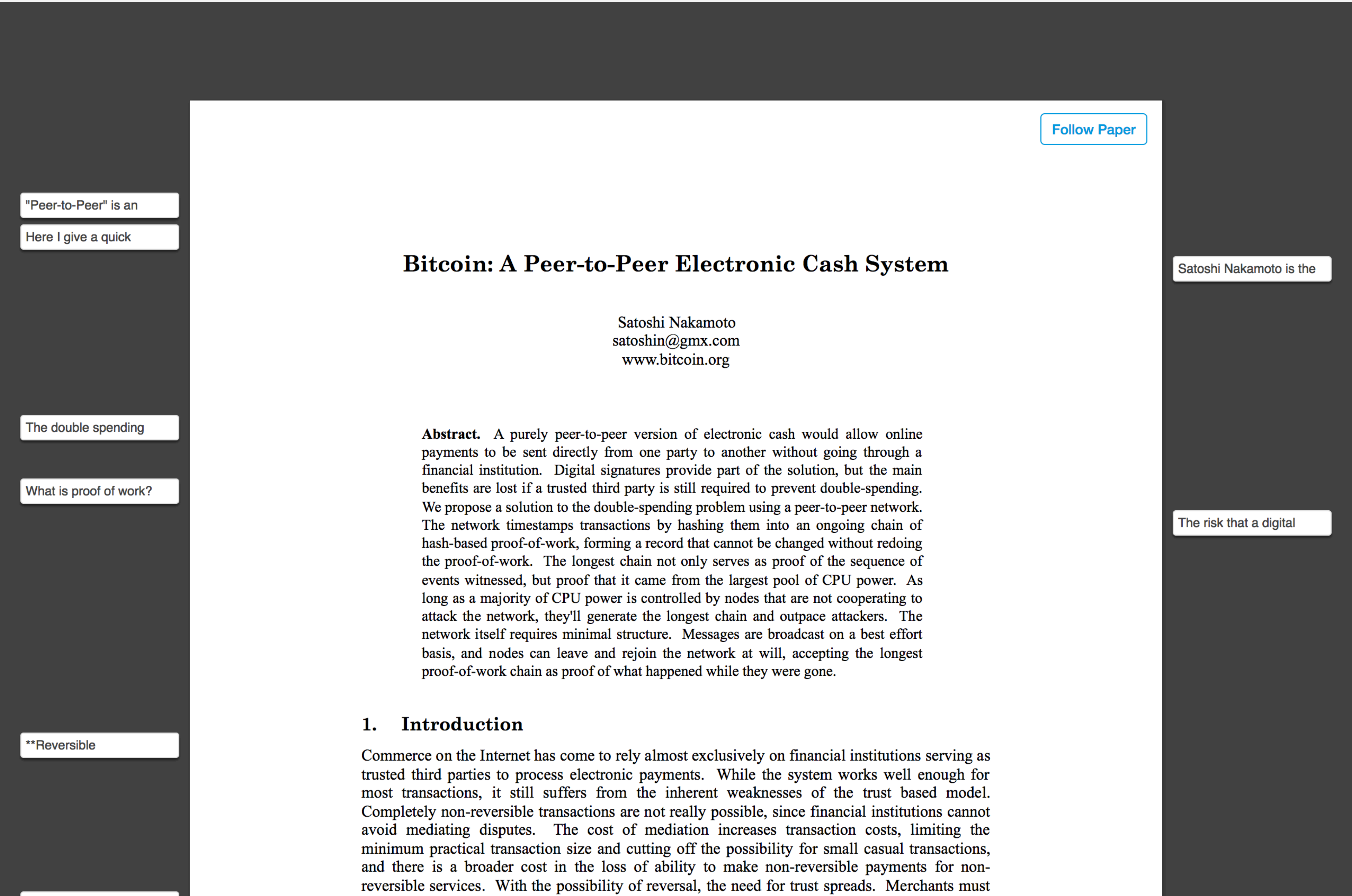
Fermat's Library

Luis Batalha

Data Science Symposium
Lisbon, March 2018

What it looks like



Bitcoin: A Peer-to-Peer Electronic Cash System

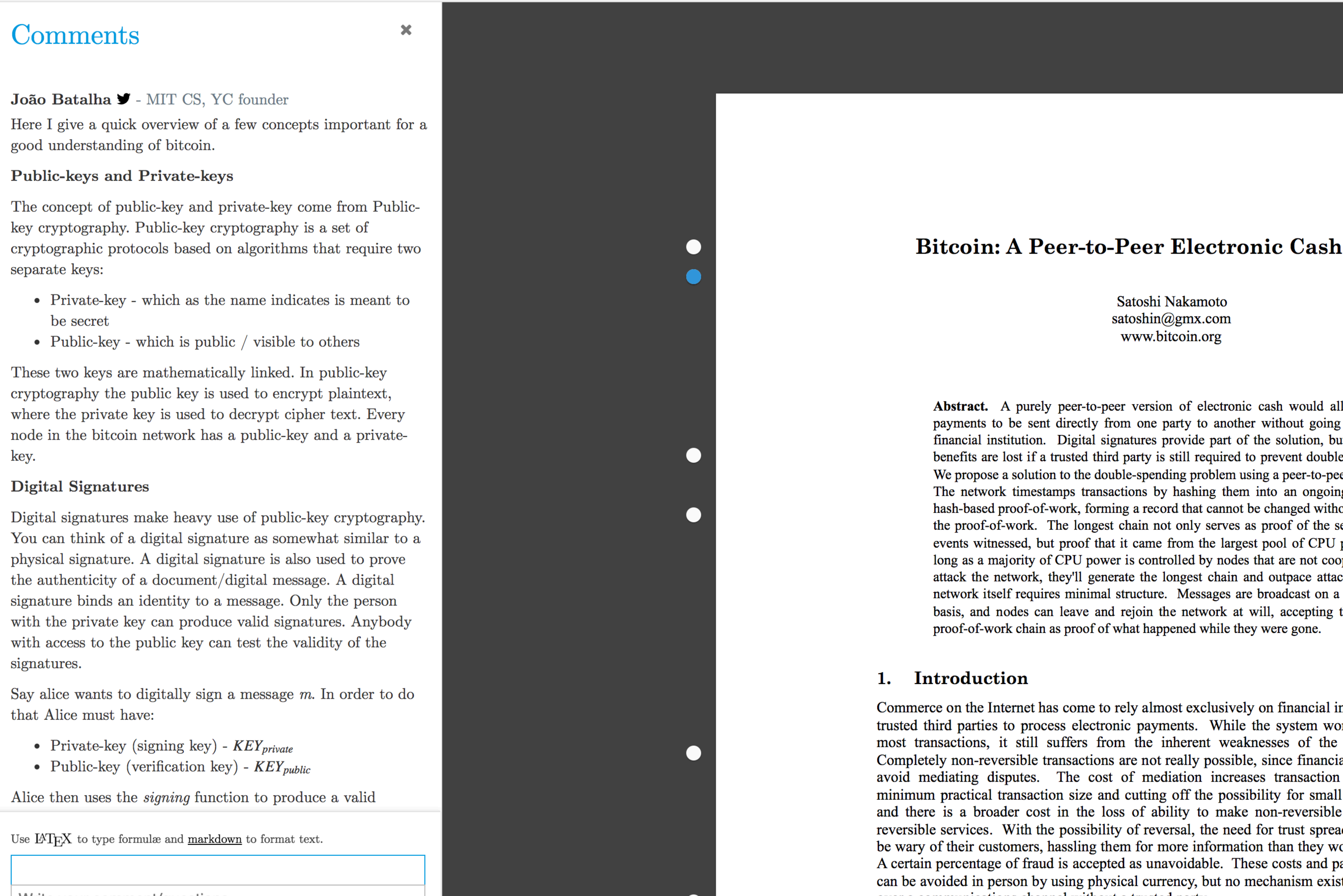
Satoshi Nakamoto
satoshin@gmx.com
www.bitcoin.org

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

What it looks like



Comments

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 formulæ and markdown to format text.

Bitcoin: A Peer-to-Peer Electronic Cash

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

Abstract. A purely peer-to-peer version of electronic cash would all
payments to be sent directly from one party to another without going
financial institution. Digital signatures provide part of the solution, bu
benefits are lost if a trusted third party is still required to prevent double
We propose a solution to the double-spending problem using a peer-to-pee
The network timestamps transactions by hashing them into an ongoin
hash-based proof-of-work, forming a record that cannot be changed witho
the proof-of-work. The longest chain not only serves as proof of the se
events witnessed, but proof that it came from the largest pool of CPU p
long as a majority of CPU power is controlled by nodes that are not coo
attack the network, they'll generate the longest chain and outpace attac
network itself requires minimal structure. Messages are broadcast on a
basis, and nodes can leave and rejoin the network at will, accepting t
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 in
trusted third parties to process electronic payments. While the system wo
most transactions, it still suffers from the inherent weaknesses of the
Completely non-reversible transactions are not really possible, since financia
avoid mediating disputes. The cost of mediation increases transaction
minimum practical transaction size and cutting off the possibility for small
and there is a broader cost in the loss of ability to make non-reversible
reversible services. With the possibility of reversal, the need for trust sprea
be wary of their customers, hassling them for more information than they wo
A certain percentage of fraud is accepted as unavoidable. These costs and pa
can be avoided in person by using physical currency, but no mechanism exist

Motivation

Motivation

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

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

$$\le C\|x^*\|^2 .$$

Motivation

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

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

What trivial really means...

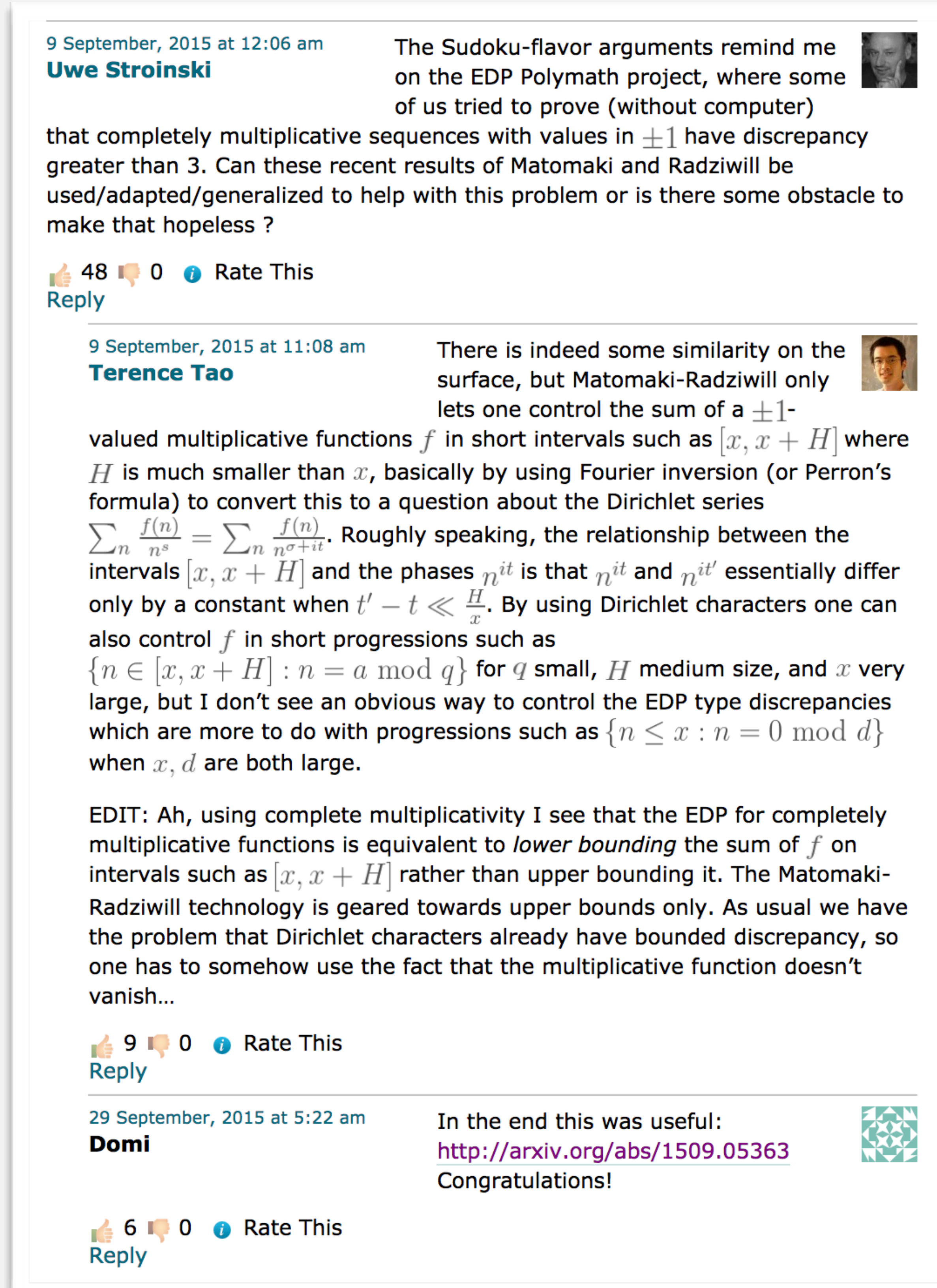

Ts: Jsou-li a, b libovolná přirozená čísla, n prvočíslo, pak číslo $(a+b)^n - a^n - b^n$ je dělitelné číslem $a+b$.
D: Ze zápisu po odečtení a^n a b^n platí:
$(a+b)^n - a^n - b^n = \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \dots + \binom{n}{n-2}a^2b^{n-2} + \binom{n}{n-1}ab^{n-1}$
Protože $\binom{n}{k} = \binom{n}{n-k}$ Spojením stejných binomických koeficientů dostaneme dále uvedené dvojčleny.
$(a+b)^n - a^n - b^n = \left[\binom{n}{1}a^{n-1}b + \binom{n}{n-1}ab^{n-1}\right] + \binom{n}{2}\left[a^{n-2}b^2 + a^2b^{n-2}\right] + \dots + \binom{n}{\frac{n-1}{2}}\left[a^{\frac{n+1}{2}} + b^{\frac{n+1}{2}}\right]$
Poznámka: Při n lichém je těchto dvojčlenů sudý počet
Vytkneme-li z každého z dvojčlenů číslo $(ab)^k$, dostaneme při lichých exponentech dvojčleny dělitelné číslem $a+b$.
Tedy $(a+b)^n - a^n - b^n = ab\cdot\left[\binom{n}{1}a^{n-2} + \binom{n}{n-1}b^{n-2}\right] + \binom{n}{2}a^2b^2\left[a^{n-4} + b^{n-4}\right] + \dots + \binom{n}{k}\cdot a^kb^k\left[a^{n-2k} + b^{n-2k}\right] + \dots + \binom{n}{\frac{n-1}{2}}\left[a^{\frac{n-1}{2}}b^{\frac{n+1}{2}}\right](a+b)$
Z uvedeného vyplývá, že $(a+b)^n - a^n - b^n = (a+b)\cdot U$, kde U je č. přirozené.
Předpokládejme, že existují přirozená čísla x, y, z a přirozené číslo $n > 2$ pro která platí $x^n + y^n = z^n$.
V úvodu uvedeme bez důkazu několik vět, dokazatelných většinou sporem.
Věta č. 1 Čísla x, y, z nemohou být sobě rovna a to ani ve dvojicích
Věta č. 2 Čísla x, y, z musí být nesoudělná; $D(x, y, z) = 1$
Věta č. 3 Ze vztahu $x^n + y^n = z^n$ lze zařídit, aby $x < y < z$ tedy aby existovala přirozená čísla v, d taková, že $z = x + v$, $z = y + d$
Věta č. 4 Protože při různých přirozených číslech x, y, z může být nejmenší vzdálenost čísel x a z číslo 2 a nejmenší vzdálenost čísel x a y číslo 1. pak z této úvahy vyplývá, že $v \geq 2$ a $d \geq 1$.
Věta č. 5 Platí-li pro přirozená nesoudělná čísla vztah $x^n + y^n = z^n$, pak existuje přirozené číslo p (nebo nula) takové, že platí $x + y = z + p$
Věta č. 6 Ze vztahů $z = x + v$, $z = y + d$, $x + y = z + p$ lze odvodit:
1) $x + y = z + p = x + v + p$ tedy $y = p + v$

Motivation

Erdös discrepancy problem (1932)

Motivation

Erdös 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 Radziwill 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-Radziwill 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 $\sum_n \frac{f(n)}{n^s} = \sum_n \frac{f(n)}{n^{\sigma+it}}$. 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 $t' - t \ll \frac{H}{x}$. By using Dirichlet characters one can also control f in short progressions such as $\{n \in [x, x+H] : n = a \bmod q\}$ 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 $\{n \leq x : n = 0 \bmod d\}$ 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-Radziwill 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

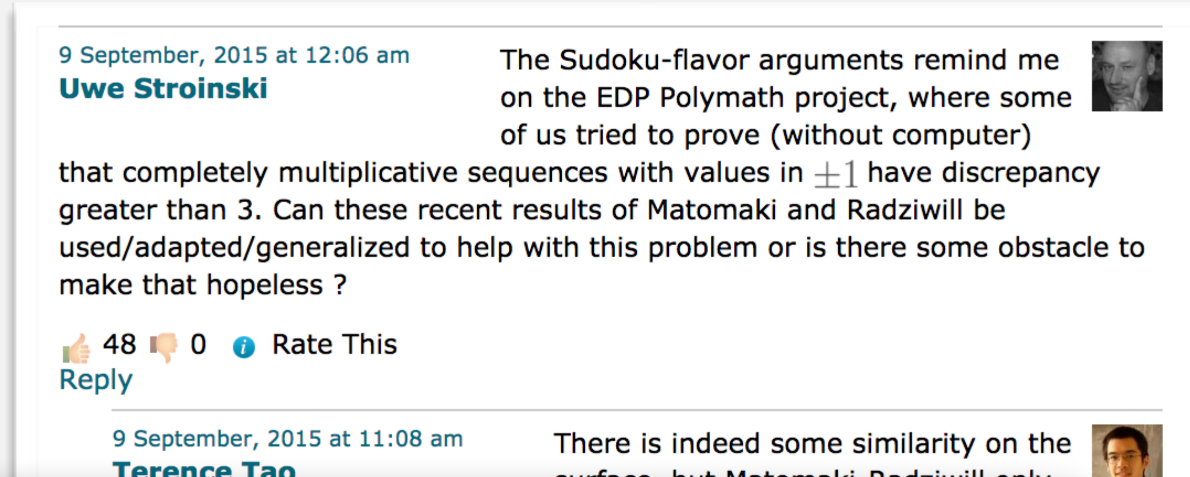
In the end this was useful:
http://arxiv.org/abs/1509.05363
Congratulations!

6 0 Rate This
Reply

From terrytao.wordpress.com

Motivation

Erdös discrepancy problem (1932)



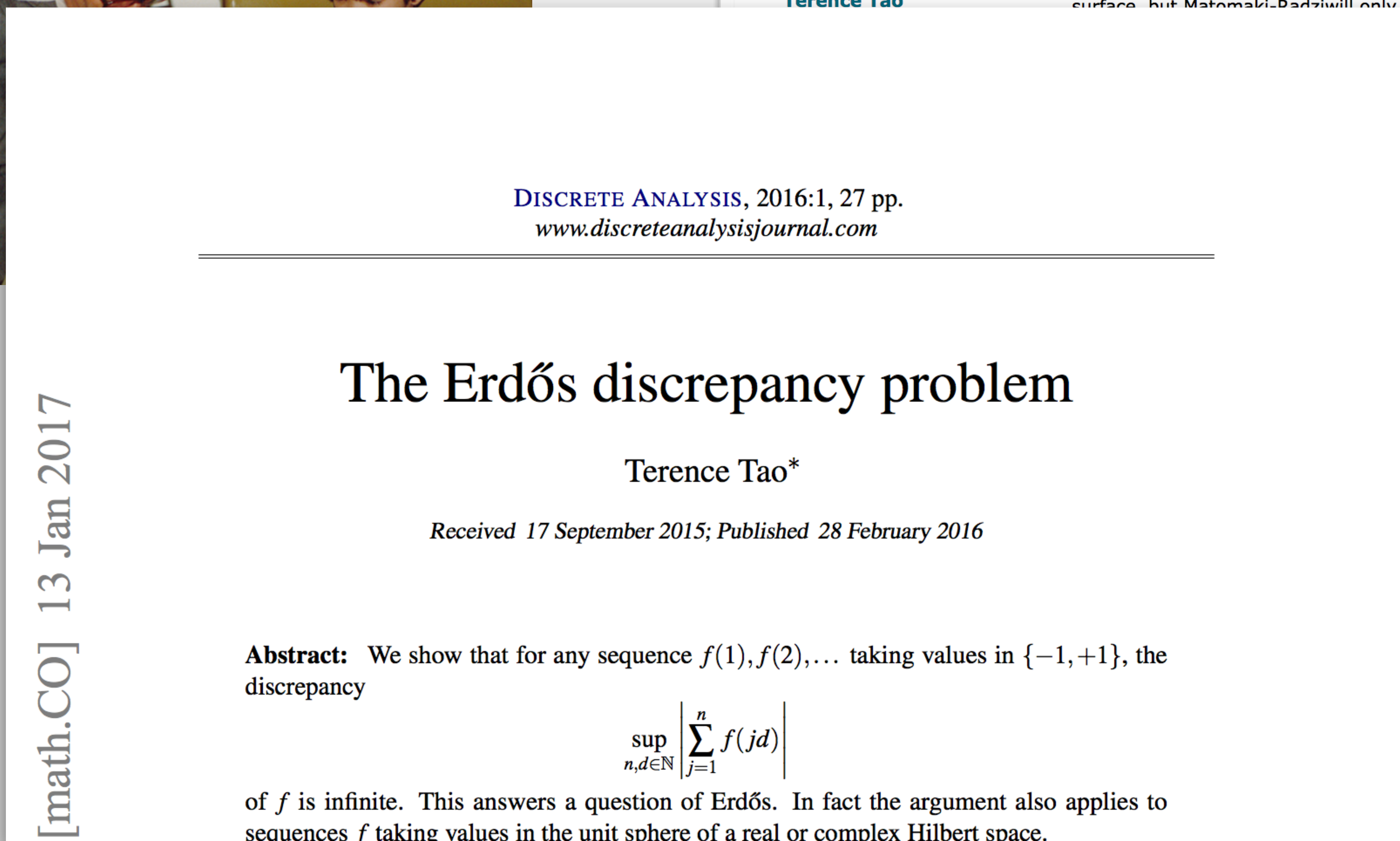
9 September, 2015 at 12:06 am
Uwe Stroinski

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 Radziwill 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



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

[math.CO] 13 Jan 2017

The Erdős discrepancy problem

Terence Tao*

Received 17 September 2015; Published 28 February 2016

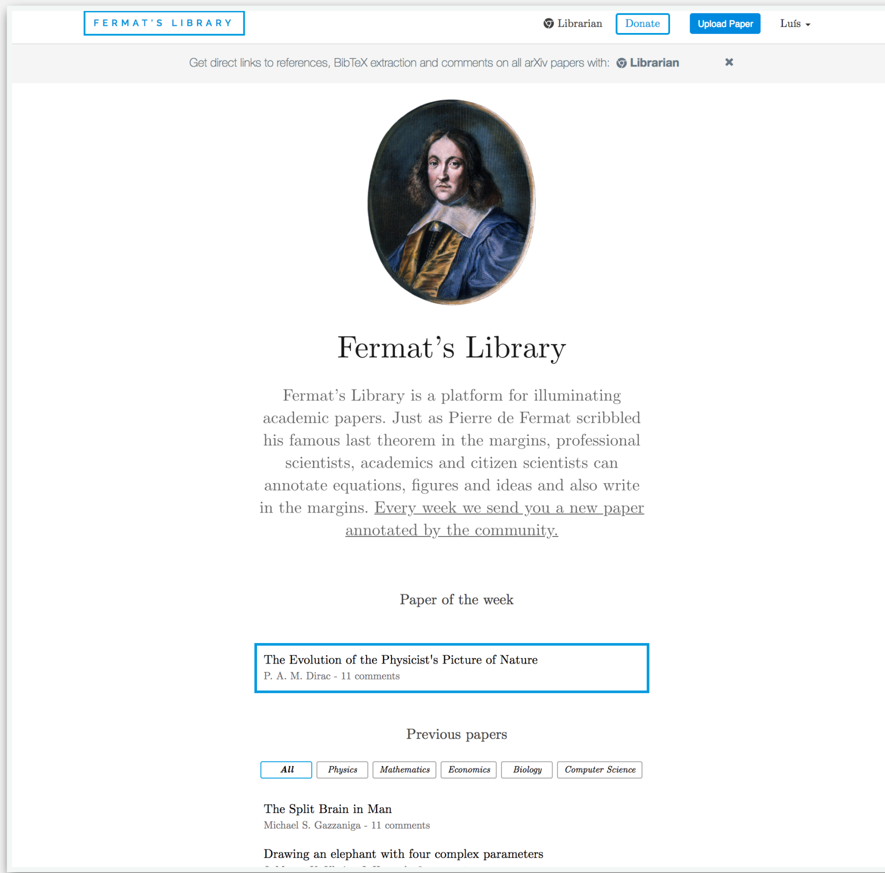
Abstract: We show that for any sequence $f(1), f(2), \ldots$ taking values in $\{-1,+1\}$, the discrepancy

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

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.

How to start a platform

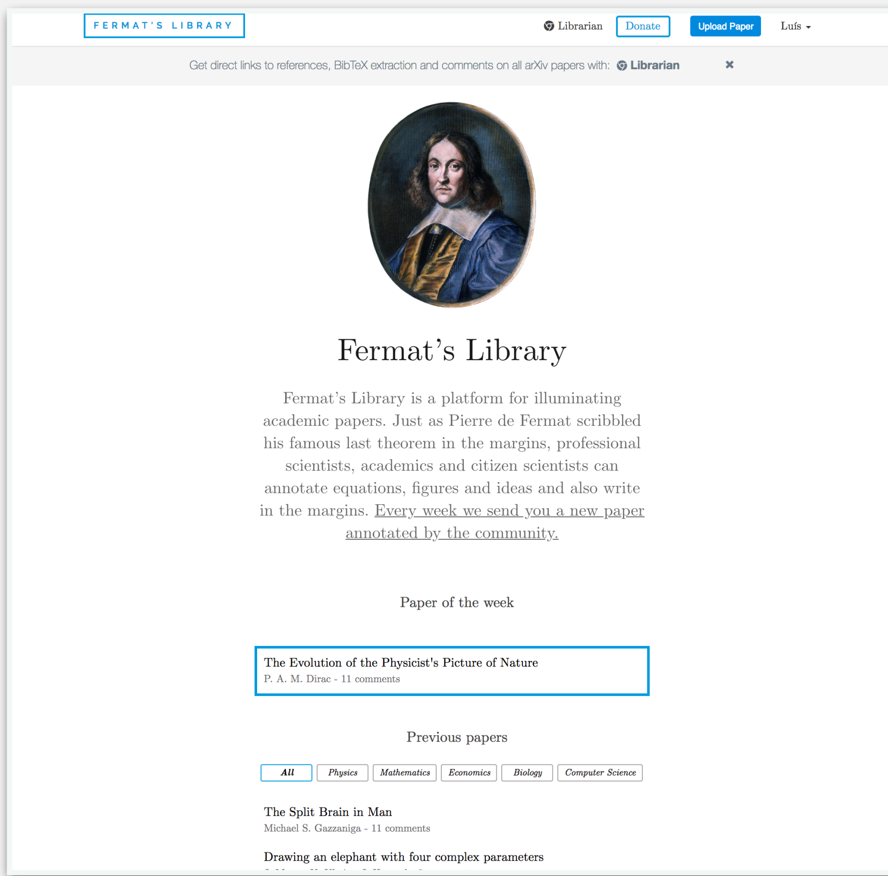
V1 - Journal Club



1 paper per week

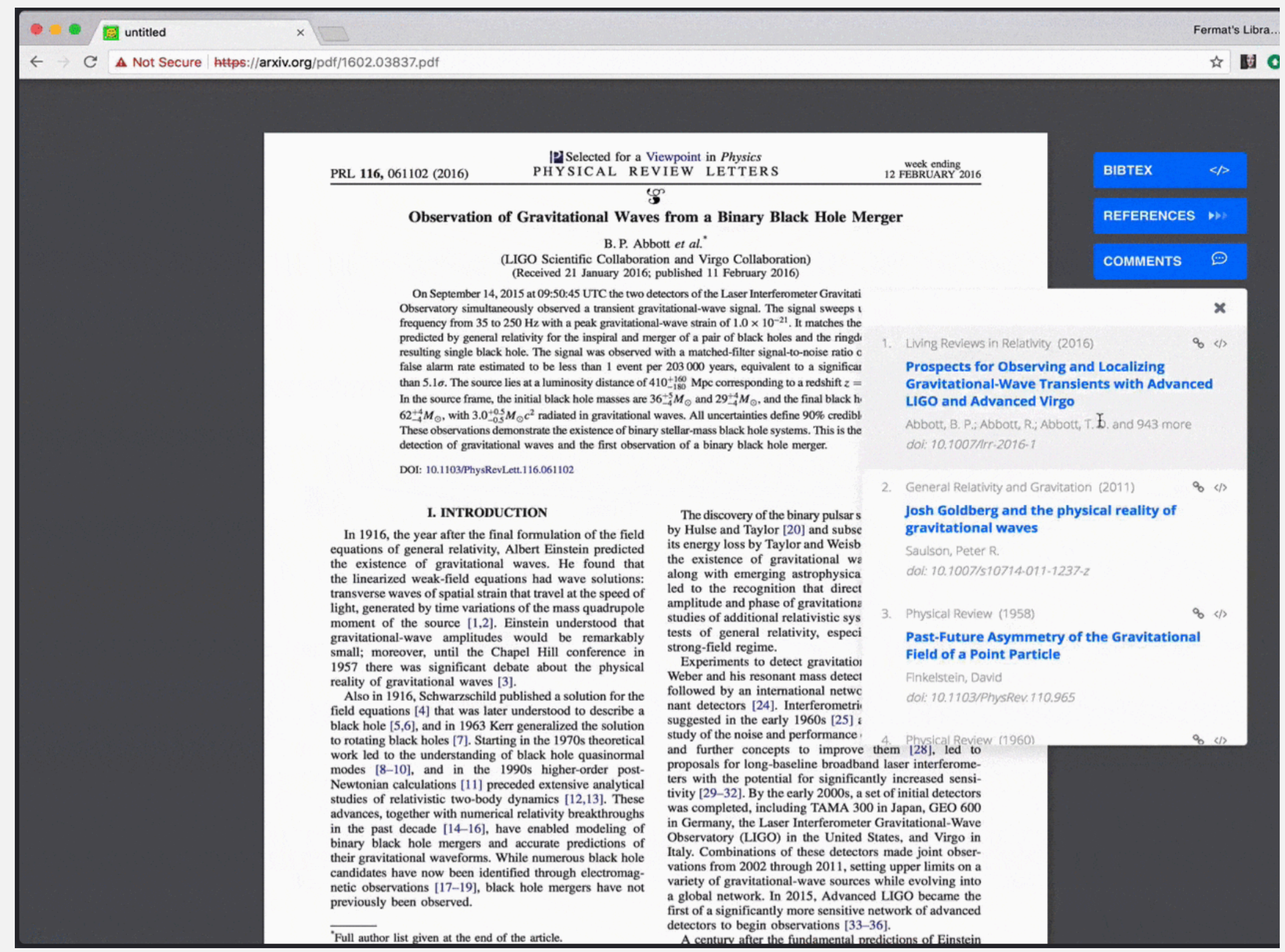
How to start a platform

V1 - Journal Club



1 paper per week

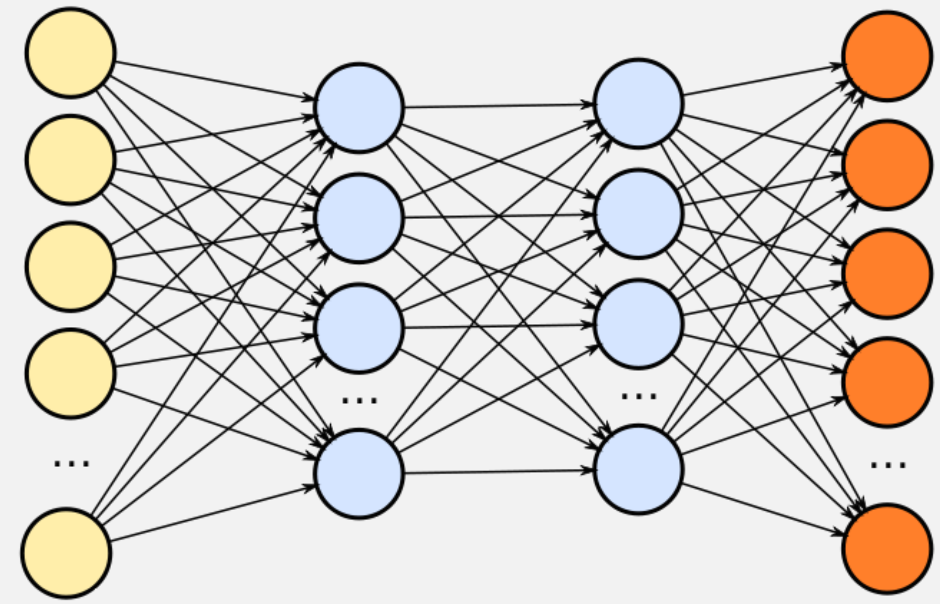
V2 - Platform



- Chrome Extension for arXiv
1.3M Pre-prints
- You can upload your own papers

Data Science Problems

Reference Extraction



References are not uniform!

5. References

References

Baker, G. P., jun 1992. Incentive Contracts and Performance Measurement. Journal of Political Economy 100 (3), 598–614.

Baker, M., may 2016. 1,500 scientists lift the lid on reproducibility. Nature 533 (7604), 452–454.

Begley, C. G., Ellis, L. M., mar 2012. Drug development: Raise standards for preclinical cancer research. Nature 483 (7391), 531–3.

[37] Y. Kubota et al. (CLEO), *Nucl. Instr. and Meth.* A **320** (1992) 66. H. Albrecht et al. (ARGUS), *Nucl. Instr. and Methd.* A **275** (1989) 1.

[38] I. Caprini and M. Neubert, *Phys. Lett.* B **380** (1996) 376; M. Shifman et al., *Phys. Rev.* D **51** (1995) 2217; Erratum-ibid. D52 (1995) 3149; A. Czarnecki, *Phys. Rev. Lett.* **76** (1996) 4124; T. Mannel, *Phys. Rev.* D **50** (1994) 428; A. F. Falk and M. Neubert, *Phys. Rev.* D **47** (1993) 2965 and 2982.

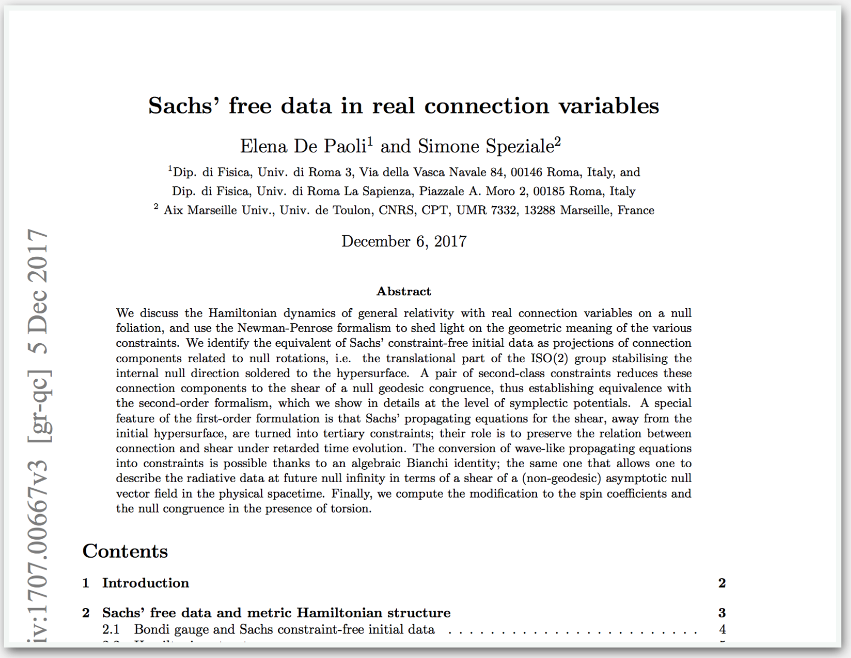
[39] M. Margoni et al., (DELPHI), *Measurement of* V_{cb} *Using the Identified Charged Pion in* $\overline{B}^o \rightarrow D^{*+}\ell^-\bar{\nu}$, DELPHI 98-140, (1998).

Baker, G. P., jun 1992. Incentive Contracts and Performance Measurement. Journal of Political Economy 100 (3), 598–614

Authors | Date | Title | Journal | Issue # | Page

Data Science Problems

Paper Recommendations


iv:1707.00667v3 [gr-qc] 5 Dec 2017

Sachs' free data in real connection variables

Elena De Paoli[1] and Simone Speziale[2]
[1]Dip. di Fisica, Univ. di Roma 3, Via della Vasca Navale 84, 00146 Roma, Italy, and
Dip. di Fisica, Univ. di Roma La Sapienza, Piazzale A. Moro 2, 00185 Roma, Italy
[2] Aix Marseille Univ., Univ. de Toulon, CNRS, CPT, UMR 7332, 13288 Marseille, France

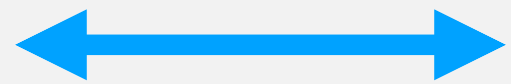
December 6, 2017

Abstract

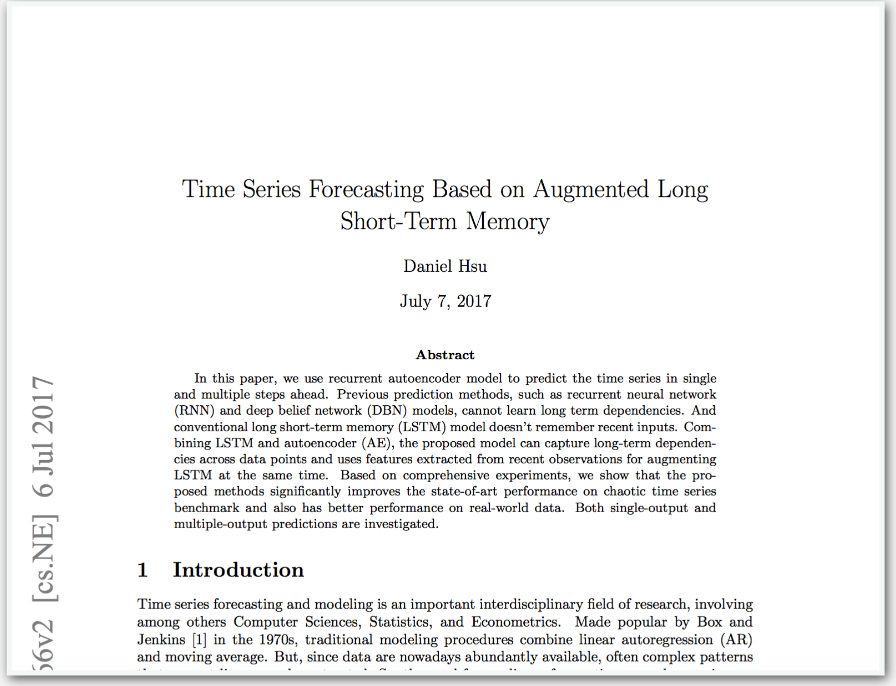
We discuss the Hamiltonian dynamics of general relativity with real connection variables on a null foliation, and use the Newman-Penrose formalism to shed light on the geometric meaning of the various constraints. We identify the equivalent of Sachs' constraint-free initial data as projections of connection components related to null rotations, i.e. the translational part of the ISO(2) group stabilising the internal null direction soldered to the hypersurface. A pair of second-class constraints reduces these connection components to the shear of a null geodesic congruence, thus establishing equivalence with the second-order formalism, which we show in details at the level of symplectic potentials. A special feature of the first-order formulation is that Sachs' propagating equations for the shear, away from the initial hypersurface, are turned into tertiary constraints; their role is to preserve the relation between connection and shear under retarded time evolution. The conversion of wave-like propagating equations into constraints is possible thanks to an algebraic Bianchi identity; the same one that allows one to describe the radiative data at future null infinity in terms of a shear of a (non-geodesic) asymptotic null vector field in the physical spacetime. Finally, we compute the modification to the spin coefficients and the null congruence in the presence of torsion.

Contents

1 **Introduction** 2
2 **Sachs' free data and metric Hamiltonian structure** 3
2.1 Bondi gauge and Sachs constraint-free initial data 4



- Tf-idf vectors of bigrams from full text of each paper followed by L2 lookups for similarity ranking
- Train personalized SVMs for people for recommendations


6v2 [cs.NE] 6 Jul 2017

Time Series Forecasting Based on Augmented Long Short-Term Memory

Daniel Hsu

July 7, 2017

Abstract

In this paper, we use recurrent autoencoder model to predict the time series in single and multiple steps ahead. Previous prediction methods, such as recurrent neural network (RNN) and deep belief network (DBN) models, cannot learn long term dependencies. And conventional long short-term memory (LSTM) model doesn't remember recent inputs. Combining LSTM and autoencoder (AE), the proposed model can capture long-term dependencies across data points and uses features extracted from recent observations for augmenting LSTM at the same time. Based on comprehensive experiments, we show that the proposed methods significantly improves the state-of-art performance on chaotic time series benchmark and also has better performance on real-world data. Both single-output and multiple-output predictions are investigated.

1 Introduction

Time series forecasting and modeling is an important interdisciplinary field of research, involving among others Computer Sciences, Statistics, and Econometrics. Made popular by Box and Jenkins [1] in the 1970s, traditional modeling procedures combine linear autoregression (AR) and moving average. But, since data are nowadays abundantly available, often complex patterns

Better paper relevance classification

- Other links to papers
- User Rating
- Citations

Future Perspectives

Growth in online collaboration

Open discussion

More knowledge sharing

Thank You!

Luis Batalha, João Batalha, Micael Oliveira, Tymor Hamamsy

team@fermatslibrary.com