Programming Research Group

EPISTEMIC MOTION IN QUANTUM SEARCHING

Keye Martin

PRG-RR-03-06

Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford OX1 3QD

Abstract

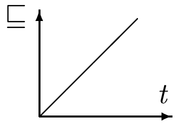
We formalize elementary kinematics of motion on domains with measurements, and then use these ideas to calculate the complexity of Grover's algorithm for quantum searching. In particular, it is now possible to analyze classical and quantum searching in a single semantic framework.

Introduction

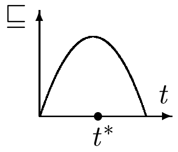
This paper is about motion that occurs not in ordinary space, but in spaces of informatic objects called *domains*. For example, if $f : D \rightarrow D$ is a Scott continuous selfmap on a domain with least element $\perp$, the iteration which produces its least fixed point

$$\mathrm{fix}(f) := \bigsqcup_{n \geq 0} f^n(\perp)$$

implicitly defines a discrete trajectory $x : \mathbb{N} \rightarrow D$ given by $x(n) := f^n(\perp)$. As we will see with classical searching, sometimes a trajectory like x has a unique extension to *all* times $t \in [0, \infty)$. If we graph x in the $t-\sqsubseteq$ plane, it looks something like this:



Even if we replace f above by the kind of nonmonotonic mappings discussed in [12], we still get the same picture. This is the basic difficulty which prevented the author from including this idea as part of the initial study of measurement [11]: No examples of nontrivial motion on domains. For instance, nothing like this:



In the context of algorithms, this might seem like an absurd picture: An algorithm which achieves the exact answer after t^* units of time but is guaranteed *in theory* to completely lose the answer after an additional t^* units. More is less, genuine nonmonotonicity. For example, a numerical algorithm is not like this: These are only problematic in practice; in theory, the more you do, the closer to the answer you get. It is counterintuitive that an algorithm like this should exist. And yet, one does one exist, and this absurd aspect is something we must take into account when analyzing it. It is Grover's algorithm for quantum searching.

We begin this paper with the elementary kinematics of motion on domains: We study trajectories $x : [0,\infty) \to D$, their time derivatives with respect to a measurement, qualitative maxima and minima for trajectories, and how to find them using the derivative. The domain $D = [0,\infty)^*$ yields the usual kinematics of straight line motion. Grover's algorithm causes an *epistemic trajectory,* a trajectory on the domain of classical states in their Bayesian order [6]. In order to understand when a trajectory moves with respect to this order and especially when it is that domain theoretic kinematics applies, we benefit greatly from new results involving the *Wronski bracket,* a Lie algebra closely tied to the Bayesian order.

In the last section, we calculate the complexity of Grover's algorithm for quantum searching by finding the point in time where the derivative of its epistemic trajectory is zero. It is a matter of theory developed within that this is also the point where its entropy is a minimum: A *domain theoretic maximum* simultaneously minimizes all variables (measurements), so one can minimize the entropy without using it in our calculations. We obtain a qualitative explanation for the complexity of quantum searching in a semantic framework where its classical counterpart is already known to reside naturally. This raises a few questions that we mention at the end.

1 Kinematics on a domain

A *domain* is a dcpo with a notion of approximation. We remain intentionally vague about approximation until the idea is explicitly needed. A variable lies in between a measurement and a map that measures the entire domain.

Definition 1.1 A *variable* on a domain is a measurement $v : D \to [0,\infty)^*$ such that for all $x, y \in D$, we have $x \sqsubseteq y \ \& \ vx = vy \Rightarrow x = y$.

A variable v measures a significant quality of the objects in D, such as uncertainty (or information content).

Definition 1.2 A *curve* on a domain D is a function $x : \mathrm{dom}(x) \to D$ where $\mathrm{dom}(x)$ is a nontrivial interval of the real line.

Each curve x determines a value of v at time t, which is the number $vx(t)$.

Definition 1.3 For a curve x and variable v on a domain,

$$\dot{x}_v(t) := \lim_{s \to t} \frac{vx(s) - vx(t)}{s - t}.$$

We then define

$$\ddot{x}_v := \frac{d\dot{x}_v}{dt}$$

and so on for higher order.

Because $\dot{x}_v : [0, \infty) \to \mathbb{R}$ is an ordinary function, higher order derivatives are calculated as usual – its the first derivative that requires theory.

Proposition 1.4 *Let x be a curve with $\dot{x}_v$ defined on (a, b).*

(i) *x is monotone increasing on $[a, b]$ iff $\dot{x}_v \leq 0$ on (a, b) and $x[a, b]$ is a chain.*

(ii) *x is monotone decreasing on $[a, b]$ iff $\dot{x}_v \geq 0$ on (a, b) and $x[a, b]$ is a chain.*

(iii) *x is constant on $[a, b]$ iff $\dot{x}_v = 0$ on (a, b) and $x[a, b]$ is a chain.*

Proof. (i) If x is monotone increasing on $[a, b]$, then $x[a, b]$ is a chain, and the sign of $vx(s) - vx(t)$ is opposite the sign of $s - t$, for $s, t \in (a, b)$, which means $\dot{x}_v \leq 0$ on (a, b). For the converse, we apply the mean value theorem to vx: Given $s < t$ with $s, t \in [a, b]$, there is $c \in (a, b)$ with

$$vx(s) - vx(t) = \dot{x}_v(c)(s - t) \geq 0$$

which means $vx(s) \geq vx(t)$, and since $x[a, b]$ is a chain, $x(s) \sqsubseteq x(t)$, by the strict monotonicity of v. The proof of (ii) is similar, and (iii) follows from (i) and (ii). $\Box$

Notice that the sign of $\dot{x}_v$ is an indicator of how *uncertainty* behaves: If $\dot{x}_v \leq 0$, then uncertainty is decreasing, so we are moving up in the order.

Definition 1.5 A curve x has a *relative maximum* at an interior point $t \in \mathrm{dom}(x)$ if there is an open set U_t containing t such that $x(s) \sqsubseteq x(t)$ for all $s \in U_t$. *Relative minimum* is defined dually, and these two give rise to *relative extremum.*

Notice that a *qualitative* relative maximum is a point in time where the *quantitative* uncertainty is a local minimum.

Lemma 1.6 *If a curve x has a relative extremum at interior point $t \in \mathrm{dom}(x)$, then for all variables v, either $\dot{x}_v(t) = 0$, or it does not exist.*

Proof. If $x(t)$ is a relative maximum on $U_t = (t - \varepsilon, t + \varepsilon)$ for some $\varepsilon > 0$, then $vx(s) - vx(t) \geq 0$ for all $s \in U_t$. Thus, if $\dot{x}_v(t)$ exists, then

$$\dot{x}_v(t) = \lim_{s \to t^+} \frac{vx(s) - vx(t)}{s - t} \geq 0 \ \text{ and } \ \dot{x}_v(t) = \lim_{s \to t^-} \frac{vx(s) - vx(t)}{s - t} \leq 0,$$

which finishes the proof. $\Box$

A nice illustration of why the qualitative idea $\sqsubseteq$ is important: If a curve has a derivative with respect to *just one* variable v, then its set of extreme points is contained in the set $\{t : \dot{x}_v(t) = 0\}$. This is quite valuable: We are free to choose the variable which makes the calculation as simple as possible.

Once we have the extreme points there is also a systematic way in the informatic setting to determine which (if any) are maxima or minima: The second derivative test, whose formalization requires one to acknowledge the qualitative structure on which it is implicitly founded.

Definition 1.7 A curve x is a *trajectory* if for all $t \in \mathrm{dom}(x)$ there is an open set U_t containing t such that

$$x(s) \sqsubseteq x(t) \text{ or } x(t) \sqsubseteq x(s)$$

for all $s \in U_t$.

Thus, a *trajectory* is a curve x with underlying qualitative structure; it is called C_v^2 when $\ddot{x}_v$ is continuous, with respect to variable v.

Proposition 1.8 *Let x be a C_v^2 trajectory. If $\dot{x}_v(t) = 0$ and $\ddot{x}_v(t) \neq 0$ for some interior point $t \in \mathrm{dom}(x)$, then x has a relative extremum at t.*

(i) *If $\ddot{x}_v(t) > 0$, then $x(t)$ is a relative maximum.*

(ii) *If $\ddot{x}_v(t) < 0$, then $x(t)$ is a relative minimum.*

Proof. (i) By the continuity of $\ddot{x}_v$, there is an open set $V_t = (t - \varepsilon, t + \varepsilon)$ for $\varepsilon > 0$ on which $\ddot{x}_v > 0$. Then $\dot{x}_v$ is strictly increasing on V_t, and since $\dot{x}_v(t) = 0$, we know that

$$(\dot{x}_v(s) < 0 \text{ for } t - \varepsilon < s < t) \text{ and } (\dot{x}_v(s) > 0 \text{ for } t < s < t + \varepsilon).$$

Because x is a trajectory in the formal sense, there is $\delta > 0$ such that $x(s) \sqsubseteq x(t)$ or $x(t) \sqsubseteq x(s)$ for all $s \in (t - \delta, t + \delta)$. Let $\alpha = \min\{\delta, \varepsilon\} > 0$ and $U_t := (t - \alpha, t + \alpha)$. Then

$$(\forall s \in U_t)\ x(s) \sqsubseteq x(t)$$

which means $x(t)$ is a relative maximum. To see this, given $s \in U_t$, we apply the mean value theorem to vx,

$$vx(s) - vx(t) = \dot{x}_v(c)(s - t) \geq 0$$

and since v is a variable and x is a trajectory, we get $x(s) \sqsubseteq x(t)$. $\square$

In this work we will be mostly concerned with the strongest form of extrema on domains:

Definition 1.9 A curve x has an *absolute maximum* at $t \in \mathrm{dom}(x)$ if

$$x(s) \sqsubseteq x(t)$$

for all $s \in \mathrm{dom}(x)$. Absolute minimum is defined similarly.

Here is a simple but surprisingly useful way of establishing the existence of absolute extrema.

Proposition 1.10 *Let v be a variable on D and $x : [a,b] \to D$ a curve whose image is a chain. If $vx : [a,b] \to \mathbb{R}$ is Euclidean continuous, then*

(i) *The map x is continuous from the Euclidean to the Scott topology, and*

(ii) *The map x assumes an absolute maximum and an absolute minimum on $[a,b]$. In particular, its absolute maximum is*

$$x(t^*) = \bigsqcup_{t\in[a,b]} x(t)$$

for some $t^* \in [a,b]$, with a similar expression for the absolute minimum.

Proof. (i) Let $U \subseteq D$ be Scott open with $x(t) \in U$. If there is no open set around t on which x remains in U, we can find a sequence $t_n \to t$ with $x(t) \not\sqsubseteq x(t_n)$. Because $\mathrm{Im}(x)$ is a chain, this means $x(t_n) \sqsubseteq x(t)$. But vx is Euclidean continuous so we have

$$\bigsqcup_{n\geq 1} x(t_n) \sqsubseteq x(t) \text{ and } \lim_{n\to\infty} vx(t_n) = vx(t)$$

which means $\bigsqcup x(t_n) = x(t)$ and this puts some $x(t_n) \in U$ using the definition of Scott open – a contradiction.

(ii) Since vx is Euclidean continuous, it takes an absolute maximum $vx(t_*)$ and absolute minimum $vx(t^*)$. Because $\mathrm{Im}(x)$ is a chain and v is variable, its absolute maximum is $x(t^*)$ and absolute minimum is $x(t_*)$. $\Box$

A valuable property of absolute maxima: If $x(t^*)$ is an absolute maximum, then for all variables v,

$$vx(t^*) = \inf\{vx(t) : t \in \mathrm{dom}(x)\}.$$

That is, an absolute maximum is a point on a curve which simultaneously minimizes *all* variables.

Example 1.11 *Binary search.* For a continuous $f : \mathbb{R} \to \mathbb{R}$, let $\mathrm{split}_f : \mathbf{I}\mathbb{R} \to \mathbf{I}\mathbb{R}$ be the bisection on the interval domain defined by

$$\mathrm{split}_f[a,b] := \begin{cases} \mathrm{left}[a,b] & \text{if } f(a)\cdot f((a+b)/2) \leq 0; \\ \mathrm{right}[a,b] & \text{otherwise.} \end{cases}$$

A given $x \in \mathbf{I}\mathbb{R}$ leads to a *trajectory* $x : [0,\infty) \to \mathbf{I}\mathbb{R}$ defined on integers by

$$x(n) = \mathrm{split}_f^n(x)$$

and then extended to all intermediate times $n < t < n+1$ by declaring $x(t)$ to be the *unique* element satisfying

$$x(n) \sqsubseteq x(t) \sqsubseteq x(n+1) \ \text{ and } \ \mu x(t) = \frac{\mu x}{2^t}.$$

Then we have

$$\mu x(t) = e^{-(\ln 2)t} \cdot \mu x(0)$$

which means

$$\dot{x}_\mu(t) = (-\ln 2)\mu x(t) < 0$$

reflecting the fact that $x : [0, \infty) \to \mathbf{I}\mathbb{R}$ is monotone *increasing.*

In particular, the last example can be used to model binary search. To search a list $L : \{1, \ldots, n\} \to \mathbb{N}$ for some element k, we need the ability to find a zero of an unknown polynomial $p(x) = x - i$, where either $L(i) = k$ or $i = 0$. Binary search is the trajectory which results from iterating split_p beginning from $x = [0, n]$.

As we said in the introduction, algorithms represented by monotone maps, and even those represented by nonmonotonic splittings like the one above, do not provide examples of nontrivial motion. If we take $D = [0, \infty)^*$ with $vx = x$, we get interesting examples of motion (like the height of a projectile) that are not a priori algorithmic. The trajectories caused by the class of quantum searching algorithms provide examples of nontrivial motion on the domain of classical states.

2 Domains of classical states in one symmetry

For an integer $n \geq 2$, the *classical states* are

$$\Delta^n := \left\{ x \in [0,1]^n : \sum_{i=1}^{n} x_i = 1 \right\}.$$

A classical state $x \in \Delta^n$ is *pure* when $x_i = 1$ for some $i \in \{1, \ldots, n\}$; we denote such a state by e_i. Pure states $\{e_i\}_i$ are the actual states a system can be in, while general mixed states x and y are epistemic entities. If we know x and by some means determine that outcome i is not possible, our knowledge improves to

$$p_i(x) = \frac{1}{1 - x_i}(x_1, \ldots, \hat{x}_i, \ldots, x_{n+1}) \in \Delta^n,$$

where $p_i(x)$ is obtained by first removing x_i from x and then renormalizing. The partial mappings which result, $p_i : \Delta^{n+1} \rightharpoonup \Delta^n$ with $\mathrm{dom}(p_i) = \Delta^{n+1} \setminus \{e_i\}$, are called the *Bayesian projections* and lead one directly to the following relation on classical states.

Definition 2.1 For $x, y \in \Delta^{n+1}$,

$$x \sqsubseteq y \equiv (\forall i)(x, y \in \mathrm{dom}(p_i) \Rightarrow p_i(x) \sqsubseteq p_i(y)).$$

For $x, y \in \Delta^2$,

$$x \sqsubseteq y \equiv (y_1 \leq x_1 \leq 1/2) \text{ or } (1/2 \leq x_1 \leq y_1).$$

The relation $\sqsubseteq$ on Δ^n is called the *Bayesian order.*

The relation $\sqsubseteq$ was introduced in [6] where the following is proven:

Theorem 2.2 *$(\Delta^n, \sqsubseteq)$ is a domain with least element $\perp := (1/n, \ldots, 1/n)$ and*

$$\max(\Delta^n) = \{e_i : 1 \leq i \leq n\}.$$

It has $vx = 1 - x^+$ and $\sigma x = -\sum x_i \log x_i$ as variables where $x^+ = \max x_i$.

The Bayesian order has a more direct description: The *symmetric formulation* [6]. Let $S(n)$ denote the group of permutations on $\{1, \ldots, n\}$ and

$$\Lambda^n := \{x \in \Delta^n : (\forall i < n)\, x_i \geq x_{i+1}\}$$

denote the collection of *monotone* decreasing classical states.

Theorem 2.3 *For $x, y \in \Delta^n$, we have $x \sqsubseteq y$ iff there is a permutation $\sigma \in S(n)$ such that $x \cdot \sigma, y \cdot \sigma \in \Lambda^n$ and*

$$(x \cdot \sigma)_i (y \cdot \sigma)_{i+1} \leq (x \cdot \sigma)_{i+1} (y \cdot \sigma)_i$$

for all i with $1 \leq i < n$.

Thus, $(\Delta^n, \sqsubseteq)$ can be thought of as $n!$ many copies of the domain $(\Lambda^n, \sqsubseteq)$ identified along their common boundaries, where the Bayesian order on Λ^n is

$$x \sqsubseteq y \equiv (\forall i < n)\, x_i y_{i+1} \leq x_{i+1} y_i.$$

But a word of caution about such reasoning: Ordering Λ^n and Δ^n are essentially unrelated problems. The Bayesian order extends to n dimensional quantum mixed states Ω^n in such a way that the classical and quantum logics of Birkhoff and von Neumann [3] arise as the irreducible elements of $(\Delta^n, \sqsubseteq)$ and $(\Omega^n, \sqsubseteq)$, respectively. This result cannot be achieved with a single symmetry Λ^n.

The crucial paradigm shift in this work is that we stop thinking '$x \sqsubseteq y$' and start thinking '*why* $x \sqsubseteq y$?' We suppose that a possible cause is a process at least partially represented by a curve $\pi : [0, 1] \to \Lambda^n$ with $\pi(0) = x$ and $\pi(1) = y$. In many ways, π is nothing more than an elaboration, a more detailed account of the situation '$x \sqsubseteq y$'.

The first thing we are going to prove is that movement in the Bayesian order implies differentiability (a.e.). A curve $\pi : [0, 1] \to \Lambda^n$ can be written as $\pi = (\pi_1, \ldots, \pi_n)$ where the functions $\pi_i : [0, 1] \to [0, 1]$ satisfy $\pi_i \geq \pi_{i+1}$ for $i < n$ and

$$\sum_{i=1}^{n} \pi_i(t) = 1$$

for all $t \in [0, 1]$.

Lemma 2.4 *If $x \sqsubseteq y$ in Λ^n, then*

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i$$

for all $1 \leq k \leq n$.

Proof. Because $x \sqsubseteq y$ in Λ^n, there is an integer $1 < m \leq n$ such that $x_i \leq y_i$ for $i < m$ and $x_i \geq y_i$ for $i \geq m$. Then for $k < m$ the claim is obvious, while for $k \geq m$ we have

$$\sum_{i=1}^{k} x_i = 1 - \sum_{i=k+1}^{n} x_i \leq 1 - \sum_{i=k+1}^{n} y_i = \sum_{i=1}^{k} y_i,$$

using $x_i \geq y_i$ for $i \geq m$. $\square$

The property of the Bayesian order expressed by Lemma 2.4 defines an order $\leq$ on Λ^n in its own right,

$$x \leq y \equiv (\forall k \in \{1, \ldots, n\}) \sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i$$

called *Majorization* [2][9][13]. It too yields a domain $(\Lambda^n, \leq)$ as noted in [10]; however, it has no natural extension to all of Δ^n. Despite this sharp difference, a monotone curve in either order has a decent amount of analytic structure.

Proposition 2.5 *If $\pi = (\pi_1, \ldots, \pi_n) : [0,1] \to (\Lambda^n, \leq)$ is a monotone curve, then each π_i is of bounded variation. Thus,*

(i) *The curve π has a 'length.'*

(ii) *It is continuous except on a countable set.*

(iii) *It is differentiable except on a set of measure zero.*

Proof. Define $f_k : [0,1] \to [0,1]$ by

$$f_k(t) = \sum_{i=1}^{k} \pi_i(t)$$

for $k \geq 0$. By the definition of $\leq$, the f_k are monotone increasing functions, where we note that $f_k = 0$ for $k = 0$. Since $\pi_i = f_i - f_{i-1}$, the map π_i is the difference of monotone increasing maps and thus of bounded variation. This establishes (i), (ii) and (iii). $\square$

From the qualitative *follows* differentiability. We study differentiable curves.

3 The Wronski bracket

Let X denote the set of differentiable real valued functions defined on $[0,1]$. For $x, y \in X$,

$$[x,y] := \dot{x}y - \dot{y}x.$$

This is called the *Wronski bracket.*

Lemma 3.1 *Let $x, y, z \in X$. Then*

(i) $y[x,z] = z[x,y] + x[y,z]$,

(ii) $\dot{y}[x,z] = \dot{z}[x,y] + \dot{x}[y,z]$,

(iii) $y\dot{[x,z]} = z\dot{[x,y]} + x\dot{[y,z]}$.

In addition, $[x,1] = \dot{x}$.

Proof. (iii) Use the identity $\dot{[x,y]} = \ddot{x}y - \ddot{y}x$. □

The perfection of form noted in the last lemma suggests that X with $[\,,\,]$ has more structure than the usual Lie algebra. Nevertheless:

Lemma 3.2 $[\,,\,]$ *is a Lie bracket.*

(i) $[\,,\,]$ *is bilinear,*

(ii) $[x,y] = -[y,x]$,

(iii) *The Jacobi identity*

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$$

holds.

Thus, the smooth maps in X with $[\cdot,\cdot]$ form a Lie algebra.

Proof. (iii) What is beautiful here is that the Jacobi identity follows from (ii) and (iii) of Lemma 3.1. □

In the next proof, most details center on resolving the cases of boundary points: This is far simpler if the functions have *continuous* derivatives; then we can just take limits.

Proposition 3.3 *Let $f, g, h : [0,1] \to [0,\infty)$ be differentiable maps with $f \geq g \geq h$.*

(i) *If $[f,g] \geq 0$ and $[g,h] \geq 0$, then $[f,h] \geq 0$.*

(ii) *If $[f,g] = 0$ and $[g,h] = 0$, then $[f,h] = 0$.*

Proof. (i) Let $s \in [0,1]$. If $g(s) > 0$, then the equation

$$g[f,h] = h[f,g] + f[g,h]$$

of Lemma 3.1(i) applies, which immediately yields $[f,h](s) \geq 0$. Thus, we need only verify the assertion for $g(s) = 0$.

But if $g(s) = 0$, then we must also have $h(s) = 0$. For $s \in (0,1)$, we have $\dot{h}(s) = 0$, which immediately gives $[f,h](s) = 0$. Likewise, $f(s) = 0$ also gives $[f,h](s) = 0$, so we may assume that (i) $s \in \{0,1\}$, (ii) $f(s) > 0$, and (iii) $g(s) = h(s) = 0$. If $s = 1$, then

$$\dot{h}(s) = \lim_{t \to 1^-} \frac{h(t)}{t-1} \leq 0,$$

so $[f,h](s) = -\dot{h}(s)f(s) \geq 0$. If $s = 0$, then $\dot{h}(s) \geq 0$. Since $g(s) = 0$, $[f,g](s) = -\dot{g}(s)f(s) \geq 0$, so $f(s) > 0$ yields $\dot{g}(s) \leq 0$. But $g \geq h$ gives

$$\dot{h}(s) = \lim_{t \to 0^+} \frac{h(t)}{t} \leq \lim_{t \to 0^+} \frac{g(t)}{t} = \dot{g}(s) \leq 0,$$

and hence $\dot{h}(s) = 0$. Then $[f,h](s) = 0$, which finishes the proof.

(ii) Applying (i) to $f \geq g \geq h$ shows $[f,h] \geq 0$. But we can also apply (i) to $f \geq g \geq g - h$ since $[g, g-h] = 0$, which gives

$$[f, g-h] \geq 0$$

and that is exactly $[f,h] \leq 0$. $\square$

The *Wronski bracket* $[x,y] = \dot{x}y - \dot{y}x$ determines when two solutions of a second order equation are independent, the area swept out by a curve in the plane, angular momentum in mechanics, its sign classifies the direction of motion as being either clockwise or counterclockwise, and Boris Zilber [16] has communicated to the author that it arises in model theory as the only known example of a 'stable' infinite dimensional Lie algebra. It is also intimately connected to the Bayesian order.

Theorem 3.4 *A differentiable curve* $\pi = (\pi_1, \ldots, \pi_n) : [0,1] \to (\Lambda^n, \sqsubseteq)$ *is increasing iff* $[\pi_i, \pi_{i+1}] \geq 0$ *for all* $i < n$.

Proof. (i) $\Rightarrow$ (ii): To prove $[\pi_i, \pi_{i+1}](s) \geq 0$, we can assume $\pi_i(s) > 0$ (otherwise, we trivially have $[\pi_i, \pi_{i+1}](s) = 0$). By the continuity of π_i, there must be an open interval $(a,b) \subseteq \mathbb{R}$ containing s such that $\pi_i > 0$ on $U := (a,b) \cap [0,1]$. By the monotonicity of π, the map π_{i+1}/π_i is monotone decreasing on U. Since it is also differentiable, its derivative cannot be positive. But

$$\dot{\left(\frac{\pi_{i+1}}{\pi_i}\right)} = \frac{[\pi_{i+1}, \pi_i]}{\pi_i^2} \leq 0,$$

which makes it clear that $[\pi_i, \pi_{i+1}] = -[\pi_{i+1}, \pi_i] \geq 0$.

(ii) $\Rightarrow$ (i): We need to show that

$$\pi_i(s)\pi_{i+1}(t) \leq \pi_{i+1}(s)\pi_i(t)$$

whenever $s \leq t$ and $1 \leq i < n$. We can assume $s < t$. In the simple case, $\pi_i > 0$ on $[s,t]$, and then as before

$$\dot{\left(\frac{\pi_{i+1}}{\pi_i}\right)} = \frac{[\pi_{i+1}, \pi_i]}{\pi_i^2} = -\frac{[\pi_i, \pi_{i+1}]}{\pi_i^2} \leq 0$$

which means that π_{i+1}/π_i is monotone decreasing on $[s,t]$. In particular, $(\pi_{i+1}/\pi_i)(s) \geq (\pi_{i+1}/\pi_i)(t)$, which finishes this case.

Suppose now that $\pi_i(c) = 0$ for some $c \in [s,t]$. We claim that $\pi_i(t) = 0$. For the proof, a simple induction based on Prop. 3.3 gives $[\pi_i, \pi_j] \geq 0$ for $1 \leq i \leq j \leq n$. This means that $f_k : [0,1] \to [0,1]$ given by

$$f_k = \sum_{i=1}^{k} \pi_i$$

is monotone increasing since

$$\dot{f_k} = [f_k, 1] = \left[f_k, f_k + \sum_{j>k} \pi_j\right] = 0 + \sum_{j>k}[f_k, \pi_j] = \sum_{j>k}\sum_{i=1}^{k}[\pi_i, \pi_j] \geq 0.$$

Because $\pi(c)$ is a monotone state, $\pi_i(c) = 0$ implies that $f_i(c) = 1$. But if $\pi_i(t) > 0$, then the monotonicity of f_i gives $1 = f_i(c) < f_i(t)$, which contradicts the fact that $\pi(t)$ is a classical state. Then $\pi_i(t) = 0$, so $\pi_i \geq \pi_{i+1}$ gives $\pi_{i+1}(t) = 0$, which proves the desired inequality holds. $\square$

The previously mentioned uses of the Wronski bracket provide many new ways to explain the Bayesian order. We can also characterize it analytic terms.

Corollary 3.5 *For $x, y \in \Lambda^n$, $x \sqsubseteq y$ iff there is a differentiable curve $\pi : [0,1] \to \Lambda^n$ from $x = \pi(0)$ to $y = \pi(1)$ with $[\pi_i, \pi_{i+1}] \geq 0$ for all $i < n$.*

Proof. Given $x \sqsubseteq y$, let $\pi : [0,1] \to \Lambda^n$ be the straight line path from x to y. The other direction follows from the last theorem. $\square$

An interesting point is that for a differentiable curve $\pi : [0,1] \to \Lambda^n$, we have $[\pi_i, \pi_j]$ constant for all $1 \leq i, j \leq n$ iff $\pi(t) = (1-t) \cdot \pi(0) + t \cdot \pi(1)$. Finally, all results in this section apply equally well to curves $\pi : [a,b] \to \Lambda^n$. We simply took $a = 0$ and $b = 1$ because we like them.

4 Epistemic motion

We call the elements of a domain D *epistemic* only when it is necessary to remember that they represent knowledge. For example, the elements of Λ^n are epistemic; the elements of $\mathbf{I}\mathbb{R}$ are intervals. The Wronski bracket provides a way of determining when a curve moves in the Bayesian order.

Theorem 4.1 *Let $\pi : [a,b] \to \Lambda^n$ be a differentiable curve. If there is an integer $m \in \{1,\ldots,n\}$ with $m < n$ such that*

(i) *For all $i \neq m$, $[\pi_i, \pi_{i+1}] = 0$, and*

(ii) *For each $i \geq 1$, there is a constant d_{m+i} with*

$$[\pi_m, \pi_{m+i}] = d_{m+i} \cdot [\pi_m, \pi_{m+1}],$$

then the image of π is a chain in the Bayesian order. Thus, π has both an absolute maximum and an absolute minimum on $[a,b]$.

Proof. By Prop. 3.3(ii) and induction, $[\pi_i, \pi_j] = 0$ whenever $i, j \leq m$ or $m+1 \leq i, j$. First let $1 \leq k \leq m$, then from $\pi_1 > 0$ and

$$\dot{\left(\frac{\pi_k}{\pi_1}\right)} = \frac{[\pi_k, \pi_1]}{\pi_1^2} = 0,$$

there must be a constant $c_k \in [0,1]$ with $\pi_k = c_k \cdot \pi_1$. On the other hand, if $k \geq m+1$, then using

$$\pi_k = 1 - \sum_{i<k} \pi_i - \sum_{i>k} \pi_i$$

we have

$$\begin{aligned} [\pi_k, \pi_k] &= \left[\pi_k, 1 - \sum_{i<k} \pi_i - \sum_{i>k} \pi_i\right] \\ &= [\pi_k, 1] - \sum_{i<k} [\pi_k, \pi_i] - 0 \\ &= \dot{\pi}_k - \sum_{i \leq m} [\pi_k, \pi_i] \end{aligned}$$

and since $[\pi_k, \pi_k] = 0$,

$$\dot{\pi}_k = \sum_{i \leq m} [\pi_k, \pi_i] = c \cdot [\pi_k, \pi_1]$$

where $c := \sum_{i=1}^{m} c_i$ is a positive constant. Now we can write

$$c_m \cdot \dot{\pi}_k = c_m \cdot c \cdot [\pi_k, \pi_1] = c \cdot [\pi_k, \pi_m] = c \cdot d_k \cdot [\pi_{m+1}, \pi_m]$$

and in particular $c_m \cdot \dot{\pi}_{m+1} = c \cdot [\pi_{m+1}, \pi_m]$, which gives

$$c_m \cdot \dot{\pi}_k = c_m \cdot d_k \cdot \dot{\pi}_{m+1}.$$

We claim that $\pi_k = c_k \cdot \pi_{m+1}$, for some constant c_k. If $c_m = 0$, then $\pi_m \equiv 0$, which means $\pi_{m+1} \equiv \pi_k \equiv 0$, and the equation holds. Otherwise, $c_m > 0$, so we can divide through by it and integrate to get

$$\pi_k = d_k \cdot \pi_{m+1} + e$$

for some constant e. If $e = 0$, the equation holds. If $e \neq 0$, then from

$$[\pi_{m+1}, \pi_k] = 0 \Rightarrow \dot{\pi}_{m+1} \cdot e = 0$$

we get that π_{m+1} is constant, while

$$[\pi_{m+1}, \pi_k] = 0 \Rightarrow \dot{\pi}_k \pi_{m+1} = 0$$

gives either that $\pi_{m+1} \equiv 0 \geq \pi_k$, in which case the equation must hold, or that π_{m+1} is a positive constant, and then we can divide by it to conclude that π_k is a constant. Setting

$$c_k := \frac{\pi_k}{\pi_{m+1}}$$

shows for the last time that the desired equation holds. We have then demonstrated the existence of constants $(c_i)_{i=1}^n$ such that

$$\pi_i := \begin{cases} c_i \cdot \pi_1 & \text{if } i \leq m; \\ c_i \cdot \pi_{m+1} & \text{if } m+1 \leq i \leq n. \end{cases}$$

This makes it clear that the image of π is a chain: For $s < t$, we have

$$\pi_m(s)\pi_{m+1}(t) \leq \pi_{m+1}(s)\pi_m(t) \Rightarrow \pi(s) \sqsubseteq \pi(t)$$

and

$$\pi_m(s)\pi_{m+1}(t) \geq \pi_{m+1}(s)\pi_m(t) \Rightarrow \pi(t) \sqsubseteq \pi(s)$$

which happens because

$$\pi_i(s)\pi_{i+1}(t) = \pi_{i+1}(s)\pi_i(t)$$

for all $i \neq m$. Finally, $v\pi$ is Euclidean continuous for the standard variable $v = 1 - x^+$, so Prop. 1.10 guarantees the existence of absolute extrema. $\square$

In particular, the bracket relations (i) and (ii) in the last result are equivalent to saying that π has the form $\pi = (c_1 f, \ldots, c_m f, c_{m+1} g, \ldots, c_n g)$. In the study of the second order ordinary linear differential equation [14], two *solutions* $y_1, y_2 : (a, b) \to \mathbb{R}$ of

$$\ddot{y} + p\dot{y} + qy = 0,$$

for $p, q : (a, b) \to \mathbb{R}$ continuous, will have $[y_1, y_2] = 0$ iff there are constants c_1, c_2 not both zero such that $c_1 y_1 + c_2 y_2 = 0$.

Corollary 4.2 *If $\pi : [a,b] \to \Lambda^n$ is a differentiable curve such that*

(i) *For all $i \neq m$, $[\pi_i, \pi_{i+1}] = 0$, and*

(ii) *For all $i \neq m$, π_i and π_{i+1} are both solutions of the same second order equation,*

then the image of π is a chain in the Bayesian order.

Notice however that Theorem 4.1 is strictly more general than the situation described in the last corollary: It does not presume the existence of a second derivative.

Example 4.3 Let $f \geq g$ be nonnegative functions with f positive. Then

$$\pi = \frac{1}{h} \cdot (\,(3/4)f, (1/2)f, (1/3)g, (1/4)g, (1/10)g\,)$$

normalized by

$$h = (3/4)f + (1/2)f + (1/3)g + (1/4)g + (1/10)g$$

must have both an absolute maximum and an absolute minimum on an interval $[a,b]$ by Theorem 4.1. To find it, we check the endpoints, and then calculate the set

$$\{t : \dot{\pi}_v(t) = 0\}$$

with respect to the variable $v = 1 - x^+$, which is easily seen to be

$$\{t : [f,h](t) = 0\} = \{t : [f,g](t) = 0\}.$$

But suppose we didn't use v. Instead, suppose we try to use entropy

$$\sigma = -\sum_{i=1}^{n} x_i \log x_i.$$

Then instead of $\{t : [f,g](t) = 0\}$ we'd have to calculate $\{t : (\exists i)\, \pi_i(t) = 0\}$ and

$$\left\{ t : \sum_{i=1}^{n} \frac{\partial \sigma}{\partial x_i}(\pi_i(t)) \cdot \dot{\pi}_i(t) = 0 \right\}.$$

Thankfully, we are free to use v in place of σ, so this isn't necessary. And if we use v to calculate the absolute maximum $\pi(t^*)$, we still get to conclude that

$$\sigma\pi(t^*) = \inf\{\sigma\pi(t) : t \in \mathrm{dom}(\pi)\}.$$

That is, $\pi(t^*)$ is the unique state which minimizes the entropy! The reason is that absolute maxima simultaneously minimize *all* variables, and both σ and v are variables in the Bayesian order.

We see then a pragmatic consequence of the qualitative aspect: We can learn where crucial quantities like entropy are minimum *without* involving them in our calculations. And another significant difference between $\sqsubseteq$ and $\leq$ emerges: $v = 1 - x^+$ is a variable for the Bayesian order; for majorization *it is not.*

Example 4.4 While both orders have the entropy σ as a variable, the standard variable for the Bayesian order $v = 1 - x^+$ is *not* a variable for majorization: For $n = 3$,

$$v(1/2, 1/4, 1/4) = v(1/2, 1/2, 0)$$

despite the fact that $(1/2, 1/4, 1/4) < (1/2, 1/2, 0)$.

Finally, let us get to why absolute maxima of epistemic curves are good things to know how to calculate.

5 Quantum searching

Grover's algorithm [4][7] searches a list L of length n (which is a power of two) for an element k known to occur in L precisely m times with $n > m \geq 1$. The register begins in the pure state

$$|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle$$

and k iterations of the Grover operator G takes it to

$$G^k|\psi\rangle = \frac{\sin(2k\theta + \theta)}{\sqrt{m}} \sum_{L(i)=k} |i\rangle + \frac{\cos(2k\theta + \theta)}{\sqrt{n-m}} \sum_{L(i)\neq k} |i\rangle$$

where $\sin^2\theta = m/n$. The probability that a measurement of the register in the computational basis yields the result i after k iterations is $\sin^2(2k\theta + \theta)/m$ if $L(i) = k$ and $\cos^2(2k\theta + \theta)/(n - m)$ if $L(i) \neq k$. In particular, the probability of obtaining a solution to the search problem is $\sin^2(2k\theta + \theta)$. Thus, the epistemic trajectory of this algorithm is

$$x(t) = \left(\frac{\sin^2(2t\theta + \theta)}{m}, \ldots, \frac{\sin^2(2t\theta + \theta)}{m}, \frac{\cos^2(2t\theta + \theta)}{n-m}, \ldots, \frac{\cos^2(2t\theta + \theta)}{n-m}\right)$$

which defines a monotone state for $t \in \mathrm{dom}(x) = [a, b]$, where $a = 0$ and $b = \pi/2\theta - 1$. Theorem 4.1 immediately gives:

Proposition 5.1 *The image of the Grover trajectory $x : [a, b] \to \Lambda^n$ is a chain in the Bayesian order.*

Thus, x has both an absolute maximum and an absolute minimum on $\mathrm{dom}(x)$. Using the variable $v = 1 - \sqrt{x^+}$, we know by Lemma 1.6 that the absolute maximum is taken either at $t = a$, $t = b$, or some point in the set $\{t : \dot{x}_v(t) = 0\} = \{b/2\}$. The second derivative test is not needed here; the absolute maximum of x has to be

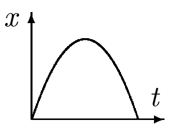
$$x(b/2) = (1/m, \dots, 1/m, 0, \dots, 0)$$

because for the other points we find an absolute minimum of

$$x(a) = x(b) = \perp.$$

Thus, the complexity of Grover's algorithm is of the same order as $b/2$ – the amount of time required to move to a state from which the likelihood of obtaining a correct result by measurement is maximized. As usual, $\theta \geq \sin\theta \geq \sqrt{m/n}$ gives $b/2 \leq (\pi/4)\sqrt{n/m} - 1/2$ so the algorithm is in $O(\sqrt{n/m})$. (We should do about $\lfloor b/2 + 1/2 \rfloor$ iterations [4]). But something unusual is going on: x takes its minimum *both* at $t = a$ and at $t = b$. What does this mean?

We can see that $\dot{x}_v(t) \leq 0$ on $[a, b/2]$ and $\dot{x}_v(t) \geq 0$ on $[b/2, b]$. By Prop. 1.4, x is monotone increasing on $[a, b/2]$ and monotone decreasing on $[b/2, b]$: Doing too many iterations *decreases* the likelihood of obtaining a solution to the search problem! The epistemic trajectory for quantum searching looks like



The space where this motion takes place is $(\Lambda^n, \sqsubseteq)$.

Based on this picture, it seems like there ought to be an "x-axis," a sort of epistemic coordinate used to describe the motion of Grover's algorithm. There is. Each *irreducible element* $y \in \mathrm{Ir}(\Lambda^n)$, which is a state y with no more than two distinct probabilities in it, gives rise to a *y-axis*: An order theoretic entity that just happens to coincide with the line joining $\perp$ to y. These *epistemic axes* suffice to determine the Bayesian order using a construction given in [5]. In the case of Grover's algorithm, the motion occurs along the

$$x(b/2) = (1/m, \dots, 1/m, 0, \dots, 0) \in \mathrm{Ir}(\Lambda^n)$$

axis. We even have $x(t) \ll x(b/2)$ for all $t \neq b/2$, which in the formal sense of exact domains [6] says that Grover's algorithm generates *approximations* of 'the answer' $x(b/2)$. Some additional points:

- By Lemma 2.4, Grover's algorithm moves in majorization as well, and when we set $m = 1$, we obtain the result of [8]. But to say that an algorithm moves up in the Bayesian order is a far stronger statement, where 'stronger' is in the sense of [8].

- One pragmatic advantage to the Bayesian order in the case of Grover's algorithm is that $v = 1 - \sqrt{x^+}$ is a variable; on majorization it is not (Example 4.4), so calculating the extreme points (complexity) would be more involved as illustrated by Example 4.3.

We have a feeling that a fun and pragmatic method for analyzing the complexity of quantum algorithms is within reach. But what is far from clear at this point in time is whether such a method can be carried out using a *single* domain of classical states, or whether different domains of classical states will have to be used, each perhaps serving to distinguish a different class of quantum algorithms.

To illustrate, an elegant connection between majorization and phase estimation is given in [15]. We believe the present work suggests that Grover's algorithm has a stronger connection to the Bayesian order than it does to majorization, a belief somewhat strengthened by a remark of the authors in [15], who write: "... both algorithms work in a different manner. Majorization is more 'natural' in the former (phase estimation) than in the latter (Grover's algorithm)." Nevertheless, the formal setting for each is a domain of classical states on which one can formulate the elements of epistemic motion, and ultimately it seems that this is where our attention should be focused.

Denouement

We have used the Bayesian order to provide a fundamental qualitative explanation for why the complexity of quantum searching is $O(\sqrt{n/m})$ in a semantic framework where its classical counterpart is already known to reside naturally. This qualitative explanation not only establishes that the classical state of the register moves in an especially rigid manner, it also reveals that calculating the complexity of this algorithm *implicitly* minimizes all variables in the Bayesian order, entropy among them. However, while this work definitely establishes the pragmatic significance of many of the ideas associated with the Bayesian order, our only conclusion is that domains of classical states with measurements may be capable of providing a systematic approach to the complexity of quantum algorithms.

We have also established but not discussed other aspects which allow one to compare classical and quantum searching. For example, if we set $y := vx$ when x is a trajectory on a domain with variable v, then the equation satisfied by binary search is

$$\ddot{y} = cy, \ c > 0$$

while for quantum searching in the variable $v = 1 - \sqrt{x^+}$ we get $\ddot{y} = cy - c, c < 0$, which setting $u = y - 1$ reveals is

$$\ddot{u} = cu, \ c < 0.$$

The first is a (radioactive) decay equation; the second is a spring equation. The case $c = 0$ is classical *linear* search. This captures an essential difference between these algorithms.

Looking beyond any particular form of computation, here is a question we should begin thinking about: To what extent can we understand 'algorithm' as a *cause* of epistemic motion? This paper has given two concrete examples where this is clearly the case, and also suggests that the mathematics of epistemic motion may only be an extension of the usual kinematics from physics. But the question of a genuine dynamics – formally connecting cause to motion – seems in the best case difficult, and in the other case – let's not think about the other case.

References

[1] S. Abramsky and A. Jung. *Domain theory.* In S. Abramsky, D. M. Gabbay, T. S. E. Maibaum, editors, Handbook of Logic in Computer Science, vol. III. Oxford University Press, 1994.

[2] P. M. Alberti and A. Uhlmann. *Stochasticity and partial order: doubly stochastic maps and unitary mixing.* Dordrecht, Boston, 1982.

[3] G. Birkhoff and J. von Neumann. *The logic of quantum mechanics.* Annals of Mathematics, **37**, 823–843, 1936.

[4] M. Boyer, G. Brassard, P. Hoyer and A. Tapp. *Tight bounds on quantum searching.* In Fourth Workshop on Physics and Computation, p.36–43, 1998. arXiv:quant-ph/9605034.

[5] B. Coecke. *Entropic geometry from logic.* Proceedings of MFPS 19, Electronic Notes in Theoretical Computer Science, 2003, to appear.

[6] B. Coecke and K. Martin. *A partial order on classical and quantum states.* 2002.

[7] L. K. Grover. *Quantum mechanics helps in searching for a needle in a haystack.* Physical Review Letters, **78**:325, 1997.

[8] Latorre, J. I. and M. A. Martín-Delgado. *The majorization arrow in quantum algorithm design.* Physical Review A, **66**:022305, 2002.

[9] A. W. Marshall and I. Olkin. *Inequalities: Theory of majorization and its applications.* Academic Press Inc., 1979.

[10] K. Martin. *Entropy as a fixed point.* 2003.

[11] K. Martin. *A foundation for computation.* Ph.D. Thesis, Department of Mathematics, Tulane University, 2000.

[12] K. Martin. *The measurement process in domain theory.* Lecture Notes in Computer Science, vol. 1853, 2000.

[13] R. F. Muirhead. *Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters.* Proc. Edinburgh Math. Soc., 21:144-157, 1903.

[14] S. Ross. *Differential equations.* Ginn-Blaisdell Publishing, 1964.

[15] R. Orús, J. I. Latorre and M. A. Martín-Delgado. *Natural majorization of the quantum fourier transformation in phase estimation algorithms.* arXiv:quant-ph/0206134, 2003.

[16] B. Zilber. *Dimensions and homogeneity in mathematical structures.* In Connections between model theory and algebraic and analytic geometry, A. Macintyre, editor, Quaderni di Matematica, v.6, 2001.