Nested perfect necklaces and normal numbers

Verónica Becher

Universidad de Buenos Aires

Moscow Seminar in Diophantine Analysis - March 23, 2021

A question asked by Korobov

A real x is normal to base b if the fractional parts of $x, bx, b^2x, \ldots$are uniformly distributed in the unit interval. That is, if $(b^n x \mod 1)_{n\geq 0}$ is u.d.

A sequence $(x_n)_{n\geq 1}$ of real numbes in $[0,1)$ is u.d. if the discrepancy

$$D_N((x_n)_{n\geq 1}) = \sup_{\gamma\in[0,1)} \left| \frac{1}{N} \#\{n \leq N : x_n < \gamma\} - \gamma \right|$$ goes to 0 as N to ∞.

Schmidt 1972 proved there is a constant C such that for every $(x_n)_{n\geq 1}$ there are infinitely many Ns, $D_N((x_n)_{n\geq 1}) > C\frac{\log N}{N}$. This is optimal (the van der Corput, the Halton, the Sobol sequences have this discrepancy).

Korobov 1956 asked what is the optimal order of discrepancy achievable by $(b^n x \mod 1)_{n\geq 0}$ for some real x. It is still unknown.

The lowest known $D_N((b^n x \mod 1))_{n\geq 0}$ is $O((\log N)^2/N)$ for a real x constructed by M.Levin 1999 using Pascal triangle matrix modulo 2.

In this talk

Present nested perfect necklaces.

Theorem 1 (Becher and Carton 2019)

For every number x whose base-b expansion is the concatenation of nested $(2^d, 2^d)$-perfect necklaces for $d = 0, 1, 2 \ldots$, $D_N((b^n x)_{n \geq 0})$ is $O((\log N)^2/N)$.

Theorem 2 (Becher and Carton 2019)

The base b-expansion of the number defined by M. Levin 1999 for base b using Pascal triangle matrix modulo 2 is the concatenation of nested $(2^d, 2^d)$-perfect necklaces for $d = 0, 1, 2, \ldots$.

Theorem 3 (Becher and Carton 2019)

For each $d = 0, 1, 2, \ldots$ there are $2^{2^{d+1}-1}$ binary nested $(2^d, 2^d)$-perfect necklaces.

Our observation

Consider all blocks of length n, concatenated in lexicographical order, view it circularly. Each block of length n occurs exactly n times at positions with different modulo n.

For example, for alphabet $\{0, 1\}$

$n = 2$ position

$12\ 34\ 56\ 78$

$00\ 01\ 10\ 11$

$00\ 01\ 10\ 11$ 00 occurs twice, at positions different modulo 2

$00\ 01\ 10\ 11$

$00\ 01\ 10\ 11$ 01 occurs twice, at positions different modulo 2

$00\ 01\ 10\ 11$

$00\ 01\ 10\ 11$ 10 occurs twice, at positions different modulo 2

$00\ 01\ 10\ 11$

$00\ 01\ 10\ 11$ 11 occurs twice, at positions different modulo 2

Our observation

$n = 3$

000 001 010 011 100 101 110 111 000 occurs three times,
000 001 010 011 100 101 110 111 at positions different modulo 3
000 001 010 011 100 101 110 111

000 001 010 011 100 101 110 111 001 occurs three times
000 001 010 011 100 101 110 111 at postions different modulo 3
000 001 010 011 100 101 110 111

$\vdots$

Observation

Not every permutation of the blocks of length n has the property:

$\mathbf{00}$ $10\ 11\ 01$

$\mathbf{000}$ $101\ 001\ 010\ 011\ 100\ 110\ 111$

Perfect necklaces

Definition (Alvarez, Becher, Ferrari and Yuhjtman 2016)

A necklace over a b-symbol alphabet is (n, k)-perfect if each block of length n occurs k times, at positions with different modulo k, for any convention of the starting point.

De Bruijn necklaces are exactly the $(n, 1)$-perfect necklaces.

The (n, k)-perfect necklaces have length kb^n.

Arithmetic progressions yield perfect necklaces

Identify the blocks of length n over a b-symbol alphabet with the set of non-negative integers modulo b^n according to representation in base b.

Theorem (Alvarez, Becher, Ferrari and Yuhjtman 2016)

Let r coprime with b. The concatenation of blocks corresponding to the arithmetic sequence $0, r, 2r, ..., (b^n - 1)r$ yields an (n, n)-perfect necklace.

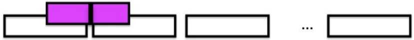
With $r = 1$ we obtain the lexicographically ordered sequence.

Arithmetic progressions yield perfect necklaces

Lemma

Let $\sigma : \{0,..,b-1\}^n \to \{0,..,b-1\}^n$ *be such that for any block* v *of length* n $\{\sigma^j(v) : j = 0,...,b^n - 1\}$ *is the set of all blocks of length* n.

The necklace $[\sigma^0(v)\sigma^1(v)\ldots\sigma^{b^n-1}(v)]$ *is* (n,n)*-perfect if and only if for every block* u *of length* n*, for every* $\ell = 0,\ldots,n-1$ *there is a unique block* v *of length* n *such that* $v(n-\ell-1\ldots n) = u(1\ldots\ell)$ *and* $(\sigma(v))(1\ldots n-\ell) = u(\ell+1\ldots n)$.

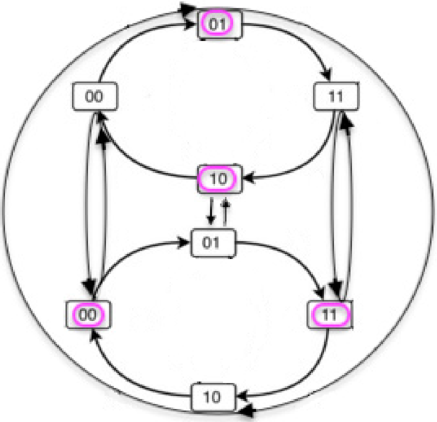


For every length-n block splitted in two parts, there is exactly one matching (a tail of a block and the head of next block).

Astute graphs

Fix b-symbol alphabet. The astute graph $G_{b,n,k}$ is directed, with kb^n vertices.
The set of vertices is $\{0,..b-1\}^n \times \{0,..,.k-1\}$.
An edge $(w,m) \to (w',m')$ if $w(2,,n) = w'(1..n-1)$ and $(m+1) \mod k = m'$

$G_{2,2,2}$



Astute graphs

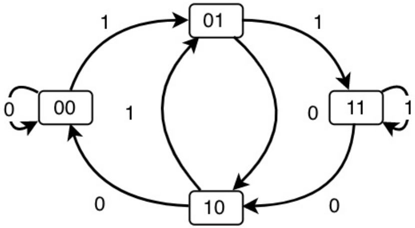
Observation

$G_{b,n,k}$ is Eulerian because it is strongly regular and strongly connected.

Observation

$G_{b,n,1}$ is the de Bruijn graph of blocks of length n over b-symbols.

$G_{2,2,1}$



Eulerian cycles in astute graphs

Each Eulerian cycle in $G_{b,n-1,k}$ gives one (n,k)-perfect necklace.

Each (n,k)-perfect necklace can come from many Eulerian cycles in $G_{b,n-1,k}$

Theorem (Alvarez, Becher, Ferrari and Yuhjtman 2016)

The number of (n,k)-perfect necklaces over a b-symbol alphabet is

$$\frac{1}{k}\sum_{d_{b,k}|j|k} e(j)\varphi(k/j)$$

where

- $d_{b,k} = \prod p_i^{\alpha_i}$*, such that* $\{p_i\}$ *is the set of primes that divide both* b *and* k*, and* α_i *is the exponent of* p_i *in the factorization of* k*,*
- $e(j) = (b!)^{jb^{n-1}} b^{-n}$ *is the number of Eulerian cycles in* $G_{b,n-1,j}$
- φ *is Euler's totient function*

Normal sequences as sequences of Eulerian cycles

Theorem (proved first by Ugalde 2000 for de Bruijn)

The concatenation of (n, k)-perfect necklaces over a b-symbol alphabet, for arithmetically increasing (n, k) is normal to the b-symbol alphabet.

The proof is a direct application of Piatetski-Shapiro theorem.

In worst case, $D_N((b^n x)_{n\geq 0}) = \Theta(\sqrt{(\log\log N)/\log N})$, Cooper and Heitsch, 2010

Corollary

The concatenation of lexicografically ordered (n, n)-perfect necklaces for $n = 1, 2, \ldots$ is normal; Champernowne's sequence is normal.

Nested perfect necklaces

Definition

An (n, k)*-perfect necklace over a* b*-symbol alphabet is* *nested*
if $n = 1$ *or it is the concatenation of* b *nested* $(n-1, k)$*- perfect necklaces.*

For example, for alphabet $\{0, 1\}$, a nested $(2, 2)$-perfect necklace

$$\underbrace{0011}_{\text{(1,2)-perfect}}\ \underbrace{0110}_{\text{(1,2)-perfect}}$$

The lexicographic order yields a perfect necklace but not nested,

$$\underbrace{00\ 01\ 02}_{\text{not (1,2)-perfect}}\ \underbrace{10\ 11\ 12}_{\text{not (1,2)-perfect}}\ \underbrace{20\ 21\ 22}_{\text{not (1,2)-perfect}}$$

Nested perfect necklaces

These following 8 blocks are $(1,4)$-perfect necklaces:

$$\begin{array}{ll} 00001111 & 01011010 \\ 00111100 & 01101001 \\ 00011110 & 01001011 \\ 00101101 & 01111000 \end{array}$$

The concatenation in each row is a $(2,4)$-perfect necklace.
The concatenation of the first two rows is a nested $(3,4)$-perfect necklace.
The concatenation of the last two rows is a nested $(3,4)$-perfect necklace.
The concatenation of all rows is a nested $(4,4)$-perfect necklace.

Proof sketch of Theorem 1

Theorem 1

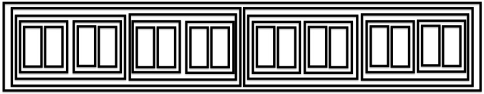
For every number x whose base-b expansion is the concatenation of nested $(2^d, 2^d)$-perfect necklaces for $d = 0, 1, 2\ldots$, $D_N((b^n x)_{n\geq 0})$ is $O((\log N)^2/N)$.

Proof sketch of Theorem 1

Observation

Assume a b-symbol alphabet. For a nested (n, n)-perfect necklace x,

- *each block of length n occurs n times in x, at positions with different congruence modulo n.*
- *for every $i = 1, \dots n$, x is the concatenation of b^{n-i} nested (i, n)-perfect necklaces. So, in every segment of length nb^i starting at a position multiple of nb^i, each block of length i occurs $1 \pm 2\varepsilon$ times, for $\varepsilon \leq 1$ at positions in each congruence class.*



Proof sketch of Theorem 1

Given N, we need to bound $D_N((b^n x)_{n\geq 1})$. Let m and M be such that N is the sum of length of nested $(2^d, 2^d)$-perfect necklaces, $i = 0, ..2^m - 1$, plus M,

$$N = \Big(\sum_{i=0}^{2^m-1} 2^i b^{2^i}\Big) + M, \qquad 0 \leq M < 2^m b^{2^m}$$

$$x = \underbrace{\boxed{2^0 b^{2^0}}\ \boxed{2^1\ b^{2^1}}\ \ldots\ \boxed{2^{m-1} b^{2^{m-1}}}\ \boxed{M}}_{N}\ \ 2^m b^{2^m} - M$$

Since segment M is an incomplete nested perfect necklace, its discrepancy determines the discrepancy of segment N.
Since N is $O\big(2^{2^m}\big)$ then $O\big(\log N\big) = O\big(2^m\big)$.

Proof sketch of Theorem 1

Write M as the sum of length of n_i nested $(i, 2^m)$-perfect necklaces, $i = 0, .., 2^m - 1$, plus M_0

$$M = M_0 + 2^m \sum_{i=0}^{2^m-1} n_i b^i, \qquad M_0 < 2^m \text{ and } n_i \in \{0, ..., b-1\}$$

$$\boxed{M} = \boxed{M_0}\boxed{n_0 2^m b^0}\boxed{n_1 2^m b^1}\boxed{n_2 2^m b^2}\ldots\boxed{n_{2^m-1} 2^m b^{2^m-1}}$$

In segment M,

- at most $b2^m$ nested $(i, 2^m)$-perfect necklaces counting $i = 0, .., 2^m - 1$
- in each, we consider positions in 2^m congruence classes
- for each $(i, 2^m)$-necklace and congruence class, difference between actual and expected number of occurrences of any block of length i is at most 2.
  This is at most $b2^m \times 2^m \times 2 = O(2^m \times 2^m)$.

We conclude $D_N((b^n x)_{n\geq 0}) = O\big((\log N)^2/N\big)$. □

Proof sketch of Theorem 2

Theorem 2

The base b-expansion of the number defined by M. Levin 1999 for base b using the Pascal triangle matrix modulo 2 is the concatenation of nested $(2^d, 2^d)$-perfect necklaces for $d = 0, 1, 2, \ldots$.

Levin's construction

- Levin's constant λ is the number whose base b-expansion is

$$\lambda = 0.\lambda_0\lambda_1\lambda_2\ldots$$

- For $d = 0, 1, 2, \ldots$ define the matrix M_d in $\mathbb{F}_2^{2^d \times 2^d}$ and consider the elements of $\mathbb{F}_b^{2^d}$ in increasing order

$$w_0, w_1, \ldots, w_{b^{2^d}-1}$$

  Identify vectors of $\mathbb{F}_b$ with blocks of symbols in $\{0, .., b-1\}$. Thus, each $(M_d w_i)$ is identified with a block of length 2^d.

- For $d = 0, 1, 2, \ldots$ define λ_d as

$$\lambda_d = (M_d w_0)\ldots(M_d w_{b^{2^d}-1})$$

Pascal triangle matrices modulo 2

Define a family of matrices using Pascal triangle modulo 2,

$$\begin{array}{cccccc} \ldots & 1 & 1 & 1 & 1 & 1 \\ \ldots & 5 & 4 & 3 & 2 & 1 \\ \ldots & 15 & 10 & 6 & 3 & 1 \\ \ldots & 35 & 20 & 10 & 4 & 1 \\ \ldots & 70 & 35 & 15 & 5 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \longrightarrow \begin{array}{cccccc} \ldots & 1 & 1 & 1 & 1 & 1 \\ \ldots & 1 & 0 & 1 & 0 & 1 \\ \ldots & 1 & 0 & 0 & 1 & 1 \\ \ldots & 1 & 0 & 0 & 0 & 1 \\ \ldots & 0 & 1 & 1 & 1 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

Pascal triangle matrices modulo 2

Define a family of matrices using Pascal's triangle modulo 2,

$$\begin{array}{cccccc} \ldots & 1 & 1 & 1 & 1 & 1 \\ \ldots & 5 & 4 & 3 & 2 & 1 \\ \ldots & 15 & 10 & 6 & 3 & 1 \\ \ldots & 35 & 20 & 10 & 4 & 1 \\ \ldots & 70 & 35 & 15 & 5 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \longrightarrow \begin{array}{cccccc} \ldots & 1 & 1 & 1 & 1 & 1 \\ \ldots & 1 & 0 & 1 & 0 & 1 \\ \ldots & 1 & 0 & 0 & 1 & 1 \\ \ldots & 1 & 0 & 0 & 0 & 1 \\ \ldots & 0 & 1 & 1 & 1 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

For $d = 0$, M_d has dimension $2^0 \times 2^0$

$$M_0 = (1)$$

Matrices de Pascal Módulo 2

Define a family of matrices using Pascal's triangle modulo 2,

$$\begin{array}{cccccc} \ldots & 1 & 1 & 1 & 1 & 1 \\ \ldots & 5 & 4 & 3 & 2 & 1 \\ \ldots & 15 & 10 & 6 & 3 & 1 \\ \ldots & 35 & 20 & 10 & 4 & 1 \\ \ldots & 70 & 35 & 15 & 5 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \longrightarrow \begin{array}{cccccc} \ldots & 1 & 1 & 1 & 1 & 1 \\ \ldots & 1 & 0 & 1 & 0 & 1 \\ \ldots & 1 & 0 & 0 & 1 & 1 \\ \ldots & 1 & 0 & 0 & 0 & 1 \\ \ldots & 0 & 1 & 1 & 1 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

For $d = 1$, M_d has dimension $2^1 \times 2^1$

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Matrices de Pascal Módulo 2

Define a family of matrices using Pascal's triangle modulo 2,

$$\begin{array}{cccccc} \dots & 1 & 1 & 1 & 1 & 1 \\ \dots & 5 & 4 & 3 & 2 & 1 \\ \dots & 15 & 10 & 6 & 3 & 1 \\ \dots & 35 & 20 & 10 & 4 & 1 \\ \dots & 70 & 35 & 15 & 5 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \longrightarrow \begin{array}{cccccc} \dots & 1 & 1 & 1 & 1 & 1 \\ \dots & 1 & 0 & 1 & 0 & 1 \\ \dots & 1 & 0 & 0 & 1 & 1 \\ \dots & 1 & 0 & 0 & 0 & 1 \\ \dots & 0 & 1 & 1 & 1 & 1 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

For $d = 2$, M_d has dimension $2^2 \times 2^2$

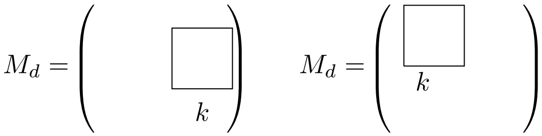
$$M_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Alternative formulation Pascal triangle matrices modulo 2

$$M_0 = (1), \qquad M_{d+1} = \begin{pmatrix} M_d & M_d \\ 0 & M_d \end{pmatrix}$$

- M_d in $\mathbb{F}_2^{2^d \times 2^d}$.
- M_d is invertible.
- The first row of M_d is the vector of 1s
- The last column of M_d is the vector of 1s

Invertible submatrices



Lemma (Levin 1999 from Bicknell and Hoggart 1978)

For $d \geq 0$, the following submatrices of M_d are invertible

- k *rows and the last* k *columns*
- *the first* k *rows and* k *columns*

Levin's number

Observation

For every $d \geq 0$, λ_d is the concatenation of all blocks of length 2^d in some order.

$$\lambda = 0.\underbrace{0\,1}_{\lambda_0}$$
$$\underbrace{00\,11\,10\,01}_{\lambda_1}$$
$$\underbrace{0000\,1111\,1010\,0101\,1100\,0011\,0110\,1001\,1000\,0111\,0010\,1101\,0100\,1011\,1110\,0001}_{\lambda_2}$$
$$\ldots$$

Levin's number

Observation

Assume $b = 2$. For every d and for every even n, $M_d w_n$ and $M_d w_{n+1}$ are complementary blocks.

λ = 0. 0 1
00 11 10 01
0000 1111 1010 0101 1100 0011 0110 1001 1000 0111 0010 1101 0100 1011 1110 0001
...

Sketch of proof of Theorem 3

Theorem 3

For each $d = 0, 1, 2, \ldots$ *there are* $2^{2^{d+1}-1}$ *binary nested* $(2^d, 2^d)$*-perfect necklaces.*

Matrices like Levin's

Definition

For $d = 0, 1, 2, \ldots$, a tuple $\nu = (\nu_1, \ldots, \nu_{2^d})$ of 2^d non-negative numbers is suitable if $\nu_{2^d} = 0$ and for every i, ν_{i+1} is equal to ν_i or $\nu_i - 1$.

- $(1, 1, 1, 0)$ is suitable;
- $(4, 3, 1, 0)$ is not suitable;
- $(3, 2, 1, 0)$ is suitable.

Observation

For each $d = 0, 1, 2, \ldots$ *there are* $2^{2^d - 1}$ *suitable tuples.*

Matrices like Levin's

Assume σ is the rotation operation.
If $\nu = (\nu_1, \ldots, \nu_{2^d})$ is suitable and $C_1, \ldots, C_{2^d}$ are columns of M_d,

$$M_d^\nu = (\sigma^{\nu_1}(C_1), \ldots, \sigma^{\nu_{2^d}}(C_{2^d}))$$

Matrices like Levin's

For $d = 2$ there are $2^{2^d-1} = 8$ suitable tuples, hence 8 matrices,

$$M_2^{(0,0,0,0)} \quad M_2^{(1,0,0,0)} \quad M_2^{(1,1,0,0)} \quad M_2^{(2,1,0,0)}$$

$$\begin{pmatrix}1&1&1&1\\0&1&0&1\\0&0&1&1\\0&0&0&1\end{pmatrix} \begin{pmatrix}0&1&1&1\\1&1&0&1\\0&0&1&1\\0&0&0&1\end{pmatrix} \begin{pmatrix}0&0&1&1\\1&1&0&1\\0&1&1&1\\0&0&0&1\end{pmatrix} \begin{pmatrix}0&0&1&1\\0&1&0&1\\1&1&1&1\\0&0&0&1\end{pmatrix}$$

$$M_2^{(1,1,1,0)} \quad M_2^{(2,1,1,0)} \quad M_2^{(2,2,1,0)} \quad M_2^{(3,2,1,0)}$$

$$\begin{pmatrix}0&0&0&1\\1&1&1&1\\0&1&0&1\\0&0&1&1\end{pmatrix} \begin{pmatrix}0&0&0&1\\0&1&1&1\\1&1&0&1\\0&0&1&1\end{pmatrix} \begin{pmatrix}0&0&0&1\\0&0&1&1\\1&1&0&1\\0&1&1&1\end{pmatrix} \begin{pmatrix}0&0&0&1\\0&0&1&1\\0&1&0&1\\1&1&1&1\end{pmatrix}$$

The number of binary perfect necklaces

For each suitable ν and for each vector $z \in \mathbb{F}_2^{2^d}$,

$$(M_d^\nu w_0 \oplus z) \dots (M_d^\nu w_{2^{2^d}-1} \oplus z).$$

is a nested $(2^d, 2^d)$-perfect necklace.

Since there are $2^{2^d} - 1$ suitable tuples ν and there are 2^{2^d} different vectors $z \in \mathbb{F}_2^{2^d}$, the number of binary nested $(2^d, 2^d)$-perfect necklaces is at least

$$2^{2^d-1} \times 2^{2^d}$$

By a graph theoretical argument we know that there can be no more.

Nested marvelous necklaces

Definition

A necklace over a a b-symbol alphabet is nested (n, k)-marvelous if all blocks of length n occur exactly k times, and in case $n > 1$ it is the concatenation of b nested $(n-1, k)$-marvelous necklaces.

This is nested $(3, 3)$-marvelous, not perfect,

$$000111\ 011001\ 000111\ 101010$$

Theorem (Becher and Carton 2020)

For every number x whose base-b expansion is the concatenation of nested $(2^d, 2^d)$-marvelous necklaces, $D_N((b^n x)_{n \geq 0})$ is $O\left((\log N)^2/N\right)$.

References

- N. Alvarez, V. Becher, P. Ferrari and S. Yuhjtman. "Perfect necklaces", *Advances of Applied Mathematics* 80:48–61, 2016.
- V. Becher and O. Carton. "Normal numbers and nested perfect necklaces", *Journal of Complexity* 54:101403, 2019
- D. Champernowne. "The Construction of Decimals Normal in the Scale of Ten." *Journal London Mathematical Society*, S1-8(4):254-260, 1933.
- J. Cooper, C. Heitsch. The discrepancy of the lex-least de Bruijn sequence. *Discrete Mathematics*, 310 (67): 1152-1159, 2010.
- M. B. Levin. On the discrepancy estimate of normal numbers. *Acta Arithmetica* 88(2):99–111, 1999.
- N. Korobov. On completely uniform distributions and jointly normal numbers. *Izv. AN SSSR, ser. matem.*, 20, 1956.
- W. Schmidt. Irregularities of distribution VII. *Acta Arithmetica*, 21:45–50, 1972.
- D.D.Wall."Normal numbers",Ph.D.Thesis University of California Berkeley,1949.