Positive-definite Functions, Exponential Sums and the Greedy Algorithm: a Curious Phenomenon

Louis Brown

Yale University

August 12, 2020

The Big Picture

Let $f(x) = x^2 - x + 1/6$ on the torus (identifying 0 with 1 on the real line)

Let $f(x) = x^2 - x + 1/6$ on the torus (identifying 0 with 1 on the real line).

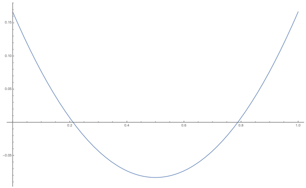


Figure: $f(x)$

The Big Picture

Let $f(x) = x^2 - x + 1/6$ on the torus (identifying 0 with 1 on the real line).

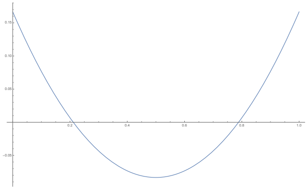


Figure: $f(x)$

We define a sequence of points by starting with an arbitrary initial set $\{x_1, \ldots, x_m\} \subset [0, 1]$

The Big Picture

Let $f(x) = x^2 - x + 1/6$ on the torus (identifying 0 with 1 on the real line).

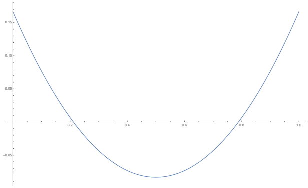


Figure: $f(x)$

We define a sequence of points by starting with an arbitrary initial set $\{x_1, \ldots, x_m\} \subset [0,1]$ and then greedily setting

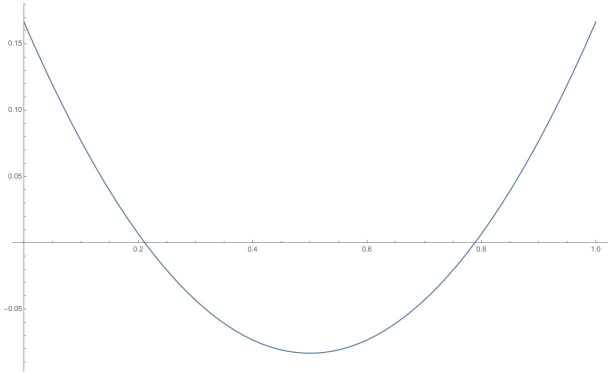
$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

The Big Picture

What happens if we start with $\{0\}$?

What happens if we start with $\{0\}$?

Figure: $f(x)$

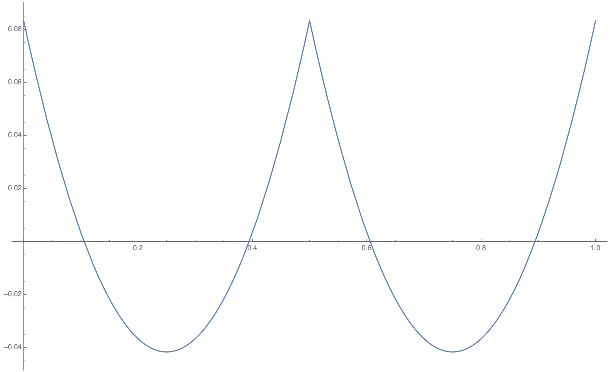


0

The Big Picture

What happens if we start with $\{0\}$?

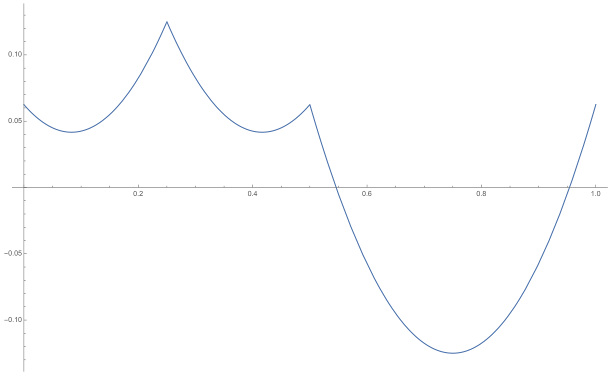
Figure: $f(x) + f(x - .5)$



$$0, .5$$

What happens if we start with $\{0\}$?

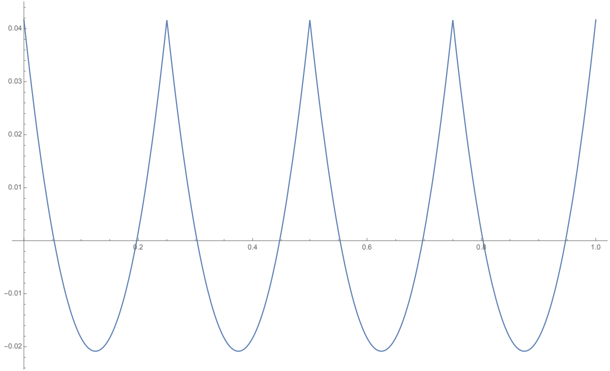
Figure: $f(x) + f(x - .5) + f(x - .25)$



$$0, .5, .25$$

What happens if we start with $\{0\}$?

Figure: $f(x) + f(x - .5) + f(x - .25) + f(x - .75)$

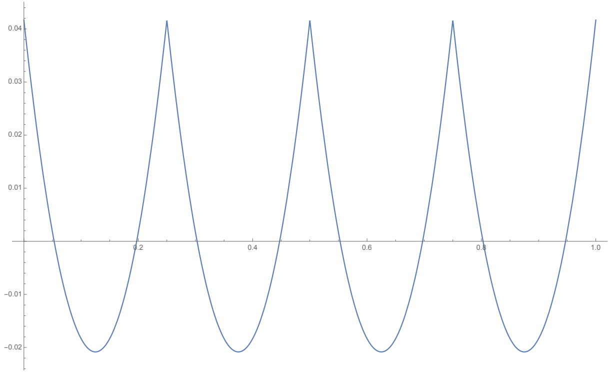


$$0, .5, .25, .75, \ldots$$

The Big Picture

What happens if we start with $\{0\}$?

Figure: $f(x) + f(x - .5) + f(x - .25) + f(x - .75)$



$$0, .5, .25, .75, \ldots$$

We see that this produces an **extremely** regular sequence.

Talk Structure

1. Two Friendly Sequences

Talk Structure

1. Two Friendly Sequences
2. Notions of Regularity

Talk Structure

1. Two Friendly Sequences
2. Notions of Regularity
3. A Fun Game

Talk Structure

1. Two Friendly Sequences
2. Notions of Regularity
3. A Fun Game
4. A Curious Phenomenon

Talk Structure

1. Two Friendly Sequences
2. Notions of Regularity
3. A Fun Game
4. A Curious Phenomenon
5. Higher Dimensions

Two Friendly Sequences

Two Friendly Sequences

We introduce two very structurally different examples of sequences.

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part).

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1$$

$$.5$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01$$

$$.5, .25$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11$$

$$.5, .25, .75$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11, .001$$

$$.5, .25, .75, .125$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11, .001, .101$$

$$.5, .25, .75, .125, .625$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11, .001, .101, .011$$

$$.5, .25, .75, .125, .625, .375$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11, .001, .101, .011, .111$$

$$.5, .25, .75, .125, .625, .375, .875$$

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{n\sqrt{2}\right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11, .001, .101, .011, .111$$

$$.5, .25, .75, .125, .625, .375, .875$$

Base 2 can be replaced by any other prime base.

Two Friendly Sequences

We introduce two very structurally different examples of sequences. These two competing examples are

1. the *Kronecker sequence* given by $x_n = \left\{ n\sqrt{2} \right\}$ (where $\{\cdot\}$ denotes the fractional part). Here, $\sqrt{2}$ could be replaced by any other number with bounded continued fraction expansion.
2. The *van der Corput sequence* given by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n when written in binary.

$$.1, .01, .11, .001, .101, .011, .111$$

$$.5, .25, .75, .125, .625, .375, .875$$

Base 2 can be replaced by any other prime base.

These two sequences were both cooked up to be optimally regular, but they are very differently built–let's take a closer look.

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{n\sqrt{2}\right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1.

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{n\sqrt{2}\right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{ n\sqrt{2} \right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

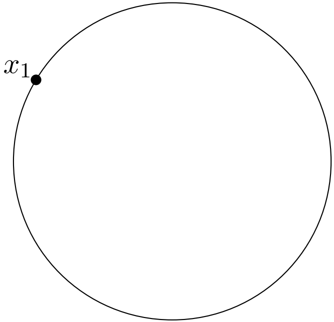


Figure: 7 terms of the Kronecker sequence

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{ n\sqrt{2} \right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

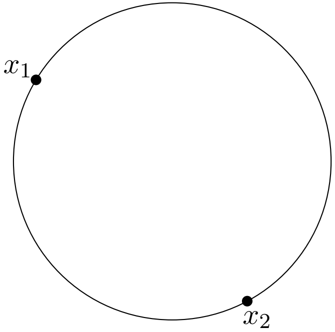


Figure: 7 terms of the Kronecker sequence

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{ n\sqrt{2} \right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

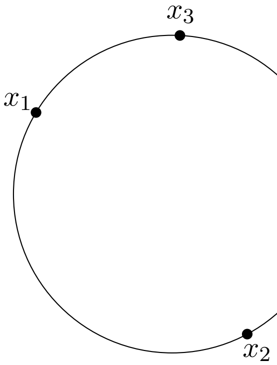


Figure: 7 terms of the Kronecker sequence

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{ n\sqrt{2} \right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

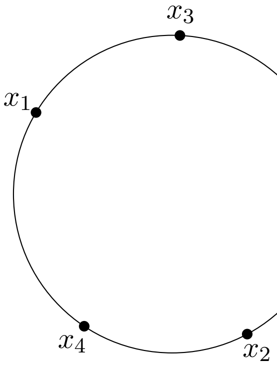


Figure: 7 terms of the Kronecker sequence

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{n\sqrt{2}\right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

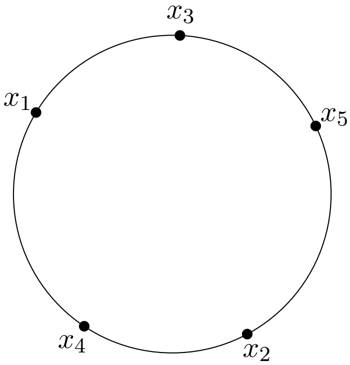


Figure: 7 terms of the Kronecker sequence

The Kronecker Sequence

The Kronecker sequence is defined by $x_n = \left\{n\sqrt{2}\right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

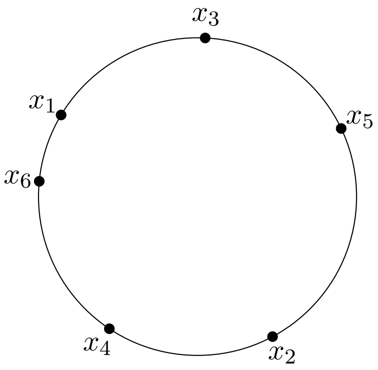


Figure: 7 terms of the Kronecker sequence

The Kronecker sequence is defined by $x_n = \left\{ n\sqrt{2} \right\}$, where $\{\cdot\}$ denotes taking the fractional part, or mod 1. The Kronecker sequence uses irrational rotations on a torus to produce uniformity:

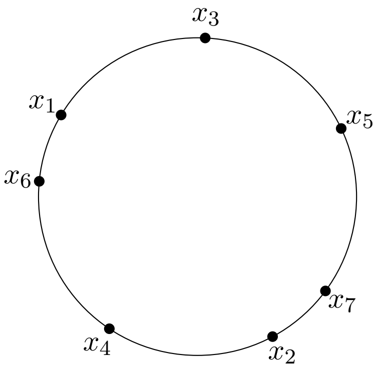


Figure: 7 terms of the Kronecker sequence

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

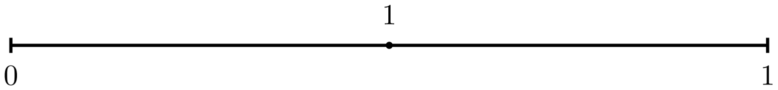


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

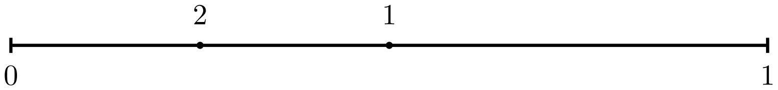


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

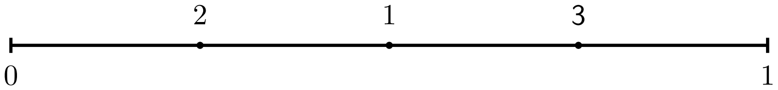


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

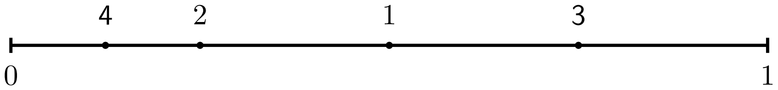


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

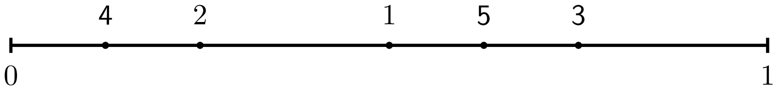


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

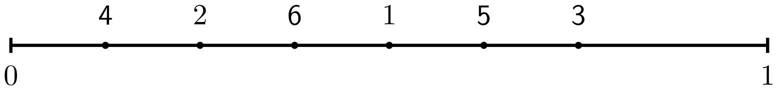


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

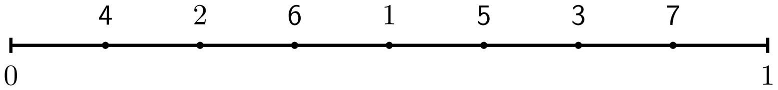


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

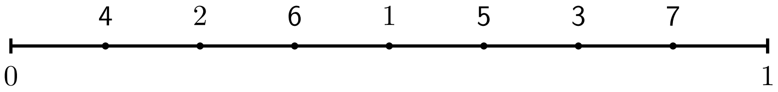


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput sequence uses the regularity of binary expansions of numbers to produce uniformity.

The van der Corput Sequence

The van der Corput sequence is defined by taking x_n to be the rational number whose binary expansion is the reversed string of bits of n, written in binary.

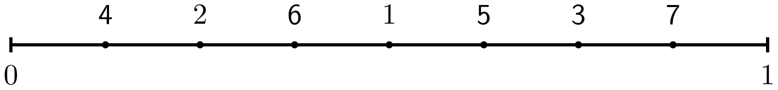


Figure: The first 7 terms of the van der Corput sequence.

The van der Corput sequence uses the regularity of binary expansions of numbers to produce uniformity. Note that it greedily "fills in the gaps"–at each step, it places a point at the midpoint of the longest empty interval. We'll come back to this...

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"?

Notions of Regularity

But what does it mean to be “optimally uniformly distributed”? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial
2. Analytical

Notions of Regularity

But what does it mean to be “optimally uniformly distributed”? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial
2. Analytical
3. Numerical

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial
2. Analytical
3. Numerical
4. Geometric

Notions of Regularity

But what does it mean to be “optimally uniformly distributed”? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial
2. Analytical
3. Numerical
4. Geometric

You do not need to memorize these notions for the talk!

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial
2. Analytical
3. Numerical
4. Geometric

You do not need to memorize these notions for the talk! In fact, they're are all, *very* loosely, equivalent

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0, 1]$.

1. Combinatorial
2. Analytical
3. Numerical
4. Geometric

You do not need to memorize these notions for the talk! In fact, they're are all, *very* loosely, equivalent: they are optimizing the same things.

Notions of Regularity

But what does it mean to be "optimally uniformly distributed"? We introduce a number of different quantitative approaches to measuring regularity of sequences in $[0,1]$.

1. Combinatorial
2. Analytical
3. Numerical
4. Geometric

You do not need to memorize these notions for the talk! In fact, they're are all, *very* loosely, equivalent: they are optimizing the same things. This is why the Kronecker and van der Corput sequences are able to perform optimally on all of them simultaneously.

Combinatorial Regularity

Combinatorial.

Combinatorial Regularity

Combinatorial. For every $n \in \mathbb{N}$, the set $\{x_1, \ldots, x_n\}$ has the property that for every interval $J \subset [0,1]$, the number of elements in J is $|J| \cdot n$ with a very small error.

Combinatorial Regularity

Combinatorial. For every $n \in \mathbb{N}$, the set $\{x_1, \ldots, x_n\}$ has the property that for every interval $J \subset [0, 1]$, the number of elements in J is $|J| \cdot n$ with a very small error. That is, each sub-interval of $[0, 1]$ should have the "correct" number of the sequence's points inside it

Combinatorial. For every $n \in \mathbb{N}$, the set $\{x_1, \ldots, x_n\}$ has the property that for every interval $J \subset [0,1]$, the number of elements in J is $|J| \cdot n$ with a very small error. That is, each sub-interval of $[0,1]$ should have the "correct" number of the sequence's points inside it, where "correct" means proportional to the length of the sub-interval.

Combinatorial Regularity

Combinatorial. For every $n \in \mathbb{N}$, the set $\{x_1, \ldots, x_n\}$ has the property that for every interval $J \subset [0,1]$, the number of elements in J is $|J| \cdot n$ with a very small error. That is, each sub-interval of $[0,1]$ should have the "correct" number of the sequence's points inside it, where "correct" means proportional to the length of the sub-interval.

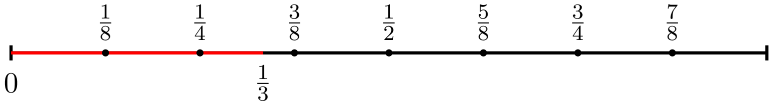


Figure: 7 terms of the van der Corput sequence; 2 lie in $(0, 1/3)$.

Analytical Regularity

Analytical (Erdős-Turán '48)**.**

Analytical Regularity

Analytical (Erdős-Turán '48)**.** The sequence has the property that $\{x_1, \ldots, x_n\}$ satisfy favorable exponential sum estimates on expressions of the form

$$\sum_{k=1}^{n} \frac{1}{k} \left| \sum_{\ell=1}^{n} e^{2\pi i k x_\ell} \right| \qquad \text{and} \qquad \sum_{k=1}^{n} \frac{1}{k^2} \left| \sum_{\ell=1}^{n} e^{2\pi i k x_\ell} \right|^2 .$$

Analytical (Erdős-Turán '48)**.** The sequence has the property that $\{x_1, \ldots, x_n\}$ satisfy favorable exponential sum estimates on expressions of the form

$$\sum_{k=1}^{n} \frac{1}{k} \left| \sum_{\ell=1}^{n} e^{2\pi i k x_\ell} \right| \quad \text{and} \quad \sum_{k=1}^{n} \frac{1}{k^2} \left| \sum_{\ell=1}^{n} e^{2\pi i k x_\ell} \right|^2.$$

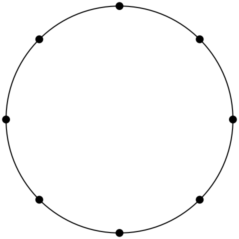
The exponential sum $\sum_{\ell=1}^{n} \exp(2\pi i k x_\ell)$ is 'small' for 'small' values of k.

Analytical Regularity

Analytical (Erdős-Turán '48)**.** The sequence has the property that $\{x_1, \ldots, x_n\}$ satisfy favorable exponential sum estimates on expressions of the form

$$\sum_{k=1}^{n} \frac{1}{k} \left| \sum_{\ell=1}^{n} e^{2\pi i k x_\ell} \right| \qquad \text{and} \qquad \sum_{k=1}^{n} \frac{1}{k^2} \left| \sum_{\ell=1}^{n} e^{2\pi i k x_\ell} \right|^2.$$

The exponential sum $\sum_{\ell=1}^{n} \exp(2\pi i k x_\ell)$ is 'small' for 'small' values of k.



Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$

Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$:

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k (\ell\sqrt{2})}\right|$$

Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$:

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k(\ell\sqrt{2})}\right| = \left|\frac{1 - e^{2\pi i k n\sqrt{2}}}{1 - e^{2\pi i k\sqrt{2}}}\right|,$$

this is just a geometric series!

Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$:

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k(\ell\sqrt{2})}\right| = \left|\frac{1 - e^{2\pi i k n\sqrt{2}}}{1 - e^{2\pi i k\sqrt{2}}}\right|,$$

this is just a geometric series! In particular, the numerator has norm at most 2 (by the triangle inequality), so the problem reduces to bounding the denominator.

Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$:

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k(\ell\sqrt{2})}\right| = \left|\frac{1 - e^{2\pi i k n\sqrt{2}}}{1 - e^{2\pi i k\sqrt{2}}}\right|,$$

this is just a geometric series! In particular, the numerator has norm at most 2 (by the triangle inequality), so the problem reduces to bounding the denominator. $\sqrt{2}$ is badly approximable

Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$:

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k(\ell\sqrt{2})}\right| = \left|\frac{1 - e^{2\pi i k n\sqrt{2}}}{1 - e^{2\pi i k\sqrt{2}}}\right|,$$

this is just a geometric series! In particular, the numerator has norm at most 2 (by the triangle inequality), so the problem reduces to bounding the denominator. $\sqrt{2}$ is badly approximable, meaning that integer multiples of it don't get too close to integers too quickly

Analytical Regularity

It's very easy to compute the exponential sum

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k x_\ell}\right|$$

on the Kronecker sequence $x_\ell = \{\ell\sqrt{2}\}$:

$$\left|\sum_{\ell=1}^{n} e^{2\pi i k(\ell\sqrt{2})}\right| = \left|\frac{1-e^{2\pi i k n\sqrt{2}}}{1-e^{2\pi i k\sqrt{2}}}\right|,$$

this is just a geometric series! In particular, the numerator has norm at most 2 (by the triangle inequality), so the problem reduces to bounding the denominator. $\sqrt{2}$ is badly approximable, meaning that integer multiples of it don't get too close to integers too quickly, so the denominator doesn't get too small.

Numerical Regularity

Numerical (Koksma-Hlawka '61).

Numerical Regularity

Numerical (Koksma-Hlawka '61)**.** The set $\{x_1, \ldots, x_n\}$ is a good set for numerical integration

Numerical Regularity

Numerical (Koksma-Hlawka '61). The set $\{x_1, \ldots, x_n\}$ is a good set for numerical integration: we have

$$\int_0^1 f(x)dx \sim \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

with a 'small' error for 'smooth' functions f (we'll make this more precise later).

Numerical Regularity

Numerical (Koksma-Hlawka '61)**.** The set $\{x_1, \ldots, x_n\}$ is a good set for numerical integration: we have

$$\int_0^1 f(x)dx \sim \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

with a 'small' error for 'smooth' functions f (we'll make this more precise later). This is extremely relevant for applications, where taking an integral is often expensive (or impossible) and we need to instead pick a good sampling of points to average.

Numerical Regularity

Numerical (Koksma-Hlawka '61). The set $\{x_1, \ldots, x_n\}$ is a good set for numerical integration: we have

$$\int_0^1 f(x)dx \sim \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

with a 'small' error for 'smooth' functions f (we'll make this more precise later). This is extremely relevant for applications, where taking an integral is often expensive (or impossible) and we need to instead pick a good sampling of points to average.
If we use $f(x) = x$ and the first 7 terms of the van der Corput sequence, we have

$$\frac{1}{7}(1/2 + 1/4 + 3/4 + 1/8 + 5/8 + 3/8 + 7/8)$$

Numerical Regularity

Numerical (Koksma-Hlawka '61). The set $\{x_1, \ldots, x_n\}$ is a good set for numerical integration: we have

$$\int_0^1 f(x)dx \sim \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

with a 'small' error for 'smooth' functions f (we'll make this more precise later). This is extremely relevant for applications, where taking an integral is often expensive (or impossible) and we need to instead pick a good sampling of points to average.
If we use $f(x) = x$ and the first 7 terms of the van der Corput sequence, we have

$$\frac{1}{7}(1/2 + 1/4 + 3/4 + 1/8 + 5/8 + 3/8 + 7/8) = 1/2$$

Numerical Regularity

Numerical (Koksma-Hlawka '61)**.** The set $\{x_1, \ldots, x_n\}$ is a good set for numerical integration: we have

$$\int_0^1 f(x)dx \sim \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

with a 'small' error for 'smooth' functions f (we'll make this more precise later). This is extremely relevant for applications, where taking an integral is often expensive (or impossible) and we need to instead pick a good sampling of points to average.
If we use $f(x) = x$ and the first 7 terms of the van der Corput sequence, we have

$$\frac{1}{7}(1/2 + 1/4 + 3/4 + 1/8 + 5/8 + 3/8 + 7/8) = 1/2,$$

integrating f over the unit interval exactly.

Geometric (Roth '54).

Geometric Regularity

Geometric (Roth '54)**.** The two-dimensional set

$$\left\{\left(\frac{i}{n}, x_i\right) : 1 \leq i \leq n\right\} \subset [0,1]^2$$

is uniformly distributed in the unit square

Geometric Regularity

Geometric (Roth '54). The two-dimensional set

$$\left\{\left(\frac{i}{n}, x_i\right) : 1 \leq i \leq n\right\} \subset [0,1]^2$$

is uniformly distributed in the unit square: every cartesian box $[a,b] \times [c,d]$ contains roughly $(b-a)(d-c)n$ elements with a small error.

Geometric Regularity

Geometric (Roth '54)**.** The two-dimensional set

$$\left\{\left(\frac{i}{n}, x_i\right) : 1 \leq i \leq n\right\} \subset [0,1]^2$$

is uniformly distributed in the unit square: every cartesian box $[a,b] \times [c,d]$ contains roughly $(b-a)(d-c)n$ elements with a small error. This is *very* closely related to the Combinatorial notion.

Geometric Regularity

Geometric (Roth '54). The two-dimensional set

$$\left\{\left(\frac{i}{n}, x_i\right) : 1 \leq i \leq n\right\} \subset [0,1]^2$$

is uniformly distributed in the unit square: every cartesian box $[a,b] \times [c,d]$ contains roughly $(b-a)(d-c)n$ elements with a small error. This is *very* closely related to the Combinatorial notion.

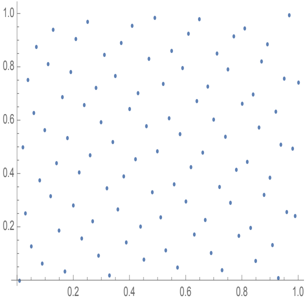


Figure: The first 100 terms of the van der Corput sequence

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \leq i \leq N : x_i \in J\}}{N} - |J| \right|.$$

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \leq i \leq N : x_i \in J\}}{N} - |J| \right|.$$

Remarks:

- This is a quantitative measure of the Combinatorial notion from earlier: the Discrepancy is the largest “error” of any interval after n terms.

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \le i \le N : x_i \in J\}}{N} - |J| \right|.$$

Remarks:

- This is a quantitative measure of the Combinatorial notion from earlier: the Discrepancy is the largest “error” of any interval after n terms.
- We trivially have $D_N \ge 1/N$ for all N and any sequence.

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \leq i \leq N : x_i \in J\}}{N} - |J| \right|.$$

Remarks:

- This is a quantitative measure of the Combinatorial notion from earlier: the Discrepancy is the largest "error" of any interval after n terms.
- We trivially have $D_N \geq 1/N$ for all N and any sequence. Simply consider the interval $[x_1 - \epsilon, x_1 + \epsilon]$ as $\epsilon \to 0$: it contains (at least) one point, but has arbitrarily small length.

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \leq i \leq N : x_i \in J\}}{N} - |J| \right|.$$

Remarks:

- This is a quantitative measure of the Combinatorial notion from earlier: the Discrepancy is the largest "error" of any interval after n terms.
- We trivially have $D_N \geq 1/N$ for all N and any sequence. Simply consider the interval $[x_1 - \epsilon, x_1 + \epsilon]$ as $\epsilon \to 0$: it contains (at least) one point, but has arbitrarily small length.
- The question naturally arises: can a discrepancy asymptotically on the order of $1/N$ be achieved?

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \leq i \leq N : x_i \in J\}}{N} - |J| \right|.$$

Remarks:

- This is a quantitative measure of the Combinatorial notion from earlier: the Discrepancy is the largest "error" of any interval after n terms.
- We trivially have $D_N \geq 1/N$ for all N and any sequence. Simply consider the interval $[x_1 - \epsilon, x_1 + \epsilon]$ as $\epsilon \to 0$: it contains (at least) one point, but has arbitrarily small length.
- The question naturally arises: can a discrepancy asymptotically on the order of $1/N$ be achieved?
- van der Corput sequence is always perfectly uniform after 2^n terms, attaining this bound.

Discrepancy

We define the *discrepancy* D_N of a sequence $\{x_n\}_{n=1}^{\infty}$ by

$$D_N = \sup_{\text{interval } J \subset [0,1]} \left| \frac{\#\{1 \leq i \leq N : x_i \in J\}}{N} - |J| \right|.$$

Remarks:

- This is a quantitative measure of the Combinatorial notion from earlier: the Discrepancy is the largest "error" of any interval after n terms.
- We trivially have $D_N \geq 1/N$ for all N and any sequence. Simply consider the interval $[x_1 - \epsilon, x_1 + \epsilon]$ as $\epsilon \to 0$: it contains (at least) one point, but has arbitrarily small length.
- The question naturally arises: can a discrepancy asymptotically on the order of $1/N$ be achieved?
- van der Corput sequence is always perfectly uniform after 2^n terms, attaining this bound. But! in between powers of 2, it accumulates a logarithmic error term.

Discrepancy

So, can any sequence beat van der Corput, and remain uniform even while still “filling in the gaps”?

Discrepancy

So, can any sequence beat van der Corput, and remain uniform even while still "filling in the gaps"? Johannes van der Corput himself posed this question originally.

So, can any sequence beat van der Corput, and remain uniform even while still "filling in the gaps"? Johannes van der Corput himself posed this question originally. The question was first answered by Tatyana van Aardenne-Ehrenfest in her 1945 paper "Proof of the Impossibility of a Just Distribution of an Infinite Sequence Over an Interval".

Discrepancy

So, can any sequence beat van der Corput, and remain uniform even while still "filling in the gaps"? Johannes van der Corput himself posed this question originally. The question was first answered by Tatyana van Aardenne-Ehrenfest in her 1945 paper "Proof of the Impossibility of a Just Distribution of an Infinite Sequence Over an Interval". In particular, we have

Theorem (Schmidt '72)

For any sequence $\{x_n\}_{n=1}^{\infty}$ *there are infinitely many integers* N *such that*

$$D_N \geq \frac{1}{100}\frac{\log N}{N}.$$

Discrepancy

So, can any sequence beat van der Corput, and remain uniform even while still "filling in the gaps"? Johannes van der Corput himself posed this question originally. The question was first answered by Tatyana van Aardenne-Ehrenfest in her 1945 paper "Proof of the Impossibility of a Just Distribution of an Infinite Sequence Over an Interval". In particular, we have

Theorem (Schmidt '72)

For any sequence $\{x_n\}_{n=1}^{\infty}$ there are infinitely many integers N such that

$$D_N \geq \frac{1}{100}\frac{\log N}{N}.$$

Up to constants, the Kronecker and van der Corput sequences achieve this lower bound.

A Fun Game

We present the following potential theoretic game:

A Fun Game

We present the following potential theoretic game: Start with a set of electrons, all with equal negative charge.

A Fun Game

We present the following potential theoretic game: Start with a set of electrons, all with equal negative charge. A pair of electrons at points x, y generate a potential of $|x - y|^{-1}$.

A Fun Game

We present the following potential theoretic game: Start with a set of electrons, all with equal negative charge. A pair of electrons at points x, y generate a potential of $|x - y|^{-1}$. Pick a new point to add an electron to by minimizing the energy of the system (which we can think of as minimizing the work we must do to keep it in place).

A Fun Game

We present the following potential theoretic game: Start with a set of electrons, all with equal negative charge. A pair of electrons at points x, y generate a potential of $|x - y|^{-1}$. Pick a new point to add an electron to by minimizing the energy of the system (which we can think of as minimizing the work we must do to keep it in place). The function we are minimizing is

$$f(x) = \sum_{i=1}^{n} |x - x_i|^{-1}.$$

We can imagine a shifted $1/|x|$ function placed over each point:

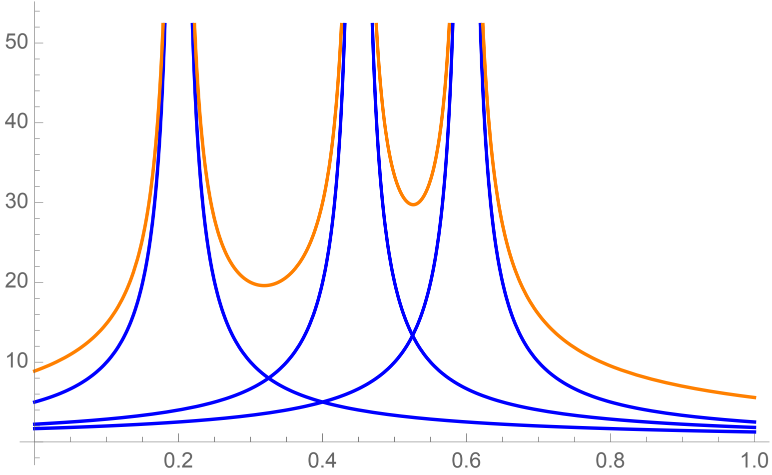


Figure: The potential of a point when charges are placed at .2,.6, and .45

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0.

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle: compact set, minimum exists, nice symmetry.

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle: compact set, minimum exists, nice symmetry.
Let's consider the second Bernoulli polynomial

$$f(x) = x^2 - x + \frac{1}{6}$$

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle: compact set, minimum exists, nice symmetry.
Let's consider the second Bernoulli polynomial

$$f(x) = x^2 - x + \frac{1}{6}$$

(identifying $\mathbb{T} \cong [0, 1]$).

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle: compact set, minimum exists, nice symmetry.
Let's consider the second Bernoulli polynomial

$$f(x) = x^2 - x + \frac{1}{6}$$

(identifying $\mathbb{T} \cong [0, 1]$). We define a sequence of points by starting with an arbitrary initial set $\{x_1, \ldots, x_m\} \subset [0, 1]$

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle: compact set, minimum exists, nice symmetry.

Let's consider the second Bernoulli polynomial

$$f(x) = x^2 - x + \frac{1}{6}$$

(identifying $\mathbb{T} \cong [0, 1]$). We define a sequence of points by starting with an arbitrary initial set $\{x_1, \ldots, x_m\} \subset [0, 1]$ and then greedily setting

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

A Fun Game

The game, as stated, is incoherent: f has no minimum, since we can place our charge further and further out, and the energy will approach 0. To resolve this, we simply play on a circle: compact set, minimum exists, nice symmetry.
Let's consider the second Bernoulli polynomial

$$f(x) = x^2 - x + \frac{1}{6}$$

(identifying $\mathbb{T} \cong [0,1]$). We define a sequence of points by starting with an arbitrary initial set $\{x_1, \ldots, x_m\} \subset [0,1]$ and then greedily setting

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

The resulting sequence has remarkable distribution properties!

A Fun Game

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

A Fun Game

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

What's the intuition behind this?

A Fun Game

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

What's the intuition behind this? Essentially, this is adding horizontal shifts of the function $f(x)$ together

A Fun Game

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

What's the intuition behind this? Essentially, this is adding horizontal shifts of the function $f(x)$ together, collecting in an aggregate function

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k)$$

which "fills in the gaps."

A Fun Game

$$x_n = \arg\min_{x\in\mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

What's the intuition behind this? Essentially, this is adding horizontal shifts of the function $f(x)$ together, collecting in an aggregate function

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k)$$

which "fills in the gaps." That is, by picking the argmin to shift by next, we add a function $f(x - x_{n+1})$ that will push up the value at the lowest point, and smooth out the aggregate.

A Fun Game

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k).$$

What's the intuition behind this? Essentially, this is adding horizontal shifts of the function $f(x)$ together, collecting in an aggregate function

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k)$$

which "fills in the gaps." That is, by picking the argmin to shift by next, we add a function $f(x - x_{n+1})$ that will push up the value at the lowest point, and smooth out the aggregate. Let's look at our example again.

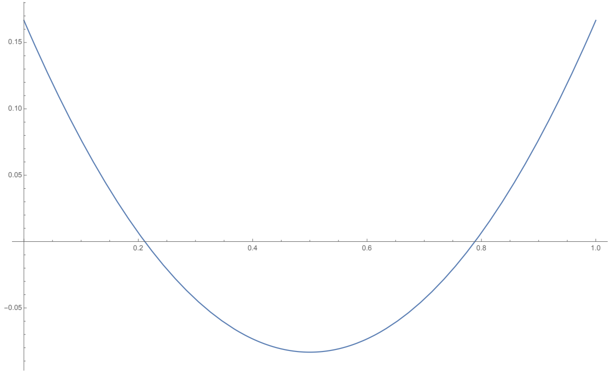
A Fun Game

What happens if we start with $\{0\}$?

A Fun Game

What happens if we start with $\{0\}$?

Figure: $f(x)$

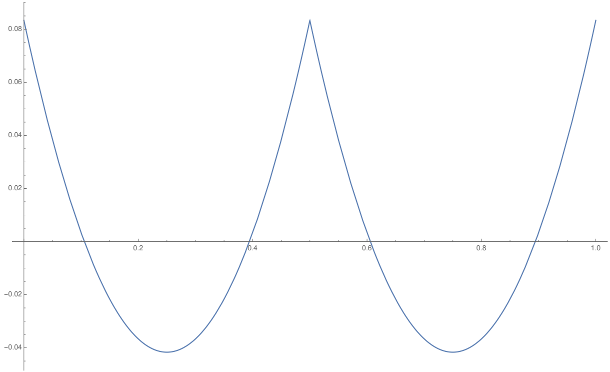


0

A Fun Game

What happens if we start with $\{0\}$?

Figure: $f(x) + f(x - .5)$

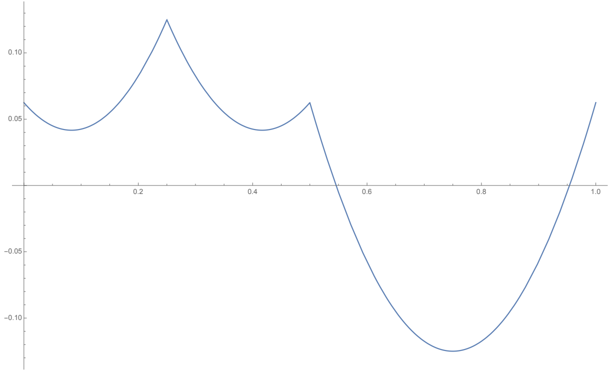


$$0, .5$$

A Fun Game

What happens if we start with $\{0\}$?

Figure: $f(x) + f(x - .5) + f(x - .25)$

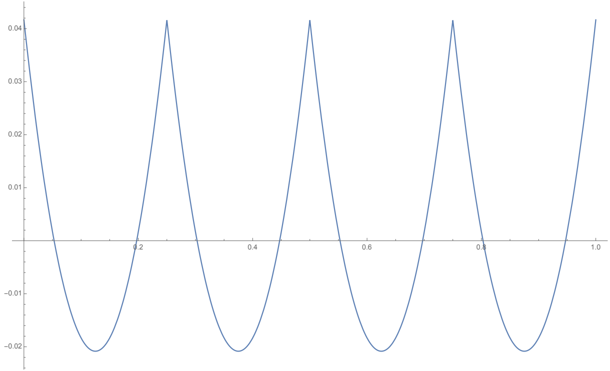


$$0, .5, .25$$

A Fun Game

What happens if we start with $\{0\}$?

Figure: $f(x) + f(x - .5) + f(x - .25) + f(x - .75)$



$$0, .5, .25, .75, \ldots$$

A Fun Game

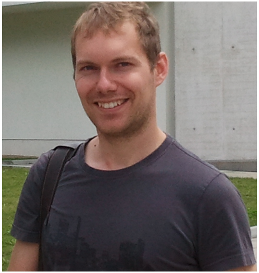


Figure: Florian Pausinger

Theorem (Pausinger '20)

A Fun Game

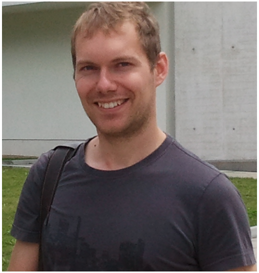


Figure: Florian Pausinger

Theorem (Pausinger '20)

Let $f : [0, 1] \to \mathbb{R}$ *be a function satisfying*

- $f(1 - x) = f(x)$

A Fun Game

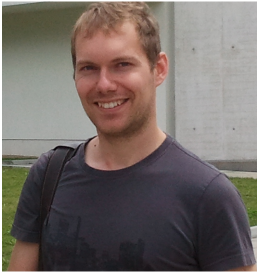


Figure: Florian Pausinger

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ be a function satisfying

- $f(1-x) = f(x)$ *(f is even)*

A Fun Game

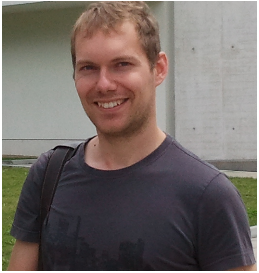


Figure: Florian Pausinger

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ *be a function satisfying*

- $f(1-x) = f(x)$ *(*f *is even)*
- $f''(x)$ *exists and is positive on* $(0,1)$.

A Fun Game

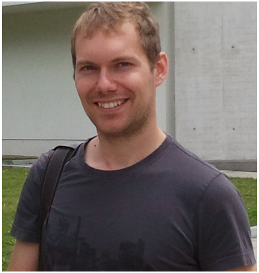


Figure: Florian Pausinger

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ *be a function satisfying*

- $f(1-x) = f(x)$ *(*f *is even)*
- $f''(x)$ *exists and is positive on* $(0,1)$.

Then the greedy algorithm running on f *and the initial set* $\{0\}$ *yields a van der Corput sequence.*

A Fun Game

We don't know what the sequence looks like if we just start with two elements though!

A Fun Game

We don't know what the sequence looks like if we just start with two elements though! But we can bound the discrepancy:

A Fun Game

We don't know what the sequence looks like if we just start with two elements though! But we can bound the discrepancy:

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ be bounded and symmetric about $1/2$. Further, assume $\widehat{f}(k) > c|k|^{-2}$ for some $c > 0$ and all $k \neq 0$.

A Fun Game

We don't know what the sequence looks like if we just start with two elements though! But we can bound the discrepancy:

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ *be bounded and symmetric about* $1/2$. *Further, assume* $\widehat{f}(k) > c|k|^{-2}$ *for some* $c > 0$ *and all* $k \neq 0$. *Then all sequences defined via the greedy algorithm on an arbitrary initial set satisfy*

$$D_N \leq \frac{\tilde{c}}{N^{1/3}}$$

A Fun Game

We don't know what the sequence looks like if we just start with two elements though! But we can bound the discrepancy:

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ be bounded and symmetric about $1/2$. Further, assume $\widehat{f}(k) > c|k|^{-2}$ for some $c > 0$ and all $k \neq 0$. Then all sequences defined via the greedy algorithm on an arbitrary initial set satisfy

$$D_N \leq \frac{\tilde{c}}{N^{1/3}},$$

where $\tilde{c} > 0$ depends on the initial set.

A Fun Game

We don't know what the sequence looks like if we just start with two elements though! But we can bound the discrepancy:

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ be bounded and symmetric about $1/2$. Further, assume $\widehat{f}(k) > c|k|^{-2}$ for some $c > 0$ and all $k \neq 0$. Then all sequences defined via the greedy algorithm on an arbitrary initial set satisfy

$$D_N \leq \frac{\tilde{c}}{N^{1/3}},$$

where $\tilde{c} > 0$ depends on the initial set.

Florian Pausinger, "Greedy Energy Charges can Count in Binary: Point Charges and the van der Corput Sequence" (January 2020).

A Fun Game

We don't know what the sequence looks like if we just start with two elements though! But we can bound the discrepancy:

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ *be bounded and symmetric about* $1/2$*. Further, assume* $\widehat{f}(k) > c|k|^{-2}$ *for some* $c > 0$ *and all* $k \neq 0$*. Then all sequences defined via the greedy algorithm on an arbitrary initial set satisfy*

$$D_N \leq \frac{\tilde{c}}{N^{1/3}},$$

where $\tilde{c} > 0$ *depends on the initial set.*

Florian Pausinger, "Greedy Energy Charges can Count in Binary: Point Charges and the van der Corput Sequence" (January 2020).

Later in the talk, we will present the (very slick!) proof.

A Fun Game

How does the discrepancy of these sequences compare with van der Corput numerically?

A Fun Game

How does the discrepancy of these sequences compare with van der Corput numerically?

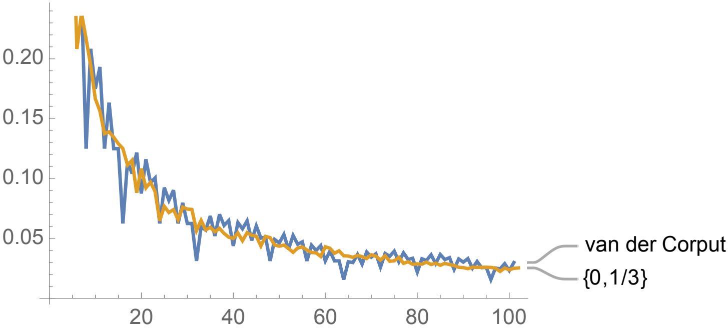


Figure: Discrepancy of van der Corput vs Algorithm running on $\{0, 1/3\}$

A Fun Game

How about something weirder?

A Fun Game

How about something weirder?

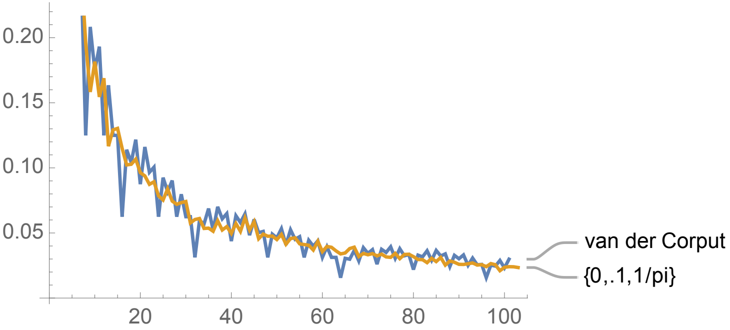


Figure: Discrepancy of van der Corput vs Algorithm running on $\{0, 1/10, 1/\pi\}$

A Fun Game

Maybe we can mess it up in the middle?

A Fun Game

Maybe we can mess it up in the middle?

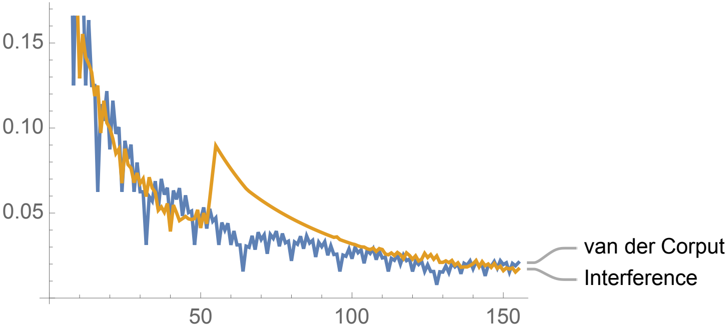


Figure: Discrepancy of van der Corput vs Algorithm running on $\{0, .6\}$ for 50 points, then add $\{.5, .51, .52\}$ to the sequence and run for another 100

A Fun Game

Maybe the Bernoulli polynomial is just a fluke.

A Fun Game

Maybe the Bernoulli polynomial is just a fluke. Let's try with another function: Let $f(x) = -\ln|2\sin \pi x|$, with Fourier series

$$\sum_{k=1}^{\infty} \frac{1}{k} \cos(2\pi k x) = \sum_{k \neq 0} \frac{1}{2|k|} e^{2\pi i k x}$$

A Fun Game

Maybe the Bernoulli polynomial is just a fluke. Let's try with another function: Let $f(x) = -\ln|2\sin\pi x|$, with Fourier series

$$\sum_{k=1}^{\infty} \frac{1}{k}\cos(2\pi kx) = \sum_{k\neq 0} \frac{1}{2|k|} e^{2\pi ikx},$$

i.e. the Green's function of the fractional Laplacian.

A Fun Game

Maybe the Bernoulli polynomial is just a fluke. Let's try with another function: Let $f(x) = -\ln|2\sin\pi x|$, with Fourier series

$$\sum_{k=1}^{\infty} \frac{1}{k}\cos(2\pi kx) = \sum_{k\neq 0} \frac{1}{2|k|} e^{2\pi ikx},$$

i.e. the Green's function of the fractional Laplacian.

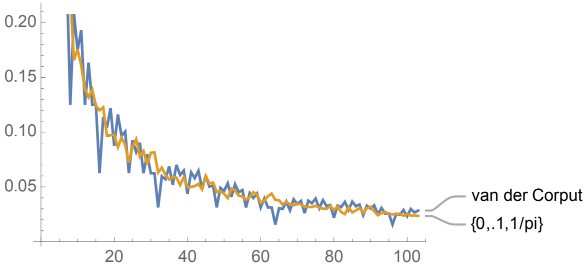


Figure: Discrepancy of van der Corput vs Algorithm running on $\{0, .1, 1/\pi\}$

Curious Phenomena

Conjecture 1 For all initial sets $\{x_1, \ldots, x_m\}$, and all even f such that $\hat{f}(k) > ck^{-2}$ for all $k \neq 0$, the greedy algorithm

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k)$$

will produce a sequence such that

$$\sum_{k,\ell=1}^{n} f(x_k - x_\ell) \lesssim \log n.$$

Curious Phenomena

Conjecture 1 For all initial sets $\{x_1, \ldots, x_m\}$, and all even f such that $\hat{f}(k) > ck^{-2}$ for all $k \neq 0$, the greedy algorithm

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k)$$

will produce a sequence such that

$$\sum_{k,\ell=1}^{n} f(x_k - x_\ell) \lesssim \log n.$$

(Proinov showed that this is optimal, no sequence can do better.)

Conjecture 2 For all initial sets $\{x_1, \ldots, x_m\}$, and all even f such that $\hat{f}(k) > ck^{-2}$ for all $k \neq 0$, the greedy algorithm

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k)$$

will produce a sequence such that

$$\left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \lesssim \log n.$$

Curious Phenomena

Conjecture 2 For all initial sets $\{x_1, \ldots, x_m\}$, and all even f such that $\hat{f}(k) > ck^{-2}$ for all $k \neq 0$, the greedy algorithm

$$x_n = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n-1} f(x - x_k)$$

will produce a sequence such that

$$\left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \lesssim \log n.$$

We really suspect both these to be true based on the numerics, but the results we can prove are much looser bounds.

Curious Phenomena

Let

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k).$$

Curious Phenomena

Let

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k).$$

Since $\|f(x - x_k)\|_{L^\infty} = \|f(x)\|_{L^\infty}$, we certainly have $\|f_n\|_{L^\infty} \lesssim n$.

Curious Phenomena

Let

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k).$$

Since $\|f(x - x_k)\|_{L^\infty} = \|f(x)\|_{L^\infty}$, we certainly have $\|f_n\|_{L^\infty} \lesssim n$. To make a stronger claim, we can apply the Koksma-Hlawka inequality, which bounds the error of numerical integration

Let

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k).$$

Since $\|f(x - x_k)\|_{L^\infty} = \|f(x)\|_{L^\infty}$, we certainly have $\|f_n\|_{L^\infty} \lesssim n$. To make a stronger claim, we can apply the Koksma-Hlawka inequality, which bounds the error of numerical integration:

Theorem (Koksma, '42)

For a function f with finite total variation

$$V(f) = \int_0^1 |f'(x)| dx$$

Curious Phenomena

Let

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k).$$

Since $\|f(x - x_k)\|_{L^\infty} = \|f(x)\|_{L^\infty}$, we certainly have $\|f_n\|_{L^\infty} \lesssim n$. To make a stronger claim, we can apply the Koksma-Hlawka inequality, which bounds the error of numerical integration:

Theorem (Koksma, '42)

For a function f with finite total variation

$$V(f) = \int_0^1 |f'(x)| dx$$

and any set of points $S = \{x_1, \ldots, x_N\}$

Curious Phenomena

Let

$$f_n(x) = \sum_{k=1}^{n} f(x - x_k).$$

Since $\|f(x-x_k)\|_{L^\infty} = \|f(x)\|_{L^\infty}$, we certainly have $\|f_n\|_{L^\infty} \lesssim n$. To make a stronger claim, we can apply the Koksma-Hlawka inequality, which bounds the error of numerical integration:

Theorem (Koksma, '42)

For a function f with finite total variation

$$V(f) = \int_0^1 |f'(x)|dx$$

and any set of points $S = \{x_1, \ldots, x_N\}$, we have

$$\left| \frac{1}{N} \sum_{i=1}^{N} f(x_i) - \int_0^1 f(x)dx \right| \leq V(f) D_N(S).$$

Curious Phenomena

$$\left| \frac{1}{N} \sum_{i=1}^{N} f(x_i) - \int_0^1 f(x)dx \right| \leq V(f)D_N(S).$$

$$\left|\frac{1}{N}\sum_{i=1}^{N} f(x_i) - \int_0^1 f(x)dx\right| \leq V(f)D_N(S).$$

To apply the Koksma-Hlawka inequality to our setting, we first note that our $f(x) = x^2 - x + 1/6$ is mean 0.

$$\left|\frac{1}{N}\sum_{i=1}^{N} f(x_i) - \int_0^1 f(x)dx\right| \leq V(f)D_N(S).$$

To apply the Koksma-Hlawka inequality to our setting, we first note that our $f(x) = x^2 - x + 1/6$ is mean 0. (We can always shift an energy function up or down to be mean 0, and it won't affect the algorithm, since it's a minimization process.)

$$\left|\frac{1}{N}\sum_{i=1}^{N} f(x_i) - \int_0^1 f(x)dx\right| \leq V(f)D_N(S).$$

To apply the Koksma-Hlawka inequality to our setting, we first note that our $f(x) = x^2 - x + 1/6$ is mean 0. (We can always shift an energy function up or down to be mean 0, and it won't affect the algorithm, since it's a minimization process.) Thus, the integral of f is 0, and the inequality simplifies to

$$\left|\frac{1}{N}\sum_{i=1}^{N} f(x_i)\right| \leq V(f)D_N(S).$$

$$\left| \frac{1}{N} \sum_{i=1}^{N} f(x_i) - \int_0^1 f(x)dx \right| \leq V(f)D_N(S).$$

To apply the Koksma-Hlawka inequality to our setting, we first note that our $f(x) = x^2 - x + 1/6$ is mean 0. (We can always shift an energy function up or down to be mean 0, and it won't affect the algorithm, since it's a minimization process.) Thus, the integral of f is 0, and the inequality simplifies to

$$\left| \frac{1}{N} \sum_{i=1}^{N} f(x_i) \right| \leq V(f)D_N(S).$$

$V(f)$ is a constant (only depends on f, not the set of points).

Curious Phenomena

Further, since discrepancy is translation invariant

Curious Phenomena

Further, since discrepancy is translation invariant, we have

$$D_n(x - x_1, \ldots x - x_n) = D_n(x_1, \ldots x_n).$$

Further, since discrepancy is translation invariant, we have

$$D_n(x - x_1, \ldots x - x_n) = D_n(x_1, \ldots x_n).$$

Thus, the inequality tells us that

$$\begin{aligned} \|f_n(x)\|_{L^\infty} &= \left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \\ &= n \left\| \frac{1}{n} \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \\ &\lesssim nD_n. \end{aligned}$$

Further, since discrepancy is translation invariant, we have

$$D_n(x - x_1, \ldots x - x_n) = D_n(x_1, \ldots x_n).$$

Thus, the inequality tells us that

$$\begin{aligned} \|f_n(x)\|_{L^\infty} &= \left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \\ &= n \left\| \frac{1}{n} \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \\ &\lesssim nD_n. \end{aligned}$$

Schmidt tells us this bound is, at best, $\sim \log n$.

Curious Phenomena

Further, since discrepancy is translation invariant, we have

$$D_n(x - x_1, \dots x - x_n) = D_n(x_1, \dots x_n).$$

Thus, the inequality tells us that

$$\begin{aligned}
\|f_n(x)\|_{L^\infty} &= \left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \\
&= n \left\| \frac{1}{n} \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \\
&\lesssim n D_n.
\end{aligned}$$

Schmidt tells us this bound is, at best, $\sim \log n$. Our Conjecture 2 posits that this is achieved in our setting.

Curious Phenomena

Empirically, $\|f_n(x)\|_{L^\infty}$ is extremely small in n (with $\|f_n(x)\|_{L^\infty}$ barely exceeding $\|f(x)\|_{L^\infty}$):

Curious Phenomena

Empirically, $\|f_n(x)\|_{L^\infty}$ is extremely small in n (with $\|f_n(x)\|_{L^\infty}$ barely exceeding $\|f(x)\|_{L^\infty}$):

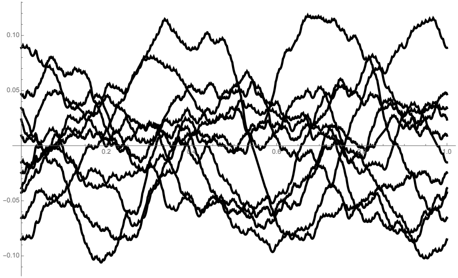


Figure: The functions $f_{100}, f_{110}, f_{120} \dots, f_{200}$.

Curious Phenomena

Empirically, $\|f_n(x)\|_{L^\infty}$ is extremely small in n (with $\|f_n(x)\|_{L^\infty}$ barely exceeding $\|f(x)\|_{L^\infty}$):

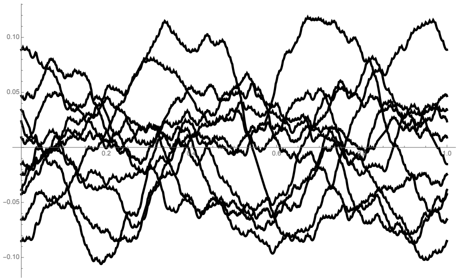


Figure: The functions $f_{100}, f_{110}, f_{120} \ldots, f_{200}$.

Somehow, all the shifted functions f balance out extremely nicely in such a way that the energy of the system stays low.

Curious Phenomena

In fact, the sequences defined via the algorithm appear, empirically, to be competitive with van der Corput and Kronecker sequences

Curious Phenomena

In fact, the sequences defined via the algorithm appear, empirically, to be competitive with van der Corput and Kronecker sequences (which are known to be optimal)

Curious Phenomena

In fact, the sequences defined via the algorithm appear, empirically, to be competitive with van der Corput and Kronecker sequences (which are known to be optimal) on all of the regularity measures discussed at the beginning

Curious Phenomena

In fact, the sequences defined via the algorithm appear, empirically, to be competitive with van der Corput and Kronecker sequences (which are known to be optimal) on all of the regularity measures discussed at the beginning: Combinatorially, Analytically, Numerically, and Geometrically.

Curious Phenomena

In fact, the sequences defined via the algorithm appear, empirically, to be competitive with van der Corput and Kronecker sequences (which are known to be optimal) on all of the regularity measures discussed at the beginning: Combinatorially, Analytically, Numerically, and Geometrically. However, a proof is evasive, and it is an open question whether or not they truly are.

Zinterhof's Diaphony

We can turn our sequence into a measure by setting

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k}.$$

Zinterhof's Diaphony

We can turn our sequence into a measure by setting

$$\mu_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{x_k}.$$

We may define Zinterhof's Diaphony of a measure μ as

$$F_N(\mu) = \left(\sum_{k\neq 0}\frac{|\widehat{\mu}(k)|^2}{k^2}\right)^{1/2}.$$

Zinterhof's Diaphony

We can turn our sequence into a measure by setting

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k}.$$

We may define Zinterhof's Diaphony of a measure μ as

$$F_N(\mu) = \left(\sum_{k \neq 0} \frac{|\widehat{\mu}(k)|^2}{k^2} \right)^{1/2}.$$

Note that

$$F_N(\mu) = \|\mu\|_{\dot{H}^{-1}}$$

Zinterhof's Diaphony

We can turn our sequence into a measure by setting

$$\mu_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{x_k}.$$

We may define Zinterhof's Diaphony of a measure μ as

$$F_N(\mu) = \left(\sum_{k\neq 0} \frac{|\widehat{\mu}(k)|^2}{k^2}\right)^{1/2}.$$

Note that

$$F_N(\mu) = \|\mu\|_{\dot{H}^{-1}},$$

people in number theory/combinatorics think of this as "diaphony", whereas the Sobolev norm is more analytical and shows up in PDEs.

Proof of Pausinger's Theorem

We now present a proof of Pausinger's Theorem:

Theorem (Pausinger '20)

Let $f : [0,1] \to \mathbb{R}$ be bounded and symmetric about $1/2$. Further, assume $\widehat{f}(k) > c|k|^{-2}$ for some $c > 0$ and all $k \neq 0$. Then all sequences defined via the greedy algorithm on an arbitrary initial set satisfy

$$D_N \leq \frac{\tilde{c}}{N^{1/3}},$$

where $\tilde{c} > 0$ depends on the initial set.

Proof of Pausinger's Theorem

Assume wlog that f is mean 0.

Proof of Pausinger's Theorem

Assume wlog that f is mean 0. Then

$$\sum_{m,\ell=1}^{n} f(x_m - x_\ell) = nf(0) + 2 \sum_{\substack{m,\ell=1 \\ m<\ell}}^{n} f(x_m - x_\ell)$$

$$= nf(0) + 2 \sum_{\ell=2}^{n} \sum_{m=1}^{\ell-1} f(x_\ell - x_m) \leq nf(0)$$

Proof of Pausinger's Theorem

Assume wlog that f is mean 0. Then

$$\sum_{m,\ell=1}^{n} f(x_m - x_\ell) = nf(0) + 2 \sum_{\substack{m,\ell=1 \\ m<\ell}}^{n} f(x_m - x_\ell)$$

$$= nf(0) + 2 \sum_{\ell=2}^{n} \sum_{m=1}^{\ell-1} f(x_\ell - x_m) \leq nf(0)$$

since, by definition of the greedy algorithm,

$$\sum_{m=1}^{\ell-1} f(x_\ell - x_m) = \min_x \sum_{m=1}^{\ell-1} f(x - x_m) \leq \int_{\mathbb{T}} \sum_{m=1}^{\ell-1} f(x - x_m) dx = 0.$$

Proof of Pausinger's Theorem

Assume wlog that f is mean 0. Then

$$\sum_{m,\ell=1}^{n} f(x_m - x_\ell) = nf(0) + 2\sum_{\substack{m,\ell=1\\ m<\ell}}^{n} f(x_m - x_\ell)$$
$$= nf(0) + 2\sum_{\ell=2}^{n}\sum_{m=1}^{\ell-1} f(x_\ell - x_m) \le nf(0)$$

since, by definition of the greedy algorithm,

$$\sum_{m=1}^{\ell-1} f(x_\ell - x_m) = \min_x \sum_{m=1}^{\ell-1} f(x - x_m) \le \int_{\mathbb{T}} \sum_{m=1}^{\ell-1} f(x - x_m)dx = 0.$$

Thus

$$\sum_{m,\ell=1}^{n} f(x_m - x_\ell) \le nf(0). \qquad (\diamond)$$

Proof of Pausinger's Theorem

On the other hand, we have

$$\begin{aligned}
\sum_{m,\ell=1}^{n} f(x_m - x_\ell) &= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \sum_{m,\ell=1}^{n} e^{2\pi i k(x_m - x_\ell)} \\
&= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left(\sum_{m=1}^{n} e^{2\pi i k x_m}\right) \left(\sum_{m=1}^{n} e^{2\pi i k(-x_m)}\right) \\
&= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left(\sum_{m=1}^{n} e^{2\pi i k x_m}\right) \overline{\left(\sum_{m=1}^{n} e^{2\pi i k x_m}\right)} \\
&= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left|\sum_{m=1}^{n} e^{2\pi i k x_m}\right|^2.
\end{aligned}$$

Proof of Pausinger's Theorem

On the other hand, we have

$$\begin{aligned}
\sum_{m,\ell=1}^{n} f(x_m - x_\ell) &= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \sum_{m,\ell=1}^{n} e^{2\pi i k(x_m - x_\ell)} \\
&= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left(\sum_{m=1}^{n} e^{2\pi i k x_m}\right) \left(\sum_{m=1}^{n} e^{2\pi i k(-x_m)}\right) \\
&= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left(\sum_{m=1}^{n} e^{2\pi i k x_m}\right) \overline{\left(\sum_{m=1}^{n} e^{2\pi i k x_m}\right)} \\
&= \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left|\sum_{m=1}^{n} e^{2\pi i k x_m}\right|^2.
\end{aligned}$$

Thus, combining with $\diamond$ from the previous slide,

$$\sum_{k\in\mathbb{Z}} \widehat{f}(k) \left|\sum_{m=1}^{n} e^{2\pi i k x_m}\right|^2 \leq n f(0).$$

Proof of Pausinger's Theorem

Now, we use the fact that $\widehat{f}(k) \geq c|k|^{-2}$

Proof of Pausinger's Theorem

Now, we use the fact that $\widehat{f}(k) \geq c|k|^{-2}$:

$$
\begin{aligned}
nf(0) &\geq \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left| \sum_{m=1}^{n} e^{2\pi i k x_m} \right|^2 \\
&\geq n^2 \sum_{k\neq 0} \frac{c}{k^2} \left| \frac{1}{n} \sum_{m=1}^{n} e^{2\pi i k x_m} \right|^2 \\
&= cn^2 \|\mu_n\|_{\dot{H}^{-1}}^2,
\end{aligned}
$$

Proof of Pausinger's Theorem

Now, we use the fact that $\widehat{f}(k) \geq c|k|^{-2}$:

$$
\begin{aligned}
nf(0) &\geq \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left| \sum_{m=1}^{n} e^{2\pi i k x_m} \right|^2 \\
&\geq n^2 \sum_{k\neq 0} \frac{c}{k^2} \left| \frac{1}{n} \sum_{m=1}^{n} e^{2\pi i k x_m} \right|^2 \\
&= cn^2 \|\mu_n\|_{\dot{H}^{-1}}^2,
\end{aligned}
$$

so $\|\mu_n\|_{\dot{H}^{-1}} \lesssim n^{-1/2}$.

Proof of Pausinger's Theorem

Now, we use the fact that $\widehat{f}(k) \geq c|k|^{-2}$:

$$\begin{aligned} nf(0) &\geq \sum_{k\in\mathbb{Z}} \widehat{f}(k) \left| \sum_{m=1}^{n} e^{2\pi i k x_m} \right|^2 \\ &\geq n^2 \sum_{k\neq 0} \frac{c}{k^2} \left| \frac{1}{n} \sum_{m=1}^{n} e^{2\pi i k x_m} \right|^2 \\ &= cn^2 \|\mu_n\|_{\dot{H}^{-1}}^2, \end{aligned}$$

so $\|\mu_n\|_{\dot{H}^{-1}} \lesssim n^{-1/2}$. Finally, by LeVeque's Inequality, we can bound the discrepancy as

$$D_n \lesssim \|\mu_n\|_{\dot{H}^{-1}}^{2/3} \lesssim n^{-1/3}$$

and we have the desired result. □

Higher Dimensions

Minimizing energy on a higher dimensional manifold is much harder.

Higher Dimensions

Minimizing energy on a higher dimensional manifold is much harder. The setup is the same as before: we want to place point charges such that the total energy is minimized

Higher Dimensions

Minimizing energy on a higher dimensional manifold is much harder. The setup is the same as before: we want to place point charges such that the total energy is minimized (i.e. each charge is far from the others, this is a regularization procedure).

Higher Dimensions

Minimizing energy on a higher dimensional manifold is much harder. The setup is the same as before: we want to place point charges such that the total energy is minimized (i.e. each charge is far from the others, this is a regularization procedure). A common energy function to use is the *Riesz kernel*:

$$k_s(x, y) = \begin{cases} -\log|x - y| & s = 0 \\ |x - y|^{-s} & s > 0 \end{cases}.$$

Higher Dimensions

Minimizing energy on a higher dimensional manifold is much harder. The setup is the same as before: we want to place point charges such that the total energy is minimized (i.e. each charge is far from the others, this is a regularization procedure). A common energy function to use is the *Riesz kernel*:

$$k_s(x,y) = \begin{cases} -\log|x-y| & s=0 \\ |x-y|^{-s} & s>0 \end{cases}.$$

It is known that minimizers of the Riesz potential are optimal with respect to the $\dot{H}^{-d/2}$ norm [Marzo, Mas '19] and uniformly distributed with respect to the Hausdorff measure (Poppy-seed bagel Theorem, [Hardin, Saff '04]).

Higher Dimensions

Minimizing energy on a higher dimensional manifold is much harder. The setup is the same as before: we want to place point charges such that the total energy is minimized (i.e. each charge is far from the others, this is a regularization procedure). A common energy function to use is the *Riesz kernel*:

$$k_s(x,y) = \begin{cases} -\log|x-y| & s=0 \\ |x-y|^{-s} & s>0 \end{cases}.$$

It is known that minimizers of the Riesz potential are optimal with respect to the $\dot{H}^{-d/2}$ norm [Marzo, Mas '19] and uniformly distributed with respect to the Hausdorff measure (Poppy-seed bagel Theorem, [Hardin, Saff '04]). It is also known that minimizers of the Green's kernel are asymptotically uniformly distributed [Beltràn, Corral, Criado del Ray '17].

Higher Dimensions

The problem on the sphere with Riesz kernel $|x - y|^{-1}$ dates back to physicist J.J. Thomson in 1904, yet, to this day, only a handful of cases are known:

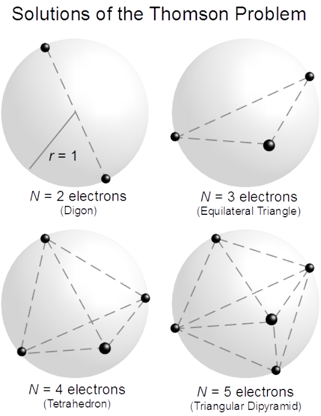


Figure: From Wikipedia

Higher Dimensions

If even the sphere poses a challenge, how can we hope to solve this on general manifolds?

Higher Dimensions

If even the sphere poses a challenge, how can we hope to solve this on general manifolds? Maybe we can play the same type of game, greedily picking the best point at each step?

Higher Dimensions

If even the sphere poses a challenge, how can we hope to solve this on general manifolds? Maybe we can play the same type of game, greedily picking the best point at each step?
This has been researched, and greedy sequences constructed this way (with any kernel) are called *Leja Points*.

Higher Dimensions

If even the sphere poses a challenge, how can we hope to solve this on general manifolds? Maybe we can play the same type of game, greedily picking the best point at each step?
This has been researched, and greedy sequences constructed this way (with any kernel) are called *Leja Points*. López-García and Wagner have a wealth of results on the 1-dimensional circle alone.

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary.

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian, satisfying

$$-\Delta_x \int_M G(x, y) f(y) dy = f(x).$$

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian, satisfying

$$-\Delta_x \int_M G(x, y) f(y) dy = f(x).$$

Green's kernel behaves similarly to Riesz kernels

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian, satisfying

$$-\Delta_x \int_M G(x, y) f(y) dy = f(x).$$

Green's kernel behaves similarly to Riesz kernels: in dimension $d = 2$, it is comparable to k_0, and for $d \geq 3$ is on the order of k_{d-2}.

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian, satisfying

$$-\Delta_x \int_M G(x, y) f(y) dy = f(x).$$

Green's kernel behaves similarly to Riesz kernels: in dimension $d = 2$, it is comparable to k_0, and for $d \geq 3$ is on the order of k_{d-2}. Unlike Riesz, Green's kernel is intrinsic and does not depend on the embedding of the manifold.

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian, satisfying

$$-\Delta_x \int_M G(x,y) f(y) dy = f(x).$$

Green's kernel behaves similarly to Riesz kernels: in dimension $d = 2$, it is comparable to k_0, and for $d \geq 3$ is on the order of k_{d-2}. Unlike Riesz, Green's kernel is intrinsic and does not depend on the embedding of the manifold. We now define our sequence greedily as

$$x_n = \arg\min_{x \in M} \sum_{k=1}^{n-1} G(x, x_k).$$

Higher Dimensions

Let (M, g) be a smooth compact manifold without boundary. Let $G(x, y)$ be the Green's function of the Laplacian, satisfying

$$-\Delta_x \int_M G(x,y)f(y)dy = f(x).$$

Green's kernel behaves similarly to Riesz kernels: in dimension $d = 2$, it is comparable to k_0, and for $d \geq 3$ is on the order of k_{d-2}. Unlike Riesz, Green's kernel is intrinsic and does not depend on the embedding of the manifold. We now define our sequence greedily as

$$x_n = \arg\min_{x \in M} \sum_{k=1}^{n-1} G(x, x_k).$$

The scaling of the proof is fundamentally different in higher dimensions, and yields stronger bounds!

Wasserstein Distance

The Wasserstein distance, also known as the Earth Mover's Distance, measures how much work it takes to "carry" the mass of one distribution to another.

Wasserstein Distance

The Wasserstein distance, also known as the Earth Mover's Distance, measures how much work it takes to "carry" the mass of one distribution to another.

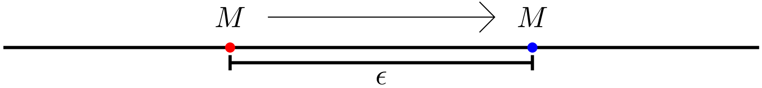


Figure: Transporting a point mass of weight M by distance ϵ incurs a cost of $M \cdot \epsilon$.

Wasserstein Distance

The Wasserstein distance, also known as the Earth Mover's Distance, measures how much work it takes to "carry" the mass of one distribution to another.

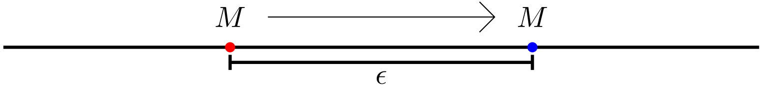
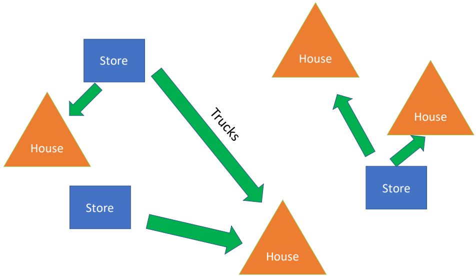


Figure: Transporting a point mass of weight M by distance ϵ incurs a cost of $M \cdot \epsilon$.

Figure: The Wasserstein distance between the blue and red point distributions which each have 2 point masses of weight $1/2$ is $\frac{1}{2}(.1) + \frac{1}{2}(.5) = .3$

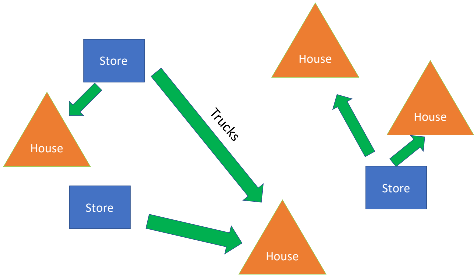
Wasserstein Distance

Figure: Transporting between a point distribution on stores and houses



Wasserstein Distance

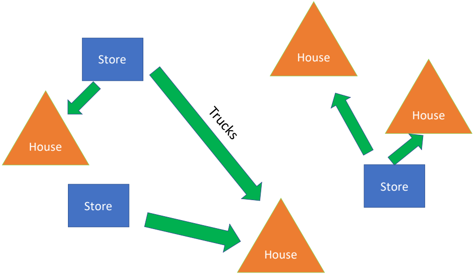
Figure: Transporting between a point distribution on stores and houses



In our setting, every point on the unit interval has a “house”

Wasserstein Distance

Figure: Transporting between a point distribution on stores and houses



In our setting, every point on the unit interval has a "house": We are interested in measuring the Wasserstein distance between the point measure μ_N from our sequence and the uniform distribution dx.

Wasserstein Distance

Here's the van der Corput sequence mapped to $[0,1]^2$ by $(\frac{i}{100}, x_i)$:

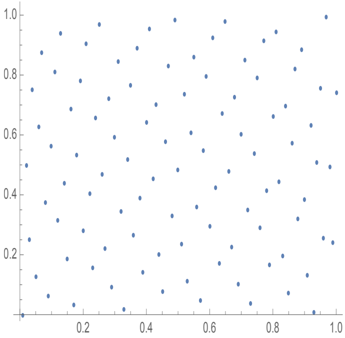


Figure: The first 100 terms of the van der Corput sequence

Wasserstein Distance

Here's the van der Corput sequence mapped to $[0,1]^2$ by $(\frac{i}{100}, x_i)$:

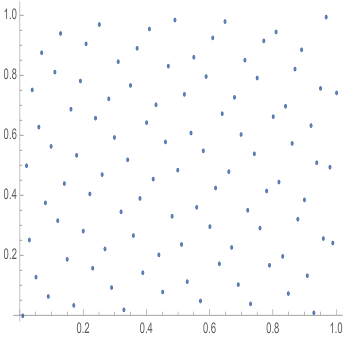


Figure: The first 100 terms of the van der Corput sequence

We can imagine taking the Wasserstein distance between the (normalized) sum of Dirac measures on the points and the uniform distribution on the unit square

Wasserstein Distance

Here's the van der Corput sequence mapped to $[0,1]^2$ by $(\frac{i}{100}, x_i)$:

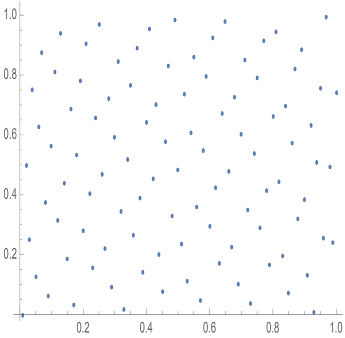


Figure: The first 100 terms of the van der Corput sequence

We can imagine taking the Wasserstein distance between the (normalized) sum of Dirac measures on the points and the uniform distribution on the unit square: we would need to "smudge" each point to transport its mass continuously over nearby points, and W_1 measures how much smudging we need.

Wasserstein Distance

Formally, the $p-$Wasserstein distance between two measures μ and ν is defined as

$$W_p(\mu,\nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{M \times M} |x-y|^p d\gamma(x,y) \right)^{1/p}$$

Wasserstein Distance

Formally, the $p-$Wasserstein distance between two measures μ and ν is defined as

$$W_p(\mu,\nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{M \times M} |x-y|^p d\gamma(x,y) \right)^{1/p},$$

where $|\cdot|$ is the metric and $\Gamma(\mu,\nu)$ denotes the collection of all measures on $M \times M$ with marginals μ and ν (also called the set of all couplings of μ and ν).

Wasserstein Distance

Formally, the $p-$Wasserstein distance between two measures μ and ν is defined as

$$W_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{M \times M} |x - y|^p d\gamma(x, y) \right)^{1/p},$$

where $|\cdot|$ is the metric and $\Gamma(\mu, \nu)$ denotes the collection of all measures on $M \times M$ with marginals μ and ν (also called the set of all couplings of μ and ν).
In 1-d, we have

$$W_2(\mu, dx) \lesssim \|\mu\|_{\dot{H}^{-1}},$$

so W_2 seems like a good generalization of diaphony to higher dimensions.

Wasserstein Distance

For any d-dimensional manifold M, there is a constant $c > 0$ such that, for any set of N points $\{x_1, \ldots, x_N\}$ on M, we have

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq cN^{-1/d}.$$

Wasserstein Distance

For any d-dimensional manifold M, there is a constant $c > 0$ such that, for any set of N points $\{x_1, \dots, x_N\}$ on M, we have

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq cN^{-1/d}.$$

The proof is a basic observation about the scaling in d dimensions

Wasserstein Distance

For any d-dimensional manifold M, there is a constant $c > 0$ such that, for any set of N points $\{x_1, \ldots, x_N\}$ on M, we have

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq cN^{-1/d}.$$

The proof is a basic observation about the scaling in d dimensions: If we place balls of radius $r = \epsilon N^{-1/d}$ around each x_i, then the total volume of the balls is at most

$$N(\omega_d r^d) = N(\omega_d \epsilon^d N^{-1}) = \omega_d \epsilon^d.$$

Wasserstein Distance

For any d-dimensional manifold M, there is a constant $c > 0$ such that, for any set of N points $\{x_1, \ldots, x_N\}$ on M, we have

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq cN^{-1/d}.$$

The proof is a basic observation about the scaling in d dimensions: If we place balls of radius $r = \epsilon N^{-1/d}$ around each x_i, then the total volume of the balls is at most

$$N(\omega_d r^d) = N(\omega_d \epsilon^d N^{-1}) = \omega_d \epsilon^d.$$

We may pick ϵ small enough that this quantity is less than half the volume of M.

Wasserstein Distance

For any d-dimensional manifold M, there is a constant $c > 0$ such that, for any set of N points $\{x_1, \ldots, x_N\}$ on M, we have

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq cN^{-1/d}.$$

The proof is a basic observation about the scaling in d dimensions: If we place balls of radius $r = \epsilon N^{-1/d}$ around each x_i, then the total volume of the balls is at most

$$N(\omega_d r^d) = N(\omega_d \epsilon^d N^{-1}) = \omega_d \epsilon^d.$$

We may pick ϵ small enough that this quantity is less than half the volume of M. Thus, we will need to transport most of the probability mass a distance of more than $\epsilon N^{-1/d}$

Wasserstein Distance

For any d-dimensional manifold M, there is a constant $c > 0$ such that, for any set of N points $\{x_1, \ldots, x_N\}$ on M, we have

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq cN^{-1/d}.$$

The proof is a basic observation about the scaling in d dimensions: If we place balls of radius $r = \epsilon N^{-1/d}$ around each x_i, then the total volume of the balls is at most

$$N(\omega_d r^d) = N(\omega_d \epsilon^d N^{-1}) = \omega_d \epsilon^d.$$

We may pick ϵ small enough that this quantity is less than half the volume of M. Thus, we will need to transport most of the probability mass a distance of more than $\epsilon N^{-1/d}$, so

$$W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, dx\right) \geq \epsilon N^{-1/d}/2.$$

Wasserstein Distance

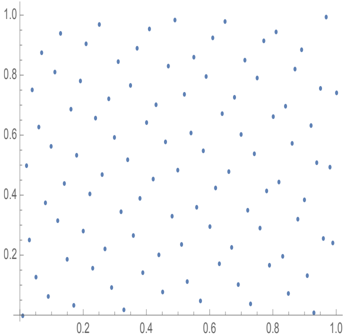


Figure: The first 100 terms of the van der Corput sequence

We may imagine that each point on the plot is actually a very small disc, and we see that the vast majority of the area is outside these discs

Wasserstein Distance

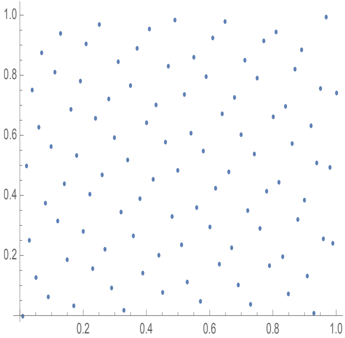


Figure: The first 100 terms of the van der Corput sequence

We may imagine that each point on the plot is actually a very small disc, and we see that the vast majority of the area is outside these discs; thus, we would need to carry most of the point mass by (much) more than the radius of the discs.

Wasserstein Distance

Theorem (B & Steinerberger '20)

Let x_n be a sequence obtained on a $d-$dimensional compact manifold, by starting with an arbitrary set $\{x_1, \dots, x_m\}$ and greedily setting

$$x_n = \arg\min_{x \in M} \sum_{k=1}^{n-1} G(x, x_k).$$

Then

$$W_2\left(\frac{1}{n}\sum_{k=1}^{n} \delta_{x_k}, dx\right) \lesssim_M \begin{cases} n^{-1/2}\sqrt{\log n} & \text{if } d = 2 \\ n^{-1/d} & \text{if } d \geq 3. \end{cases}$$

Wasserstein Distance

Theorem (B & Steinerberger '20)

Let x_n be a sequence obtained on a $d-$dimensional compact manifold, by starting with an arbitrary set $\{x_1, \dots, x_m\}$ and greedily setting

$$x_n = \arg\min_{x \in M} \sum_{k=1}^{n-1} G(x, x_k).$$

Then

$$W_2\left(\frac{1}{n}\sum_{k=1}^{n} \delta_{x_k}, dx\right) \lesssim_M \begin{cases} n^{-1/2}\sqrt{\log n} & \text{if } d = 2 \\ n^{-1/d} & \text{if } d \geq 3. \end{cases}$$

This result is optimal in $d \geq 3$, but nobody knows what the best discrepancy is (or if this implies that these sequences obtain it)!

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...
- Maybe such sequences are also optimal in higher dimensions?

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...
- Maybe such sequences are also optimal in higher dimensions?
- Wasserstein ✓

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...
- Maybe such sequences are also optimal in higher dimensions?
- Wasserstein ✓
- Discrepancy ??

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...
- Maybe such sequences are also optimal in higher dimensions?
- Wasserstein ✓
- Discrepancy ??
- Numerically challenging to compute in high dimensions.

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...
- Maybe such sequences are also optimal in higher dimensions?
- Wasserstein ✓
- Discrepancy ??
- Numerically challenging to compute in high dimensions.
- Nice connections to potential theory (Green's function)?

Open Questions

This is a completely new type of sequence, with lots of mysteries even just in $[0, 1]$.

- Seems numerically to be optimal...
- Maybe such sequences are also optimal in higher dimensions?
- Wasserstein ✓
- Discrepancy ??
- Numerically challenging to compute in high dimensions.
- Nice connections to potential theory (Green's function)?
- Other types of functions that work?

Thank you!