PHYSICS IN THE WORLD OF IDEAS:

COMPLEXITY AS ENERGY

Yuri I. Manin

PLAN

PART I: MODERN META–PHYSICS: STRUCTURE OF PHYSICAL LAWS

PART II: KOLMOGOROV COMPLEXITY

PART III: ZIPF's LAW AS "MINIMIZATION OF EFFORT"

PART IV: ERROR–CORRECTING CODES AND PHASE TRANSITIONS

PART V: COMPLEXITY IN QUANTUM COMPUTING ?

PART I. MODERN META–PHYSICS: STRUCTURE OF PHYSICAL LAWS

- **An isolated system** : configuration/phase space X.
- **Energy/action** : function(al)s $E, S : X \to \mathbf{R}$.
- **Classical partition function/quantum evolution** :

$$Z_T := \int_X e^{-E(x)/T} Dx \qquad \text{vs} \qquad Z = \int_X e^{itS(x)} Dx$$

- **Probability density/quantum evolution operator** :

$$\frac{1}{Z_T} e^{-E(x)/T} Dx \qquad \text{vs} \qquad \frac{1}{Z} e^{itS(x)}$$

NB Inverse temperature T <=> imaginary time it !

- **Symmetries, scale invariance etc.**

PART II: KOLMOGOROV COMPLEXITY

Zoo of complexities :

Logarithmic		**of combinatorial objects**
	Kolmogorov complexity	
Exponential		**of computable functions**

An intuitive description :

- **Logarithmic Kolmogorov complexity of** ω **is defined as**

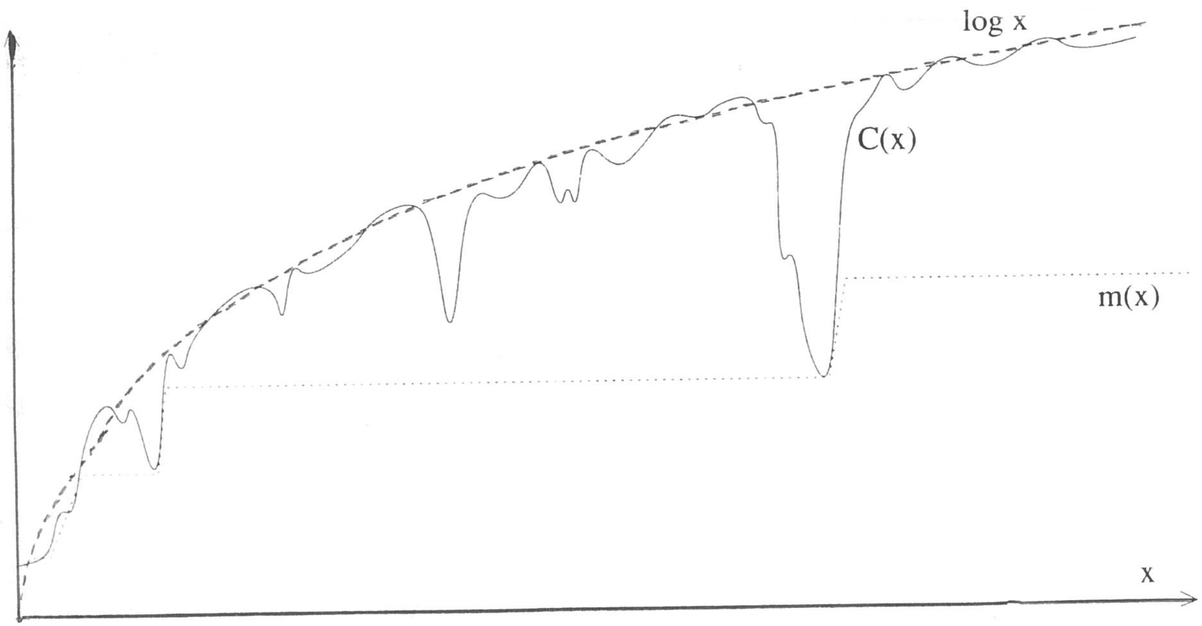
---

the measure of compressibility of ω :=
the length of the shortest program that can generate ω

A representative example :

ω := an integer $N > 0$, represented by N written dashes.

The length of its compressed description is $\leq \log_2 N + c$, **and its exponential Kolmogorov complexity is** $\leq C\,N$.



GRAPH OF THE Log–COMPLEXITY OF NATURALS

Symmetry group and fractality

• **Symmetry group S_∞^{rec}: for any totally recursive permutation $\sigma : \mathbf{Z}_+ \to \mathbf{Z}_+$, there exists a constant $c = c(\sigma)$ such that difference of log–complexities of x and $\sigma(x)$ is $< c$.**

• **Fractality: For any infinite decidable subset $D \subset \mathbf{Z}_+$, the graph of log–complexity restricted upon D has, up to additive $O(1)$, the same form as the total graph.**

Example: $D := \{n^{n^{\cdot^{\cdot^{\cdot^{n}}}}}\ (n \text{ times}) \,|\, n = 1, 2, 3, \dots \}$

• **The standard application of symmetry: one can define complexity for any objects of any infinite "constructive world" X, for example, a language in the sense of comp. sci.**

X comes with a computable numbering, and arbitrariness in its choice (almost) does not influence the size of complexity.

• **Exponential complexity and Kolmogorov order.**

Let X be a constructive world. For any (semi)–computable function $u:\mathbf{Z}_+\to X$, the (exponential) complexity of an object $x\in X$ relative to u is

$$K_u(x) := \min\{m\in\mathbf{Z}_+\,|\,u(m)=x\}.$$

If such m does not exist, we put $K_u(x)=\infty$.

• **CLAIM: there exists such u ("an optimal Kolmogorov numbering", or "decompressor") that for each other v, some constant $c_{u,v} > 0$, and all $x \in X$,**

$$K_u(x) \leq c_{u,v} K_v(x).$$

This $K_u(x)$ is called exponential Kolmogorov complexity of x.

A Kolmogorov order of a constructive world X is a bijection $\mathbf{K} = \mathbf{K}_u : X \to \mathbf{Z}$ arranging elements of X in the increasing order of their complexities K_u.

• **WARNINGS :**

– **Any optimal numbering is only partial function, and its definition domain is not decidable.**

– **Kolmogorov complexity K_u itself is not computable. It is the lower bound of a sequence of computable functions.**

– **Kolmogorov order of $\mathbf{Z}_+$ cardinally differs from the natural order in the following sense: it puts in the initial segments very large numbers that can be at the same time Kolmogorov simple.**

– **Example: let $a_n := n^{n^{\cdot^{\cdot^{\cdot^{n}}}}}$ (n times).**

Then $K_u(a_n) \leq cn$ for some $c > 0$.

• **MY CENTRAL ARGUMENT IN THIS TALK:**

I will argue that there are natural observable and measurable phenomena in the world of information that can be given a mathematical explanation, if one postulates that logarithmic Kolmogorov complexity plays a role of energy.

I will consider two examples: Zipf's Law and asymptotic bounds in the theory of error–correcting codes.

PART III: ZIPF LAW AS “MINIMIZATION OF EFFORT”

• Consider a corpus of texts in a given language, make the list of words occurring in them and the numbers of occurrences. Range these words in the order of *diminishing* frequencies. Define the Zipf rank of a word as its number in this ordering.

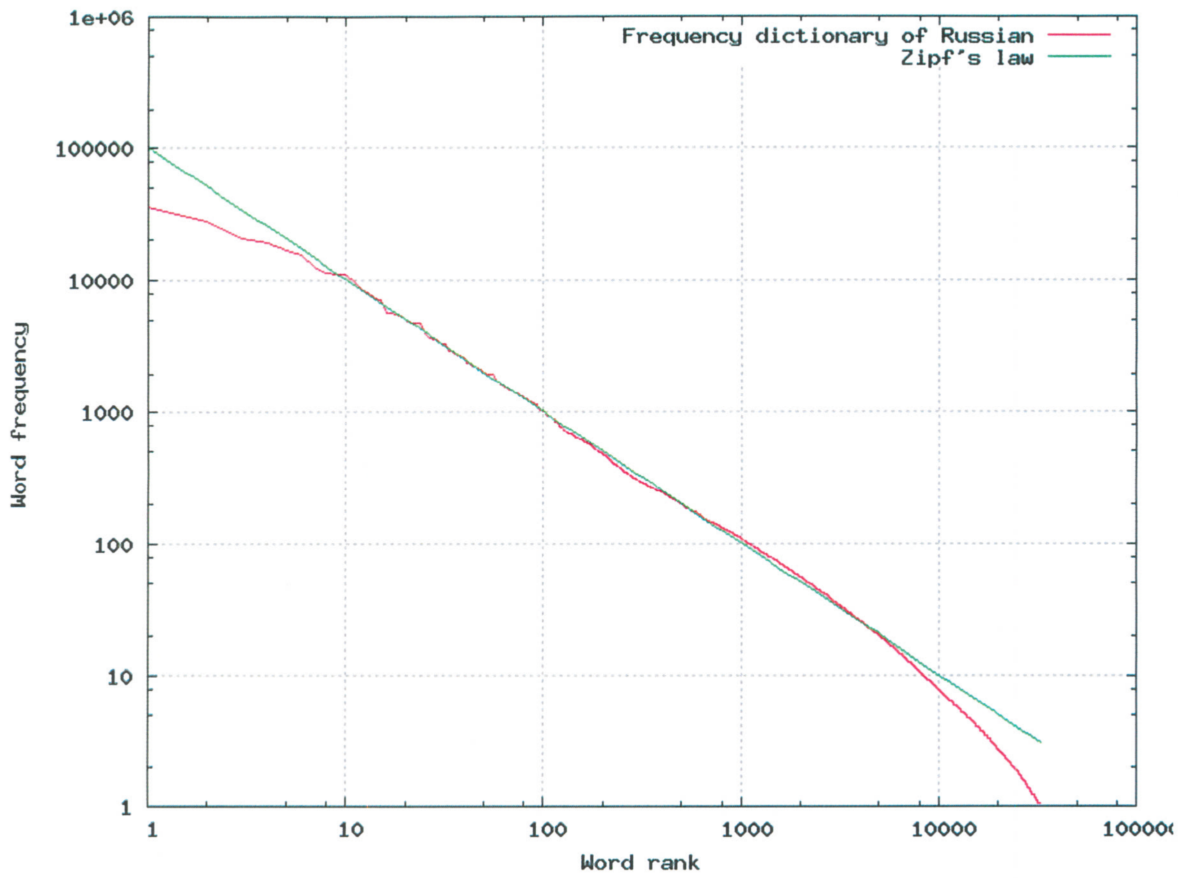
• *Zipf's Law* (1935, 1949) :

FREQUENCY

IS INVERSELY PROPORTIONAL TO THE RANK

PICTURE:

Zipf's distribution of Russian words(logarithmic scale)



• **<u>Universality of Zipf′s law</u> : the law is empirically observed in very different databases, that allow one to calculate frequency of occurrence of certain** *patterns* **("words") in certain** *massifs of data.*

• **Example on the next page: patterns in** *financial audit data.*

• *"Unlike the central limit theorem [. . .] this law is primarily an empirical law; it is observed in practice, but mathematicians still do not have a fully satisfactory and convincing explanation for how the law comes about, and why it is so universal. "*

Terence Tao

Decision Support Systems 46 (2008) 70–83

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

An investigation of Zipf's Law for fraud detection (DSS#06-10-1826R(2))

Shi-Ming Huang [a], David C. Yen [b,*], Luen-Wei Yang [a], Jing-Shiuan Hua [c]

[a] *Department of Accounting & Information Technology, National Chung-Cheng University, Chia-Yi, Taiwan, ROC*
[b] *Department of DSC & MIS, Miami University, Oxford, OH 45056, United States*
[c] *Department of Information Management, National Chung-Cheng University, Chia-Yi, Taiwan, ROC*

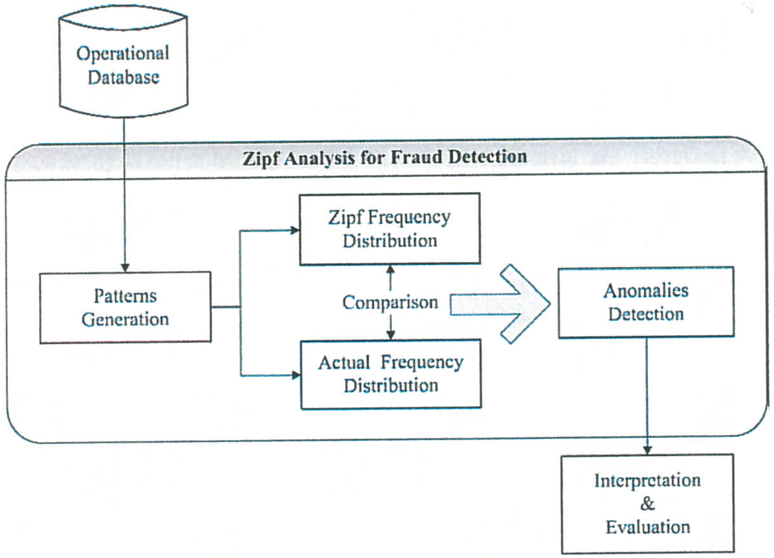


Fig. 1. Fraud detection model of Zipf Analysis.

ZIPF's RANK AND ZIPF's LAW FROM COMPLEXITY

• I suggest that (at least in some situations) Zipf's law emerges as the combined effect of two factors:

(A) Rank ordering coincides with the ordering with respect to the growing (exponential) Kolmogorov complexity $K(w)$ *up to a factor* $exp\,(O(1))$.

(B) The probability distribution producing Zipf's law is (an approximation to) the L. Levin a priori distribution.

If we accept (A) and (B), then Zipf's law follows from two basic properties of Kolmogorov complexity:

(a) rank of w defined according to (A) is $exp\,(O(1)) \cdot K(w)$.

(b) L. Levin's a priori distribution assigns to an object w probability $\sim KP(w)^{-1}$ *where KP is the exponentiated prefix Kolmogorov complexity, and we have, up to* $exp\,(O(1))$*–factors,*

$$K(w) \preceq KP(w) \preceq K(w) \cdot \log^{1+\varepsilon} K(w)$$

with arbitrary $\varepsilon > 0$.

• NB A probability distribution on infinity of objects cannot be constructed directly from K: the series $\sum_m K(m)^{-1}$ diverges. However, on finite sets of data the small discrepancy is additionally masked by the dependence of both K and KP on the choice of an optimal encoding.

• COMPLEXITY AS EFFORT. The picture described above agrees with Zipf's motto "minimization of effort", but reinterprets the notion of effort: its role is now played by the logarithm of the Kolmogorov complexity that is, by the length of the maximally compressed description of an object.

Such a picture makes sense especially if the objects satisfying Zipf's distribution, are *generated* **rather than simply** *observed.*

PART IV: ERROR–CORRECTING CODES AND PHASE TRANSITIONS

• BASIC NOTATION:

Alphabet A:= **a finite set of cardinality** $q \geq 2$.

Code $C \subset A^n$:= **a subset of words of length** n.

Hamming distance between two words:

$$d((a_i),(b_i)) := \mathrm{card}\{i \in (1,\dots,n) \mid a_i \neq b_i\}.$$

Code parameters: **cardinality of the alphabet q and**

$$n(C) := n, \quad k(C) := k := [\log_q \mathrm{card}(C)],$$

$$d(C) := d = \min\{d(a,b) \mid a, b \in C, a \neq b\}.$$

Relative distance and Transmission rate :

$$\delta(C) := \frac{d(C)}{n(C)}, \quad R(C) = \frac{k(C)}{n(C)}.$$

Briefly, C is an $[n,k,d]_q$–code.

SOURCE DATA

Encoding: ↓

A sequence of code words

Noisy channel: ↓

Sequence of (corrupted) code words

Error correction: ↓

(Ideally) sequence of initial code words

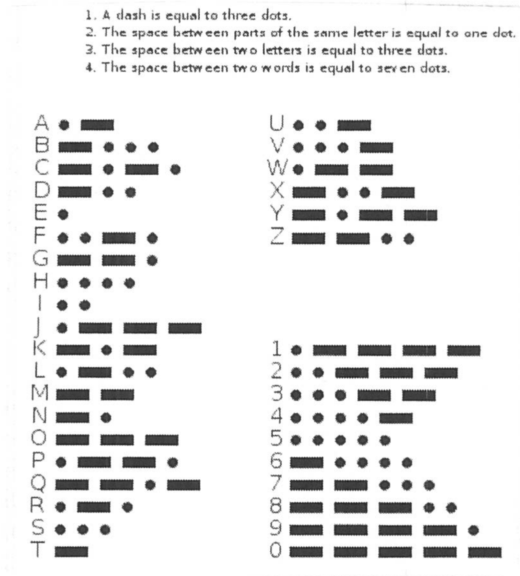
Decoding: ↓

TRANSMITTED DATA

Examples: Morse code and Barcodes

Samuel F. B. Morse
(from 1836 on)

Alphabet: $\{dash,\ dot,\ space\}$,
$q = 3$.
Block length: $n = 7$
$d =?$: (*Exercise*)



Norman J. Woodland
(from 1949 on):
"His [...]inspiration came from Morse code, and he formed his first barcode from sand on the beach.

I just extended the dots and dashes downwards and made narrow lines and wide lines out of them"

ISBN 978-3-16-148410-0

9 783161 484100

A 13-digit ISBN, 978-3-16-148410-0, as represented by an EAN-13 bar code.

- **Explaining terms:**

(Minimal) relative distance and Transmission Rate :

$$\delta(C) := \frac{d(C)}{n(C)}, \quad R(C) = \frac{[k(C)]}{n(C)}.$$

Minimal Relative Distance **must match channel's** *noisiness*: **probability of corruption of one letter.**

Transmission rate **is the share of** *meaningful* **(code) words; their number must be maximized for any given relative distance.**

• **<u>A good code</u> must maximize minimal relative distance when the transmission rate is chosen.**

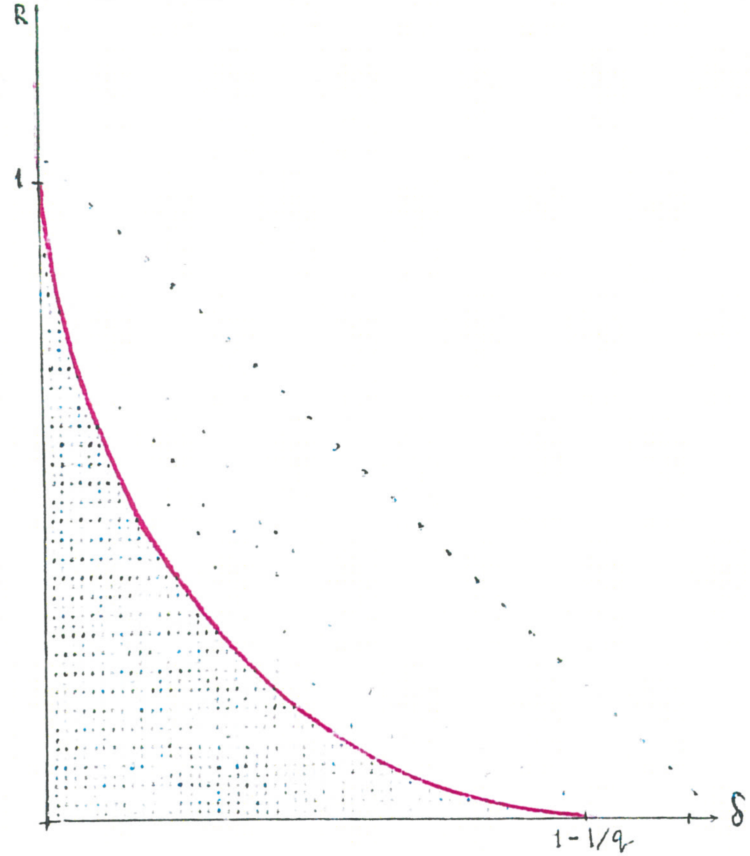
• **One more property of good codes: they must admit** *efficient algorithms of encoding, decoding and error–correction.*

How this can be achieved: consider structured codes. Typical choice:

• **<u>Linear codes</u> := linear subspaces of $\mathbf{F}_q^n$.**

- **Code points:**

$$[n,k,d]_q - code\ C \mapsto P_C := (R(C), \delta(C)) = \left(\frac{[k(C)]}{n(C)}, \frac{d(C)}{n(C)}\right)$$



How a finite pixel plot of all code points might look (q fixed)

Explanations to the picture :

• **DEFINITION.** *Multiplicity* **of a code point is the number of codes that project onto it.**

• **THEOREM (Yu.M.,1981 + 2011). There exists such a continuous function $\alpha_q(\delta)$, $\delta \in [0,1]$, that**

(i) The set of code points of infinite multiplicity is exactly the set of rational points $(R,\delta) \in [0,1]^2$ satisfying $R \leq \alpha_q(\delta)$.

The curve $R = \alpha_q(\delta)$ is called the asymptotic bound.

(ii) Code points x of <u>finite multiplicity</u> all lie above the asymptotic bound and are called <u>isolated ones</u>:

for each such point there is an open neighborhood containing x as the only code point.

(iii) The same statements are true for linear codes, with, a possibly, different asymptotic bound $R = \alpha_q^{lin}(\delta)$.

ASYMPTOTIC BOUNDS FROM COMPLEXITY

• **Oracle assisted approximate computation of the asymptotic bound.**

– The set $Codes_q$ of all q–ary codes in a fixed alphabet A is a constructive world.

– CLAIM. If an oracle produces for us elements of $Codes_q$ in their Kolmogorov order, then we can write an oracle assisted algorithm that for each "pixel size" N^{-1} enumerates all code points of the form

$$(k/N, d/N), \quad a, d \in \mathbf{Z}_+$$

CF. PICTURE ON PAGE 24

• <u>**Partition function for codes involving complexity.**</u>

– The function $\alpha_q(\delta)$ is continuous and strictly decreasing for $\delta \in [1, 1 - q^{-1})$.

Hence the limit points domain $R \le \alpha_q(\delta)$ can be equally well described by the inequality $\delta \le \beta_q(R)$ where β_q is the function inverse to α_q.

– Fix an $R \in \mathbf{Q} \cap (0,1)$. For $\Delta \in \mathbf{Q} \cap (0,1)$, put

$$Z(R, \Delta; \beta) := \sum_{C:\, R(C)=R, \Delta \le \delta(C) \le 1} K_u(C)^{-\beta + \delta(C) - 1},$$

where K_u is an exponential Kolmogorov complexity on $Codes_q$.

• **Theorem.** (i) If $\Delta > \beta_q(R)$, then $Z(R, \Delta; \beta)$ is a real analytic function of β.

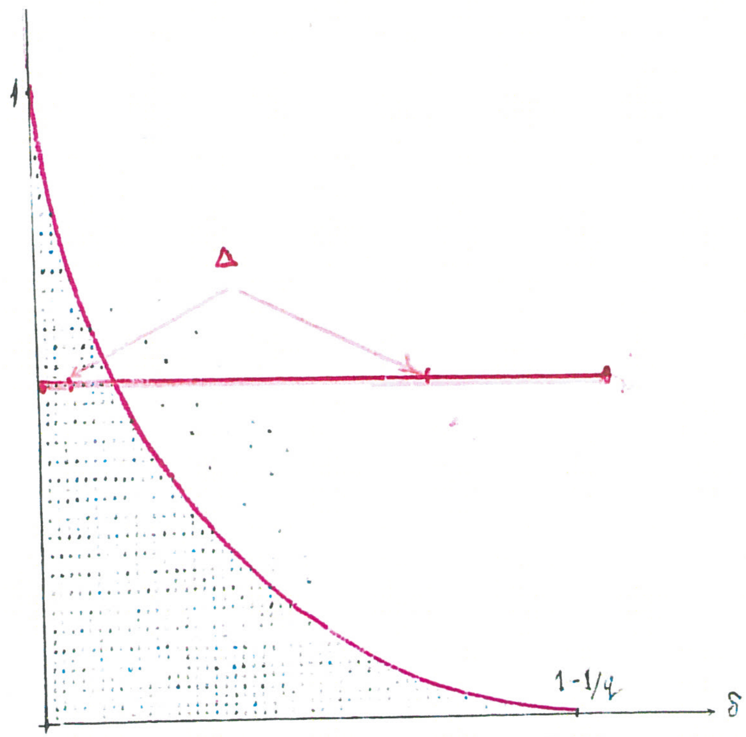
(ii) If $\Delta < \beta_q(R)$, then $Z(R, \Delta; \beta)$ is a real analytic function of β for $\beta > \beta_q(R)$ such that its limit for $\beta - \beta_q(R) \to +0$ does not exist.

• **Thermodynamical analogies.**

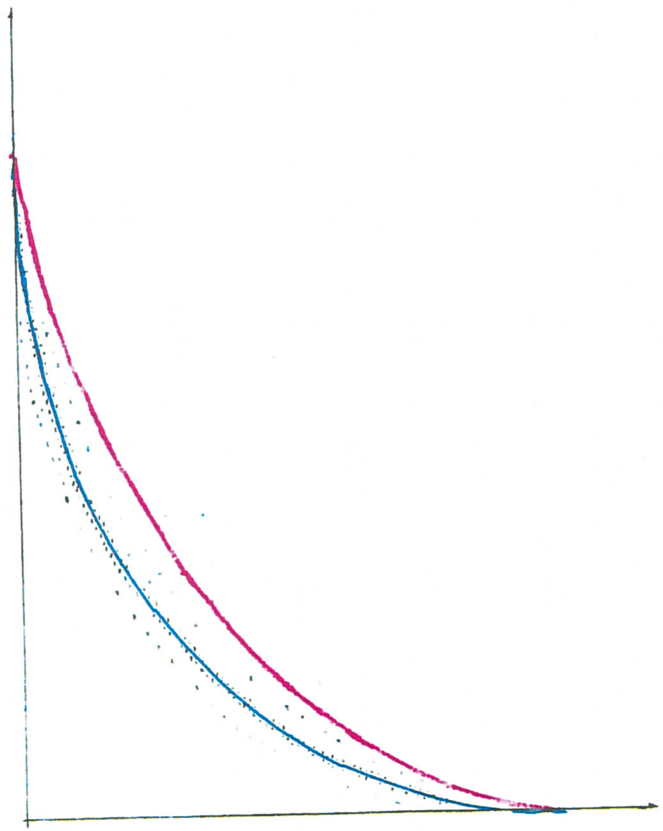
– The argument β of the partition function corresponds to the inverse temperature.

– The transmission rate R corresponds to the density ρ.

– Our asymptotic bound transported into $(T = \beta^{-1}, R)$–plane as $T = \beta_q(R)^{-1}$ becomes the phase transition boundary in the (temperature, density)–plane.


1
Δ
1 - 1/q
δ

Can we see the asymptotic bound
plotting the set of (linear) code points of bounded size?



NO, we will see a cloud of points
concentrating near the Varshamov–Gilbert bound

PART V. COMPLEXITY IN QUANTUM COMPUTING ?

REFERENCES

[A] S. Aaronson. *NP–complete problems and physical reality.* arXiv:quantum-ph/0502072

[Lev] L. A. Levin, *Various measures of complexity for finite objects (axiomatic description)*, Soviet Math. Dokl. Vol.17 (1976) N. 2, 522–526.

[LiVi] M. Li, P. M. B. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications*, 3rd edn. Springer, New York, 2008.

[Man1] Yu. I. Manin, *What is the maximum number of points on a curve over* $\mathbf{F}_2$*?* J. Fac. Sci. Tokyo, IA, Vol. 28 (1981), 715–720.

[Man2] Yu. I. Manin. *Zipf's law and L. Levin's probability distributions.* Functional Analysis and its Applications, vol. 48, no. 2, 2014. DOI 10.107/s10688-014-0052-1. arXiv:1301.0427

[Man3] Yu. I. Manin, *A course in Mathematical Logic for Mathematicians*, 2nd ed., Springer, New York, 2010.

[ManMar] Yu. I. Manin, M. Marcolli. *Error–correcting codes and phase transitions.* Mathematics in Computer Science, Vol. 5 (2011) 133-170. arXiv:0910.5135

[ManVla] Yu. I. Manin. S.G. Vladut, *Linear codes and modular curves*. J. Soviet Math., Vol. 30 (1985), 2611–2643.

[VlaNoTsfa] S. G. Vladut, D. Yu. Nogin, M. A. Tsfasman. *Algebraic geometric codes: basic notions.* Mathematical Surveys and Monographs, 139. American Mathematical Society, Providence, RI, 2007.

[Z1] G. K. Zipf. *The Psychobiology of Language.* Houghton-Mifflin, 1935

[Z2] G. K. Zipf. *Human Behavior and the Principle of Least Effort.* Addison-Wesley, 1949.