

Title: TBA

Date: Mar 26, 2018 02:00 PM

URL: <http://pirsa.org/18030101>

Abstract:

Legendre transform in mechanics

- How many undergraduate math/physics majors in audience?

Legendre transform in mechanics

- How many undergraduate math/physics majors in audience?
- Can you remember what the Legendre transform is?
- Hint: Legendre transform converts between the Lagrangian and Hamiltonian

$$L(q, \dot{q}) \leftrightarrow H(q, p)$$

Legendre transform in mechanics

- How many undergraduate math/physics majors in audience?
- Can you remember what the Legendre transform is?
- Hint: Legendre transform converts between the Lagrangian and Hamiltonian

$$L(q, \dot{q}) \leftrightarrow H(q, p)$$

- Typical answer: $H(q, p) = p\dot{q} - L(q, \dot{q})$
- But this isn't a *mathematical function transform*. Not unless you include your secret physics knowledge:

$$p = \partial L(q, \dot{q}) / \partial \dot{q}$$

Legendre transform in mechanics

- So we're defining a mathematical function transform in terms of a weird substitution procedure...

$$H(q, p) = p\dot{q} - L(q, \dot{q}) \qquad p = \partial L(q, \dot{q}) / \partial \dot{q}$$

Legendre transform in mechanics

- So we're defining a mathematical function transform in terms of a weird substitution procedure...

$$H(q, p) = p\dot{q} - L(q, \dot{q}) \qquad p = \partial L(q, \dot{q}) / \partial \dot{q}$$

- Why is it an involution? Is it always uniquely defined?

Legendre transform in mechanics

- So we're defining a mathematical function transform in terms of a weird substitution procedure...

$$H(q, p) = p\dot{q} - L(q, \dot{q}) \qquad p = \partial L(q, \dot{q}) / \partial \dot{q}$$

- Why is it an involution? Is it always uniquely defined?
- Surprise! The Lagrangian has to be convex in the velocity \dot{q}

Legendre transform in mechanics

- So we're defining a mathematical function transform in terms of a weird substitution procedure...

$$H(q, p) = p\dot{q} - L(q, \dot{q}) \qquad p = \partial L(q, \dot{q}) / \partial \dot{q}$$

- Why is it an involution? Is it always uniquely defined?
- Surprise! The Lagrangian has to be convex in the velocity \dot{q}
- Hmm...isn't the Legendre transform also used everywhere in thermodynamics? Is it really this ugly?

Legendre transform in mechanics

$$f(x) \leftrightarrow g(p)$$

(Suppressing q dependence)

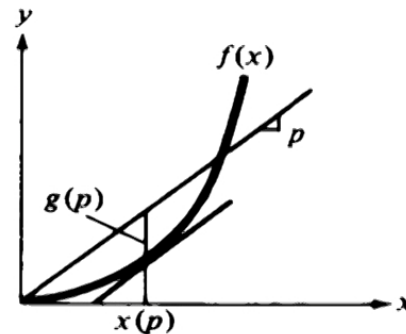
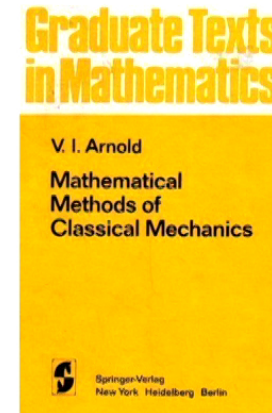


Figure 43 Legendre transformation



Let $y = f(x)$ be a convex function, $f''(x) > 0$.

The *Legendre transformation* of the function f is a new function g of a new variable p , which is constructed in the following way (Figure 43). We draw the graph of f in the x, y plane. Let p be a given number. Consider the straight line $y = px$. We take the point $x = x(p)$ at which the curve is farthest from the straight line in the vertical direction: for each p the function $px - f(x) = F(p, x)$ has a maximum with respect to x at the point $x(p)$. Now we define $g(p) = F(p, x(p))$.

The point $x(p)$ is defined by the extremal condition $\partial F / \partial x = 0$, i.e., $f'(x) = p$. Since f is convex, the point $x(p)$ is unique.²⁸

Legendre transform in mechanics

- What is the key, defining mathematical property of the Legendre transform?

Legendre transform in mechanics

- What is the key, defining mathematical property of the Legendre transform?
- cf. the Fourier transform: unique expansion of arbitrary f in eigenstates of translation operator:

$$f(x) = \sum_k \tilde{f}(k) e^{ikx}$$

Legendre transform in mechanics

- What is the key, defining mathematical property of the Legendre transform?
- cf. the Fourier transform: unique expansion of arbitrary f in eigenstates of translation operator:

$$f(x) = \sum_k \tilde{f}(k) e^{ikx}$$

- Remember: this is *the* link between Lagrangian and Hamiltonian mechanics, the two most important formulations of both classical and quantum physics. **This transformation binds together the fundamental operating system of the universe.**

Legendre transform in mechanics

- Equivalent definition: f and g are Legendre transforms of each other iff

$$f' = (g')^{-1}$$

Legendre transform in mechanics

- Equivalent definition: f and g are Legendre transforms of each other iff

$$f' = (g')^{-1}$$

- Bold claim: **essentially no physicist is taught this sensibly**

Legendre transform in mechanics

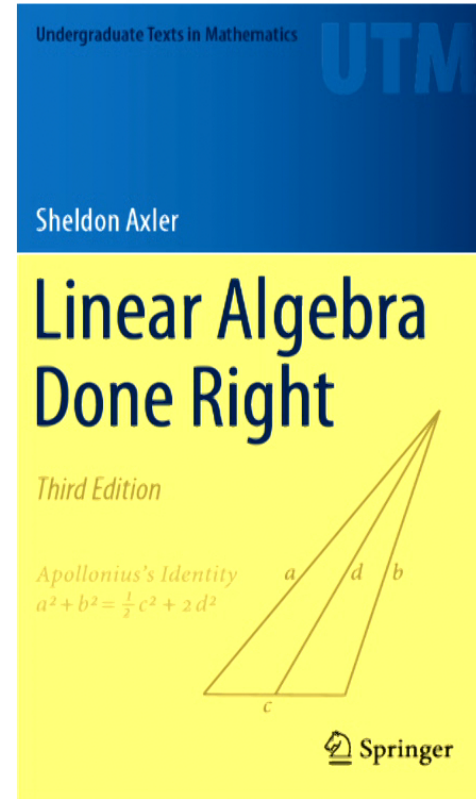
- Equivalent definition: f and g are Legendre transforms of each other iff

$$f' = (g')^{-1}$$

- Bold claim: **essentially no physicist is taught this sensibly**
- We have no reasons to think this is an exception
- (For my longer rant, google “Legendre transform Riedel”)

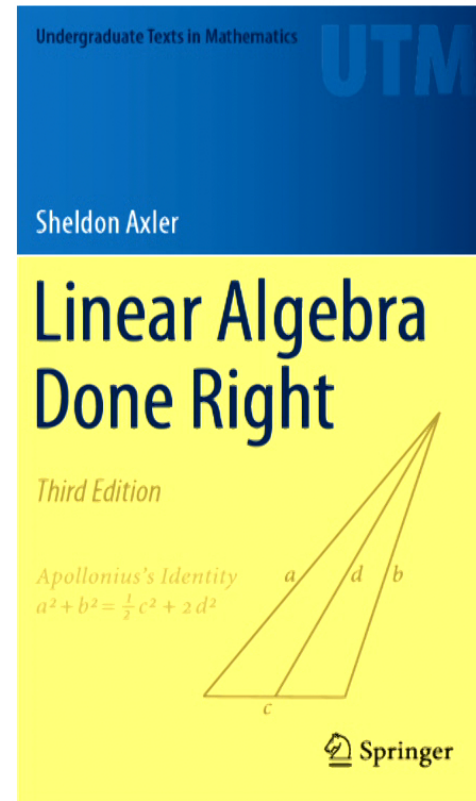
Determinants in linear algebra

- Claim: Nobody wrote a linear algebra textbook that motivates and defines determinants in a sensible way until **1997**.
- *Linear Algebra Done Right* by Sheldon Axler



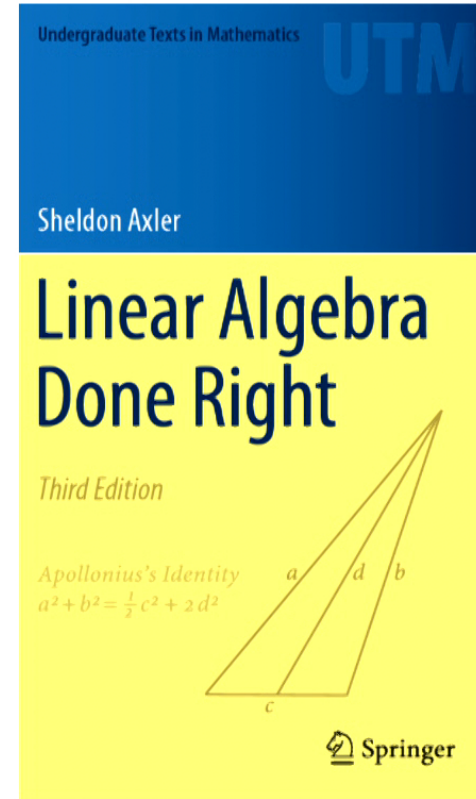
Determinants in linear algebra

- No revolutionary new proofs or ideas required



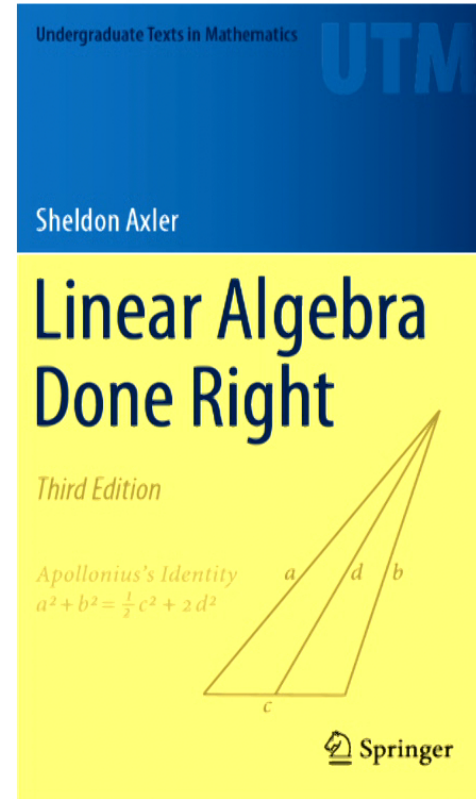
Determinants in linear algebra

- No revolutionary new proofs or ideas required
- Now “winning”: 300+ schools after 20+ years



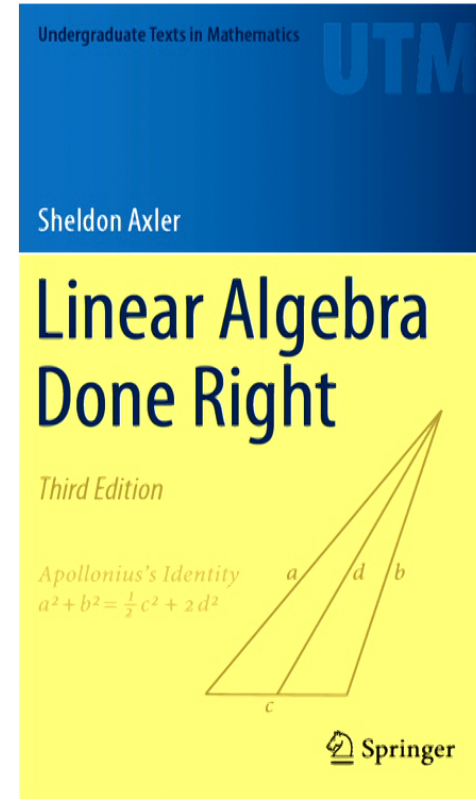
Determinants in linear algebra

- No revolutionary new proofs or ideas required
- Now “winning”: 300+ schools after 20+ years
- Consider: it would have been hopeless to just post a paper on the arXiv and expect the to work it's way into the curriculum; Axler had to write an entire new textbook!



Determinants in linear algebra

- No revolutionary new proofs or ideas required
- Now “winning”: 300+ schools after 20+ years
- Consider: it would have been hopeless to just post a paper on the arXiv and expect the to work it's way into the curriculum; Axler had to write an entire new textbook!
- This is way too hard. We are ossifying



Textbook ages

- Classical Mechanics:
 - Goldstein - 1st ed. **67** years; 3rd ed. **17** years
 - Landau & Lifshitz vol. 1 - 1st Russian ed. **58** years; 3rd ed. **42** years
 - Arnold - 1st Russian ed. **44** years; 2nd ed. **21** years
- Quantum mechanics
 - Landau & Lifshitz vol. 3 - 1st Russian ed. **60** years; 3rd ed. **41** years
 - Sakurai - **33** years
 - Shankar - 1st ed. **38** years; 2nd ed. **6** years (!)
- Electromagnetism
 - Griffiths - 1st ed. **37** years; 4th ed. **1** year (!)
 - Jackson - 1st ed. **56** years; 3rd ed. **19** years

Textbook ages

- Classical Mechanics:
 - Goldstein - 1st ed. **67** years; 3rd ed. **17** years
 - Landau & Lifshitz vol. 1 - 1st Russian ed. **58** years; 3rd ed. **42** years
 - Arnold - 1st Russian ed. **44** years; 2nd ed. **21** years
- Quantum mechanics
 - Landau & Lifshitz vol. 3 - 1st Russian ed. **60** years; 3rd ed. **41** years
 - Sakurai - **33** years
 - Shankar - 1st ed. **38** years; 2nd ed. **6** years (!)
- Electromagnetism
 - Griffiths - 1st ed. **37** years; 4th ed. **1** year (!)
 - Jackson - 1st ed. **56** years; 3rd ed. **19** years

Textbook ages

- I am *not* criticizing oldness per se; newness not valuable for its own sake; don't fix what isn't broke, etc.

Textbook ages

- I am *not* criticizing oldness per se; newness not valuable for its own sake; don't fix what isn't broke, etc.
- Key question: do you think these books are *ideal*?

Textbook ages

- I am *not* criticizing oldness per se; newness not valuable for its own sake; don't fix what isn't broke, etc.
- Key question: do you think these books are *ideal*?
- *More* physicists are being educated than ever before, and basic physics is changing *less*

Textbook ages

- I am *not* criticizing oldness per se; newness not valuable for its own sake; don't fix what isn't broke, etc.
- Key question: do you think these books are *ideal*?
- *More* physicists are being educated than ever before, and basic physics is changing *less*
- We should be investing *more* in carefully optimizing pedagogical materials

Textbook ages

- I am *not* criticizing oldness per se; newness not valuable for its own sake; don't fix what isn't broke, etc.
- Key question: do you think these books are *ideal*?
- *More* physicists are being educated than ever before, and basic physics is changing *less*
 - We should be investing *more* in carefully optimizing pedagogical materials
 - Instead, we are allowing our researcher pipeline to slowly decay

Stagnation

- If Landau & Lifshitz has a flaw, currently you can write up some supplementary notes, or assign alternate reading

Stagnation

- If Landau & Lifshitz has a flaw, currently you can write up some supplementary notes, or assign alternate reading
- But these improvements rarely propagate; people can copy your reading list or pass around your notes, but it's impractical for them to improve on top of *that*

Stagnation

- If Landau & Lifshitz has a flaw, currently you can write up some supplementary notes, or assign alternate reading
- But these improvements rarely propagate; people can copy your reading list or pass around your notes, but it's impractical for them to improve on top of *that*
- If you want something that can propagate and displace the single flawed section of Landau & Lifshitz, you need to write *an entire textbook net-better than Landau & Lifshitz*

Stagnation

- If Landau & Lifshitz has a flaw, currently you can write up some supplementary notes, or assign alternate reading
- But these improvements rarely propagate; people can copy your reading list or pass around your notes, but it's impractical for them to improve on top of *that*
- If you want something that can propagate and displace the single flawed section of Landau & Lifshitz, you need to write *an entire textbook net-better than Landau & Lifshitz*
- If your textbook is better on net, but worse in certain parts, then teaching of those parts will get *worse*

Not just Legendre and Det

- These are elementary examples familiar to much of the current audience

Not just Legendre and Det

- These are elementary examples familiar to much of the current audience
- Review articles and research articles have similar or worse deficiencies; textbooks are just egregious

Not just Legendre and Det

- These are elementary examples familiar to much of the current audience
- Review articles and research articles have similar or worse deficiencies; textbooks are just egregious
- Things get worse as things get more specialized

Not just Legendre and Det

- These are elementary examples familiar to much of the current audience
- Review articles and research articles have similar or worse deficiencies; textbooks are just egregious
- Things get worse as things get more specialized
 - E.g., the infinitely-repeated misconception that “Classical sensitivity to initial conditions doesn’t extend to quantum chaos because the Schrödinger equation is linear”

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”
- Who thinks Wikipedia is useful?

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”
- Who thinks Wikipedia is useful?
- Almost every fact in Wikipedia is accessible elsewhere on the internet

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”
- Who thinks Wikipedia is useful?
- Almost every fact in Wikipedia is accessible elsewhere on the internet
 - The fact that it’s *out there, somewhere*, is not enough

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”
- Who thinks Wikipedia is useful?
- Almost every fact in Wikipedia is accessible elsewhere on the internet
 - The fact that it’s *out there, somewhere*, is not enough
- Bringing everything together is critical:

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”
- Who thinks Wikipedia is useful?
- Almost every fact in Wikipedia is accessible elsewhere on the internet
 - The fact that it’s *out there, somewhere*, is not enough
- Bringing everything together is critical:
 - *Much* faster to find

“It is known that...”

- Common response: “This information is *known*. It’s not a *secret*. It’s available on Google for goodness’ sake.”
- Who thinks Wikipedia is useful?
- Almost every fact in Wikipedia is accessible elsewhere on the internet
 - The fact that it’s *out there, somewhere*, is not enough
- Bringing everything together is critical:
 - *Much* faster to find
 - Central location allows iterative improvement

“It is known that...”

- “What the community knows” is not good as a discrete category

“It is known that...”

- “What the community knows” is not good as a discrete category
- Not just pedantry: It has large implications for how we organize expert knowledge

“It is known that...”

- “What the community knows” is not good as a discrete category
 - Not just pedantry: It has large implications for how we organize expert knowledge
- Researchers sometimes get upset or dismissive when someone “reinvents an old idea” in another field, and gets tons of citations

“It is known that...”

- “What the community knows” is not good as a discrete category
 - Not just pedantry: It has large implications for how we organize expert knowledge
- Researchers sometimes get upset or dismissive when someone “reinvents an old idea” in another field, and gets tons of citations
 - Researchers are hungry for useful techniques! If they don’t deploy an idea until it is reinvented in their language, **it was not practically accessible before**

“It is known that...”

- “What the community knows” is not good as a discrete category
 - Not just pedantry: It has large implications for how we organize expert knowledge
- Researchers sometimes get upset or dismissive when someone “reinvents an old idea” in another field, and gets tons of citations
 - Researchers are hungry for useful techniques! If they don’t deploy an idea until it is reinvented in their language, **it was not practically accessible before**
- The goal of research is **not** for one person to know; if it was, we wouldn’t require publication

Possibilities

Increasingly ambitious —

- For any document on the arXiv, suppose the reader could...
 1. Give one-click feedback to draw author's attention to issues (confusing, unjustified, etc.)

— Cutting edge
Pedagogy —

One-click feedback

Reader's
view:

But since $D \cdot \Delta \leq 1/4$, and, by assumption $\mu \leq \sqrt{2D}$, it follows that $D(\mu^2 + \Delta) < 1$, and so the overlap of $|L_i\rangle|R_i\rangle$ with the ground state is larger than μ . ■

With this bound in place, we start from the product state with the maximal overlap with the ground state, and use any AGSP to obtain controlled approximations of the ground state, from which an upper bound on its entropy can be found. A very similar argument was used in Hastings' proof of the 1D area law.¹³

Lemma III.3. If there exists a product state whose overlap with the ground state is at least μ , together with a (D, Δ) -AGSP, then the entanglement entropy of $|\Omega\rangle$ is bounded by

$$S \leq \mathcal{O}(1) \cdot \frac{\log \mu^{-1}}{\log \Delta^{-1}} \log D. \quad (14)$$

The proof can be found in the appendix. The brief overview is that we begin with the accepted product state and repeatedly apply the AGSP to it, increasing the number of vectors with increasing SR by a factor of Δ at each step. That approach the ground state at a rate of μ per step. Using these vectors and the Young's inequality (see Appendix II.2) provides an adequate upper bound on the entropy. The Young's inequality provides an adequate upper bound on the entropy of the ground state to be approximated.

Lemmas III.3 and III.2 can be combined to give

Corollary III.4. If there exists an (D, Δ) -AGSP such that $D \cdot \Delta \leq \frac{1}{2}$, the ground-state entropy is bounded by

$$S \leq \mathcal{O}(1) \cdot \log D. \quad (15)$$

the $(D,$

with $(D$

log

We r

With
cut in th
As a re
operato
will mo

We l
the m e
projecti
them by

the pro
jections
 P_{α} wh

One-click feedback

Reader's
view:

But since $D \cdot \Delta \leq 1/4$, and, by assumption $\mu \leq \sqrt{2D}$, it follows that $D(\mu^2 + \Delta) < 1$, and so the overlap of $|L_i\rangle|R_i\rangle$ with the ground state is larger than μ . ■

With this bound in place, we start from the product state with the maximal overlap with the ground state, and use any AGSP to obtain controlled approximations of the ground state, from which an upper bound on its entropy can be found. A very similar argument was used in Hastings' proof of the 1D area law.¹³

Lemma III.3. If there exists a product state whose overlap with the ground state is at least μ , together with a (D, Δ) -AGSP, then the entanglement entropy of $|\Omega\rangle$ is bounded by

$$S \leq \mathcal{O}(1) \cdot \frac{\log \mu^{-1}}{\log \Delta^{-1}} \log D. \quad (14)$$

The proof can be found in the appendix. The brief overview is that we begin with the asserted product state, repeatedly apply the AGSP to it. The result is a series of vectors with increasing SR by a factor D each time, they approach the ground state at a rate quantified by powers of Δ . Using these vectors and the Young-Eckart theorem (Fact II.2) provides adequate upper bounds for the high Schmidt coefficients of the ground state to bound the entropy.

Lemmas III.3 and III.2 can be combined to give

Corollary III.4. If there exists an (D, Δ) -AGSP such that $D \cdot \Delta \leq \frac{1}{2}$, the ground-state entropy is bounded by

$$S \leq \mathcal{O}(1) \cdot \log D. \quad (15)$$

the $(D,$

with $(D$

log

We r

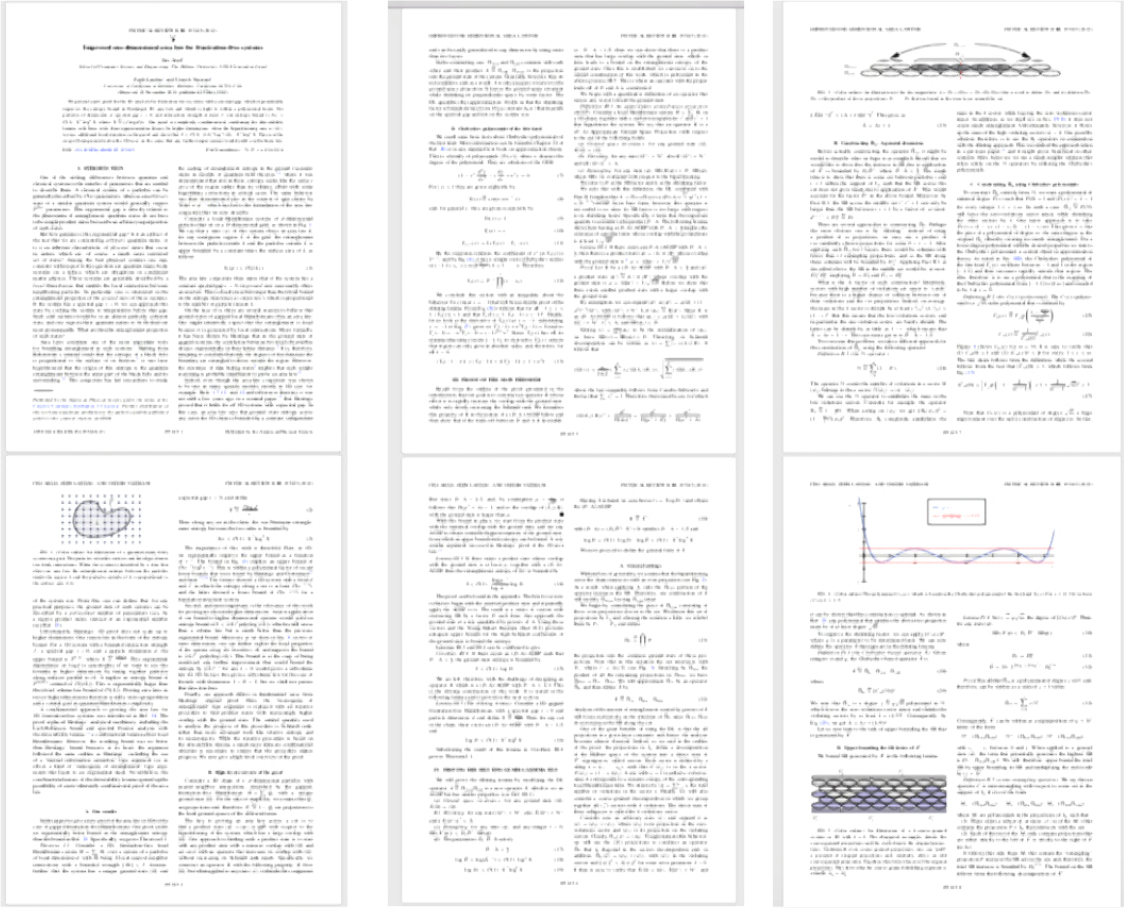
With
cut in th
As a re
operato
will mo

We l
the m e
projecti
them by

the pro
jections
 P_{α} wh

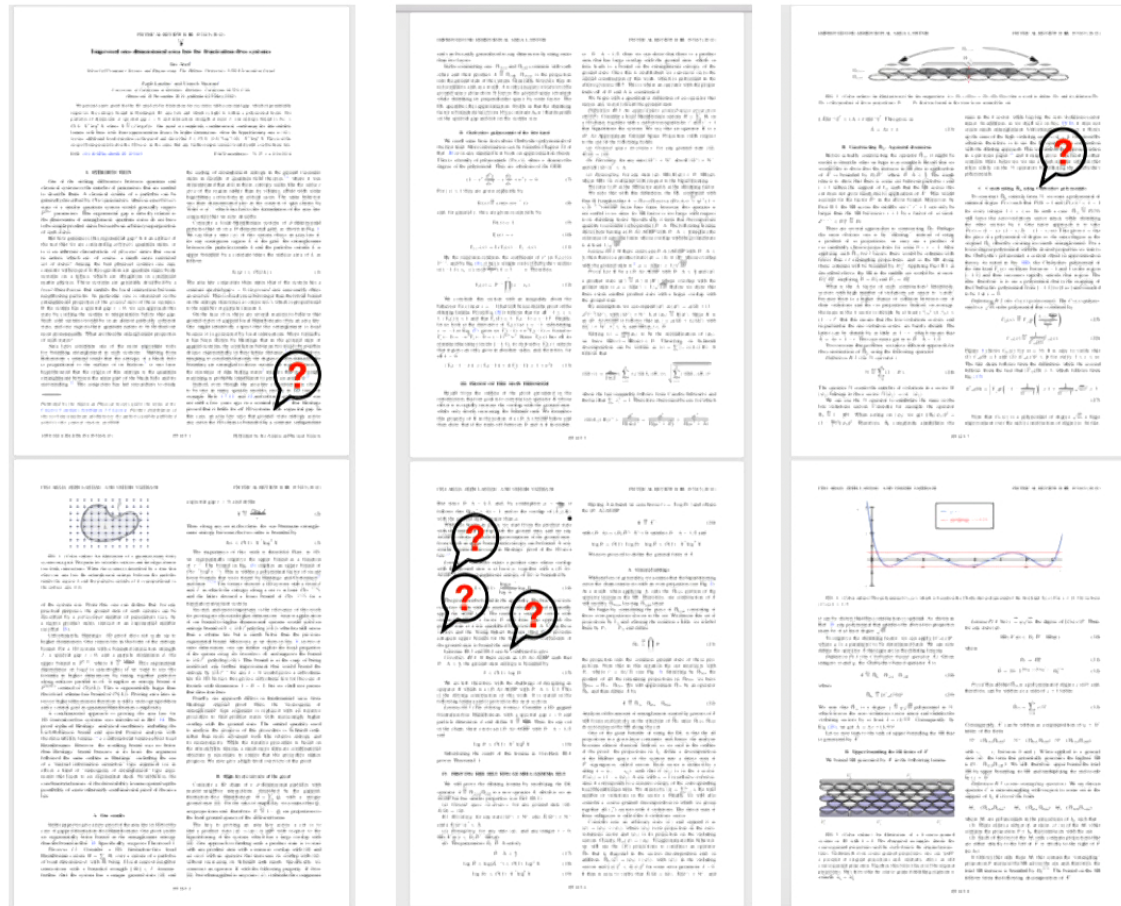
One-click feedback

Author's
view:



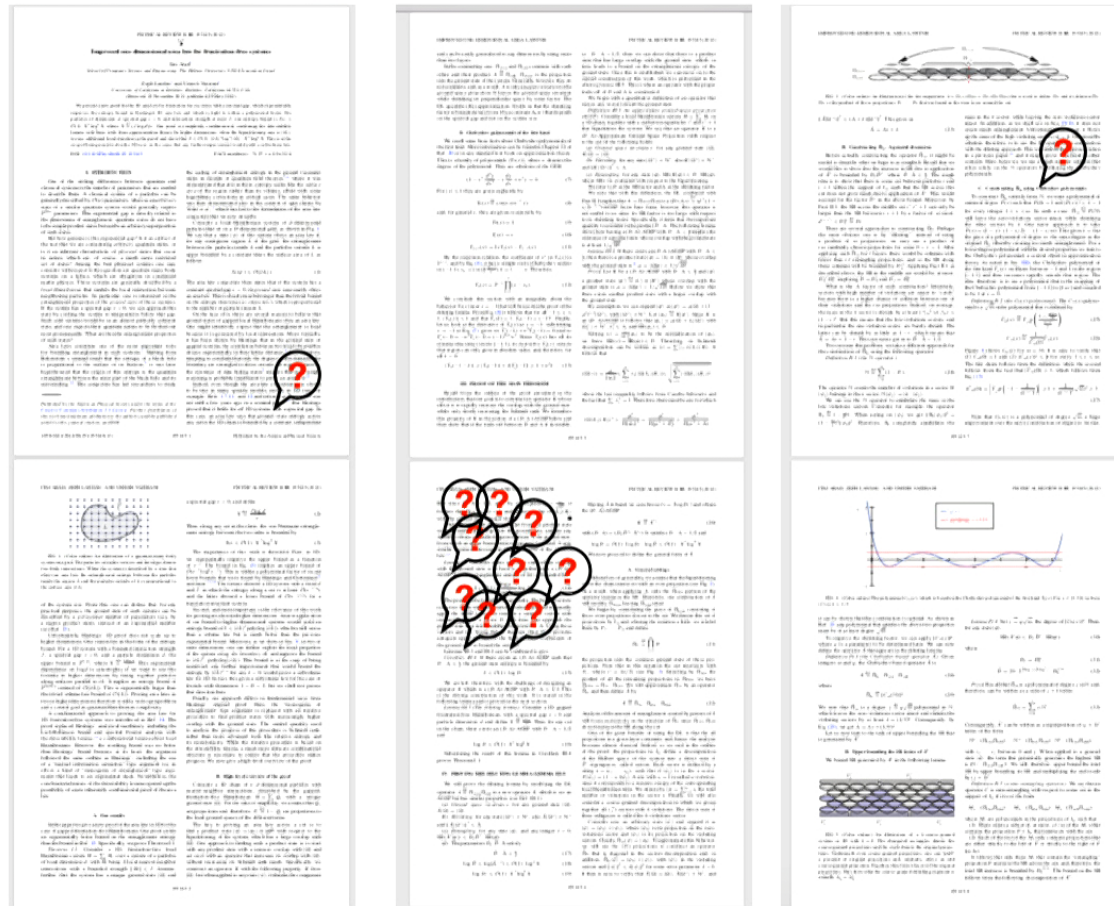
One-click feedback

Author's
view:



One-click feedback

Author's
view:



Possibilities

Increasingly ambitious —

- For any document on the arXiv, suppose the reader could...
 1. Give one-click feedback to draw author's attention to issues (confusing, unjustified, etc.)
 2. Instantly write a (possibly anonymous) message to the author: "There is a simpler derivation in..."

— Cutting edge
Pedagogy —

Message feedback

Reader's
view:

follows that $D(\mu^2 + \Delta) < 1$, and so the overlap of $|L_i\rangle|R_i\rangle$ with the ground state is larger than μ . ■

With this bound in place, we start from the product state with the maximal overlap with the ground state, and use any AGSP to obtain controlled approximations of the ground state, from which an upper bound on its entropy can be found. A very similar argument was used in Hastings' proof of the 1D area law.¹³

Lemma III.3. If there exists a product state whose overlap with the ground state is at least μ , together with a (D, Δ) -AGSP, then the entanglement entropy of $|\Omega\rangle$ is bounded by

$$S \leq \mathcal{O}(1) \cdot \frac{\log \mu^{-1}}{\log \Delta^{-1}} \log D. \quad (14)$$

The proof can be found in the appendix. The is that we begin with the selected product state apply the AGSP to increasing SR by a ground state at a rate vectors and the You adequate upper bound the ground state to be

Lemmas III.3 and III.2 can be combined to give

Corollary III.4. If there exists an (D, Δ) -AGSP such that $D \cdot \Delta \leq \frac{1}{2}$, the ground-state entropy is bounded by

$$S \leq \mathcal{O}(1) \cdot \log D. \quad (15)$$

the $(D,$

with $(E$

log

We r

With
cut in th
As a re

There is a simpler
derivation of this
in 1713.4629

the pro
jections
 P_{α} wh

Message feedback

Author's
view:

follows that $D(\mu^2 + \Delta) < 1$, and so the overlap of $|L_i\rangle|R_i\rangle$ with the ground state is larger than μ . ■

With this bound in place, we start from the product state with the maximal overlap with the ground state, and use any AGSP to obtain controlled approximations of the ground state, from which an upper bound on its entropy can be found. A very similar argument was used in Hastings' proof of the ID area law.¹³


Lemma III.3. If there exists a product state whose overlap with the ground state is at least μ , together with a (D, Δ) -AGSP, then the entanglement entropy of $|\Omega\rangle$ is bounded by

$$S \leq \mathcal{O}(1) \cdot \frac{\log \mu^{-1}}{\log \Delta^{-1}} \log D. \quad (14)$$

The proof can be found in the appendix. The brief overview is that we begin with the asserted product state and repeatedly apply the AGSP to it. The result is a series of vectors with increasing SR by a factor D each time, that approach the ground state at a rate quantified by powers of Δ . Using these vectors and the Young-Eckart theorem (Fact II.2) provides adequate upper bounds for the high Schmidt coefficients of the ground state to bound the entropy.

Lemmas III.3 and III.2 can be combined to give

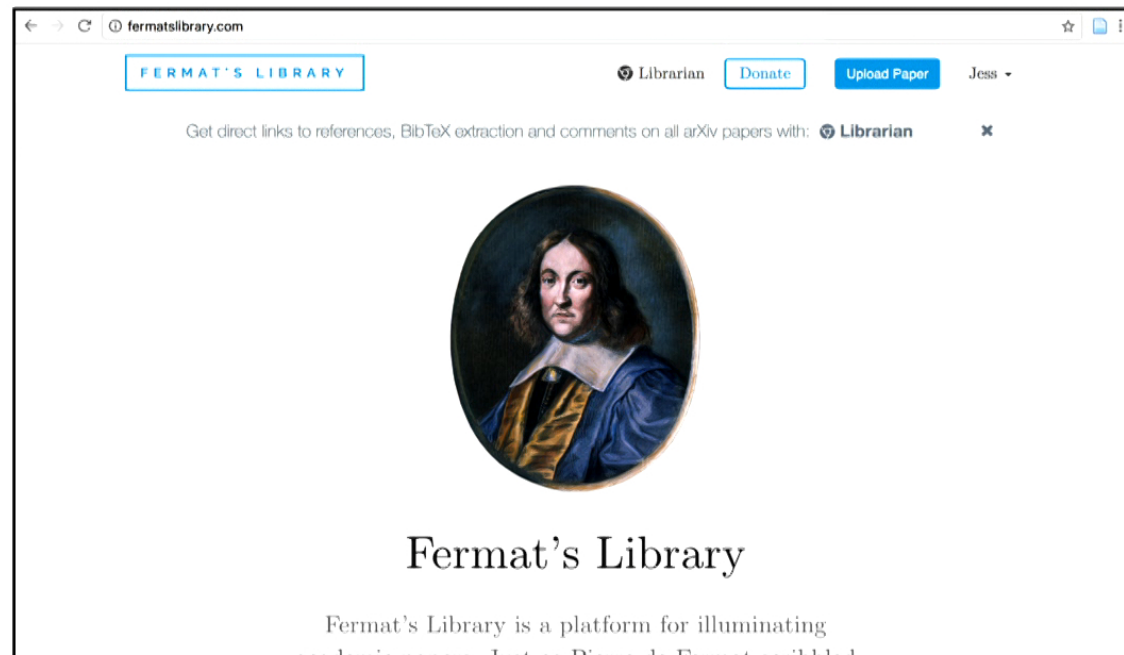
Corollary III.4. If there exists an (D, Δ) -AGSP such that $D \cdot \Delta \leq \frac{1}{2}$, the ground-state entropy is bounded by

 **Starr Sackstein**
4:29 PM Yesterday

There is a simpler derivation of this in 1713.4629

 **Jess Riedel**
4:29 PM Yesterday

Message feedback...recently available!



FermatsLibrary.com

see the “Librarian” plugin for Chrome



2
KLHPJ



Fermat's Library

fermatslibrary.com

Perimeter Institute, March 2018



Luis

...t it looks like



2
KLHPJ

"Peer-to-Peer" is an

Here I give a quick

The double spending

What is proof of work?

**Reversible

Follow Paper

Bitcoin: A Peer-to-Peer Electronic Cash System

Satoshi Nakamoto
satoshin@gmx.com
www.bitcoin.org

Abstract. A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a best effort basis, and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone.

1. Introduction

Commerce on the Internet has come to rely almost exclusively on financial institutions serving as trusted third parties to process electronic payments. While the system works well enough for most transactions, it still suffers from the inherent weaknesses of the trust based model. Completely non-reversible transactions are not really possible, since financial institutions cannot avoid mediating disputes. The cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions, and there is a broader cost in the loss of ability to make non-reversible payments for non-reversible services. With the possibility of reversal, the need for trust spreads. Merchants must

Satoshi Nakamoto is the

The risk that a digital



Luis

What it looks like

João Batalha - MIT CS, YC founder

Here I give a quick overview of a few concepts important for a good understanding of bitcoin.

Public-keys and Private-keys

The concept of public-key and private-key come from Public-key cryptography. Public-key cryptography is a set of cryptographic protocols based on algorithms that require two separate keys:

- Private-key - which as the name indicates is meant to be secret
- Public-key - which is public / visible to others

These two keys are mathematically linked. In public-key cryptography the public key is used to encrypt plaintext, where the private key is used to decrypt cipher text. Every node in the bitcoin network has a public-key and a private-key.

Digital Signatures

Digital signatures make heavy use of public-key cryptography. You can think of a digital signature as somewhat similar to a physical signature. A digital signature is also used to prove the authenticity of a document/digital message. A digital signature binds an identity to a message. Only the person with the private key can produce valid signatures. Anybody with access to the public key can test the validity of the signatures.

Say Alice wants to digitally sign a message m . In order to do that Alice must have:

- Private-key (signing key) - $KEY_{private}$
- Public-key (verification key) - KEY_{public}

Alice then uses the *signing* function to produce a valid

Use $LaTeX$ to type formulae and *markdown* to format text.

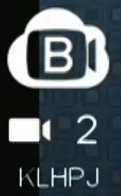
Bitcoin: A Peer-to-Peer Electronic Cash

Satoshi Nakamoto
satoshi@gmx.com
www.bitcoin.org

Abstract. A purely peer-to-peer version of electronic cash would allow payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a peer-to-peer basis, and nodes can leave and join the network at will, accepting the proof-of-work chain as proof of what happened while they were gone.

1. Introduction

Commerce on the Internet has come to rely almost exclusively on financial institutions to process electronic payments. While the system works well for most transactions, it still suffers from the inherent weaknesses of the trusted third party: a central point of failure and a costly network. Completely non-reversible transactions are not really possible, since financial institutions avoid mediating disputes. The cost of mediation increases transaction size and cutting off the possibility for small transactions. There is a broader cost in the loss of ability to make non-reversible payments for irreversible services. With the possibility of reversal, the need for trust spreads. Merchants are wary of their customers, hassling them for more information than they would otherwise need. A certain percentage of fraud is accepted as unavoidable. These costs and pains can be avoided in person by using physical currency, but no mechanism exists



2
KLHPJ



Luis

Motivation

This gives (ii). The implication (ii) \rightarrow (iii) is trivial since, for example,

$$\int_0^1 \left\| \sum_k r_k(u) \psi(k+t-A)^* x^* \right\|^2 du \leq \sup_{\epsilon_k = \pm 1} \left\| \left[\sum_k \epsilon_k \psi(k+t-A) \right]^* \right\|^2 \|x^*\|^2 \leq C \|x^*\|^2.$$



2
KLHPJ



motivation

Erdős discrepancy problem (1932)



2

KLHPJ



Luis

motivation

Erdős discrepancy problem (1932)



9 September, 2015 at 12:06 am
Uwe Stroinski

The Sudoku-flavor arguments remind me on the EDP Polymath project, where some of us tried to prove (without computer) that completely multiplicative sequences with values in ± 1 have discrepancy greater than 3. Can these recent results of Matomaki and Radziwiłł be used/adapted/generalized to help with this problem or is there some obstacle to make that hopeless?

48 0 Rate This
Reply

9 September, 2015 at 11:08 am
Terence Tao

There is indeed some similarity on the surface, but Matomaki-Radziwiłł only lets one control the sum of a ± 1 -valued multiplicative functions f in short intervals such as $[x, x+H]$ where H is much smaller than x , basically by using Fourier inversion (or Perron's formula) to convert this to a question about the Dirichlet series $\sum_n \frac{f(n)}{n^s} = \sum_n \frac{f(n)}{n^{s+1}}$. Roughly speaking, the relationship between the intervals $[x, x+H]$ and the phases n^{it} is that n^{it} and $n^{it'}$ essentially differ only by a constant when $t' - t \ll \frac{1}{x}$. By using Dirichlet characters one can also control f in short progressions such as $\{n \in [x, x+H] : n \equiv a \pmod{q}\}$ for q small, H medium size, and x very large, but I don't see an obvious way to control the EDP type discrepancies which are more to do with progressions such as $\{n \leq x : n \equiv 0 \pmod{d}\}$ when x, d are both large.

EDIT: Ah, using complete multiplicativity I see that the EDP for completely multiplicative functions is equivalent to *lower bounding* the sum of f on intervals such as $[x, x+H]$ rather than upper bounding it. The Matomaki-Radziwiłł technology is geared towards upper bounds only. As usual we have the problem that Dirichlet characters already have bounded discrepancy, so one has to somehow use the fact that the multiplicative function doesn't vanish...

9 0 Rate This
Reply

29 September, 2015 at 5:22 am
Domi

In the end this was useful:
<http://arxiv.org/abs/1509.05363>
Congratulations!

6 0 Rate This
Reply

From terrytao.wordpress.com



2
KLHPJ



Luis

motivation

Erdős discrepancy problem (1932)



9 September, 2015 at 12:06 am

Uwe Stroinski

The Sudoku-flavor arguments remind me on the EDP Polymath project, where some of us tried to prove (without computer) that completely multiplicative sequences with values in ± 1 have discrepancy greater than 3. Can these recent results of Matomaki and Radziwiłł be used/adapted/generalized to help with this problem or is there some obstacle to make that hopeless?

48 0 Rate This
Reply

9 September, 2015 at 11:08 am

Terence Tao

There is indeed some similarity on the surface, but Matomaki-Radziwiłł only

where
on's

differ
can

very
cies
(d)

tely
aki-
have
r, so
t

DISCRETE ANALYSIS, 2016:1, 27 pp.
www.discreteanalysisjournal.com

The Erdős discrepancy problem

Terence Tao*

Received 17 September 2015; Published 28 February 2016

Abstract: We show that for any sequence $f(1), f(2), \dots$ taking values in $\{-1, +1\}$, the discrepancy

$$\sup_{n,d \in \mathbb{N}} \left| \sum_{j=1}^n f(jd) \right|$$

of f is infinite. This answers a question of Erdős. In fact the argument also applies to sequences f taking values in the unit sphere of a real or complex Hilbert space.

[math.CO] 13 Jan 2017



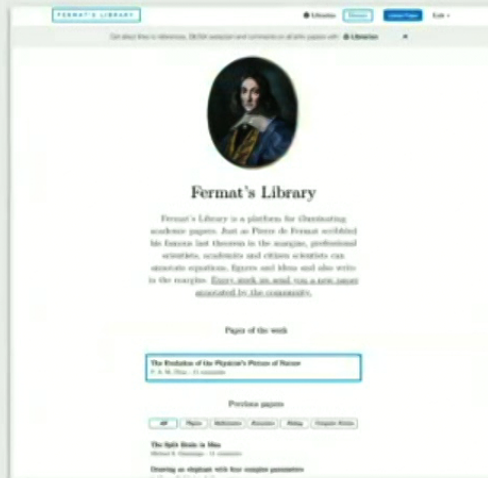
2
KLHPJ



Luis

Evolution of Fermat's Library

V1 - Journal Club (end of 2015)



1 paper per week

Newsletter (13k subscribers)

Twitter (92k Followers)

Hacker News

Reddit



10M people reached every month



2

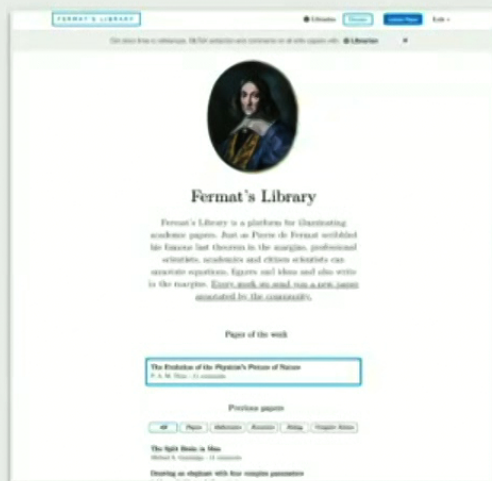
KLHPJ



Luis

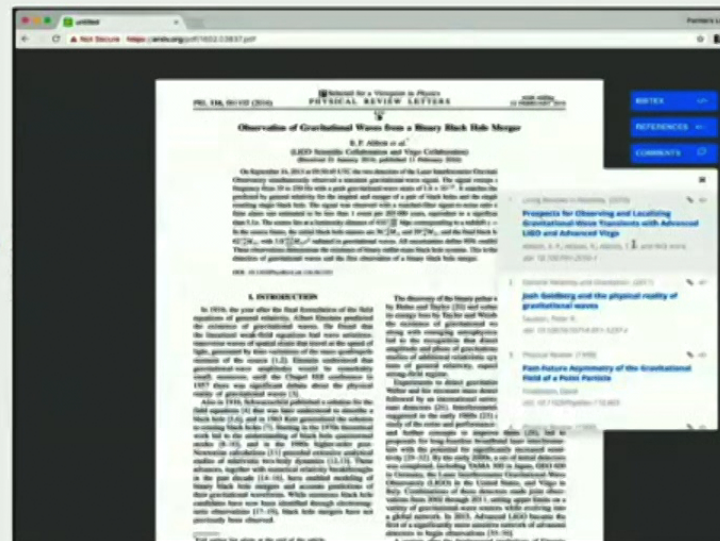
Evolution of Fermat's Library

V1 - Journal Club (end of 2015)



1 paper per week

V2 - Platform (end of 2017)



- Librarian: Chrome Extension for arXiv - 1.3M Pre-prints (7k users)
- Fermat's Library Core: You can upload your own papers



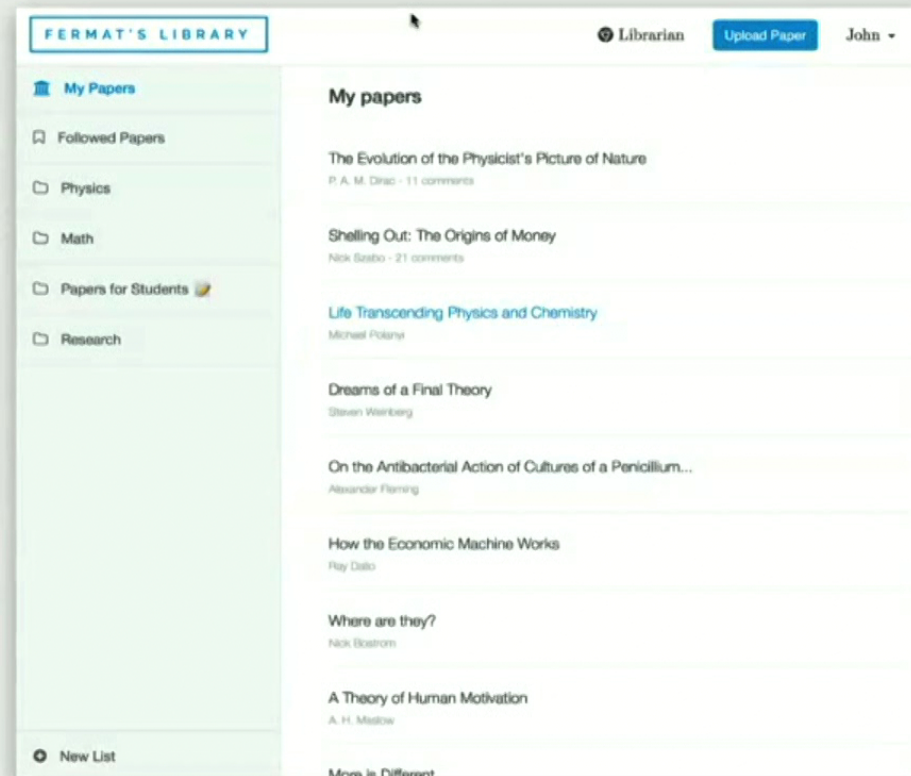
2
KLHPJ



Luis

Fermat's Library Core

- A Notes app (or Evernote) for all of your papers
- Scientists/Students or Researchers can upload and annotate their papers and share them with peers
- University Journal Clubs can use it to collect questions and foster the discussion around papers



2
KLHPJ



Luis

arXiv Chrome Extension - Librarian

- Support for comments and other features requested by arXiv users (reference extraction, bibtex etc) to incentivize the installation of the extension
- Easy to use without having to leave arXiv
- AMA's (Ask Me Anything) with authors to promote Librarian

Comments for 1406.2661

Avhirup Chakraborty
Are there papers / work out there which test transferability of generative models. What I mean suppose we have a model to generate images of dogs and cats, and we later run some of the earlier layers to generate images if wolves and tigers

Ian Goodfellow
Replying to "Are there papers / work out there which test transferability of generative models. What I mean suppose we have a model to generate images of dogs and cats, and we later run some of the earlier layers to generate images if wolves and tigers? I don't know of any, but I assume that would work, for the generator, it's probably the later layers (closer to the output) that are easier to share."

Generative Adversarial Nets

John Pong-Akshid, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio
Department of Mathematics and Statistics, Université de Montréal
Montreal, QC H3C 3J7

Abstract

We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that takes noise as input and generates data-like samples, and a discriminator model D that takes samples or data generated by G and tries to distinguish between them. The discriminator D is trained to maximize the probability of D taking a sample from G as input to be a sample from the training data. This corresponds to a minimax two-player game. In the space of arbitrary functions G , a simple solution exists, with G recovering the training data of D (up to a permutation). In the case where G and D are defined as parametric models, the entire system can be trained with backpropagation. For any Markov chain or stochastic approximation (including neural networks), we show that the system can be trained with backpropagation. Experiments demonstrate the effectiveness of the framework through qualitative and quantitative evaluation of samples.

Learning to discriminate data, adversarial models [2] that approximate probability distributions over data, are used in a wide range of applications, such as natural language processing, speech, and computer vision. In the case of natural language, the task is to learn to discriminate between data generated by a generative model and data from the training set. This is a challenging task, as the generative model is often trained to maximize the likelihood of the training data, which is a different objective than maximizing the probability of the training data being generated by the model. In this paper, we propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that takes noise as input and generates data-like samples, and a discriminator model D that takes samples or data generated by G and tries to distinguish between them. The discriminator D is trained to maximize the probability of D taking a sample from G as input to be a sample from the training data. This corresponds to a minimax two-player game. In the space of arbitrary functions G , a simple solution exists, with G recovering the training data of D (up to a permutation). In the case where G and D are defined as parametric models, the entire system can be trained with backpropagation. For any Markov chain or stochastic approximation (including neural networks), we show that the system can be trained with backpropagation. Experiments demonstrate the effectiveness of the framework through qualitative and quantitative evaluation of samples.

Learning to discriminate data, adversarial models [2] that approximate probability distributions over data, are used in a wide range of applications, such as natural language processing, speech, and computer vision. In the case of natural language, the task is to learn to discriminate between data generated by a generative model and data from the training set. This is a challenging task, as the generative model is often trained to maximize the likelihood of the training data, which is a different objective than maximizing the probability of the training data being generated by the model. In this paper, we propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that takes noise as input and generates data-like samples, and a discriminator model D that takes samples or data generated by G and tries to distinguish between them. The discriminator D is trained to maximize the probability of D taking a sample from G as input to be a sample from the training data. This corresponds to a minimax two-player game. In the space of arbitrary functions G , a simple solution exists, with G recovering the training data of D (up to a permutation). In the case where G and D are defined as parametric models, the entire system can be trained with backpropagation. For any Markov chain or stochastic approximation (including neural networks), we show that the system can be trained with backpropagation. Experiments demonstrate the effectiveness of the framework through qualitative and quantitative evaluation of samples.

Ian Goodfellow
Replying to "This is probably a bit off topic, but what do you think re-thinking the way 'neurons' work in deep learning models today is the only way to be robust against adversarial examples? In a couple of your videos, you'd mentioned that the existence of such universal examples exists because of the 'linearity' of such models (as opposed to 'non-linearity', as many people claim). Are we barking up the wrong tree by pursuing with our current model of these neurons? Or is it something that a mixture of well-placed activation functions and data-augmentation / defenses applied to the model as a whole can fix?"

Ian Goodfellow
Replying to "This is probably a bit off topic, but what do you think re-thinking the way 'neurons' work in deep learning models today is the only way to be robust against adversarial examples? In a couple of your videos, you'd mentioned that the existence of such universal examples exists because of the 'linearity' of such models (as opposed to 'non-linearity', as many people claim). Are we barking up the wrong tree by pursuing with our current model of these neurons? Or is it something that a mixture of well-placed activation functions and data-augmentation / defenses applied to the model as a whole can fix?"

Use DFTB to explore the structure and properties of the system.

View your comments/answers

Page 13 of 15

Ian Goodfellow, a world renowned Machine Learning researcher, answers dozens of questions about one of his most famous papers in an AMA on Librarian



Luis

uture Perspectives

Growth in online collaboration

Open discussion

Open peer review system

Ranking of relevant research

More knowledge sharing



2
KLHPJ



2
KLHPJ

Thank You! ,



fermatslibrary.com

@fermatslibrary

team@fermatslibrary.com

Luis Batalha, João Batalha, Micael Oliveira, Tymor Hamamsy

Concrete steps - copyleft

- Understand the difference between open access and copyleft; be able to explain it to others