On a recursive hierarchy for Numbers algorithmically emerging from processing binary sequences

Françoise Chaitin-Chatelin[1]

CERFACS Report TR/PA/03/08

[1]Université Toulouse 1 and CERFACS, 42 Ave. G. Coriolis, 31057 Toulouse Cedex 1, France. E-mail: chatelin@cerfacs.fr

Abstract

In the first part of the report, we show how binary sequences of length n, $n \geq 1$ can be algorithmically processed to produce arithmetical order, weight, measure and connectedness. The process is based primarily on hypermultiplication for $n = 2^k$, $k \geq 0$, and secondarily on addition for arbitrary n. Analysis of the values $k = 0$ to 4 is the basis for the iterative process, $k \geq 4$. The dynamics of the process requires to explore sequences at least as long as $n = 2^{16}$. In the second part, we show that the use of different models for the Cantor space $\{0, 1\}^{\mathbb{N}}$ allows to tame the combinatorial complexity. The report ends by suggesting directions for future research.

Key words: binary sequence, Cantor space, order, weight, measure, connectedness, hypermultiplication, hypercomplex algebra, Mersenne number, Fermat number, Sierpinski triangle, spectral analysis, Fourier analysis, modular arithmetic, Newcomb-Borel paradox.

"Small is beautiful" *E. F. Schumacher*

1. Introduction

Using the notion of program-size complexity [1] G. J. Chaitin has shown that lack of structure, or randomness, is the rule for sets of binary sequences (or bit strings) of length $n \geq 1$, $n \in \mathbb{N}^\star$. This is well illustrated by the Ω number whose digits are not computable in the classical sense of Turing [2].

In this report, we wish to explore the other side of the coin: we look for an algorithmic process by which the natural multilevel pattern present in such sets can emerge. The process orders naturally the integers in $\mathbb{N}$ and yields the continuous interval $[0,1]$ in $\mathbb{R}$ and the unit disk in $\mathbb{C}$ when we let $n \to \infty$. To do this, we use two remarkable and basic properties:

i) the recursive structure of the triangle of ones of side length $n \geq 1$, starting from the number 1 for $n = 1$, a fact which accounts for many convergence proofs of algorithms in Linear Algebra [4],

ii) the possibility, for $n = 2^k$, $k \in \mathbb{N}$, to equip the set C_n of binary sequences with an inductively defined *multiplication* [3, 5, 6] which confers to C_n a structure of hypercomplex algebra on $\mathbb{Z}_2 = \{0, 1\}$.

The natural order is easily established with i). The novelty comes from ii), that is, from considering numbers as vectors of dimension 2^k in algebras $B_k (= C_{2^k})$. This induces a natural hierarchy of levels k which connects the level k to its predecessor $k-1$ (sequences of half length).

The existence of such a hierarchy allows either to use inductive reasoning (k increases for $k \geq 4$) or to use recursivity (k decreases), which is the modern name for the descent method of Pythagoras and Fermat.

By careful inspection of the small values $k = 0$ to 4 (i.e., $n = 1, 2, 4, 8, 16$) we find that the four algebras B_0 to B_3 are special cases for one reason or another.

To get a completely generic behavior, we need to consider sequences having at least $16 = 2^4$ bits. There are at least $2^{2^4} = 65{,}536$ of them. It is only at the level $k = 4$ that the algebra B_k does begin to display the complete variety of complexity which enables to base on it an inductive ordering process for $k \geq 4$. We shall see why this is just the beginning, that is, why more powerful sieves than order call for no less than two levels of exponentials.

2. The set C_n of binary sequences of length n, $n \geq 1$

The set of binary sequences of length n, $n \geq 1$, is the set $C_n = \{0,1\}^n$, also known as the n-cube. It contains $\mathrm{card}(C_n) = 2^n$ different sequences. For n arbitrary, we can put an order on a small subset of sequences in C_n, in the following way.

2.1. Partial order by the triangle rule

We label $0, 1, 2, \ldots, n$ the $n+1$ first sequences displayed in Figure 2.1. The remaining sequences (marked with $\times$) are unordered.

$$\left.\begin{array}{ccccccccccc}
0 & 1 & 1 & . & . & . & 1 & \times & \times & \ldots & \times \\
. & 0 & 1 & . & & & 1 & . & . & & . \\
. & . & 0 & & . & & . & . & . & & . \\
. & . & . & & & . & 1 & . & . & & . \\
0 & 0 & 0 & . & . & 0 & 1 & \times & \times & \ldots & \times
\end{array}\right\} n$$

Figure 2.1 *Partial order on C_n by the triangle rule*

We call this rule to define a partial order the *triangle rule.*

For any ordered sequence, its ordinal number is equal to its number of consecutive ones.

The recursive structure of the triangle implies that, as n increases, more sequences are ordered. The ones already ordered keep their ordinal: only zeros are appended to them. Therefore the natural order on $\mathbb{N}$ emerges algorithmically from the triangle rule applied to C_n, $n = 1, 2, 3, \ldots$. The order is a map defined on a subset C'_n of C_n onto $\{0, 1, 2, \ldots, n\} \subset \mathbb{N}, \mathrm{card}(C'_n) = n+1$. For $n = 1 = 2^0$ we get the total order defined by the identity map: $C_1 = C'_1 = \{0,1\} \to \{0,1\}$. For $n = 2 = 2^1$, we also get a total order on $C_2 = C'_2$ defined by: $00 \to 0, 10 \to 1, 11 \to 2, 01 \to 3$. For $n = 3$ only half of the $2^3 = 8$ sequences are ordered by $000 \to 0, 100 \to 1, 110 \to 2, 111 \to 3$, and the other 4 sequences remain unordered.

2.2. Left and right weights in base 2

In addition to an ordinal, we can give each sequence a *weight* by considering its integral numerical value in base 2 (that is, using the successive powers of

2, namely $1, 2, 2^2, \ldots, 2^{n-1}$). We find it useful to use the metaphor of **weight** to describe the representation of a number in a given base. Such a metaphor was in common use amongst mathematicians of the 16th and 17th centuries [9, 10].

A given sequence receives two numerical values in $\mathbb{N}$ (weights) Lw and Rw, depending whether the sequence is considered as written i) from *left* to right (the exponents of 2 increase) as in the dyadic representation Lw, or ii) from *right* to left (the exponents decrease) as in the binary representation Rw. For example, (11000) is weighted i) $Lw = 1+2 = 3$ and ii) $Rw = 2^3+2^4 = 24$. One remarks that Rw requires to know $n = 5$, which is *not* the case for Lw.

The sequence of all ones (denoted $\mathbf{e}_n$) with ordinal n has weight $Lw = Rw = 1+2+\cdots+2^{n-1} = 2^n - 1 = M_n$, the Mersenne number with exponent n. It is the sequence of maximal weight in C_n.

2.3. Left and right measures in base 2

We consider now *negative* values for the exponents of 2, and we associate to $b = (b_0, \ldots, b_{n-1}) \in C_n$ a **measure** $\lambda(b)$ which is a rational value between 0 and 1:

$$L\lambda(b) = \sum_{i=0}^{n-1} b_i 2^{-(i+1)}, \quad R\lambda(b) = \sum_{i=1}^{n} b_{n-i} 2^{-i}.$$

We note that $\lambda(\mathbf{e}_n) = 2^{-1} + \cdots + 2^{-n} = \frac{M_n}{2^n} < 1$: $\mathbf{e}_n$ is the sequence of maximal measure in C_n.

Lemma 1.1 *For* $n \in \mathbb{N}^\star$, $\lambda_{\max}^{(n)} = \frac{w_{\max}^{(n)}}{\mathrm{card}(C_n)} \to 1$ *as* $n \to \infty$

Proof. Clear. The upper index (n) indicates that λ and w are defined relative to C_n.

2.4. Up and down connectedness

The **connectedness** of a sequence $b = (b_0, \ldots, b_{n-1}) \in C_n$ is defined in two versions, up (U) and down (D): Uc (resp. Dc) $= 0$ if b_0 (resp. b_{n-1}) $= 0$ else it is the total number of consecutive ones.

For $\mathbf{e}_n$, $c(\mathbf{e}_n) = \mathrm{ord}(\mathbf{e}_n) = n$, its ordinal value.

In summary, we have defined on C_n, $n \geq 1$, four arithmetic applications:

1) a partial order $\mathrm{ord} : C'_n \subset C_n \to \{0, 1, \ldots, n\}$

2) a connectedness $c : C_n \to \{0, \ldots, n\}$

3) a weight $w : C_n \to \{0, \ldots, M_n\}$

4) a measure $\lambda : C_n \to [0, \frac{M_n}{2^n}]$

For $n = 1, 2$, $C'_n = C_n$. The three maps ord, c, w take values in $\mathbb{N}$, and λ takes values in $\mathbb{Q}$. In general, the functions c, w, λ are not symmetric in the following sense: for w and λ (resp. c) left and right (resp. up and down) values differ. They are identical for 0 and $\mathbf{e}_n$.

The *triangle rule* appears as a geometric mechanism to order the sequences according to increasing up connectedness: $Uc(b) = \text{ord}(b)$ from 0 to n.

For any $n \neq 2^k$, $k \geq 0$, the structure of C_n is that of a linear vector space on $\mathbb{Z}_2$. But for $n = 2^k$, the structure of C_{2^k} is *much richer:* C_{2^k} can be equipped with a multiplication which makes it a binary algebra on $\mathbb{Z}_2$, that we denote B_k, $k \geq 0$, with $B_0 = \mathbb{Z}_2$. In the sequel, when $n = 2^k$, we automatically consider C_{2^k} with the algebraic structure induced by $+$ and $\times$. Therefore we identify $C_{2^k} = B_k$, $k \geq 0$. We observe that $C_1 = B_0 = \mathbb{Z}_2$ and $C_2 = B_1$ have been totally ordered by the triangle rule.

3. The binary algebras B_k, $k \geq 0$

We look at the hierarchical connection between 2 (resp. 3) successive algebras induced by multiplication (resp. order) on B_k.

3.1. The Dickson-Albert doubling process on $\mathbb{Z}_2$

L. E. Dickson [3] proposed to define an unbounded sequence of algebras of dimension 2^k, $k \geq 0$, on the basis field $\mathbb{R}$, by a doubling process which uses addition, conjugation and multiplication defined at the preceding level $k-1$ [5]. A. A. Albert [11] extended the idea to a general basis field K, including K of characteristic 2.

The idea was developed in [5, 6, 7, 8] for the particular basis fields $K = \mathbb{R}$ (infinite and continuous) and $\mathbb{Z}_2$ (finite and discrete). In the later case, conjugation is reduced to identity.

From $B_0 = \mathbb{Z}_2 = \{0, 1\}$, we define inductively the binary algebra B_k of dimension 2^k, $k \geq 1$, with $B_k = B_{k-1} \times B_{k-1}$ by means of the two operations $+$ and $\times$:

addition $(a, b) + (c, d) = (a + c, b + d)$,

multiplication $(a, b) \times (c, d) = (ac + bd, bc + ad)$,

where $a, b, c, d \in B_{k-1}$, and ac denotes $a \times c$ in B_{k-1}.

All algebras defined by the Dickson-Albert process are *quadratic* : $x^2 = \alpha x + \beta 1$, where $\alpha, \beta \in K$ and 1 represents the unit vector $(1 \times 1 = 1)$.

This yields, for $K = \mathbb{R}$, $x^2 = -x \times \overline{x} + (x + \overline{x})x$, where $\overline{x}$ is the conjugate of x, $x + \overline{x} = 2(\mathrm{Re}\, x) \in \mathbb{R}$ is twice the real part of x, $x \times \overline{x} = \| x \|^2 \in \mathbb{R}^+$ is the square of the euclidean norm of x.

As k increases from 0 to 3, the four real hypercomplex algebras ($\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ and $\mathbb{G}$) offer a decreasing variety of possibilities with respect to commutativity, associativity, isometry and division by 0 [5]. For $k \geq 4$, they are not isometric, they are without division; they remain power associative and quadratic.

For $K = \mathbb{Z}_2$, we get $x^2 = 0$ or 1. All B_k are associative, commutative, self-conjugate finite algebras. Only B_0 is a field, therefore a division algebra. The structure of the set $Z(a)$ of zero divisors for $a \in B_k$, $k \geq 1$ is studied in [7, 8]. The difficulty of the analysis of B_k comes from the exponential explosion : $N_k = \mathrm{card}(B_k) = 2^{2^k}$ reaches, already for $k = 4$, the respectable value $N_4 = 2^{16} = 65536$.

We shall see how the four arithmetic maps ord, c, w and λ introduced on C_n, can be used as sieves for $n = 2^k$ which exploit the inductive structure $B_k = B_{k-1} \times B_{k-1}$, for $k \geq 1$.

3.2. $n = 2^k$: order by the extended triangle rule, $k \geq 1$

For $k = 1$, $n = 2$, the triangle rule yields the total order $0 < 1$.

1) For $k = 2$, we extend the triangle of side length 4 into a parallelogram to order $8 = 5 + 3$ sequences out of 16, as shown in Table 3.1.

	0	1	1	1	1	0	0	0	×	×	×	×	×	×	×	×
	0	0	1	1	1	1	0	0	×							×
B_2	0	0	0	1	1	1	1	0	×							×
	0	0	0	0	1	1	1	1	×	×	×	×	×	×	×	×
ord	0	1	2	3	4	5	6	7	–	–	–	–	–	–	–	–

Table 3.1 *Order by parallelogram in B_2.*

The ordered sequences 0 to 7 can be expressed as pairs of ordinals in $B_1 = \{0, 1, 2, 3\}$:

B_2	0	1	2	3	4	5	6	7
$B_1 \times B_1$	(0, 0)	(1, 0)	(2, 0)	(2, 1)	(2, 2)	(3, 2)	(0, 2)	(0, 3)

This is displayed by the 4×4 lattice in Figure 3.1 (left).

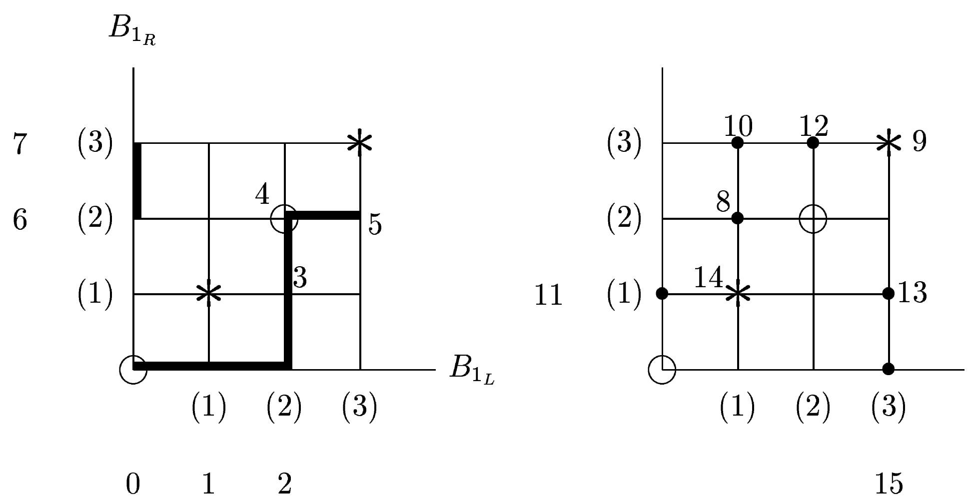

Figure 3.1 *B_2 as the lattice $B_{1_L} \times B_{1_R}$*

The sequences ordered 0 to 5 are displayed contiguously, there is a discontinuity from 5 to 6, followed by the unit segment 6 to 7. On the right, are displayed the ordinals 8 to 15 assigned in [7] to the (yet) unordered sequences, according to the following table :

$B_1 \times B_1$	(1, 2)	(3, 3)	(1, 3)	(0, 1)	(2, 3)	(3, 1)	(1, 1)	(3, 0)
B_2	8	9 $*$	10	11	12	13	14 $*$	15

Out of the 16 sequences of B_2, 4 display a distinguished pattern :

1) $0 = (0000)$ and $\mathbf{e}_4 = (1111)$ correspond to 0 and 1 in B_0 with *period 1.* They are marked with $\circ$.

2) $(0101) = \eta_2$ and $(1010) = \eta_2 + \mathbf{e}_4 = \overline{\eta}_2$, in the notation of [8]. They correspond respectively to $(01) = 3$ and $(10) = 1$ in B_1 with *period 2.* They are marked above with $*$ (see Fig. 3.1). The order given in [7] yields $\mathrm{ord}(\eta_2) = 9$ and $\mathrm{ord}(\overline{\eta}_2) = 14$. We recall that the order was chosen to make the multiplication table in B_2 differ (in its 8×8 left corner) from the multiplication table mod 8 on $\{0, 1, \ldots, 7\}$, by only 4 places. This yields an error rate of 2^{-4} [7, 8].

2) For $k = 3$, we extend the triangle of side length 8 into the truncated parallelogram below : it ends with 3 trailing ones rather than 1.

	0	1	1	1	1	1	1	1	1	0	0	0	0	0	$\times \ldots$
	0	0	1	1	1	1	1	1	1	1	0	0	0	0	$\times$
	0	0	0	1	1	1	1	1	1	1	1	0	0	0	$\times$
	0	0	0	0	1	1	1	1	1	1	1	1	0	0	$\times$
B_3	0	0	0	0	0	1	1	1	1	1	1	1	1	0	$\times$
	0	0	0	0	0	0	1	1	1	1	1	1	1	1	$\times$
	0	0	0	0	0	0	0	1	1	1	1	1	1	1	$\times$
	0	0	0	0	0	0	0	0	1	1	1	1	1	1	$\times \ldots$
ord	0	1	2	3	4	5	6	7	8	9	10	11	12	13	——

This allows us to order $14 = 9 + 5$ sequences out of $N_3 = 2^{2^3} = 256$. See Figure 3.2, which displays the thick line 0 to 11, then the unit segment 12 to 13. It also displays, with underlined numbers, the order <u>14</u> to <u>31</u> given in [8]. This order allows the emergence of multiplication mod 16 with an error rate of the order of $\frac{1}{2}$ [8].

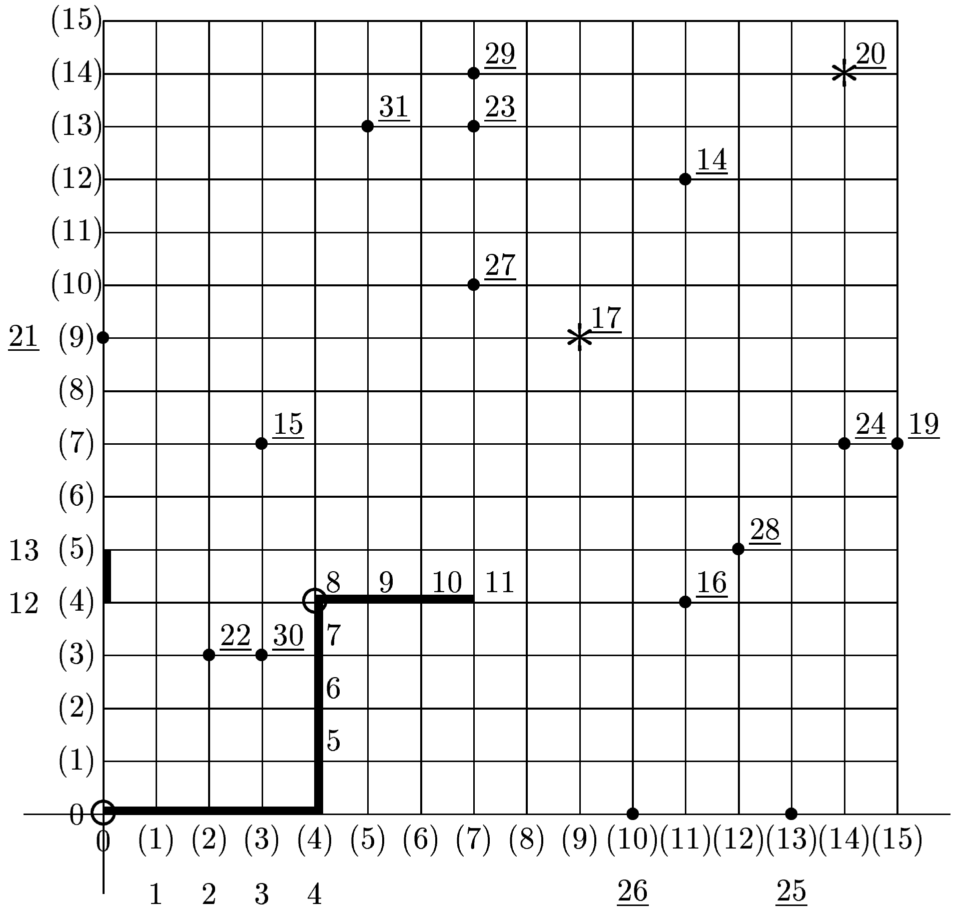

Figure 3.2 *Partial order on* $B_3 = B_2 \times B_2$

The distinguished sequences 0, $\mathbf{e}_8$, η_3, $\overline{\eta}_3$ are respectively labeled by ord(0) = 0, ord($\mathbf{e}_8$) = 8, ord(η_3) = 17 and ord($\overline{\eta}_3$) = 20. Because $\eta_3 = (9,9)$ and $\overline{\eta}_3 = (14,14)$, we see a pattern emerging on which we can base the iterative ordering process for the pair $(\eta_k, \overline{\eta}_k)$.

3) Iterative order for $k \geq 4$ *for* $(\eta_k, \overline{\eta}_k)$. The function ord is given at $(\eta_k, \overline{\eta}_k)$ by Table 3.2.

k	1	2	3	$k \geq 4$
ord(η_k)	(01)=3 ‖	(3,3)=9	(9,9)=17	$2^{k+1}+1$
ord($\overline{\eta}_k$)	(10)=1	(1,1)=14	(14,14)=20 ‖	4ord(η_{k-2})

Table 3.2 *The map ord at* $(\eta_k, \overline{\eta}_k)$

For $k \geq 2$, $\text{ord}(\eta_k) = 2^{k+1} + 1$. The pattern for $\overline{\eta}_k$ becomes regular for $k = 4$. If we set $\text{ord}(\overline{\eta}_k) = \text{ord}(\eta_k) + 3$ for $k \geq 3$, we get $\text{ord}(\overline{\eta}_k) = 4\text{ord}(\eta_{k-2})$ valid for $k \geq 4$.

The broken line in Table 3.2 materializes the border between the general rule and the exceptional cases. The 3 values 1, $14 = 2 \times 7$, $20 = 4 \times 5$ do *not* obey the rule $\text{ord}(\overline{\eta}_k) = 4\text{ord}(\eta_{k-2})$. Their value come from the specific order imposed to B_2 and B_3. Therefore they could be interpreted as *accidental,* true for no other reason than an arbitrary choice. This is a rational conclusion at this local level. However, from the broader perspective of *computation* we get a logical explanation that cannot be found at the local level of *order.* Their values are related to the emergence of multiplication mod 8 and 16 [8].

3.3. Generic and particular features of B_k, $k \geq 0$

1) The previous analysis confirms the dominant role (over addition) that the multiplication operator plays in Scientific Computing, as this was empirically discovered by the astronomer S. Newcomb in 1881. The resulting **Newcomb-Borel paradox** is explained by the Lévy law of large numbers (on the uniform limit distribution of the sum mod 1 of random variables) [14, 8]. However B_0 is an exception to this rule. For $k = 0$, multiplication alone cannot distinguish between 0 and 1 ($0^2 = 0$, $1^2 = 1$, $1 \times 0 = 0 \times 1 = 0$). The distinction requires addition: $1 + 0 = 0 + 1 = 1, 1 + 1 = 0 + 0 = 0$. The arithmetic parity is conveyed by $+$ mod 2. The natural order on $\mathbb{N}$ is immediate by applying repeatedly $+1$ to 0.

However, for $k \geq 1$, the parity of x in B_k is expressed by multiplication alone, depending whether $x^2 = 0$ (even) or 1 (odd) [8].

2) The aspect which makes B_1 an exception is related to the order in C_n. For $n \geq 3$, $C'_n \subset C_n$ is mapped onto $\{0, \ldots, n+1\}$ with $n + 1 < 2^n - 1$. But for $n = 2 = 2^1$, that is for B_1, $3 = 2^2 - 1$. Moreover, $B_1 = B_0 \times B_0$ has the structure of $\mathbb{Z}_4$ (multiplication mod 4). B_1 is exceptional because it is a cyclic group on $\{0, 1, 2, 3\}$, yielding a rotation of $\pi/2$ centered at $w_1 = (\frac{1}{2}, \frac{1}{2})$, Figure 3.3 (top). See [7] for the application to Quantum Computing.

3) We have already seen that looking at $(\eta_k, \overline{\eta}_k)$ makes B_1, B_2, B_3 appear special. We look into this in more detail.

We first look at $\mathbf{e}_{2^k}$ which distinguishes B_0 from B_k, $k \geq 1$. For $k = 0$ $e_1 = 1$ is odd ($1^2 = 1$), but for $k \geq 1$, $\mathbf{e}_{2^k}$ is *even* ($\mathbf{e}_{2^k} = 0$). The sequence $\mathbf{e}_{2^k}$

is on the diagonal of $B_k = B_{k-1} \times B_{k-1}$ for $k \geq 1$, with $\mathrm{ord}(\mathbf{e}_{2^k}) = 2^k$. The diagonal which passes through 0 and $\mathbf{e}_{2^k}$ defines the LR symmetry axis for B_k. We call $\mathbf{e}_{2^k}$ the *axial point* for the algebra B_k, and recall that it enjoys special properties such as, for any $x \in B_k$,

$$x \times \mathbf{e}_{2^k} = 0 \text{ iff } x^2 = 0, \quad x \times \mathbf{e}_{2^k} = \mathbf{e}_{2^k} \text{ iff } x^2 = 1 \quad [8].$$

We turn to the pair $(\eta_k, \overline{\eta}_k)$. For $k = 1$, $\eta_1 = (01) = 3$ and $\overline{\eta}_1 = (10) = 1$ are *not* on the diagonal which connects $0 = (00)$ to $2 = (11) = \mathbf{e}_2$. The line which joins 3 and 1 intersects the diagonal at w_1, Fig. 3.3 top. The pair (1, 3) is connected rightwise (resp. leftwise) to the axial point 2 of B_1 by $3 = 2 + 1$. It is equally connected to $0 = (00)$ the origin of B_1 mod (11),

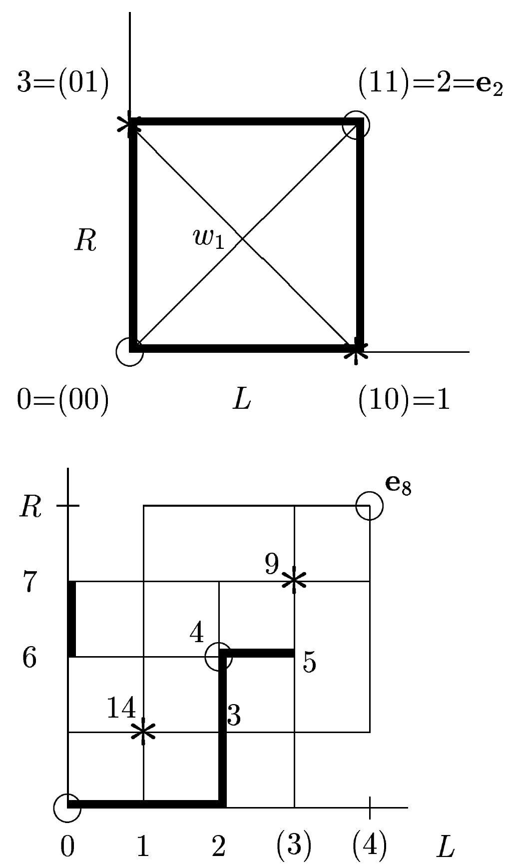

Figure 3.3 B_1 *and* $B_2 = B_{1_L} \times B_{1_R}$

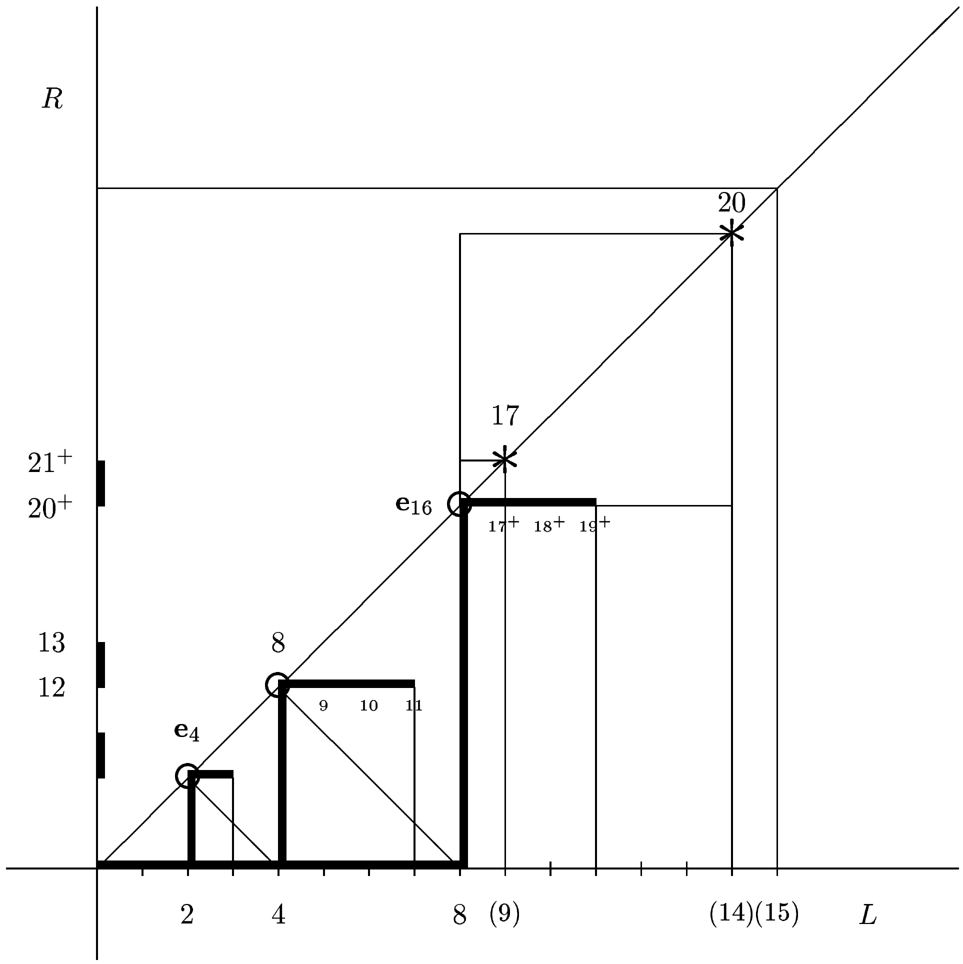

Figure 3.4 *Partial superposition of B_3 with B_2 and B_4*

that is $(01) \equiv (10) \bmod (11)$, because B_1 is a cyclic group. Note that 0 is *not* the axial point, but the origin of B_1. The connection can be leftwise (L) or rightwise (R) according to the table

connection	0	$\mathbf{e}_2 = 2$
$\eta_1 = (01) = 3$	R	L
$\overline{\eta}_1 = (10) = 1$	L	R

The distance is 1 in all four cases.

For $k \geq 2$, the pair $(\eta_k, \overline{\eta}_k)$ is on the diagonal. We look at its diagonal connections to the 3 axial points $\mathbf{e}_{2^{k-1}}$, $\mathbf{e}_{2^k}$ and $\mathbf{e}_{2^{k+1}}$ for B_{k-1}, B_k and B_{k+1}

respectively. Let γ_l (resp. δ_l) represent the left or right distance from $\mathbf{e}_{2^l}$ to η_k (resp. $\overline{\eta}_k$), for $l = k-1,\ k$ and $k+1$. Particular values and general formulae are given for η_k in Table 3.3 and for $\overline{\eta}_k$ in Table 3.4, separated by a vertical double line.

k	2	$k \geq 3$
γ_{k-1}	2	$1 + 3 \times 2^{k-2}$
γ_k	1	$1 + 2^{k-1}$
γ_{k+1}	-1	1

Table 3.3 γ_l *for* $l = k-1,\ k,\ k+1$.

k	2	3	$k \geq 4$
δ_{k-1}	0	12	$\delta_l = \gamma_l + 3$
δ_k	-1	10	
δ_{k+1}	-3	6	

Table 3.4 δ_l *for* $l = k-1,\ k,\ k+1$.

On Figure 3.4 we have displayed the 16×16 lattice B_3 with B_2 and a part of B_4 superimposed, in order to represent schematically the dynamics of the ordering process. Three continuous thick lines are shown (each followed by a vertical unit segment). They correspond to the order given by the truncated parallelogram rule for $k = 2, 3, 4$. The axial points $\mathbf{e}_4$, $\mathbf{e}_8 = 8$ and $\mathbf{e}_{16}$ are marked with $\circ$. The distinguished pair $(\eta_3, \overline{\eta}_3)$ is marked with $*$. The superscript $+$ on numbers assigned to points in B_3, such as $17^+, 18^+, \ldots, 21^+$, indicate that the ordinal will be assigned by B_4, that is for $k = 4$.

This review has shown why the four first algebras B_0 to B_3 are special, each of them in a specific way. This analysis concerns one level of exponential complexity $n = 2^k$. We shall be led to consider more levels as we progress in the discovery of interesting features for C_n.

3.4. The functions w, λ, c at $\mathbf{e}_{2^k}$, η_k, $\overline{\eta}_k$

The values are obtained by repeated use of the identity

$$(1 + a + \cdots + a^{n-1})(a-1) = a^n - 1.$$

1) Values at $\mathbf{e}_n$, $n \geq 1$.

The three sieves w, λ, c, take their maximal values at $\mathbf{e}_n$, respectively

$$M_n, \quad \frac{M_n}{\operatorname{card}(C_n)} = 1 - \frac{1}{\operatorname{card}(C_n)}, \quad \text{and } n.$$

There is L-R symmetry. When $n = 2^k$, $k \geq 1$, $\mathbf{e}_{2^k}$ is an axial point.

2) Values at the pair $P_k = (\eta_k, \overline{\eta}_k)$, $k \geq 1$, $n = 2^k$.
The following notations are helpful:

$$D_k = 2^k, \quad N_k = 2^{2^k}, \quad W_k = M_{2^k} = 2^{2^k} - 1 = N_k - 1,$$

$$M_k = 2^k - 1 = D_k - 1.$$

The U-D values of c at P_k are the pair (0, 1).

The L-R values of w (resp. λ) at P_k are the pair $(\frac{1}{3}W_k, \frac{2}{3}W_k)$ (resp. $(1 - \frac{1}{3N_k}, 1 - \frac{2}{3N_k})$). They are related to the values at the axial point $\mathbf{e}_{2^k}$ by the numerical identity $\frac{1}{3} + \frac{2}{3} = 1$.

3.5. The case $k = 2^p - 1 = M_p$, $p \geq 0$

When $k \geq 2$ and $k + 1 = 2^p$, $\operatorname{ord}(\eta_k) = 2^{k+1} + 1 = 2^{2^p} + 1 = F_p$, the Fermat number with exponent $p \geq 2$. $F_0 = 3$ is equal to $\operatorname{ord}(\eta_1)$, but $F_1 = 5$ does not appear ($\operatorname{ord}(\eta_2) = 9$).

The five first consecutive Fermat numbers $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$ are *prime*. Primality is a powerful sieve on Numbers, known at least since Eratosthenes, a famous librarian at the Alexandrina during the 3rd Century BC. We explore one of the many facets of this sieve by means of the Sierpinski triangle in Section 4.

Before that, we remark that $k = 2^p - 1 = M_p$ represents the number of zeros in F_p, the number of places which are vacant in the binary representation of F_p. We call M_p the **capacity** of F_p.

4. The Sierpinski triangle

4.1. Definition

The *arithmetical triangle* on $\mathbb{N}$ gives the coefficients of the binomial expansion. Often called the Pascal Triangle, it was well-known to algebraists of the

Islamic World, in particular to the famous Al Karaji born in Persia (Karaj) who worked at the celebrated House of Wisdom in Bagdad, at the turn of the 10th Century [13].

The binary version on $\mathbb{Z}_2$ (where C_k^m is reduced mod 2) is known as the *Sierpinski triangle.* It is a much more recent object with a recursive structure which makes it a self-similar fractal [2, 12]. This structure has a rich spectrum of possible interpretations. For example, in Theoretical Computer Science, it plays an essential role in the construction of the exponential Diophantine equation associated by Chaitin [2] to the kth bit of the real number Ω. In the context of a theory of Computation based on the model of a universal Turing Machine, Ω is the halting probability and is algorithmically random.

For a completely different perspective related to Mathematical Analysis, the reader is referred to the insightful paper of Robert [12] in **Le Labyrinthe du Continu.** The interpretation of the triangle as a *discrete parity wave* leads to three possible partial differential equations (of parabolic, hyperbolic or elliptic type).

4.2. Arithmetic connection with Mersenne and Fermat

However, the facet of the Sierpinski triangle most relevant to our present work lies in the little known but remarkable following property [15, p. 140]:

> *The* $2^5 = 32$ *first rows of the Sierpinski triangle give the binary expansion of the* 32 *divisors of the product* $G_4 = F_0 \times F_1 \times F_2 \times F_3 \times F_4$.

The property is easy to check directly. It uses the fact that for $p \geq 1$, $M_{2^{p+1}} = G_p = \prod_{i=0}^{p} F_i$. For $p \geq 1$, each cycle of length 2^p in the triangle starts at the Fermat number F_p on the row $2^p + 1$ and ends at the Mersenne number $M_{2^{p+1}} = F_p - 2$ on row 2^{p+1}.

Because F_5 is composite, the rows 33 to 64 will not give all the divisors of $G_5 = \prod_{i=0}^{5} F_i$: F_5, although composite, is treated as prime. The capacity M_p of F_p takes prime values (besides 1) for $p = 2, 3, 5$. The corresponding *perfect* numbers are $\Delta_2 = 6$, $\Delta_3 = 28$, $\Delta_5 = 496$, with $\Delta_p = \frac{1}{2} M_p(M_p + 1)$.

Depending whether the Sierpinski triangle is used for *one* goal A (producing the list of divisors of $M_{2^{p+1}}$) or for *two* goals A and B (B = the construction of an even perfect number from M_p), the process should be stopped either at row 32, or else at row 63. Taking into account that $n = 2^k$ and $k + 1 = 2^p$, letting $p = 5$ or 6 means two levels of exponential for the length n and three for card(C_n). For $p = 5$, we get $n = 2^k = 2^{(2^p-1)} = \frac{1}{2}N_5$ of the order of 2×10^9.

The classical way to come to terms with combinatorial complexity is an appeal to the continuous.

4.3. Computational meaning assigned to the Sierpinski triangle

We return to the two goals A and B which have been assigned to the Sierpinksi triangle, as B may seem, so far, more related to Ancient Greece arithmetical folklore than to Computation.

1) Goal A is reached by a finite process.

It is known that a regular polygon with n sides can be constructed with ruler and compass iff n is a finite product of powers of 2 and of prime Fermat numbers. We write $n = 2^l s$, s odd. Therefore the 32 first rows of the Sierpinski triangle give all the possible values for s, so that the n vertices of a regular polygon are computable within a quadratic algebra.

This property is to be related to the emergence of patterns in C_n by looking at the solutions of $a \times x = b$ in B_k, $n = 2^k$, for a, b even. The structure of the solutions $S(a, b)$ in $B_{k+1} = B_k \times B_k$ is related to that of the sets of zero divisors for a and b. Specific choices for a and b favor collective invariance in $S(a, b)$, where the solutions can be seen as vertices of a regular polygon [8].

2) Is goal B within finite reach?

It is not known whether the sequence of prime Mersenne numbers is finite or not. The largest known p (as of 1996) is $p = 1{,}257{,}787$ leading to M_p with 378,632 digits. It appears that the goal B could be pursued for very large values of p. Therefore, the question of its role for computation needs to be addressed.

$\Delta_p = \frac{1}{2}M_p(M_p + 1)$ represents the number of zeros in the large central zero triangle of side M_p, generated by F_p at row $2^p + 1$, which ends with 1 zero on row $2^{p+1} - 1$, $p \geq 1$. For $p = 1$, the odd $\Delta_1 = 1$ is obviously exceptional. For $p \geq 2$, the even Δ_p can be interpreted as a discrete value for the area of the triangle. Δ_p is such that $\Delta_p \leq \sum_p$, the sum of its true divisors, counting 1 (the aliquot parts) by Proposition 4.1, with equality iff M_p is prime (Euler).

We set $\rho_p = \frac{\Sigma_p}{\Delta_p}$.

Proposition 4.1. *For $p \in \mathbb{N}$, $p \geq 2$, the even number $\Delta_p = M_p 2^{p-1}$ is such that $\rho_p \geq 1$.*

Proof. Let n be an arbitrary integer $n \in \mathbb{N}$, $\sigma(n)$ represents the sum of its divisors, including 1 and n.

For $n \geq 2$, $\sigma(n)$ is minimum when n is prime : $\sigma(n) = 1 + n$. When n is perfect, $\sigma(n) = 2n$. Otherwise, n is deficient ($\sigma(n) < 2n$) or abundant ($\sigma(n) > 2n$). The ratio $\frac{\sigma(n)}{n} = 1 + \rho(n)$ is strictly monotonically increasing by multiplication : for any $m > 1$, $m \in \mathbb{N}$, $\rho(mn) > \rho(n)$.

By the formula for $\sigma(n)$, [10], p. 239, it is clear that, for $\Delta_p = \frac{1}{2} M_p 2^p = M_p 2^{p-1}$ such that M_p is prime, $\sigma(\Delta_p) = (M_p + 1) M_p = 2\Delta_p$.

What happens when M_p is composite? In order to prove that $\rho_p > 1$, we only have to examine the case $M_p = a \times b$ with a and b prime, not of the Mersenne form, by the monotonicity property for $\rho(n)$. We get $\sigma(\Delta_p) = (a+1)(b+1)(2^p - 1)$, then

$$\frac{\sigma(\Delta_p)}{\Delta_p} = \frac{(a+1)(b+1)}{M_p} \frac{M_p}{2^{p-1}} = \frac{2^p + a + b}{2^{p-1}} > 2.$$

Therefore $\rho_p = 1 + \frac{a+b}{2^{p-1}} > 1$.

△

The role of B is therefore to signal that ρ_p achieves its minimum value 1.

The completion of the two goals A and B simultaneously is achieved by the 63 first rows of the Sierpinski triangle. This yields the table

p	0	1	2	3	4	5
F_p	3	5	17	257	65537	–
M_p	0	1	3	7	$15 = 3 \times 5$	31
ρ_p	–	1	1	1	2	1

where ρ_p only takes the values 1 and 2. In general, ρ_p is rational.

The following relation connects, for $p \geq 0$, F_p to M_p, its capacity which is the side length of the central zero triangle generated by F_p :

$$2^p = M_p + 1 = \log_2(F_p - 1).$$

A more conservative strategy would consist in considering only the first 16 rows (stopping *before* F_4) which gives the 16 divisors of $G_3 = \prod_{i=0}^{3} F_i$ and $\Delta_3 = 28$ which is perfect. This corresponds to examining sequences of length $n = 2^{16}$.

5. Models for the Cantor space $\{0,1\}^{\mathbb{N}}$

We consider the set of infinite sequences which, equipped with the product topology, becomes the Cantor space $C_\infty = \{0,1\}^{\mathbb{N}}$, which is not countable. Depending on the aspect of C_∞ which is to be studied, several models for C_∞ will prove useful.

5.1. The field $\mathbb{Q}_{(2)}$ of dyadic (2-adic) numbers [12]

In this paragraph, we follow [12] closely. In order to avoid ambiguity with $\mathbb{Z}_2 = \{0,1\}$, we write $\mathbb{Z}_{(2)}$ for the ring of 2-adic numbers. Such a ring derives from $+$ and $\times$ with carry on infinite sequences. The sequence $(1111\dots)$ is identified with -1, because $1 = (100\dots)$ and $(111\dots)$ add to 0. Then $\mathbb{Q}_{(2)} = \mathbb{Z}_{(2)}[1/2]$ with valuation $|\,.\,|$ such that $|2| = 1/2$. This corresponds to the limit, as $n \to \infty$, of the map Lw (representation of $b \in C_n$ on non-negative powers of 2).

5.2. The unit interval $[0,1] \subset \mathbb{R}$

We turn to the limit of the map $L\lambda$, i.e. the representation of $b \in C_n$ on negative powers of 2. We identify the two sequences $(0111\dots) = (1000\dots)$ in C_∞. This yields the familiar binary representation of a real number in $[0,1]$.

The two models above for C_∞ play an essential role in Mathematics, the first one in Number Theory, and the second one in Real Analysis. For our purpose we need to go further since the field $\mathbb{R}[X]$ is not algebraically closed.

5.3. The unit disk in $\mathbb{C}$

So far we have considered *representation* sieves such as Lw and $L\lambda$. To connect with $\mathbb{C}$, we have to consider the more powerful sieve of *spectral* analysis [4].

For that purpose, we consider $b \in C_\infty$ as defining a linear operator on the Banach space $X = l^p(0,\infty)$, $1 \leq p \leq \infty$, for example ($\dim X = \infty$), by

$$x = (x_i)_1^\infty \to T_b\, x = (b_{i-1} x_{i-1}),\ b_i \in \mathbb{Z}_2,$$

with first component 0. The particular choice $\mathbf{e} = (111\ldots)$ gives the right shift operator

$$T_{\mathbf{e}} = T : (x_1, x_2, \ldots) \stackrel{T}{\to} (0, x_1, x_2, \ldots)$$

[4, Example 2.26, p. 98-99]. The left shift operator T' is defined by

$$(x_1, x_2, \ldots) \stackrel{T'}{\to} (x_2, x_3, \ldots).$$

Both T and its adjoint T' have the unit disk $D = \{z \in \mathbb{C}; |z| \leq 1\}$ for their spectrum σ. However the partition $\sigma = P_\sigma \cup C_\sigma \cup R_\sigma$ is different, where P_σ represents the set of algebraic singularities, whereas C_σ and R_σ reflect topological properties [4].

For $1 \leq p < \infty$, we get for T : $P_\sigma(T) = \overset{o}{D}$, the *open* disk, $C_\sigma(T) = C = \{z; |z| = 1\}$ the unit circle and $R_\sigma(T) = \emptyset$.

For T', on the other hand, $P_\sigma = \emptyset$, $C_\sigma = C$ and $R_\sigma = \overset{o}{D}$. It is interesting to point out that T' has no *algebraic* singularities ($P_\sigma = \emptyset$).

This shows how topology is entering the game when the structure is rich enough : here $\dim X = \infty$. For matrices ($\dim X < \infty$) the spectrum is a finite collection of isolated points in $\mathbb{C}$. And the map $T \to \sigma(T)$ is *continuous*. But when $\dim X = \infty$ this is not true anymore in general : $T \to \sigma(T)$ is only *upper semicontinuous* in the Banach space of bounded, or closed linear operators [4, 17].

Therefore, $\sigma(T)$ can shrink under the perturbation induced by finite rank approximation, for example. Here, the finite sequence $\mathbf{e}_{n-1} = (11\ldots1)$ of length $n-1$ defines the right shift represented by the $n \times n$ matrix

$$J_n = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & & \cdot \;\; 1 \\ \mathbf{0} & & & & 0 \end{pmatrix}$$

which is a Jordan block of order n. In exact arithmetic, $\sigma(J_n) = \{0\}$ which is very different from D! We mention that finite precision computation smoothes the discrepancy between 0 and D. The eigenvalues computed (with finite precision) by a reliable algorithm (such as QR) tend to cluster on the unit circle for $n > 500$. This comes from the fractal dimension $1 - \frac{1}{n}$ of the eigenvalue 0 of J_n, which tends to 1^- as $n \to \infty$ [16].

The value 0 is a singular point for the map $J_n \to J_n^{-1}$ which has zero *Baire* measure (it is not generic) in Exact Computing, but has positive *Lebesgue* measure in Inexact Computing [14].

5.4. Spectral analysis in commutative Banach algebras

There is an important case where the map $T \to \sigma(T)$ is continuous for dim $X = \infty$. This is when T belongs to a *commutative* Banach algebra [17].

An important example of this property is provided by *doubly* infinite sequences converted into Fourier series. We develop this example in the next paragraph. First we look at the relation between doubly infinite sequences in $l^p(-\infty, +\infty)$ and binary sequences in C_n.

To the sequence $\mathbf{e}_n$, with $n = 2m$, we can associate its expansion on negative powers 2^{-k}, $k = 1$ to m, **and** on positive powers 2^k, $k = 0$ to $m - 1$, at the same time. This yields a real positive number between 0 and M_m. This amounts to consider the two sieves w and λ simultaneously, rather than sequentially.

We now consider $f_n = \mathbf{e}_n - 1$, the sequence of $n - 1$ ones, and 0 at place $m + 1$. The $n \times n$ matrix K_n for the associated right shift has the structure

$$K_n = \left(\begin{array}{c|c} J_m & 0 \\ \hline & J_m \end{array} \right)$$

and the unique eigenvalue 0 of multiplicity $n = 2m$. In $X = l^p(-\infty, \infty)$, the corresponding right shift operator T_f is such that $\sigma(T_f) = D$ [17].

Consider the rank one perturbation U on T_f, which amounts to put 1 instead of 0 into f, that is $T_f + U = T_e$, where e is the doubly infinite sequence of ones, $e = (\ldots 111 \ldots)$. We get $\sigma(T_e) = C$ the unit circle [17, Example 3.8, p. 210]. In the truncated matrix version, this perturbation amounts to change K_n by putting a 1 in the $m \times (m + 1)$ position in K_n. The two copies of the Jordan block J_m are *weakly coupled* by this 1. This sheds light on a subtle numerical phenomenon reported in [16]. When one computes (with QR) the eigenvalues of a matrix with Jordan structure J_n, then for n large enough but not too large, say $100 \leq n \leq 2000$, all the computed eigenvalues are clustered on a circle whose radius rapidly tends to 1^-. The computed eigenvalues begin to fill the open disk $\overset{o}{D}$ for extremely large values of n only.

5.5. Fourier analysis

As already indicated, the extremely powerful sieve of spectral (and Fourier) analysis can be interpreted as using the two more elementary finite sieves w and λ *in conjunction.* Note that because the sequences are doubly infinite, we can use L or R versions for w and λ as we please.

The presentation of Fourier analysis to follow is taken from [17, Example 3.22, p. 217–218]. We consider the space l of doubly infinite sequences b such that $\sum_{k\in\mathbb{Z}} |b_k| < \infty$, $b_k \in \mathbb{R}$, in which the product is defined by *convolution* :

$$d = bc \Longleftrightarrow d_k = \sum_j b_j c_{k-j},\ k \in \mathbb{Z}.$$

l is a commutative Banach algebra with unit 1 equal to the sequence

$$(\ldots 00100 \ldots),$$

where the only 1 is at place 0.

To $b \in l$, we associate the complex valued function $b(e^{i\theta})$ defined on $C = \{z; |z| = 1\}$, the unit circle, by $e^{i\theta} \to b(e^{i\theta}) = \sum_{k\in\mathbb{Z}} b_k e^{ik\theta}$, a Fourier series which is absolutely convergent.

The spectrum $\sigma(b)$ of b is the range $\operatorname{Im} b(e^{i\theta})$. Therefore b is invertible in l iff $0 \notin \sigma(b)$, that is $b(e^{i\theta}) \neq 0$. This yields the Wiener theorem : if the complex function $b(e^{i\theta})$, having an absolutely convergent Fourier series, does not vanish anywhere, the function $1/b(e^{i\theta})$ has also an absolutely convergent Fourier series.

Fourier analysis appears therefore, as the **synthesis**, by spectral analysis, between the discrete sieves w and λ, synthesis which requires to take the limit as $n \to \infty$. Real analysis corresponds to the limit of λ alone (negative exponents of 2) and complex analysis emerges from the spectral analysis of either w, *or* λ. The Fourier analysis realises the **and** function.

6. Conclusion and perspectives

To explore the diversity of the world of binary sequences *without a priori simplification* requires at least two levels of exponential, which means sequences of length of the order of 10^9. This represents an ocean of information well beyond the capacity of the human mind to be processed consciously. This report has shown how glimpses of this combinatorial complexity can be obtained by using such tools as order, weight, measure and connectedness. However, the most powerful tool to transform theoretically the problem into an understandable form is the use of various models for the Cantor space $C_\infty = \{0,1\}^{\mathbb{N}}$. We have described three such models, $\mathbb{Q}_{(2)}$, $[0,1]$ in $\mathbb{R}$ and D in $\mathbb{C}$, which capture each a different aspect of C_∞. Representation by measure leads to real analysis. Linear spectral theory leads to complex analysis. The synthesis of weight and measure leads to Fourier analysis. In each of these domains, the combinatorial complexity, due to the discrete approach, has vanished.

To progress in the understanding of the Cantor space, one possibility is to look for patterns that emerge from the resolution of $a \times x = b$ for a and b even. The preliminary studies done in [7, 8] look promising.

Another important direction is modular arithmetic : mod 2 is obvious on B_0, mod 4 results from the cyclic structure of B_1, mod 8 and 16 emerge from the order imposed on B_2 and B_3. Modular arithmetic is related to the Newcomb law for leading digits so important for Scientific Computing. It is the mechanism by which the leading digits in floating point computations can carry the relevant information *in the presence of round-off* [14].

Despite its discovery more than a century ago, the law for the leading digit has not been fully appreciated before the last decade of the 20th century. This explains why, in precomputer days, such great minds as von Neumann and Hotelling misjudged that computer simulations had no future for Science. They feared that results would be fatally contaminated by round-off. They underestimated the power of the floating point representation for computer arithmetic which was designed by G. R. Stibitz at Bell Labs between 1938 and 1941.

It is not clear, some sixty years later, that this unjustified fear has been completely eradicated amongst scientists...

References

[1] G. J. Chaitin, *Exploring Randomness,* Springer-Verlag, 2001.

[2] G. J. Chaitin, *Algorithmic Information Theory,* Cambridge University Press, 1987.

[3] L. E. Dickson, "On quaternions and their generalization and the history of the 8 square theorem," *Annals of Math.*, **20**, 155–171, 1919.

[4] F. Chatelin, *Spectral Approximation of Linear Operators,* Academic Press, 1983.

[5] F. Chaitin-Chatelin, T. Meškauskas, A. N. Zaoui, "About real and binary algebras: the interplay between geometry and algebra," Cerfacs Rep. TR/PA/00/74, 2000.

[6] F. Chaitin-Chatelin, T. Meškauskas, "Computation with hypercomplex numbers," *Nonlinear Analysis,* **47**, 3391–3400, 2001.

[7] F. Chaitin-Chatelin, T. Meškauskas, A. N. Zaoui, "Hypercomplex division in the presence of zero divisors on $\mathbb{R}$ and $\mathbb{Z}_2$," Cerfacs Rep. TR/PA/02/29, 2002.

[8] F. Chaitin-Chatelin, A. N. Zaoui, "Hypercomputation on {0,1}," Cerfacs Rep. TR/PA/02/73, 2002.

[9] G. Bachet de Méziriac, *Problèmes Plaisants et Delectables qui se font par les Nombres,* A. Blanchard, Paris, 1993 (reédition).

[10] G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers,* Oxford Clarendon Press, fifth ed., 1979.

[11] A. A. Albert, "Quadratic forms permitting composition," *Annals of Math.*, **43**, 161–177 (1942).

[12] A. Robert, "Exemples de Discret-Continu en Analyse," in *Le Labyrinthe du Continu* (J. M. Salanskis, H. Sinaceur eds.), Springer-Verlag France, Paris, 1992, pp. 279–295.

[13] A. Dahan-Dalmedico, J. Pfeiffer, *Une Histoire des Mathématiques,* Seuil, Paris, 1986.

[14] F. Chaitin-Chatelin, E. Traviesas, "Qualitative Computing," to appear as a chapter of *Handbook for Computation,* B. Einarsson ed., SIAM, Philadelphia (also available as Cerfacs Rep. TR/PA/02/58, 2002).

[15] J. H. Conway, R. K. Guy, *The Book of Numbers,* Springer-Verlag NY, 1996.

[16] F. Chaitin-Chatelin, "Is non normality a serious computational difficulty in practice?," in *Quality of Numerical Software,* R. Boisvert ed., Chapmann and Hall, 1997, pp. 300–314.

[17] T. Kato, *Perturbation Theory for Linear Operators,* Second revised edition, Springer-Verlag, Berlin, 1984.

The Cerfacs reports are available from http://aton.cerfacs.fr/algor/reports/