MFCS

Ordered Fields

Klaus Sutner
Carnegie Mellon University
Fall 2022

More Numbers

You are all familiar with the standard number systems used in calculus:

$$\mathbb{Q} \qquad \mathbb{R} \qquad \mathbb{C}$$

The difference between these and the integers is that one can perform division and solve many more equations. In fact, according to the Fundamental Theorem of Algebra, all non-constant, single-variable polynomial equations with integer coefficients have a solution over the complex numbers[†].

What are the key properties of these number systems?

Can we axiomatize these structures?

[†] This is a hard result, even Gauss made a mistake in his 1799 proof.

Axiomatization

What would be the ingredients for our putative axioms? Minimally, we would need

- two arithmetic operations plus and times,
- two constants zero and one,
- possibly an order relation.

And then we have to write down the basic laws that these objects obey—and hope that the basic rules already capture everything there this to know. Take "everything" with a grain of salt, we won't get completeness of the reals, for example.

Definition

A field is an algebraic structure $\mathbb{F} = \langle F; +, *, 0, 1 \rangle$ where the following axioms hold: associativity, commutativity, neutral elements, inverses, distributivity.

$$x + (y + z) = (x + y) + z \qquad x * (y * z) = (x * y) * z$$

$$x + y = y + x \qquad x * y = y * x$$

$$x + 0 = x \qquad x * 1 = x$$

$$\forall\, x\, \exists_1\, y\, (x + y = 0) \qquad \forall\, x \neq 0\, \exists_1\, y\, (x * y = 1)$$

$$x * (y + z) = (x * y) + (x * z)$$

0 Inverse?

One also insists that $0 \neq 1$, so any field has at least two elements.

The additive and multiplicative inverses are usually written $-x$ and x^{-1} or $1/x$ in fractional notation. Note the guard $x \neq 0$ for multiplicative inverses.

There is no way around this. For suppose we have some inverse 0^{-1}. Then

$$1 = 0 * 0^{-1} = (0 + 0) * 0^{-1} = 0 * 0^{-1} + 0 * 0^{-1} = 1 + 1$$

whence $0 = 1$ and we have a contradiction.

It is easy to check (white lie) that the usual suspects are all fields:

$$\mathbb{Q} \qquad \mathbb{R} \qquad \mathbb{C}$$

Are There Finite Fields?

There are: the modular numbers $\mathbb{Z}_p$ form a field iff p is a prime.
In particular, there is a two-element field

$$\mathbb{Z}_2 = \langle \mathbf{2}; \oplus, \wedge, 0, 1 \rangle$$

where the operations are "exclusive or" and "and"[†].

We can compute a multiplicative inverse in $\mathbb{Z}_p$ using the extended Euclidean algorithm. There are also finite fields of size p^k for all $k \geq 1$, but these are much, much harder to describe.

Finite fields are hugely important in coding theory, cryptography, the theory of algorithms and complexity theory.

[†]So in this case algebra comes down to logic.

Perhaps the most important application of the reals is calculus and real analysis. For this purpose, having a field is not quite enough[†].

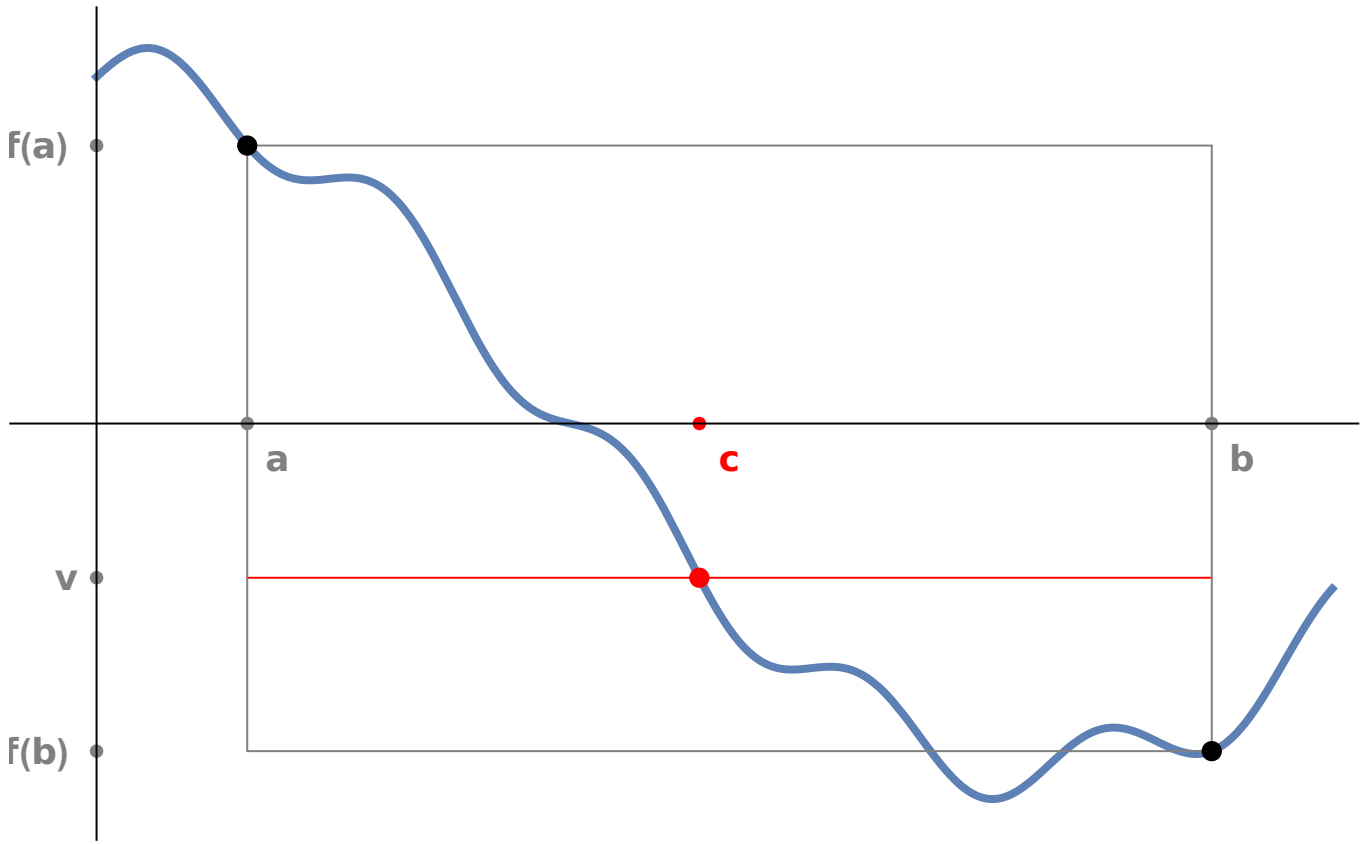
For example, we define continuity like so:

$$\forall\, \varepsilon > 0\, \exists\, \delta > 0\, \forall\, x\, \big(|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon\big)$$

Theorem (Intermediate Value Theorem, Cauchy 1821)

If the function $f(x)$ is continuous with respect to the variable x between the limits $x = x_0$ and $x = X$ and if b designates a quantity between $f(x_0)$ and $f(X)$, one can always satisfy the equation $f(x) = b$ for one or several real values of x between x_0 and X.

[†]Completeness is the central issue, a fact we will ignore here.


f(a)
a
c
b
v
f(b)

Order and Intervals

Cauchy's result directly requires order, we really need a way to define intervals:

$$[a, b] = \{\, x \in \mathbb{R} \mid a \leq x \leq b \,\}$$

With these examples in mind, we would like to add two more concepts to the idea of a field:

- order
- absolute value

Here are the requisite definitions.

Good Orders

Recall from modular arithmetic that Gaussian congruences are interesting not simply because they are equivalence relation of finite index, but because they coexist peacefully with arithmetic.

The same applies here: we do not want some random order imposed on the reals, we want an order that is compatible the arithmetic in the field.

For our purposes, we want a strict total order that has the following two properties:

$$(O1) \qquad x < y \Rightarrow x + z < y + z$$

$$(O2) \qquad 0 < x, y \Rightarrow 0 < x * y$$

It is clear that the usual order on $\mathbb{Q}$ and $\mathbb{R}$ satisfies these properties, we are just axiomatizing matters.

Ordered Fields

Definition

An ordered field is a field $\mathbb{F}$ together with a strict order $<$ that satisfies axioms $(O1)$ and $(O2)$.

The rationals and the reals are uniquely ordered.

We will prove in a moment that neither finite fields nor the complex numbers can be ordered.

There are some strange fields that can be ordered in different ways, but we won't worry about those[†].

[†] Take a course in field theory if you are interested

Positivity

Note that

$$x < y \iff 0 < y - x$$

by $(O1)$, so the order can also be expressed in terms of positive elements.

A set $P \subseteq \mathbb{F}$ is a positive set if it satisfies the following conditions:

$$(P1) \qquad x, y \in P \Rightarrow x + y \in P$$

$$(P2) \qquad x, y \in P \Rightarrow x * y \in P$$

$$(P3) \qquad x \in P \vee x = 0 \vee -x \in P$$

The last condition is meant to be exclusive or: any element is either positive, zero or negative (x is negative if $-x$ is positive).

Back and Forth

So given an order $<$ we can define a positive set by

$$P_< = \{\, x \in \mathbb{F} \mid 0 < x \,\}$$

Conversely, given a positive set P, we can define an order relation by

$$x <_P y \Leftrightarrow y - x \in P$$

One can check that this really works as advertised. And, the two operations are mutual inverses.

Some Facts

Claim 1: In any ordered field, $0 < 1$.

Proof.

Assume otherwise. Since $0 \neq 1$ and $<$ is a total order, we must have $1 < 0$. By $(O1)$, $0 < -1$ and, by $(O2)$, $0 < (-1)(-1) = 1$, a contradiction. □

Claim 2: In any ordered field, $0 \neq x \Rightarrow 0 < x^2$.

Proof.

Since $<$ is a total order and $x \neq 0$ we must have $0 < x$ or $x < 0$. In the first case we are done by $(O2)$. In the second case, by $(O1)$, $0 < -x$, and, by $(O2)$, $0 < (-x)(-x) = x^2$. □

Finite No Good

The first claim implies that no finite field can be ordered.

First note that in any finite field, we must have

$$\underbrace{1+1+\ldots+1}_{k}=0$$

for some k. The minimal such k is called the characteristic of the field.

But $0<1$, whence $1<1+1$, $1+1<1+1+1$ and so on by $(O1)$. By transitivity, $0<1+1+\ldots+1=0$, a contradiction.

Squares

The complex numbers are not an ordered field: there is no ordering of $\mathbb{C}$ that is compatible with addition and multiplication.

By the second claim, non-zero squares are positive in any ordered field. But in $\mathbb{C}$ we have $i^2 = -1 < 0$, a contradiction.

Also note that in the reals we could actually define order in terms of squares via the positive set

$$P = \{\, x \in \mathbb{R} \mid \exists\, z \neq 0\, (z^2 = x) \,\}$$

since every positive real has a square root.

This clearly fails for the rationals.

Absolute Value

We can now define the absolute value or magnitude of a field element in any ordered field:

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{otherwise.} \end{cases}$$

Thus $0 \leq |x|$ with equality only for $x = 0$.

Note that $z \leq |z|$ with equality only for non-negative z.

Multiplication is well-behaved: $|xy| = |x||y|$.

Lastly, $z^2 = |z|^2$.

Warning: Complex Numbers

There is a very useful notion of an absolute value, also called modulus, of a complex number:

$$|a + ib| = \sqrt{a^2 + b^2}$$

Geometrically, this measures the distance of the number from the origin.

But: this kind of absolute value is not based on a field order.

Triangle Inequality

Proposition

$|x + y| \leq |x| + |y|$

Proof.

We have

$$(x + y)^2 = x^2 + 2xy + y^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2$$

Hence $|x + y|^2 \leq ||x| + |y||^2$ and taking square roots yields $|x + y| \leq ||x| + |y|| = |x| + |y|$.

Chasing equality conditions we can see that equality holds in the proposition exactly when x and y have the same sign or at least one of them is 0.

□

Completeness

For any set $A \subseteq \mathbb{R}$, an upper bound for A is a real b such that

$$\forall\, x \in A\, (x \leq b)$$

Write $\mathrm{UB}(z)$ to mean: z is an upper bound for A. Then b is a least upper bound for A if

$$\mathrm{UB}(b) \wedge \forall\, z\, \big(\mathrm{UB}(z) \Rightarrow b \leq z\big)$$

Completeness of the reals means: every subset of the reals that has an upper bound also has a least upper bound.

This is the reason why calculus is based on the reals, not the rationals.

Definition

Let $\mathbb{F}$ be an arbitrary field. A vector space over $\mathbb{F}$ is an algebraic structure $\langle V; \oplus, \cdot, \mathbf{0} \rangle$ where the two operations have the format

- vector addition: $\oplus : V \times V \to V$
- scalar multiplication: $\cdot : \mathbb{F} \times V \to V$

and the properties listed below.

The vector addition is associative, commutative, has a neutral element $\mathbf{0}$ and unique inverses.

The scalar multiplication is subject to the conditions

- $a \cdot (x \oplus y) = (a \cdot x) \oplus (a \cdot y)$
- $(a + b) \cdot x = (a \cdot x) \oplus (b \cdot x)$
- $(a * b) \cdot x = a \cdot (b \cdot x)$
- $1 \cdot x = x$

Sloppy Notation

We have written vector addition as $\oplus$ to distinguish it from the addition in the field. No one does this in the RealWorld$^{\text{TM}}$, one simply writes $+$ for both. And 0 for $\mathbf{0}$, and one drops the parens and replaces the multiplication operators by concatenation.

- $a(x + y) = ax + ay$
- $(a + b)x = ax + bx$
- $(ab)x = a(bx)$
- $1x = x$

In the long run, this is better for one's sanity, trust me.

Cheap Example

Let $\mathbb{F}$ be any field, finite or infinite.

Then $\langle \mathbb{F}; +, *, 0 \rangle$ is a vector space over $\mathbb{F}$.

In other words, we let $\oplus = +$, $\cdot = *$ and $\mathbf{0} = 0$.

Then all the vector space axioms hold.

OK, this is pretty lame, but hold on.

The Mother of All Vector Spaces

Let $\mathbb{F}$ be any field, finite or infinite.

Generalizing the last example slightly, consider $\mathbb{F}^n$, the collection of all sequences over $\mathbb{F}$ of length n.

In this context, these sequences are called n-dimensional vectors.

$\mathbb{F}^n$ is a vector space over $\mathbb{F}$ using componentwise operations:

$$\begin{aligned} \boldsymbol{u} \oplus \boldsymbol{v} &= (u_i + v_i) \\ a \cdot \boldsymbol{v} &= (a v_i) \\ \mathbf{0} &= (0, 0, \ldots, 0) \end{aligned}$$

Exercise

Check that $\mathbb{F}^n$ really is a vector space over $\mathbb{F}$.

Longer Vectors

Example (Product)

$$\prod_I \mathbb{F} = \{\, \boldsymbol{x} \mid \boldsymbol{x} : I \to \mathbb{F} \,\}$$

is a vector space over $\mathbb{F}$ for any infinite index set I.

The point here is that in a standard, finite-dimensional vector space we use the perfectly natural index set $I = [n]$.

But, all the definitions work just as well if we replace $[n]$ by any infinite set I, no matter how big. For example, $\mathbb{F} = I = \mathbb{R}$ makes perfect sense, and is hugely useful: real functions form a vector space over $\mathbb{R}$.

This is a good sign, our development seems to be quite robust.

Important Variant

Example (Coproduct)

$$\coprod_I \mathbb{F} = \{\, \boldsymbol{x} : I \to \mathbb{F} \mid \boldsymbol{x} \text{ finite support} \,\}$$

is a vector space over $\mathbb{F}$ for any infinite index set I.

Finite support means that $x_i \neq 0$ for only finitely many $i \in I$ (a property that is preserved under the vector space operations).

As a consequence, every vector in a coproduct vectors space has a finite description (as long as the elements of $\mathbb{F}$ do).

Alas, we will stick with finite-dimensional Euclidean spaces $\mathbb{R}^n$.

Measuring Length

A norm on a vector space is a map $\|.\| : V \to \mathbb{R}$ with the following properties:

$$\begin{aligned}
&\|\boldsymbol{x}\| \geq 0 \\
&\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \mathbf{0} \\
&\|a\boldsymbol{x}\| = |a|\|\boldsymbol{x}\| \\
&\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|
\end{aligned}$$

Thus, we require the triangle inequality to hold for norms (for absolute value, it follows from the usual axioms as we have seen).

A norm provides a measure of distance between points in a vector space:

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$$

So the question is: can we find a norm for $\mathbb{R}^n$?

Euclidean Norm

Taking inspiration from dimension 2, we could try

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

The first three properties of a norm are trivially satisfied, but the triangle inequality requires a bit of work.

We cannot simply use the squaring trick from the absolute value case since we are now dealing with vectors, not scalars.

That naturally leads to the question of whether there might be some kind of reasonable product on vectors in $\mathbb{R}^n$ that could be exploited in a proof.

Inner Product

It is not at all clear how we might multiply vectors. Here is the axiomatic approach to the properties we would want from a product $\circ : V \times V \to \mathbb{R}$ (we will only deal with real spaces[†]).

symmetry	$\boldsymbol{x} \circ \boldsymbol{y} = \boldsymbol{y} \circ \boldsymbol{x}$
linearity	$(a\boldsymbol{x} + b\boldsymbol{x}') \circ \boldsymbol{y} = a(\boldsymbol{x} \circ \boldsymbol{y}) + b(\boldsymbol{x}' \circ \boldsymbol{y})$
positive definiteness	$\boldsymbol{x} \neq \boldsymbol{0} \Rightarrow \boldsymbol{x} \circ \boldsymbol{x} > 0$

Note that the combination of symmetry and linearity in the first argument also produces linearity in the second argument:

$$\boldsymbol{x} \circ (a\boldsymbol{y} + b\boldsymbol{y}') = a(\boldsymbol{x} \circ \boldsymbol{y}) + b(\boldsymbol{x} \circ \boldsymbol{y}')$$

[†]Other popular notation is $\boldsymbol{x} \cdot \boldsymbol{y}$ or $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

OK, but how do we concoct such a product?

A key observation is that, by linearity, it suffices to figure out what to do with unit vectors:

$$\boldsymbol{x} \circ \boldsymbol{y} = \sum_{i,j} x_i y_j \, (\boldsymbol{e}_i \circ \boldsymbol{e}_j)$$

Since we require positive definiteness, a fair guess would be

$$\boldsymbol{e}_i \circ \boldsymbol{e}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

This attempt produces

$$\boldsymbol{x} \circ \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

and we can easily check that we really have an inner product.

Having found a function that satisfies the axioms is nice, but not really enough: it should make some intuitive sense beyond just having the right formal properties. In our case, we are in luck.

Leaning on trigonometry, one can check that

$$\boldsymbol{x} \circ \boldsymbol{y} = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \cos\theta$$

where θ is the angle subtended by the two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

In particular, $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal iff $\boldsymbol{x} \circ \boldsymbol{y} = 0$. The following classical result follows immediately from this characterization, but we will give a proof avoiding trigonometry.

Cauchy-Schwartz Inequality

Theorem

$|\boldsymbol{x} \circ \boldsymbol{y}| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$

Proof. We may assume $\boldsymbol{y} \neq \boldsymbol{0}$ and define $a = (\boldsymbol{x} \circ \boldsymbol{y})\|\boldsymbol{y}\|^{-2}$.

$$\begin{aligned} 0 &\leq \|\boldsymbol{x} - a\boldsymbol{y}\|^2 \\ &= (\boldsymbol{x} \circ \boldsymbol{x}) - 2a(\boldsymbol{x} \circ \boldsymbol{y}) + a^2(\boldsymbol{y} \circ \boldsymbol{y}) \\ &= \|\boldsymbol{x}\|^2 - (\boldsymbol{x} \circ \boldsymbol{y})^2\|\boldsymbol{y}\|^{-2} \end{aligned}$$

Both steps are justified by unfolding the definition of our norm, and for the second step we use the definition of a. □

Norm Triangle Inequality

If you are still breathing, we can now tackle the triangle inequality for our Euclidean norm.

Proposition

$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

Proof.

$$\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|^2 &= (\boldsymbol{x} \circ \boldsymbol{x}) + 2(\boldsymbol{x} \circ \boldsymbol{y}) + (\boldsymbol{y} \circ \boldsymbol{y}) \\
&\leq \|\boldsymbol{x}\|^2 + 2|\boldsymbol{x} \circ \boldsymbol{y}| + \|\boldsymbol{y}\|^2 \\
&\leq \|\boldsymbol{x}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{y}\|^2 \\
&= (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2
\end{aligned}$$

where the third step is justified by Cauchy-Schwartz. □

Generalizing

One might wonder whether our Euclidean norm is the only choice for $\mathbb{R}^n$ or whether there are other, geometrically meaningful ways to measure length and distance. Since

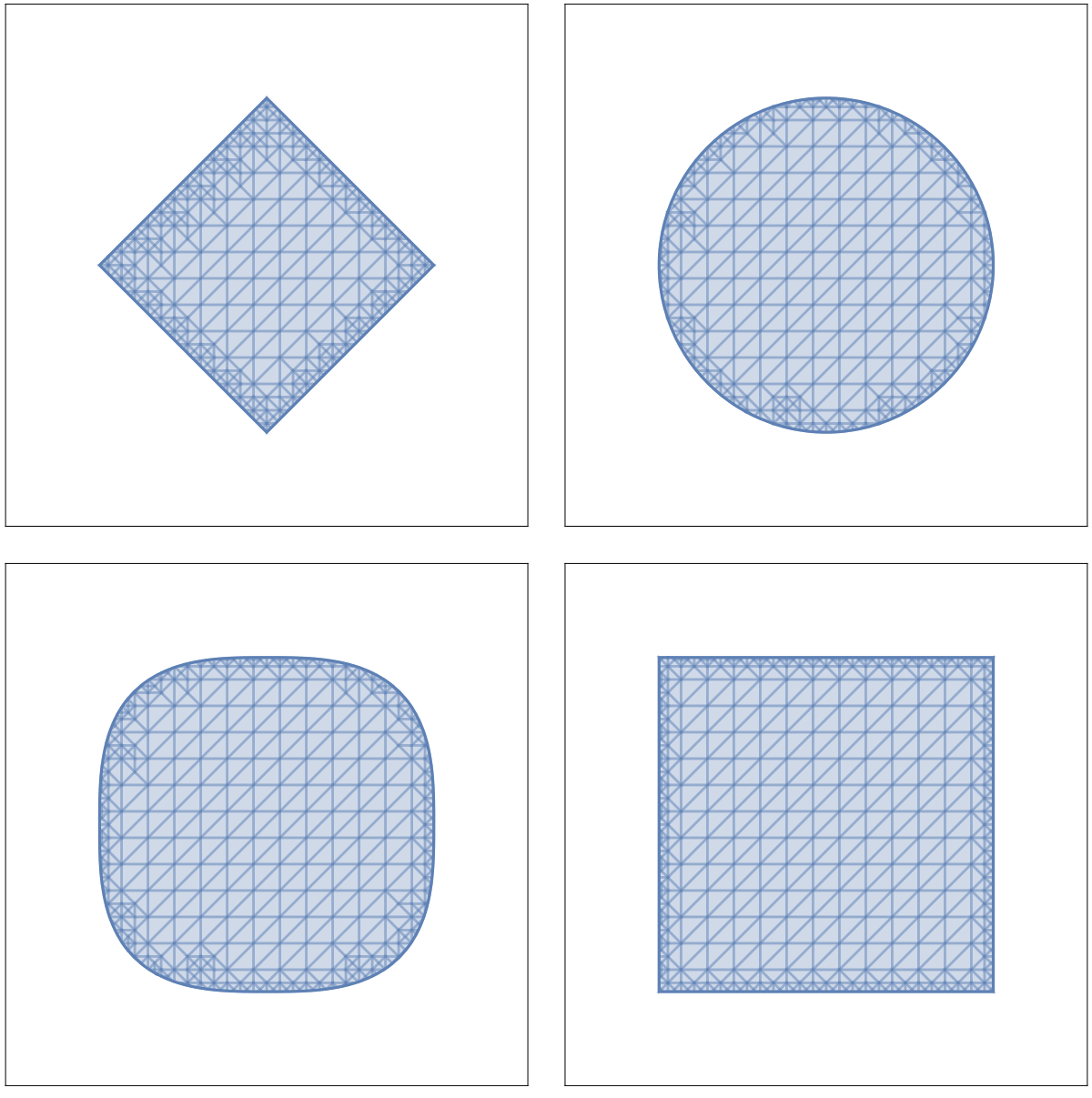
$$\|\boldsymbol{x}\| = \left(\sum x_i^2\right)^{1/2}$$

one just might conjecture that we could use so-called p-norms

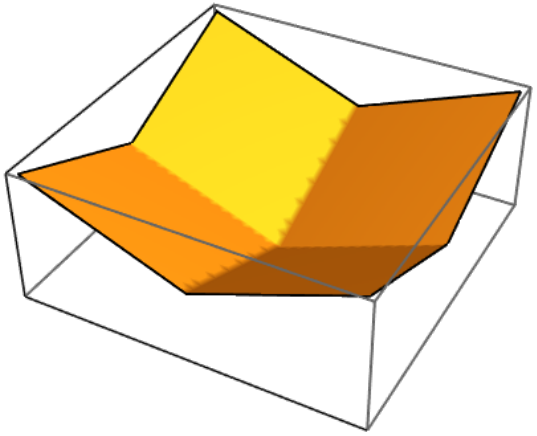
$$\|\boldsymbol{x}\|_p = \left(\sum |x_i|^p\right)^{1/p}$$

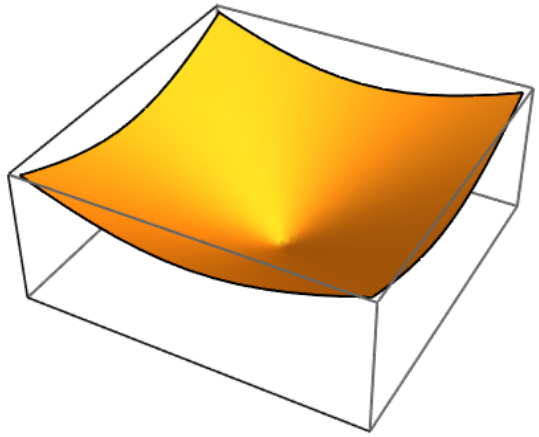
Note that we have replaced x_i by $|x_i|$ to ensure positivity.

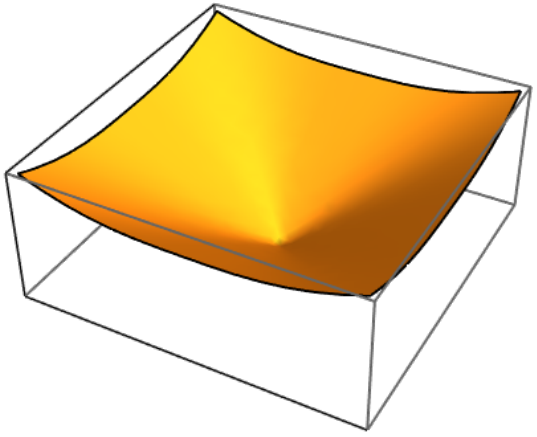
Unit Disks $p = 1, 2, 3, \infty$

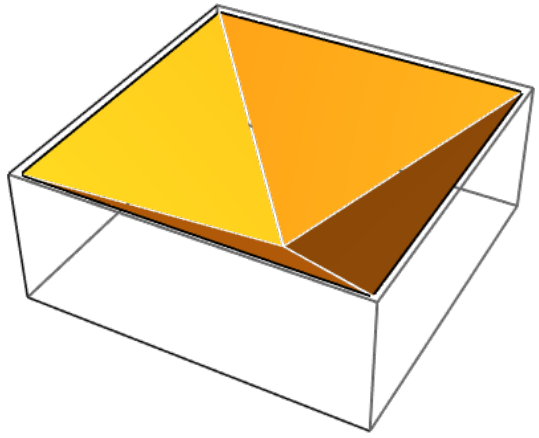


Norms $p = 1, 2, 3, \infty$







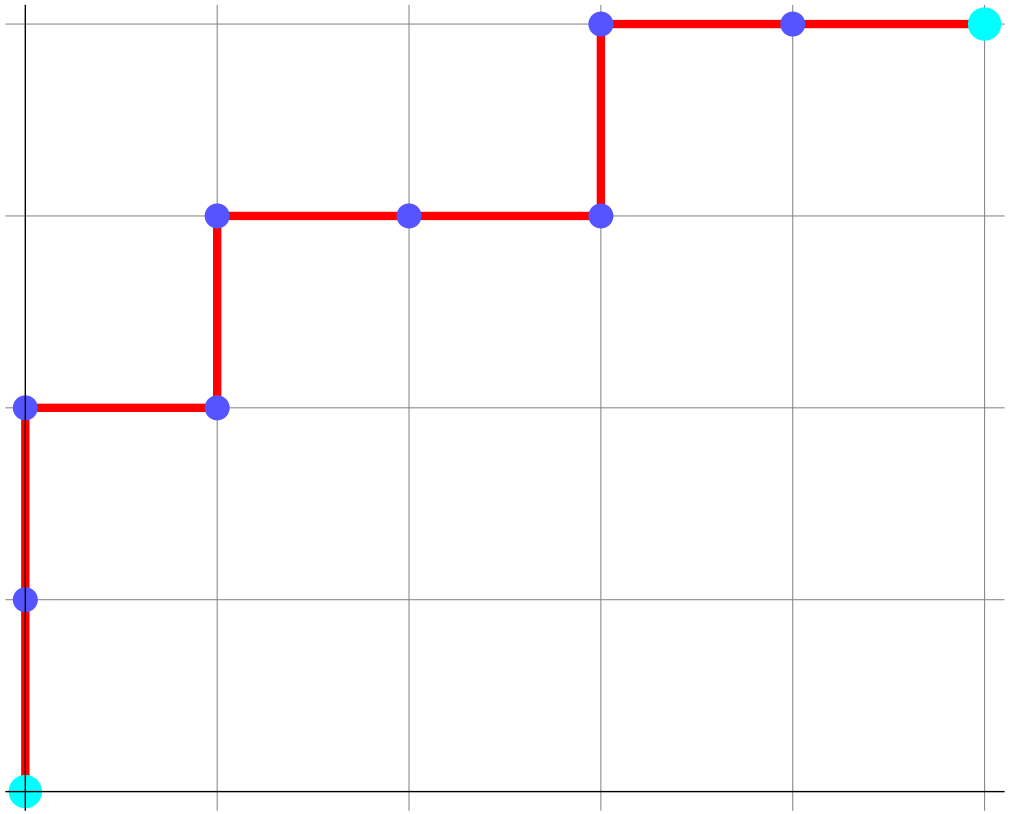


1-Norm

This is also called the Manhattan norm:

$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \ldots + |x_n|$$

It properly describes distance on a rectangular grid.



Chebyshev Norm

At the other end of the spectrum we have the infinity norm or Chebyshev norm or maximum norm:

$$\|\boldsymbol{x}\|_\infty = \max\big(|x_1|, |x_2|, \ldots, |x_n|\big)$$

One can show that is really is the limit of p-norms as p approaches infinity.

In terms of a grid walk we would have to add the ability to move diagonally from one corner of a block to the other (much like a king on a chess board)

Traveling Salesperson Problem (TSP)

Different norms appear for example in computational geometry and complexity. Here is one famous combinatorial problem that has directly to do with distances.

Suppose we have an $n \times n$ matrix D of non-negative numbers. We think of $D(i, j)$ as the cost of traveling from location i to location j.

We are interested in cheap closed tours, where by a tour we mean a permutation π of $[n]$ such that $\pi(n) = 1$: we are supposed to travel

$$1, \pi(1).\pi(2), \ldots, \pi(n-1), \pi(n) = 1$$

Thus, we start at 1, hit every location exactly once, and return to 1.

Cost

The cost of such a tour is the sum of all individual steps:

$$\text{cost}(\pi) = \sum_{i<n} D(\pi(i), \pi(i{+}1)) \; + D(\pi(n), \pi(1))$$

The goal of TSP is find a tour of minimal cost.

In a sense, this is trivial: we can simply try out all permutations and find a cheapest one. Alas, in applications n can be several hundred, so the search would never end in this universe.

Getting Serious

Since this is an algorithmic problem, one needs to be a bit more careful about specifying the input. We said D is a matrix of non-negative numbers, but that is much too vague: we should insist that the numbers are rational.

In fact, by scaling properly we can assume the distances are in $\mathbb{N}$.

Here is a particularly simple case: the metric or triangle TSP:

- D is symmetric, and
- D obeys the triangle inequality: $D(i,j) \leq D(i,k) + D(k,j)$.

In other words, D behaves somewhat like a table of geometric distances.

Bad News

Unfortunately, the Triangle TSP is not much better than the general version: no polynomial time algorithm (ie., polynomial in n) is known.

In fact, things already spin out of control even when all distances just are 1 or 2. The problem does not require complicated distances to become unmanageable.

Maybe things become manageable if we constrain things to actual geometric distances?

Euclidean TSP

Say, we assume n points p_i in the Euclidean plane, and we define $D(i,j) = \|p_i - p_j\|$, the ordinary Euclidean distance.

Hold it, hold it. "Points in the plane" sounds good, but we need rational coordinates. In fact, by scaling, we may as well assume that the points are all in $\mathbb{Z} \times \mathbb{Z}$.

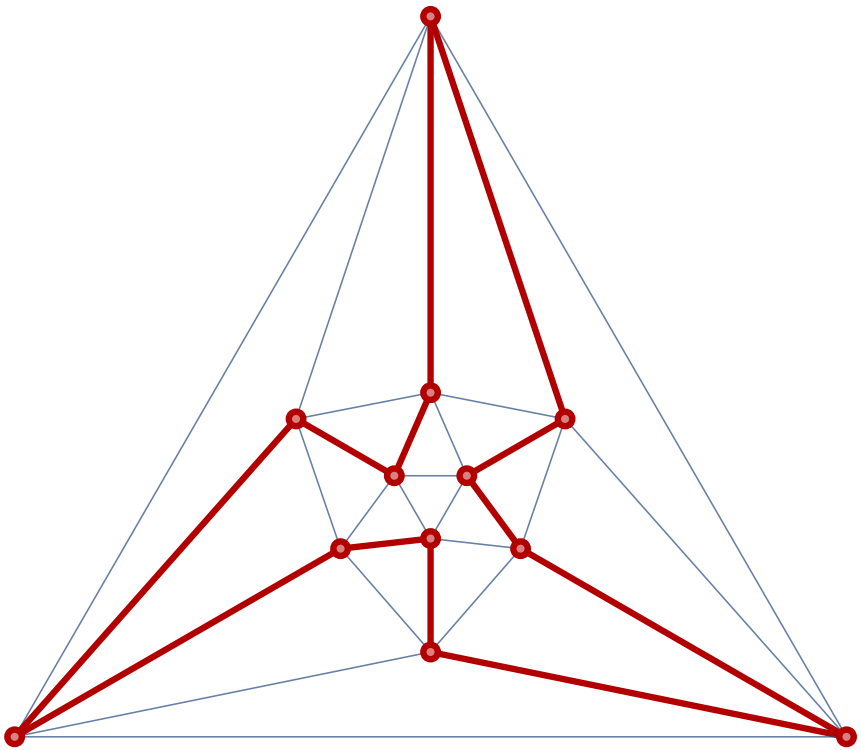
Much better, but there still is a problem: in order to find the cheapest tour, one has to compare expressions of the form

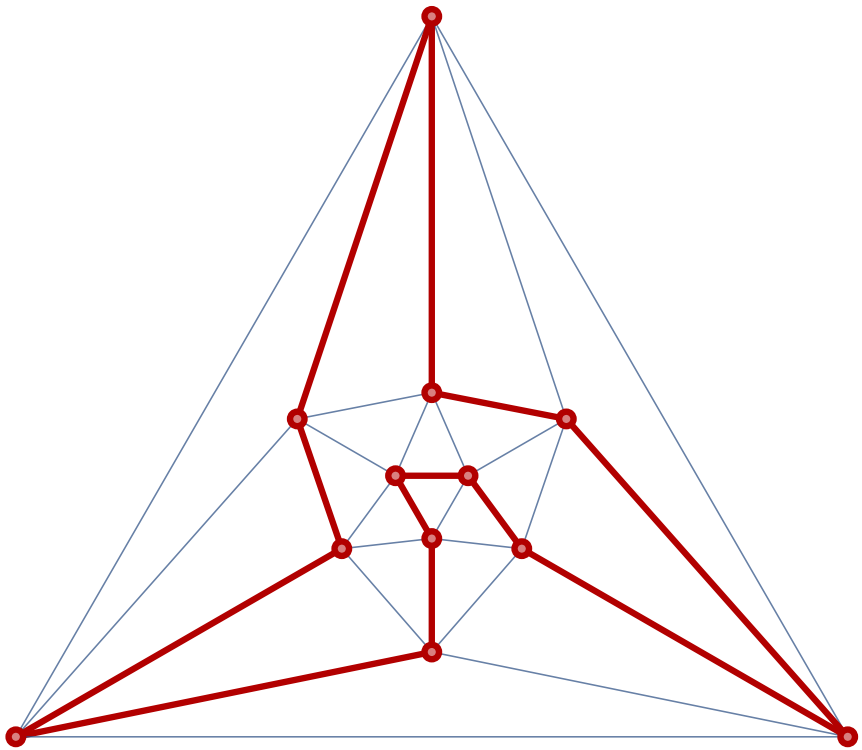
$$\sqrt{a_1} + \sqrt{a_2} + \ldots + \sqrt{a_k}$$

for integral a_i to compare tours or parts thereof.

This may look trivial, but it is not: the roots are irrational in general, and it is not known how many digits are needed.

Example





Comparing

The last two icosahedron tours are both pretty good. To figure out which is better, one needs to determine the sign of

$$544 + 5\sqrt{11833} - \sqrt{461401} - \sqrt{462113} + \\ - 2\sqrt{724645} + \sqrt{1056890} + \sqrt{1058285}$$

The numerical value is about 83.1, so the first, more symmetric tour is better. Of course, this assumes that the software I used is sufficiently accurate.

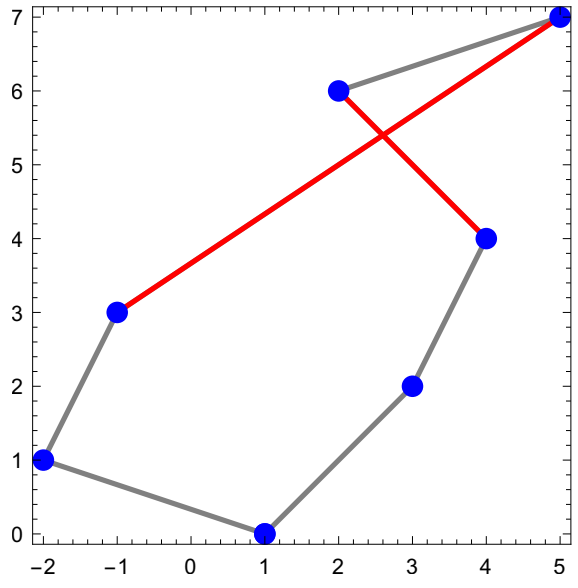
In actual applications such as routing an Amazon truck, the situation is even worse: one needs to find not just a cheapest tour, but there are additional constraints.

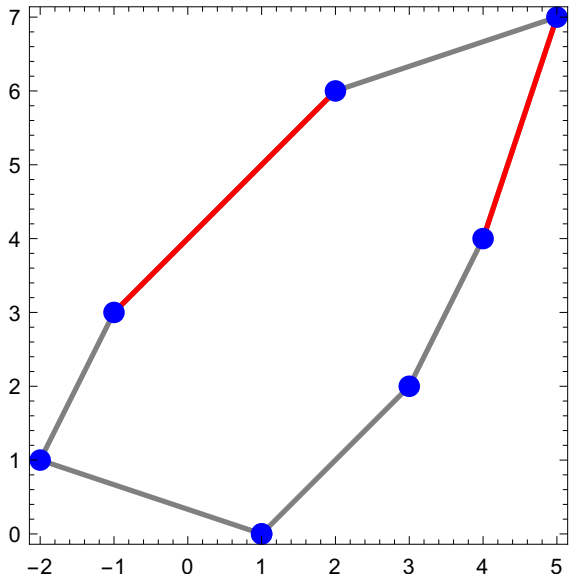
For example, one would like the truck to take about the same number of left and right turns: otherwise the wear on the tires is uneven, and poor Bezos might miss out on another billion or two. Just think about how you would even model this additional constraint.

Reasonably good approximation algorithms exist, but that's about it.

Nearest Neighbor

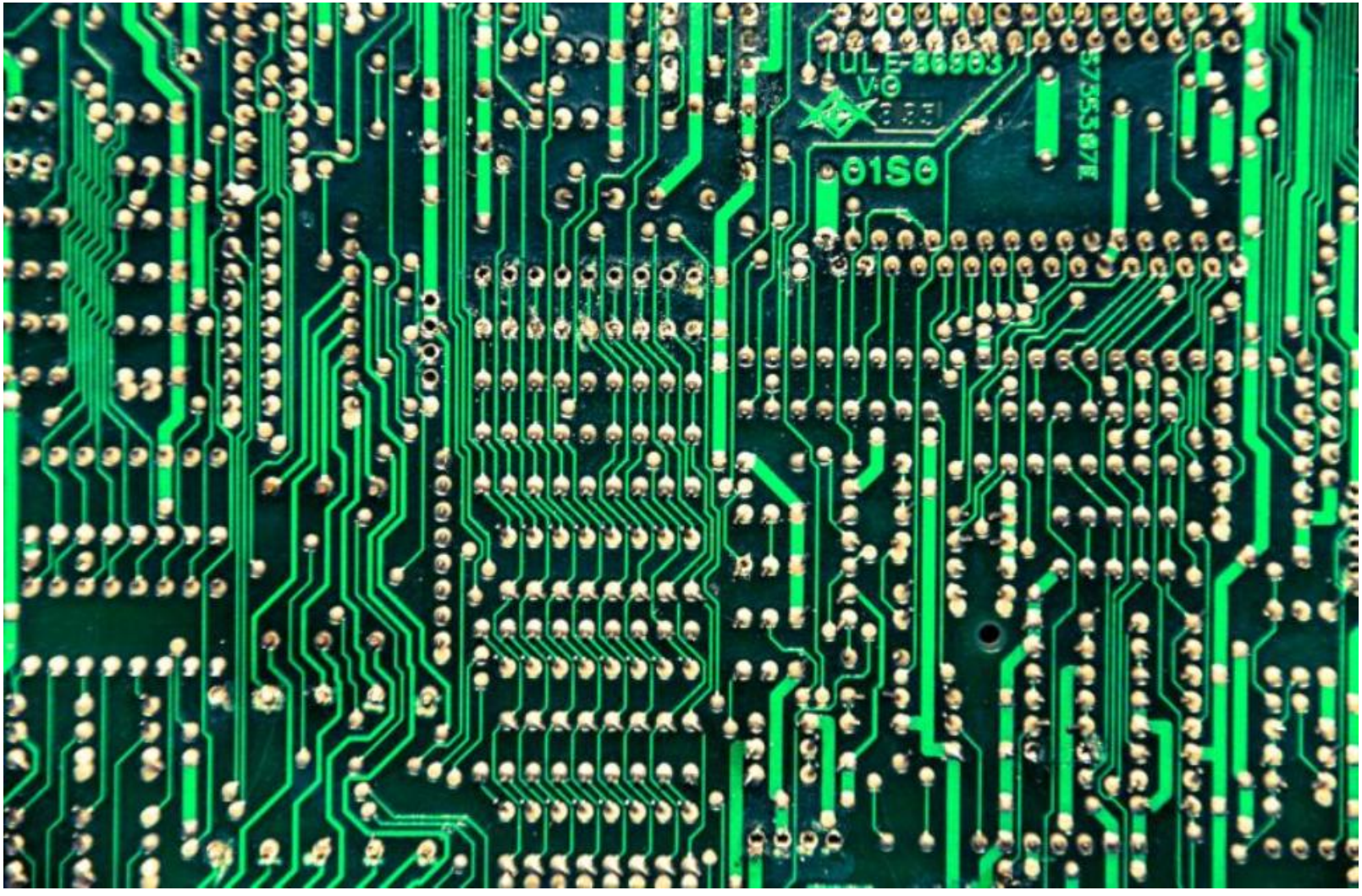
Perhaps the most tempting strategy for Euclidean TSP is to work strictly locally: start in some random place, then always go to the nearest untouched neighbor.





The example shows that this can go wrong, but the error here can be fixed with a little post-processing.

Circuit Boards



1-Norm

The drill that produces the holes in the board typically only moves along the axes, not diagonally.

So to model this situation we need to use the Manhattan norm.

Not to worry, this is just as difficult.