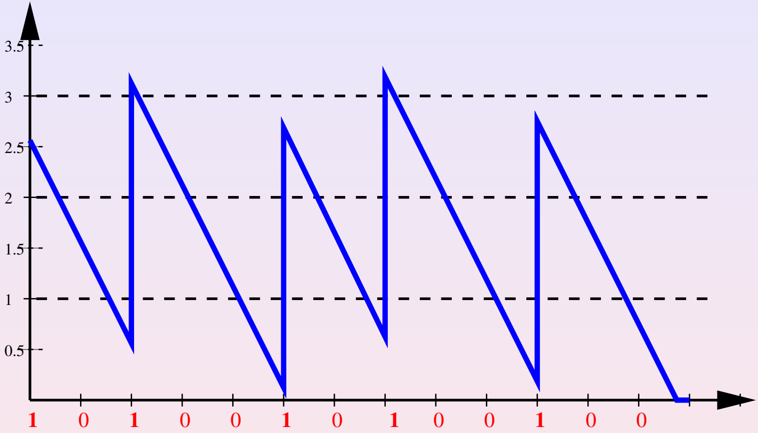
Let $I = [\alpha_1, \alpha_2]$ be a Farey interval, then for all $\alpha \in I$, m_α can be factorized into m_{α_1} and m_{α_2}.

Using the greatest integer continued fraction expansion of α and the associated factorization of m_α (below) one can show that the queue is empty at the end of each period of a mechanical word with intercept 0.

Linearity over Farey interals(II)



Linearity over Farey interals(II)



Corollary

The function $\alpha \mapsto \alpha W^1(m_\alpha)$ is linear over any Farey interal $I \subset J$.

Construction of the best control sequence

The global cost function in both queues is

$$C(\alpha) = \alpha W^1(m_\alpha) + (1-\alpha)W^2(m_{1-\alpha}).$$

Inside the stability interval $J = [1-\mu_2, \mu_1]$, let us consider the successive convergents of the lower continued fraction decomposition of μ_1 and of the upper continued fraction decomposition of $1-\mu_2$. They form Farey intervals over which the function $C(\alpha)$ is linear.

Construction of the best control sequence

The global cost function in both queues is

$$C(\alpha) = \alpha W^1(m_\alpha) + (1-\alpha)W^2(m_{1-\alpha}).$$

Inside the stability interval $J = [1-\mu_2, \mu_1]$, let us consider the successive convergents of the lower continued fraction decomposition of μ_1 and of the upper continued fraction decomposition of $1-\mu_2$. They form Farey intervals over which the function $C(\alpha)$ is linear.

By convexity and piecewise linearity, the function $C(\alpha)$ is minimal at one of these convergents.

Construction of the best control sequence

The global cost function in both queues is

$$C(\alpha) = \alpha W^1(m_\alpha) + (1-\alpha)W^2(m_{1-\alpha}).$$

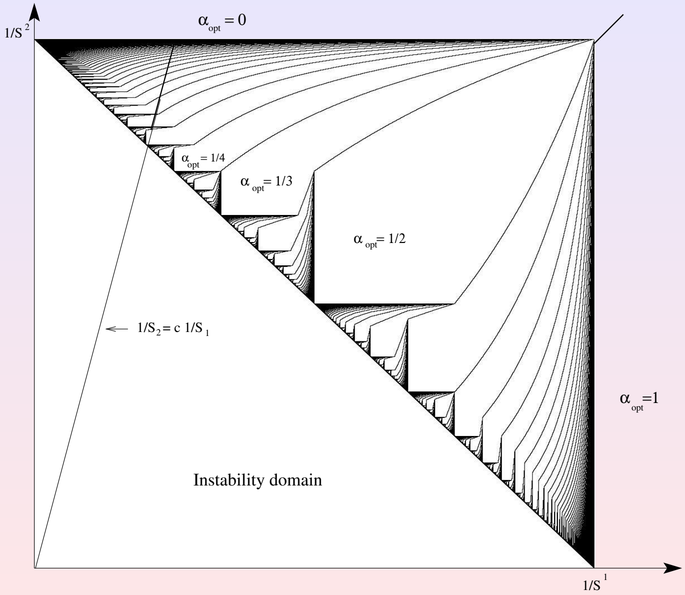
Inside the stability interval $J = [1-\mu_2, \mu_1]$, let us consider the successive convergents of the lower continued fraction decomposition of μ_1 and of the upper continued fraction decomposition of $1-\mu_2$. They form Farey intervals over which the function $C(\alpha)$ is linear.

By convexity and piecewise linearity, the function $C(\alpha)$ is minimal at one of these convergents.

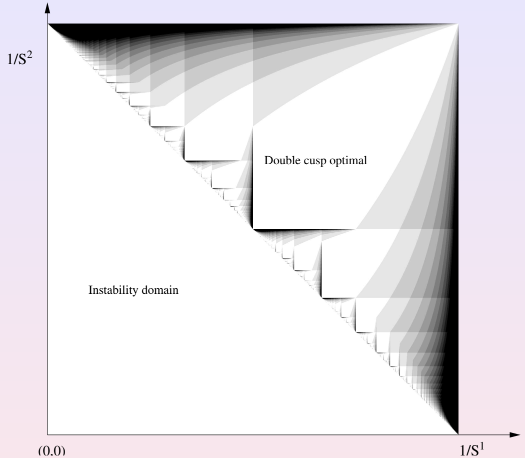
Furthermore, using the recursive factorisation of mechanical words shown below (Theorem 7) one can compute the function C at all those points.

Numerical computations(I)

Numerical computations(II)



Numerical computations(III)



The number of convergents browsed by the algorithm before reaching the minimum point is small (as show in the figure).
More numerical evidences show that for 99% of the stability polytope, the optimal sequence has a period less than 10.

Nearest Integer Continued fraction

A *semi regular continued fraction* (SRCF)-expansion of a number $s \in [0, 1)$ is a finite or infinite fraction

$$s = 0 + \cfrac{\epsilon_1}{b_1 + \cfrac{\epsilon_2}{b_2 + \ddots \cfrac{\epsilon_n}{b_n + \ddots}}},$$

where, $b_n \in \mathbb{N}$ and $\epsilon_n \in \{-1, 1\}$.
($b_n + \varepsilon_n \geqslant 1$, $b_n + \varepsilon_{n+1} \geqslant 1$ and $b_n + \varepsilon_{n+1} \geqslant 2$ infinitely often.)

Nearest Integer Continued fraction

A *semi regular continued fraction* (SRCF)-expansion of a number $s \in [0,1)$ is a finite or infinite fraction

$$s = 0 + \cfrac{\epsilon_1}{b_1 + \cfrac{\epsilon_2}{b_2 + \ddots \cfrac{\epsilon_n}{b_n + \ddots}}},$$

where, $b_n \in \mathbb{N}$ and $\epsilon_n \in \{-1, 1\}$.
($b_n + \varepsilon_n \geqslant 1$, $b_n + \varepsilon_{n+1} \geqslant 1$ and $b_n + \varepsilon_{n+1} \geqslant 2$ infinitely often.)
Let $\mathcal{U}_\gamma : [0,1] \to [0,1]$,
$\mathcal{U}_\gamma(s) = \left|\frac{1}{s}\right| - \lfloor \left|\frac{1}{s}\right| + 1 - \gamma \rfloor$ if $s \neq 0$ and $\mathcal{U}_\gamma(0) = 0$ otherwise.

The coefficients of the SRCF-expansion are

$$\epsilon_n(s) = sign\left(\mathcal{U}_\gamma^{n-1}(s)\right) \;; \quad b_n(s) = \lfloor \frac{1}{\left|\mathcal{U}_\gamma^{n-1}(s)\right|} + 1 - \gamma \rfloor .$$

Nearest Integer Continued fraction (II)

Some SRCFs are quite remarkable.

- $\gamma = 1$ corresponds to the regular continued fraction expansion (RCF).
- $\gamma = 0$ corresponds to an expansion where, for all n, $\epsilon_n = -1$. It appears in the litterature under the name *greatest integer continued fraction* (GICF).
- $\gamma = 1/2$ is the *Nearest Integer Continued Fraction* (NICF), on which we will mainly focus.

Nearest Integer Continued fraction (III)

Let $\alpha = 13/31$.

$$\mathcal{U}_{1/2}(13/31) = |\frac{31}{13}| - \lfloor |\frac{31}{13}| + \frac{1}{2} \rfloor = \frac{5}{13},$$

$$b_1 = \lfloor |\frac{31}{13}| + \frac{1}{2} \rfloor = 2, \epsilon_2 = 1.$$

Continuing the expansion gives $\mathcal{U}^2_{1/2} = -2/5$, $b_2 = 3$, $\epsilon_3 = -1$ and $\mathcal{U}^3_{1/2} = -1/2$, $b_3 = 3$, $\epsilon_4 = -1$ finally $b_4 = 2$.
The NICF expansion of $13/31$ is $[\![2, 3, -3, -2]\!]$.
The sequence of convergents is $[\![2]\!] = 1/2$, $[\![2, 3]\!] = 3/7$, $[\![2, 3, -3]\!] = 8/19$ and $[\![2, 3, -3, -2]\!] = 13/31$.

NICF and word factorization

Let $\alpha \in (0, 1/2)$ such that its partial NICF of order $n+1$ is $[\![0, b_1, \epsilon_2 b_2, \ldots, \epsilon_n b_n, \epsilon_{n+1} b_{n+1} + \epsilon_{n+2}\alpha_{n+1}]\!]$. Let $\mathcal{I}(n) = \sum_{i=2}^{n+1} \mathbf{1}_{\{\epsilon_i = +1\}}$.

Theorem

The mechanical word m_α can be factorized only using two factors x_n and y_n for all n defined by $x_0 = 1$, $y_0 = 0$ and for $n \geqslant 1$,

$$\begin{cases} x_n = x_{n-1}(y_{n-1})^{b_n-1}, y_n = x_{n-1}(y_{n-1})^{b_n} & \text{if } \mathcal{I}(n) \text{ even } \wedge \epsilon_{n+2} = +1, \\ x_n = x_{n-1}(y_{n-1})^{b_n-2}, y_n = x_{n-1}(y_{n-1})^{b_n-1} & \text{if } \mathcal{I}(n) \text{ even } \wedge \epsilon_{n+2} = -1, \\ x_n = (x_{n-1})^{b_n} y_{n-1}, y_n = (x_{n-1})^{b_n-1} y_{n-1} & \text{if } \mathcal{I}(n) \text{ odd } \wedge \epsilon_{n+2} = +1, \\ x_n = (x_{n-1})^{b_n-1} y_{n-1}, y_n = (x_{n-1})^{b_n-2} y_{n-1} & \text{if } \mathcal{I}(n) \text{ odd } \wedge \epsilon_{n+2} = -1. \end{cases}$$

example

Let α be equal to $13/31$, its partial remainders are $5/13$, $-2/5$ and $-1/2$ and its convergents are $1/2$, $3/7$ and $8/19$. It comes

$$m_{13/31} = \underbrace{\overbrace{\underbrace{\overbrace{10}^{x_1}\overbrace{10}^{x_1}\overbrace{100}^{y_1}}_{x_2}\underbrace{\overbrace{10}^{x_1}\overbrace{10}^{x_1}\overbrace{100}^{y_1}}_{x_2}\underbrace{\overbrace{10}^{x_1}\overbrace{100}^{y_1}}_{y_2}}^{x_3}\overbrace{\underbrace{\overbrace{10}^{x_1}\overbrace{10}^{x_1}\overbrace{100}^{y_1}}_{x_2}\underbrace{\overbrace{10}^{x_1}\overbrace{100}^{y_1}}_{y_2}}^{y_3}}_{x_4},$$

since $\overline{\Phi}(m_{13/31}) = m_{1-5/13}$, $\underline{\Phi}(m_{8/13}) = m_{1-2/5}$, $\underline{\Phi}(m_{3/5}) = m_{1/2}$ and $\underline{\Phi}(m_{1/2}) = m_1$.

Algorithmic considerations

Algorithmic considerations

Result

Let $\alpha = p/q \leqslant 1/2$ *be a rational number. The number of iterations to reach the complete expansion in NICF is upper bounded by* $\log_\eta(1 + q\sqrt{8})$ *with* $\eta = 1 + \sqrt{2}$.

Algorithmic considerations

Result

Let $\alpha = p/q \leqslant 1/2$ *be a rational number. The number of iterations to reach the complete expansion in NICF is upper bounded by* $\log_\eta(1 + q\sqrt{8})$ *with* $\eta = 1 + \sqrt{2}$.

Corollary

The maximal number of regular reductions used to test if a word w *is balanced is bounded by* $\log_\eta(1 + |w|\sqrt{8})$ *with* $\eta = 1 + \sqrt{2}$.

Extension 1 : The stochastic case

The cost is now the Cesaro limit of the expected worloads. The optimal polling between two queues can be computed when the service is exponential (with rate μ_i in queue i) and the inter-arrivals are exponential (with rate λ in both queues).

Extension 1 : The stochastic case

The cost is now the Cesaro limit of the expected worloads. The optimal polling between two queues can be computed when the service is exponential (with rate μ_i in queue i) and the inter-arrivals are exponential (with rate λ in both queues).
The continuous time Markov chain X_t is a quasi birth and death process whose generator Q is given by

$$Q = \begin{bmatrix} C & A_0 & 0 & 0 & \dots \\ A_2 & A_1 & A_0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & \dots \\ 0 & 0 & A_2 & A_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

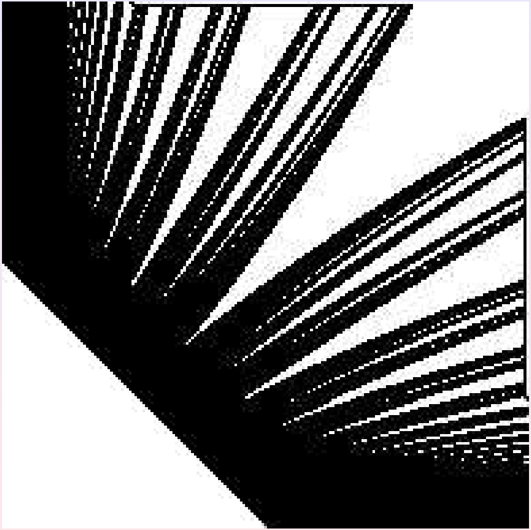
Extension 1 : The stochastic case

The cost is now the Cesaro limit of the expected worloads. The optimal polling between two queues can be computed when the service is exponential (with rate μ_i in queue i) and the inter-arrivals are exponential (with rate λ in both queues).
The continuous time Markov chain X_t is a quasi birth and death process whose generator Q is given by

$$Q = \begin{bmatrix} C & A_0 & 0 & 0 & \dots \\ A_2 & A_1 & A_0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & \dots \\ 0 & 0 & A_2 & A_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

The stationary measure can be computed using the Kernel method : $\Pi(z)K(z) = \pi_0\mu(1-z)M$.

The stochastic case(II)



EXtension 2 : Multidimensional routing problems

A natural extension of the routing problem is to consider a routing to $N > 2$ queues.

EXtension 2 : Multidimensional routing problems

A natural extension of the routing problem is to consider a routing to $N > 2$ queues.
The control sequence can be seen as a word over N letters.

EXtension 2 : Multidimensional routing problems

A natural extension of the routing problem is to consider a routing to $N > 2$ queues.
The control sequence can be seen as a word over N letters.
The optimization problem is to minimize the average waiting time of all incoming packets : $W(a) = \sum_{i=1}^{N} p_i W_i(\mathbf{1}_i(a))$.

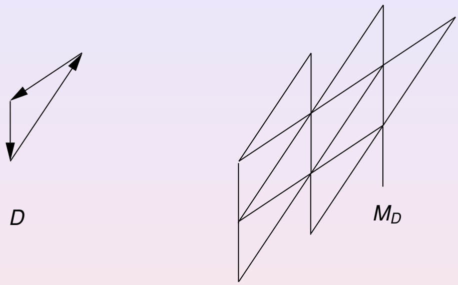
EXtension 2 : Multidimensional routing problems

A natural extension of the routing problem is to consider a routing to $N > 2$ queues.
The control sequence can be seen as a word over N letters.
The optimization problem is to minimize the average waiting time of all incoming packets : $W(a) = \sum_{i=1}^{N} p_i W_i(\mathbf{1}_i(a))$.
The problem is still open.
One of the main difficulties : the sequences $\mathbf{1}_i(a)$ cannot be all balanced (Fraenkel Conjecture Altman, G.,Hordijk, 2000 ; Tijdeman 2004).

The multimodular ordering

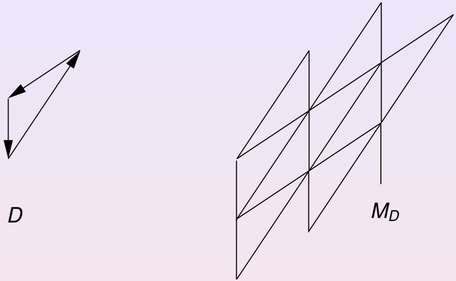
D $(n+1) \times n$ of rank N s. t. the rows $s_0 + \cdots + s_n = 0$.

The Mesh M_D is the set $\{a_0 s_0 + \cdots + a_n s_n, a_i \in \mathbb{Z}\}$.



The multimodular ordering

D $(n+1) \times n$ of rank N s. t. the rows $s_0 + \cdots + s_n = 0$.
The Mesh M_D is the set $\{a_0 s_0 + \cdots + a_n s_n, a_i \in \mathbb{Z}\}$.



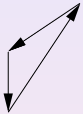
Definition (AGH, 2003)

A function f is D-multimodular if $\forall a \in \mathbb{Z}^n$ and $\forall i \neq j$,

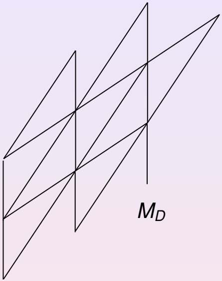
$$f(a+s_i) + f(a+s_j) \geqslant f(a) + f(a+s_i+s_k)$$

The multimodular ordering

D $(n+1) \times n$ of rank N s. t. the rows $s_0 + \cdots + s_n = 0$.
The Mesh M_D is the set $\{a_0 s_0 + \cdots + a_n s_n, a_i \in \mathbb{Z}\}$.



D



M_D

Definition (AGH, 2003)

A function f is D-multimodular if $\forall a \in \mathbb{Z}^n$ and $\forall i \neq j$,

$$f(a+s_i) + f(a+s_j) \geqslant f(a) + f(a+s_i+s_k)$$

f is D-mm iff its linear interpolation over the atoms of M_D is convex.

The multimodular ordering (II)

Let $\mathcal{F}(D, r)$ be the set of all D-mm functions minimum at $r \in M_D$.

Definition

$x \leqslant_{mm} y$ if $f(x) \leqslant f(y)$ for all $f \in \mathcal{F}(D, r)$.

The multimodular ordering (II)

Let $\mathcal{F}(D, r)$ be the set of all D-mm functions minimum at $r \in M_D$.

Definition

$x \leqslant_{mm} y$ if $f(x) \leqslant f(y)$ for all $f \in \mathcal{F}(D, r)$.

Theorem

$x \leqslant_{mm} y$ iff there exists a path over M_D from x to y, "away" from r.

The multimodular ordering (II)

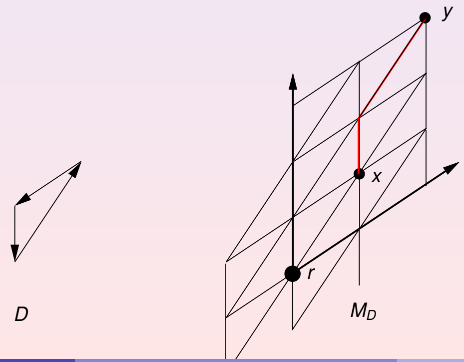
Let $\mathcal{F}(D, r)$ be the set of all D-mm functions minimum at $r \in M_D$.

Definition

$x \leqslant_{mm} y$ if $f(x) \leqslant f(y)$ for all $f \in \mathcal{F}(D, r)$.

Theorem

$x \leqslant_{mm} y$ iff there exists a path over M_D from x to y, "away" from r.



Multimodular order and shifts

Let $P(T, S)$ be the set of all integer sequences of size T that sum to S.
D' is the mesh induced from D over the hyperplane $\sum_i a_i = S$.
$\tilde{P}(T, S)$ is the set of conjugacy classes w.r.t. cyclic permutations.

Multimodular order and shifts

Let $P(T,S)$ be the set of all integer sequences of size T that sum to S.
D' is the mesh induced from D over the hyperplane $\sum_i a_i = S$.
$\tilde{P}(T,S)$ is the set of conjugacy classes w.r.t. cyclic permutations.

Definition

$x \leqslant_{mms} y$ in $\tilde{P}(T,S)$, if $f(x) \leqslant f(y)$ for all D'-mm function f which is invariant by cyclic permutations.

Multimodular order and shifts

Let $P(T, S)$ be the set of all integer sequences of size T that sum to S.
D' is the mesh induced from D over the hyperplane $\sum_i a_i = S$.
$\tilde{P}(T, S)$ is the set of conjugacy classes w.r.t. cyclic permutations.

Definition

$x \leqslant_{mms} y$ in $\tilde{P}(T, S)$, if $f(x) \leqslant f(y)$ for all D'-mm function f which is invariant by cyclic permutations.

Theorem

The balanced word $w(T, S)$ is minimal over $\tilde{P}(T, S)$ for the mms-order.

Multimodular order and shifts

Let $P(T, S)$ be the set of all integer sequences of size T that sum to S.
D' is the mesh induced from D over the hyperplane $\sum_i a_i = S$.
$\tilde{P}(T, S)$ is the set of conjugacy classes w.r.t. cyclic permutations.

Definition

$x \leqslant_{mms} y$ in $\tilde{P}(T, S)$, if $f(x) \leqslant f(y)$ for all D'-mm function f which is invariant by cyclic permutations.

Theorem

The balanced word $w(T, S)$ is minimal over $\tilde{P}(T, S)$ for the mms-order.

Corollary

$x \leqslant_{mms} y$ if the "cone"-distance to $w(T, S)$ over mesh $M_{D'}$ is larger for y.

Unbalance in $P(T, S)$.

Partial sums : $k_s(n) = s_i + \cdots + s_n$
Discrepancy : $\phi_s(n) = k_s(n) - nS/T$.

Theorem

If T, S are co-prime, among all cyclic permutations of s, only one verifies $\forall n, \phi_s(n) \geqslant 0$ and only one verifies $\forall n, \phi_s(n) \leqslant 0$, called $\overline{s}$ and $\underline{s}$ respectively.

The discrepancy also induces orders over $\tilde{P}(T, S)$:

Definition

$u \leqslant_{\overline{g}} v$ if $\forall n$, $\phi_{\overline{u}}(n) \leqslant \phi_{\overline{v}}(n)$.
$u \leqslant_{\underline{g}} v$ if $\forall n$, $\phi_{\underline{u}}(n) \leqslant \phi_{\underline{v}}(n)$.
$u \leqslant_g v$ if $u \leqslant_{\underline{g}} v$ and $u \leqslant_{\overline{g}} v$.

Theorem

$w = w(T, S)$ is minimal over $\tilde{P}(T, S)$ for the orders $g, \overline{g}, \underline{g}$.

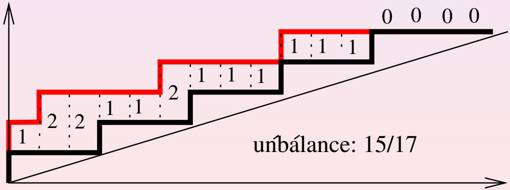
Unbalance(II)

The upper unbalance of a sequence u in $P(T, S)$ is

$$\bar{I}(u) = \frac{1}{T}\sum_{n=1}^{T} k_{\overline{u}}(n) - k_{\overline{w}}(n)$$

The lower unbalance of a sequence u in $P(T, S)$ is

$$\underline{I}(u) = \frac{1}{T}\sum_{n=1}^{T} k_{\underline{u}}(n) - k_{\underline{w}}(n)$$



Links between the orders

In $\mathbb{Z}^n$:

$$u \leqslant_{mm} v \Leftrightarrow \text{ Path } u \rightarrow v \text{ away from } r$$

In $P(T, S)$,

$$\text{Path } u \rightarrow v \text{ away from } w \Rightarrow u \leqslant_{mms} v$$

$$\text{Path } u \rightarrow v \text{ away from } w \not\Leftarrow u \leqslant_{g}$$

$$u \leqslant_{g} \not\Rightarrow u \leqslant_{mms} v$$

Links between the orders

In $\mathbb{Z}^n$:

$$u \leqslant_{mm} v \Leftrightarrow \text{ Path } u \rightarrow v \text{ away from } r$$

In $P(T, S)$,

$$\text{Path } u \rightarrow v \text{ away from } w \Rightarrow u \leqslant_{mms} v$$

$$\text{Path } u \rightarrow v \text{ away from } w \not\Rightarrow u \leqslant_{g}$$

$$u \leqslant_{g} \not\Rightarrow u \leqslant_{mms} v$$

Therefore, the g-order does not look very useful. However, the unbalance provides the good measure of the gap with an optimal sequence w.r.t. waiting times.

Unbalance and waiting times

If $W(u)$ is the long run average waiting time in a queue, where interarrival times and service times and iid sequences. then

Unbalance and waiting times

If $W(u)$ is the long run average waiting time in a queue, where interarrival times and service times and iid sequences. then

Theorem

W is D'-mm over $P(T, S)$ when D is the set of left shifts.
W is invariant by cyclic permutations.

Unbalance and waiting times

If $W(u)$ is the long run average waiting time in a queue, where interarrival times and service times and iid sequences. then

Theorem

W is D'-mm over $P(T, S)$ when D is the set of left shifts.
W is invariant by cyclic permutations.

Therefore, $u \leqslant_{mms} v$ implies $W(u) \leqslant W(v)$.

Unbalance and waiting times

If $W(u)$ is the long run average waiting time in a queue, where interarrival times and service times and iid sequences. then

Theorem

W is D'-mm over $P(T, S)$ when D is the set of left shifts.
W is invariant by cyclic permutations.

Therefore, $u \leqslant_{mms} v$ implies $W(u) \leqslant W(v)$. The density of $u \in P(T, S)$ is $d = S/T$.

Theorem

$$W(u) \leqslant W(w) + \frac{\mathbb{E}(\delta)}{d}\bar{I}(u).$$

Moreover, if the system is deterministic and $\delta = d\sigma$, then

$$W(u) = W(w) + \frac{\delta}{d}\bar{I}(u).$$

Routing to parallel queues

We consider all routing sequences a where the frequency of routing in queue i is d_i. The admission sequence in queue i, $u^i = \mathbf{1}_i(a)$ is a binary sequence with density d_i.

Routing to parallel queues

We consider all routing sequences a where the frequency of routing in queue i is d_i. The admission sequence in queue i, $u^i = \mathbf{1}_i(a)$ is a binary sequence with density d_i.
The average waiting time of all packets can be written as $C(a) = \sum d_i W(u^i)$.

Routing to parallel queues

We consider all routing sequences a where the frequency of routing in queue i is d_i. The admission sequence in queue i, $u^i = \mathbf{1}_i(a)$ is a binary sequence with density d_i.
The average waiting time of all packets can be written as $C(a) = \sum d_i W(u^i)$.

Theorem

Let $L = \sum_{i=1}^{n} d_i W(w(d_i))$, *then*

$$L \leqslant C(a) \leqslant L + \mathbb{E}(\delta) \sum_{i=1}^{n} \bar{I}(u_i).$$

Routing to parallel queues

We consider all routing sequences a where the frequency of routing in queue i is d_i. The admission sequence in queue i, $u^i = \mathbf{1}_i(a)$ is a binary sequence with density d_i.
The average waiting time of all packets can be written as $C(a) = \sum d_i W(u^i)$.

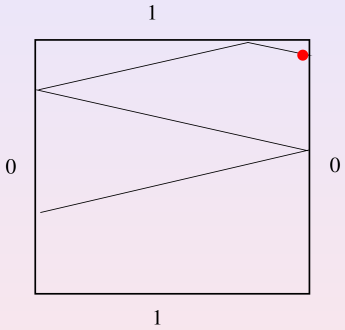
Theorem

Let $L = \sum_{i=1}^n d_i W(w(d_i))$*, then*

$$L \leqslant C(a) \leqslant L + \mathbb{E}(\delta) \sum_{i=1}^n \bar{I}(u_i).$$

Moreover, in the deterministic case, under load one, $C(a) = L + \mathbb{E}(\delta) \sum_{i=1}^n \bar{I}(u_i)$.

Billiard sequences



m = 0 0 1 0

Billiards and unbalance

In dimension n, there exists an optimal starting point for rational billiard sequences in terms of unbalance Van der Laan, 2004

Billiards and unbalance

In dimension n, there exists an optimal starting point for rational billiard sequences in terms of unbalance Van der Laan, 2004

The optimal starting point can be computed using an integer program. The optimal unbalance of a rational billiard sequence is $n/2 - 1$.

Billiards and unbalance

In dimension n, there exists an optimal starting point for rational billiard sequences in terms of unbalance Van der Laan, 2004

The optimal starting point can be computed using an integer program. The optimal unbalance of a rational billiard sequence is $n/2 - 1$.

This provides in polynomial time a routing sequence which cost is at most $\sum d_i W(w(d_i)) + \mathbb{E}(\delta) n/2 - 1$.