Codes and Complexity

Matilde Marcolli

Ma148: Algebraic and Categorical Aspects of Information
Caltech, Winter 2025

This lecture is based on:

- Yuri I. Manin, Matilde Marcolli, *Error-correcting codes and phase transitions*, Mathematics in Computer Science (2011) 5:133–170.
- Yuri I. Manin, Matilde Marcolli, *Kolmogorov complexity and the asymptotic bound for error-correcting codes*, Journal of Differential Geometry, Vol.97 (2014) 91–108
- Yuri I. Manin, Matilde Marcolli, *Asymptotic bounds for spherical codes*, arXiv:1801.01552
- M. Mezard, A. Montanari, *Information, Physics and Computation*, Oxford University Press, 2009.

Coding and information

- source of information: random variable $\mathcal{X}$ with values in a finite alphabet $\mathfrak{A}$ generating a sequence of symbols
- $\mathfrak{A}^*$ all finite sequences (arbitrary length) in the alphabet $\mathfrak{A}$
- $\mathfrak{A}^N$ all sequences of length N
- Problem: store the information contained in a given sequence $\underline{x} \in \mathfrak{A}^N$ in the most compact way
- source coding: a source code for the random variable $\mathcal{X}$ with a reference alphabet (say $\{0,1\}$ case of a *binary* code)

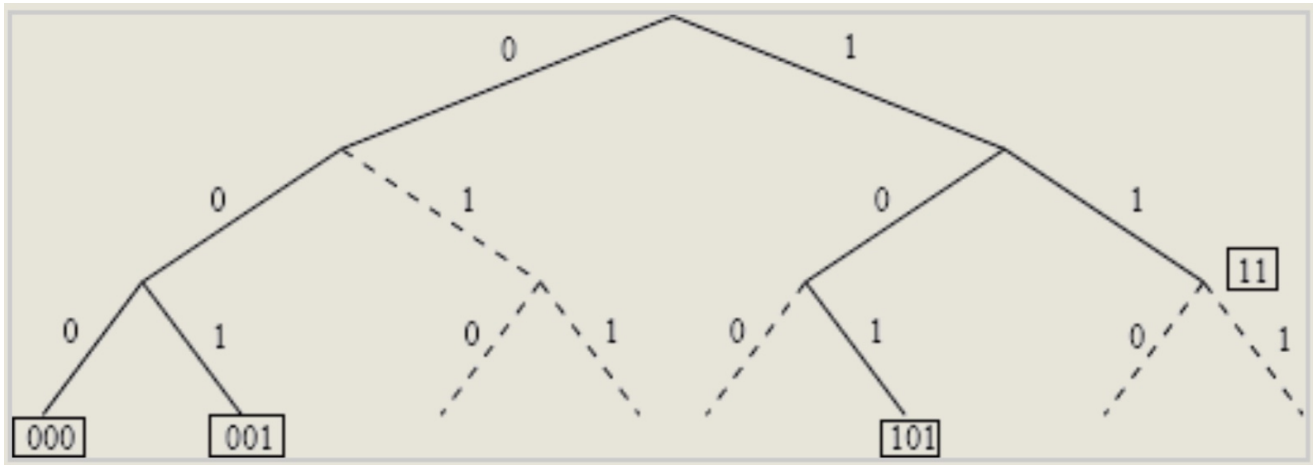
$$E : \mathfrak{A}^N \to \{0,1\}^* \quad \underline{x} \mapsto E(\underline{x}) \text{ codewords}$$

- stream of outputs of random variable $\mathcal{X}$: break into blocks in $\mathfrak{A}^N$ and apply encoding E to blocks, get sequence of codewords

$$\underline{x}_0\underline{x}_1\underline{x}_2 \cdots \underline{x}_n \cdots \mapsto E(\underline{x}_0)E(\underline{x}_1)E(\underline{x}_2) \cdots E(\underline{x}_n) \cdots$$

decoding

- usually more than one way of parsing this concatenation into codewords: ambiguities
- need code that avoids problem: any concatenation of codewords can be parsed unambiguously
- uniquely decodable code
- requirement: for any $\underline{x}, \underline{x}' \in \mathfrak{A}^N$, the codeword $E(\underline{x})$ is not a prefix of $E(\underline{x}')$: called instantaneous codes



Example of instantaneous source code: each codeword assigned to a node in a binary tree so that none is an ancestor of another

average length of encoding

- how good is a source code used to store information from a source $\mathcal{X}$?
- $\ell_E(\underline{x})$ length of the string $E(\underline{x})$
- average length

$$L(E) := \sum_{\underline{x} \in \mathfrak{A}^N} p(\underline{x})\, \ell_E(\underline{x})$$

- $p(\underline{x})$ probability that the random variable $\mathcal{X}$ produces the string $\underline{x}$
- measure of efficiency of code: a code can achieve a shorter average length by assigning shorter codewords $E(\underline{x})$ to strings $\underline{x}$ that occur more frequently (higher probability) and longer code words to sequences occurring more rarely
- can this be optimized?

optimal average length

- random variable $\mathcal{X}$ with Shannon entropy

$$S(\mathcal{X}) = -\sum_{x} \mathbb{P}(\mathcal{X} = x) \log \mathbb{P}(\mathcal{X} = x)$$

- L_N shortest average length achieved by instantaneous codes
- for all $N \geq 1$ and $\mathcal{X}_N$ with $\underline{x} \in \mathfrak{A}^N$ outputs

$$S(\mathcal{X}_N) \leq L_N \leq S(\mathcal{X}_N) + 1$$

- if source has finite entropy rate

$$\lim_{N\to\infty} \frac{S(\mathcal{X}_N)}{N} = \sigma < \infty$$

then also

$$\lim_{N\to\infty} \frac{L_N}{N} = \sigma$$

Shannon codes

- in an automatic binary code can always represent code words as leaves of a binary tree
- Kraft inequality follows

$$\sum_{\underline{x} \in \mathfrak{A}^N} 2^{-\ell_E(\underline{x})} \leq 1$$

(erase all descendants as cannot be other codewords in automatic code: total number of erased descendants $\leq$ total number of descendants)

- any set of lengths $\{\ell(\underline{x})\}$ satisfying Kraft inequality is set of lengths of an automatic binary code
- minimize average length over all $\{\ell(\underline{x})\}$ with Kraft inequality
- Lagrange multipliers $\Rightarrow \ell(\underline{x}) = -\log_2 p(\underline{x})$
- these minima may not be realizable as some not integers
- but give average length equal of Shannon entropy (lower bound $S(\mathcal{X}_N)$)
- realizable $\ell'(\underline{x}) = \lceil -\log_2 p(\underline{x}) \rceil$ give upper bound $S(\mathcal{X}_N) + 1$

Channel coding: information transmission

- redundancy helps correct some transmission errors
- level of redundancy related to maximal level of noise tolerated for error-free transmission
- here encoder is a map $E : \{0,1\}^M \to \{0,1\}^N$ with $N > M$
- channel C described by a transition probability $\mathbb{P}_C(\underline{y}|\underline{x})$ where $\underline{y} \in \{0,1\}^N$ what is received and $\underline{x} \in \{0,1\}^N$ what was transmitted
- decoder computes from $\underline{y}$ an estimate $\underline{x'}$ of the transmitted message $\underline{x}$
- memoryless channel:

$$\mathbb{P}_C(\underline{y}|\underline{x}) = \prod_{i=1}^{N} \mathbb{P}_C(y_i|x_i)$$

Mutual information

- random variables X, Y with probabilities $p(x) = \mathbb{P}(X = x)$ and $p(y) = \mathbb{P}(Y = y)$
- mutual information $\mathcal{I}_{X,Y}$ of two random variables

$$\mathcal{I}_{X,Y} = \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$$

- for a channel apply to $y \in \mathfrak{A}^N$ received message and $x \in \mathfrak{A}^N$ transmitted message

$$p(x,y) = p(x)\mathbb{P}_C(y|x)$$

- $\mathcal{I}_{X,Y}$ measures reduction in uncertainty about x by knowledge of y

Channel capacity

- channel capacity:

$$C = \max_{p(x)} \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$$

- Example: if output of the channel is pure noise y and x uncorrelated so $C = 0$
- Example: if $y = f(x)$ deterministic function then $C = \max_p S(p) = 1$ (for binary)
- Example: channel with flip probability p and source with $(q, 1-q)$ probabilities: mutual info maximized when source uniform $q = 1/2$ then $C = 1 - S(p)$

Focus here on properties of codes $C : \{0,1\}^M \hookrightarrow \{0,1\}^N$ (and more general non-binary codes) by studying their parameterizing space

Error-correcting codes

- *Alphabet*: finite set A with $\#A = q \geq 2$.
- *Code*: subset $C \subset A^n$, *length* $n = n(C) \geq 1$.
- *Code words*: elements $x = (a_1, \ldots, a_n) \in C$.
- *Code language*: $\mathcal{W}_C = \cup_{m \geq 1} \mathcal{W}_{C,m}$, words $w = x_1, \ldots, x_m$; $x_i \in C$.
- ω-*language*: Λ_C, infinite words $w = x_1, \ldots, x_m, \ldots$; $x_i \in C$.
- Special case: $A = \mathbb{F}_q$, *linear codes*: $C \subset \mathbb{F}_q^n$ linear subspace
- in general: *unstructured codes*

Code parameters

- $k = k(C) := \log_q \#C$ and $[k] = [k(C)]$ integer part of $k(C)$

$$q^{[k]} \leq \#C = q^k < q^{[k]+1}$$

- *Hamming distance*: $x = (a_i)$ and $y = (b_i)$ in C

$$d((a_i), (b_i)) := \#\{i \in (1, \ldots, n) \,|\, a_i \neq b_i\}$$

- *Minimal distance* $d = d(C)$ of the code

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

Code parameters

- $R = k/n =$ *transmission rate* of the code
- $\delta = d/n =$ *relative minimum distance* of the code

Small R: fewer code words, easier decoding, but longer encoding signal; small δ: too many code words close to received one, more difficult decoding. Optimization problem: increase R and δ... how good are codes?

- M.A. Tsfasman, S.G. Vladut, *Algebraic-geometric codes*, Mathematics and its Applications (Soviet Series), Vol. 58, Kluwer Academic Publishers, 1991.

The space of code parameters:

- $Codes_q$ = set of all codes C on an alphabet $\#A = q$
- function $cp : Codes_q \to [0,1]^2 \cap \mathbb{Q}^2$ to code parameters $cp : C \mapsto (R(C), \delta(C))$
- the function $C \mapsto (R(C), \delta(C))$ is a *total recursive map* (Turing computable)
- *Multiplicity* of a code point (R, δ) is $\#cp^{-1}(R, \delta)$

Bounds in the space of code parameters

- singleton bound: $R + \delta \leq 1$
  - from singleton bound $k \leq n - d + 1$ for $n \to \infty$
  - code words $c_1, \ldots, c_M$ this bound says $M \leq q^{n-d+1}$
  - for code word c_i prefix c_i' of length $n - d + 1$
  - for any $i \neq j$ must have $c_i' \neq c_j'$ otherwise $d_H(c_i, c_j) \leq n - (n - d + 1) = d - 1$ but d
  - so $M = \#$ prefixes of length $n - d + 1$, at most q^{n-d+1}
- Gilbert–Varshamov line: $R = \frac{1}{2}(1 - H_q(\delta))$

$$H_q(\delta) = \delta \log_q(q-1) - \delta \log_q \delta - (1-\delta) \log_q(1-\delta)$$

q-ary entropy (for linear codes GV line $R = 1 - H_q(\delta)$)

Shannon Random Code Ensemble (SRCE)

- study behavior of codes by focusing on ensembles of random codes
- case of binary codes (more general codes analogous)
- want to randomize encoding map $E : \{0,1\}^k \to \{0,1\}^n$: there are 2^{n2^k} such possible encoding maps (specify n bits for each of the 2^k codewords)
- in SRCE encoding map is picked uniformly at random from this set
- then encoding of a message: sequence of $\underline{x}_i \in \{0,1\}^k$ and corresponding sequence of codewords $E(\underline{x}_i) \in \{0,1\}^n$ obtained by tossing an unbiased coins N-times, with i-th result being the i-th coord of $E(\underline{x}_i)$
- random codes are not injective: different words can have same encoding, but such occurrences are rare in probability

decoding problem for random codes

- probability distribution $\mathbb{P}(\underline{x}|\underline{y})$ of $\underline{x}$ being the channel input if $\underline{y}$ is the received message
- suppose memoryless channel with $\mathbb{P}_C(\underline{y}|\underline{x})$
- Bayes rule:

$$\mathbb{P}(\underline{x}|\underline{y}) = \frac{1}{Z(\underline{y})} \prod_{i=1}^{n} \mathbb{P}_C(y_i|x_i)\mathbb{P}(\underline{x})$$

with $Z(\underline{y})$ determined by imposing normalization condition $\sum_{\underline{x}} \mathbb{P}(\underline{x}|\underline{y}) = 1$ and $\mathbb{P}(\underline{x})$ a priori probability of $\underline{x}$ being produced as message at the source

- if source uniform probability $\mathbb{P}(\underline{x}) = 2^{-k}$

Geometry of Shannon Random Code Ensemble

- code: set C of 2^k codewords inside ambient space $\{0,1\}^n$
- each of these points drawn with uniform probability from $\{0,1\}^n$
- how many codewords are near a given codeword?
- Hamming distance $d_H(\underline{x}, \underline{x'}) = \#\{i : \underline{x}_i \neq \underline{x'}_i\}$ number of differing coordinates

Hamming enumerator

- Hamming distance enumerator $\mathcal{N}_{\underline{x}^{(0)}}(d)$
- counting number of codewords at distance d from a chosen one $\underline{x}^{(0)}$
- average $\mathbb{E}(\mathcal{N}_{\underline{x}^{(0)}}(d))$ over the code ensemble
- since all code words drawn independently with uniform probability result should not depend on which $\underline{x}^{(0)}$ used, so pick $\underline{x}^{(0)} = (0, 0, \ldots, 0)$
- given $2^k - 1$ points chosen uniformly at random in $\{0, 1\}^n$ how many are at distance d from $(0, 0, \ldots, 0)$ corner?
- number of points $(2^k - 1)$ times fraction of Hamming volume at distance d from $(0, 0, \ldots, 0)$ (which is $2^{-n}\binom{n}{d}$), Hamming "sphere"

asymptotics of Hamming enumerator

- when $n \to \infty$ with $d/n \to \delta$ and $k/n \to R$ finite

$$\mathbb{E}(\mathcal{N}_{\underline{x}^{(0)}}(d)) = (2^k - 1)\, 2^{-n} \binom{n}{d} \sim 2^{n(R-1+H_2(\delta))}$$

$$H_2(\delta) = -\delta \log_2 \delta - (1-\delta) \log_2 (1-\delta)$$

Shannon entropy

- similar for q-ary codes, alphabet A with $\#A = q \geq 2$

$$\mathbb{E}(\mathcal{N}_{\underline{x}^{(0)}}(d)) = (q^k - 1)\, q^{-n} \binom{n}{d} (q-1)^d \sim q^{n(R-1+H_q(\delta))}$$

with q-ary entropy
$H_q(\delta) = \delta \log_q (q-1) - \delta \log_q \delta - (1-\delta) \log_q (1-\delta)$

- Hamming ball volume

$$\mathrm{Vol}_q(n,d) = \sum_{j=0}^{d} \binom{n}{j} (q-1)^j$$

estimate of Hamming ball volume

upper bound estimate

$$
\begin{aligned}
1 &= (p + (1-p))^n \\
&= \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{pn} \binom{n}{i} p^i (1-p)^{n-i} + \sum_{i=pn+1}^{n} \binom{n}{i} p^i (1-p)^{n-i} \\
&\geq \sum_{i=1}^{pn} \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{pn} \binom{n}{i} (q-1)^i \left(\frac{p}{q-1}\right)^i (1-p)^{n-i} \\
&= \sum_{i=1}^{pn} \binom{n}{i} (q-1)^i (1-p)^n \left(\frac{p}{(q-1)(1-p)}\right)^i \\
&\geq \sum_{i=1}^{pn} \binom{n}{i} (q-1)^i (1-p)^n \left(\frac{p}{(q-1)(1-p)}\right)^{pn} \\
&= \left(\frac{p}{q-1}\right)^{pn} (1-p)^{(1-p)n} \sum_{i=1}^{pn} \binom{n}{i} (q-1)^i \\
&\geq \mathrm{Vol}_q(pn, n) q^{-nH_q(p)}
\end{aligned}
$$

estimate of Hamming ball volume

Stirling formula:

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\lambda_1(n)} \leq n! \leq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\lambda_2(n)}$$

$$\begin{aligned}\binom{n}{pn} &= \frac{n!}{(pn)!((1-p)n)!} \\ &> \frac{(n/e)^n}{(pn/e)^{pn}((1-p)n/e)^{(1-p)n}} \cdot \underbrace{\frac{e^{\lambda_1(n)-\lambda_2(pn)-\lambda_2((1-p)n)}}{\sqrt{2\pi p(1-p)n}}}_{\ell(n)} \\ &= \frac{\ell(n)}{p^{pn}(1-p)^{(1-p)n}}\end{aligned}$$

then lower bound estimate

$$\begin{aligned}\mathrm{Vol}_q(pn,n) &\geq \binom{n}{pn}(q-1)^{pn} \\ &> \frac{(q-1)^{pn}}{p^{pn}(1-p)^{(1-p)n}} \cdot l(n) \\ &\geq q^{nH_q(p)+\log_q l(n)}\end{aligned}$$

Statistics of codes and the Gilbert–Varshamov bound

Known *statistical* approach to the GV bound: *random codes*

Shannon Random Code Ensemble: ω-language with alphabet A; uniform Bernoulli measure on Λ_A; choose code words of C as independent random variables in this measure

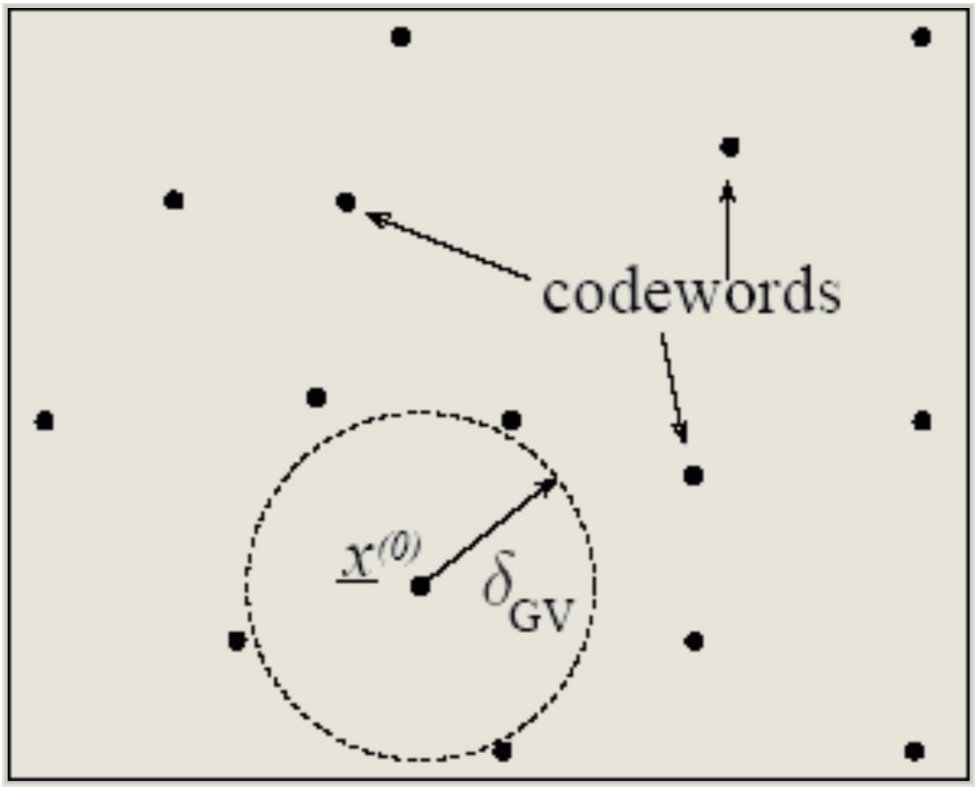
Volume estimate:

$$q^{(H_q(\delta)-o(1))n} \leq Vol_q(n, d = n\delta) = \sum_{j=0}^{d} \binom{n}{j} (q-1)^j \leq q^{H_q(\delta)n}$$

Gives probability of parameter δ for SRCE meets the GV bound with probability exponentially (in n) near 1: expectation

$$\mathbb{E} \sim \binom{q^k}{2} Vol_q(n, d) q^{-n} \sim q^{n(H_q(\delta)-1+2R)+o(n)}$$

code words distribution in random codes



for $n >> 1$ ball around a code word contains no other code words when $\delta < \delta_{GV}$ and exponentially many code words for $\delta > \delta_{GV}$

random linear codes and the GV bound

- for a linear code $d = \min_{\underline{y} \in C} \omega(\underline{y})$ with $\omega(\underline{y}) = \#\{i \,:\, y_i \neq 0\}$
- given a non-zero vector $\underline{x} \in \mathbb{F}_q^k$ and a uniformly random matrix $T \in M_{k \times n}(\mathbb{F}_q)$, the vector $\underline{y} = T\underline{x}$ is uniformly distributed over $\mathbb{F}_q^n$
- $y_i = \sum_j T_{ij} x_j$ so for $i \neq i'$ independent $y_i, y_{i'}$ as depend on different sets of entries of T (independently randomly chosen)
- each y_i uniformly distributed over $\mathbb{F}_q$: take an $x_j \neq 0$, fix other $T_{ij'}$, varying T_{ij} equiprobable, all values in $\mathbb{F}_q$ achieved for y_i
- Claim: for $k = (1 - H_q(\delta) - \epsilon)n$ (slightly below GV curve) there is some $T \in M_{k \times n}(\mathbb{F}_q)$ such that for all $\underline{x} \in \mathbb{F}_q^k \setminus \{0\}$ the $\omega(T\underline{x}) \geq d$
- using equidistribution of images and

$$\mathbb{P}(\omega(T\underline{x}) < d) = q^{-n} \mathrm{Vol}_q(n, d-1) \leq q^{n(H_q(\delta)-1)}$$

- then have

$$\mathbb{P}(\exists \underline{x} : \omega(T\underline{x}) < d) \leq q^k q^{n(H_q(\delta)-1)} = q^{(1-H_q(\delta)-\epsilon)n+n(H_q(\delta)-1)} = q^{-\epsilon n}$$

- for large n this probability very small so Claim follows
- also T has full rank: with high probability $\omega(\underline{y}) > d$ for all codewords, so since linear min of Hamming distances also $> d$, hence $C : \mathbb{F}_q^k \hookrightarrow \mathbb{F}_q^n$ injective
- this shows that random linear codes with high probability lie on the GV-curve for $n \to \infty$

- probability distribution of code words given received ouput y of channel

$$\mu_y(x) = \frac{1}{Z(y)} \prod_i \mathbb{P}_C(y_i|x_i)\mu_0(x)$$

for memoryless channel (Bayes rule)

- for a binary code and a channel that randomly flips bits with $0 < p < 1$ probability

$$\mu_y(x) = \frac{1}{Z(y)} p^{d_H(x,y)}(1-p)^{n-d_H(x,y)}$$

some (other) normalization $Z(y)$

- with $B = \frac{1}{2}\log\left(\frac{1-p}{p}\right)$ partition function counts contribution of correct codeword x_0 and of all other codewords x

$$Z = e^{-2Bd_H(x_0,y)} + \sum_{d=0}^{n} \hat{\mathcal{N}}_y(d)\, e^{-2Bd}$$

number $\hat{\mathcal{N}}_y(d)$ of incorrect code words at distance d from y

- for large n (law of large numbers) $d_H(x_0, y) \sim np$ so first term $Z_{corr} = e^{-2Bd_H(x_0,y)} \sim e^{-2nBp}$
- distance enumerator $\hat{\mathcal{N}}_y(d)$ as before exponentially large for $\delta_{GV}(R) < \delta < 1 - \delta_{GV(R)}$ and vanishes with high probability outside that interval
- also for $\delta_{GV}(R) < \delta < 1 - \delta_{GV(R)}$ concentrated at the mean value

$$\mathbb{E}(\hat{\mathcal{N}}_y(d)) \sim 2^{n(R-1+H_2(\delta))}$$

- then summation over d by saddle point

$$Z_{err} = \sum_{d=0}^{n} \hat{\mathcal{N}}_y(d)\, e^{-2Bd} \sim n \int_{\delta_G V}^{1-\delta_{GV}} e^{n((R-1)\log 2 + S(\delta) - 2B\delta)} \sim e^{n\varphi_{err}}$$

$$\varphi_{err} = \max_{\delta \in [\delta_{GV}, 1-\delta_{GV}]} ((R-1)\log 2 + H(\delta) - 2B\delta)$$

- since $B = \frac{1}{2} \log\left(\frac{1-p}{p}\right)$ max $\varphi_{err} = \varphi_{err}(p)$ (assume $p < 1/2$)
- when max inside interval $(\delta_{GV}, 1 - \delta_{GV})$ it occurs where $H'(\delta) = 2B$
- otherwise max at lower end $\delta = \delta_{GV}$ (since $B > 0$)

$$\varphi_{err}(p) = \begin{cases} -\delta_{GV}(R) \log\left(\frac{1-p}{p}\right) & p < \delta_{GV} \\ (R-1)\log 2 - \log(1-p) & \text{otherwise} \end{cases}$$

- for low noise level (small p) term Z_{err} exponentially small
- for high noise (past the δ_{GV} threshold) Z_{err} dominates

Statistical physics: finite temperature decoding

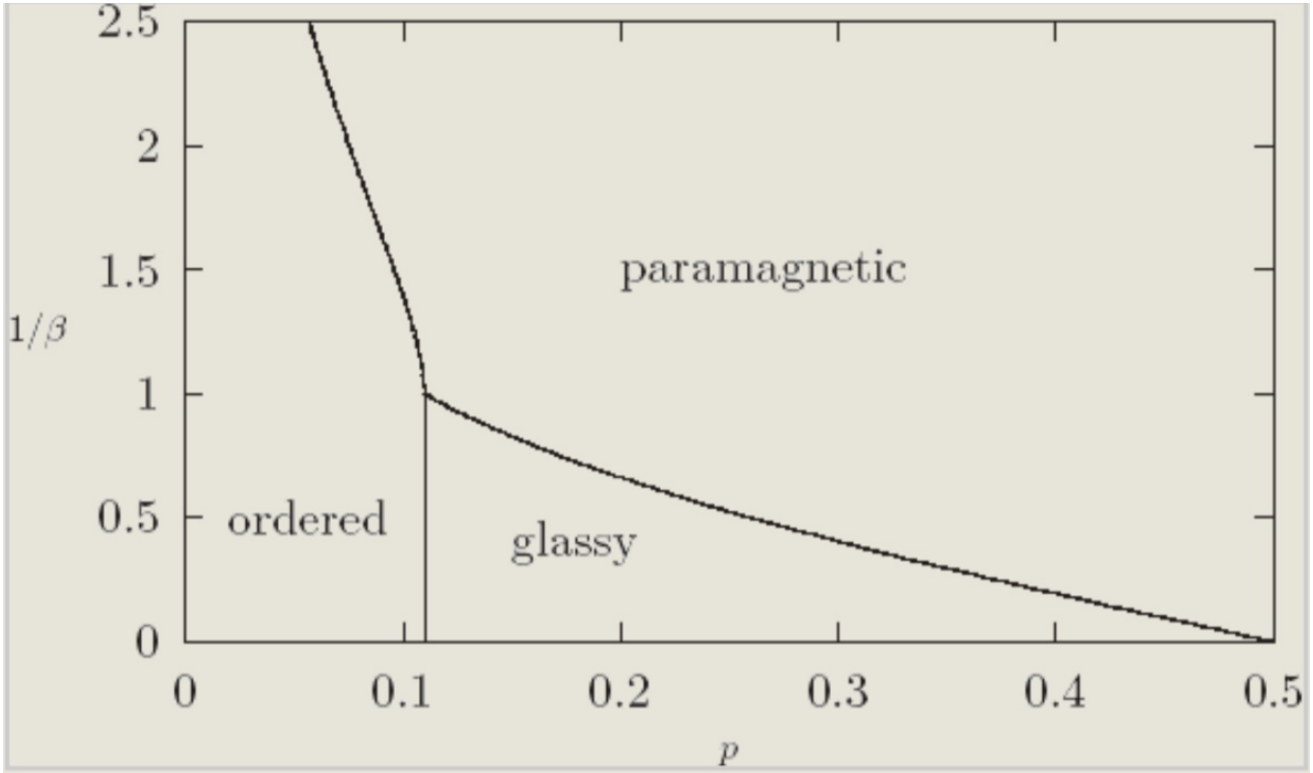
- introduce a temperature parameter $\beta = 1/T$
- probability distribution of code words given received ouput y of channel

$$\mu_{\beta,y}(x) = \frac{1}{Z_y(\beta)} e^{-2\beta B d_H(y,x)} \quad \text{with} \quad Z_y(\beta) = \sum_x e^{-2\beta B d_H(y,x)}$$

- this shows a phase transition diagram
  1. completely ordered crystal phase: low noise $p < \delta_{GV}$ and low temperature (large enough β) *good decoding* distribution $\mu_{\beta,y}(x)$ dominated by correct code word
  2. glassy phase: higher noise $p > \delta_{GV}$ still low temperature (large β) correct code word has small weight and $\mu_{\beta,y}(x)$ dominated by other code words closest to y (not correct one)
  3. entropy dominated high temperature paramagnetic gas phase: high temperature (small β) with $\mu_{\beta,y}(x)$ dominated by code word at distance $d = n\delta_*$ larger than min distance with

$$\delta_* = \frac{p^\beta}{p^\beta + (1-p)^\beta}$$

phase transition diagram



Spoiling operations on codes: C an $[n, k, d]_q$ code

- $C_1 := C *_i f \subset A^{n+1}$

$$(a_1, \ldots, a_{n+1}) \in C_1 \text{ iff } (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{n+1}) \in C\,,$$

and $a_i = f(a_1, \ldots, a_{i-1}, a_{i+1} \ldots, a_{n+1})$
C_1 an $[n+1, k, d]_q$ code (f constant function)

- $C_2 := C*_i \subset A^{n-1}$

$(a_1, \ldots, a_{n-1}) \in C_2$ iff $\exists b \in A$, $(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_{n-1}) \in C$.

C_2 an $[n-1, k, d]_q$ code

- $C_3 := C(a, i) \subset C \subset A^n$

$$(a_1, \ldots, a_n) \in C_3 \text{ iff } a_i = a.$$

C_3 an $[n-1, k-1 \leq k' < k, d' \geq d]_q$ code

Asymptotic bound

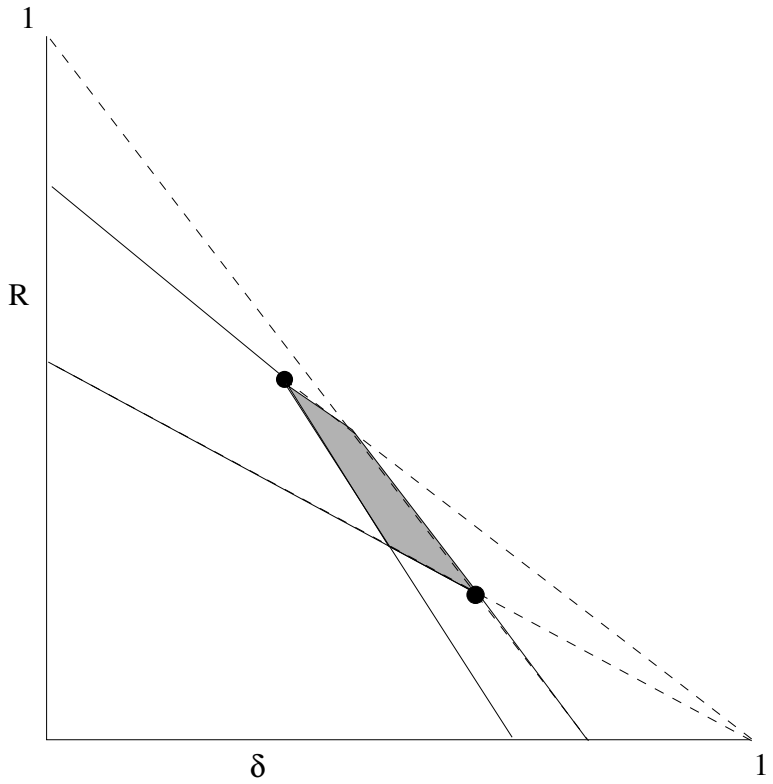
- Yu.I.Manin, *What is the maximum number of points on a curve over $\mathbb{F}_2$?* J. Fac. Sci. Tokyo, IA, Vol. 28 (1981), 715–720.

- $V_q \subset [0,1]^2$: all code points $(R, \delta) = cp(C)$, $C \in Codes_q$
- U_q: set of limit points of V_q
- Asymptotic bound: U_q all points below graph of a function

$$U_q = \{(R, \delta) \in [0,1]^2 \mid R \leq \alpha_q(\delta)\}$$

- Isolated code points: $V_q \setminus (V_q \cap U_q)$

Method: controlling quadrangles



$R = \alpha_q(\delta)$ continuous decreasing function with $\alpha_q(0) = 1$ and $\alpha_q(\delta) = 0$ for $\delta \in [\frac{q-1}{q}, 1]$; has inverse function on $[0, (q-1)/q]$;
U_q union of all lower cones of points in $\Gamma_q = \{R = \alpha_q(\delta)\}$

Characterization of the asymptotic bound

- Code points and multiplicities

- Set of code points of infinite multiplicity
$U_q \cap V_q = \{(R, \delta) \in [0,1]^2 \cap \mathbb{Q}^2 \mid R \leq \alpha_q(\delta)\}$ below the asymptotic bound

- Code points of finite multiplicity all above the asymptotic bound
$V_q \setminus (U_q \cap V_q)$ and isolated (open neighborhood containing (R, δ) as unique code point)

Questions:

- Is there a characterization of the isolated good codes on or above the asymptotic bound?

Estimates on the asymptotic bound

- Plotkin bound:

$$\alpha_q(\delta) = 0, \quad \delta \geq \frac{q-1}{q}$$

- singleton bound:

$$\alpha_q(\delta) \leq 1 - \delta$$

- Hamming bound:

$$\alpha_q(\delta) \leq 1 - H_q(\frac{\delta}{2})$$

- Gilbert–Varshamov bound:

$$\alpha_q(\delta) \geq 1 - H_q(\delta)$$

Computability question

• Note: only the asymptotic bound marks a significant change of behavior of codes across the curve (isolated and finite multiplicity/accumulation points and infinite multiplicity)

• in this sense it is very different from all the other bounds in the space of code parameters

• .... but no explicit expression for the curve $R = \alpha_q(\delta)$

• ... is the function $R = \alpha_q(\delta)$ computable?

• ... a priori no good statistical description of the asymptotic bound: is there something replacing Shannon entropy characterizing Gilbert–Varshamov curve?

- Yu.I. Manin, *A computability challenge: asymptotic bounds and isolated error-correcting codes*, arXiv:1107.4246

The asymptotic bound and Kolmogorov complexity

• while random codes are related to Shannon entropy (through the GV-bound) good codes and the asymptotic bound are related to Kolmogorov complexity

• the asymptotoc bound $R = \alpha_q(\delta)$ becomes computable given an oracle that can list codes by increasing Kolmogorov complexity

• given such an oracle: iterative (algorithmic) procedure for constructing the asymptotic bound

• ... it is at worst as "non-computable" as Kolmogorov complexity

• asymptotic bound can be realized as phase transition curve of a statistical mechanical system based on Kolmogorov complexity

- Yu.I. Manin, M. Marcolli, *Kolmogorov complexity and the asymptotic bound for error-correcting codes*, Journal of Differential Geometry, Vol.97 (2014) 91–108

Complexity

- How does one measure complexity of a physical system?
- Kolmogorov complexity: measures length of a minimal algorithmic description

... but ... gives very high complexity to completely random things

- Shannon entropy: measures average number of bits, for objects drawn from a statistical ensemble
- There are other proposals for complexity, but more difficult for formulate
- Gell-Mann complexity: complexity is high in an intermediate region between total order and complete randomness

Kolmogorov complexity

• Let $T_{\mathcal{U}}$ be a universal Turing machine (a Turing machine that can simulate any other arbitrary Turing machine: reads on tape both the input and the description of the Turing machine it should simulate)

• Given a string w in an alphabet $\mathfrak{A}$, the Kolmogorov complexity

$$\mathcal{K}_{T_{\mathcal{U}}}(w) = \min_{P: T_{\mathcal{U}}(P)=w} \ell(P),$$

minimal length of a program that outputs w

• universality: given any other Turing machine T

$$\mathcal{K}_T(w) = \mathcal{K}_{T_{\mathcal{U}}}(w) + c_T$$

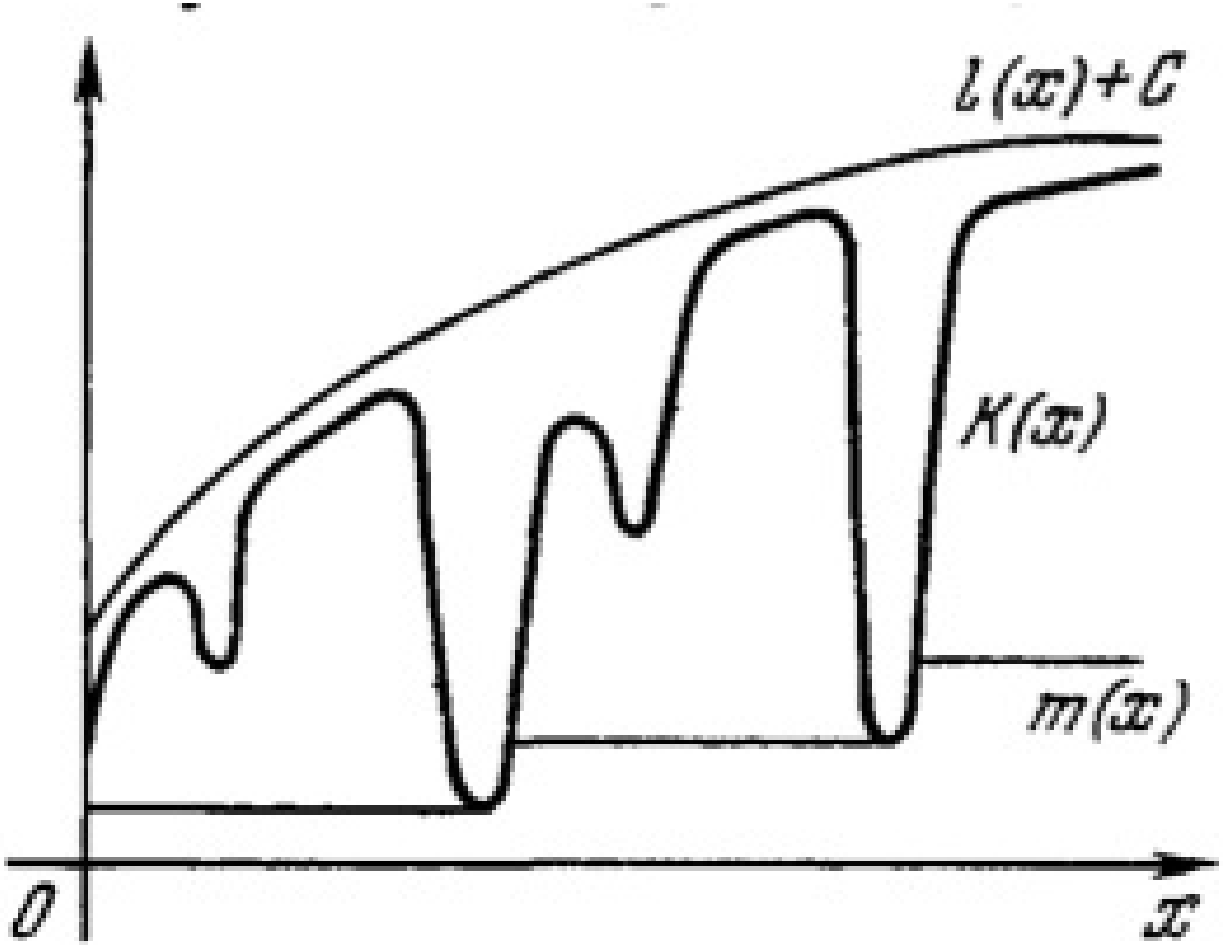
shift by a bounded constant, independent of w; c_T is the Kolmogorov complexity of the program needed to describe T for $T_{\mathcal{U}}$ to simulate it

- any program that produces a description of w is an upper bound on Kolmogorov complexity $\mathcal{K}_{T_\mathcal{U}}(w)$
- think of Kolmogorov complexity in terms of data compression
- shortest description of w is also its most compressed form
- can obtain upper bounds on Kolmogorov complexity using data compression algorithms
- finding upper bounds is easy... but NOT lower bounds

Main problem

Kolmogorov complexity is NOT a computable function

- suppose list programs P_k (increasing lengths) and run through $T_{\mathcal{U}}$: if machine halts on P_k with output w then $\ell(P_k)$ is an upper bound on $\mathcal{K}_{T_{\mathcal{U}}}(w)$
- but... there can be an earlier P_j in the list such that $T_{\mathcal{U}}$ has not yet halted on P_j
- if eventually halts and outputs w then $\ell(P_j)$ is a better approximation to $\mathcal{K}_{T_{\mathcal{U}}}(w)$
- would be able to compute $\mathcal{K}_{T_{\mathcal{U}}}(w)$ if can tell exactly on which programs P_k the machine $T_{\mathcal{U}}$ halts
- but... halting problem is unsolvable



with $m(x) = \min_{y \geq x} \mathcal{K}(y)$

Kolmogorov complexity

$X =$ *infinite constructive world*: have structural numbering

computable bijections $\nu : \mathbb{Z}^+ \to X$ principal homogeneous space over group of total recursive permutations $\mathbb{Z}^+ \to \mathbb{Z}^+$

- *Ordering*: $x \in X$ is generated at the $\nu^{-1}(x)$-th step

Optimal partial recursive enumeration $u : \mathbb{Z}^+ \to X$
(Kolmogorov and Schnorr)

$$K_u(x) := \min\{k \in \mathbb{Z}^+ \mid u(k) = x\}$$

Kolmogorov complexity

- changing $u : \mathbb{Z}^+ \to X$ changes $K_u(x)$ up to bounded (multiplicative) constants $c_1 K_v(x) \leq K_u(x) \leq c_2 K_v(x)$
- min length of program generating x (by Turing machine)

Main Idea:

• use characterization of asymptotic bound as separating code points with finite multiplicity from code points with infinite multiplicity

• given the function from codes to code parameter, want an algorithmic procedure that inductively constructs preimage sets with finite/infinite multiplicity

• choose an ordering of code points: at step m list code points in order up to some growing size N_m

• initialize A_1: a set of *a preimage* for each code point up to N_1; initialize $B_1 = \emptyset$

• want to increase at each step A_m and B_m so that the first set only contains code points with multiplicity m

• going from step m to step $m+1$: new code points listed between N_m and N_{m+1} are added to A_m, and then points (previously in A_m or added) that do not have an $m+1$-st preimage are moved to B_{m+1}

• as $m \to \infty$ the sets A_m converge to set of code points of infinite multiplicity and the B_m converge to set of code points of finite multiplicity

• key problem: need to search for the $m+1$-st preimage to detect if a code point stays in A_{m+1} or is moved to B_{m+1}

• ordinarily this would involve an *infinite search*...

• ordering and complexity: use a relation between ordering and complexity that shows that only need to search among bounded complexity codes, so a *complexity oracle* will render the search finite

X, Y infinite constructive worlds, ν_X, ν_Y structural bijections, u, v optimal enumerations, K_u and K_v Kolmogorov complexities

- total recursive function $f : X \to Y \Rightarrow \forall y \in f(X), \exists x \in X$, $y = f(x)$: $\exists$ computable $c = c(f, u, v, \nu_X, \nu_Y) > 0$

$$K_u(x) \leq c \cdot \nu_Y^{-1}(y)$$

Kolmogorov ordering

$\mathbf{K}_u(x) =$ order X by growing Kolmogorov complexity $K_u(x)$

$$c_1 K_u(x) \leq \mathbf{K}_u(x) \leq c_2 K_u(x)$$

So... if know how to generate elements of X in Kolmogorov ordering then can generate all elements of $f(X) \subset Y$ in their structural ordering

In fact... take $F(x) = (f(x), n(x))$ with

$$n(x) = \#\{x' \mid \nu_X^{-1}(x') \leq \nu_X^{-1}(x),\ f(x') = f(x)\}$$

total recursive function $\Rightarrow E = F(X) \subset Y \times \mathbb{Z}^+$ enumerable

- $X_m := \{x \in X \mid n(x) = m\}$ and $Y_m := f(X_m) \subset Y$ enumerable
- for $x \in X_1$ and $y = f(x)$: complexity $K_u(x) \leq c \cdot \nu_Y^{-1}(y)$ (using inequalities for complexity under composition)

Multiplicity: $mult(y) := \#f^{-1}(y)$

$$Y_\infty \subset \cdots f(X_{m+1}) \subset f(X_m) \subset \cdots \subset f(X_1) = f(X)$$

$Y_\infty = \cap_m f(X_m)$ and $Y_{fin} = f(X) \setminus Y_\infty$

Key Step: $y \in Y_\infty$ and $m \geq 1$: $\exists$ unique $x_m \in X$, $y = f(x_m)$, $n(x_m) = m$ and $c = c(f, u, v, \nu_X, \nu_Y) > 0$

$$K_u(x_m) \leq c \cdot \nu_Y^{-1}(y)\, m\, \log(\nu_Y^{-1}(y) m)$$

Oracle mediated recursive construction of Y_∞ and Y_{fin}

- Choose sequence (N_m, m), $m \geq 1$, $N_{m+1} > N_m$
- Step 1: $A_1 =$ list $y \in f(X)$ with $\nu_Y^{-1}(y) \leq N_1$; $B_1 = \emptyset$
- Step $m+1$: Given A_m and B_m, list $y \in f(X)$ with $\nu_Y^{-1}(y) \leq N_{m+1}$; $A_{m+1} =$ elements in this list for which $\exists\, x \in X$, $y = f(x)$, $n(x) = m+1$; $B_{m+1} =$ remaining elements in the list
- oracle: search for $x \in X$, $y = f(x)$, $n(x) = m+1$ only among those x with complexity bounded by function of $\nu_Y^{-1}(y)$ as above
- $A_m \cup B_m \subset A_{m+1} \cup B_{m+1}$, union is all $f(X)$; $B_m \subset B_{m+1}$ and $Y_{fin} = \cup_m B_m$, while $Y_\infty = \cup_{m \geq 1}(\cap_{n \geq 0} A_{m+n})$
- from A_m to A_{m+1} first add all new y with $N_m < \nu_Y^{-1}(y) \leq N_{m+1}$ then subtract those that have no more elements in the fiber $f^{-1}(y)$: these will be in B_{m+1}

Structural numbering for codes

- $X = Codes_q$, $Y = [0,1]^2 \cap \mathbb{Q}^2$ and $f : X \to Y$ is $cp : C \mapsto (R(C), \delta(C))$ code parameters map
- $A = \{0, \ldots, q-1\}$ ordered, A^n lexicographically; computable total order ν_X:
(i) if $n_1 < n_2$ all $C \subset A^{n_1}$ before all $C' \subset A^{n_2}$;
(ii) $k_1 < k_2$ all $[n, k_1, d]_q$-codes before $[n, k_2, d']_q$-codes;
(iii) fixed n and q^k: lexicographic order of code words, concatenated into single word $w(C)$ (determines code): order all the $w(C)$ lexicographically
- total recursive map $cp : Codes_q \to [0,1]^2 \cap \mathbb{Q}^2$
- fixed enumeration ν_Y of rational points in $[0,1]^2$

... inductively building the asymptotic bound using the described oracle mediated procedure

- Question: is there a statistical view of this procedure?

Partition function for code complexity

$$Z(X,\beta) = \sum_{x \in X} K_u(x)^{-\beta}$$

weights elements in constructive world X by inverse complexity;
$\beta =$ inverse temperature, thermodynamic parameter

Convergence properties

- Kolmogorov complexity and Kolmogorov ordering

$$c_1\, \mathbf{K}_u(x) \leq K_u(x) \leq c_2\, \mathbf{K}_u(x)$$

- convergence of $Z(X,\beta)$ controlled by series

$$\sum_{x \in X} \mathbf{K}_u(x)^{-\beta} = \sum_{n \geq 1} n^{-\beta} = \zeta(\beta)$$

- Partition function $Z(X,\beta)$ convergence for $\beta > 1$; phase transition at pole $\beta = 1$

Asymptotic bound as a phase transition

- $X' \subset X$ infinite decidable subset of a constructive world
- $i : X' \hookrightarrow X$ total recursive function; also $j : X \to X'$ identity on X' constant on complement

$$K_u(i(x')) \leq c_1 K_v(x') \quad \text{and} \quad K_v(j(x)) \leq c_2 K_u(x)$$

- $\delta = \beta_q(R)$ inverse of $\alpha_q(\delta)$ on $R \in [0, 1 - 1/q]$
- Fix $R \in \mathbb{Q} \cap (0, 1)$ and $\Delta \in \mathbb{Q} \cap (0, 1)$

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

Phase transition at the asymptotic bound

- $1 - \Delta > \beta_q(R)$: partition function $Z(R, \Delta; \beta)$ real analytic in β
- $1 - \Delta < \beta_q(R)$: partition function $Z(R, \Delta; \beta)$ real analytic for $\beta > \beta_q(R)$ and divergence for $\beta \to \beta_q(R)_+$

Another view of the asymptotic bound as a phase transition

- Yuri I. Manin, Matilde Marcolli, *Error-correcting codes and phase transitions*, Mathematics in Computer Science (2011) 5:133–170.

• when constructing random codes (Shannon Random Code Ensemble): choose code words as equally distributed independent random variables

• imagine passing from classical to quantum systems, where the code words remain the fundamental degrees of freedom

• the basic quantum system of this kind is a system of independent harmonic oscillators: creation/annihilation operators associated to the basic independent degrees of freedom

Single Code: algebra of creation/annihilation operators

- for a single code C: code words are degrees of freedom
- Algebra of observable of a single code: Toeplitz algebra on code words

$$\mathcal{T}_C : \quad T_x, \quad x \in C, \quad T_x^* T_x = 1$$

$T_x T_x^*$ mutually orthogonal projectors

- Fock space representation $\mathcal{H}_C$ spanned by ϵ_w, words $w = x_1, \ldots, x_N$ in code language $\mathcal{W}_C$

$$T_x \epsilon_w = \epsilon_{xw}$$

Quantum Statistical Mechanics of a single code

- algebra of observables $\mathcal{T}_C$; time evolution $\sigma : \mathbb{R} \to \mathrm{Aut}(\mathcal{T}_C)$

$$\sigma_t(T_x) = K_u(C)^{it}\, T_x$$

- Hamiltonian $\pi(\sigma_t(T)) = q^{itH}\pi(T)q^{-itH}$

$$H\,\epsilon_w = \ell(w)\, log_q K_u(C)\,\epsilon_w$$

in Fock representation, $\ell(w)$ length of word (# of code words)

- Partition function

$$Z(C,\sigma,\beta) = \mathrm{Tr}(e^{-\beta H}) = \sum_m (\#W_{C,m}) K_u(C)^{-\beta m}$$

$$= \sum_m q^{m(nR - \beta \log_q K_u(C))} = \frac{1}{1 - q^{nR} K_u(C)^{-\beta}}$$

- Convergence: $\beta > nr/\log_q K_u(C)$

QSM system at a code point (R, δ)

- Different codes $C \in cp^{-1}(R, \delta)$ as independent subsystems
- Tensor product of Toeplitz algebras $\mathcal{T}_{(R,\delta)} = \otimes_{C \in cp^{-1}(R,\delta)} \mathcal{T}_C$
- Shift on single code temperature so that

$$Z(C, \sigma, n(\beta - \delta + 1)) \leq (1 - K_u(C)^{-\beta})^{-1}$$

by *singleton bound* on codes $R + \delta - 1 \leq 0$

- Fock space $\mathcal{H}_{(R,\delta)} = \otimes \mathcal{H}_C$; time evolution $\sigma = \otimes \sigma^C$
- Partition function (variable temperature)

$$Z(cp^{-1}(R, \delta), \sigma; \beta) = \prod_{C \in cp^{-1}(R,\delta)} Z(C, \sigma, n(\beta - \delta + 1))$$

- Convergence controlled by $\prod_C (1 - K_u(C)^{-\beta})^{-1}$; in turned controlled by the classical zeta function $Z(cp^{-1}(R, \delta), \beta) = \sum_{C \in cp^{-1}(R,\delta)} K_u(C)^{-\beta}$

first versus second quantization

• Bosonic second quantization: example of primes p and integers $n \in \mathbb{N}$; independent degrees of freedom (primes) quantized by isometries $\tau_p^* \tau_p = 1$; tensor product of Toeplitz algebras $\otimes_p \mathcal{T}_p = C^*(\mathbb{N})$ semigroup algebra; $\sigma_t(\tau_p) = p^{it}\tau_p$, partition function $\zeta(\beta) = \prod_p (1 - p^{-\beta})^{-1}$ prod of partition functions individual systems

• Infinite tensor product: second quantization; finite tensor product: quantum mechanical (finitely many degrees of freedom) first quantization

• $(\mathcal{T}_{(R,\delta)}, \sigma)$ is quantum mechanical above the asymptotic bound; bosonic QFT below asymptotic bound

Asymptotic bound boundary between first and second quantization

Asymptotic bound as a phase transition (QSM point of view)

- Variable temperature partition function: $\mathcal{A} = \otimes_{s \in S} \mathcal{A}_s$, $\sigma = \otimes_s \sigma_s$; $\beta : S \to \mathbb{R}_+$; $Z(\mathcal{A}, \sigma, \beta) = \prod_s Z(\mathcal{A}_s, \sigma_s, \beta(s))$
- fix a code point (R, δ); partition function (variable β)

$$Z((R,\delta),\sigma;\beta) = \prod_{C \in cp^{-1}(R,\delta)} (1 - q^{(R-\beta)n_C})^{-1}$$

- if (R, δ) above bound finite product; if below bound convergence governed by $\sum_C q^{(R-\beta)n_C}$, for $\beta > R$.
- change of behavior of the system at $R = \alpha_q(\delta)$ asymptotic bound

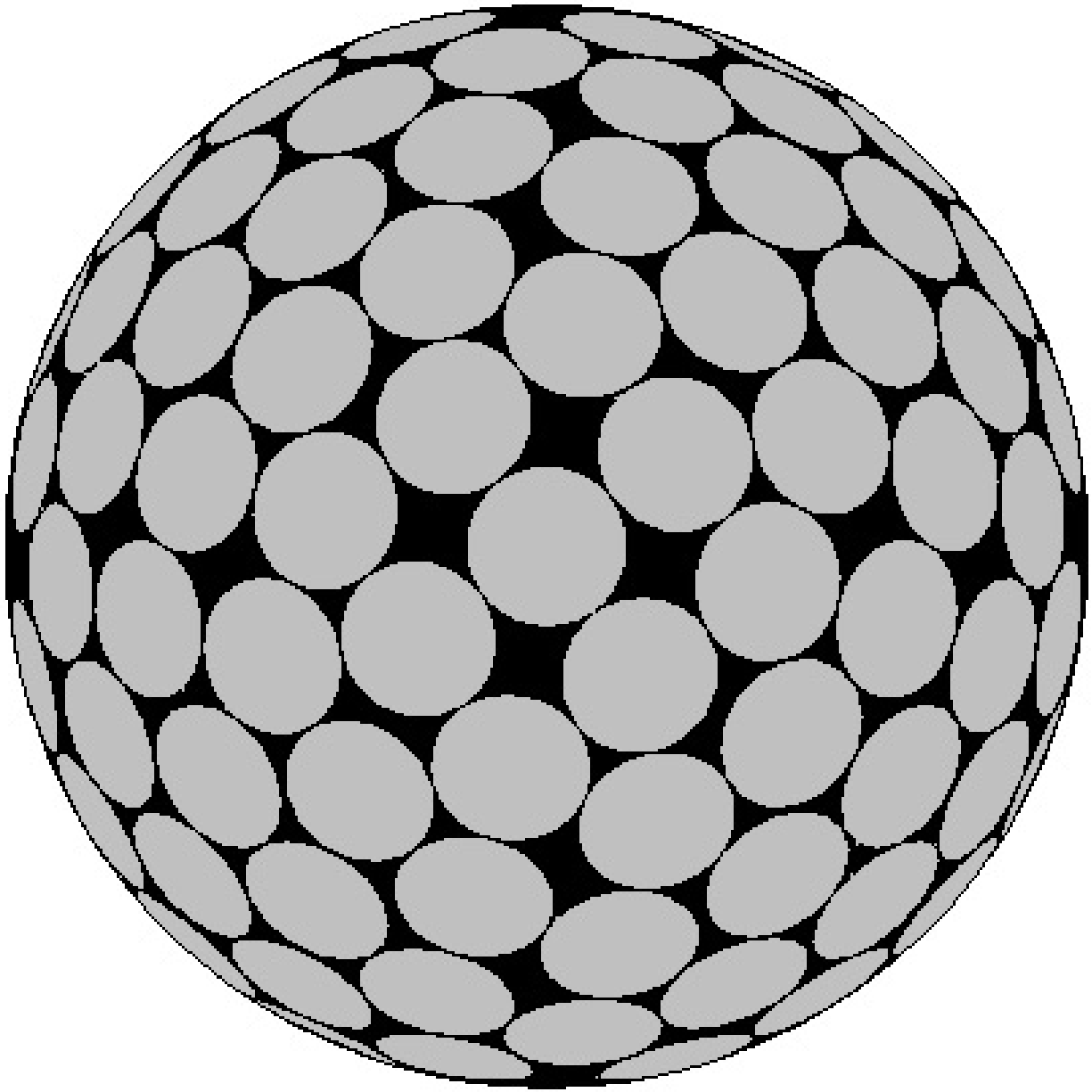
Spherical Codes

- Yuri I. Manin, Matilde Marcolli, *Asymptotic bounds for spherical codes*, arXiv:1801.01552

- spherical code: finite set X of points on unit sphere $S^{n-1} \subset \mathbb{R}^n$
- spherical code X has minimal angle ϕ if $\forall x \neq y \in X$

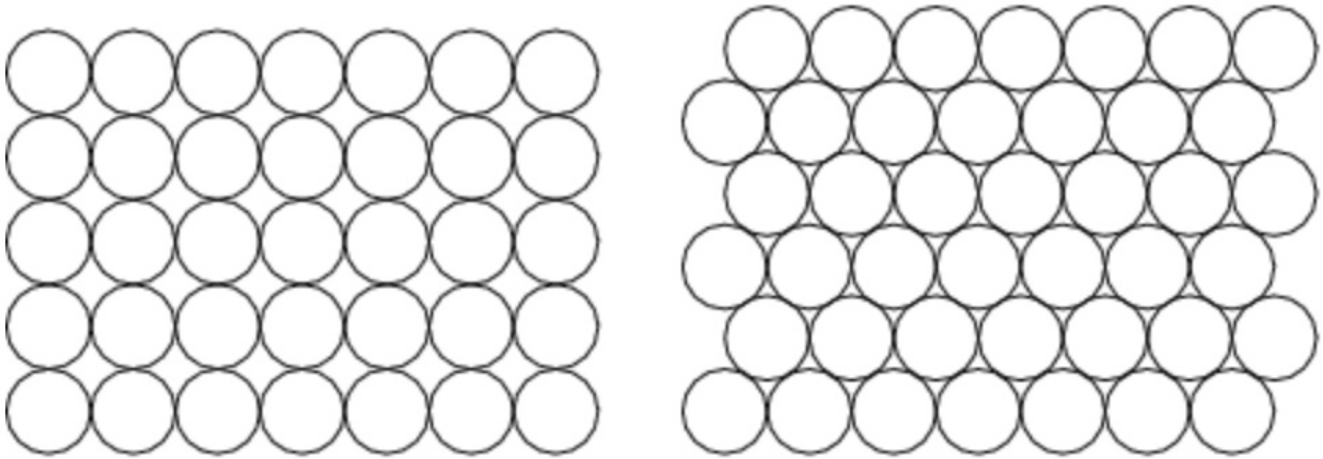
$$\langle x, y \rangle \leq \cos \phi$$

- $A(n, \phi) =$ max number of points on S^{n-1} with minimal angle ϕ



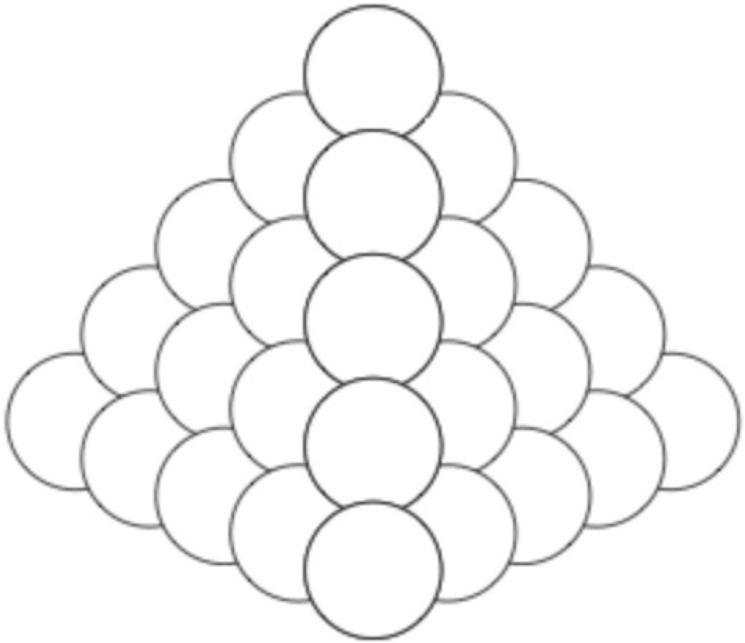
Spherical codes and sphere packings

- non-overlapping congruent balls in $\mathbb{R}^n$
- density: fraction of space covered by the balls in the packing
- ball $B_R^n(x)$ of radius R centered at x
- density of packing: limit for $R \to \infty$ (if exists) of fraction of $B_R^n(x)$ covered by spheres in the packing, independent of x if exists
- $\Delta_{\mathbb{R}^n}$ maximal packing density (actually achieved by some packing, Groemer 1963)
- Kepler conjecture proved by Hales solves sphere packing problem in $3D$
- Viazovska solved sphere packing in dim 8: unique max realized by E_8-lattice
- in dim 24 (Cohn, Kumar, Miller, Radchenko, Viazovska): unique max realized by Leech lattice
- these results use an argument based on modular forms and linear programming bounds for sphere packings



square packing *hexagonal packing*

In $\mathbb{R}^2$ two regular lattice packings of spheres, hexagonal one realizes max density of planar packings (László Fejes Tóth, 1940)



In $\mathbb{R}^3$ Kepler problem optimal sphere packing (Thomas Hales, 1998)

Examples of lattices involved in best sphere packings in low dimensions

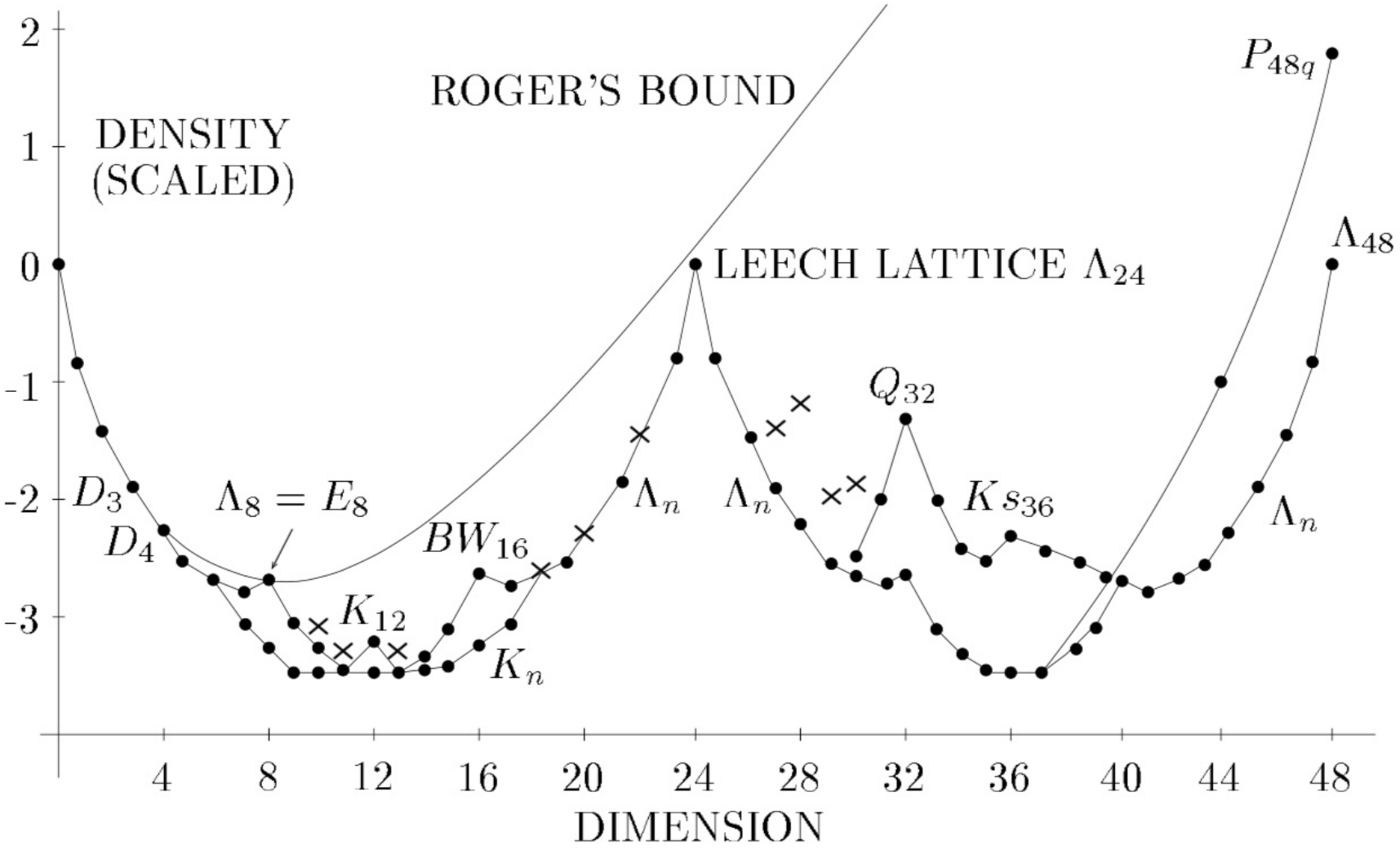
- lattice $A_n = \{x \in \mathbb{Z}^{n+1} \mid \sum_i x_i = 0\}$ simplex lattice (zero-sum hyperplane)
- checkerboard lattice $D_n = \{x \in \mathbb{Z}^n \mid \sum_i x_i \text{ even } \}$
- E_8 lattice $E_8 = D_8 \cup (D_8 + (\frac{1}{2}, \ldots, \frac{1}{2})$
- E_7 orthogonal complement of A_1 inside E_8, etc

The densest lattices in low dimensions are

n	1	2	3	4	5	6	7	8	24
Λ	A_1	A_2	A_3	D_4	D_5	E_6	E_7	E_8	Leech

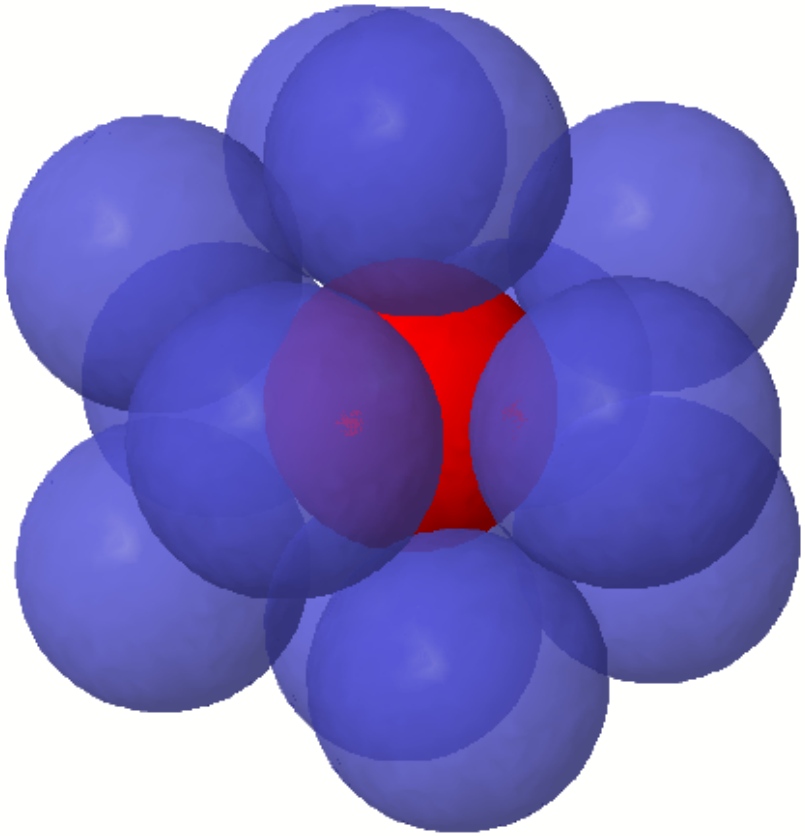
• But... densest lattice typically *not* the max density solution of all packing: in most dimensions densest packing realized by a non-lattice packing

• E_8 maximality is an actual lattice solution!



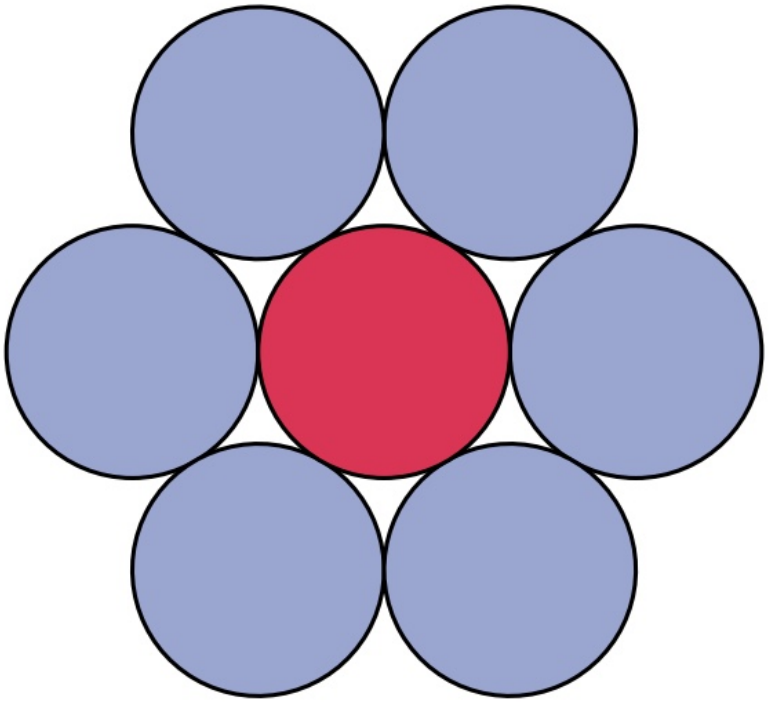
plot of densest sphere packings in low dimensions (Sloane)

Relation to sphere packings and kissing number



example of sphere configuration with kissing number 12

- kissing number: how many balls can touch one given ball at the same time if all the balls have the same size same size
- in $2D$ hexagonal planar lattice packing is optimal solution for (1) the $2D$ kissing number problem, (2) the lattice packing problem, (3) the sphere packing problem



Dim.	Densest packing	Highest kissing number	
1	$\mathbb{Z} \simeq \Lambda_1$	2	
2	$A_2 \simeq \Lambda_2$	6	
3	$A_3 \simeq D_3 \simeq \Lambda_3$	12	
4	$D_4 \simeq \Lambda_4$	24	
5	$D_5 \simeq \Lambda_5$	40	
6	$E_6 \simeq \Lambda_6$	72	
7	$E_7 \simeq \Lambda_7$	126	
8	$E_8 \simeq \Lambda_8$	240	
9	Λ_9	272	(306 from P_{9a})
10	Λ_{10} (P_{10c})	336	(500 from P_{10b})
12	K_{12}	756	(840 from P_{12a})
16	$BW_{16} \simeq \Lambda_{16}$	4320	
24	Leech $\simeq \Lambda_{24}$	196560	

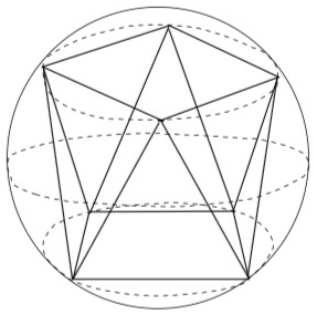
lattice packing and kissing number solutions in low dim (in brackets better non-lattice solutions of max sphere packing density)

Spherical codes

- T. Ericson, V. Zinoviev, *Codes on Euclidean Spheres*, North Holland, 2001.
- optimization questions (in a given dimension n)
  1. given $M \in \mathbb{N}$ find a spherical code with M points such that minimum distance (min angle) between points of the code is as large as possible
  2. given distance $d > 0$ (angle ϕ) find a spherical code with largest number M of points with at least this min distance
- analogs of encoding and decoding optimization questions for q-ary codes

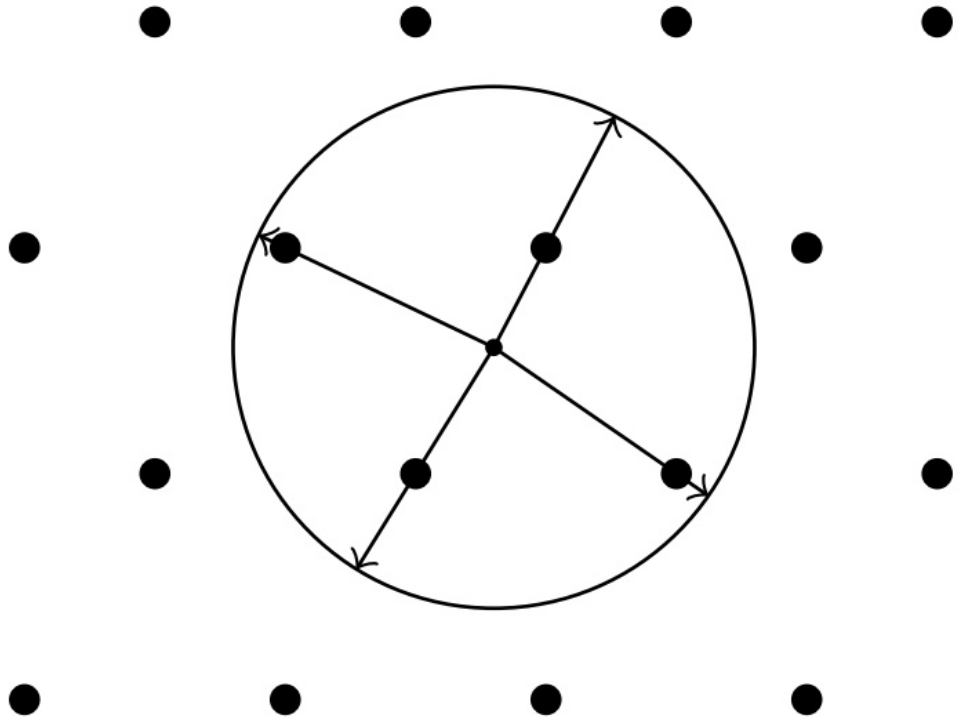
Examples in $3D$ (points on S^2)

For $M = 2, 3, 4$ (antipodal points, equilateral triangle at the equator, regular tetrahedron).
For $M = 8$:



- for angle separation $\phi = \pi/3$ can view points of a spherical code as contact points for an arrangement of touching non-overlapping equal spheres: kissing number problem (maximize M given ϕ)
- Henri Cohn, Yufei Zhao, *Sphere packing bounds via spherical codes*, Duke Math. J. 163 (2014), no. 10, 1965–2002
- upper bound for sphere packing densities are obtained from spherical codes
- by obtaining asymptotic upper bounds for $A(n, \phi)$ of spherical codes (for large n) and deducing from these the upper bounds on the density: for all $n \geq 1$ and for $\pi/3 \leq \phi \leq \pi$

$$\Delta_{\mathbb{R}^n} \leq \sin^n(\phi/2) \cdot A(n, \phi)$$



• estimate of sphere packing density: sphere radius $R \leq 2$ at randomly chosen location (not centered on lattice) contains on average $\Delta \cdot R^n$ sphere centers; project these to surface of the sphere from center; can check they are separated by at least ϕ with $\sin(\phi/2) = 1/R$; so $\Delta \cdot R^n \leq A(n, \phi)$

Spherical codes from binary codes

- C binary $[n, k, d]_2$-code
- identifying $\mathbb{Z}/2\mathbb{Z} = \{\pm 1\}$: code words as subset of the vertices of n-cube centered at origin in $\mathbb{R}^n$ inscribed in sphere S^{n-1} (normalization factor)
- binary code C gives spherical code X_C with parameters

$$\cos\phi = 1 - \frac{2d}{n} \Leftrightarrow \delta(C) = \frac{d}{n} = \sin^2(\phi/2) = \frac{1-\cos\phi}{2}$$

$$R(C) = \frac{\log_2 \#X_C}{n}$$

with maximum (for fixed n and d)

$$R(C)_{max}(n, d) = \frac{\log_2 A(n, \phi(n, d))}{n}$$

- Question: is there an asymptotic bound for spherical codes?

Space of code parameters

- binary codes: $[0,1]^2 \cap \mathbb{Q}$ coordinates (δ, R)
- spherical codes:
  - code rate $R = n^{-1} \log_2 \# X_C$
  - minimum angle $\phi = \phi_{X_C}$ (or $\cos \phi$)
- unbounded: ϕ smaller maximal number of points $A(n, \phi)$ grows, so R unbounded near $\phi \to 0$
- space $\mathbb{R}_+ \times [0, \pi]$

Regions in the space of code parameters

- code points of some spherical code X

$$\mathcal{P} = \{P = (R, \phi) \,|\, \exists X \subset S^{n-1} \,:\, (R, \phi) = (R(X) = \frac{1}{n} \log_2 \#X, \phi_X)\}$$

- accumulation points of set of code parameters

$$\mathcal{A} = \{P = (R, \phi) \,|\, \exists (R_i, \phi_i) \in \mathcal{P} \,:\, (R, \phi) = \lim_i (R_i, \phi_i),\ (R_i, \phi_i) \neq (R, \phi)\}$$

- points surrounded by a 2-ball densely filled by code parameters

$$\mathcal{U} = \{P = (R, \phi) \,|\, \exists \epsilon > 0 \,:\, B(P, \epsilon) \subset \mathcal{A}\}$$

- asymptotic bound:

$$\Gamma = \{(R = \alpha(\phi), \phi) \,|\, \alpha(\phi) = \sup\{R \in \mathbb{R}_+ \,:\, (R, \phi) \in \mathcal{U}\}\}$$

with $\alpha(\phi) = 0$ if $\{R \in \mathbb{R}_+ \,:\, (R, \phi) \in \mathcal{U}\} = \emptyset$

New phenomena with respect to binary codes

- the two regions $\mathcal{A}$ and $\mathcal{U}$ do not coincide
- asymptotic bound is the boundary of the region $\mathcal{U}$ (not of $\mathcal{A}$)
- the part of the region $\mathcal{A}$ that is not in $\mathcal{U}$ consists of sequences of horizontal segments not contained in $\mathcal{U} \cup \Gamma$
- also the asymptotic bound is only non-trivial in a "small angle region"
  - small angles region: $0 \leq \phi \leq \pi/2$
  - large angle region: $\pi/2 < \phi \leq \pi$

Large angle region $\pi/2 < \phi \leq \pi$

- Rankin bound: for $\pi/2 < \phi \leq \pi$

$$A(n, \phi) \leq (\cos\phi - 1)/\cos\phi$$

- bound realized for $-1 \leq \cos\phi \leq -1/n$ while for $-1/n \leq \cos\phi < 0$ one has $A(n, \phi) = n + 1$
- code points lie below the curve

$$R = \frac{1}{n}\log_2(\min\{n+1, \frac{\cos\phi - 1}{\cos\phi}\})$$

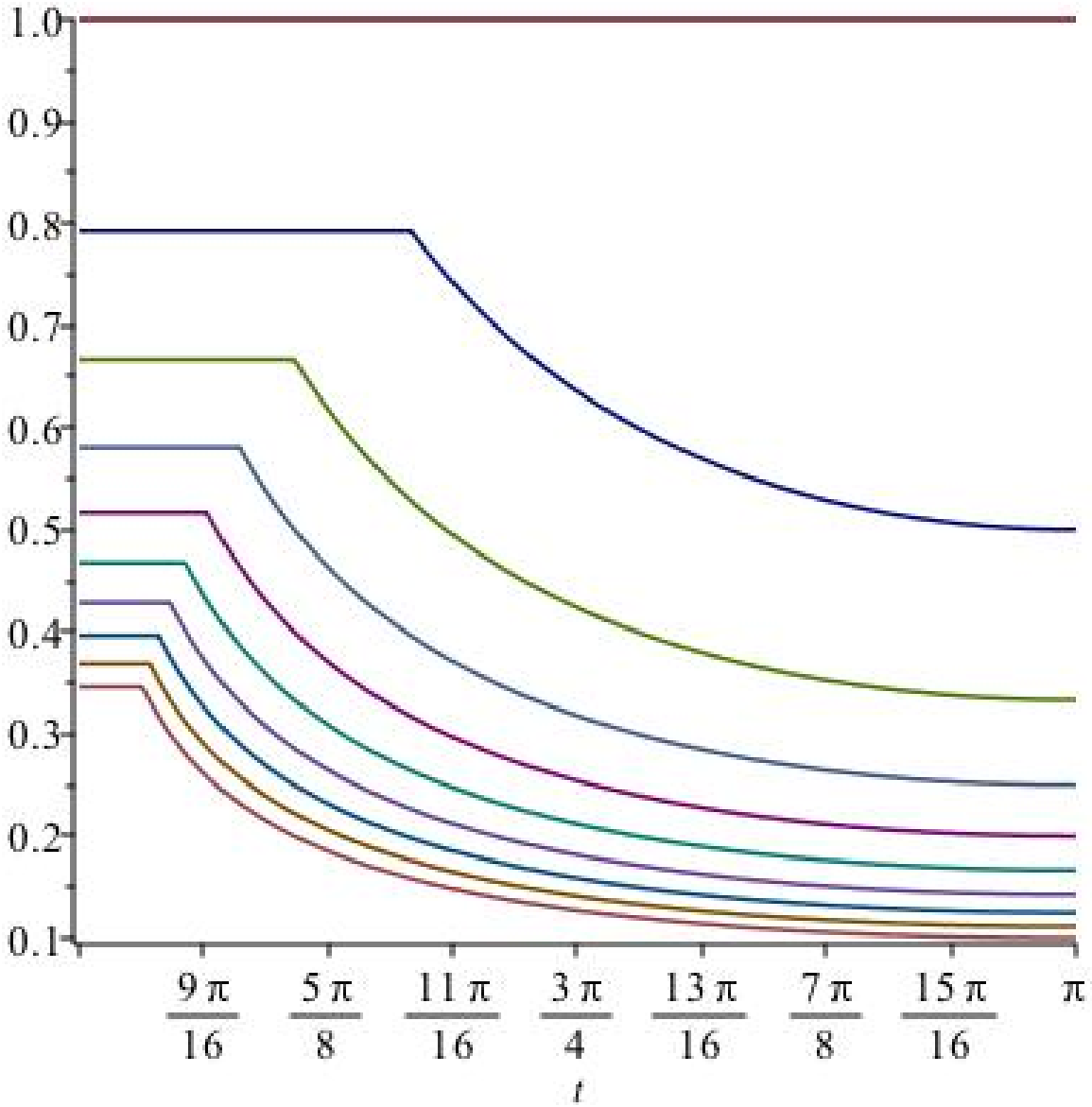
- large $n \to \infty$ behavior

$$R = \frac{\log_2 \#X}{n} \leq \frac{\log_2 A(n, \phi)}{n} \to 0, \qquad \pi/2 \leq \phi \leq \pi$$

$\Rightarrow$ no interesting asymptotic bound in this region

- still contains code points in $\mathcal{A} \setminus \mathcal{U}$ and $\mathcal{P} \setminus \mathcal{A}$

Plots for $n = 1, \ldots, 10$



Estimates in the small angle region

- Kabatiansky–Levenshtein bound: large $n \to \infty$

$$R \leq \frac{\log_2 A(n,\phi)}{n} \leq \frac{1+\sin\phi}{2\sin\phi}\log_2\left(\frac{1+\sin\phi}{2\sin\phi}\right) - \frac{1-\sin\phi}{2\sin\phi}\log_2\left(\frac{1-\sin\phi}{2\sin\phi}\right)$$

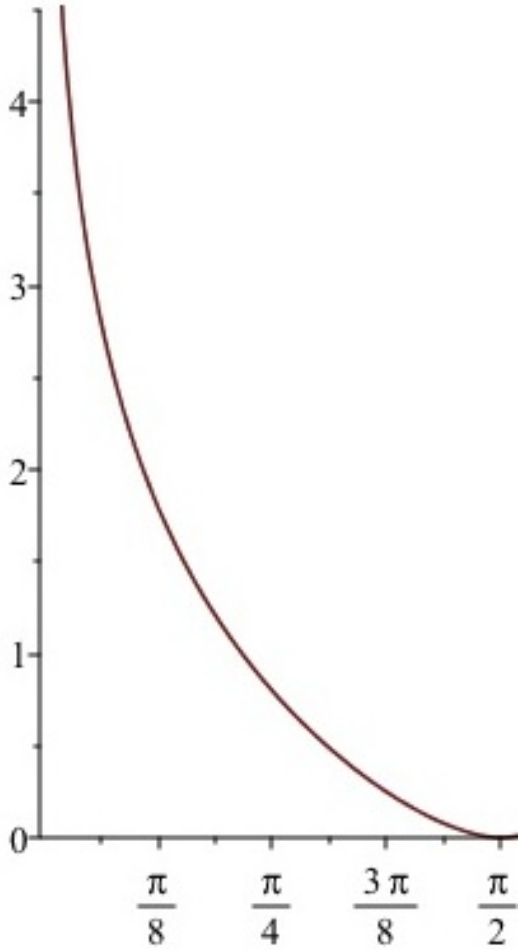
for minimum angle $0 \leq \phi \leq \pi/2$

- for large $n \to \infty$ code parameter in the undergraph

$$\mathcal{S} := \{(R,\phi) \in \mathbb{R}_+ \times [0,\pi] \,:\, R \leq H(\phi)\}$$

$$H(\phi) = \frac{1+\sin\phi}{2\sin\phi}\log_2\left(\frac{1+\sin\phi}{2\sin\phi}\right) - \frac{1-\sin\phi}{2\sin\phi}\log_2\left(\frac{1-\sin\phi}{2\sin\phi}\right)$$

Graph of $H(\phi)$:



• either cutoff on minimum angle $\phi \geq \phi_0$ (e.g. case of sphere packings) or cutoff on $R = \frac{1}{n} \log_2 \#X \leq T$ (more natural for spoiling operations)

Spoiling operations for spherical codes

1. first spoiling operations:
   - binary codes: $C_1 = C \star_i a$ associates to a word $c = (a_1, \ldots, a_n)$ of C the word $c \star_i a = (a_1, \ldots, a_{i-1}, a, a_i, \ldots, a_n)$
   - spherical codes: take code $X_C \subset S^{n-1}$ and inserts S^{n-1} as hyperplane section of S^n
2. second spoiling operation:
   - binary codes: $C_2 = C\star_i$, which is a projection of the code C in the i-th direction
   - spherical codes: $\cos\theta = \langle v_k, v_r \rangle$ angle between two points of code X_C: orthogonal projection along x_i-axis

$$\cos\tilde{\theta} = \frac{n}{n-1}\langle v_k^{\perp_i}, v_r^{\perp_i}\rangle = \frac{n}{n-1}(\cos\theta - \langle v_{k,i}, v_{r,i}\rangle)$$

3. third spoiling operation:
   - binary codes: $C_3 = C(a, i)$ code words with i-th digit a
   - spherical codes: line ℓ and orthogonal hyperplane L through origin of $\mathbb{R}^n$, with $X_3 := X_\ell^\pm = X \cap S_{\ell,\pm}^{n-1}$ intersection with one of the two hemispheres

Main differences: continuous parameters in spoiling operations

- first spoiling operation extends with *continuous parameters* (choice of a hyperplane H): scaling the sphere S^{n-1} and identifying it with the section $H \cap S^n$ to embed new code $X_1 = X \star H$ in S^n
- parameters: $k(X_1) = k(X)$, $n(X_1) = n(X) + 1$ and

$$\cos \phi_{X_1} = \rho_H^2 \cos \phi_X + (1 - \rho_H^2)$$

ρ_H radius of scaled sphere $S_\rho^{n-1} = H \cap S^n$

- second spoiling operation: L hyperplane through origin in $\mathbb{R}^n$ with orthogonal ℓ not containing code points; orthogonal projection $P_L : \mathbb{R}^n \to L \simeq \mathbb{R}^{n-1}$ and normalize vectors: $X_2 = X \star_L \subset S^{n-2}$
- code parameters: $k(X_2) = k(X)$ and $n(X_2) = n(X) - 1$

$$\cos \phi_{X_2} = (1 + u) \cos \phi_X + u, \qquad u = (1 - \xi_{X,L}^2)/\xi_{X,L}^2$$

with $\xi_{X,\ell} = \mathrm{dist}(X, \ell)$

• third spoiling operation also continuous choice of ℓ, L with $X_3 := X_\ell^\pm = X \cap S_{\ell,\pm}^{n-1}$ one hemisphere

• code parameters: $\exists \ell$ with $k(X) - 1 \leq k(X_3) < k(X)$ and minimum angle $\phi(X_3) \geq \phi(X)$

controlling cones: starting with X with code parameters $[n, k, \cos\phi]$

• use spoling operations to obtain code parameters to obtain

1. $[n+1, k, \lambda\cos\phi + 1 - \lambda]$, for all $\lambda \in [0,1]$;
2. $[n-1, k, (1+u)\cos\phi \pm u]$ for $u = (1 - \xi_{X,L})^2/\xi_{X,L}^2$;
3. $[n-1, k-a, \cos\phi]$, for $0 < a < k$.

for $0 \leq \phi \leq \pi/2$

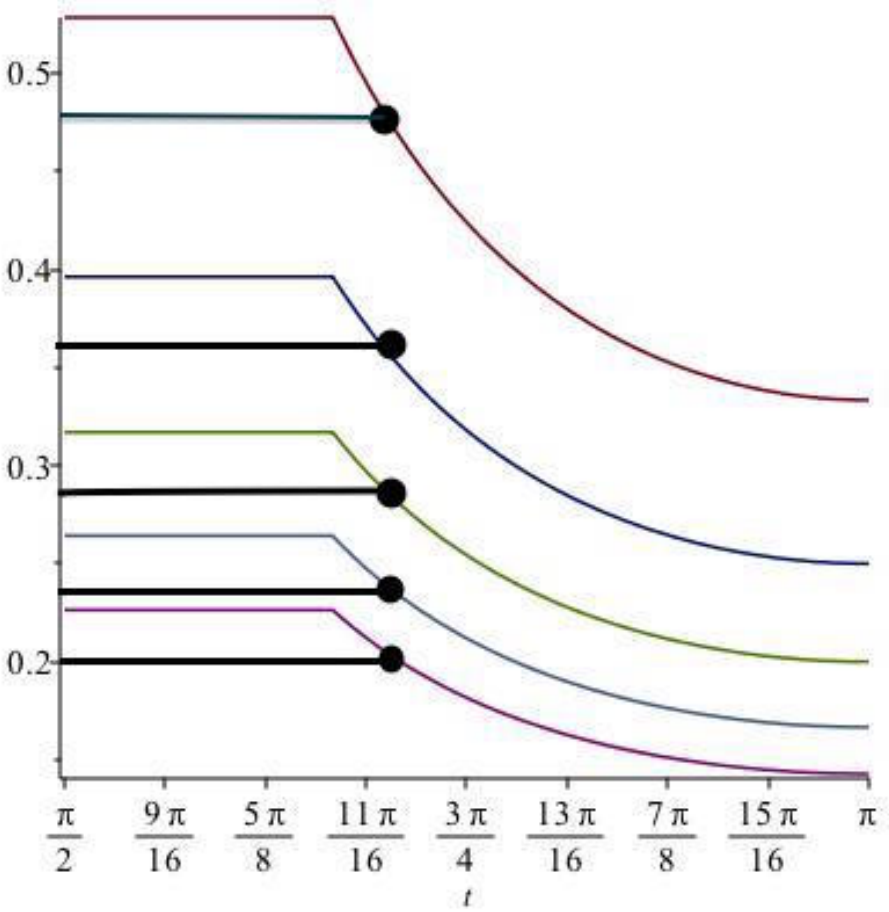
• consequence: if (R, ϕ) code point all line segment

$$\ell_{n,k,\cos\phi} = \{(\frac{n}{n+1}R, \lambda\cos\phi + 1 - \lambda) \,:\, \lambda \in [0,1]\}$$

also made of code points: in $\mathcal{A}$ not always in $\mathcal{U}$

Example of segments in $\mathcal{A}$ not in $\mathcal{U}$

- Rankin examples of spherical codes realizing bound (large angles)
$R(X) = \frac{1}{n}\log_2(\frac{\cos\phi - 1}{\cos\phi})$ for $-1 \leq \cos\phi \leq -1/n$ and
$R(X) = \frac{1}{n}\log_2(n+1)$ for $-1/n \leq \cos\phi < 0$
- apply first spoiling:



Existence of the asymptotic bound

• construct controlling regions $\mathcal{R}_{L,c}(P)$, $\mathcal{R}_{R,c}(P)$, $\mathcal{R}_{U,c}(P)$, $\mathcal{R}_{D,c}(P)$ in a cutoff of undergraph of $H(\phi)$

• use these to constrain position of the asymptotic bound: Γ graph of continuous decreasing $R = \alpha(\phi)$ with $\alpha(\phi) \to \infty$ for $\phi \to 0$ and $\alpha(\pi/2) = 0$.

• set $\mathcal{U}$ is undergraph of this function

$$\mathcal{U} = \{(R, \phi) \, : \, R \leq \alpha(\phi)\}$$

union of all the lower controlling regions $\mathcal{R}_L(P)$ of all points $P \in \Gamma$

• code point $P = (R, \phi) \notin \Gamma$ in region $\mathcal{U}$ iff infinite multiplicity and $\exists$ sequence X_i of spherical codes with $(R(X_i), \phi_{X_i}) = (R, \phi)$ and $n_i \to \infty$ and $\#X_i \to \infty$.

Questions

- applications to sphere packings? (maximal density sphere packings)
- interplay between classical binary (q-ary?) codes and spherical codes
- asymptotic bound and complexity: spherical codes and complexity
- classical to quantum codes (for binary and q-ary: CSSR algorithm): what about spherical codes?