Matrix Combinatorics and Algebra

Milan Kunz

June 18, 2006

Preface

Studying mathematics, you find that it has many branches and specialties: (algebra, geometry, topology, differential and integral calculus, combinatorics), and different theories: (number theory, group theory, set theory, graph theory, information theory, coding theory, special theories of equations, operators, etc.. It seems that there is none unifying concept.

But we know only one world, we live in, only one physics, one chemistry, one biology. There should be only one mathematics, too.

The title of this book is "Matrix Combinatorics and Algebra".

Combinatorics is an old branch of mathematics, and its essentials considered to be elementary, since they are based on good examples. But it has its limitation: There exist too many identities and still more remain to be discovered. The mind is boggled by them as Riordan [1] pointed out because there appears only disorder in abundance. Classical combinatorics contained many branches which separated later but which are essential for understanding it. You will find here themes from number and group theories as well.

Algebra is very abstract science, except for its one branch, linear algebra. There are studied operations with vectors and matrices. And on these notions the heart of the book is based.

I think that I found a path into the strange world of combinatorics. It started long ago when I accidentally discovered that two famous entropy functions H= $-\sum p_j \log p_j$, as defined by Boltzmann [2] and Shannon [3], are two distinct functions derived from two polynomial coefficients, contrary to generally accepted views and the negentropy principle.

I had a feeling as Cadet Biegler [4]. Remember his desperate ejaculation: "Jesusmarja, Herr Major, es stimmt nicht!". Senior officers quietly listened the lecture about coding, but the given example did not make sense, because they had at hand another volume than prescribed by instructions. The name of the book was "Sünde der Väter". Similarly as Švejk, I think, that a book should be read from its first volume.

It was almost impossible to publish my results, because they did not

conform with accepted views. My first attempt was rejected from reason that my explanation was ununderstable. Being angry, I wrote its essence for a technical journal for youth, where it was accepted as suitable lecture for their naive readers [5]. Till now, I was not able to publish my result explicitly. The referees did not accept my arguments. From many reasons. I was forced to continue my research, discover new relations which proved my conception. There are very elementary things about graph matrices which are not explained in textbooks on their suitable place, it means at the beginning. Thus I hope that I succeeded.

If this book were written one hundred years ago, it could save one human life, if it were published fifty years ago, it could prevent erroneous interpretation of the information theory.

Mathematical equations and identities are like pieces of a puzzle. They are arranged in a traditional way into specialties which are studied separately. If mathematicians were unable to realize that both entropy functions stem from an identity known paradoxically as the Polya-Brillouin statistics which could be found in commonly used textbooks [6], then some basic obstacles must have prevent them to interpret their abstract definitions correctly.

When I studied the entropy problem I knew that it was a combinatorial problem because Boltzmann himself connected the function H with a combinatorial identity. Moreover, I correlated it intuitively with matrices because: "I did not even know what a matrix is and how it is multiplied," as Heisenberg [7] before me. Usual explanations of matrices did not make any sense to me.

My approach is elementary: A string of symbols (a word, a text)

- "explanations of matrices did not make any sense"

  is considered as a string of consecutive vectors written in bold face letters as vectors

- "**explanations of matrices did not make any sense**"

  and vectors in the string are written using the formalism

- $\mathbf{j} = \mathbf{e}_j = (0, 0, ...1_j, ...0)$

as a vector column in the matrix form. I named matrices having in each row just one unit symbol "naive". The matrices, obtained by permutations and by finding scalar products of naive matrices with unit vectors, are counted and results tabulated. The resulting tables of combinatorial functions have the form of matrices and matrix operations as multiplication, transposition and inversion can be performed on them. Applications

of matrix operations were common in combinatorics as the Kronecker function δ_{ij}, which is an implicit application of inverse matrices. Riordan has given many examples of their use. However, matrix technique was not used systematically and combinatorial identities were not connected with intrinsic properties of vector spaces.

Sums and differences of two naive matrices are studied in the second part of the book. They are known as graphs. The block designs could form the next step. They are exploited in advanced combinatorics. Blocks have matrix form and the numbers of distinguishable blocks are searched for. Hall [8] began his book by chapters surveying classical combinatorics before he treated block designs, but no attempt was made to use an unified matrix technique to traditional combinatorics and explain combinatorial problems as counting naive blocks.

When I connected combinatorial problems with properties of countable vector spaces, I discovered another way into the Euclidean space. I can clear up, at least I hope so, how this space is built. Some of its basic properties are not explained in textbooks. Either mathematicians do not consider them important, or they simply ignore them. Of course, a possibility exists that they keep them as hermetic secrets unexplained to uninitiated. In any case, the Euclidean space has very strange properties.

This book is an elementary one. Only exceptionally results of higher mathematics are introduced, and then without proofs. Nevertheless, I do not think that it is an easy book. It shows how complicated the world is, that everything is connected with everything. I try to explain some parts of combinatorics and matrix algebra in an unconventional way. The purpose is not mathematical rigor or practical applications but the achievement of intuitive understanding of vector space complexity. I prefer full induction to generating functions and the main purpose of the book is to show that the world has not only three dimensions, we can move in. I must admit that I myself have difficulties trying to visualize some elementary things. Some solutions I found only after very long periods of thinking, as if the right way were blocked by invisible obstacles.

Before we start let us make a note about the number systems. Everybody knows the decimal one:

$$0 = 10^{-\infty};\ 1 = 10^0;\ 10 = 10^1;\ 100 = 10^2\ .$$

Somebody knows the binary one:

$$0 = 2^{-\infty};\ 1 = 1^1 = 2^0;\ 10 = 2^1;\ 11 = 3;\ 100 = 4 = 2^2\ .$$

But nobody, as it seems to me, studied the unitary number system:

$$1 = 1^1;\ 11 = 2 = 1^2;\ 111 = 3 = 1^3.$$

The difference is that the system starts from the first power of 1, which is undistinguishable from its zero power

$$1 = 1^{-1} = 1^0\ .$$

Logarithm of 1 with the base of logarithm 1 is again 1, logarithm of 111 with the base of logarithms 1 is 3.

Mathematical operations in this system are simple:

addition

$$111 + 11 = 11\ 111$$

subtraction

$$111 - 11 = 1$$

Multiplication and division should be written as powers, e. g. multiplication $(111)^{11}$, but it can be arranged as blocks

$$\begin{aligned} 11 \times 111 &= 111 \\ 111 &= 111\,111 \end{aligned}$$

and division

$$\begin{array}{rl} 111\ \ 111 \div 11 = 11\ 11\ 11 & \Big| 1 \\ 11\ 11 & \Big| 1 \\ 11 & \Big| 1 = 111 \end{array}$$

We will use this system implicitly without mentioning it. There will be some problems with notation. Not enough letters are available to use a special one for each function. We will use some letters for different functions without warning. Figures, tables, and equations are indexed separately in each chapter.

One difficulty of a systematic exposition is that you can not understand everything completely at once. It is necessary to introduce concepts consecutively. New knowledge modifies previous definitions. Therefore, some topics will be treated repeatedly, when it becomes possible to exploit newly introduced techniques. Be patient, please, when something seems to be too unimportantly detailed. If you really want to understand, reread the book many times.

I have mentioned books of Riordan [9] which were important for combinatorics. Similarly should be mentioned Harary for graphs [10, 11] and

the book of Cvetkovič, Doob and Sachs [12] for eigenvalues of adjacency matrices. Some parts of the present book are compiled from journal articles. I wish to express my acknowledgement especially to members of the Zagreb group for numerous reprints.

This can be considered as second edition of this book. I corrected some formulations and an error concerning of gamma function of negative numbers and added a new generating function of natural numbers.

Contents

List of Figures

List of Tables

Chapter 1

Euclidean, Hilbert, and Phase Spaces

1.1 Preliminary Notes

It is generally believed that we are living in a three dimensional space with 3 possible directions and their opposites: forward and backward, left and right, up and down. Sometimes the time is added as the fourth dimension with specific properties. The time is indispensable for movement. We can not move in time physically, since it is a stream which drift away everything, but our mind can move in time without any difficulties.

Our notion of the space is based on our book form: a point has three coordinates corresponding to the page, line, and column numbers, respectively[1].

Three dimensions of a book are formed by the given convention from a string of symbols. Too long strings are cut into lines, too long strings of lines are cut into pages and eventually too long sequences of pages are cut into volumes forming the fourth dimension since we must determine at first positions of symbols in lines. There exist different forms of books, as for examples scrolls. Strings of symbols can be wound on reels, or rolled up into balls, and they remain essentially unchanged. Similarly the points of the space can be indexed in different ways.

Books exist without any movement but when we read them, we need time to transfer their symbols into our brain, to remember essential facts and thoughts, to transcribe the book into our brain. The world is a word,

[1]There exist polar coordinates giving positions as on reels, too, but they are outside our study.

a very long one, in a foreign language. We must learn, how to understand it.

There exist one essential difference between a book and our world. The world is moving. As if a book itself were constantly transcribed. Some parts seem to us to be constant, but somewhere invisible corrections are made constantly. The world is in instant a the book A, in the next instant b the book B. All possible states of the world form a library.

But we will analyze at first the simpler case, the unmoving text.

The three dimensions of the space are not equivalent. To move forward is easy, backward clumsy, left or right movements, as crabs do, are not normal, up and down we can move only in short jumps (long falls down end dangerously). In books eyes must jump on the next line, a page turned, a new volume opened. Increasing efforts are needed in each step.

Mathematics abstracted these differences. The three dimensions of the space are considered to be equivalent and orthogonal.

Our world seems to be limited by these three dimensions. We are not able to find the fourth geometrical dimension which would be orthogonal to the first three. This is a source of many difficulties and misunderstandings. Mathematicians try to avoid them by concealing decently our inabilities as a shame.

From ancient times the ortogonality means that between two straight lines the right angle R exists. Actually there must be always 4 R if two lines cross

R	R
R	R

The third straight line in the plane must be either parallel to one them, and then it crosses the other one, or it crosses both of them, and then they form a triangle, except lines going through the cross of the first two lines.

The most important property of right triangles is, that the squares of their hypotenuses are equal to the sums of squares of both other sides as on Fig. 2.8

The smallest right triangle which sides are whole numbers has sides 3, 4, 5 and their squares are $9 + 16 = 25$. The relation between the sides of right triangles is known as the *Pythagorean theorem*. The knowledge of right triangles was one from the first mathematical achievements of mankind. The pyramids have square bases. Their triangulation was very accurate due to exploitation of this knowledge.

But similarly as we are not able to find the fourth dimension, we are not

Figure 1.1: Pythagorean theorem. $a^2 + b^2 = c^2$

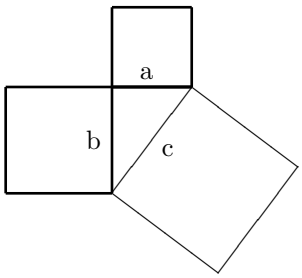
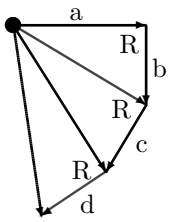


Figure 1.2: Consecutive Pythagorean addition. New vectors are added as orthogonal to the sum of previous ones.



able to decide if a set of numbers does not correspond to a set of orthogonal straight lines which lengths are corresponding to the given numbers. We form consecutively right triangles as on Fig.1.2

Each new line is orthogonal to the right triangle sum of all preceding lines. Try to form a three dimensional model, putting a the third straight line orthogonal to the plane in which the first two lines lie. Then rotate this line in the plane the orthogonal to the hypotenuse, folding it down till it touches the plane. Now there there appears place for the fourth vector again orthogonal to the hypotenuse of the first three vectors. We get the general equation

$$L^2 = \sum m_j^2 \,, \tag{1.1}$$

where m_j^2 stands for n different abscissa and L^2 is the square of the length of all n abscissa. We can rotate consecutively each vector of the plane in such a way that it forms a right triangle with the sum of all other

$(n\ -\ 1)$ vectors, but we must not consider simultaneously more lines or we find that they are not orthogonal as we clearly see on Fig.1.2 where a series of right triangles was drawn.

If we had at disposal more dimensions, we could decompose such a sum into n orthogonal directions.

The sum of n squares with side values (lengths) m_j^2 can be in its turn decomposed into a Pythagorean triangle which squared sides a, b, and c are

$$a^2 = n\overline{m} \tag{1.2}$$

$$b^2 = \sum\ m^2\ -\ n\overline{m} \tag{1.3}$$

and

$$c^2 = \sum m_j^2, \tag{1.4}$$

respectively. $\overline{m}$ in 1.1 is known as the *arithmetical mean.* Actually, the arithmetical mean can be identical with one from n summands. The arithmetical mean is calculated usually by finding the sum of all m values and dividing it by n

$$\overline{m} = \sum m_j/n. \tag{1.5}$$

The straight length of the side is its square root. Here the square root of n appears somewhat surprisingly, but it is the length of the diagonal of n dimensional cube. Similarly the third side of the triangle (1.3) can be normalized by dividing it with n. Then we get the value σ^2 known as the *dispersion*

$$\sigma^2 = 1/n \sum (m_j^2 - n\overline{m}^2). \tag{1.6}$$

Its square root, comparable with the mean, is the *standard deviation* σ. For example, take the values 1, 2, 3, 4. Their mean is 2.5, the sum of squares $30 = 1 + 4 + 9 + 16$. The dispersion is $1/4(30 - 4 \times 6.25) = 1.25$.

Calculating the mean and the standard deviation we need not to know the directions of both legs, since they are determined automatically by their lengths, as when a triangle is constructed from the known lengths of its three sides. We draw two circles with diameters a and b on both ends of the side c. Where the circles cross, the third vertex lies. The direction of all sides in the multidimensional space is abstract for us.

1.2 Euclidean space

The space of right triangles and never crossing parallel lines is known as the *Euclidean space*. Its generalization to infinite many dimensions $n \to \infty$ for the sum 1.1 is known as the *Hilbert space*. An Euclidean space with n dimensions forms a subspace in it. Euclides based his geometry on five postulates:

1. To draw a straight line from any point to another.
2. To produce a finite straight line continuously to a straight line.
3. To describe a circle with any center and distance.
4. That all right angles are equal to each other.
5. That if a straight line falling on two straight lines makes interior angles on the same side less than two right angles, if produced infinitely, meet on that side on which are angles less than two right angles.

The fifth postulate is superfluous. It follows directly from applications of the first four postulates for the following construction. We take a square ABCD. All its right angles are according to the 4. postulate right, and all its sides are straight lines. We add to this square ABCD a new square CDEF and align the sides AE and BF according to the 2. postulate. To the obtained rectangle ABEF, we and a new square EFGH and align again the sides AG and BH according to the 2. postulate. In such a way we continue with adding of squares infinitely, eventually on the other shorter side of the rectangle.

In such a way we produce a pair of parallel straight lines. There are two possibilities that the long sides of the infinite rectangle meet or diverge. Either these long sides are not straight lines meeting the demands of the 2. postulate, or the right angles of consecutive squares did not meet the demands of the 4. postulate. The fifth postulate is a consequence of an application of postulates for an infinite construction.

The problem of ortogonality is loosing its importance in the Hilbert space. If you have a store of infinite many vectors, you can pick any two as the first ones. You can be sure, that you find the third one which is orthogonal to the first two. So you continue. You will be exhausted before you will be able to empty the store. Or you can be lazy, and to use alternatively vectors with angles greater and smaller than orthogonal. The errors will compensate. Euclides introduced axioms into mathematics. Space and its elements are defined by a set of propositions. A disadvantage of this approach is that we don't know a priori which elements form the space. We will use another approach and generate elements consecutively. We encounter spaces of many dimensions by recognizing that we are not alone in the space. There are other people living and many things exist in this particular space. Each entity has its own position in the space. Listing these

positions, we need for each object its specific line with its coordinates. In each line must be as many coordinates as the space, entities are embedded in, has dimensions. If there were m entities in n dimensional space it would be necessary to know mn coordinates, in the 3 dimensional space we need for m entities $3m$ coordinates and in 1 dimensional space still m coordinates to determine positions of all objects. Spaces with m objects are known in physics as phase spaces[2]. They have curious properties and we can sense directly some of them, as for example temperature and the wind velocity of a system of molecules of air are, which correspond to mathematical notions. Each molecule has at ambient temperature mean velocity several hundred meters per second. Impacts of tiny molecules on walls of the container produce the pressure. Chaotic collisions of the molecules moving in different directions lead to a stable distribution of particle velocities.

These velocities decompose into two components. One component is formed by a part of movement which all particles in the given volume have common. This component, the mathematical arithmetical mean, usually great only few meters per second, as compared to the above mentioned hundred meters per second, we feel, when we are inside it as its part, as the wind, the physical property of the system of molecules. The other component is the dispersion from the mean vector velocity. It is known as the thermal motion of molecules, the temperature.

We will show that all phase spaces are *isomorphic*. Some of their properties do not depend on the dimensionality n of the space the system of m entities is embedded, but they are given only by the number of the entities. The phase space is thus a reality and not a mathematical construction. Unfortunately our experience is limited by properties of the cave we are living in as Plato wrote. It is extremely difficult to overcome this handicap and see ideal spaces behind shadows they produce. Shadows of the outer world, our eyes project on the wall of our scull cave (retina) are two dimensional. The third dimension we recognize by efforts of eye muscles focusing the images on the retina. This is done automatically. Higher dimensions we recognize by efforts of our brains or their extensions, computers. It will take a long time, before we accommodate us and get accustomed to the notion of the higher dimensions in the same way as our vision is used to the three dimensions. Our visual organs were developing for hundred millions of years. The notion of a possibility of invisible spaces was born about 2500 years ago, study of their properties started about 250 years ago, and computers which made their understanding easier appeared about 50 years ago. It takes time.

[2]The number of objects is there given as n. To avoid confusion with n dimensions, we use the symbol m.

1.3 Unit Vectors $\mathbf{e}_j$

Now it is time to introduce the notion of the linear translation. If a microparticle moves in the Wilson chamber (or a plane in stratosphere), it leaves a trace of ionized particles and molecules condensing on them. Imagine that even an abstract point, when moving, leaves such a trace. We call it a *vector* and draw it as a line with arrow showing the direction of the movement $\longrightarrow$. To shift a point, we must apply this trace connecting both positions of the point, the initial and the final ones, respectively. We discussed the ortogonality and it is obvious that the vectors can be orthogonal. But we defined ortogonality only between straight lines and thus we suppose that the vectors are straight. Of course, motion in space need not to be limited exclusively to the motion along straight lines but we try to keep our space as simple as possible. A method could be to divide bent vectors into tiny straight vectors with slightly different directions. This are methods of differential and integral calculus. We can assume, that spaces with bent vectors are isomorphic to the space with straight vectors. Next we introduce a special place in our n dimensional space from which we will measure all translations. This point we will call the *center of the coordinate system*. Then we define n points on a sphere (circle) with its center in the center of the coordinate system. We accept the radius of the sphere as the unit length. We can imagine that the points on the sphere are the translated center of coordinate system and we will call each vector connecting the center of the coordinate system with the defined n points on a sphere the *unit vector* $\mathbf{e}_j$. The notation of the unit vector $\mathbf{e}_j$ is a row in round brackets with n elements. In physics symbols with arrows are used as $\vec{\mathrm{j}}$. $(n\ -\ 1)$ elements of the vector $\mathbf{e}_j$ are zeroes and there is only one unit element on the j-th place

$$\mathbf{e}_j = (0_1, 0_2, \ldots, 1_j, \ldots, 0_n). \tag{1.7}$$

Equal length of all unit vectors $\mathbf{e}_j$ in (1.3) is not an essential condition of the existence of a vector space. We could define unit vectors $\mathbf{e}_j$ in (1.3) as having different lengths, make all operations as with vectors having equal lengths, and only then modify results according the defined lengths. A cube which sides are not equal is a rectangular parallelepiped. Its volume, for example, can be calculated as

- side $a = 4.4$ cm
- side $b = 3.9$ cm
- side $c = 0.4$ cm

- volume $= 4.4 \times 3.9 \times 0.4 = 6.864$.

The other possibilities are

- side $a = 2.2\times$ vector $\mathbf{a} = 2$ cm
- side $b = 1.3\times$ vector $\mathbf{b} = 3$ cm
- side $c = 0.4\times$ vector $\mathbf{c} = 1$ cm
- volume $= 2.2 \times 1.3 \times 0.4 = 1.144$
- volume of the parallelepiped $= 2 \times 3 \times 1 = 6$
- total volume $1.144 \times 6 = 6.864$.

Vectors which begin in other places of the space are compared with these specimen vectors $\mathbf{e}_j$ beginning in the center. They are considered to be *identical* with unit vectors $\mathbf{e}_j$, if they are *collinear*. Of course, vectors can be shorter or longer, can have opposite directions, but such differences will be remedied by algebraic means later. Sometimes vectors do not coincide with unit vectors, but with their linear combinations. We suppose, that the unit vectors $\mathbf{e}_j$ are orthogonal by definition.

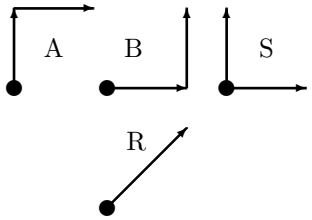
We will subject these unit vectors to different mathematical operations. We find their sums, differences, products and we will try even to divide them. These operations will be done in algebraic form. But before we proceed, we must investigate the results of vector translations on some examples, to interpret the algebraic results correctly.

How can be two vectors added? Suppose that the center $\mathbf{0}$ was at first translated into a point a by the vector $\mathbf{e}_a$ and then to the point with coordinates ab by the translation $\mathbf{e}_b$. There are another possibilities how to reach the same point. We can at first make translation $\mathbf{e}_b$ and then $\mathbf{e}_a$. In textbooks of algebra you can find that the summation is a *commutative* operation. This word means that the result of the operation does not depend on the ordering of terms in the operation. It is true: The final position in space does not contain information about the way, how it was reached. But there is still another possibility how vectors can be added: Both vectors can act simultaneously and the point is shifted directly in direction between both of them as pulling a car by two ropes as on Fig.1.3

1.4 Matrices

Thus we need three possibilities how to write a sum of two vectors. We must have the opportunity to write them as consecutively acting vectors or

Figure 1.3: Vector action. Consecutive actions A and B and the simultaneous action S of two vectors **a** and **b** lead to the same final position R



as simultaneously acting vectors. Simultaneously acting unit vectors can be written easily as a sum of two unit vectors in a single row. The rule is simple, elements are added in their places:

$$(1,\ 0,\ 0) + (0,\ 1,\ 0) = (1,\ 1,\ 0).$$

In this notation we have already n simultaneously acting vectors in a row. Thus we must write consecutive vectors in such a sum as a column of row vectors. We get two different columns for our examples

$$\begin{pmatrix} (1, & 0, & 0) \\ (0, & 1, & 0) \end{pmatrix} \qquad \begin{pmatrix} (0, & 1, & 0) \\ (1, & 0, & 0) \end{pmatrix}.$$

Such columns of m vector-rows having in each row n elements are known as *matrices*. The row brackets and commas are wiped out of matrices. Notice that in a matrix its elements are arranged in columns similarly as in rows. It is thus possible to use the convention that a matrix is formed from n consecutive m dimensional vector-columns. Since we have introduced for individual columns the lower index j going from 1 to n, we can use for rows of matrices the index i going from 1 to m. Remember, the index i is present in texts implicitly, as the natural order of consecutive symbols. It need not to be given explicitly.

Sometimes it is convenient to let both indices start from zero. Then they go to $(m\ -1)$ or to $(n\ -1)$, respectively. It can be found one matrix index written over the other one. But it is better to reserve the upper index for powers. When two same symbols follow, For example: **aa**, is written shortly $\mathbf{a}^2$. Doing so, we treat consecutive vectors as if they were multiplied, and the multiplication is a *noncommutative* operation. The result depends on the ordering of terms in the operation. We will not use any symbol for multiplication of vectors and or matrices.

We have in our examples the small round brackets in all rows of matrices within the larger brackets used for matrices. Matrices are also bordered by double vertical lines or they are written into a frame. We will write them sometimes without any trimmings, but when they touch, we separate them by simple lines.

It is still necessary to consider different matrices with unit vectors:

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

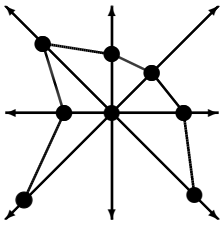
Matrices with empty rows are, maybe, superfluous, since no action corresponds to the given row, but notice that the third matrix can be obtained from the second one by rotating the elements around to the main diagonal, or changing the row i and column j indices. Matrices $\mathbf{M}$ are transposed into matrices $\mathbf{M}^{\mathrm{T}}$. A matrix with two identical unit vectors in consecutive rows can be interpreted as two consecutive translations going in the same direction. The resulting position in space can be obviously described by the vector $(2,\ 0)$. But if we try to interpret this vector with another numbers than 0 and 1, keeping in mind our convention that vectors in a row are simultaneous, we have some difficulties with the interpretation of these elements. We can imagine that the translation requires a greater force to be performed and that it has double intensity as in music a forte. To be consistent, we can not interpret other matrix elements than 0 and 1 simply as the length of a vector, unless we introduce such vectors by some algebraic operation which will make such multivectors allowed elements in our space.

The exclusion principle exists in quantum mechanics, formulated by Pauli. It states that in a system can not be two identical particles. From our experience we know that in one place can not be two things simultaneously. We will apply such principle for vectors, too. Lets limit at first on matrices having just one unit vector $\mathbf{e}_j$ not only in each position but in each row. We will use the symbol $\mathbf{e}_j$ not only for geometrical translations in space but also for different objects, e.g. for letters of this book. (I write letters in rows, therefore the text is a row of columns and each letter j must be substituted by the corresponding unit vector-column $\mathbf{e}_j^{\mathrm{T}}$). Matrices having one unit element in each row seem to be too "naive" to be studied, but we will see that they have quite interesting properties[3].

One of the useful properties of naive matrices $\mathbf{N}$ is that they can be interpreted either as a string of m unit vectors $\mathbf{e}_j$ going from the center of the coordinate system to some point in n dimensional space or as a position vector in m dimensional space. To keep our convention about con-

[3]It was difficult to find a name for them, because other suitable names, as primitive, elementary, were exploited.

Figure 1.4: A face in 8 dimensional space. The ends of individual vectors are connected with their neighbors by straight lines.



secutivity and simultaneity of vector translations, we transpose the naive matrix $\mathbf{N}$ into $\mathbf{N}^{\mathrm{T}}$. We write it as a row of the unit vectors-columns $\mathbf{e}_j$. The unit symbol will appear in the j-th row of the n dimensional column instead in the j-th column of the n dimensional row. The index system of the unit element is a convenient mark of the length of the vector $\mathbf{e}_i$ that goes from the center of coordinates in m dimensional space. There is no element which could interfere with this interpretation. But distances from the center can be zeroes. Therefore the row indices need to be counted from zero, subtracting one from each original index i. In such an interpretation the matrices $\mathbf{N}^{\mathrm{T}}$ correspond to *faces* (Fig. 1.4).

Drawing m vectors on the paper, indexing them consecutively, marking the length of each vector, and connecting the marks by straight lines we get a figure suggesting a face. Each face represents a point of the m dimensional space and there are as many faces as there are points in this space. Do you know your face? It is formed differently in different spaces.

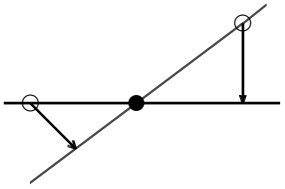
Only after we get acquainted with all naive matrices by counting them, we will study their sums and differences, that means the properties of matrices having in each row a sum or a difference of two unit vectors. Before we move to matrix arithmetic we are going to learn, how to operate with matrices. At first we introduce matrix products.

1.5 Scalar Products and Quadratic Forms

Having two vectors $\mathbf{a}$, $\mathbf{b}$, we can find mutual projections of both vectors as on Fig. 1.5.

The projections are known as the *scalar products*. If both vectors are orthogonal, the scalar product is $\mathbf{0}$, if they are collinear, the scalar product is, after normalization, 1. The unnormalized scalar product of a vector with

Figure 1.5: Scalar products. Both vectors are projected on the other one.



itself is known as the *quadratic form*. The *normalization* means that the scalar product is compared with the unit length of the projected vector. A scalar product seems to be therefore just cosine of the angle between both vectors. But it is not as simple as it seems to be.

The word product is connected with the operation of *multiplication*. How do we multiply two vectors? Take a vector column $\mathbf{v}$ and multiply it by a vector row $\mathbf{v}^{\mathrm{T}}$. Each element j of the column is multiplied by the matching element i of the row and the products are summed into one number. For example: $(1,\ 1,\ 1) \times (3,\ 1,\ 0)^{\mathrm{T}}$ is written in the form

			3
			1
			0
1	1	1	4

The result was obtained as $1\times3+1\times1+1\times0 = 4$. Otherwise multiplying matching elements vertically

$$\begin{array}{ll} (1,\ 1,\ 1) & \times \\ (3,\ 1,\ 0) & \end{array}$$

$$(3,\ 1,\ 0) \quad = 4$$

Changing the row and column position, we get the same result

			1
			1
			1
3	1	0	4

$3\times1+1\times1+0\times1 = 4$. When multiplied, the elements of one vector are weighted by the elements of other vector. All weights in the first example were 1. I think that you already know scalar products of vector columns

with the unit vector-rows, since they are used for finding sums of more numbers written in columns. In the second example the weights were 3, 1 and 0, respectively. The unit elements got different weights. Or the operation was simply $3 + 1 + 0 = 4$.

If a vector is weighted by itself, we get its *quadratic form*

			1
			1
			1
1	1	1	3

and

			3
			1
			0
3	1	0	10

Here we have $1 \times 1 + 1 \times 1 + 1 \times 1 = 3$ and $3 \times 3 + 1 \times 1 + 0 \times 0 = 10$, respectively. Corresponding elements of both vectors are multiplied and from products their sum is made. You already know the result of the first example, since it is simply a sum of n units (here $n = 3$). It seems to be elementary but it is not. Recall what was said about the Hilbert space and analyze the scalar product of the *unit vector* $\mathbf{J}$ with the unit vector $\mathbf{J}^{\mathrm{T}}$. (The unit vector $\mathbf{J}$ is the vector column, the unit vector $\mathbf{J}^{\mathrm{T}}$ is the vector row. All their elements are 1). The scalar product is just the sum of vector elements, the quadratic form is the square of their Euclidean length. If you think that we should work with square roots of quadratic forms, imagine that the unit vector $\mathbf{J}$ represents n people. The quadratic form just counts these people. Should we determine their number as $\sqrt{n}$ (square root from n)? We introduced the Hilbert space and we will work with scalar products and square forms as with basic vectors without finding roots.

We obtained in the scalar products from two n (m) dimensional vectors just one number. The multiplication decreased the dimensionality, we got just one number (scalar) determining the length of the first vector. Therefore the product of a vector row multiplied by a vector column from the right (natural order of both vectors, the vector column was multiplied by a vector row from the left) is called the *inner product*. There exists the *outer product*. This is obtained when we change the positions of both vectors and multiply a vector column with a vector row from the right:

	1	1	1
1	1	1	1
1	1	1	1
1	1	1	1

	3	1	0
3	9	1	0
1	3	1	0
0	0	0	0

Here three one dimensional vector columns acted on three one dimensional vector rows. The whole vector column was weighted by all elements of the vector column, and as the result matrices of dimension 3×3 were obtained. Instead two numbers we got two matrices, each having 9 matrix

elements. The outer product matrix is called *tensor*[4]. Notice that the elements of both inner products appeared as diagonal elements of the outer products. Their sum, known as the *trace of the matrix*, is identical with the final form of the inner product.

The scalar products can be made from matrix vectors, too. Scalar products of matrix vectors multiplied by vector-rows from the left are just vector-rows and matrix vectors multiplied by vector-columns from the right are just vector-columns, respectively. The multiplication is decreasing the dimensionality of the matrix vector:

$$\text{vector-row} \times \mathbf{M} = \text{vector-row} \tag{1.8}$$

$$\mathbf{M} \times \text{vector-column} = \text{vector-column}\ . \tag{1.9}$$

The vector-row is multiplied from the right consecutively by all columns of the matrix and the result has as many places as the matrix columns. The vector-column is multiplied from the left consecutively by all rows of the matrix and the result has as many places as the matrix rows.

If both vectors are matrices, the multiplication must be made for all combinations of rows and columns. The product is again a matrix. In the case of square matrix vectors, both products have identical dimensions and the distinction between inner and outer space is lost.

1.6 Matrices in unit frames

The quadratic form $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ counts the elements of the unit vector $\mathbf{J}$. It is simultaneously an operator

$$\mathbf{J}^{\mathrm{T}}(*)\mathbf{J}. \tag{1.10}$$

If we insert inside this product a matrix $\mathbf{M}$

$$\mathbf{J}^{\mathrm{T}}(\mathbf{M})\mathbf{J}\,, \tag{1.11}$$

we get the sum of elements of the matrix $\mathbf{M}$. $\mathbf{J}^{\mathrm{T}}\mathbf{M}$ is a n dimensional vector row and $\mathbf{MJ}$ is m dimensional vector column. The next multiplication by $\mathbf{J}$ (or by $\mathbf{J}^{\mathrm{T}}$) sums the elements of these vectors, respectively.

When we insert in (1.10) instead $\mathbf{M}$ as in (1.11) the quadratic forms $(\mathbf{MM})^{\mathrm{T}}$ or $(\mathbf{M}^{\mathrm{T}}\mathbf{M})$, we get the quadratic forms of the scalar products $\mathbf{J}^{\mathrm{T}}\mathbf{M}$ and $\mathbf{MJ}$.

[4]Tensor is a muscle that extends a part to which it is fixed. Tonsor is a barber.

We noticed that a vector column is transposed into a vector row and the other way around. If we repeat this operation, we obtain back the original vector form:

$$(\mathbf{v}^{\mathrm{T}})^{\mathrm{T}} \;=\; \mathbf{v}. \tag{1.12}$$

A matrix is transposed in such a way that all vector columns are transposed into vector rows and the other way around. It means that in the transposed matrix the indices i and j are exchanged.

At a transposition of a product of two matrices (a vector row or column is a matrix which has either $m = 1$ or $n = 1$, respectively) both matrices exchange their places, thus

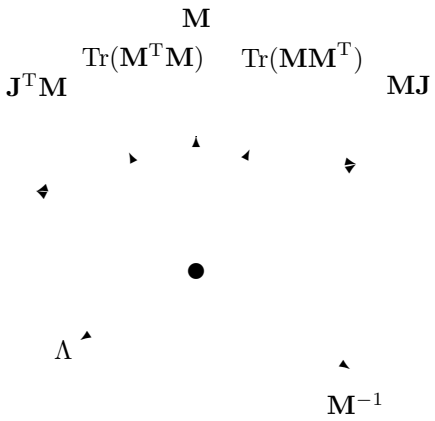
$$(\mathbf{J}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{MJ} \text{ and } (\mathbf{MJ})^{\mathrm{T}} = \mathbf{J}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}} \; . \tag{1.13}$$

We obtain two quadratic forms : $\mathbf{J}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{MJ}$ and $\mathbf{J}^{\mathrm{T}}\mathbf{MM}^{\mathrm{T}}\mathbf{J}$. We see that both products have the same frames $\mathbf{J}^{\mathrm{T}}(*)\mathbf{J}$, which act on a matrix product which is inside. This frame just counts elements of the inner matrix. The quadratic forms $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ and $\mathbf{MM}^{\mathrm{T}}$ are more important and interesting than the final product, because each of them contains more information.

We supposed that the original matrix $\mathbf{M}$ had m rows and n columns, both m and n being different. Therefore the transposed matrix $\mathbf{M}^{\mathrm{T}}$ had n rows and m columns and was different from the original matrix $\mathbf{M}$. We say that such matrices are *asymmetrical*. Both quadratic forms are symmetrical matrices. $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ has n rows and n columns, $\mathbf{MM}^{\mathrm{T}}$ has m rows and m columns. On the traces of both product matrices there are the sums of squared elements m_{ij}^2 of the matrix $\mathbf{M}$. This is the Hilbert length of the matrix vector and both traces which have the same length lie on a sphere with the diameter of the matrix vector. Off diagonal elements of both quadratic forms form with their traces the right triangles having both unit projections $\mathbf{J}^{\mathrm{T}}\mathbf{M}$ and $\mathbf{MJ}^{\mathrm{T}}$ as hypotenuses (Fig. 1.6).

Both scalar products transform a matrix into a vector, row or column. They count simply the elements in rows or columns of the matrix $\mathbf{M}$. They give us the final results of all translations, $\mathbf{MJ}$ in the m dimensional space of rows, $\mathbf{J}^{\mathrm{T}}\mathbf{M}$ in n dimensional space of columns. Finding these sums, we are reducing dimensionality of the space, instead of mn elements we have only m or n elements, respectively. When we reduced the dimensionality of the matrix space, we simplified the matrix vector, but we lost information about original order of vectors in the matrix. And moreover, at least in one quadratic scalar product, we joined together different vectors. If these vectors represented different things, we counted together apples with pears as fruits.

Figure 1.6: Matrix vector system. $\mathbf{M}$ – matrix vector, $\mathbf{J}^{\mathrm{T}}\mathbf{M}$ – matrix vector projection into columns, $\mathbf{MJ}$ – matrix vector projection into rows, $Tr(\mathbf{M}^{\mathrm{T}}\mathbf{M})$ – trace vector of the inner quadratic form, $Tr(\mathbf{M})\mathbf{M}^{\mathrm{T}})$ – trace vector of the outer quadratic form, Λ – eigenvalue vector, $\mathbf{M}^{-1}$ – inverse matrix vector.



The matrix vector system on Fig. 1.6 is composed from the matrix $\mathbf{M}$ itself and its two projections. $\mathbf{J}^{\mathrm{T}}\mathbf{M}$ and $\mathbf{MJ}$. These projections decompose into trace vectors, $\mathrm{Tr}(\mathbf{J}^{\mathrm{T}}\mathbf{M})$ and $\mathrm{Tr}(\mathbf{MJ})$, respectively. These trace vectors have an important property: They have the same length as the matrix vector $\mathbf{M}$ itself. Even the eigenvalue vector Λ has the same length as the matrix $\mathbf{M}$ and it can substitute both quadratic forms. The inverse matrix vector $\mathbf{M}^{-1}$, if exist, belongs to the matrix vector system (sometimes it can be substituted by the generalized inverse).

A matrix vector has mn elements. It is simplified by its projections into the separated spaces of rows and columns. We disregard in this projection some properties of the matrix vector. We are gaining some information but the price for it is loosing another information. Finding of quadratic forms corresponds to the logical abstraction. The most important property of both quadratic forms is their ability to replace matrix vectors. This ability is not a mere mathematical construction. It is based on physical experience because the world we live in is simply constructed in such a way.

To conclude: *A matrix corresponds to an action and its quadratic form to the result of this action.*

Both quadratic forms have this important property: They split the space

and its elements. Let the matrix $\mathbf{M}$ be a list of n different books (with an unknown number of copies) belonging to m different persons. Each row is a catalogue of the i-th personal library, each column is a list of occurrences, it registers in which libraries the j-th book can be found. The quadratic form $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ is the space of n books, on the diagonal there are numbers of libraries in which each book can be found. $\mathbf{M}\mathbf{M}^{\mathrm{T}}$ is the space of libraries but its elements are books. Compare it with ancient sayings that there is a measure in everything or that the measure of everything is man.

Chapter 2

Construction of the Vector Space

2.1 Number and Vector Scales

From the history of mathematics we know how carefully mathematicians constructed the number axis, introducing consecutively natural numbers, rational numbers, irrational numbers. It is not necessary to remember all problems connected with the notion of continuum and with the different axiomatic systems. The number axis forms one dimensional space. The next steps, the formation of two, three and more dimensional spaces were made as a audacious jump by so called Cartesian products.

The recipe seems to be simple: Take at least two one dimensional spaces and multiply them together. The set theory remedied some faults but it did not connect its set spaces with the vector spaces and both disciplines remained separated.

When we consider the natural number scale[1]
(0) — (1) — (2) — (3) — (4) — (5)
and compare it with a unit vector $\mathbf{e}_j$ scale

$$(0) \longrightarrow (1) \longrightarrow (2) \longrightarrow (3) \longrightarrow (4) \longrightarrow (5)$$

we see that the only difference is, that the vector scale is oriented and the number scale is not.

[1] Whole positive numbers, zero including.

2.2 Formal Operations with Vector Sets

We have introduced the unit vectors $\mathbf{e}_j$ in Sect. 1.2 as basic units of our space. At first, we will allow only positive translations corresponding to natural numbers. This means that a matrix vector can go only forwards from the center of coordinates and never back. A string of consecutive vectors forms a *path*. All possible paths in this space form a *lattice*. We already know, that the distinction must be made between a path and its final point, the *position vector*. This distinction is the same as between reading and mere counting words. Suppose that we have two vector strings for example

$$\mathbf{aababac}$$

and

$$\mathbf{abcaaba}\ .$$

Both lead to the point with coordinates (4, 2, 1, 0, 0, ...). We will write it sometimes as $(a^4b^2c^1d^0e^0\ldots)$. Such a notation is useful at some operations as

$$(a+b)^2 = a^2 + 2ab + b^2\ ,$$

where we need to distinguish the meaning of terms $2a$ and a^2. The multiplier gives the number of the strings, the power determines the length of the vector. Now it is convenient, that the base of the unit vectors is 1. The upper indexes, having meaning as powers of the unit vectors do not change them. When we accept that $x^0 = 1$, the zero power vector is just a multiplier one. Thus it is not necessary to write this 1, because it does not change the product as $a\times 1 = 1\times a = a$.

All vector strings ending in a point, as represented by the naive matrices $\mathbf{N}$, are *equivalent*. There are defined mathematical operations, which transform a naive matrix into another equivalent matrix. If this transformation does not give identical results, then both matrices belong to different classes. Two equivalent naive matrices have the identical quadratic form $\mathbf{N}^{\mathrm{T}}\mathbf{N}$ and lead to one point. For example

$$(\mathbf{aaba} = (\mathbf{a}^3\mathbf{b}) = (\mathbf{baaa})$$

Here we have the first example how useful the introduction of quadratic forms was. Later we formulate another equivalence classes of naive matrices.

To be able to distinguish between $2a$ and a^2 (between parallel and consecutive translations), we need the same difference also for construction of the multidimensional space from the unit vectors. Therefore, for vector sets, unit vectors, and their strings existing *simultaneously*, we will use the symbol of *summation* $\sum$. For *consecutive* vector sets, we will use the symbol for *multiplication* Π. The multiplication is transformed into the summation on a logarithmic scale. Using the unit base of logarithms, the number and its logarithm coincide, or do they not? For example

$$\mathbf{aaaaa} = \mathbf{a}^5,\ \lg_a \mathbf{a}^5 = 5$$

This convention inverses the order of both operations in the space construction. Classical way was to have two axis, say $(1 + a + a^2 + \ldots)$ and $(1+b+b^2+\ldots)$ and to multiply them. As the result we get positions points of a square

$$\begin{array}{cccc} 1 & a & a^2 & \ldots \\ b & ab & a^2b & \ldots \\ b^2 & ab^2 & a^2b^2 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

This square can be multiplied later by the third axis and the 3 dimensional cube is obtained, then the fourth axis can be applied and so higher dimensional cubes, sometimes called hypercubes, are obtained. We could speak about hyperplanes, hyperedges and so on, but we will not use this prefix because it would hyperinflate our text.

The space is constructed consecutively in layers from the sets of n unit vectors representing n dimensional space. For example:

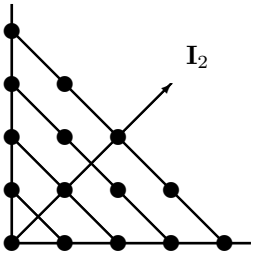
$$(a + b)^0 + (a + b)^1 + (a + b)^2 + (a + b)^3 + \ldots\ . \tag{2.1}$$

The individual products in the sum are vector strings ending on the lines orthogonal to the diagonal vector $\mathbf{I}$. The square with the side $0 - 2$ is obtained from these points by truncating points a^3 and b^3 from the incomplete layer 3 and adding 6 strings a^2b^2 from the fourth level product $(a + b)^4$:

$$\begin{array}{ccc} 1 & a & a^2 \\ b & 2ab & 3a^2b \\ b^2 & 3ab^2 & 6a^2b^2\ . \end{array}$$

The numbers at the coordinates give count of different vector strings going to the given natural point of the square. For example: 3 strings **aab**, **aba**, **baa** lead to the point with coordinates a^2b. The commutative

Figure 2.1: Two dimensional space. The unit vector $\mathbf{I}_2$ is orthogonal to the plane simplices.



algebra is obtained from the noncommutative one by the algebraic operation transforming vector strings into position vectors.

Vector strings created in the 2 dimensional space by the multiplication's $(a + b)^m$ go to the points lying on a straight line orthogonal to the diagonal of the complex as on Fig. 2.1. Later we fill this arranement with numbers, and we will obtain Pascal triangle.

The sum of n unit vectors $\mathbf{e}_j$ multiplied m times is the *generator* of the vector space. When the three dimensional generator is applied, the vector strings go to the triangle planes (Fig. 2.2). It ia known in mathematical literature as Pascal pyramid or tetrahedron.

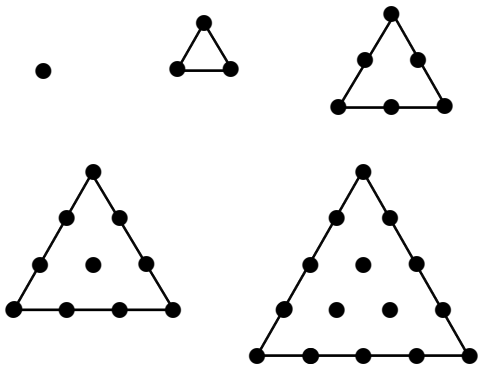
In higher dimensions these would be hyperplanes. Again, we truncate their names and modestly call them simply planes in all dimensions. But it inversely means that a line is a plane in the 2 dimensional space and a point is a plane in the 1 dimensional space. This might seem strange but an unlimited plane cuts its space into two parts. A point cuts the line similarly as a line divides a 2 dimensional plane into two parts.

Our vectors were limited only to natural numbers and therefore planes are generated by the operator

$$[\sum_{j=1}^{n} \mathbf{e}_j]^m \tag{2.2}$$

are the elements of the *natural space*. It includes its limit, points with the coordinates a^0b^0, ab^0, a^0b and so on. The elements of the natural space are countable and are formed by the vector strings going to points with nonnegative coordinates. We will call the individual layers the *plane simplices*. If you have heard something about simplices than you know that a simplex in n dimensional space should be determined by $(n + 1)$ points

Figure 2.2: The first five 3 dimensional plane simplices.



and we have just n points. But remember that we speak about planes. A plane in n dimensional space is a body[2] in $(n\ -\ 1)$ dimensional space and the missing point is restored.

The planes mentioned above are orthogonal to the *diagonal unit vector* **I**. It is necessary to explain, why there are three unit vectors: **I**, **J** and $\mathbf{J}^{\mathrm{T}}$. We have shown that the unit vector row $\mathbf{J}^{\mathrm{T}}$ and the unit vector column **J** have different effects on the naive matrices **N**, which are basic elements of our space, or generally on any matrices **M**. They transform them into vector rows or columns, respectively. Therefore we need a new unit vector invariant to the matrices. This vector is the unit diagonal vector **I**. It is the square matrix having the unit elements on the diagonal, where both indices are equal, $i\ =\ j$.

When the unit diagonal matrix **I** multiplies any matrix either from the left or from the right, it leaves the matrix unchanged:

$$\mathbf{IM} = \mathbf{MI} = \mathbf{M}. \tag{2.3}$$

The unit diagonal matrix **I** is known as the *identity matrix* and was already mentioned its sophisticated formulation as the Kronecker symbol δ_{ij}, where $\delta_{ij} = 1$, if $i = j$, and $\delta_{ij} = 0$ otherwise.

Lets continue with the construction of space using the plane simplices and laying consecutive layers as in an onion. The sums of plane simplices

[2]My son suggested here to add the adjective 'solid'. But a solid body is a solid body, whereas the plain term 'body' includes a code, a system, thus it is a more abstract notion.

form the *plane complex*. It is determined by three symbolical operations

$$\sum_{i=0}^{m} [\sum_{j=1}^{m} \mathbf{e}_j]^i \,. \tag{2.4}$$

If m goes to infinity we obtain the whole natural vector space of the given dimension.

Compare the sequence of operations with the traditional prescription

$$\prod_{j=1}^{n} [\sum_{i=0}^{m} \mathbf{e}^i]_j \tag{2.5}$$

and you will see that we just inverted the ordering of formal operations. The multiplication in (2.2) is done by upper index i. But we obtained another kind of space. Our space of vector strings is *noncommutative*, whereas the space formed by a lattice of points is *commutative*. The transition between both spaces is made by finding of scalar products. This formal operation corresponds to logical abstraction as it was shown in the previous Chapter.

2.3 Properties of Plane Simplices

One and two dimensional plane simplices are trivial. Our investigation starts with initial 3 dimensional plane simplices as on Fig. 2.2. The 3 dimensional plane simplices are triangles with 1, 3, 6 and 10 points. Each upper simplex has $(m + 1)$ more points than its predecessor and it is relatively easy to arrange them into the 3 dimensional complex. This forms the positive cone of the 3 dimensional octagon as on Fig.2.3.

The higher simplices differ from lower ones not only by an addition of a new edge but also by increased number of strings leading to all points except vertices.

If you compare the 3 dimensional plane simplex with the 2 dimensional complex, the difference between them consist in the number of strings going to different points. The origin (0, 0) gets the coordinate (0, 0, 3), points a, b are transformed into ac^2 and bc^2, respectively, and so on.

The 4 dimensional simplices are bodies in 3 dimensional space. They are regular tetrahedrons. If we try to draw them on a 2 dimensional surface, we must deform them as on Fig.2.4, where their edges have different lengths. And on a drawing, the inside of the tetrahedron does not appear, unless we do not draw it in a stereoscopical projection.

Figure 2.3: Three dimensional plane complex.

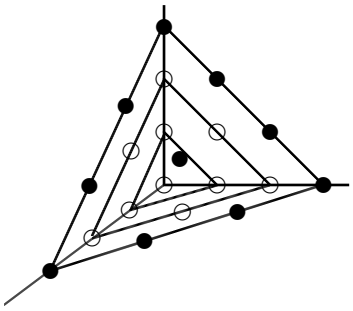
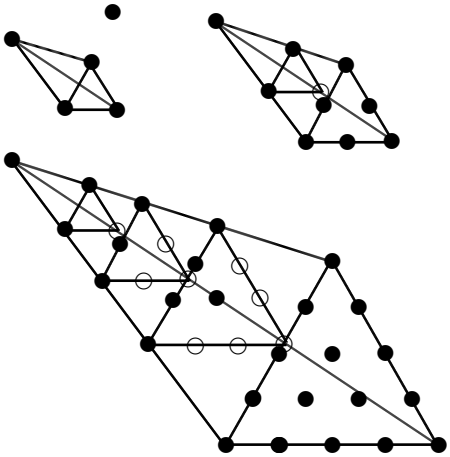


Figure 2.4: The first three 4 dimensional plane simplices and the fifth one.



The first difficulty appears: We are not able to form from 4 dimensional planes their complex. Why? All vertices of a tetrahedron must be in equal distances from the center of the coordinate system. An appropriate point seems to lie inside the tetrahedron, but the center of the tetrahedron has the coordinate (1/4, 1/4, 1/4, 1/4). The center of the system with the coordinate (0, 0, 0, 0) can not be inside the plane, it must lie outside it. The task of finding this point is similar to task to locate Nirvana. Breathing exercises do not help. Somewhat more useful is time. We just shift the whole plane from its original place for one unit length in our thought. Since this operation has no geometrical coordinate, it solves the task.

Even greater obstacles must be overcome when we try to imagine a five dimensional plane simplex as on Fig. 2.5.

Its envelope composes from five four dimensional planes, tetrahedrons having one coordinate zero:

$$\begin{array}{ccccc} a & b & c & d & 0 \\ a & b & c & 0 & e \\ a & b & 0 & d & e \\ a & 0 & c & d & e \\ 0 & b & c & d & e\,. \end{array}$$

In 3 dimensional space the tetrahedron sides interfere. If we draw the simplex as a trigonal bipyramide (Fig. 2.5 A), we can see in one moment two tetrahedrons, say abcd and abce, having the common side abc as the base of two trigonal pyramides, and in other moment three tetrahedrons having a common edge de, which goes through the bipyramide. But these are only sides of the simplex and its inside lies between these five tetrahedrons. We must move them aside before we come inside. It demands a concentration to enter inside planes of higher dimensions.

Or one tetrahedron can be flattened, say abcd (Fig. 2.5 B), and over this deformed base four tetrahedrons have place which cover the pyramid twice, once as two tetrahedrons abce and acde, once as two tetrahedrons abde and bcde. Into 2 dimensional plane the 5 dimensional plane simplex is projected as the pentagram (Fig. 2.5 C). In all cases the plane simplex is distorted by squeezing it into the lower dimensional space. In the ideal state all edges should have equal length.

The 5 dimensional plane simplices of the 6 dimensional plane simplex cover their 3 dimensional projection trice. The projection in the form of the tetragonal bipyramide can be divided into two pyramides having the common side abcd as the base, and then into four simplices along the axis ef as before at the 5 dimensional simplex.

Or one 5 dimensional plane simplex can be flattened into the regular pentagon and over this base five 5 dimensional plane simplices have

place which cover the base of the pentagonal pyramid 3 times, the corners of the pentagram 4 times its center 5 times. This makes the analysis of the 7 dimensional plane simplex difficult, since the pentagonal bipyramide is its simplest model.

Time and patience are essential when analyzing planes of higher dimensions. Decompose them into subplanes and decrease their dimensionality as your homework.

A conjecture outside mathematics: Multidimensional objects can appear in lower dimensional space only changing constantly their configurations. Thus microparticles emerge in the wave form.

2.4 Construction of the Number Scale

Until now we used only natural numbers and vectors. But we will need fractional numbers and vectors, too. Now we are able to introduce them because we have enough space for necessary constructive operations.

Recall the 2 dimensional complex (Fig. 2.1).

The position vector (1,1) goes through the plane simplex $(a + b)^1$ in a point which has until now no name in our world. We introduce it by finding its coordinates on both axes. This is done using parallel lines with both axes. The new numbers are defined as the ratio of the coordinate a of the position vector and the power of its simplex, or as the ratio of the coordinate b of the position vector and the power of its simplex, respectively. In the example the ratio is 1/2.

When this operation is done with all simplices going to infinity (or equivalently with the infinite simplex), we obtain infinite many points in the interval $< 0,\ 1 >$. All these points are countable by indices i of the infinite plane. They are known as *rational numbers*. The rational numbers outside the interval $< 0,\ 1 >$ are obtained by adding the rational number and the natural number (or by multiplying).

The infinite plane simplex itself remained at this operation undivided, as it was in its natural state. We use again one of the properties of the Euclidean space, namely that parallel lines never meet and translate the fine division of rational numbers from the first simplex onto the infinite plane (Fig. 2.7).

New position vectors appear on it. They cross the unit simplex in points, which all lie before the first rational number. They divide the angle between the first countable rational vector from the primary infinite division of the unit interval and therefore form a new set of infinite many points on the number scale. The operation can be repeated ad infinitum. The first set of the irrational numbers is sufficient for representation of

Figure 2.5: Three projections of the 5 dimensional plane simplex. A – the bipyramide, B – one tetrahedron side flattened, C – the whole simplex is flattened.

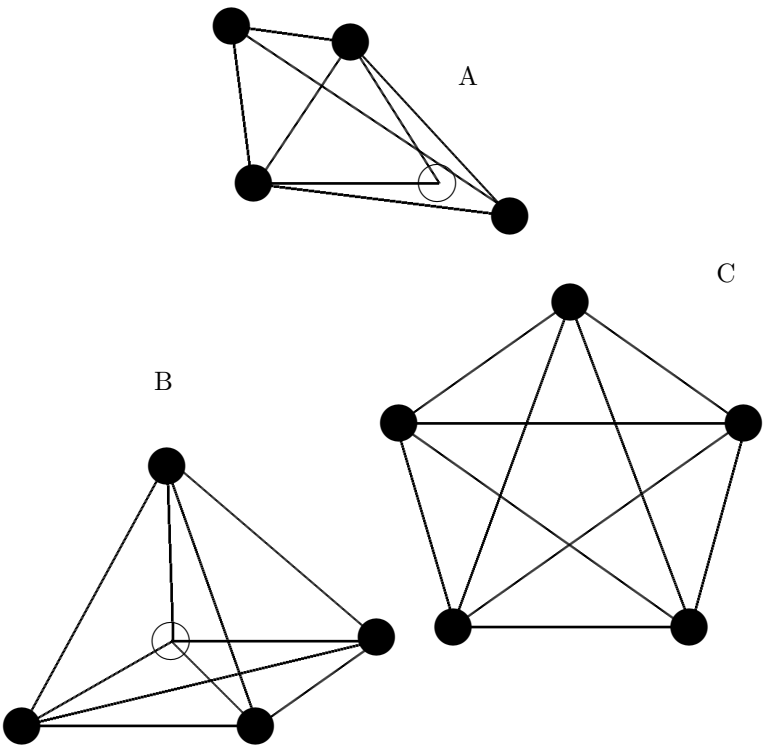


Figure 2.6: Construction of the rational numbers. Vector (1, 1) intersects the first plane simplex in the point with the coordinate (0.5, 0.5).

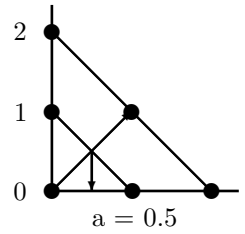
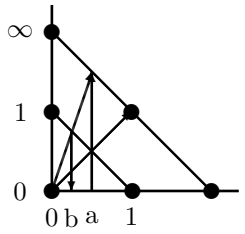


Figure 2.7: Construction of irrational numbers. The vector leading to the projection of the first rational number a onto the infinite plane simplex has as the coordinate the irrational number b.



continuum. Its elements are not be countable because the infinite store of numbers is exhausted by counting the first crop of the rational numbers. The uncountable numbers of the second crop are *irrational numbers*.

We already need such numbers which we are not able to write explicitly. If we return to the simplex plane and try to measure the length of the vector leading to the point (0.5, 0.5), or (1, 1), rotated onto the axis, we will not find it between the rational numbers. Square root from 2 ($\sqrt{2}$) is not a rational number.

The numbers obtainable by the consecutive divisions of continuum and eliminating the decimal point of are known as *aleph numbers*. In the Euclidean space everywhere and always is true that $1 \times 1 = 1$. Nowhere the product is an irrational number greater or lesser than 1.

2.5 Complex Numbers

We have shown that a matrix vector $\mathbf{M}$ can be projected onto the unit vector row $\mathbf{J}^{\mathrm{T}}$ or column $\mathbf{J}$, and that the quadratic forms $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ and $\mathbf{M}\mathbf{M}^{\mathrm{T}}$ can be separated into right triangles. This is true for matrix vectors in which all elements are either positive or negative. If a matrix vector contains numbers of both signs, its projections are shorter than the matrix vector itself. Then the hypotenuse of the right triangle (Fig. 1.2), represented by the trace of the quadratic form, is longer than the outer product, where the off-diagonal elements form the legs. The off-diagonal elements can be either positive or negative. For example:

	-3	-2	1	Σ
3	9	-6	3	6
-2	-6	4	-2	-4
1	3	-2	1	2
Σ	6	-4	2	4
				Trace = 14.

The diagonal vector length (trace) is 14, the off-diagonal vector length (sum off-diagonal elements) is -10, the outer product length (the projection on the unit vector is 4, it means that it is shorter than the vector itself. The negative sum of the off-diagonal elements indicates that their sum must be subtracted from the trace, not added to it. This changes the construction of the triangle.

You have probably heard about imaginary numbers i, square roots from the negative number $\sqrt{-1}$. When they appeared as possible solutions of quadratic equations, mathematicians feared them as ghosts. Only later Euler showed, how they can be exorcised by mapping them onto a complex plane (Fig. 2.8).

Now, if we have a number z in the form

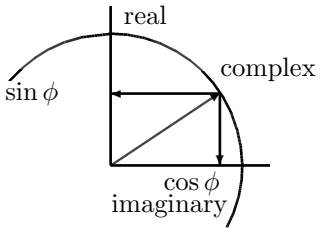
$$z = (x + iy) \text{ or } z = r(\cos\phi + i\sin\phi)\ , \tag{2.6}$$

we can divide it into a right triangle and replace a linear vector by a plane vector, which is always composed from two elements, one real and one imaginary. There are specific rules for calculating with complex numbers and especially with matrices containing complex numbers.

2.6 Generating Functions

We have shown how a complex is constructed from its simplices. This technique is used intensively in combinatorics for generating functions. A

Figure 2.8: Complex numbers. They are composed from the real and imaginary parts.



space is defined by some functional relation, usually a sum or a product, which argument goes from 0 to ∞. The generating function is evaluated with a dummy variable, for example t, and coefficients at different powers of t are calculated.

Because strings $\mathbf{x}_a\mathbf{x}_b$ and $\mathbf{x}_b\mathbf{x}_a$ are undistinguishable in the commutative process, it was considered as impossible to formulate a generating function which would exhibit the order of symbols in products (permutations). Nevertheless, the enumerators are easy to find in the form

$$\sum_{k=0}^{n} t^k/k! \; . \tag{2.7}$$

These enumerators are known as exponential generating functions.

It is possible to make different algebraic operations with generating functions, For example: to find their sums, products, etc.. The corresponding operations are known as Cauchy and Blissard algebra's. There are many conceptual problems connected with the convergence of infinite series for different arguments. We simplify them by using unit vectors and the properties of Euclidean space. Only exceptionally, we mention some deformations of the ideal space.

2.7 Generalized Unit Vectors

By using unit vectors $\mathbf{e}_j$ we narrowed the possibilities of the calculus. The simplification has many advantages but they must be paid for. Some formulas in next chapters are true even if $\mathbf{e}_j$ is not 1 but any number. For example: $(a+b+c)^k$ can be evaluated as $(1+2+3)^k$ as well as $(2.1+0.1+5)^k$ depending on actual values of variables. It is true even for geometrical representations of variables. It is possible to imagine, as if the space were

elastic and its lattice could be stretched as required. Each specific case can be divided into a part which is isomorphic with the ideal case, and into the specific distortion of the unit vectors.

2.8 Trigonometric Functions

We discuss shortly trigonometric functions, sinus, cosines, tangents and cotangents. They connect values of angles in the right triangle with ratios of legs to hypotenuse. If α is the angle of the leg b and the hypotenuse c, its opposite side being a, the definitions of trigonometric functions are

- $\sin \alpha \ = a/c$
- $\cos \alpha \ = b/c$
- $\tan \alpha \ = a/b$
- $\cot \alpha \ = b/a = 1/tan\,\alpha$
- $\sin \alpha \ = cos\,\beta$.

The sides of both angles change their positions.
The formula

$$sin^2\alpha + cos^2\alpha = 1$$

is in fact the Pythagorean sentence in the form:

$$(a/c)^2 + (b/c)^2 = (c/c)^2 \ .$$

2.9 Natural Numbers and Numerals

Two basic definitions of the natural numbers are Peano's axiomatic one and the von Neumann's set model. Both definition are strictly functional, they do not provide for the relations between the numbers and the numerals as natural names of the natural numbers and their written form, their notation.

Peano defined natural numbers by the algorithm which forms from a number k a greater number by adding one $(k + 1)$. It is a nice approach and we already exploited it for generating the space where instead 1 new simplex layers were added. Later we derive a generalization of Peano definition: a natural number is any sum of natural numbers.

The von Neumann set model generates numbers by counting sets. The empty set {0} has one element, it generates 1. The set containing {0, 1} has two elements, it generates 2, and so on.

All languages I know, have numerals k for numbers 0 to ten. Numerals for 11 – 19 are formed as $(10 + k)$ For example: fourteen. Eleven and twelve are corrupted because they were used often.

Multiplets of tens are expressed by one numeral formed as ($k\times$ty=ten), For example: forty. Hundreds and thousands are counted separately, then only kilomultiplets of thousands (million, ...) have their own numerals. Numbers between these pivots are expressed as linear combinations of basic numerals.

Of course exceptions exist, as mentioned 11 and 12. For example, corruptions and exceptions of numerals appear to one hundred in Hindí. Ancient Egyptians had specific names and hieroglyphs for decimals.

Number notations had different forms: In the primitive form, one cut on a stick corresponded to each counted object. Egyptians introduced specific signs for powers of 10 to 10^7, but numerals one to nine expressed primitively by the corresponding number of signs. Phoenicians introduced letters for 1 – 9, 10 – 90 and 100 – 900. It shortened the notation considerably. This system has been taken over by Hebrews and Greeks. Romans used their own system. Specific symbols were reduced on I, V, X, L, C, D, and M and the number of necessary symbols in one number by using a position system IV = one hand without one, IX = two hands without one. Finally, we have Indian-Arabic decimal position system.

The Mayan score system should be mentioned with position notation, where zero with a numeral signified multiplication by 20 (quatre-vingt in French four twenties) and the Babylonian hexadecimal system (German Schock, Czech kopa), where powers of three scores were expressed by size of their symbol (compare dozen – gross – great gross).

Numerals, that is the names of numbers, are generated by a modular system which is based on our fingers. We count sets by grabbing them with our hands and it is the natural way we speak and think about numbers in the decimal system. The definition of the natural numbers should express this fact. Therefore I propose the following definition:

The natural numbers are generated by a series of modular operations, comparing two sets, the compared set $\{n\}$ *and the modular set* $\{m\}$.

The empty set {0} is for obvious reasons unsuitable as the modular set {m}.

The set {1} as the modular set $\{m\}$ generates the natural number 0, only, since

$$\{n\} \bmod \{1\} \equiv 0 \ .$$

The set $\{2\}$ generates the natural numbers 0 and 1.

Using great enough modular set $\{m\}$ we obtain in one modular operation all natural numbers. But it is inconvenient because we do not have an unlimited store of simple symbols and numerals for them. Therefore, a series of modular comparisons must be used which result in a series of modular identities. The position notation leads to the modular equalities:

$$\{135\} \text{ mod } \{10\} = 135$$

$$\{135\} \text{ mod } \{4\} = 2013$$

The written form of the number is obtained by the series of consecutive divisions with the modulo rests

- $135 : 4 = 33 + 3$
- $33 : 4 = 8 + 1$
- $8 : 4 = 2 + 0$
- $2 : 4 = 0 + 2$

The resulting number modulo 4 is formed as the position combination all modular rests written from the first one from the right to left, where the last rest is written as[3] = 2013.

Although the set $\{1\}$ seems to be a natural base for a number system, and the objects in sets already exist in such a form, a series of modular comparisons with 1 gives only a series of zeroes. A division by 1 does not decrease the digit size of a number and it does not compress the notation. Therefore, such a number system is impractical. The binary system is the first applicable.

The modular operation is essentially a mechanical one. In the first step the line of elements is cut into rows by the given modulo. The last line which is incomplete (it can be empty) is the result of the modular operation

*****	mod	**:	**	
			**	
Rest			*	= 1.

One column of the complete rows is transposed into the row and the operation is repeated

[3]This Semitic writing was accepted from Phoenicians.

**	mod	**:	**
Rest			0 = 0.

One full column obtained by the second modular operation is again compared similarly until all elements are exhausted

*	mod	**:	0 (the number of complete rows)
Rest			* = 1.

The result is the binary notation ***** = 101. The third modular operation was in fact the division by the second power of 2, the third rest gives the number of fours in the original set. In the binary notation, they are determined by their third position from the last digit giving the number of $1 = 2^0$. A number of a smaller modulo is simultaneously a number of a greater modulo. The binary number four looks like the decadic number hundred ($4 = 100$).

Two natural numbers are equal if they are obtained from the same set $\{n\}$, and comparable if they are determined using the same modular set $\{m\}$.

Compared with the von Neumann set model, where joined sets $\{\{0\}, \{1\}\}$ produce the number 2, here the generating set $\{2\}$ covers the numbers 0 and 1.

The advantages of the proposed definition are obvious: It connects the natural numbers with the cardinal numerals by the algorithm which shows how the names and notations of the natural numbers are formed from the numerals. It is logical: Numbers which are described in natural languages by combinations of the cardinal numerals are the natural numbers.

Later we will show a generalization of the Peano algorithm.

Chapter 3

Linear Operators

3.1 Introduction

Vectors are operators which shift a point on another place in a space. In this chapter special operators will be discussed which act on sets of points or on sets of vectors as if they were one point or a solid body. Some of these operations were already mentioned but now they will receive more systematic treatment. Nevertheless, some important properties of operators will become clear only later, after graph operators will be introduced and exploited to obtain practical results.

The operators can be divided into *additive*, as vector translations are, and *multiplicative*, as scalar products are. Another aspect of classification is possible depending on how many matrices or vectors are affected. The operation can proceed inside one matrix, or one matrix operator can act onto another vector or matrix. Remember that a vector row or a vector column are matrices with just one row or column, respectively.

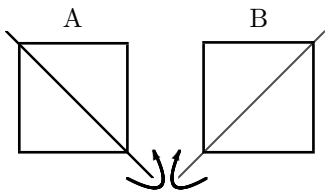
3.2 Transposing and Transversing

Transposition of matrix vectors was already defined. It changes simply the row indices i and column indices j of all matrix elements

$$\mathbf{M}^{\mathrm{T}} \rightarrow \mathrm{m}_{ij}^{\mathrm{T}} = \mathrm{m}_{ji} \ . \tag{3.1}$$

If $\mathbf{M}^{\mathrm{T}} = \mathbf{M}$ the matrix is *symmetrical*. This property has important consequences for other properties of matrices. It is interesting that the transposition changes the ordering of terms in matrix products:

Figure 3.1: Transposing (A) and transversing (B) of matrices.



$$(\mathbf{ABC})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \ . \tag{3.2}$$

A conceptual problem is connected with transpositions. We accepted the convention that rows of a matrix mean ordering in time, the consecutiveness, whereas columns are ordered in space as orthogonal vectors. Transposition changes this ordering. But remember a book. All words exist simultaneously, we are only forced to read them consecutively. The similar function has time in vector space but it is not conventional time which is measured by clock. All matrix elements exist simultaneously in all instants. Otherwise we would need another algebra.

The second operation introduced here, *transvesing*, is not used in textbooks but we need it to prove simply without calculations some combinatorial identities. The transversing changes the ordering of both indices, that means rows and columns are counted backwards. If transposing rotates matrix elements around the main diagonal $m_{11} \longrightarrow m_{nn}$, transversing rotates them around the diagonal (its name will be transversal) $m_{1n} \longrightarrow m_{n1}$ (Fig. 3.1). We look on the matrix's most distant corner as its starting point.

3.3 Translating and Permuting

We translate a sentence from a language into another, or we translate it as a block from a place in a file in another place. Similarly we can translate vectors, or their strings. Now we must find techniques to express different kinds of translations in an abstract way. Essentially there are two possibilities how such translations can be achieved. The operators can be additive or multiplicative.

The additive operator is formulated as a difference. We take two states of a matrix vector, the original $\mathbf{M}_1$ and the final $\mathbf{M}_2$ and the searched operator $\mathbf{S}$ is just their difference:

$$\mathbf{S} = \mathbf{M}_2 - \mathbf{M}_1 \ . \tag{3.3}$$

For example

$$\mathbf{N}_1 \qquad \mathbf{N}_2 \qquad \mathbf{S}$$

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

It looks trivial but a special branch of mathematics, graph theory, studies only these operators and vectors orthogonal to them. According to our convention, a row shifts one symbol into another. It corresponds to coding a message, in transposed form it grimaces faces shown on Fig. 1.4.

Each row of an operator $\mathbf{S}$ is the difference of two unit vectors $\mathbf{e}_j$. The negative $\mathbf{e}_a$ is going from the vertex a back to the center and the path through the space continues by the vector $\mathbf{e}_b$ to the vertex b. The resulting simultaneous translation is a vector going directly from the vertex a to the vertex b without touching the center (Fig. 3.2).

The unit vectors $\mathbf{e}_j$ are *primary* vectors, their sums or differences s_{ij} are *secondary* vectors. Their space is bent in angle 45^0 to the primary space. To each sum $(i+j)$ belong two differences, $(i-j)$ and $(j-i)$.

The operator $\mathbf{S}$ is a string of such secondary vectors. These vectors form edges of the plane simplex n^1. They do not go from the center to some point of the space, but they change a vector string into another one going to the same simplex. Since both vector strings are continuous paths, the operator that translates one into another lies on a surface in the n dimensional space (Fig. 3.3).

The sum of two unit vectors $(\mathbf{e}_j + \mathbf{e}_i)$ is orthogonal to the difference $(\mathbf{e}_j - \mathbf{e}_i)$ and the corresponding matrices $\mathbf{G} = \mathbf{N}_1 + \mathbf{N}_2$ differ from matrices $\mathbf{S}$ only by positive signs relating both unit vector strings. Since each consecutive element in the string is orthogonal, the $\mathbf{G}$ represent vectors orthogonal to the operators $\mathbf{S}$. Matrices $\mathbf{G}$ are linear vectors orthogonal to the surface of the operator $\mathbf{S}$. They form a secondary vector space, which is not complete, as we will see in the second part of this book.

The first multiplicative operators allowed to form our space, are determined by properties of a special class of naive matrices $\mathbf{N}$, which have one unit symbol not only in each row but also in each column. These matrices $\mathbf{P}$ are known as the *unit permutation matrices*. The unit diagonal matrix $\mathbf{I}$ belongs to them. All permutation matrices are square matrices

Figure 3.2: Representation of arcs and edges as vector sums or differences.

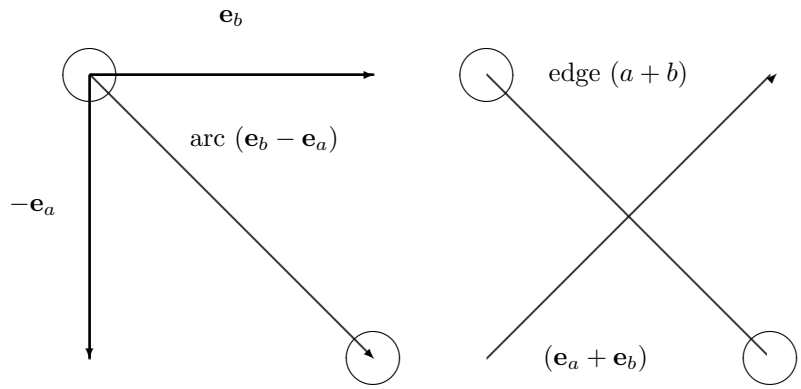
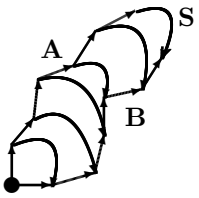


Figure 3.3: Difference of vector strings **A** and **B** forms the surface **S**.



and they form groups S_n of permutation matrices with n rows and columns. When a matrix is multiplied with a permutation matrix from the right, this operation changes the ordering of columns of the multiplied matrix. For example

$$\begin{array}{cccc|cccc} & & & & 0 & 1 & 0 & 0 \\ & & & & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \\ & & & & 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}$$

The first column appears as the second one in the product since the matrix **P** has 1 in the second column in the first row. The last (zero) column is similarly shifted on the first place by the last unit element in the first column.

The multiplication from the left changes the ordering of rows of the multiplied matrix. For example

$$\begin{array}{cccc|cccc} & & & & 1 & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 \\ & & & & 0 & 1 & 0 & 0 \\ & & & & 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}$$

On Fig. 3.4, where the effects of 6 permutation matrices of the group S_3 on a three dimensional plane simplex are depicted, we can see the effect of such multiplication of columns. The unit diagonal matrix leaves the simplex unchanged, two matrices rotate it along its center and three matrices change the positions of only two vertices as the triangle were mirrored along the plane orthogonal to the corresponding edge (or rotated along an axis lying in the plane). These are *symmetry operations*. They will be studied later in more detail.

All permutation matrices with n rows and columns are obtained as consecutive rotations and form the cycle group S_n. They rotate vectors in after given number of repeated operations the vectors return back to their original position.

Figure 3.4: Symmetry group S_3. A – the identity, all elements remain on their places; B, C, D – reflections, pair of elements interchange their places; E, F – rotations, three elements exchange their places in cycles.

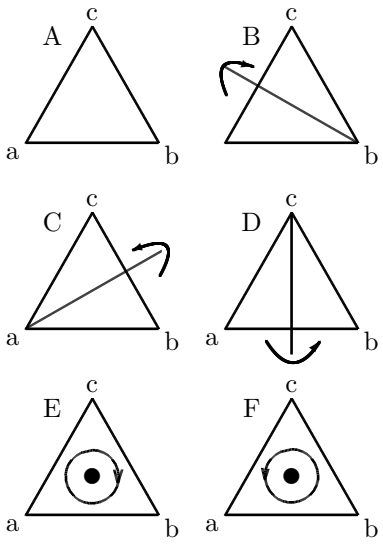
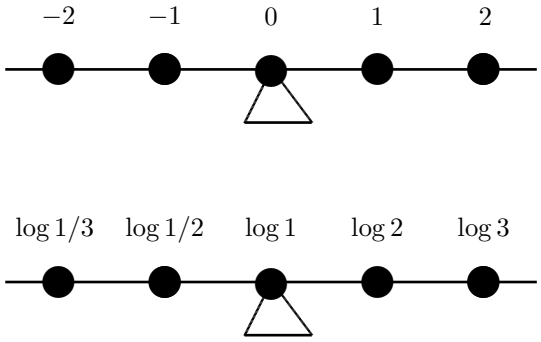


Figure 3.5: Additive and multiplicative balancing of numbers.



3.4 Inverse Elements

When we have a number, say 5, we can define its inverse element again by two modes, additive and multiplicative. Similar elements can be defined for vectors.

The inverse operation to addition is subtraction. The number 5 was obtained from 0 by adding 5 and we restore the original situation by subtracting 5: $5 + (-5) = 0$. The inverse additive element of 5 is (-5) and the inverse additive element of (-5) is 5. We can imagine these numbers on a balance (Fig. 3.5). Additive inverses are just vector collinear with the parent vectors, having equal length but opposite direction. They are formed by changing sign of the vector.

Now we can consider the inverse element for the multiplication operation:

$$a \times a^{-1} = a^0 = 1 \, .$$

When we apply the logarithmic scale we get

$$\log a + \log a^{-1} = \log 1 = 0 \, .$$

From it we find that $a^{-1} = 1/a$. On the number scale, the inverses of numbers greater than 1 are in the range (0,1), which seems to be unbalanced, see Fig. 3.5, but it is balanced on the logarithmic scale.

It seems that it is easy to find the inverse vectors to vectors-columns (or vector-rows). They must give the unit scalar product, For example:

			3
			1/2
			1
1/6	1	0	1

			1/6
			1
			0
3	1/2	1	1

But such inverses have one essential disadvantage: They are not unique. There exist infinite many such inverses which balance each vector-column (or each vector-row), therefore they are *undetermined*, For example: another suitable solution is:

			3
			1/2
			1
1/9	2/3	1/3	1

			1/9
			2/3
			1/3
3	1/2	1	1

If we try to find a left (right) inverse for a matrix, its rows must be left (right) inverses for corresponding columns (rows), but simultaneously zero vectors for other columns (rows). In the given case the zero vector is again undetermined:

			3
			1/2
			1
1	0	-3	0
-4/3	2	3	0

			1	-4/3
			0	2
			-3	3
3	1/2	1	0	0

Another difficulty with inverse elements of vectors is, that one can not find a right inverse to a vector-column (left inverse to a vector-row):

	?	?	?
	?	?	?
	?	?	?
3	1	0	0
1/2	0	1	0
1	0	0	1

			3	1/2	0
?	?	?	1	0	0
?	?	?	0	1	0
?	?	?	0	0	1

There were 1/3 as the first inverse element m_{ij}^{-1}, but it can not be nullified in following rows of the first column. For its nullification we needed

some nonzero elements in the second and third columns of the left matrix. For matrix vectors we can, at least sometimes, find matrices which transform all their vector columns into a diagonal matrix. One vector column does not have any inverse from the right, but their system has. Which properties a matrix must have for being inversifiable will be shown later. If a matrix has inverses both from the left and from the right, then both inverses are identical and there exist only one inverse which action is equal from both sides. This is the true inverse of the given matrix.

We could search inverses trying haphazardly suitable vectors. Better is to use some verified algorithms, which will be introduced later.

A matrix having the inverse is *regular* or *nonsingular*. Nonsingular matrices have none zero eigenvalues and eigenvectors and singular matrices have at least one zero eigenvalue and eigenvector. The *eigenvector* is a vector in which all elements are multiplied, if the vector is multiplied by the given matrix, by the same value which is called *eigenvalue*. For example

			1	1	1
			1	-2	0
			1	1	-1
		Π	0	3	-1
1	-1	0	0	3	1
-1	2	-1	0	-6	0
0	-1	1	0	3	-1

The first column is the zero eigenvector, all values in its column product are zero, and the second eigenvector eigenvalue is 3, the eigenvalue of the last eigenvector is 1. There is yet another condition on eigenvectors, see the next Section.

Some nonsingular matrices are easily recognizable. If a matrix has all nonzero elements below or over the diagonal and all diagonal elements are unit elements, then it is nonsingular. The inverse in this case can be simply found by a technique known as the *principle of inclusion and exclusion*. Suppose that k rows were already balanced. In the next row the scalar products of vector rows with the inverse matrix columns (multiplication from the right is supposed) will be unbalanced by some value. We must add or subtract as many elements to it for obtaining zero off-diagonal elements. For example (zero symbols are omitted)

			1		
			-2	1	
			1	-2	1
1			1		
2	1			1	
3	2	1			1

.

Here the second row balances are $1 \times 2 - 2 \times 1 = 0$, and $1 \times 1 = 1$. In detail, the problem of inverse matrices will be treated in Chapt. 16.

3.5 Diagonalization of Matrices

An inverse matrix transforms a matrix $\mathbf{M}$ into the diagonal unit matrix $\mathbf{I}$ but there is still another form of diagonalization. This operation demands a simultaneous action of two matrices from both sides of the matrix which has to be diagonalized

$$\mathbf{L}(\mathbf{M})\mathbf{R} = \Delta(\mathbf{M}) \,. \tag{3.4}$$

$\Delta(\mathbf{M})$ is a diagonal matrix which has all off-diagonal elements zero. The matrix in the brackets is the source of the diagonal elements.

The product $\mathbf{M}\mathbf{M}^{-1}$ were an example of a matrix diagonalization where one from diagonalizing matrices is the unit diagonal matrix $\mathbf{I}$. It is required from diagonalizing matrices that the action of a matrix $\mathbf{L}$ from the left were balanced by the multiplication by a matrix $\mathbf{R}$ from the right. The diagonalization matrices form a frame for the matrix $\mathbf{M}$.

Imagine, that you observe the matrix as between two polarizing filters. When the filters rotate, the view clears or becomes dark, but at one position the filter is transparent. Such transparency of matrices we look for. Both diagonalizing matrices function as polarizing filters, they decrease off-diagonal elements and increase diagonal ones. A diagonal matrix is transparent since diagonal elements are not obscured with the off-diagonal ones. Recall Fig. 1.6. The obtained diagonal matrix is equivalent to the matrix of $\mathbf{M}$.

The especially useful effect is obtained, when the product of both diagonalizing matrices $\mathbf{L}$ and $\mathbf{R}$ is the unit diagonal matrix

$$\mathbf{L}\mathbf{R} = \mathbf{I} \,, \tag{3.5}$$

or equivalently when their action does not change the unit diagonal matrix in their frame:

$$\mathbf{LIR} = \mathbf{I} \, .$$

Then, if moreover

$$\mathbf{L} = \mathbf{R}^{\mathrm{T}} \, , \tag{3.6}$$

we say that these matrices are *eigenvectors* of the given matrix. The diagonal matrix obtained as the result of such a multiplication is known as the matrix of *eigenvalues*. The sum of eigenvalues is equal to the trace of the diagonalized matrix and the diagonal matrix of eigenvalues is equivalent to the matrix vector of the diagonalized matrix.

The vector set used for finding of eigenvalues gives the diagonal matrix, but not the unit matrix **I**:

			1	1	1
			1	-2	0
			1	1	-1
1	1	1	3	0	0
1	-2	1	0	6	0
1	0	-1	0	0	2

The eigenvectors must be now normalized by dividing with square roots of 1/3, 1/6 and 1/2, respectively. The normalized eigenvectors are

$$\begin{pmatrix} \sqrt{1/3} & \sqrt{1/6} & \sqrt{1/2} \\ \sqrt{1/3} & \sqrt{-4/6} & \sqrt{0} \\ \sqrt{1/3} & \sqrt{1/6} & \sqrt{-1/2} \end{pmatrix}$$

The eigenvalues and eigenvectors are not an abstract mathematical construction, but a result of practical experience. Eigenvalues are known from physical and technical sciences. Eigenvectors are known as factors when used in the Chapt. 15.

3.6 Matrix Arithmetic

Sums and differences of vectors were already discussed. It is important to examine arithmetic of matrices more thoroughly, because in textbooks you can find different restrictions on how matrices can be combined.

Arithmetical operations with matrices are limited usually to matrices of identical dimensions, having equal number of rows and columns. It is a too rigid rule. Before a less strict rule will be introduced, we inspect all possible cases, how matrices can be related, if their indices obtain their true values, as if two documents are compared and ordered (Fig.3.7).

Figure 3.6: Matching of matrices according their indices.

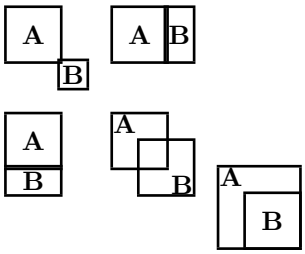
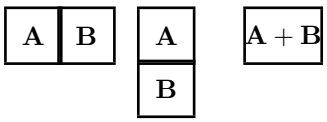


Figure 3.7: Matrix addition and subtraction possibilities.



The row indices go in each matrix from 1 till m, the column indices go in each matrix from 1 till n. This is internal counting. Similarly as Jewish, Christian and Islamic epochs, sets of indices in compared matrices can be unequal, or one set be the same, or both sets can match. Thus the rule of matrix arithmetic for addition and subtraction of matrices is simply addition and subtraction of individual matrix elements according the rule:

$$\text{if } \mathbf{A} \pm \mathbf{B} = \mathbf{C}, \text{ then } \mathrm{a}_{ij} \pm \mathrm{b}_{ij} = \mathrm{c}_{ij} \ . \tag{3.7}$$

The difficulty is based on the question, what to do with unknown matrix elements. If they are zero, the results can be as on Fig. 3.7. Before the arithmetical operation is done, one or both matrices are completed to equal dimensions by adding zero elements in missing rows and columns. The cases of arithmetical operations in blocks is known as the direct sum or difference of matrices. If unknown matrix elements are not zero, the operations lead to errors.

Essentially the same conditions hold for matrix multiplication's. We have explained the effect of permutation matrices and scalar products of vectors. If we multiply a matrix by a vector column from the right, the row elements $\mathbf{v}$ of the matrix multiply all elements of the column. If elements of

v are smaller than 1, they shorten all elements of this column, if elements of **v** are greater than 1, they increase them. Two simultaneous processes occur at multiplication: the elements of matrix rows are weighted and summed, or if vector elements are negative, they are subtracted. Multiplication from the left has the transposed effect. The multiplication of a matrix by a vector transforms the matrix into the vector. Usually, it is defined otherwise, a matrix transforms a vector into an another.

A simultaneous multiplication of a matrix by a vector row from the left and by a vector column from the right, transforms the matrix into one element. If both vectors are unit vectors $\mathbf{J}^{\mathrm{T}}$ and $\mathbf{J}$, they just sum the matrix elements.

It is useful to define also a *direct product* of two matrices. To distinguish it from the scalar product, it is written with the multiplication sign $\times$:

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} \ .$$

In the direct product only elements of both matrices having both indices identical are multiplied:

$$\mathrm{c}_{ij} = \mathrm{a}_{ij}\mathrm{b}_{ij} \ .$$

It is the same as if both matrices were nm dimensional diagonal vectors and components of their scalar product were found:

$$\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -3 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -6 \\ 10 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Similarly can be explained the matrix addition. Both matrices are decomposed into nm dimensional diagonal vectors and the sums found:

$$\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 7 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

3.7 Normalization of matrices

We have discussed the problem of simultaneous action of more vectors or vectors having another intensity than 1. This can be sometimes solved by normalization of vectors. The aim of this shaping of vectors and matrices is to make them comparable. The normalization of vectors is done by eigenvectors, which must give the unit diagonal matrix $\mathbf{I}$. We introduced unit vectors $\mathbf{e}_j$. A row vector is comparable with the unit vector if it has the same length. The Euclidean length is the criterion, therefore a vector is normalized if its elements are divided by the square root of its Euclidean length. For example, the vector $(2,1,1,0)^{\mathrm{T}}$ is normalized by dividing it with $\sqrt{6}$. The length of its scalar product is then 1. A matrix vector is normalized by multiplying it by square root diagonal matrices from both sides. Here we have two possibilities. Either we normalize only diagonal elements or all rows and columns. For the normalization, the matrix must be symmetrical.

By normalization of diagonal elements, the matrix vector is oriented in direction of the unit vector $\mathbf{I}$. This has some consequences on properties of such normalized matrices.

3.8 Matrix Roots

We defined scalar products and quadratic forms of vectors and matrix vectors. Now we formulate the problem backwards: a matrix $\mathbf{M}$ has roots if it can be decomposed into a product of transposed matrices. For example, the unit diagonal matrix has many roots:

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

The unit diagonal matrix forms root to itself, since we can not distinguish forms

$$\mathbf{I} = \mathbf{I}^2 = \mathbf{I}^{-1} = \mathbf{I}^{\mathrm{T}}\mathbf{I} = \mathbf{I}\mathbf{I}^{\mathrm{T}} \,. \tag{3.8}$$

Its roots are symmetrical permutation matrices and asymmetrical permutation matrices. Moreover there are matrices with negative signs, since $(-1) \times (-1) = 1$. Our effort to find the natural space is somewhat complicated by this fact but we already introduced complex numbers and so

we can find the root even for the matrix of the last example[1]. It is simultaneously the fourth root from $\mathbf{I}_3$. Then there are eigenvectors to all nonsingular matrices. Our efforts to generate space by an supervised way is going out of control.

[1] The roots of permutation matrices can be compared to quarks in physics: Elementary particles are split into their components.

Chapter 4

Partitions

4.1 Preliminary Notes

Partitions of a natural number m into n parts were introduced into mathematics by Euler. The analytical formula for finding the number of partitions was derived by Ramanudjan and Hardy [13]. Ramanudjan was a mathematical genius from India. He was sure that it was possible to calculate the number of partitions exactly for any number m. He found the solution in cooperation with his tutor, the English mathematician Hardy. It is rather complicated formula derived by higher mathematical techniques. We will use only simple recursive methods for different relations between partitions.

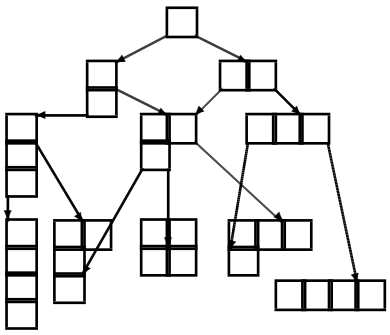
Steve Weinberg in his lecture [13] about importance of mathematics for physics mentioned that partitions got importance for theoretical physics, even if Hardy did not want to study practical problems. But partitions were used in physics before Hardy by Boltzmann [2]. He used this notion for splitting m quanta of energy between n particles in connection with his notion of entropy. He called partitions complexions, considering them to be orbits in phase space. His idea was forgoten.

A partition splits a number m into n parts which sum is equal to the number m, say $7: \quad 3, 2, 1, 1$. A partition is an ordered set. Its objects, *parts*, are written in a row in decreasing order:

$$m_{j-1} \geq m_j \geq m_{j+1} \ .$$

If we close a string of parts into brackets, we get a n dimensional vector row $p = (3, 2, 1, 1)$. From a partition vector, another vectors having equivalent structure of elements, for example. $r = (1, 2, 1, 3)$, are obtained by permuting, simple changing of ordering of vector elements. The partitions

Figure 4.1: Ferrers graphs construction. New boxes are added to free places.



are thus indispensable for obtaining combinatorial identities, for ordering points of plane simplices having constant sums of its constituting vectors.

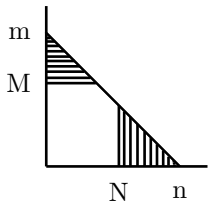
All unit permutations of a vector have the same length. Therefore different partitions form bases for other vectors composed from the same parts. Vectors belonging to the same partition of p into three parts are connected with other points of the three dimensional simplex by circles. In higher dimensions the circles become spheres and therefore we will call an ordered partition the *partition orbit* or simply orbit.

The number of vectors in partitions will be given as n, the size of the first vector as m_1. The bracket (m, n) means all partitions of the number m into at most n parts. Because we write a partition as a n dimensional vector we allow zero parts in a partition to fill empty places of the vector. It is a certain innovation against the tradition which will be very useful. But it is necessary to distinguish strictly both kinds of partitions, with zeroes and without them.

4.2 Ferrers Graphs

Ferrers graphs are used in the theory of partitions for many proofs based simply on their properties. Ferrers graphs are tables (see Fig. 4.1) containing m objects, each object in its own box. The square boxes are arranged into columns in nonincreasing order $m_j \geq m_{j+1}$ with the sum

Figure 4.2: Truncation of partitions by restrictions of rows and columns.



$$\sum_{j=1}^{n} m_j = \sum_{k=0}^{\infty} n_k m_k = m \; . \tag{4.1}$$

If partitions contain equal parts, it is possible to count them together using the index k and their number n_k.

It is obvious that a Ferrers graph, completed to an quadrangle with zero positions, is a matrix $\mathbf{F}$ which has its unit elements arranged consecutively in the initial rows and columns.

Introducing Ferrers graphs as matrices, we come necessarily to the notion of *restricted partitions*. The parts of a partitions can not be greater than the number of rows of the matrix and the number of parts greater than the number of its columns.

The interpretation of restrictions is geometrical. The part m_{max} determines the side of a cube, n is its dimension, see Fig. 4.2.

A sophistication of the notation distinguishes the partitioned number M and the number of rows m in the matrix $\mathbf{F}$. The unrestricted number of partitions p(M) is equal to the number of restricted partitions when restricting conditions are loose then $m \geq M$ and $n \geq M$:

$$p(M)_{unrestricted} = p(M, M, M) \; . \tag{4.2}$$

We write here first the number of rows m, then the number of parts n, here considered as equal to m and at last the sum of unit elements (the number of filled boxes) M.

An important property of restricted partitions is determined by transposing Ferrers graphs $\mathbf{F} \rightarrow \mathbf{F}^{\mathrm{T}}$:

$$p(m, n, M) = p(n, m, M) \; . \tag{4.3}$$

The partitions are conjugated. The number of partitions into exactly n parts with the greatest part m is the same as the number of partitions into m parts having the greatest part n. This is simple transposing of $\mathbf{F}$.

A Ferrers graph can be subtracted from the matrix containing only unit elements (defined as $\mathbf{J}\mathbf{J}^{\mathrm{T}}$, $\mathbf{J}$ being the unit column), and the resulting matrix transversed (Tr), For example:

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

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{\mathrm{Tr}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

The relation between the number of restricted partitions of two different numbers is obtained according to the following equation

$$p(m,n,M) = p(n,m,mn-M)\ . \tag{4.4}$$

This identity was derived by an operation very useful for acquiring elements of partition schemes (see later) and restricted partitions of all kinds. A restricted partition into exactly n parts, having m as the greatest part, has $(m+n-1)$ units bounded by elements forming the first row and column of the corresponding Ferrers graph (Fig. 4.1). Only $(M-m-n+1)$ elements are free for partitions in the restricted frame $(m-1)$ and $(n-1)$. Therefore:

$$p(m,n,M) = p(m-1,n-1,M-m-n+1)\ . \tag{4.5}$$

For example: p(4,3,8) = p(3,2,2) = 2. The corresponding partitions are 4,3,1 and 4,2,2; or 2,0 and 1,1; respectively. This formula can be used for finding all restricted partitions.

It is rather easy when the difference $(M-m-n+1)$ is smaller than the restricting values m and n or at least one from the restricting values. The row and column sums of partially restricted partitions having the other constrain constant (shown as an asterics), where either n or m can be 1 till M are:

$$p(m,*,M) = \sum_{j=1}^{M} p(m,j,M) \tag{4.6}$$

$$p(*,n,M) = \sum_{i=1}^{M} p(i,n,M)\ . \tag{4.7}$$

Before we examine restricted partitions in more detail, tables of unrestricted and partially restricted partitions will be introduced.

4.3 Partition Matrices

Partially restricted partitions can be obtained from unrestricted partitions by subtracting a row of n units or a column of m units. This gives us the recursive formula for the number of partitions as a sum of two partitions

$$p(*, N, M) = p(*, N-1, M-1) + p(*, N, M-N-1) \; . \tag{4.8}$$

All partitions into exactly N parts are divided into two sets. In one set are partitions having in the last column 1, their number is counted by the term $p(*, N-1, M-1)$ which is the number of partitions of the number $(M-1)$ into exactly $(N-1)$ parts to which 1 was added on the n th place and in other set are partitions which have in the last column 2 and more. They were obtained by adding the unit row $\mathbf{J}^{\mathrm{T}}$ with n unit elements to the partitions of $(M-N)$ into N parts. Their number can be found in the same column column n places above.

A similar formula can be deduced for partitions of M into at most N parts. These partitions can have zero at least in the last column or they are partitioned into n parts exactly:

$$p(*, * = N, M) = p(*, * = N-1, M) + p(*, * = N, M-N) \; . \tag{4.9}$$

The term p(*, *=N-1, M) are partitions of M into $(N-1)$ parts transformed in partitions into N parts by adding zero in the n-th column, the term p(*, *=N, M-N) are partitions of $(M-1)$ into N parts to which the unit row was added.

To formulate both recursive formulas more precisely, we had to define an apparently paradoxical partitionat first:

$$p(0, 0, 0) = 1 \; .$$

What it means? A partition of zero into zero number of parts. This partition represents the empty space of dimension zero. This partition is justified by its limit. We write $n = 0^0$ and find the limit:

$$\lim 0^0 = \lim_{x \to \infty} (1/x)^0 = 1/x^0 = 1 \; . \tag{4.10}$$

We get two following tables of partitions

Table 4.2 is obtained from the Table 4.1 as partial sums of its rows, it means, by multiplying with the unit triangular matrix $\mathbf{T}^T$ from the right. The elements of the matrix $\mathbf{T}^{\mathrm{T}}$ are

$$h_{ij} = 1 \text{ if } j \geq i \;\; h_{ij} = 0 \text{ if } j > i \; . \tag{4.11}$$

Table 4.1: Partitions into exactly n parts

n	0	1	2	3	4	5	6	Σ
m=0	1							1
1		1						1
2		1	1					2
3		1	1	1				3
4		1	2	1	1			5
5		1	2	2	1	1		7
6		1	3	3	2	1	1	11

Table 4.2: Partitions into at most n parts

n	0	1	2	3	4	5	6
m=0	1	1	1	1	1	1	1
1		1	1	1	1	1	1
2		1	2	2	2	2	2
3		1	2	3	3	3	3
4		1	3	4	5	5	5
5		1	3	5	6	7	7
6		1	4	7	9	10	11

On the other hand, the Table 4.2 is obtained from the Table 4.1 by multiplying with a matrix $\mathbf{T}^{-\mathrm{T}}$ from the right. The inverse elements are

$$h^{-1}_{ii} = 1 \ , \ h^{-1}_{i,i+1} = -1 \ , \ h_{ij} = 0 \ , \text{otherwise} \ . \tag{4.12}$$

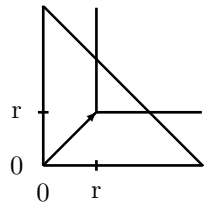
Notice, that the elements of the Table 4.2 right of the diagonal remain constant. They are equal to the row sums of the Table 4.1. Increasing the number of zeroes does not change the number of partitions.

When we multiply Table 4.1 by the matrix $\mathbf{T}^{-\mathrm{T}}$ again, we obtain partitions having as the smallest allowed part the number 2. The effect of these operators can be visualized on the 2 dimensional complex, the operators shift the border of counted orbits (Fig. 4.3). The operator $\mathbf{T}^{\mathrm{T}}$ differentiates n dimensional complexes, shifting their border to positive numbers and cutting lover numbers. Zero forms the natural base border.

4.4 Partitions with Negative Parts

Operations with tables of partitions lead to a thought, what would happen with partitions outside the positive cone of nonnegative numbers. Thus let

Figure 4.3: Limiting of partition orbits. The lowest allowed part r shifts the plane simplex.



us allow the existence of negative numbers in partitions, too[1].

If the number of equal parts n_k is written as the vector row under the vector formed by the number scale, the number of partitions is independent on shifts of the number scale, see Table 4.3. Partitions, shown in the bottom part of the table, are always derived by shifting two vectors, one 1 position up, the other 1 position down. Each partition corresponds to a vector. If we write them as columns then their scalar product with the number scale, forming the vector row $\mathbf{m}^{\mathrm{T}}$, gives constant sum:

$$\mathbf{m}^{\mathrm{T}}\mathbf{p} = \sum_{k \geq r} m_k n_k = m \; . \tag{4.13}$$

There is an inconsistency in notation, elements of the vector $\mathbf{p}$ are numbers of vectors having the same length and the letter n with an index k is used for them. For values of the number scale the letter m is used with the common index k which goes from the lowest allowed value of parts r till the highest possible value. The index k runs to infinity but all too high values n_k are zeroes.

Using different partition vectors and different vectors $\mathbf{m}$ we get the following examples:

$$\begin{aligned} (4 \times -2) + (1 \times 3) &= -5 \\ (3 \times -1) + (1 \times 0) + (1 \times 3) &= \; 0 \\ (3 \times 0) + (1 \times 2) + (1 \times 3) &= \; 5 \end{aligned}$$

[1]The negative parts can be compared in physics with antiparticles. Since an annihilation liberates energy, it does not annihilate it, the energy of the Universe is infinite. Speculations about existence of antiworlds, formed only by antiparticles balancing our world, can be formulated as doubt if the Universe is based in the natural cone of the space.

Table 4.3: Partitions as vectors

Parameter	r							
Vector **m**	-2	-1	0	1	2	3	**mp** =	-5
	-1	0	1	2	3	4		0
	0	1	2	3	4	5		5
	1	2	3	4	5	6		10
	2	3	4	5	6	7		15
Vector p	4					1		
	3	1			1			
	3		1	1				
	2	2		1				
	2	1	2					
	1	3	1					
	1	2	2					
		5						

$$(2 \times 1) + (1 \times 2) + (2 \times 3) = 10$$
$$(1 \times 2) + (3 \times 3) + (1 \times 4) = 15.$$

The parameter r shifts the table of partitions, its front rotates around the zero point. If r were $-\infty$, then $p(-\infty, 1) = 1$ but $p(-\infty, 2)$ were undetermined, because a sum of a finite number with an infinite number is again infinite. The parameter r will be written to a partition as its upper index to show that different bases of partitions are differentiating plane simplices.

4.5 Partitions with Inner Restrictions

Partitions were classified according to the minimal and maximal allowed values of parts, but there can be restrictions inside the number scale, it can be prescribed that some values are be forbidden. It is easy to see what this means:

The plane simplex has holes, some orbits cannot be realized and its $(n-1)$ dimensional body is thinner than the normal one. It is also possible to arrange partitions in a plane in nonincreasing order.

It is easy to find the number of partitions in which all parts are even. It is not possible to form an even partition from an uneven number, therefore:

$$p_{\text{even}}(2n) = p_{\text{unrestricted}}(n) \; . \tag{4.14}$$

Table 4.4: Odd, even, and mixed partitions

	Number of odd partitions									Sums			
n	1	2	3	4	5	6	7	8	9	Odd	Even	Mixed	p(m)
m=1	1									1	0	0	1
2		1								1	1	0	2
3	1		1							2	0	1	3
4		1		1						2	2	1	5
5	1		1		1					3	0	4	7
6		2		1		1				4	3	4	11
7	1		2		1		1			5	0	10	15
8		2		2		1		1		6	5	11	22
9	1		3		2		1		1	8	0	22	30

A more difficult task is finding the number of partitions in which all parts are odd. The rejected partitions contain mixed odd and even parts. The relation between different partitions is etermined as

$$p_{\text{unrestricted}}(n) = p_{\text{odd}}(n) + p_{\text{even}}(n) + p_{\text{mixed}}(n) \; . \tag{4.15}$$

The corresponding lists are given in Table 4.4

Notice how the scarce matrix of odd partitions is made from Table 4.1. Its elements, except the first one in each column, are shifted down on cross diagonals. An odd number must be partitioned into an odd number of odd parts and an even number into even number of odd parts. Therefore the matrix can be filled only in half. The recurrence is given by two possibilities how to increase the number m. Either we add odd 1 to odd partitions of $(m-1)$ with exactly $(j-1)$ parts or we add $2j$ to odd numbers of partitions of $(m-2j)$ with exactly j parts. The relation is expressed as

$$o(i,j) = p[(i+j)/2, j] \; . \tag{4.16}$$

Partitions with all parts unequal are important, because their transposed Ferrers graphs have the greatest part odd, when the number of parts is odd, and even, when the number of parts is even. For example

Table 4.5: Partitions with unequal parts

n	1	2	3	4	Σ	Difference ($n_{odd} - n_{even}$)
m=1	1				1	1
2	1				1	1
3	1	1			2	0
4	1	1			2	0
5	1	2			3	-1
6	1	2	1		4	0
7	1	3	1		5	-1
8	1	3	2		6	0
9	1	4	3		8	0
10	1	4	4	1	10	0
11	1	5	5	1	12	0
12	1	5	7	2	15	1

10

9,1

8,2

7,3 7,2,1

6,3 6,3,1

5,4,1

5,3,2

4,3,2,1

The partitions with unequal parts can be tabulated as in Table 4.5. Notice that the difference of the even and odd columns partitions is mostly zeroes and only sometimes ±1. The importance of this phenomenon will be explained later. The number of partitions with unequal parts coincide with the partitions which all parts are odd.

The differences are due to Franklin blocks with growing minimal parts and growing number of parts (their transposed notation is used), which are minimal in that sense that their parts differ by one, the shape of corresponding Ferrers graphs is trapeze:

Table 4.6: Partitions Differentiated According to Unit Parts

n	0	1	2	3	4	5	6
m=0	1						
1	0	1					
2	1	0	1				
3	1	1	0	1			
4	2	1	1	0	1		
5	2	2	1	1	0	1	
6	4	2	2	1	1	0	1

(1)	(11)	$1, 2$
$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$	$5, 7$
$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$	$12, 15$

4.6 Differences According to Unit parts

We have arranged restricted partitions according to the number of nonzero parts in Table 1. It is possible to classify partitions according the number of vectors in the partition having any value. Using value 1, we get another kind of partition differences as in Table 4.6.

The elements of the table are:

$$p_{i0} = p(i) - p(i-1),\ p_{ij} = p_{i-1,j-1}\ , \text{otherwise}\ . \tag{4.17}$$

Table 4.6 is obtained from the following Table 4.7 of rows of unrestricted partitions by multiplying it with the matrix $\mathbf{T}^{-1}$. The zero column of the Table 4.6 is the difference of two consecutive unrestricted partitions according to m. To all partitions of $p(m-k)$ were added k ones. The partitions in the zero column contain only numbers greater than 1. These partitions can not be formed from lower partitions by adding ones and they are thus a difference of the partition function according to the number n_1. Since Table 4.6 is composed, it is the product of two matrices, its inverse is composed, too.

Table 4.7: Partitions and their Euler inversion

	Partition table						Euler inversion					
j	0	1	2	3	4	5	0	1	2	3	4	5
i=0	1						1					
1	1	1					-1	1				
2	2	1	1				-1	-1	1			
3	3	2	1	1			0	-1	-1	1		
4	5	3	2	1	1		0	0	-1	-1	1	
5	7	5	3	2	1	1	1	0	0	-1	-1	1

4.7 Euler Inverse of Partitions

If we write successive partitions as column or row vectors as in Table 7, which elements are

$$p_{ij} = p(i - j + 1) \, , \tag{4.18}$$

we find rather easily its inverse matrix which is given in the second part of the same Table.

The nonzero elements in the first column of the Euler inversion (and similarly in the next columns which are only shifted down one row) appear at indices, which can be expressed by the Euler identity concerning the coefficients of expansion of

$$(1-t)(1-t^2)(1-t^3)... = 1 + \sum_{i=1}^{\infty} (-1)^i \left[t^{3i^2-i)/2} + t^{3i^2+i)/2} \right] . \tag{4.19}$$

For example: the last row of the partition Table 4.7 is eliminated by multiplying it with the Euler inversion as:

$$(7 \times 1) + (5 \times -1) + (3 \times -1) + (2 \times 0) + (1 \times 0) + (1 \times 1) = 0$$

when $i = 1$, there is the pair of indexes at t 1, 2 with negative sign; for $i = 2$ the pair is 5, 7; for $i = 3$ the pair is $-12, -15$ and so on. These numbers are the distances from the base partition. The inverse matrix becomes scarcer as $p(m)$ increases, as it was already shown in Franklin partitions above. All inverse elements are $-1, 0, 1$. The nonzero elements of the Euler polynomial are obtained as sums of the product

$$\prod_{i=1}^{\infty} (1 - t^i) \, . \tag{4.20}$$

This is verified by multiplying several terms of the infinite product. If we multiply the Euler polynomial with its inverse function

$$\prod_i = 1^{\infty}(1-t^i)^{-1}\ , \tag{4.21}$$

we obtain 1. From this relation follows that partitions are generated by the inverse Euler function which is the *generating function* of partitions. Terms t^i must be considered as representing unequal parts.

The Euler function has all parts t^i different. We have constructed such partitions in Table 4.5. If the coefficient at t^i is obtained as the product of the even number of $(1-t^i)$ terms then the sign is positive, and if it is the result of the uneven number of terms then the sign is negative. The coefficients are determined by the difference of the number of partitions with odd and even number of unequal parts. This difference can be further explained according to Franklin using Ferrers graphs.

All parts in $p(n)$ having as at least one part equal to 1 are obtained from $p(n-1)$. The difference $p(n)-p(n-1)$ is due to some terms of $p(n-2)$. We must add 2 to each partition of $p(n-2)$, except all partitions of $p(n-2)$ containing 1. These must be either removed or used in transposed form using transposed Ferrers graphs, since big parts are needed. One from the pair of conjugate partitions is superfluous. These unused partitions must be subtracted. For example: for $p(8)$:

6;	1^6;	*Formed* :	8;	62;
$\underline{51}$;	21^4;		53;	
42;	2^21^2;		44;	2^4;
$\underline{33}$;	2^3;		3^22;	
41^2;	$\underline{31^3}$;		42^2;	
$\underline{321}$;				

Leftovers (underlined above):

$$\text{p(1) + 5: 51;} \qquad \text{p(3) + 3: 33; 321; } 31^3$$

are obtained by subtracting the largest part from corresponding partition. Two must be added to the subtracted part. We get p(8-5) and p(8-7) as the corrections.

4.8 Other Inverse Functions of Partitions

We already met other tables of partitions which have inverses because they are in lower triangular form. The inverse to the Table 4.1 is Table 4.8.

Table 4.8: Inverse matrix to partitions into n parts

n	1	2	3	4	5	6
m=1	1					
2	-1	1				
3	0	-1	1			
4	1	-1	-1	1		
5	0	1	-1	-1	1	
6	0	1	0	-1	-1	1

Table 4.9: Inverse matrix of unit differences

n	1	2	3	4	5	6
m=1	1					
2	0	1				
3	-1	0	1			
4	-1	-1	0	1		
5	-1	-1	-1	0	1	
6	0	-1	-1	-1	0	1

The inverse to Table 4.6 is Table 4.9.

Whereas the columns of the Table 4.8 are irregular and elements of each column must be found separately, columns of the Table 4.9 repeat as they are only shifted in each column one row down, similarly as the elements of their parent matrix are. They can be easily found by multiplying the matrix of the Euler function (Table 4.7) by the matrix **T** from the left.

4.9 Partition Orbits in m Dimensional Cubes

Restricted partitions have a geometric interpretation: They are orbits of n dimensional plane complices truncated into cubes with the sides $(m-1)$ as on Fig. 3.

We can count orbits even in cubes. It is a tedious task if some special techniques are not applied, since their number depends on the size of the cube. For example: for the 3 dimensional space we get orbits as in Table 4.10.

The Equation 3 can be applied for cubes. It shows their important property, they are symmetrical along the main diagonal, going from the center of the coordinates, the simplex n^0 to the most distant vertex of the cube in which all n coordinates are $(m-1)$. The diagonal of the cube is represented on Table 4.10 by k indices. Moreover, a cube is convex,

Table 4.10: Orbits in 3 dimensional cubes

Edge size	0	1	2	3
m=0	000	000	000	000
1		100	100	100
2		110	200; 110	210; 110
3		111	210; 111	300; 210; 111
4			220; 211	310; 220; 211
5			221	320; 311; 221
6			222	330; 321; 222
7				331; 322
8				332
9				333

therefore

$$M \leq mn/2 \text{ then } p(m,n,M) \geq p(m,n,M-1) \tag{4.22}$$

and if

$$M \geq mn/2 \text{ then } p(m,n,M) \leq p(m,n,M-1)\ . \tag{4.23}$$

Here we see the importance of restricted partitions. From Table 10, we find the recurrence, which is given by the fact that in a greater cube the lesser cube is always present as its base. New orbits which are on its enlarged sides are added to it. But it is enough to know orbits of one enlarged side, because the other sides are formed by these orbits. The enlarged side of a n dimensional cube is a $(n-1)$ dimensional cube. The recurrence relation for partitions in cubes is thus

$$p(m,n,M) = p(m-1,n,M) + p(m,n-1,M)\ . \tag{4.24}$$

This recurrence will be explained later more thoroughly.

4.10 Generating Functions of Partitions in Cubes

The generating function of partitions is simply the generating function of the infinite cube in the Hilbert space, which sides have different meshes:

$$\text{Parts } 1 : (1 + t_1^1 + t_1^2 + \dots\ t_1^\infty) \tag{4.25}$$

$$\text{Parts } 2 : (1 + t_2^1 + t_2^2 + \dots\ t_2^\infty) \tag{4.26}$$

and so on till

$$\text{Parts } \infty : \ (1 + \ldots \ t_{\infty}^{1}) \ . \tag{4.27}$$

When the multiplication's for all parts are made and terms on consecutive plane simplices counted, we get:

$$1 + t_1^1 + [t_2^1 + t_1^2] + [t_3^1 + \ldots \ . \tag{4.28}$$

The generating function of restricted partitions is obtained by canceling unwanted (restricted) parts. Sometimes the generating function is formulated in an inverse form. The infinite power series are replaced by the differences $(1 - t_k^{-1})$. This is possible if we consider t to be only a dummy variable. For example, the generating function of the partitions with unequal unrepeated parts is given by the product

$$u(t) = \prod_{k=1}^{\infty} (1 - t_k) \ . \tag{4.29}$$

The mesh of the partition space is regular, it covers all numbers. The number of partitions is obtained by recursive techniques. But it is a very complicated function, if it is expressed in one closed formula, as the Ramanudjan-Hardy function is. The partitions form a carcass of the space. We will be interested, how the mesh of partitions is filled into the space which all axes have unit mesh and which contains also vector strings.

Chapter 5

Lattices of Orbits

5.1 Partition Schemes

Multidimensional plane simplices are complicated objects and it is necessary to find tools how to analyze them. To draw them is impossible, as it was mentioned, because their parts are layered in our 3 dimensional world over themselves.

We already classified orbits in plane simplices according to the number k of nonzero parts. This number shows the dimensionality of subsimplices, their vertices, edges, and (k-1) dimensional bodies. Lately we introduced the number of unit vectors as a tool differentiating the simplex. Now we arrange partitions as two dimensional tables. These tables will be called *partition schemes*.

Analyzing a 7 dimensional plane simplex with $m = 7$, we can start with its 3 dimensional subsimplices. We see that they contain points corresponding to partitions: 7,0,0; 6,1,0; 5,2,0; 4,3,0; 5,1,1; 4,2,1; 3,3,1; 3,2,2. The points corresponding to partitions are connected with other points of the simplex by circles. In higher dimensions the circles become spheres and this is the reason why we call a partition an *orbit*. The other points on each orbit have only different ordering of the same set of the coordinates.

Arranging partitions into tables (Table 5.1), the column classification is made according to the number of nonzero parts of partitions. Another classifying criterion is needed for rows. This will be the length of the longest vector m_1. From all partition vectors having the same dimensionality the longest vector is that one with the longest first vector. It dominates them. But there can exist longer orbits nearer to the surface of the simplex with a lesser number of nonzero parts. For example, vector (4,1,1) has equal

Figure 5.1: Lattice of partition orbits (7,7)

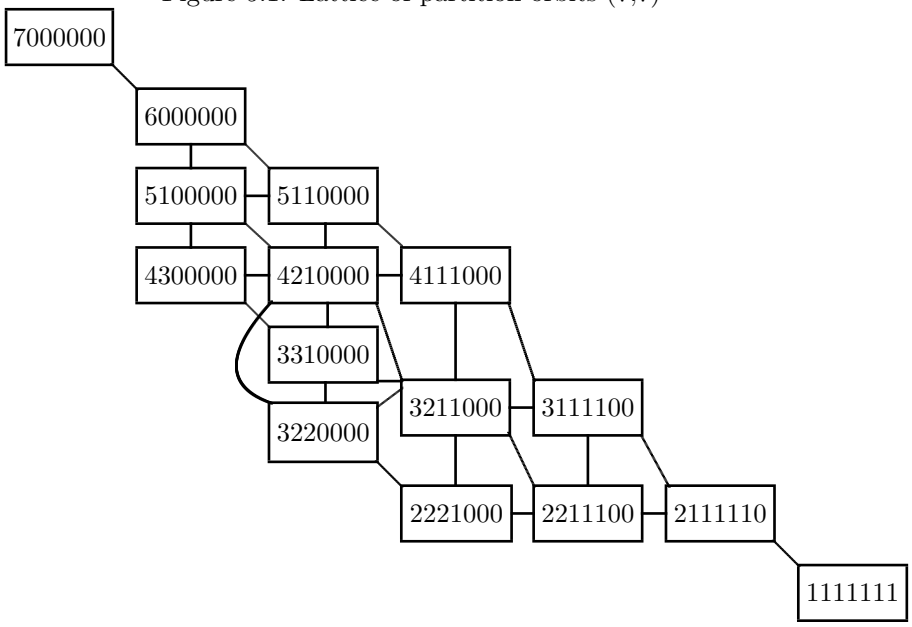


Table 5.1: Partition scheme (7,7)

n	1	2	3	4	5	6	7	Σ
m = 7	1							1
6		1						1
5		1	1					2
4		1	1	1				3
3			2	1	1			4
2				1	1	1		3
1							1	1
Σ	1	3	4	3	2	1	1	11

length as (3,3,0) but vector (4,1,1,1,1) is shorter than (3,3,2,0,0). Such an arrangement is on Table 5.1. Orbits with three nonzero parts lie inside the 3 dimensional simplex, with two nonzero parts lie on its edges. Orbits with four nonzero parts are inside tetrahedrons, it is on a surface in the fourth dimension. There exist these partitions: 4,1,1,1; 3,2,1,1; 2,2,2,1. Similarly columns corresponding to higher dimensions are filled.

The rows of partition schemes classify partitions according to the length of the first and longest vector $\mathbf{e}_1$. It can be shown easily that all vectors in higher rows are longer than vectors in lover rows in corresponding columns. In the worst case it is given by the difference

$$(x+1)^2 + (x-1)^2 > (2x)^2 \,. \tag{5.1}$$

A three dimensional plane simplex to be can be considered as a truncated 7 dimensional simplex, and after completing the columns of the Tab. 5.1) by the corresponding partitions, we get a crossection through the 7 dimensional plane. The analysis is not perfect, an element is formed by two orbits, but nevertheless the scheme gives an insight how such high dimensional space looks like. We will study therefore properties of partitions schemes thoroughly.

The number of nonzero vectors in partitions will be given as n, the size of the first vector as m. Zeroes will not be written to spare work. The bracket (m, n) means all partitions of the number m into at most n parts. Because we write a partition as a vector, we allow zero parts to complete the partition as before.

5.2 Construction of Partition Schemes

A partition scheme is divided into four blocks. Diagonal blocks repeat the Table 4.1 (the left upper block), the right lower one is written in the transposed form for $n > m/2$. Odd and even schemes behave somewhat differently, as can be seen on Tables 5.2 and 5.3.

In the left lower block nonzero elements indicated by asterisks * can be placed only over the line which gives sufficient great product mn to place all units into the corresponding Ferrers graphs and their sums must agree not only with row and column sums, but with diagonal sums, as we show below. This can be used for calculations of their numbers, together with rules for restricted partitions.

The examples show three important properties of partition schemes:

- Partition schemes are symmetrical according to their transversals, due to the conjugated partitions obtained by transposing Ferrers graphs.

Table 5.2: Partition scheme m = 13

n	1	2	3	4	5	6	7	8	9	10	11	12	13
m=13	1												
12		1											
11		1	1										
10		1	1	1									
9		1	2	1	1								
8		1	2	2	1	1							
7		1	3	3	2	1	1						
6			3	4	3	2	1	1					
5			2	4	5	3	2	1	1				
4				3	4	4	3	2	1	1			
3					2	3	3	2	2	1	1		
2							1	1	1	1	1	1	
1													1
Σ	1	6	14	18	18	14	11	7	5	3	2	1	1

Table 5.3: Partition scheme m = 14

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
m=14	1													
13		1												
12		1	1											
11		1	1	1										
10		1	2	1	1									
9		1	2	2	1	1								
8		1	3	3	2	1	1							
7		1	3	4	3	2	1	1						
6			3	*	*	*	2	1	1					
5			1	*	*	*	3	2	1	1				
4				3	*	*	4	3	2	1	1			
3					2	*	3	3	2	2	1	1		
2							1	1	1	1	1	1	1	
1														1
Σ	1	7	16	23	23	20	15	11	7	5	3	2	1	1

Table 5.4: Partition scheme (7,7) and its inversion

n	1	2	3	4	5	6	7	1	2	3	4	5	6	7
m = 7	1							1						
6		1							1					
5		1	1						0	1				
4		1	1	1					0	-1	1			
3			2	1	1				2	-1	-1	1		
2				1	1	1			-2	2	0	-1	1	
1							1		0	0	0	0	0	1

- The upper left quarter (transposed lower right quarter) contain elements of the Table 4.1 of partitions into exactly n parts shifted one column up.

- The schemes have form of the matrix in the lower diagonal form with unit diagonal. Therefore, they have inverses. It is easy to find them, for example for $n = 7$ (Table 5.4).

The partitions in rows must be balanced by other ones with elements of inverse columns. The third column includes or excludes 331 and 322 with 3211 and 31^4; 2^31 and 2^21^3 with 2×21^5, respectively.

5.3 Lattices of Orbits

Partition orbit is a sphere which radius r is determined by the Euclidean length of the corresponding vector: $r = (\sum p_j^2)$. Radiuses of some partition orbits coincide, For example: $r(3,3,0)^2 = r(4,1,1)^2 = (18)$. It is thus impossible to determine distances between orbits using these radii (Euclidean distances) since the distance between two different orbits cannot be zero.

We have shown in Sect. 4.4 that one orbit can be obtained from another by shifting just two vectors, one up and other down on the number scale. We can imagine that both vectors collide and exchange their values as two particles of the ideal gas exchange their energy. If we limit the result of such an exchange to 1 unit, we can consider such two orbits to be the nearest neighbor orbits. The distance inside this pair is $\sqrt{2}$. We connect them in the scheme by a line. Some orbits are thus connected with many neighbor orbits, other have just one neighbor, compare with Fig. 5.1. Orbits (3,3,0) and (4,1,1) are not nearest neighbors, because they must be transformed in two steps:

Table 5.5: Right hand One-unit Neighbors of Partition Orbits

n	1	2	3	4	5	6	Σ
m=2	1						1
3	1	1					2
4	1	2	1				4
5	1	3	2	1			7
6	1	4	4	2	1		12
7	1	5	6	4	2	1	19
D(7-6)	0	1	2	2	1	1	7

$$(3,3,0) \leftrightarrow ((3,2,1) \leftrightarrow (4,1,1)$$

or

$$(3,3,0) \leftrightarrow (4,2,0) \leftrightarrow (4,1,1)\ .$$

Partition schemes are generally not suitable for construction of orbit lattices, because at $m = n > 7$ there appear several orbits on some table places. It is necessary to construct at least 3 dimensional lattices to show all existing connections. For example:

$$\begin{array}{ccccc} (5,2,1) & \leftrightarrow & (4,3,1) & \leftrightarrow & (3,3,2) \\ & \searrow\nwarrow & \updownarrow & \swarrow\nearrow & \\ & & (4,2,2) & & \end{array}$$

Sometimes stronger condition are given on processes going at exchanges, namely, that each collision must change the number of empty parts, as if they were information files which can be only joined into one file or one file separated into two or more files, or as if a part of a file transferred into an empty file. Here also the nearest neighbor is limited on unifying of just 2 files or splitting a file into two (Fig.5.2). In this case the path between two orbits must be longer, For example:

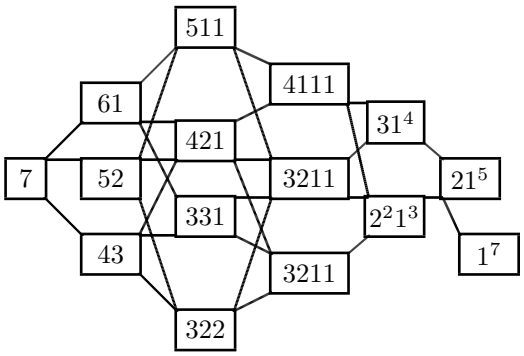
$$(3,3,0) \leftrightarrow (6,0,0) \leftrightarrow (4,2,0) \leftrightarrow (4,1,1)$$

or

$$(3,3,0) \leftrightarrow (3,2,1) \leftrightarrow (5,1,0) \leftrightarrow (4,1,1)\ .$$

In a lattice it is possible to count the number of nearest neighbors. If we investigate the number of one unit neighbors or connecting lines between columns of partition schemes, we obtain an interesting Table 5.5.

Figure 5.2: Lattice of file partitions. A file can be split into two new ones or two files can be combined into one.



The number of right hand neighbors is the sum of two terms. The right hand diagonal neighbors exist for all $p(m, n-1)$. We add 1 to all these partitions and decrease the largest part. Undetermined remain right hand neighbors in rows. Their number is equal to the number of partitions $p(m-2)$. To each partition $p(m-2, n-1)$ are added two units, one in the n th column, the second in the (n- 1) the column.

The number of right hand neighbors $P(n)$ is the sum of the number of unrestricted partitions

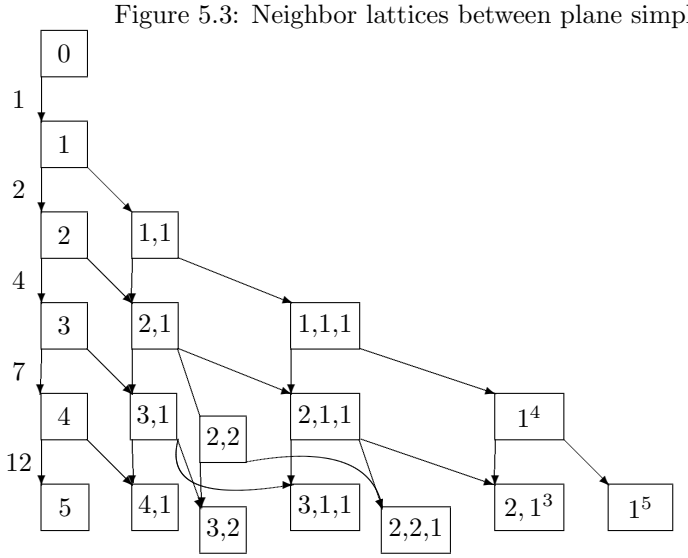
$$P(n) = \sum_{k=0}^{n-2} p(k) \; . \tag{5.2}$$

To find all neighbors, we must add neighbors inside columns. The number of elements in columns is the number of partitions into exactly n parts $p(m, n)$, the difference of each column must be decreased by 1 but there exist additional connections, see Fig. 5.3.

These connections must be counted separately. The resulting numbers are already known. The construction of partition schemes gives the result which we know as Table 4.1 read from the diagonal to the left.

The other interpretation of right hand one-unit neighbors of partitions is the plane complex as on Fig. 5.3. Vectors connect nearest neighbors in layers.

Figure 5.3: Neighbor lattices between plane simplices.



5.4 Diagonal Differences in Lattices

In lattices, we can count orbits on side diagonals going consecutively parallel to the main diagonal. They count orbits having the form $[n-k]^k$. Their Ferrers graphs have a L form

$$\begin{matrix} x & x & x & x \\ x & & & \\ x & & & \\ x & & & \end{matrix}$$

Side diagonal elements counts partitions which have in this layer smaller number of units, the other are inside this base.

The corresponding Table is 5.6.

The initial k column values have these analytical forms:

- 1n counts elements in n columns (rows) having the form $(n-k)1^k$, $k = 0$ – $(n-1)$;

- 1(n-3) counts elements in (n - 2) columns (rows) obtained from the basic partition 2,2 by adding units in the first row and column;

Table 5.6: Diagonal Sums of Partitions

k	1	2	3	4	5	6	7	8	9	Σ
n= 1	1									1
2	2									2
3	3									3
4	4	1								5
5	5	2								7
6	6	3	2							11
7	7	4	4							15
8	8	5	6	3						22
9	9	6	8	6	1					30
10	10	7	10	9	6					42
11	11	8	12	12	11	2				56
12	12	9	14	15	16	9	2			77
13	13	10	16	18	21	16	7			101
14	14	11	18	21	26	23	18	4		135
15	15	12	20	24	31	30	29	12	3	176

- 2(n-5) counts elements in (n - 2) columns (rows) obtained from the basic partitions 3,3 and 2,2,2 by adding units in the first row and column;

- 3(n-7) counts elements in (n - 2) columns (rows) obtained from the basic partitions 4,4; 3,3,2, and 2,2,2,2 by adding units in the first row and column;

- 5(n-9) + 1. On this level appears the partition 3,3,3 where elements start to occupy the third L layer;

- 7(n-11) + 2.

The values in brackets are the numbers for partitions which lie inside the L frame having $(2k-1)$ units. At higher diagonal layers appear these possibilities to add new elements later. Partitions 4, 4, 4 and 3, 3, 3, 3, for $n = 12$, are counted in the seventh layer. For $n = 13$, the layer counts seven partitions:

Table 5.7: Binomial Ordering of Partitions

1	2	3	4	5	Σ
(1)					1
(1,1)	(2)				2
(1,1,1)	(2,1)	(3)			
	(2,2)				4
(1,1,1,1)	(2,1,1)	(3,1)	(4)		
	(2,2,1)	(3,2)			
	(2,2,2)	(3,3)			8
(1,1,1,1,1)	(2,1,1,1)	(3,1,1)	(4,1)	(5)	
	(2,2,1,1)	(3,2,1)	(4,2)		
	(2,2,2,1)	(3,2,2)	(4,3)		
	(2,2,2,2)	(3,3,1)	(4,4)		
		(3,3,2)			
		(3,3,3)			16

$5, 5, 3;$
$5, 4, 4;$
$4, 4, 4, 1;$
$4, 4, 3, 2;$
$4, 3, 3, 3;$
$3, 3, 3, 3, 1;$
$3, 3, 3, 2, 1.$

There appears a very interesting property of partition lattices. The side diagonals being on side diagonals of the Table 5.6 have equal length n, and the number of partitions $p(d)$ lying on them is equal to

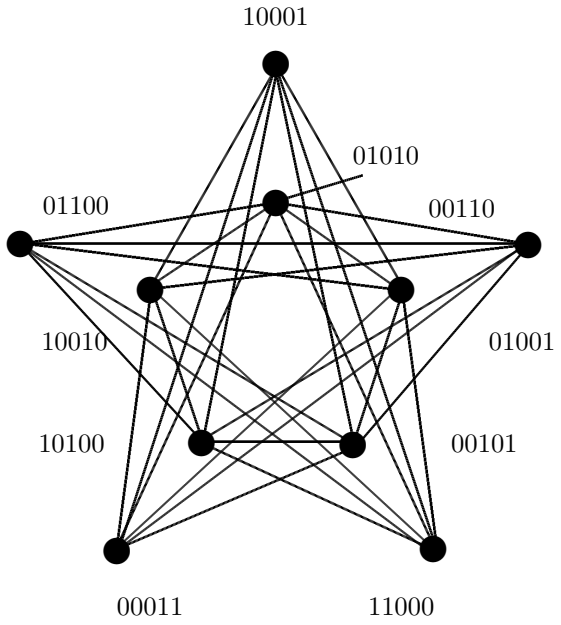
$$p(d) = 2^{(n-1)} \tag{5.3}$$

this is true for all complete diagonals in the table, also the seventh diagonal sum is completed by the partition (4,4,4,4). It can be conjectured, that it is a general property of lattices. There are counted partitions which superposed Ferrers graphs can be situated into isoscele triangular form ($M = N$) ending in the transversal which were not ounted before. The condition is that all Ferrers graphs are superposed from the same starting place, otherwise Ferrers graphs of each partition can fill their isoscele triangular form.

The partitions can be ordered in the following way (see Table 5.7).

Counting of partitions is changed into a combinatorial problem of finding of all ordered combinations of $(k-1)$ numbers with the greatest part equal

Figure 5.4: Nearest neighbors in 00111 lattice.

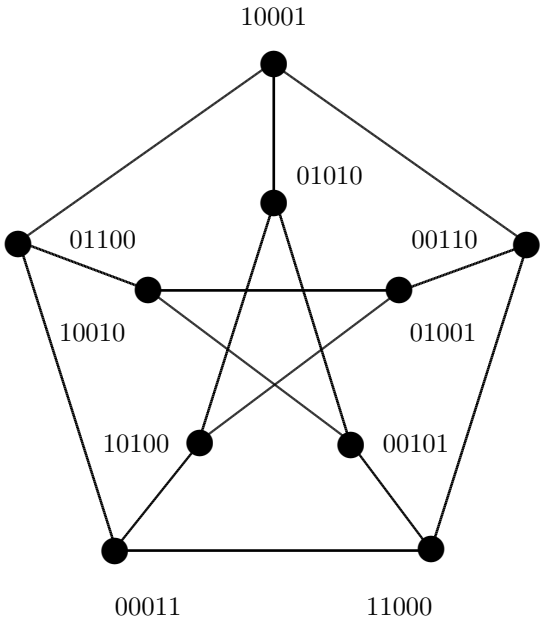


to k. The partitions are formed by a fixed part which is equal to the number of the column and starts in the corresponding row. To this fixed part are added two movable parts from the previous partitions, the whole upper predecessor and the movable part of the left upper predecessor. The resulting counts are the binomial numbers.

5.5 Generalized Lattices

The notion of lattices can be used also for possible transformations of points having specific properties among themselves, For example: between all 10 permutations of a 5 tuple composed from 3 symbols of one kind and 2 symbols of another kind. When the neighbors differ only by one exchange of the position of only anyone pair of two kinds symbols we obtain lattice as on Fig.5.4. Each from three unit symbols has two possibilities to change 0 into 1. Rearrange these ten points as a simple triangle. The simultaneous

Figure 5.5: Petersen graph. Adjacent vertices are in distances 4.



exchange of two pairs (or two consecutive changes of one pair give a pattern as on Fig.5.5, known as the Pettersen graph.

Lattices are formed by vertices of n dimensional cubes. The nearest vertices differ only by one coordinate. The lattices of the 3 dimensional cube is on Fig. 5.6. Compare lines of the graphs with a real 3 dimensional cube and try to imagine the 4 dimensional cube (Fig. 5.7).

A classical example of relation lattices is Aristotle's attribution of four properties: **warm**, **cold**, **dry**, and **humid** to four elements: fire, air, water and earth, respectively. It can be arranged in a form

$$\begin{array}{ccccc} air & & \textbf{humid} & & water \\ & \textbf{warm} & 0 & \textbf{cold} & \\ fire & & \textbf{dry} & & earth\ . \end{array}$$

Figure 5.6: Lattice of the three dimensional unit cube.

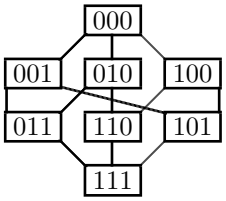
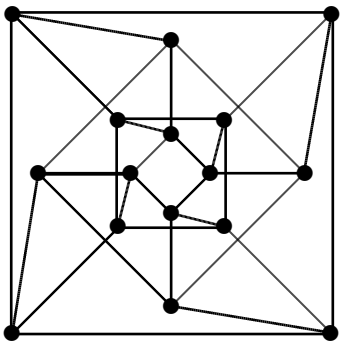


Figure 5.7: Four dimensional cube projection. One 3 dimensional cube is twisted 45^0.



The elements have always only two properties. The properties adjacent vertically and horizontally exclude themselves. Something can not be simultaneously warm and cold, or humid and dry[1].

[1]More precisely, it is necessary to draw a borderline (point zero) between these properties. Depending on its saturation, water vapor can be dry as well as wet.

Chapter 6

Erasthothenes Sieve and its Moebius Inversion

6.1 Divisors of m and Their Matrix

In this chapter an important notion will be introduced, the *divisor*. A number k is a divisor of the number m if $m \equiv 0 \mod k$, it means m is identical with 0 when divided with k. Or otherwise, $m = kn$, number m splits into n equal parts k. It follows that each number has at least two divisors, the number 1, which leaves the number unchanged and the number itself, when the division gives 1 as the result. If only these two divisors exist, such a number is called the *prime*.

It is possible to find prime numbers p by the Erasthothenes sieve. This algorithm works like a sieve. A number put on the first column of the sieve falls through its columns. If it gets the diagonal without meeting a divisor, it is a prime. The divisors j represented by units in divisor rows of corresponding columns work as meshes in a sieve. Thus the Erasthothenes sieve is a matrix which elements are

$$e_{ij} = 1 \ ,$$

if the number j is the divisor of the number i, and

$$e_{ij} = 0 \ ,$$

otherwise. On Table 6.1 is the Erasthothenes sieve with its Moebius inverse function.

The divisors form a regular pattern, they are in all rows $i \equiv 0 \mod j$. The prime numbers become scarcer, as the matrix grows, but it is always

Table 6.1: Erasthothenes sieve and its Moebius inversion

	Erasthothenes sieve							Moebius inversion						
j	1	2	3	4	5	6	7	1	2	3	4	5	6	7
i=1	1							1						
2	1	1						-1	1					
3	1	0	1					-1	0	1				
4	1	1	0	1				0	-1	0	1			
5	1	0	0	0	1			-1	0	0	0	1		
6	1	1	1	0	0	1		1	-1	-1	0	0	1	
7	1	0	0	0	0	0	1	-1	0	0	0	0	0	1

possible to find another prime number $p(n+1)$ as the product of all n previous prime numbers increased by 1.

$$p(n+1) = \prod_{j=1}^{n} p_j + 1 \; . \tag{6.1}$$

This equation does not generate all prime numbers. Between $p(2) = 3$ and $p(3) = 7$ is $p = 5$.

The row sums of the Erasthothenes sieve (**EJ**) are the numbers of divisors. They appear on the diagonal of the quadratic form $\mathbf{E}\mathbf{E}^{\mathrm{T}}$ of the matrix **E**. They are known as the *Euler function* $\sigma^0\ (n)$. This function is related with logarithms of divisors. If we use as the base of logarithms the number n itself, we get (except n = 1)

$$\sigma^0(n) = 2\sum \lg(d|n) \tag{6.2}$$

or for any base of logarithms

$$\sigma^0(n) = 2\sum \lg(d|n)/\lg a \; . \tag{6.3}$$

The divisors appear in pairs, $d_i d_j = n$, except the lone divisor which is the square root of n. The sum of logarithms with the base n is thus only a half of the number of divisors of the number n. The sum of divisors values $\sigma^1(n)$ sometimes gives twice the number itself as $2 \times 6 = 6 + 3 + 2 + 1$ or $2 \times 28 = 28 + 14 + 7 + 4 + 2 + 1$. Such numbers are known as the *perfect numbers*.

6.2 Moebius Inversion of the Erasthothenes Sieve

In Table 6.1 the Moebius function was shown as the inverse matrix $\mathbf{E}^{-1}$. The elements of its first column are

- $e_{i1}^{-1} = 1, \text{if} i = 1$, or in the case of the product of an even number of prime numbers;
- $e_{i1}^{-1} = -1$, if i is a prime number or a product of an odd number of prime numbers, and
- $e_{i1}^{-1} = 0$, if i is product of higher powers of prime numbers as $4 = 2^2$ in the Table 6.1.

These elements appear in other columns on places, where the ratio i/j is a whole number, otherwise there are zeroes. The unit elements are scarcer in higher columns.

The Moebius inversion is the classical example of the combinatorial *inclusion and exclusion principle*. Some objects are counted in their combinations twice or more times and then these overbalanced parts are subtracted in other combinations for obtaining the true value. We formulated this principle by a sophisticated technique of matrix products. This technique can be applied to all matrices which have the unit diagonal and all nonzero elements under or on the diagonal. The unit matrix $\mathbf{I}$ is subtracted from such a matrix and then the difference is multiplied with itself till all nonzero elements disappear (at most n times). For example

$$(\mathbf{E}-\mathbf{I}) \qquad (\mathbf{E}-\mathbf{I})^2 \qquad (\mathbf{E}-\mathbf{I})^3$$

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

Developing the product $(\mathbf{E}-\mathbf{I})^k$ when it is equal $\mathbf{0}$, the unit diagonal matrix is expressed as

$$\sum_{i=1}^{n} (-1)^i \binom{n}{k} \mathbf{E}^i = \mathbf{I}\,. \tag{6.4}$$

Multiplying both sides by $\mathbf{E}^{-1}$ and eliminating $\mathbf{E}^{-1}\mathbf{E} = \mathbf{I}$ we get

$$\mathbf{E}^{-1} = \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{k} \mathbf{E}^{i-1} \ . \tag{6.5}$$

Objects $\binom{n}{k}$ looking as one column matrices in both equations are known as binomial coefficients. They count all possibilities how to choose k objects from n objects. The inverse matrix $\mathbf{E}^{-1}$ is a sum of positive and negative multiples of positive powers $\mathbf{E}^k$. It sounds quite occult.

6.3 Divisor Functions

The number of divisors $\sigma^0(n)$ and the sum of divisors values are rather irregular functions. Their sequence and consecutive sums of $\sigma^0(n)$ are

n	1	2	3	4	5	6	7	8	9	10	11
$\sigma^0(n)$	1	2	2	3	2	4	2	4	3	4	2
$\sigma^1(n)$	1	3	4	7	6	12	8	15	13	16	12
$\sum[\sigma^0(n)]$	1	3	5	8	10	14	16	20	23	27	29

The sums $\sum[\sigma^0(n)]$ are obtained as traces of corresponding matrix products of growing Erasthothenes sieves $\mathbf{E}\mathbf{E}^{\mathrm{T}}$, or simply by counting elements of the matrix $\mathbf{E}$:

$$\sum[\sigma^0(n)] = \sum_{j=1}^{n} [n/j] \ , \tag{6.6}$$

where [n/j] means the whole part of the given ratio. Therefore the sum $\sum[\sigma^0(n)]$ has as a limit the product $n \sum_{j=1}^{n} n/j$. For example

$$\sum[\sigma^0(3)] = 5 < 3(1 + 1/2 + 1/3) = 11/2 \ .$$

If we arrange the elements of traces $\mathbf{E}^{\mathrm{T}}\mathbf{E}$ (this is the second quadratic form of the Erasthothenes sieve), or by counting consecutively elements in columns of the matrix $\mathbf{E}$ into a table and find its inverse, than its row sums give the values of the Moebius function (Table 6.2).

The row elements of the previous matrix $\mathbf{M}$ are $\mathbf{J}^{\mathrm{T}}\mathbf{E}$, thus the Moebius function is $\mathbf{M}^{-1}\mathbf{J}$.

A still more important function is the sum of the divisor values. It can be expressed as the matrix product having in the frame $\mathbf{E}(*)\mathbf{E}^{\mathrm{T}}$ the diagonal matrix of indices $\Delta(j)$. $\mathbf{E}\Delta(j)$ is the matrix of divisor values. The sums of divisor values $\sigma^1(n)$ are the diagonal elements of the matrix $\mathbf{E}\Delta(j)\mathbf{E}^{\mathrm{T}}$:

Table 6.2: Erasthothenes sieve diagonal values and their Moebius inversions.

	Diagonal values							Σ	Moebius inversion							Σ
j	1	2	3	4	5	6	7		1	2	3	4	5	6	7	
i=1	1							1	1							1
2	2	1						3	-2	1						-1
3	3	1	1					5	-1	-1	1					-1
4	4	2	1	1				8	1	-1	-1	1				0
5	5	2	1	1	1			10	-1	0	0	-1	1			-1
6	6	3	2	1	1	1		14	2	0	-1	0	-1	1		1
7	7	3	2	1	1	1	1	16	-1	0	0	0	0	-1	1	-1

$$\mathbf{E}\Delta(j) \qquad\qquad \mathbf{E}\Delta(j)\mathbf{E}^{\mathrm{T}}$$

$$\begin{pmatrix} 1 & & & \\ 1 & 2 & & \\ 1 & 0 & 3 & \\ 1 & 2 & 0 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 3 \\ 1 & 1 & 4 & 1 \\ 1 & 3 & 1 & 7 \end{pmatrix}$$

The number of divisors j which also gives ratios n/d is obtained as another matrix product:

$$\Delta(j)\mathbf{E}[\Delta(j)]^{-1} \tag{6.7}$$

The rows of $\mathbf{E}$ are multiplied by the corresponding index i and the columns are divided by the corresponding index j. The elements of the matrix product are $e_{ij} = i/j$, if $i \equiv 0 \bmod j$, and $e_{ij} = 0$ otherwise.

$$\begin{pmatrix} 1 & & & & \\ 2 & 1 & & & \\ 3 & 0 & 1 & & \\ 4 & 2 & 0 & 1 & \\ 5 & 0 & 0 & 0 & 1 \end{pmatrix} .$$

If we multiply this matrix by the inverse $\mathbf{E}^{-1}$, we get the matrix which elements count the numbers of those numbers between 1 and n that are divided by the given divisor, provided, that they were not already divided by a greater divisor. Thus the row sums of the table are always n.

For example: 1 in the sixth row divides 1,5; 2 divides 2,4; 3 and 6 divide themselves and 4 and 5 are not divisors.

This inverse function has again the table form (see Table 6.4).

Table 6.3: Numbers of numbers divided by the given divisors

n	1	2	3	4	5	6	7	8	Σ
m=1	1								1
2	1	1							2
3	2	0	1						3
4	2	1	0	1					4
5	4	0	0	0	1				5
6	2	2	1	0	0	1			6
7	6	0	0	0	0	0	1		7
8	4	2	0	1	0	0	0	1	8

Table 6.4: Inverse function of numbers of numbers

n	1	2	3	4	5	6	7	8	Σ
m=1	1								1
2	-1	1							0
3	-2	0	1						-1
4	-1	-1	0	1					-1
5	-4	0	0	0	1				-3
6	2	-2	-1	0	0	1			0
7	-6	0	0	0	0	0	1		-5
8	-1	-1	0	-1	0	0	0	1	-2

It is necessary to find elements d_{i1}^{-1} of the first column, since in further columns are only the elements of the first column diluted by zeroes as in the basic matrix. It is obvious that the elements $(1-p)$ must appear in prime rows, there are zeroes in the following columns. Values for i = 4, 6 show, that powers of prime numbers are just products of these elements. Value 2 in the sixth row is interpreted as $(-1) \times (-2)$, the product of two divisors. To check it, we try to find the solution for 30, the product of three prime divisors

Divisors	1	2	3	5	6	10	15	30	Σ
Divided numbers d_{i1}	8	8	4	2	4	2	1	1	30
d_{i1}^{-1}	1	-1	-2	-4	2	4	8	-8	
$d_{i1}d_{i1}^{-1}$	8	-8	-8	-8	8	8	8	-8	0

where $d_{30}^{-1} = -8 = -1 \times -2 \times -4$, or -4×2, if we express 30 as 5×6.

Another division function is the function $\varphi(n)$. This function counts the numbers, which are not divisible by the divisors of n except 1. They are

n	1	2	3	4	5	6	7
$\varphi(n)$	1	1	2	2	4	2	6
Counted numbers	1;	1;	1,2;	1,3;	1 – 4;	1,5;	1 – 6

The values $\varphi(n)$ appeared as elements in the first column Table 6.4. It was shown that $\varphi(n)$ are easily found as the product:

$$\varphi(n) = n \prod_{p=2}^{n} (1 - 1/p) \ , \tag{6.8}$$

where p are prime numbers that are divisors of n. The ratio n/p is split from the number n by each inverse of the prime number $1/p$. The sum counts all subtracted parts from the total n. The function $\varphi(n)$ of the product of two numbers is simply the product of the values for each number

$$\varphi(nm) = \varphi(n)\varphi(m) \ . \tag{6.9}$$

The following relation is very interesting

$$\sum_{n_{d|n}} \varphi(d) = n \ . \tag{6.10}$$

For example: for $n = 6$: $\varphi(1) + \varphi(2) + \varphi(3) + \varphi(6) = 1 + 1 + 2 + 2 = 6$.

6.4 Relation Between Divisors and Partitions

The reason why the Erasthothenes sieve was introduced, is its involvement in counting of partitions. In each unrestricted plane simplex are $p(m)$ partitions of the number m. The sum of their parts is $m \times p(m)$. This product is obtained from the Erasthothenes sieve, if this is multiplied from the left by the diagonal matrix Δ of unrestricted partitions written in the decreasing order: $p(i) = p(m-i)$ and from the right by the diagonal matrix Δ of indices i. For example

$$\begin{pmatrix} 5 & & & & \\ & 3 & & & \\ & & 2 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 0 & 1 & & \\ 1 & 1 & 0 & 1 & \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 3 & & \\ & & & 4 & \\ & & & & 5 \end{pmatrix}$$

$$= \begin{array}{c} \begin{pmatrix} 5 & & & & \\ 3 & 6 & & & \\ 2 & 0 & 6 & & \\ 1 & 2 & 0 & 4 & \\ 1 & 0 & 0 & 0 & 5 \end{pmatrix} \\ \hline 12 \quad 8 \quad 6 \quad 4 \quad 5 \end{array}$$

The sum of elements of the product is $35 = 5 \times 7$. The partition $p(5)$ was obtained from values of parts added to lower simplices which were counted. Ones are counted in the first column. They were added to $p(m-1)$ partitions. But this set contains simultaneously all ones from lower partitions enlarged by such a way in lower steps, till one representing p(1). In the second column two is added to 3 partitions of 3. One of them, (2,1) already contained one 2, when this partition was obtained from $p(1)$. Similarly, other numbers are counted in following columns.

This product of 3 matrices can be inserted into the frame $\mathbf{J}^{\mathrm{T}}(*)\mathbf{J}$ which sums up the elements of the framed matrix. The insert in the frame is:

$$\mathbf{J}^{\mathrm{T}}\Delta[p(m-i)]\mathbf{E} \ \times \{\Delta(i)\mathbf{J} \tag{6.11}$$

Consecutive vectors form matrices in lower and upper triangular form and products of 3 matrices are replaced by a product of only 2 matrices:

					1	1	1	1	1
						2	2	2	2
							3	3	3
								4	4
									5
1					1	1	1	1	1
2	1				2	4	4	4	4
4	1	1			4	6	9	9	9
7	3	1	1		7	13	16	20	20
12	4	2	1	1	12	20	26	30	35

The left matrix counts the numbers n_k in the partitions, the right matrix weights them as m_k.

The diagonal elements $mp(m)$ can be decomposed into another pairs of vectors and so another product of 2 matrices exists having identical diagonal. The left matrix is the matrix of successive partitions (Table 4.8), the right matrix is the matrix of sums of divisors $\sigma^1(i)$, written similarly as the matrix of successive partitions, but in upper triangular form, in columns

$$\mathbf{S} : s_{ij} = \sigma^1(i) \text{ if } i \leq j, s_{ij} = 0, \text{ otherwise .} \tag{6.12}$$

					1	1	1	1	1
						3	3	3	3
							4	4	4
								7	7
									6
1					1	1	1	1	1
1	1				1	4	4	4	4
2	1	1			2	5	9	9	9
3	2	1	1		3	9	13	20	20
5	3	2	1	1	5	14	22	29	35

.

The numbers $mp(m)$ appear again on the diagonal of the product. This elementary relations can be expressed in a formal abstract form. We write the generating function of unrestricted partitions

$$P(x) = \sum_{m=1}^{\infty} p(m)\, q^m \tag{6.13}$$

and find its derivation

Table 6.5: Numbers of parts in partitions

n	0	1	2	3	4	5	6	Σ
m=0	1							1
1	0	1						1
2	1	2	1					4
3	3	4	1	1				9
4	8	7	3	1	1			20
5	15	12	4	2	1	1		35
6	31	19	8	4	2	1	1	66

$$dP(x) = \sum_{m=1}^{\infty} mp(m)\, q^{m-1} \,. \tag{6.14}$$

The partition function $P(x)$ is represented in the rows of the left matrix. The difference $dP(x)$ appears on the diagonal of the product. When we find the ratio of both matrices the result can be formulated as

$$d\lg[P(x)] = dP(x)/P(x) = \sum_{m=1}^{\infty} \varphi(m) q^m \,. \tag{6.15}$$

The ratio $dP(x)/P(x)$ is the difference of the logarithm of the function $P(x)$. Divisor sums are thus the differences of the logarithmic measure of the generating function of partitions. It relates divisor and partition functions and it was used for finding of the asymptotic behavior of the $p(m)$ function.

6.5 Zeroes in partitions

If the sum of values of all parts of the plane simplex is $mp(m)$, we can find also the number of zeroes in all parts $n_0(m)$. These numbers form the first column in the table which counts the number of all parts classified according to their values (Table 6.2)

Matrix elements of the Table 6.2, except its first column, are obtained as the partial result of the matrix product used for finding the sum of values of parts in partitions exploiting equation 5.3. They are elements of products of two matrices $\Delta[p(m-i)]\mathbf{E}$. The first column is in its turn again a matrix product of the matrix of partitions into exactly n parts (Table 4.2) and the matrix of positive $(j-i)$ elements, and the unit vector column $\mathbf{J}$, which sums the row values of the intermediate product. It is easy to explain this

relation: In each partition, when m splits into exactly n parts, there are $(m-n)$ zeroes. For example: for $m = 4 : 8 = 3\times1+2\times2+1\times1+0\times1$. The number of zeroes is balanced by other numbers. This leads to the simple form of some elements of inverse matrix

$$m_{i0}^{-1} = (1 - i) \; .$$

Chapter 7

Groups of Cyclic Permutations

7.1 Notion of Cyclic Permutations

Lets suppose that we have n objects labeled by an index, arranged in natural order, and we change their positions. It is convenient to describe this operation of *permutation* in two lines, the first one corresponding to the starting position, the second one giving the final position. E.g.

- Start 0: 1 2 3 4 5 6
- Step 1: 2 3 1 5 4 6

The first three objects are permuted in a *cycle of length 3*, the first object appeared after the third one, next two objects form a cycle of length 2, they exchanged their places, and the last object remained in its place. By repeating the procedure we obtain permutations 2 to 6:

- Step 2: 3 1 2 4 5 6
- Step 3: 1 2 3 5 4 6
- Step 4: 2 3 1 4 5 6
- Step 5: 3 1 2 5 4 6
- Step 6: 1 2 3 4 5 6
- Step 7: 2 3 1 5 4 6

The string returns in the 6 th step into the initial order and a new cycle starts in the 7 th step.

The index labeling objects is the column index j. The position in the permutation is labeled by the row index i at the element 1_{ij}. Thus permutations are isomorphic with matrices. The starting position, corresponding to the diagonal unit matrix $\mathbf{I}$, can be considered as the zero order. The last element remained in all steps in its original position and the first three elements returned to their positions twice and two elements made three turns. The length of the total cycle is the product of individual cycles:$3\times 2\times 1 = 6$. The elements belonging to the same cycles are usually written in brackets: $(2, 3, 1)(5, 4)(6)$.

The number of elements n splits into k cycles, k going from 1 to n. The cycle structure is described by *partition orbits*.

We could map the cycle changes by the additive operators $\mathbf{S}$ having -1_{ij} for the leaving object j, 1_{ij} for the becoming object j, zero rows for unmoved objects (+1 and -1 appear on the same place). This operator was introduced in Chapt. 3 and in more detail will be studied in Chapt. 12. Now we will study the *multiplication operators* $\mathbf{P}$. Their matrices, *unit permutation matrices*, are naive, they have in each row only one unit element and moreover they have only one unit element in each column. The matrices $\mathbf{P}$ are simultaneously notations of permutations, since their row unit elements p_{ij} correspond to indexes (or equivalently to alphabetical symbols) j.

Using multiplicative operators, permutations are the results of multiplication's of the row vectors by the unit permutation matrix $\mathbf{P}$ from the right and column vectors by the multiplication with the unit permutation matrix $\mathbf{P}$ from the left. Different steps can be written as powers of these matrices $\mathbf{P}^i$. The unit diagonal matrix is $\mathbf{I} = \mathbf{P}^0$.

The last but one power of any permutation matrix is its inverse (Fig. 7.1). It is rather easy to find this matrix, because it is identical with the transposed matrix $\mathbf{P}^{\mathrm{T}}$:

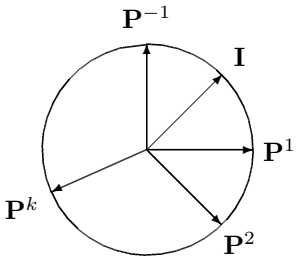
$$\mathbf{P}^{n-1} = \mathbf{P}^{-1} = \mathbf{P}^{\mathrm{T}} \ . \tag{7.1}$$

The set of all permutation matrices $\mathbf{P}$, with n rows and n columns, represents all possible permutations. A special class of permutation matrices are the symmetrical ones. For them the following relations are true:

$$\mathbf{P} = \mathbf{P}^{-1} = \mathbf{P}^{\mathrm{T}} \ . \tag{7.2}$$

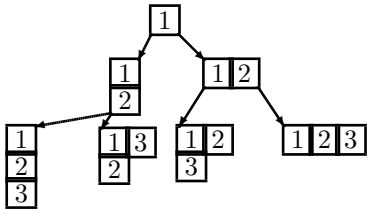
Such matrices have all unit elements either on the diagonal, or

Figure 7.1: Cycle of permutation matrices. Positive powers become negative ones.



.

Figure 7.2: Sequence of Young tables



otherwise they form cycles of the length 2. These permutations are known as *convolutions*. We will show a surprisingly simple technique for their generating.

7.2 Young Tables

We will reconstruct the sequence of the Ferrers graphs, finding all ways they can be formed from lower graphs by adding one new element. To do this the order, each box was added to a smaller Ferrers graph enlarging it into the larger Ferrers graph, is indexed. Equivalent boxes will have different indices, because they can be reached in different steps. Such labeled Ferrers graphs are known as Young tables (Fig. 7.2).

Young tables are connected with permutations by the following algorithm:

- 1 If a greater index follows a lesser one, it is written in the next free

column of the Young table.

- 2. If a lesser index follows a greater one in a permutation, it replaces it in its column of the Young table and shifts it down to the next row. For example:

$$\begin{array}{cccccccc} 3412 & \rightarrow & 34 & \rightarrow & 14 & \rightarrow & 12 \\ & & & & 3 & & 34 \end{array}$$

The third element 1 jumps in the first column and shifts 3 down, then 2 shifts 4 down. Or:

$$\begin{array}{ccccccccc} 4231 & \rightarrow & 4 & \rightarrow & 2 & \rightarrow & 2\,3 & \rightarrow & 1\,3\,. \\ & & & & 4 & & 4 & & 2 \\ & & & & & & & & 4 \end{array}$$

One property of the algorithm seems to be disadvantageous but this property only reproduces relations between permutations. It allows an asymmetric permutation matrix to be decomposed differently according to its rows and columns. But both Young tables belong to the same type of Ferrers graphs. For example:

							Σ
0	0	0	1	0	0	0	4
0	0	0	0	0	1	0	6
0	1	0	0	0	0	0	2
0	0	0	0	1	0	0	5
0	0	0	0	0	0	1	7
1	0	0	0	0	0	0	1
0	0	1	0	0	0	0	3
6	3	7	1	4	2	5	

$$\begin{array}{cccc} & 1 & 2 & 5 \\ \textit{Columns} & 3 & 4 & \\ & 6 & 7 & \end{array} \qquad \begin{array}{cccc} & 1 & 3 & 7 \\ \textit{Rows} & 2 & 5 & \\ & 4 & 6 & \end{array}$$

Remember that convolutions have symmetrical matrices and that then column and row readings are identical. A permutation matrix or a Young table is always a product of two permutation matrices or Young tables of the same type. They can be identical in case of convolutions, but mostly they differ in rows and columns, as

			1	0	0
			0	0	1
			0	1	0
0	1	0	0	0	1
1	0	0	1	0	0
0	0	1	0	1	0

$$(a, c, b) \times (b, a, c) = (c, a, b)$$

A relation there appears between the number of partitions orbits $p(n)$, the number of Young tables $Y(n)$, and the number of permutation matrices $P(n)$.

The Young tables are formed from Ferrers graphs by a recursive algorithm. If we use for the number of Young tables corresponding to a Ferrers graph with an index k the notation $y(k)$, then $y^0(k) = 1$, and we have the relation between the number of partitions $p(n)$ and the number of Young tables $Y(n)$. Similarly, if we square all y(k), we obtain all permutations of n elements. Therefore

$$\sum y^0(k) = p(n) \; ; \sum y(k) = Y(n) \; ; \sum y^2(k) = P(n) = n! \tag{7.3}$$

Here $n!$ means n *factorial*. It is a product of successive natural numbers:

$$\prod_{k=1}^{n} k = n! \; . \tag{7.4}$$

We will explain this function later, when we will look for other formulas determining the number of permutations. Before then we will study convolutions. Here an example is given how equation (7.4) works:

Partition:	5	4,1	3,2	$3, 1^2$	$2^2 1$	$2, 1^3$	1^5	Σ
$y^0(k)$	1	1	1	1	1	1	1	7
$y^1(k)$	1	4	5	6	5	4	1	26
$y^2(k)$	1	16	25	36	25	16	1	120

7.3 The Number of Convolutions

The number of convolutions is the number of all possible connections in a telephone network. We classify all convolutions according to the number of elements which remain on their places, that means unconnected. It is easy to fill in the following table

Table 7.1: Distribution of convolutions

On diagonal	0	1	2	3	4	5	6	Σ
n=0	1							1
1	0	1						1
2	1	0	1					2
3	0	3	0	1				4
4	3	0	6	0	1			10
5	0	15	0	10	0	1		26
6	15	0	45	0	15	0	1	76

The recurrence of the table elements is

$$y_{00} = 1 \; ; y_{ij} = (i-1)y_{i-2,j} + y_{i-1,j-1} \; . \tag{7.5}$$

The inverse table has the same elements, only the signs of elements which indices i differ from indices j by the value $(4k+2)$ are negative. Their recurrence is

$$y_{00}^{-1} = 1 \; ; y_{ij}^{-1} = (1-i)y_{i-2,j} + y_{i-1,j-1} \; . \tag{7.6}$$

All convolution matrices are obtained in two ways. Either by adding 1 to the last place of diagonal. These convolutions are counted by the term $y_{i-1,j-1}$. Or the unit element is added in the last row off-diagonal position. It is inserted between existing columns into a new one. Then an unit element must simultaneously added in the last column into a new row, $i = j$. There are $(n-1)$ off diagonal places where it is possible to form a new convoluted pair to j already existing pairs in a matrix of convolutions. This new pair occupies two rows and columns and therefore it is formed in matrices with $(n-2)$ rows and columns. It does not increase the number of elements on the diagonal, so it increases the number of elements in the same column.

A similar recurrence is applied to the row sums counting the total number of convolutions

$$Y(n) = (n-1)Y(n-2) + Y(n-1) \; . \tag{7.7}$$

It is possible to determine the elements of the Table 7.1 directly, because they are calculated according to the formula

$$y_{ij} = i!/j!t!2^t \; , \tag{7.8}$$

where $t = (i - j)/2$ is the number of cycles of length 2. This formula contains 3 factorials. The term 2^t equilibrates the string of t divisors. Notice, that equation (7.3) counts together Young tables of different formats, they must have only the same number of columns. Still another expression of the number of convolutions is a formal binomial equation

$$Y(n) = (1 + y_i)^a \ , \tag{7.9}$$

where the terms in the first column of the Table 7.1 y_{k0} are considered as powers of y^k when the sum $(1+y)$ is multiplied with itself. For example

$$(1 + y)^6 = 1 \times 1 + 6 \times 0 + 15 \times 1 + 20 \times 0 + 15 \times 3 + 6 \times 0 + 1 \times 15 = 76 \ .$$

The convolutions counted by these terms have no elements on the main diagonal and are obtained by multiplying odd numbers. They are *odd factorials*, since they are obtained by consecutive multiplication's of odd numbers: $1 \times 3 \times 5 \times 7 \times 9 \times 11 \times 13 \times 15 \times$ and so on.

7.4 Factorial and Gamma Function

The number of all permutation matrices $P(n)$ is determined easily by counting possibilities of arrangements of the units in rows and columns in a permutation matrix. In the first row there are n possibilities, in the second row one column is blocked by the element of the first row. The second row element can not be in the same column. The possibilities decrease regularly. In each row $(n - i)$ remaining places are free. These possibilities are independent and therefore they multiply for all rows. We get the factorial

$$P(n) = n \times (n-1) \times \ldots \times 2 \times 1 = \prod_{j=1}^{n} j = n! \tag{7.10}$$

The factorial function has an interesting property. If p is a prime number, then $(p-1)!$ mod $p = (p-1)$ and simultaneously $(p-2)!$ mod $p = 1$. The factorial is divisible by all its factors, say b. If the modular value were different, say a, then this value could be chosen in such a way that $a+b = p$. The factorial were divisible by the prime number greater than its factors, which is impossible. For example: $p = 7,\ 720 \bmod 7 \equiv 6;\ 120 \bmod 7 \equiv 1$.

The factorial function is defined for natural numbers, zero including. We complete its definition by the term $0! = 1$. We already did something similar, defining the empty partition.

Combinatorial functions are defined for counting objects, which must be whole. There can emerge questions, what is a object, or an animal,

or a man, when they start to correspond to their definitions and when they are something different. In mathematics, such small differences can be expressed by numbers. In higher mathematics the factorial function is only the special case of the *gamma function*, defined by Euler as

$$\Gamma(z+1) = z\Gamma(z) \tag{7.11}$$

When $\Gamma(1) = 1$, then

$$\Gamma(2) = 1\Gamma(1) = 1,$$

$$\Gamma(3) = 2\Gamma(1) = 2,$$

and

$$\Gamma(4) = 3\Gamma(3) = 6\ .$$

Therefore

$$\Gamma(n+1) = n!\ .$$

Drawing the graph of the gamma function, we can interpolate it for any real number. The gamma function is defined by the integral[1].

$$\Gamma(z+1) = \int_0^\infty x^z e^{-x} dx\ . \tag{7.12}$$

We will not deal with problems connected with evaluation of such integrals, and we introduce the function e in the next chapter. Now we accept only the result giving for

$$\Gamma(1/2) = \sqrt{\pi}. \tag{7.13}$$

From it, other $n/2$ values of the gamma function are calculated easily which fits excellently in holes between factorials to plot one smooth function (Fig. 7.3).

When we interpolate the gamma function to negative values:

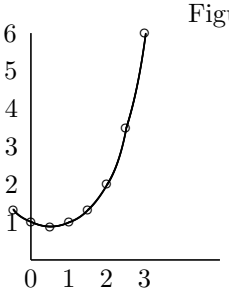
$$\Gamma(1) = 0\Gamma(0)$$

we get

$$\Gamma(0) = \Gamma(1)/0 = \infty$$

[1] e in the integral is the base of the natural logarithms. Logarithms can be decadic $\lg a$, binary a, natural $\ln a$, or with any base b $\log_b a$.

Figure 7.3: Plot of the function Gamma.



$$\Gamma(0) = (-1)\Gamma(-1)$$

$$\Gamma(-1) = \Gamma(0)/(-1) = -\infty \, .$$

The gamma function oscillates for consecutive negative numbers from $+\infty$ to $-\infty$, and than it starts with an opposite sign again in infinity. The functional relation is no more solid but it behaves as the see at storm under clouds. The world of mathematical functions is not symmetrical to the sign inversion, similarly as our physical world, where antiparticles are rare events which annihilate immediately.

The Euler gamma function can be used for finding approximations of the factorial function for large n. The Stirling approximation is

$$n! = n^n e^{-n} \sqrt{2\pi n} \, . \tag{7.14}$$

7.5 Index of cyclic permutations

After this transgression, we now return to permutations finding formulas to determine the numbers of each cycle structure. The cycle structure forms an orbit of permutations and the sum over all orbits gives the factorial. A partition orbit counts all permutations of a cycle s_k of the length k. If there are more cycles of equal length, their lengths s_k multiply. This gives the terms $s_k^t k$, where t_k is the number of cycles s_k. Different cycles of equal length are permuted between themselves with each other when their elements interchange. This interchanges are counted by partial factorials $t_k!$. The index of cyclic permutations is

$$n!/\prod n_k! s_k^t k \ . \tag{7.15}$$

For example: for n = 4:

Orbit	Cycle index	Value	
4	$4!/1!^4$	6	One cycle of length 4
31	$4!/1!^1 3!^1$	8	One cycle of length 3, one cycle of length 1
22	$4!/2!2$	3	Two cycles of length 2
211	$4!/1!^1 2!^1$	6	One cycles of length 2, two cycles of length 1
1	$4!/4!^1$	1	Four cycles of length 1
Σ		24	

7.6 Permutation Schemes

We introduced orbit schemes and now we have the first opportunity to use them for calculating partial sums of cyclic indices. These partial sums are known as different *combinatorial identities*. At first, we will arrange partition schemes according to the number of cycles in permutations and the length of the longest cycle k. For example for $n = 6$ we get:

n	1	2	3	4	5	6
$k = 6$	120					
5		144				
4		90	90			
3		40	120	40		
2			15	45	15	
1						1
Σ	120	274	225	85	15	1

The row sums of consecutive schemes give the Table 7.2. Its elements are known as the *Stirling numbers of the first kind*. Their name suggests that there are more kinds of Stirling numbers. They are related by many ways as we will see later.

The recurrence of Stirling numbers is

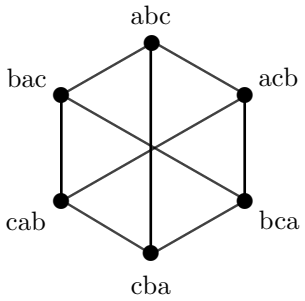
$$s_{ij} = (n-1)s_{i-1,j} + s_{i-1,j-1} \ . \tag{7.16}$$

The formula was explained by describing how permutation matrices $\mathbf{P}_{n-1}$ are enlarged with the new row and column. We have $(n-1)$ off-diagonal positions in the last row which split $(n-1)$ dimensional permutation matrices and prolong some existing cycle but do not change their

Table 7.2: Stirling numbers of the first kind

t	1	2	3	4	5	6	Σ
n=1	1						1
2	1	1					2
3	2	3	1				6
4	6	11	6	1			24
5	24	50	35	10	1		120
6	120	274	225	85	15	1	720

Figure 7.4: Central orbit in the 3 dimensional cube with the sides 0-2. Lines connect points with distances 2.



number. Then the unit element can be added on the diagonal but this operation increases the number of cycles of unit length. In this way we obtain the intermediate sums of several cycle indices directly without changes of all corresponding orbits. Remember that these sums correspond to vertices, edges, surfaces and generally n dimensional subsimplices of the surface simplex. But here they split only one original orbit in the center of the plane simplex or the central orbit in the cube (Fig. 7.4).

7.7 Rencontres Numbers

Another possibility to count permutations is to use the number of unit cycles, this is to determine the unit elements on the main diagonal of the unit permutation matrices, known as *unmoved elements*. The counts of partitions can be obtained according to the number of ones in partitions. Using this technique for tabulating permutation indices, we obtain the column sums known as the *rencontres numbers*. They are shown in the Table 7.3.

Table 7.3: Rencontre numbers

s	0	1	2	3	4	5	6	Σ
n=0	1							1
1	0	1						1
2	1	0	1					2
3	2	3	0	1				6
4	9	8	6	0	1			24
5	44	45	20	10	0	1		120
6	265	264	135	40	15	0	1	720

The recurrence is somewhat surprising; the rencontres numbers are obtained from the zero column by multiplying it with binomial coefficients

$$r_{ij} = \binom{i}{j} r_{i-j,0} \tag{7.17}$$

Compare this with the Table 4.6 of partitions ordered according to the number of unit parts. Now these parts are just combined with other parts. The elements of the Table 7.3 are obtained as terms of a somewhat complicated expression

$$n! = 1 + (1 - 1/1!)n + (1 - 1/1! + 1/2!)n(n-1) + \ldots \tag{7.18}$$

which can be formulated as

$$n! = \sum_{k=0}^{n} (-1/k!)^k (n)_k \ . \tag{7.19}$$

For example: $4! = 1 + 0 + 1/2 \times 12 + 2/6 \times 24 + 9/24 \times 24$.

Now it is necessary at least to explain, that the binomial coefficient $\binom{i}{j}$ is a ratio of 3 factorials $i!/j!(i-j)!$. How a binomial coefficient is obtained, we will see later. Here we give an example how the 5-th row of the Table 7.3 are obtained by equation (7.19): $1\times44+5\times9+10\times2+10\times1+5\times0+1\times1 = 120$.

The rencontres numbers r_{i0} count permutations matrices with i rows and columns having no unit elements on the diagonal (no unmoved object). These matrices are combined with the diagonal unit matrices $\mathbf{I}$ with $(i-j)$ rows and columns in all possible ways counted by the binomial coefficient.

The rencontres numbers r_{i0} are known also as *subfactorials*, because they produce factorials by the following equation which terms were determined by 7.7, Now they are inserted as formal powers of subfactorials $r^i = r_i$:

$$n! = (r_i + 1)^a \ . \tag{7.20}$$

It is possible to formulate equation (7.19) also in the matrix form as the direct product

$$\Delta(n!) = \mathbf{R} \times \mathbf{B} \ , \tag{7.21}$$

where $\mathbf{R}$ is the matrix of subfactorials in rows and $\mathbf{B}$ is the matrix of binomial coefficients. Inverting the formal powers we have $r(n)_0 = (k!^n - 1)$. Inserting $(k!)^n = n!$ we obtain the formula

$$n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \ldots \pm \binom{n}{n}(n-n)! = (k!)^a \ . \tag{7.22}$$

This becomes for n going to infinity

$$n![1 - 1 + 1/2! - 1/3! + \ldots] \approx n^n/e^n \ , \tag{7.23}$$

where e is the base of natural logarithms. This approximate formula gives the rough Stirling approximation for factorials of large numbers. Compare with the exact formula (7.4).

We should mention still another formal notation for subfactorials. It is the notation of the theory of finite differences[2].

$$r_0(n) = [E - 1]^n 0! = \Delta^n 0! \ . \tag{7.24}$$

Here Δ^n is not a diagonal matrix but a difference of the n-th degree, or n times repeated difference of the basic state E.

We rencontres the rencontres numbers again in Chapt. 14.

There exists still another recurrence for subfactorials

$$r_{n0} = n r_{n-1,0} + (-1)^a \ , \tag{7.25}$$

For example: $5 \times 9 - 1 = 44$; $6 \times 44 + 1 = 245$. When we return to the partition scheme in Table 7.3 and reclassify permutations without unit cycles according to the number of cycles, or when we delete from the original scheme (Table 7.2) all permutations with unit cycles, we obtain the Table 7.4 of the adjoined Stirling numbers of the first kind.

The recurrence is

$$a_{i+1,j} = i[a_{ij} + a_{i-1,j-1}] \ . \tag{7.26}$$

[2]This will be explained in Sect. 9.4.

Table 7.4: Adjoined Stirling numbers of the first kind

j	0	1	2	3	$\Sigma = r$
i=0	1				1
1		0			0
2		1			1
3		2			2
4		6	3		9
5		24	20		44
6		120	130	15	265

The recurrence is justified again by possibilities to insert a new element to existing cycles. Either we can insert it into an existing cycle or a new cycle can be formed. It is simpler to formulate this for $(i+1)$ matrices. There we have in $(i+1)$ dimensional matrix i off-diagonal possibilities to insert a new element to existing cycles. Or we can add a new 2 dimensional cycle to matrices with $(i-1)$ rows with the same number of possibilities.

7.8 Euler Numbers

We have not exhausted all possibilities of classifying permutations. Another statistics counts the number of *segments* of a permutations in which its elements are arranged according to their natural order as increasing indices. For example: a permutation (357168942) is split into four segments 357/1689/4/2. The recurrence of this statistics, known as Euler numbers, is:

$$e_{11} = 1 \, ; e_{ij} = j e_{i-1,j} + (i - j + 1) e_{i-1,j-1} \, . \tag{7.27}$$

If we give the i-th element to the end of each segment, the number of segments remains unchanged. If we put it in the first place, we increase the number of segments. Similarly, if we put it inside an existing segment this is then split into two segments. There are $(i-j)$ places inside segments. An alternative explanation is that this statistics counts elements of permutation matrices, which are over the main diagonal. Here the index j goes from 0 to $(n-1)$. The corresponding matrix is Table 7.8.

A question: How the inverse function of the Euler numbers can be interpreted?

Second order Eulerian triangle is also known. Its recurrence equation is

$$t_{11} = 1 \, ; t_{ij} = j e_{i-1,j} + (2i - j) t_{i-1,j-1} \, . \tag{7.28}$$

Table 7.5: Euler numbers

j	1	2	3	4	5	6	Σ
i=1	1						1
2	1	1					2
3	1	4	1				6
4	1	11	11	1			24
5	1	26	66	26	1		120
6	1	57	302	302	57	1	720

Table 7.6: Mac Mahon numbers

k	0	1	2	3	4	5	6	7	8	9	10
n=1	1										
2	1	1									
3	1	2	2	1							
4	1	3	5	6	5	3	1				
5	1	4	9	15	20	22	20	15	9	4	1

The first column is formed by ones, on the diagonal are factorials, and the row sums are odd factorials.

7.9 Mac Mahon Numbers

Till now we have counted permutations as objects. Now we will determine their moments, expressed by the number of inversions in a permutation. They are counted by zero elements over unit elements which are below the main diagonal as in the example, where 4 on the first place has 3 inversions and 3 on the second place only 2

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

Permutations classified according to this method give the Mac Mahon numbers as in Table 7.6.

Notice that here the parameter k does not end at the number n but continues to the value $n(n-1)/2$. It is as if we counted these values on the diagonal of a square. The maximal moment k is just the sum of values $(i-1)$ where i is going from 1 to n

$$\sum_{i=1}^{n} (i-1) = \binom{n}{2} . \tag{7.29}$$

The distribution of moments is symmetric and therefore the matrix elements are related as

$$m_{ik} = m_{i,[i(i-1)/2]-k} . \tag{7.30}$$

The recurrence of Mac Mahon numbers m_{ij} is

$$m_{ij} = \sum_{k=0}^{n} (m-k, n-1) \; ; \; m_{10} = 1 \tag{7.31}$$

or for the $k \leq i$:

$$m_{ij} = m_{i-1,j} + m_{i,j-1} . \tag{7.32}$$

If we add to a lesser permutation matrix the unit element on the last diagonal place, it does not change the sum of displacements. It yields the term $m_{i-1,j}$. Matrices counted by the term $m_{i,j-1}$ are sums of elements of previous rows which moments are increased by adding a new element into the corresponding column. They have the required dimensionality. Their moments are increased by permutations of the last element into the first column.

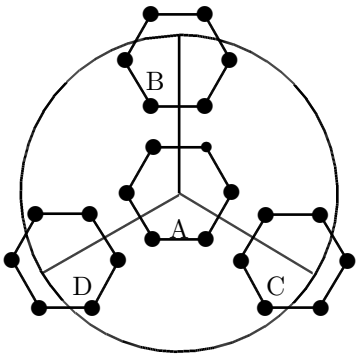
7.10 Spearman Correlation Coefficient

The sum of differences of positions of all objects permuted as compared to the basic unit permutation is always 0. These differences can be either positive or negative. The sum of squared differences must be necessarily positive. These differences of positions can be treated as distances in the cubes (see Fig. 7.4) for the three dimensional case and Fig. 7.5 where the four dimensional case is drawn).

Reference point:	1	2	3	4	5	Σ
Permutation point:	5	2	4	3	1	
(2-1)	4	0	1	-1	-4	0
Squares	16	0	1	1	16	34

If we divide obtained values by the largest possible sum of squares which is 40 for $n = 5$, we obtain values going from 0 to 1 which characterize permutations and are known as the *Spearman correlation coefficient*. It is used for evaluation of probability of obtained rank statistics.

Figure 7.5: 24 permutations of the string **abcd**. They are divided into four sets beginning by the capitals. Arrange the remaining three symbols and draw all permutations on the sphere.



7.11 Reduced groups of cyclic permutations

Till now we have worked with permutations which were read from one side, only. Most of our symbols determine from which side they must be read (with some exceptions as W, A, T, 8, and other symmetric symbols are). Imagine now, that a permutation is represented by a string of colored beads as

(red)-(blue)-(white)-(green)-(yellow)

If we find such a string accidentally, we can not tell from which side we should read it. The result is that we can not distinguish a half of permutations as:

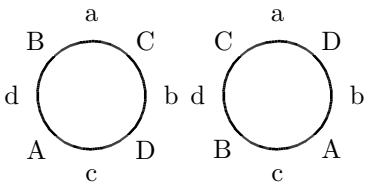
$$123 \leftrightarrow 321;\ 213 \leftrightarrow 312;\ 132 \leftrightarrow 231\ .$$

The name of such a group, which is undistinguishable by readings from both sides is the *dihedral*.

Still more complicated situation is if the string of colored beads forms a necklage. Then we can not find neither the reading direction neither the beginning of a permutation. Thus we have undistinguishable permutations:

$$(123-231-312) \leftrightarrow (213-132-321)\ .$$

Figure 7.6: Menage problem. Two sitting plans for four couples.



From problems connected with these reduced groups, we mention only the task of *menages*: n married couples should be seated at a round table in such a way that a woman should seat between 2 males but not alongside her husband. For $n = 4$ there are 2 seating orders (Fig. 7.6).

The menage numbers M(n) are:

n	0	1	2	3	4	5	6
M(n)	2	-1	0	1	2	13	80

The negative value at $n = 1$ is necessary for complying with the recurrent relation:

$$(n-2)U_n = n(n-2)U_{n-1} + nU_{n-2} + 4(-1)^{n+1} \; . \tag{7.33}$$

For example: $3U_5^6 = 15 \times 2 + 5 \times 1 + 4(-1) = 39;\; U_5^6 = 13.$

7.12 Groups of Symmetry

Till now we have supposed that vectors are defined in multidimensional space and that the number of permutations is determined by the dimensionality of the space. It is possible to define groups which are only isomorphic with some groups S_n of cyclic permutations. As an example we introduce the group of 6 matrices with 2 rows and columns, which is isomorphic with S_3:

$$\mathbf{I} \qquad \mathbf{D} \qquad \mathbf{E}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} \qquad \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$

$$\mathbf{A} \qquad \mathbf{G} \qquad \mathbf{P}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix} \qquad \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$$

If we multiply by these 2 dimensional matrices a vector-row from the right (or a vector-column from the left) its Euclidean length remains constant but not the sum of their elements. The effect of these matrices can be shown on the unit circle. Operators **I** and **A** are mutually orthogonal, the other matrices rotate vectors for $(2/3)\pi$, that is for 120 degrees, 0.5 being $\cos 60\ ^0$, $\sqrt{3}/2 = 0.866 = 30\ ^0$.

Instead of cycles of different lengths, new symmetry elements appear in the three dimensional geometry.

There are *rotation axes*. If a figure has k dimensional rotation axis, it has k equivalent positions and returns to its original position after k translations which rotate it around the axis.

The other kind of symmetry elements is the *reflection plane* which reflects a figure as a double sided mirror.

These basic symmetry elements combine in different ways and their systems are known under different names.

7.13 Vierer Gruppe

One system of 4 unit permutation matrices 4×4 is:

$$\mathbf{I} \qquad \mathbf{B}$$

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

$$\mathbf{A} \qquad \mathbf{C}$$

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

If we imagine that these matrices permute vertices of a square labeled a, b, c, and d, then $\mathbf{I}$ is the identity, which leaves the positions of corners square unchanged, $\mathbf{A}$ and $\mathbf{C}$ reflect according to the planes perpendicular with the sides of the square and $\mathbf{B}$ rotates the corners the square around the center. The group contains all possible products of these four matrices.

With the group of these 4 matrices such groups of matrices are isomorphic, which are obtained by multiplying the unit permutation matrices $\mathbf{P}$ from the left by a suitable matrix and from the right by its inverse

$$\mathbf{UPU}^{-1} = \mathbf{P}_a \; ; \; \mathbf{UU}^{-1} = \mathbf{I} \, . \tag{7.34}$$

Using Hadamard matrices we get another group of four matrices

$$\mathbf{I} \qquad \mathbf{B}$$

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} .$$

$$\mathbf{A} \qquad \mathbf{C}$$

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix}$$

Notice, that corresponding matrices in both groups have identical traces, which are known as *characters of the group*.

Chapter 8

Naive Matrices in Lower Triangular Form

8.1 Another Factorial Function

Before we study all naive matrices $\mathbf{N}$, we will deal at first with the naive matrices in the *lower triangular form* which form a subgroup of naive matrices. The remaining naive matrices can be obtained from them by permuting columns with the unit permutation matrices $\mathbf{P}$ from right. Recall that naive matrices $\mathbf{N}$ have one unit element in each row. If a matrix is in the lower triangular form then all its nonzero elements must be on or below the main diagonal. Similarly, if a matrix is in the upper triangular form, then all its nonzero elements must be on or over the main diagonal. From all permutation matrices $\mathbf{P}$ only the *identity matrix* $\mathbf{I}$ has the triangular form. But it exists simultaneously in both triangular forms as all diagonal matrices do.

There is only one place in the first row of the lower triangular form for the unit element, two places are in the second row and always one place more in each consecutive row for the unit element. This situation is just opposite to the construction of permutation matrices. There the possibilities, of placement of the unit element decreased in every row. Nevertheless both approaches give the same result. Therefore there are n! naive matrices in the lower triangular form (or in the case of transposed naive matrices $\mathbf{N}^{\mathrm{T}}$ in the upper triangular form). The transposed naive matrices can be mapped onto points with natural coordinates in m dimensional cubes.

If we leave the first column as a dummy variable (indexed as zero column) for the center of the coordinate system: $\mathbf{e}_{0j} = 1$, the naive matrices

in the lower triangular form can be compared with terms of the formal multiplication

$$(1)(1+a)(1+a+b)(1+a+\ldots) = \prod_{j=1}^{n}(\sum_{j=1}^{n} \mathbf{e}_j) \; . \tag{8.1}$$

All transposed matrices **N** are placed in m dimensional rectangular parallelepiped which sides are $0, 1, 2, 3, .., (n-1)$. With these matrices all classifications as with the permutation matrices will be repeated. Matrices with m rows form in these parallepides factorial plane simplices. Compare them with the generating function of partitions, explained in Sect. 4.10, where the number of elements also decreased but from other reasons.

8.2 Decreasing Order Classification

In the preceding Chapter Young tables were introduced and compared with the convolutions having cycles of length 1 and 2, only. The Young tables correspond to naive matrices which partial column sums are always ordered in the decreasing order:

$$\sum_{i=1}^{k} n_{ij} \geq \sum_{i=1}^{k} n_{i,j+1} \; . \tag{8.2}$$

For example: two naive matrices $n = 3$ are excluded by this rule:

$$\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \; . \end{array}$$

A is excluded since $b^2 > a$, **B** is excluded since $c > b^0$.

8.3 Stirling Numbers of the First Kind

These numbers count the naive matrices classified according to the number k of elements on the main diagonal

$$s_{nk} = (n-1)s_{n-1,k} + s_{n-1,k-1} \; . \tag{8.3}$$

There are $(n-1)$ places under the diagonal in the n th row which can be added to each naive matrix with k elements on the main diagonal without changing k. This multiplies the first term.

If we add 1_{nn}, we increase the number of elements on the main diagonal counted by the second term. See Table 7.2.

This is not all what can be said about the Stirling numbers of the first kind. If we multiply the Stirling numbers of the first kind directly with powers 2^{j-1} we get a table which row sums are equal to the half of the higher factorial $(i+1)!/2$ as in

	1	2	3	4	5	Σ
m=1	1					1
2	1	2				3
3	2	6	4			12
4	6	22	24	8		60
5	24	100	140	80	16	360

When we multiply the Stirling numbers with powers $2^(i-j)$, then the row sums give $(2i-1)!/i!2^i$ or the products of m odd numbers $1 \times 3 \times 5 \times \ldots$. The factorial is depleted of even numbers.

	1	2	3	4	5	Σ
m=1	1					1
2	2	1				3
3	8	6	1			15
4	48	44	12	1		105
5	384	400	140	20	1	945

If the columns are multiplied with column ± signs alternatively, then in the first case the row sums are zero, except $m = 2$, and in the second they give lower odd factorials[1].

8.4 Euler Polynomials

The Euler numbers (Table 7.5) classify naive matrices according to the number k of nonzero columns. We can add the new element in the last row into k already occupied columns or we can put it into $(n-k)$ unoccupied columns. It is clear that the index k can not be here 0, as it was convenient at permutation matrices.

The Euler numbers are just the beginning of a series of polynomials $E_n(r)$ where $r = 1$. The Euler polynomial $E_n(2)$ is obtained by multiplying each previous column, except the first one, by powers of 2 as if elements of naive matrices in columns had signs ± and all combinations of signs were acceptable and finding the differences of the consecutive columns. The

[1]This is given without a proof. The proof will be shown later.

Table 8.1: Euler polynomials $E_n(2)$

k	1	2	3	4	5	6	Σ
n=1	1						1
2	1	2					3
3	1	8	4				13
4	1	22	44	8			75
5	1	52	264	208	16		541
6	1	114	1208	2416	912	32	4683

resulting numbers are given in the Table 8.1 which elements are differences of the matrix, obtained by multiplying the matrix of the Euler numbers with the matrix of m-th powers of 2:

				1	1	1	1	1	-1		
					2	2	2		1	-1	
						4	4			1	
							8				1
1				1	1	1	1	1			
1	1			1	3	3	3	1	2		
1	4	1		1	9	13	13	1	8	4	
1	11	11	1	1	23	67	75	1	22	44	8

The row sums of the Table 7.1 are interesting. They are generated directly by the formal equation

$$[1 + E(k)]^m = 2E(m) \; , \tag{8.4}$$

where $E(k)^i = E(i)$ and $E(0) = 1$. Then

$$2E(1) = \binom{1}{0} E(0) + \binom{1}{1} E(1) \; , \tag{8.5}$$

from it E(1) = 1 and so on. These numbers will appear later many times, as differences of complete plane simplices, but here they appear as an extension of the factorial simplices or as the product of the matrix of the Euler numbers with the diagonal matrix of powers 2^{j-1}.

8.5 Mac Mahon Numbers

This statistics (Table 7.6) counts naive matrices according their moments (or counting empty spaces in all rows to the first unit element) which are

obtained by multiplying naive matrices by the diagonal matrix with indices $\Delta(j-1)$. The recurrence is obtained from lesser matrices by repeating n times terms of the next to the last row of the table of Mac Mahon numbers with equal or increased moments. If the unit element in the last row is placed in the first column the moment remains the same, and it is increased to $(n-1)$ if it is placed in the n-th column. From each matrix with $(n-1)$ rows n new matrices are produced. Their moments are counted as For example: for $n = 5$:

Moments:	0	1	2	3	4	5	6	7	8	9	10
4 rows and columns	1	3	5	6	5	3	1				
Term 6							1	1	1	1	1
Term 5						3	3	3	3	3	
Term 4					5	5	5	5	5		
Term 3				6	6	6	6	6			
Term 2			5	5	5	5	5				
Term 1		3	3	3	3	3					
Term 0	1	1	1	1	1						
Mac Mahon numbers	1	4	9	15	20	22	20	15	9	4	1

This scheme gives automatically the factorials.

8.6 Stirling Numbers of the Second Kind

When we look at the sequence *aac*, we see that in its matrix the column b is missing. We excluded such strings as faulty Young tables or convolutions, but the strings as *abb* were also not allowed, where b appeared twice, and a only once. There is a difference in these two cases: If not all columns are occupied successively, we jump in the space over some positions. Therefore we will now count all naive matrices in the lower triangular form with successively occupied columns. Their recurrence is

$$s_{11} = 1;\ s_{ij} = j s_{i-1,j} + s_{i-1,j-1}\ . \tag{8.6}$$

It is possible to place a new element into j already occupied columns and there is only one possibility, how to increase the number of occupied columns. In this way we obtain a table of numbers that are known as the *Stirling numbers of the second kind* (Table 8.2).

Stirling numbers of the second kind are inverses of the Stirling numbers of the first kind. Similarly Stirling numbers of the first kind are inverses of the Stirling numbers of the second kind. The inverse is obtained when

Table 8.2: Stirling numbers of the second kind.

j	1	2	3	4	5	6	Σ
n=1	1						1
2	1	1					2
3	1	3	1				5
4	1	7	6	1			15
5	1	15	25	10	1		52
6	1	31	90	65	15	1	203

one from the two matrices (Table 7.2 and Table 8.2) is multiplied with alternating signs $(-1)^{i-j}$.

Stirling found numbers bearing his name when he compared powers of any number t with its *factorial moments* $(t)_k$ defined by the products

$$(t)_k = t(t-1)...(t-k+1)\ . \tag{8.7}$$

Stirling numbers of the first kind transform sums and differences of powers into factorial moments as in: $(4)_3 = 24 = 2 \times 4 - 3 \times 16 + 1 \times 64$. The Stirling numbers of the second kind invert sums of factorial moments into powers as in : $4^3 = 64 = 1 \times 4 + 3 \times 12 + 1 \times 24$. Here t can substitute rational (irrational) numbers.

The row sums of Stirling numbers of the second kind, which count naive matrices in the lower triangular form with successively occupied columns are obtained as selfgenerating function

$$S(n) = (S_i + 1)^{n-1},\ \text{where}\ S^i = S_i\ . \tag{8.8}$$

Another possibility, how to these sum is with help of the Bell triangle, also known as Aitken's array or Pierce triangle:

	1	2	3	4	5	Σ
m=1	1					1
2	1	2				3
3	2	3	5			10
4	5	7	10	15		37
5	15	20	27	37	52	151

The first column is given by diagonal elements, which are obtained as sums of two elements in the preceeding column. Another possibility is by using matrices. The first one is the matrix having the numbers S(n) on the diagonal and under it, the second one is the matrix of binomial coefficients $\mathbf{B}^{\mathrm{T}}$. Then the numbers S(n) are obtained on the diagonal of the product.

We can multiply the product further by the diagonal index matrix to obtain moments

				1	1	1	1	1			
					1	2	3		2		
						1	3			3	
							1				4
1				1	1	1	1	1	2	3	4
1	1			1	2	3	4	1	4	9	16
1	1	2		1	2	5	10	1	4	15	40
1	1	2	5	1	2	5	15	1	4	15	60

Notice that the matrix of the Stirling sums begins with two 1 and on the diagonal of the product is only one 1 and then immediately the higher sums follow. In the final product the sums are multiplied with corresponding powers. Since we labelled the matrix of binomial coefficients as $\mathbf{B}^{\mathrm{T}}$, we can consider its product with the diagonal matrix $\Delta(i)$ as the logarithmic difference $d(log\mathbf{S})$, similarly as it was derived in Sect. 6.4. The inverse matrix of sums of Stirling numbers of the second kind has elements:

$$s^{-1}_{jj} = 1;\ s^{-1}_{j-1,j} = -[S_{j-1}/S_j];\ s_{ij} = 0,\ \text{otherwise}\,. \tag{8.9}$$

We have shown one relation between Stirling numbers of the second kind and the binomial coefficients. But there appears still another relation. The difference of two successive sums of Stirling numbers is generated again by a binomial

$$\Delta_n S_n = S_n - S_{n-1} = (S_{k+1} + 1)^{n-2}\,, \tag{8.10}$$

where we put again $S^k = S_k$. For example: $S_6 - S_5 = 1 \times 1 + 4 \times 2 + 6 \times 5 + 4 \times 15 + 1 \times 52 = 151$.

The Stirling numbers of the second kind are defined by the formal relation

$$\Delta^1_n(m)^n = m^{n-1}[(1 + 1/m)^{n-1} + (1 + 1/m)^{n-2} \ldots + (1 - 1/m)^0]\,. \tag{8.11}$$

Inserting $m = 1$, we obtain numbers $S(n, 2) : \Delta_m 1^n = 1^n[2^{n-1} + 2^{n-2} \ldots + 2^0]$. The other numbers are derived by the relation

$$\Delta^m 1^n = (m + 1)\Delta^m 1^{n-1} + \Delta^{m-1} 1^{n-1} \tag{8.12}$$

under the condition $\Delta^0 1^0 = 1$.

The differences of the Stirling numbers of the second kind

Table 8.3: Differences of Stirling numbers of the second kind

j	1	2	3	4	5	6	Σ
m=1	1						1
2	0	1					1
3	0	2	1				3
4	0	4	5	1			10
5	0	8	19	9	1		37
6	0	16	65	55	14	1	151

Table 8.4: Substirlings

n	0	1	2	3	4	5	Σ
n=0	1						1
1	0	1					1
2	1	0	1				2
3	1	3	0	1			5
4	4	4	6	0	1		15
5	11	20	10	10	0	1	52

$$S(m,n) - S(m-1,n) = \Delta^{n-1}2^m \tag{8.13}$$

form the Table 8.3.

For example: $\Delta^2 2^6 = 8[(3/2)^3 + (3/2)^2 + (3/2)^1 + (3/2)^0] = 65$. This number counts naive matrices in lower triangular form with 3 occupied columns and 6 rows, obtained from 15 matrices counted by $S(5,2)$ by adding the unit element into the third column and 2×25 matrices counted by $S(5,3)$ increased by adding the new unit element into one from the two first columns.

8.7 Substirlings

In analogy with subfactorials defined in Sect. 7.6, we introduce numbers which we will call *substirlings*. They count naive matrices in lower triangular form with successively occupied columns in another order, according to the number of columns containing just one nonzero element. We have shown that such orbits are differences of plane simplices, therefore also now these matrices form differences. Their matrix is in the Table 8.4 which is completed by the row and column indexed from 0. For example: $s_{40} = 4$ counts $\mathbf{N}:\ a^4,\ a^2b^2,\ abba,\ abab.$

Now again the binomial coefficients appeared here as generating factors. Naive matrices without any columns containing only one unit element are combined with n columns with only one unit element and the result gives the matrix elements. Therefore the sums of Stirling numbers of the second kind are obtained by the formal binomial:

$$S_n = (s_{n0} + 1)^n, \text{ where } s^k = s_{n0} \, . \tag{8.14}$$

Another possibility, how the Stirling numbers of the second kind are obtained, is the direct count of the corresponding matrices arranged according to the powers of a. For example:

$$\begin{array}{ccc} a & & \\ ab & aa & \\ abb,\ abc; & aab,\ aba; & aaa \end{array} \, .$$

We get a table where the naive matrices are arranged according to the rows containing the symbol a. Again these matrices are obtained by multiplying lower matrices (without this symbol) by binomial coefficients, showing combinatorial possibilities:

j	1	2	3	4	5	6	Σ
m=1	1						1
2	1	1					2
3	2	2	1				5
4	5	6	3	1			15
5	15	20	12	4	1		52
6	52	75	50	20	5	1	203

Substirlings are in their turn the sums of the *associated Stirling numbers of the second kind* which count naive matrices in the lower triangular form without empty columns, having column sums at least $m_k = 2$. Their recurrence is given by the formula

$$a_{ij} = j a_{i-1,j} + (i-1) a_{i-2,j-1} \tag{8.15}$$

and their values are given in Table 8.5.

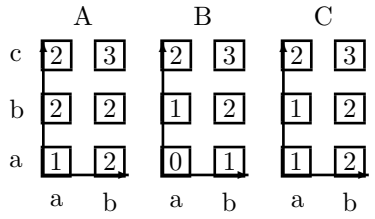
8.8 Space of Four Statistics

We have mapped naive matrices in the lower triangular form on points of rectangular n dimensional parallelepipeds. These points are classified by three different statistics Euler, Stirling and Mac Mahon, in 3 directions. They split the space behind the Euclidean one. These statistics distribute

Table 8.5: Associated Stirling numbers of the second kind

j	0	1	2	3	Σ
m=0	1				1
1	0	0			0
2		1			1
3		1			1
4		1	3		4
5		1	10		11
6		1	25	15	41

Figure 8.1: Three statistics. A is Euler's, B is Mac Mahon's, and C is Stirling's. Arranged strings are a, horizontal symbol, vertical symbol.



points differently, as it is shown on the Fig. 8.1 for three dimensional space and on the scheme for four dimensional space (Table 8.6).

We have compared three statistics, but a fourth one appeared here and that is on the diagonal of the Stirling and Euler statistics. The Euler numbers divide naive matrices in the lower triangular form according to the number of occupied columns. The Stirling numbers of the second kind count matrices with rows occupied consecutively and, these matrices appear on the crossection of both statistics. The Euler statistics splits 6 naive matrices in the lower triangular form with one unit element on the diagonal in three groups the Stirling numbers evaluate differently naive matrices in the lower triangular form with two unit elements on the diagonal.

I do not know what you think about these coincidences. The Euclidean space is full of surprises. It seems to be alive and if we try to analyze it, new layers appear just on elementary levels. Euclides was wrong when he told to the king Ptolemaos that there was no other way to his space then his axioms. Different combinatorial functions lead through this maze as the Ariadne's thread.

Table 8.6: Scheme of four statistics for $\mathbf{N}_4$ in the lower triangular form.

		Stirling I				Mac Mahon						
		6										
			8									
			3	6								
					1	1	3	5	6	5	3	1
Euler	1	**1**										1
	11	4	**7**					1	2	5	3	
	11	1	4	**6**			3	4	4			
	1				**1**	1						
		6	11	6	1	↖			Stirling II			

Chapter 9

Combinatorics of Natural Vectors

9.1 The Binomial Coefficient

The *binomial coefficient* is a special case of the polynomial coefficient. This definition is invalid but it corresponds to the facts. The two dimensional space is a special case of the multidimensional space.

When a binomial, say $(a + b)$ is multiplied by itself m times and the terms in the product grouped we get For example:

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \ .$$

The first term 4 counts strings **aaab**, **aaba**, **abaa**, and **baaa**, the third term 4 counts strings **abbb**, **babb**, **bbab**, and **bbba**. The binomial coefficient is written as the number m stacked over the number k in the brackets

$$\binom{m}{k} \ . \tag{9.1}$$

The binomial coefficient is the product of three factorials $m!$, $k!_{-1}$, and $(m-k)!_{-1}$. Therefore

$$\binom{m}{k} = \binom{m}{m-k} \ . \tag{9.2}$$

The binomial coefficients have many interesting properties. For example, if n is a prime number, all elements of the binomial, except 1, are divisible by n.

Another curious property is distribution of even coefficients. If we write The Pascal triangle in the isoscele form, than even coefficients , e. g. $\binom{8}{k}$, propagate as the isoscele triangles, and there appear such subsidiary triangles.

There are many relations concerning the binomial coefficients. One from them:

$$\binom{n}{2} + \binom{n+1}{2} = n^2 \; . \tag{9.3}$$

9.2 The Polynomial Coefficient

A partition of the number m into n parts is an n dimensional vector $\mathbf{m}$ which elements are ordered in the decreasing order, $m_j \geq m_{j+1}$. From this vector all other vectors on the given orbit can be generated when its elements are permuted by the unit permutation matrices acting on the partition vector from the right. These vectors correspond to the scalar products of naive matrices $\mathbf{N}$ with the unit vector rows $\mathbf{J}^{\mathrm{T}}$ or to the quadratic forms $\mathbf{N}^{\mathrm{T}}\mathbf{N}$, because

$$\mathbf{J}^{\mathrm{T}}\mathbf{N} = \mathbf{J}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{N} \; . \tag{9.4}$$

There are $n!$ permutation matrices, but not as many permuted vector columns, when some elements of the vector row are not distinguishable. Vectors with equal length m_k point to the sphere and if rotated, their permutations are undistinguishable. If all elements of the vector are equal, then no permutations have any effect on the partition vector.

We divide vector elements into two groups, one with all zero elements, that is n_0 elements, and the second group with all remaining $(n - n_0)$ elements. The number of possible permutations will be reduced from the factorial $n!$ to the binomial coefficient $\binom{n}{n_0}$, or $n!/n_0!(n - n_0)!$.

In the next step we single out from the second group the vectors with the length 1, their number is n_1. All other vectors will be counted by the third term $(n - n_0 - n_1)$, and corresponding permutations by the binomial coefficient $(n - n_0)!/n_1!(n - n_0 - n_1)!$. In this way we proceed till all possible values of m_k are exhausted. If some $n_k = 0$, then conveniently $0! = 1$ and the corresponding term is ineffective. At the end we obtain a product of binomial coefficients:

$$\left(\frac{n!}{n_0!(n-n_0)!}\right)\left(\frac{(n-n)!}{(n-n_0-n_1)!}\right)\left(\frac{(n-n_0-n_1)!}{n_2!(n-n_0-n_1n_2)!}\right)\cdots$$

$$\left(\frac{(n-\sum_{k=0}^{m-1} n_k)!}{n_m!0!}\right) \quad (9.5)$$

Equal factorials appear consecutively as dividends and divisors. When they cancel, the *polynomial coefficient* remains from the product of binomial coefficients

$$n!/\prod_{k\geq 0} n_k! \,. \quad (9.6)$$

Lets call it the *polynomial coefficient for n permutations* because it is obtained by permuting n columns. Later we will construct another polynomial coefficient for permutations of rows of naive matrices.

We limited the index k by the lower limit 0. The coefficient could be used actually also for vectors with negative elements. The numbers n_k of equal vectors are always positive even if the vectors themselves are negative. We count by the polynomial coefficient (9.2) points on the partition orbits of the positive cone of the n dimensional space.

Please note the importance of this step. We know the vector **m** exactly, but we replace it by the corresponding partition. All points on the given orbit are considered to be *equivalent*. Replacing the vector **m** by the partition is a logical abstraction. We can proceed further, the partition is compared with an analytical function and the orbit is described by a density distribution.

9.3 Simplex Sums of Polynomial Coefficients

Now it is possible to apply again partition schemes and to study sums of polynomial coefficients on all orbits of plane simplices, it means all natural n dimensional vectors with constant sums m.

The overall sum is known in combinatorics as the distribution of m undistinguishable things (objects) into n boxes. It is counted by a binomial coefficient

$$\sum_{k\geq 0} n!/\prod n_k! = \binom{m+n-1}{m} = \binom{m+n-1}{n-1} \,. \quad (9.7)$$

Both binomial coefficient are in reality different forms of one coefficient. Most easily this binomial coefficient is obtained by following all possibilities distributing m things into a row of $(n-1)$ bars (the objects of the second kind) representing dividing walls of compartments. There are $(m+n-1)$

objects of two kinds and the result is simply given by a binomial coefficient. Who is not satisfied with this explanation, can prove (9.3) by the *full induction*.

We tested the relation at simple cases and it functioned well. Thus we suppose that it is true for all n dimensional vectors with $(m-1)$ elements and to all $(n-1)$ dimensional vectors with m elements. We use the proposition for counting points with sums m in n dimensions. These points we divide into two distinct subsets. In one subset will be all points having as the last element 0. Clearly, they all are in $(n-1)$ dimensional subspace and they are counted by the binomial coefficient $\binom{m+n-2}{m}$. In the second subset vectors having as the last element at least 1 are counted. They are obtained from partitions of $(m-1)$ things into exactly n parts by adding 1 to the first element. This addition does not change the corresponding number of points $\binom{m+n-2}{m-1}$. The result is formed by a sum of 2 binomial coefficients and verified by calculations

$$\left(\frac{(m+n-2)!}{m!(n-2)!}\right)+\left(\frac{(m+n-2)!}{(m-1)!(n-1)!}\right)= \\ \left(\frac{(m+n-2)![(n-1)+m]}{m!(n-1)!}\right)=\binom{m+n-1}{m}. \tag{9.8}$$

As was said that we will not be interested in vectors with negative values, but it is instructive to show results according to the lover limit of the value r, which appears as the parameter $(1-r)$ of the term n in the binomial coefficients. The value r can be considered as differentiating of the simplex

Lower limit	-1	0	1	2
Points on the simplex	$\binom{m+2n-1}{n-1}$	$\binom{m+n-1}{n-1}$	$\binom{m-1}{n-1}$	$\binom{m-n-1}{n-1}$

The binomial coefficients $\binom{m+3-1}{m}$ are known as triangle numbers. They count points of 3 dimensional planes which are regular triangles.

9.4 Differences of Normalized Simplices

We have counted points of the plane simplices directly, now we apply partition schemes and insert into them polynomial coefficients, similarly as we

Table 9.1: Van der Monde identity

k	1	2	3	4	5	6	Σ
m=1	1						1
2	2	1					3
3	3	6	1				10
4	4	18	12	1			35
5	5	40	60	20	1		126
6	6	75	200	150	30	1	462

did for the cycle indices in Chapt. 7. We limit ourselves to cases when $m = n$. As an example, we give the scheme for $m = n = 6$:

n	1	2	3	4	5	6
m=6	6					
5		30				
4		30	60			
3		15	120	60		
2			20	90	30	
1						1
$\sum$	6	75	200	150	30	1

In the first column vertices of the plane simplex are counted, in the second column points on 2 dimensional edges, in the third column points of its 3 dimensional sides. Only the last point lies inside of the 6 dimensional plane, all other 461 points are on its borders.

This is a rather surprising property of the high dimensional spaces that the envelope of their normal plane simplices is such big. But we can not forget, that usually $m \gg n$ and then there are more points inside than on the border.

Column sums of consecutive normalized plane simplices can be arranged into Table 9.1 which rows are known as the *Van der Monde identity*.

The elements in each row can be written as products of two binomial coefficients, For example: $75 = (6!/4!2!) \times (5!/4!1!)$. This is a special case of the identity

$$\sum_{i=0}^{m-k} \binom{m}{k+i} \binom{m-k}{i} = \binom{m+k}{m} = \binom{m+n-1}{n-1} . \tag{9.9}$$

The sum of products of two binomial coefficients can be written as a formal power of a binomial

$$\left(\binom{m}{i}+1\right)^n = \binom{m+n}{m}. \tag{9.10}$$

This relation counts the points of the plane simplices in one direction. Its special case is the *Wallis identity* for $m = n$:

$$\sum_{i=0}^{n/2} \binom{n}{i}^2 = \binom{2n}{n}. \tag{9.11}$$

We interpret it as the simplex in which the first vector is rooted and only $(n-1)$ other vectors are permuted. For example:

Orbits	4000	3100	1300	2200	2110	1210	1111	$\sum$
Points	1	3	3	3	3	6	1	20
Counts	1	9			9		1	20

9.5 Difference According to Unit Elements

When we arrange the partition scheme according to the number of unit vectors n_1, we obtain a difference of the plane simplex. For example for $m = n = 5$:

n_1	0	1	2	3	4	5
m=5	5					
4		20				
3	20		30			
2		30		20		
1						1
$\sum$	25	50	30	20	0	1

The resulting column sums of polynomial coefficients are tabulated in Table 9.2.

The numbers b_{i0} are formed by vectors without any unit elements. They can be called *subplane numbers*, because they generate the number of points of the normal plane simplex by multiplying with binomial coefficients:

$$(b_i+1)^m = \binom{m+n-1}{m}. \tag{9.12}$$

They are $(n-k)$ dimensional vectors without unit elements but with zero elements. Their $(n-k)$ elements are combined with k unit elements. When $m \neq n$, then these relations are more complicated. Corresponding

Table 9.2: Unit elements difference

n_1	0	1	2	3	4	5	6	Σ
m=0	1							1
1	0	1						1
2	2	0	1					3
3	3	6	0	1				10
4	10	12	12	0	1			35
5	25	50	30	20	0	1		126
6	71	150	150	60	30	0	1	462

subplane numbers are obtained by calculations of partitions without unit parts. The beginning of the table is

n	0	1	2	3	4	5	6
m=0	1	1	1	1	1	1	1
1	0	0	0	0	0	0	0
2	0	1	2	3	4	5	6
3	0	1	2	3	4	5	6
4	0	1	3	6	10	15	21
5	0	1	4	9	16	25	36
6	0	1	5	13	26	45	71

Its values $b(i, j)$ for small m are:

- b(0,n) = 1;
- b(1,n) = 0;
- b(2,n) = $\binom{n}{1}$;
- b(3,n) = $\binom{n}{1}$;
- b(4,n) = $\binom{n}{1} + \binom{n}{2} = \binom{n+1}{2}$;
- b(5,n) = $\binom{n}{1} + 2\binom{n}{2} = n^2$;
- b(6,n) = $\binom{n}{1} + 3\binom{n}{2} + 3\binom{n}{3} = (n^3 - n)/2$.

The subplane numbers appear here on the diagonal. An example of their application for $m = 4,\ n = 6$:

$$21 + 6 \times 5 + 15 \times 4 + 20 \times 0 + 15 \times 1 = 126 = \binom{9}{4}.$$

Vectors without unit elements are combined with unit vectors.

9.6 Differences According to One Element

In partition schemes the points are counted in spherical orbits. We orient the plane simplex in the direction of one vector and then differentiate the plane according to only one specific vector **x**. It can be shown on the 2 dimensional complex:

m_a		0	1	2	3	4	5	Orbit
Points	0	*	*	*	*	*	*	0,m
	1		*	*	*	*	*	1,(m-1)
	2			*	*	*	*	2,(m-2)
	3				*	*	*	3,(m-3)
	4					*	*	4,(m-4)
	5						*	5,(m-5)
Number		1	2	3	4	5	6	

The 2 dimensional complex forms a 3 dimensional simplex and its points for different values of the vector **a** are counted by column sums. It is similar to a situation as when points of the (n-1) dimensional complex are counted for different values of m, m_k going from 0 to m. The points are counted by binomial coefficients $\binom{m+k-2}{k}$. For example: for $n = m = 7$:

m_k	0	1	2	3	4	5	6	7
Binomial coefficient	792	462	252	123	56	21	6	1

We obtain the identity

$$\sum_{k=0}^{m} \binom{m+k-2}{k} = \binom{m+n-1}{m} . \tag{9.13}$$

Now we introduce another difference.

9.7 Difference $\Delta(n)$ of Plane Simplices

Till now zero elements were permuted with the other elements. We exclude the zero element and count only existing (nonzero) vectors and not virtual vectors. It means, that we count consecutively all k dimensional vectors ($k = 1$ to n) with constant sums m. If we draw the tetrahedron (Fig. 9.1), then the counted set of points is formed by one vertex, one edge without the second vertex, the inside of one side and by the four dimensional core. In combinatorics these vectors are known as *compositions*. They can be arranged onto partition schemes. For $m = 5$ we get:

Figure 9.1: Difference of the plane simplex. It is formed by one vertex, one incomplete edge, one incomplete side, etc.

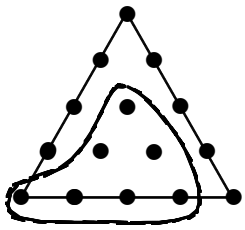


Table 9.3: Binomial coefficients (matrix **B**).

k	1	2	3	4	5	6	Σ
m=1	1						1
2	1	1					2
3	1	2	1				4
4	1	3	3	1			8
5	1	4	6	4	1		16
6	1	5	10	10	5	1	32

n	1	2	3	4	5	Σ
m=5	5					1
4		41;14				2
3		32;23	311;131;113,			5
2			221;212;122;	2111;1211;1121;1112		7
1					11111	1
$\sum$	1	4	6	4	1	16

The column sums of the normal plane simplices give Table 9.3.

Both indices in Table 9.3 were decreased by one, to obtain the true binomial coefficient $\binom{k-1}{m-1}$. We had difficulties with the binomial coefficient before, when it appeared as $\binom{m+n-1}{m}$. In that case they fill the matrix otherwise, as in the Table 9.4:

In both tables of binomial coefficients, their elements were obtained similarly that is a sum of two neighbors, the left one and the upper one, except that in the Table 9.3 the left element is added only if $j \geq i$.

Recall the transactions with partitions and their counts according to the lower allowed limit of parts. Here a similar shift of values of Tables

Table 9.4: Matrix $\mathbf{B}\mathbf{B}^{\mathrm{T}}$ of binomial coefficients.

k	1	2	3	4	5	6
m=0	1	1	1	1	1	1
1	1	2	3	4	5	6
2	1	3	6	10	15	21
3	1	4	10	20	35	56
4	1	5	15	35	70	126
5	1	6	21	56	126	252

9.3 and 9.4 occurred, but the operation is done by the matrix of binomial coefficients $\mathbf{B}^{\mathrm{T}}$. We permute k nonzero elements with $(n-k)$ zero elements and from a part of the plane simplex we obtain the whole simplex. Therefore this part is the difference $\Delta(n)$. Because there are more differences, this is the difference according to the number of vectors n. In the tetrahedron one vertex is multiplied four times, one edge six times, one side four times, and the inside only once.

Now we can return to Table 9.3. Its elements have the recurrence

$$b_{11} = 1;\ b_{ij} = b_{i-1,j} + b_{i-1,j-1}\ . \tag{9.14}$$

They are generated by the binomial

$$(1_i + 1)^m = 2^m\ . \tag{9.15}$$

We already formulated the recurrence formula of the Table 9.4 in (9.8). Notice that the elements of the Table 9.4 are sums of all elements of its preceding row or column, which is the consequence of the consecutive applications of (9.8).

The inverse matrix $\mathbf{B}^{-1}$ to the matrix $\mathbf{B}$ is obtained from the formal binomial

$$(1_i - 1)^m = 0\ . \tag{9.16}$$

It is just the matrix $\mathbf{B}$ which elements are multiplied by alternating signs $(-1)^{j-i}$.

9.8 Difference $\Delta(m)$

When we arranged the vector compositions in the table, we treated only its column sums. There are also row sums which count compositions classified

Table 9.5: Composition of vectors with m parts

n	1	2	3	4	5	6	7	8	9
m =1	1	1	1	1	1	1	1	1	1
2		1	2	4	7	12	20	33	54
3			1	2	5	11	23	47	94
4				1	2	5	12	25	59
5					1	2	5	12	28
6						1	2	5	12
							1	2	5
								1	2
									1
$\sum$	1	2	3	8	16	32	64	128	256

according to the greatest vector m_k. The consecutive results for $n = m$ can be arranged into the Table 9.5

The elements c_{ij} of the Table 9.5 are sums of the polynomial coefficients counting compositions. Their column sums are 2^{j-1}. For $j \leq j/2$ the elements c_{ij} become constant. For example

Orbit	Number of compositions
$m-3,3$	2
$m-3,2,1$	6
$m-3,1^3$	4
$\sum$	12.

For $i = 2$ the elements c_{2j} are sums of the binomial coefficients and their recurrence is

$$c_{2j} = \sum_{k=1}^{j/2} \binom{j-k}{k} = 2c_{2,j-1} - c_{2,j-3} \ , \tag{9.17}$$

where k is the number of 2.

9.9 The Second Difference – the Fibbonacci Numbers

When we admit as the smallest element 2, we get the Table 9.6 of points of truncated plane simplices. Its row sums are known as the *Fibbonacci numbers*. In a medieval arithmetic book they appeared as the answer on a number of rabbit pairs in the consecutive litters.

Table 9.6: Fibbonacci numbers

n	1	2	3	Σ
m=2	1			1
3	1			1
4	1	1		2
5	1	2		3
6	1	3	1	5
7	1	4	3	8

The vectors counted for $m = 7$ are: 7; 52, 25, 43, 34; 322, 232, 223. Notice, that the elements of the Table 9.6 are binomial coefficients shifted in each column for 2 rows. Fibbonacci numbers F_m have the recurrence

$$F_m = F_{m-1} + F_{m-2} \; . \tag{9.18}$$

The elements of the Table 9.6, f_{ij} are obtained by adding 2 to each vector with $(j-1)$ nonzero elements or 1 to the greatest element of the j dimensional vectors

$$f_{21} = 1; \; f_{ij} = f_{i-2,j-1} + f_{i-1,j} \; . \tag{9.19}$$

In each row all elements of both preceding rows are repeated which gives the recurrence of the Fibbonacci numbers.

Another way to obtain the Fibbonacci numbers is to count the compositions in which all elements are odd. We get a scarce Pascal triangle:

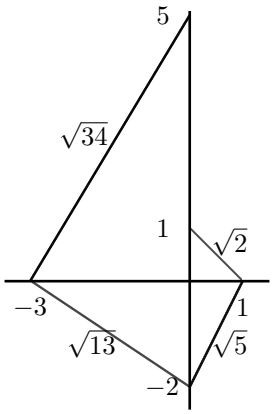
k	1	2	3	4	5	6	Σ
m=1	1						1
2	0	1					1
3	1	0	1				2
4	0	2	0	1			3
5	1	0	3	0	1		5
6	0	3	0	4	0	1	8

For example, the last row counts the compositions: 51, 15, 33; $4 \times$ (3111); 111111.

9.10 Fibbonacci Spirals

If we draw on two orthogonal axes consecutive Fibbonacci numbers, then the hypotenuses connecting consecutive points of the corresponding right

Figure 9.2: Fibbonacci spiral. Squared hypotenuses of right triangles with consecutive Fibbonacci legs are odd Fibbonacci numbers.



triangles are square roots of the squared Fibbonacci numbers F_{2k+1} (Fig. 9.2). This implies the identity

$$F_{2k+1} = F_{k+1}^2 + F_k^2 \ . \tag{9.20}$$

A similar identity is obtained for even numbers from the difference of two squared Fibbonacci numbers, For example: $F_8 = F_5^2 - F_3^2 = 21 = 25 - 4$. This difference can be written as a sum of products of the Fibbonacci numbers.

$$F_{2k} = F_{k+1}^2 - F_{k-1}^2 = F_k^2 + F_k F_{k-1} \ . \tag{9.21}$$

We decompose the higher Fibbonacci numbers consecutively and express coefficients by the lower Fibbonacci numbers as:

$$F_{2k+1} = F_2 F_{2k} + F_1 F_{2k-1} = F_3 F_{2k-1} + F_2 F_{2k-2} = \ldots \tag{9.22}$$

There appears still another formula

$$F_{n+1} F_{n-1} - F_n^2 = (-1)^a \ . \tag{9.23}$$

For example: at $n = 5 : 3 \times 8 - 25 = -1$.

This relations be formulated in the matrix form (using knowledge what the matrix determinant is) as

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n .$$

This relation leads to two things. The first one is the eigenvalues of the matrix, see later Chapters, the second one is the zero power of this matrix:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} F_1 & F_0 \\ F_0 & F_{-1} \end{pmatrix} .$$

On the diagonal the values F_{n+1} and F_{n-1} are. This fact gives a possibility to prolongate the Fibbonacci numbers to negative indices. This series must be: $1, -1, 2, -3, 5, -8, \ldots$. We obtain these numbers again as the sums of the two consecutive Fibbonacci numbers, row sums of the elements of $\mathbf{B}^{-1}$ or as the elements of their generating matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}^n .$$

Chapter 10

Power Series

10.1 Polynomial Coefficients for m Permutations

The polynomial coefficients were defined for permutations of columns of vector rows. It is clear, that such a coefficient must be applicable to transposed vector-rows, it means to vector-columns. It seems that it is not necessary to have some special coefficients for permutations of rows of vector columns, when the only difference would be, that corresponding permutation matrices acted on the vector from the left instead from the right. But a different situation appears for strings of symbols, For example: $(\mathbf{aaabbccdef})^{\mathrm{T}}$. We determine easily the number of produced strings by a polynomial coefficient 10!/3!2!2!1!1!1!. We cannot distinguish equal symbols and therefore their mutual permutations are ineffective as permutations of vectors having equal length. But this polynomial coefficient is different from the polynomial coefficient for n permutations.

The polynomial coefficient for n permutations permutes the numbers n_k of vectors having the same value (frequency) m_k. Now the appearances of individual vectors $\mathbf{j}$, counted as m_j, are permuted. It is clear from the example that some values m_j can be equal for more vectors (1 for three, 2 for two). Thus a new index k is useful (its value coincides with the number m_k itself. The number of vectors with the value m_k is n_k, and the *polynomial coefficient for m permutations* is written as

$$m!/\prod_{j=1}^{n} m_j! = m!/\prod_{k\geq 0} m_k!^{n_k};\ \text{where } m = \sum_{j=1}^{n} m_j = \sum_{k\geq 0} n_k m_k\ . \qquad (10.1)$$

The m permutations transform the sequence of symbols for example $(\mathbf{dagfabcace})^{\mathrm{T}}$, whereas n permutations act as *substitutions*, For example: $(\mathbf{abcceeefgg})^{\mathrm{T}}$. The substitution **a** into **e** was not direct, but it was a part of a cycle, moreover **g** appeared (which was not in the example) but as a column with zero elements in the alphabet matrix.

10.2 Naive Products of Polynomial Coefficients

In Chapt. 7 we studied symmetry of a special class of the naive matrices, having one unit element not only in rows but simultaneously in columns. They all go to an orbit consisting only from one point.

Now we shall find the symmetry index of two groups of cyclic permutations acting simultaneously on other naive matrices from the left and from the right:

$$\mathbf{P}_m\mathbf{N}\mathbf{P}_n\ . \tag{10.2}$$

The action of the permutation matrices from the left is counted by the polynomial coefficient for m permutations (10.1), the action of the permutation matrices from the right is counted by the polynomial coefficient for n permutations (9.1). The effect of permutations from the right is identical with the n permutations of columns of the vector-row **m** of the column sums of naive matrices:

$$\mathbf{J}^{\mathrm{T}}\mathbf{N}\mathbf{P}_n\ =\ \mathbf{m}\mathbf{P}_n\ . \tag{10.3}$$

Both actions are independent and therefore the final result is just the product of both coefficients

$$\sum(n!/\prod_{k\geq 0} n!)(m!/\prod_{k\geq 0} m_k^{n_k}!) = n^m\ . \tag{10.4}$$

The sum is made over all partition orbits. It is a special case of the Newton polynomial formula, where coefficients having the same partition structure are counted together by the polynomial for n permutations[1]. The final result is obtained easily. Exactly n columns are placed in each row, where one element can be put. Individual choices in m rows are independent and therefore they multiply.

The right side result is known as the *distribution of m distinguishable objects into n boxes*. Objects are distinguished by their index i. This index

[1]The identity is known in physics as the Polya-Brillouin statistics. But Brillouin and others did not recognized its key importance.

Table 10.1: Power series sequence

k	1	2	3	4	5	6	Σ
m=1	1						1
2	2	2					4
3	3	18	6				27
4	4	84	144	24			256
5	5	300	1500	1200	120		3125
6	6	930	10800	23400	10800	720	46656

is lost in a sum. The distinguishability is not a property of things but circumstances[2]. All 1 in naive matrices are identical, only their positions vary. If they were different, it were necessary to introduce a third index, which gives another statistics (see later).

The difference against the cycle index (Equation 7.15) is the second factorial $m!$ and factorials of m_k instead their first powers. When we use (10.2) for the partition 1^m we obtain $(n!/n_1!)(m!/1!^{n_1}) = m!$. The cycle index splits the S_m group according to the cycle structure.

10.3 Differences in Power Series

When we arrange polynomial coefficients into partition schemes we obtain again column sums as for $m = n = 6$:

k	1	2	3	4	5	6	Σ
m =6	6						6
5		180					180
4		450	1800				2250
3		300	7200	7200			14700
2			1800	16200	10800		18800
1						720	720
Σ	6	930	10800	23800	10800	720	$46656 = 6^6$

From consecutive schemes we obtain Table 10.1.

In Table 10.1, only the first column and the row sums are clearly connected with m and n^m. Moreover there appear factorials but other elements grow too fast to be analyzed directly. But all elements are divisible by m.

[2]This has an important philosophical consequence. In previous century, a question was disputed if the microparticles are distinguishable or not. But the notion of distinguishability was ill defined.

Table 10.2: Differences $\Delta^n \mathbf{0}^m$.

n	0	1	2	3	4	5	6	$\Delta^n \mathbf{0}^m$
m=0	1							1
1		1						1
2		1	2					3
3		1	6	6				13
4		1	14	36	24			75
5		1	30	150	240	120		541
6		1	62	540	1560	1800	720	4683

In this way the Table 10.1 is decomposed into the direct product of two matrices. One from them is the matrix of binomial coefficients $\binom{m}{k}$. This is the matrix $\mathbf{B}^{\mathrm{T}}$. The other one is the matrix of differences $\Delta^n \mathbf{0}^m$:

We already encountered the row sums $\Delta^n \mathbf{0}^m$ in Table 9.1 as the Euler polynomials $E_n(2)$. These numbers count the naive matrices in lower triangular form with nonempty columns according to the number of columns. For example: for $m = n = 4$:

n	1	2			3			4
Basic string	aaaa	aaab	aabb	abbb	aabc	abbc	abcc	abcd
Permutations	1	4	6	4	12	12	12	24
Counts	1	14			36			24

The binomial coefficients $\binom{m}{k}$ permute nonzero columns with zero columns. The table of differences has the first row and column indexed with zero indices. But they contain, except the element 1_{00}, only zeroes. This eliminates the effect of the first row of the binomial matrix in the direct product.

The recurrence in the Table 10.2 is simple

$$m_{00} = 1;\ m_{ij} = j(m_{i-1,j-1} + m_{i-1,j})\ . \tag{10.5}$$

In each column we have j possibilities how to add the new element. Either it is added to the occupied columns, or it is added into a new column. Then other column are only shifted without permuting.

Table 10.1 is the direct product $c_{ij} = a_{ij} \times b_{ij}$. When we find the normal product $(\Delta^n 0^m)\mathbf{B}^{\mathrm{T}}$, we obtain the matrix which elements are powers j^i. For example

Table 10.3: Differences of power series

k	0	1	2	3	4
m=0	1				
1	0	1			
2	1	1	1		
3	14	3	1	1	
4	181	13	3	1	1

				1	1	1	1	1
					1	2	3	4
						1	3	6
							1	4
1				1	1	1	1	1
	1				1	2	3	4
	1	2			1	4	9	16
	1	6	6		1	8	27	64

Even the Table 10.2 is not an elementary one. It can be decomposed again into the matrix of the Stirling numbers of the second kind (Table 8.2) and the diagonal matrix of factorials $\Delta(j!)$ which multiply the Stirling matrix from the right. The Stirling numbers of the second kind count naive matrices in the lower triangular form. This condition assures that all columns form a base for column permutations, when the restriction of the lower triangular form is removed.

In another arrangement, we can form the table of finite differences as in Table 10.3.

In the zero column are counted strings of the simplex which are not in its difference. The elements in other columns are consecutive differences. For example: the elements in $d_{30} = 14$ are: b^3, c^3, b^3, $3b^2c$, $3bc^2$, $3a^2c$, $3ac^2$. The column indices correspond to the powers of the first index, For example: $d_{41} = 13 = ab^3 + 3ab^2c + 3abc^2 + 6abcd$, $d_{42} = 3 = a^2b^2 + 2a^2bc$. When we multiply this matrix with the transposed matrix of binomial coefficients $\mathbf{B}^{\mathrm{T}}$, we get on the diagonal of the product corresponding powers n^n. The binomial coefficient permutes the first vector with other already permuted vectors.

10.4 Operator Algebra

We used the operator notation many times. Now we shall explain its notation. There exist the *identity function* E and the *difference function* Δ.

Moreover there are formal powers 0^n. These functions are defined reciprocally as

$$\Delta^m 0^n = [E^m 0^n - 1]^m = \sum_{j=0}^{m} \binom{m}{j} (-1)^j (m-j)^m \ . \tag{10.6}$$

This gives for the corresponding matrix elements sums of powers of the index m:

- $\Delta^m 0^1 = 1 \times 1^m$,
- $\Delta^m 0^2 = 1 \times 2^m - 2 \times 1^m$,
- $\Delta^m 0^3 = 1 \times 3^m - 3 \times 2^m + 3 \times 1^m$.

We calculate for n=3:

$$\begin{array}{lll} \Delta^m\ 0^3 & \text{m=1=} & 1{\times}3 - 3 \times 2 + 3 \times 1 \quad =0, \\ & \text{m=2=} & 1{\times}9 - 3 \times 4 + 3 \times 1 \quad =0, \\ & \text{m=3=} & 1{\times}27 - 3 \times 8 + 3 \times 1 \quad =6, \\ & \text{m=4=} & 1{\times}81 - 3 \times 16 + 3 \times 1 \quad =36\ . \end{array}$$

The original function is restored by the product of $\Delta^m 0^n$ with the matrix of binomials. This corresponds to the formal equation

$$n^m = E^m\ 0^n = (1 + \Delta^m\ 0^n)^m \ . \tag{10.7}$$

The row sums of the Table 10.2 taken with alternating signs (the difference of even and odd columns) gives $(-1)^i$. Let suppose that this is true for some row. The elements of the next row are just multiplied sums of the preceding row:

$$d_{ij} = j(d_{i-1,j-1} + d_{i-1,j}) \ . \tag{10.8}$$

When we make the difference $d_1 - 2(d_1 + d_2) + 3(d_2 + d_3) - \ldots = -d_1 + d_2 - d_3 \ldots$, we get the elements of the preceding row with the other signs which sum was +/-1.

10.5 Differences dx and Sums of n^m

The power n^m is the binomial, if we write n as a sum $n = (n-1)+1$. Then

$$n^m = [(n-1)+1]^m = \sum_{k=0}^{m} \binom{m}{k} (n-1)^k \ . \tag{10.9}$$

For example: $3^4 = (1\times1+4\times2+6\times4+4\times8+1\times16) = 81$. The terms of the binomial are differences of the number of strings of the plane simplices according to one vector (this vector must have the prescribed value).

The function n^m can be differentiated still in another way. When we look at its Table 10.3, we see that the powers can be defined by their row differences

$$(n^m - 1) = (n-1)\sum_{i=0}^{m} n^i \; . \tag{10.10}$$

For example: $27 - 1 = 2(1 + 3 + 9)$. We can write this as the sum of differences of an infinite sequence $1/n^k$. We add 1 to both sides of (10.5) and write it as

$$n^m = (n-1)\sum_{k=1}^{\infty} n^{m-k} \; . \tag{10.11}$$

This equation is true even for $m = 1$ and we therefore have

$$n/(n-1) = (n-1)\sum_{i=0}^{\infty} n^{-i} \; . \tag{10.12}$$

This infinite sequence is hidden in the zero simplex because the numbers with negative powers $1/a^i$ cannot be interpreted as geometrical points with have negative sign, a^{-1} is not identical with $-a$.

For the sums of the first rows the following identities are found easily

$$\sum_{k=1}^{n} k^0 = n; \; \sum_{k=1}^{n} k^1 = \binom{n+1}{2}; \; \sum_{k=1}^{n} k^3 = \binom{n+1}{2}^2 \; . \tag{10.13}$$

All identities are easily proven by the full induction. Especially if the last one is true for n, then for $(n + 1)$ we have

$$\binom{n+1}{2}^2 + \binom{n+1}{1}^3 = \binom{n+2}{2}^2 \; .$$

This is verified by direct calculations.

It should be noted that the i-th row of the Table 10.2 is obtained consecutively by multiplying this matrix by the **Q** from the right from the (i-1)-th row. **Q** is the diagonal matrix of indices which repeat once again just under the main diagonal as in the following example

Table 10.4: Rencontres numbers of differences

k	0	1	2	3	4	5	Σ
m=0	1						1
1	0	1					1
2	1	1	1				3
3	4	6	2	1			13
4	27	28	16	3	1		75
5	187	214	104	31	4	1	541

				1			
				1	2		
					2	3	
						3	4
1				1			
	1			1	2		
	1	2		1	6	6	
	1	6	6	1	14	36	24 .

10.6 Some Classification Schemes

We can classify naive matrices similarly as it was done for the permutation matrices. Such classifications lead sometimes to complicated recurrences. For example if we imitate the rencontres numbers and count the number of elements on the main diagonal in vector strings, we obtain for $(3, 3)$ following two classifications

	The difference	Σ	*The rest of the simplex*	Σ	Σ
k=0	bca, cab, bab, baa	4	ccb, bcb, caa, cca	4	8
1	$aaa, aab, bba, acb, bac, cba$	6	$bbb, ccc, cbb, bcc, aca, cac$	6	12
2	aba, abb	2	bbc, cbc, aac, acc	4	6
3	abc	1		0	1
Σ		13		14	27

The Table 10.3 shows the rencontres numbers in the difference simplices, the Table 10.4 gives the counts for all naive matrices

We will not analyze these recurrences, but show another one. If the strings in plane simplices are classified according to the number of unit vectors n_1, we obtain the difference Table 10.5.

Table 10.5: Rencontres numbers in power series

k	0	1	2	3	4	Σ
m=0	1					1
1	0	1				1
2	1	2	1			4
3	8	12	6	1		27
4	85	104	54	12	1	256

Table 10.6: Differences of powers according to n_1

k	0	1	2	3	4	5	6	Σ
m=0	1							1
1	0	1						1
2	2	0	2					4
3	3	18	0	6				27
4	40	48	144	0	24			256
5	205	1000	600	1200	0	120		3125
6	2556	7380	18000	7200	10800	0	720	46656

The first column elements of the Table 10.5 can be named *subpowers*, because they generate the other elements in rows which sums give the powers n^n. The recurrence is

$$p_{i0} = 1 \; p_{ij} = p_{i-j,0} \; [i!/(i-j)!]^2 \times 1/j! = p_{i-j,0} j! \binom{i}{j}^2 . \qquad (10.14)$$

This recurrence can be divided into two steps. At first to naive matrices with $(i-j)$ elements j unit elements are added and the rows are permuted using the binomial coefficient $\binom{i}{j}^2$. Then we repeat permutations with columns using the same binomial coefficient. The result must be corrected for the permutations of th-added j unit elements between themselves, this is done by the factorial term $1/j!$.

10.7 Classification According to Two Vectors

All points in the partition diagrams of the simplices were divided into the orbits. They were classified according to the size of the largest vector. It is possible to count points and strings according to the size of one specific vector. This can be done for more vectors simultaneously, conveniently only

for two vectors, when the classification is planar. We abandon the spherical perspective and scan a simplex according to two axis. As an example we show the classification of the triangle 3^3

m_b	0	1	2	3	σ
$m_a = 0$	c^3	$3bc^2$	$3b^2c$	b^3	8
1	$3ac^2$	6abc	$3ab^2$		12
2	$3a^2c$	$3a^2b$			6
3	a^3				1
σ	8	12	6	1	27

For $(4, 4)$ simplex the following scheme is obtained similarly

m_b	0	1	2	3	4	Σ
$m_a = 0$	16	32	24	8	1	81
1	32	48	24	4		108
2	24	24	6			54
3	8	4				12
4	1					1
Σ	81	108	54	12	1	256

The zero row and column correspond to simplices 3^4, their crossection s_{00} and diagonal to 2^4. The elements are calculated as products of two binomial coefficients and corresponding the powers

$$\binom{m_a + m_b}{m_a}\binom{m}{m_a}(n-2)^{m-m_a-m_b} \ . \tag{10.15}$$

The row and column sums of two vector schemes give the one vector classification

$$\binom{m}{m_a}(n-1)^{m-m_a} \ . \tag{10.16}$$

10.8 Falling and Rising Factorials

In (10.6) a ratio of two factorials $i!/(i-j)!$ appeared. It was obtained from the corresponding binomial by multiplying it with the factorial $j!$. This ratio is known as the *falling factorial* and it is noted as $(n)_k$. The meaning of this notation of the falling factorial is that it is the product of k terms $(n-k)$, k going from 0 to $(k-1)$. When we arrange falling factorials into the Table 10.7 the falling factorial has a very simple inverse matrix.

The falling factorials can be obtained formally from the binomial

Table 10.7: Falling factorial and its inverse matrix.

k	0	1	2	3	4	5		0	1	2	3	4	5
m=0	1							1					
1	1	1						-1	1				
2	2	2	1						-2	1			
3	6	6	3	1						-3	1		
4	24	24	12	4	1						-4	1	
5	120	120	60	20	5	1						-5	1

$$(k!+1)^n \text{ substituting for } k!^j = j! \; . \tag{10.17}$$

We have mentioned the problem of distinguishability of things in distributions of things into distinguishable boxes. The distribution of the undistinguishable things, obtained as a sum of the polynomial coefficients for n permutations, led to the binomial coefficient $\binom{m+n+1}{m}$. Then we divided m ones into m rows and obtained the polynomial coefficient for m permutations, because these ones were equivalent. The sum of products of both coefficients gave n^m. Now we add the third index k. We can distinguish, if on the row i in the column j is 1_α or 1_β. There appears constant number $m!$ of permutations of m objects for all points counted by the sum of polynomial coefficients for n permutations. The result is

$$\sum_{k\geq 0} m!n!/\prod n_k! = (m+n-1)!/(n-1)! \tag{10.18}$$

This identity is known as the *rising factorial* and the notation $(n)^m$ is used. Both rising and falling factorials are related as

$$(n+m-1)_m = (n)^m \; . \tag{10.19}$$

It is possible to define the rising factorial as the falling factorial of negative numbers

$$(n)^m = (-1)^m(-n)_m \; . \tag{10.20}$$

For example: $(n)^2 = (n+2)(n+1)n = (-1)^3(-n)(-n-1)(-n-2)$.

10.9 Matrices $\mathbf{NN}^{\mathrm{T}}$

We have already counted quadratic forms $\mathbf{N}^{\mathrm{T}}\mathbf{N}$. Now we shall study the other quadratic forms $\mathbf{NN}^{\mathrm{T}}$. In them blocks $\mathbf{JJ}_k^{\mathrm{T}}$ obtained as outer products of the unit vector columns $\mathbf{J}_k$ appear.

For example: the block matrix

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

is permuted as

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

These blocks can not distinguish sequences $(\mathbf{ababa})^{\mathrm{T}}$ and $(\mathbf{babab})^{\mathrm{T}}$. They only register that on places $1, 3, 5$ one vector was and the other vector was on places 2 and 4.

The difference between both quadratic forms can be compared to two observers of trains. $\mathbf{N}^{\mathrm{T}}\mathbf{N}$ is an observer sitting on a train. He registers how many times his train moved but he can not tell when. $\mathbf{N}\mathbf{N}^{\mathrm{T}}$ is an observer on rails registering intervals when rails were used but he can not tell by which train.

The quadratic forms[3] $\mathbf{N}\mathbf{N}^{\mathrm{T}}$are counted by an index known as the *Bell polynomial*

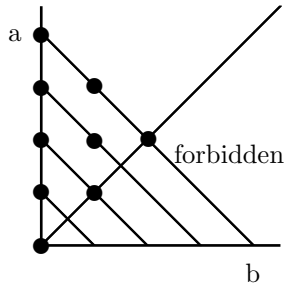
$$m!/\prod_{k\geq 0} n_k!m_k!^{n_k} \tag{10.21}$$

When we compare it with the product of two polynomial coefficients, we see that this was divided by the term $n!/n_0!$. This term appeared as the operand multiplying all Stirling numbers of the second kind to give differences $\Delta^m 0^n$ (Sect. 10.3). Therefore the Bell polynomials are counted by the Stirling numbers of the second kind and their sums. The number of quadratic forms $\mathbf{N}\mathbf{N}^{\mathrm{T}}$ is identical with the number of naive matrices in the lower triangular form without empty intermediate columns.

When the Bell polynomials are compared with the cycle index (7.15), we see that here instead of the simple m terms their factorials appear. The elements in columns do not form cycles but undistinguishable subsets. The Stirling numbers generate differences, if multiplied with the matrix of

[3]The quadratic forms $\mathbf{N}\mathbf{N}^{\mathrm{T}}$ of long strings form very interesting patterns.

Figure 10.1: Balloting numbers cone. Coordinates **a** are always greater then coordinates **b**.



factorials, and with the matrix of powers, if multiplied with falling factorials:

				1	1	1	1	1	1	1	1
					2	2	2		1	2	3
						6	6			2	6
							24				6
1				1	1	1	1	1	1	1	1
1	1			1	3	3	3	1	2	3	4
1	3	1		1	7	13	13	1	4	9	16
1	7	6	1	1	15	51	75	1	8	27	64

When the lowest allowable value $m_j = 2$, the polynomials give associated Stirling numbers of the second kind which recurrence is

$$a_{ij} = ja_{i-1,j} + (i-1)a_{i-2,j-1} \text{ with } a_{00} = 1\ . \tag{10.22}$$

10.10 Balloting Numbers

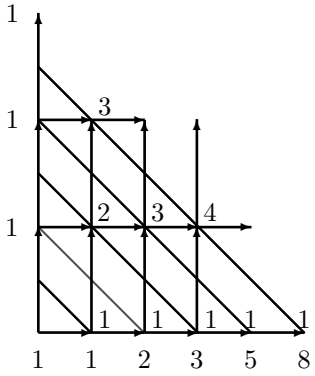
In Sect. 9.8 the Fibbonacci numbers were introduced with a singular matrix. If we rearrange its elements as in Table 9.7, we obtain a matrix which can be inverted. The positive numbers of the inverse matrix are known as the *balloting numbers*.

Actually, the balloting numbers are all positive. The negative signs appear by multiplication with $\mathbf{I}*$ from both sides. They count binary strings in which one side has always an advantage given by the sieve rule $m_{ai} \geq m_{bi}$. The counted strings lead only in a half of the two dimensional cone (Fig. 10.1). The inverse Fibbonacci matrix counts strings which elements are **b**

Table 10.8: Fibbonacci and balloting numbers

	Fibbonacci numbers								Balloting numbers						
k	1	2	3	4	5	6	7		1	2	3	4	5	6	7
m=1	1								1						
2		1								1					
3	1		1						-1		1				
4		2		1						-2		1			
5	1		3		1				2		-3		1		
6		3		4		1				3		-4		1	
7	1		6		5		1		-2		9		-5		1

Figure 10.2: Fibbonacci lattice. Odd vectors **a** are not formed. The Fibbonacci numbers count the restricted strings.



and two consecutive $\mathbf{aa} = \mathbf{a}^2$. For example: $f_{75} = 5$ counts strings $\mathbf{b}^5\mathbf{a}^2$, $\mathbf{b}^4\mathbf{a}^2\mathbf{b}$, $\mathbf{b}^3\mathbf{a}^2\mathbf{b}^2$, $\mathbf{b}^2\mathbf{a}^2\mathbf{b}^3$, $\mathbf{ba}^2\mathbf{b}^4$.

The corresponding lattice is depicted on Fig. 10.2. The Fibbonacci numbers f_{ij} are generated by the recursion

$$f_{11} = 1;\ f_{ij} = f_{i-1,j-1} + f_{i,j-2}\ . \tag{10.23}$$

The balloting numbers b_{ij} are generated by the recursion

$$b_{11} = 1;\ b_{ij} = b_{i-1,j-1} + b_{i-1,j+1}\ . \tag{10.24}$$

We can formulate also a table, which elements count strings in which $m_a \geq m_b^2$. Its elements are $b_{ij} = b_{i-1,j-1} + b_{i-2,j}$, and it is again a rarefied matrix of binomial coefficients.

Table 10.9: Differences of binomial coefficients

j	0	1	2	3	4	5
m=0	1					
1	2	1				
2	4	3	1			
3	8	7	4	1		
4	16	15	11	5	1	
5	32	31	26	16	6	1

The inverse matrix with positive signs is

	1	2	3	4	5	6	7	8	9
n=0	1								
1		1							
2			1						
3	1			1					
4		2			1				
5			3			1			
6	3			4			1		
7		7			5			1	
8			12			6			1

The matrix elements, the numbers b_{ij}, are generated by the recursion

$$b_{11} = 1;\ b_{ij} = b_{i-1,j-1} + b_{i-1,j+2}\ . \tag{10.25}$$

They count strings with the elements a^3, and b.

10.11 Another Kind of Differences

For all points (elements) of space we can measure distances (differences) from the other points. These distances are induced by their special functions. As an example we introduce differences $[2^m + 1] - \binom{m}{j}$ tabulated in Table 10.9. The inverse matrix has elements as in Table 10.10.

When we do not consider the signs $(-1)^{m+j}$, the squares of $(m-1)$ are in the third column, in the second column its first differences $(m+1)^2 - m^2$ and in the first column its second differences which are from the second row up constant. The higher column elements are differences of the elements of the previous columns

$$m_{ij} = m_{i-1,j-1} - m_{i-1,j}\ , \tag{10.26}$$

Table 10.10: Differences of m^2

j	0	1	2	3	4	5
m=0	1					
1	-2	1				
2	2	-3	1			
3	-2	5	-4	1		
4	2	-7	9	-5	1	
5	-2	9	-16	14	-6	1

Table 10.11: Lah numbers $\mathbf{L}$

n	1	2	3	4	5	Σ
m=1	1					1
2	2	1				3
3	6	6	1			13
4	24	36	12	1		73
5	120	240	120	20	1	501

similarly as all sums in the matrix of binomial coefficients.

Another possible decomposition of the Table 10.10 is on the sum of two tables of binomial coefficients $\mathbf{B}_{m,j} + \mathbf{B}_{m-1,j-1}$.

Such difference tables can be constructed for any powers of consecutive numbers. Their inverse matrices have no simple interpretation as differences of squared numbers.

10.12 Lah Numbers

It is difficult to show all relations between all space functions. The Lah numbers $\mathbf{L}$ are introduced simply by their Table 10.11.

Actually, the original Lah numbers $\mathbf{L}$ have odd rows negative signs and then

$$\mathbf{L}^2 = \mathbf{I}; \text{ or } \mathbf{L}^{-1} = (-1)^{i+j}\mathbf{L} \,. \tag{10.27}$$

The elements of the Table 10.11 are direct products of falling factorials with the binomial coefficients

$$l_{ij} = i!/j!\binom{i-1}{j-1} \,. \tag{10.28}$$

The recurrence of the Lah numbers is

Table 10.12: Differences as product $\mathbf{S}_2\mathbf{S}_1$.

n	1	2	3	4	5	Σ
m=1	1					1
2	2	1				3
3	6	6	1			13
4	26	36	12	1		75
5	150	250	120	20	1	541

$$l_{i+1,j} = (i+j)l_{ij} + l_{i,j-1} \tag{10.29}$$

Another possibility to produce the Lah numbers, is the product of matrices of Stirling numbers of both kinds. The matrix of the Stirling numbers of the second kind is multiplied by the matrix of the Stirling numbers of the first kind from the right:

$$\mathbf{L} = \mathbf{S}_1\mathbf{S}_2 \; . \tag{10.30}$$

Due to the relations of both kinds of Stirling numbers the inverse of the Lah matrix is identical with the matrix itself.

The transposed order of the Stirling numbers multiplication gives another Table 10.12, this time of differences $\Delta(n)n^n$

Multiplying the matrix of the Stirling numbers of the first kind by the matrix of the Stirling numbers of the second kind gives the same result as the permutations of columns of the naive matrices in the lower triangular form with j columns with nonzero elements by the permutation matrices **P** with i rows and columns and j cycles. The arrangement assures that empty columns are not permuted from their position. The Table 10.12 counts strings according to the number of columns in the lower triangular form, which were not permuted from their positions. The elements of its first column are, except the first element, $2\Delta^{n-1}0^{n-1}$. There are counted matrices in the lower triangular form with the leading first element **a** and the second element either **a** or **b**.

Chapter 11

Multidimensional Cubes

11.1 Introduction

As an introduction to this chapter, we repeat some of the facts about cubes, which were already explained previously. We used as the generating function of the powers of vector sets $(\Sigma \mathbf{e}_j)^m$. We obtained vector strings $\mathbf{N}$ leading to the points on the planes orthogonal to the unit diagonal vector $\mathbf{I}$. We found mathematical operations which arranged these matrix vectors $\mathbf{N}$ onto spherical orbits and mentioned some possibilities to form from the plane simplices their complexes, that is the positive cones in vector space. We have also shown that cubes or generally any parallelepipeds are formed from plane complexes by truncating too long vectors. The traditional approach, the Cartesian product of n one dimensional complexes gives only points, no vector strings

$$(1+a+a^2)\times(1+b+b^2) = 1+a+a^2+b+ab+b^2+a^2b+ab^2+a^2b^2\ . \quad (11.1)$$

These n dimensional cubes are formed usually by the Cartesian products of n one dimensional complexes, For example:

$$(1+a+a^2)\times(1+b+b^2) = 1+(a+b)+a^2+ab+b^2+a^2b+ab^2+a^2b^2\ . \quad (11.2)$$

The first three simplices are complete, but the last two are truncated. Moreover, not all strings are produced. Now we will treat cubes systematically. Especially we will show how the vector strings are transformed into points of cubes and points of plane simplices into orbits. This transformation is possible by interpretation of the transposed naive matrices $\mathbf{N}^{\mathrm{T}}$ as

faces (Fig. 1.4), the vectors determining the coordinates of the points in m dimensional space. Each vector string corresponds to one point and all strings of the plane simplex n^m are mapped onto points of m dimensional cube which side is $(n-1)$. This transformation is not a simple task. It can be demonstrated on mapping a 3 dimensional plane onto 4 dimensional cube with the sides 0-2.

Moments:		0	1	2	3	4	5	6	7	8	Σ
Plane strings:	b=0	1		4		6		4		1	16
	1		4		12		12		4		32
	2			6		12		6			24
	3				4		4				8
	4					1					1
Cube points:	Σ	1	4	10	16	19	16	10	4	1	81

The strings from different orbits are counted together because they have equal *moments*. New orbits go from 0 to $m(n-1)$. Some of the known functions receive a new interpretation, but it still will be necessary to introduce some new functions.

For plane simplices we have introduced differences which have somewhat curious properties. They include one vertex, one incomplete edge, one incomplete side. But when we transpose the naive matrices **N** and interpret them as faces, we see, that these properties mean that the difference of a cube is occupied by its points touching its surfaces nearest to the center of the coordinates, having at least one coordinate zero, at least one coordinate one and so on in the higher dimensional cubes (Fig. 11.1).

11.2 Unit Cubes

The unit cubes are most instructive to start with. They have n sides and on each side there are just two points, 0 and 1. They are generated by the function

$$\prod_{j=1}^{n}(1+\mathbf{e}_j) = 2^n \ . \tag{11.3}$$

For example, for $n = 3$ we get points: $1, a, b, c, ab, ac, bc, abc$ (Fig. 11.2). One from the most interesting properties of the unit cubes, in which only whole coordinates are allowed, is that they are formed only by a surface. There is no point inside them representing their center.

Figure 11.1: Difference of the three dimensional cube with the sides $0-2$. The difference is made from points touching the surfaces of the cube nearest to the center of coordinates. The points of the difference have the coordinates (permuted): $(0,0,0)$, $(0,0,1)$, $(0,1,1)$, and $(0,1,2)$.

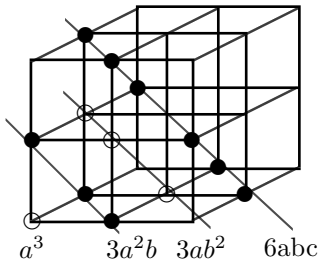


Figure 11.2: Three dimensional cube with the sides 0-1.

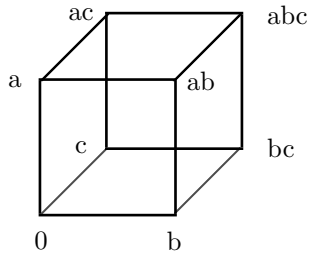


Table 11.1: Strings of unit cubes **F**.

k	0	1	2	3	4	5	Σ
m=0	1						1
1	1	1					2
2	1	2	2				5
3	1	3	6	6			16
4	1	4	12	24	24		65
5	1	5	20	60	120	120	326

There are $(m+1)$ partition orbits in the unit cubes, from each plane simplex there is just one orbit. The number of points on each orbit is determined by the corresponding binomial coefficient. What remains to be determined is the number of strings in the unit cubes, but we have studied even this function, and this number is given by the falling factorial $(i)_{(i-j)}$. We will look on it again. We write it in the inverse order against the Table 11.1

The elements of the Table 11.1 f_{ij} are obtained as the product of the binomial matrix **B** and the diagonal matrix of factorials $\Delta(j!)$:

$$\mathbf{F} = \mathbf{B}\Delta(j!) \; . \tag{11.4}$$

We can choose k objects (vectors) from n objects and then to permute them. This is done as the formal binomial, when consecutive factorials are treated as powers

$$(n)_m = [(k)_i + (n-k)_i]^m \text{ where } (k)_i^j = (k)_j \; . \tag{11.5}$$

For example: if $n = 5, m = 3$, and we choose $k = 2$, the result is

$$(5)_3 = 60 = \binom{3}{0}(2)_3(3)_0 + \binom{3}{1}(2)_2(3)_1 + \binom{3}{2}(2)_1(3)_2 + \binom{3}{3}(2)_0(3)_3 =$$

$$1 \times 0 \times 1 + 3 \times 2 \times 3 + 3 \times 2 \times 6 + 1 \times 1 \times 6 \; .$$

It counts 18 permutations of strings with two symbols, say

$$a, b : 6(abc, abd, abe) \; ;$$

36 permutations with either a or b: $6(acd, ace, ade, bcd, bce, bde)$, and 6 permutations of the string cde. $(2)_3 = 0$; it is not possible to form a sequence of three symbols from only two symbols. The row sums are given simply as

$$S_m = m(S_{m-1}) + 1 \ . \tag{11.6}$$

It is possible to add a new object to the preceding strings in m ways, except the zero string. Another possibility to obtain the matrix 11.1, is to multiply the matrix of rencontres numbers **R** (Table 7.3) with the matrix of the binomial coefficients

$$\mathbf{F} = \mathbf{RB} \ . \tag{11.7}$$

Otherwise, the strings of the unit cubes are generated similarly as the factorials from subfactorials by the Apple polynomial D. Here it is the polynomial of the second order, $(D+2)^2$, For example:

$$44 + 5 \times 9 \times 2 + 10 \times 2 \times 4 + 10 \times 1 \times 8 + 5 \times 0 \times 16 + 1 \times 1 \times 32 = 326 \ .$$

It is well known, that in Nature many events are described by the binomial distribution. When you toss n coins simultaneously, then the results will fill vertices of the unit cube evenly, especially if the experiment is repeated many times. At least that is what the probability theory supposes. Less known is the derivation of another statistics generated by the unit cubes. Suppose that we are registering accidents. Let us have S_m persons with at most m accidents and the mean accident rate 1 per person. At such conditions, we choose as the registration tool strings of k ones from m symbols if the other $m-k$ places are exploited for indexing the persons. Such a register will have the following capacity: There will be $m!$ persons with no accident, $m!$ persons with one accident, m persons with $(m-1)$ accidents, and at last only one person with m accidents. Such distribution of accidents is known as the *Poisson distribution*. It is applied usually to low accident rates and it is then necessary to change the conditions. Nevertheless, if Einstein said that God does not play dice, we can say, that he himself is the Dice. The tossing of coins or dices models only the ideal space.

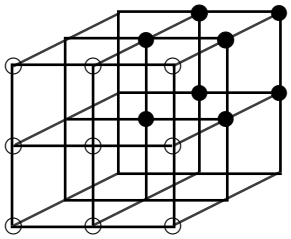
11.3 Partition Orbits in Cubes

The partition orbits in cubes correspond to points of plane simplices. Thus we know their total number. We have also shown above, how differently these points are mapped on the plane simplices and the cubes. We already found that counting of the orbits is very simple in the unit cubes.

Table 11.2: Partition orbits in cubes 0-2

k	0	1	2	3	4	5	6	7	8	9	10	11	12	Σ
m=0	1													1
1	1	1	1											3
2	1	1	2	1	1									6
3	1	1	2	2	2	1	1							10
4	1	1	2	2	3	2	2	1	1					15
5	1	1	2	2	3	3	3	2	2	1	1			21
6	1	1	2	2	3	3	4	3	3	2	2	1	1	28

Figure 11.3: Formation of three dimensional cube with the side 0-2 from the square with the side 0-2 (empty circles). The unit three dimensional cube with the side 0-1 is added (filled circles) and sides are completed.



The partition orbits in the m dimensional cubes which sides are 0-2 are easily found. The results are given in the Table 11.2. It was shown in Sect. 11.1, how its row m=4 is obtained from points of the plane simplex.

Some properties of the distribution of partition orbits are clear. They are symmetrical according to the parameter k. This follows from the symmetry of the cubes. The number of orbits on planes near to the zero point does not depend on the dimensionality of cubes and remains constant. It is determined by the number k and cannot be greater than the number of unrestricted partitions p(k). If we use the constant c as the length of the sides of the cubes, the diagonal k is going from 0 to cm.

When we observe row differences in Table 11.3, we see that they are always 1 on the last $(m + 1)$ occupied places. These numbers are just the numbers of partition orbits in m dimensional unit cubes. In 3 dimensional space, it can be drawn (Fig. 11.3). To a square with the sides 0-2 the unit three dimensional cube is added, forming the back of the cube with sides 0-2. The orbit 111 is formed from the orbit 11, which was not

Table 11.3: Points in cubes with c=2.

k	0	1	2	3	4	5	6	7	8	Σ
n=0	1									1
1	1	1	1							3
2	1	2	3	2	1					9
3	1	3	6	7	6	3	1			27
4	1	4	10	16	19	16	10	4	1	81

in the square, 211 or 221 is obtained from 21, 22 generates 221 and 222. This suggests the recurrence of the partition orbits. It can be formulated graphically:

	0 < MOMENTS > mc < > m(c+1)
	Orbits of m dimensional cube of lesser size (c-1)
	Orbits of (m-1) dimensional cube of the same size
Σ:	Orbits of m dimensional cube with the size c

Because the cubes are symmetrical along their diagonals, the positions of the summands can be inverted. For example

1	1	1	1								1	1	1	1
		1	1	2	1	1	=	1	1	2	1	1		
1	1	2	2	2	1	1		1	1	2	2	2	1	1

Defining the number of orbits p(m,n,c) on the plane m of n dimensional cube with the side c, we have

$$p(m,n,c) = p(m,[n-1],c) + p([m-n],n,[c-1]) \; . \tag{11.8}$$

11.4 Points in Cubes

We know the total number of the points with the natural coordinates (otherwise the volume m^n) in the cubes, and now we want to determine their distribution according to their moments in the plane simplices. If starting simplices are not truncated, these numbers must be the binomial coefficients $\binom{m+k-1}{k}$. Similar numbers appear on the tails of distributions. From the first cubes with $c = 2$, the recurrence can be easily deduced the table 11.3, which is known in literature as the *trinomial triangle*.

Here the recurrence is simple. To each point of a $(n-1)$ dimensional cube we add a new side with c points. By adding 0, 1 we simply sum up

$(c+1)$ partition orbits of the less dimensional cube. For example the term 19 in the last row is obtained as 6 + 7+ 6. The formula is

$$c_{ij} = \sum_{k=0}^{c} c_{i-1,j-k} \ . \tag{11.9}$$

A new vector with all its allowed values is added to each partition on a suitable place.

Another possibility to produce cubes is to increase the size c of the cubes. The cubes of different dimensions m are multiplied by the transposed matrix of the binomial coefficients as follows. The number of points of greater cubes appears on the diagonal

				1	1	1	1
					1	2	3
						1	3
							1
1				1	1	1	1
3	1			3	4	5	6
9	3	1		9	12	16	21
27	9	3	1	27	36	48	64

The tree dimensional cube with $c = 2$ has 27 points. It is transformed in the dimensional cube with $c = 3$ by adding 3×2 dimensional cubes (squares), 3×1 dimensional cubes (edges) and 1×0 dimensional cube, the point with coordinates $(3, 3, 3)$. The new diagonal elements in the inverse order, $64, 16, 4, 1$, form the new baseline of the next cube. To increase the size of the cubes, it is necessary to rearrange the diagonal elements and repeat the multiplication.

11.5 Vector Strings in Cubes

In Sect. 11.2, we have shown that in the unit cubes the strings are counted by the falling factorials. For other cubes the numbers of strings are not determined as easily, but it is not as that difficult, if we make it consecutively. For example: for $c = 2$ we obtain the Table 11.4.

To show how the elements of the Table 11.4 are generated, the result for s_{45} is given: $600 = 90 + 5 \times 54 + 10 \times 24$. We obtained the points in the cubes by summing $(c+1)$ elements of the less dimensional cube (11.9). In this case it is necessary to permute added symbols with symbols of the corresponding strings with $(n-1)$ symbols. This is done by multiplying

Table 11.4: Vector strings in cubes with c=2

m	0	1	2	3	4	5	6	7	8	Σ
n=0	1									1
1	1	1	1							3
2	1	2	4	6	6					19
3	1	3	9	24	54	90	90			271
4	1	4	16	60	204	600	1440	2520	2520	7365

Table 11.5: Strings in 2 dimensional Cubes.

k	0	1	2	3	4	5	6	7	8	Σ
c=0	1									1
1	1	2	2							5
2	1	2	4	6	6					19
3	1	2	4	8	14	20	20			69
4	1	2	4	8	16	30	50	70	70	201

the corresponding numbers with binomial coefficients. The recurrence is therefore

$$s_{ij} = \sum_{k=0}^{c} \binom{m}{k} s_{i-1,j-1} \ . \tag{11.10}$$

Another possibility to obtain the greater cubes by increasing the sides of the n dimensional cubes gives also a possibility for finding recurrent formulas for the number of strings. For $n = 2$ (the squares), we obtain the Table 11.5.

The recurrence is

$$s_{i0} = 1;\ s_{ij} = s_{i-1,j-1} + s_{i,j-1};\ s_{ij} = 0 \text{ outside the cube} \ . \tag{11.11}$$

There are always two possibilities to prolong the strings, except of strings leading for the back sides. The first corresponds to the term $s_{i,j-1}$, the second possibility inside the squares is accounted by counting the strings from the lesser square $s_{i-1,j-1}$.

It is also possible to shift a cube in its space, when its point with the lowest moment is not incident with the beginning of the coordinate system. The number of orbits and points is not changed by this operation, but the number of strings is.

11.6 Natural Cubes - e Constant

We have shown that the unit cubes are generated by the formula 1.3. The term 1 in $(1 + \mathbf{e}_j)$ was interpreted as $\mathbf{e}_j^0$. The volume of a cube depends on its base m and on the its dimensionality n. Now we will study, what volume e a cube has, if its side nears to one and its dimensionality to infinity. We try to find what value has the limit

$$e = \lim_{z \to \infty} (1 + 1/z)^z \ . \tag{11.12}$$

The argument in the equation 11.6 can be either positive or negative.

The base of e cube lies between cubes with whole numbers $1 < (1 + 1/z) < 2$. When $z = 1$, the result is 1.5 instead 2^1. When $z = 2$, the result is $1.5^2 = 2.25$ instead 2^2. Evaluating the binomial development of (11.7), we obtain inequalities

$$\sum_{k=0}^{\infty} 1/k! < e = \sum_{k=0}^{\infty} \binom{z}{k} 1/k! < 1 + \sum_{k=0}^{\infty} 1/2^k = 3 \ . \tag{11.13}$$

Using sophisticated mathematical arguments, it can be proven that the number e must be greater than the sum of the inverse factorials. Because it should be simultaneously smaller, the best solution is the one where both limits coincide. The constant e is an irrational number and its first digits are $e = 2.71828\ldots$. The sum of inverse factorials approaches to the exact value fast. Thus the first seven terms give

$$e = 1 + 1 + 1/2 + 1/6 + 1/24 + 1/120 + 1/720 = 2,71805 \ .$$

The next term is $1/5040 = 0.000198$. It corrects the fourth decimal place.

If z is negative, the substitution $z = -(t + 1)$ is inserted into the formula (11.7) and than some modifications show, that again the number e is obtained:

$$\lim_{t \to \infty} [1 - 1/(t+1)]^{-(t+1)} = \lim [t/(t+1)]^{-(t+1)} = \tag{11.14}$$
$$lim(1 + 1/z)^{t+1} \times \lim(1 + 1/z) = e \times 1 = e \ .$$

The generating function of the e cube has some important properties, which make from it a useful tool. When a substitution $x = az$ is applied, the limit of the expression

$$\lim_{x\to\infty}(1+a/x)^x = e^a = exp(a) \tag{11.15}$$

is the a-th power of the number e. This property of the number e is exploited using e as the base of natural logarithms.

When we return to the function of the rising factorial (10.8) which counts strings in unit cubes, then the number of all strings in the infinite unit cube can be expressed using the constant e:

$$\lim_{n\to\infty} n!\sum_{k=0}^{\infty} 1/k! = en! \; . \tag{11.16}$$

A note: To paint a half of sides of a cubical canister in infinite dimensional space, more varnish is needed that the volume of the canister is.

Chapter 12

Matrices with Whole Numbers

12.1 Introductory Warning

This chapter will remain only sketched. The reason is that it is practically impossible to treat its content systematically as it was done with naive matrices. The remaining of the chapter will be exploited for introducing some matter which belongs to the following Chapters.

12.2 Matrices with Unit Symbols

We started our study with permutation matrices having in each row and column exactly one unit symbol. Then we added the naive matrices, having this restriction only for rows and the transposed naive matrices, where it was applied for columns. The next step is to allow units to be inserted to any available place of a matrix. We already know, that the number of these matrices will be determined by a binomial coefficient. For matrices with m columns and n rows, with k unit elements in the matrix, the number of the possible configurations will be determined by the binomial coefficient $\binom{mn}{k}$. These configurations can be counted using tables having two partition orbits in rows as well as in columns. For example: for $m = n = k = 4$ we obtain Table 12.1.

The Table 12.1 gives some perspective. In the space, new vector strings appeared. They lead to the same points as the naive matrices, but their orbits are not simple partition orbits but the *pattern orbits* which are the

Table 12.1: Distribution of unit matrices $m = n = k = 4$.

Partition $\mathbf{N}^{\mathrm{T}}$	4	31	22 $\mathbf{N}$	211	1111	Σ
4	0	0	0	0	4	4
31	0	0	0	144	48	192
22	0	0	36	144	36	216
211	0	144	144	720	144	1152
1^4	4	48	36	144	24	256
Σ	4	192	216	1152	256	1820

products of two partitions, one for rows and the other one for columns.

For example, the pattern produced by the partition product (211×310) is

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

This is a Ferrers graph. There are 6 possible permutations of the rows of this pattern (all rows are different) which are combined with permutations of the fourth zero row. Two columns are equal. Thus there are 3 possible permutations which are combined with permutations of the fourth zero columns.

The partition product (211×211) has two patterns:

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

with all possible $36 = 3! \times 3!$ permutations of rows and columns, and the second one

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

with 9 permutations of the zero element marked by *. Unit elements fill only the marked row and column. These permutations of patterns are multiplied respectively by 16 permutations of the fourth row and column with zero elements. It is easy to count all permutations to a given pattern, but it is more difficult to find all patterns generated by the given partition product.

Table 12.2: Matrices with elements $\geq$ 1

Partition	4	31	22	211	1111	Σ
4	16	48	24	48	0	136
31	48	288	144	288	0	768
22	24	144	72	144	0	384
211	48	288	144	288	0	768
Σ	136	768	384	768	0	2056
1111	4	192	216	1152	256	1820
Σ	140	960	600	1920	256	3876

The total number of unit vectors with constant sums is given by the row or column sums elements of tables similar to Table 12.1. Due to the difficulties with the notation, we will give the formula only for column sums, where we can use the symbol n_i for the number of identical binomial coefficients

$$\sum (n!/\prod n!)\binom{m}{k_j}^{n_j} = \binom{mn}{k}; \sum_{j=1}^{n} k_j = k\ . \tag{12.1}$$

The sum is made over all possible partitions. The product of the binomials is not restricted by any conditions on column sums, and therefore units in each row can be distributed independently, then the rows obtained by such a way are permuted ($n = m$) but $n!$ overestimates permutations of rows with equal sums, therefore the result must be divided by the partial factorials.

12.3 Matrices with Natural Numbers

Now the next step seems to be easy. A matrix is a mn dimensional vector and if k unit elements can be placed in it without any restrictions, the number of all possible vectors is given by the binomial coefficient (10.2) $\binom{mn+k}{k}$. The Table 12.1 should be completed by 2056 new entries to give $\binom{19}{4}$ different matrices instead $\binom{16}{4}$ matrices with the unit elements. The new patterns fill the Table differently, see the Table 12.2

It is practically impossible to follow all possible patterns of matrix vectors as we did before. One special class of them was studied systematically, matrices having in each row exactly two unit symbols. These patterns developed into a special branch of mathematics, the graph theory (see the next Chapter).

In the previous Chapters we have counted the partition vectors, that is the number of the Ferrers graphs. This is simultaneously the number of the diagonal patterns corresponding to quadratical forms of the naive matrices. This patterns can be compared with the symmetrical unit patterns of $\mathbf{J}\mathbf{J}_j^{\mathrm{T}}$ matrices with m_j elements, which is the pattern of the number Σm_j^2.

12.4 Interpretation of Matrices with Natural Numbers

If a diagonal matrix is projected onto the unit vector row $\mathbf{J}^{\mathrm{T}}$, the result is a row vector corresponding to a vector row of generalized matrices with natural numbers. It is thus possible to write such a matrix as a string of projections of quadratic forms of naive strings onto the consecutive unit vector rows.

$$(\mathbf{J}_1^{\mathrm{T}}\mathbf{N}_1^{\mathrm{T}}\mathbf{N}_1,\ \mathbf{J}_2^{\mathrm{T}}\mathbf{N}_2^{\mathrm{T}}\mathbf{N}_2,\ \mathbf{J}_3^{\mathrm{T}}\mathbf{N}_3^{\mathrm{T}}\mathbf{N}_3)^{\mathrm{T}}\ . \tag{12.2}$$

Another possibilities will be shown later. We can interpret a matrix $\mathbf{M}$ together with its transpose $\mathbf{M}^{\mathrm{T}}$, taken in the block form

$$\begin{pmatrix} \mathbf{0} & \mathbf{M}^{\mathrm{T}} \\ \mathbf{M} & \mathbf{0} \end{pmatrix},$$

as an adjacency matrix $\mathbf{A}$ of a bipartite graph with multiple edges (see the next Chapter).

12.5 Coordinate Matrices

We interpreted rows in matrices as strings of consecutive vectors. There exist still another explanation. The rows are just simultaneous vectors determining positions of different points or objects. The matrix

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

gives for two different points (or objects) the same address. This is possible, if the address $(1, 0)$ is For example: a house or a box. Thus it is necessary to study the possibility that matrices define positions of m points in space, that they are lists of coordinates in orthogonal axes. Such a list forms the *coordinate matrix* $\mathbf{C}$ which elements c_{ij} are coordinates of m points (vertices, objects) i on n axes.

The matrix column **A**

$$(0, 1, 2, 3, 4)^{\mathrm{T}}$$

determines coordinates of five points lying on the natural number axis. Between all points unit distances are.

The matrix **B**

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix}^{\mathrm{T}}$$

determines coordinates of five points rotated into two dimensional plane. Another straight configuration of five points **C** is the plane simplex

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 \end{pmatrix}^{\mathrm{T}} .$$

These are examples of the simplest regular structure of five points, evenly spaced straight chain.

If the quadratic forms $\mathbf{C}\mathbf{C}^{\mathrm{T}}$ of coordinate matrices are calculated, they have on their diagonals squared Euclidean distances of each point from the center of the coordinate a system

A

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 & 8 \\ 0 & 3 & 6 & 9 & 12 \\ 0 & 4 & 8 & 12 & 16 \end{pmatrix}$$

B

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 4 & 6 & 8 \\ 0 & 4 & 8 & 12 & 16 \\ 0 & 6 & 12 & 18 & 24 \\ 0 & 8 & 16 & 24 & 32 \end{pmatrix}$$

C

$$\begin{pmatrix} 16 & 12 & 8 & 4 & 0 \\ 12 & 10 & 6 & 3 & 0 \\ 8 & 6 & 4 & 2 & 0 \\ 4 & 3 & 2 & 10 & 0 \\ 0 & 0 & 0 & 0 & 16 \end{pmatrix}$$

Off-diagonal elements are quadratic products of both distances i and j.

Coordinates of points form structures in space. If the chain is flexible, it can be wound over edges on the unit cube

D

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

Here the four points are placed on the vertices of the three dimensional cube. Another configuration is

E

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

Here all four coordinates in the first column are zeroes. They can be thus neglected. The first point lies in the center of the coordinate system, the second one on the end of the second unit vector, the third one on the end of the third unit vector. The points are related as in the three dimensional plane complex. The distances between them are not equal. The first point is in the unit distance to the other three points, the distances between these three points are doubled.

The configuration of four points determined by the coordinate matrix

F

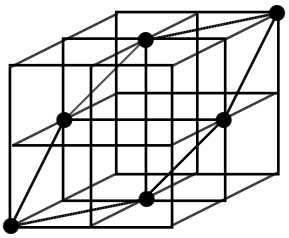
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

corresponds to the regular tetrahedron. The chain is wound over its vertices.

12.6 Oriented and Unoriented Graphs as Vector Strings

If we draw the difference of two vectors $(\mathbf{e}_b - \mathbf{e}_a)$, as on Fig. 3.2, it corresponds to the accepted convention for drawing arcs of oriented graphs (see

Figure 12.1: Two diagonal strings in three dimensional cube $0--2$. Find the remaining four.



next Chapt.). The incidence matrix $\mathbf{S}$ of a graph is just the difference of two naive matrices

$$\mathbf{S} = \mathbf{N}_a - \mathbf{N}_b \ ,$$

as it was shown in Sect 3.3. It is an operator which transfers a vector string into another. A vector string is a continuous path in the vector space, the operator transferring one vector string into another is also continuous. It seems to be the area between two strings or vector lines, and we could imagine it as a surface. But when we make consecutive differences at all pairs of unit vectors, we get a linear vector again. A loop transfers a unit vector into itself. All these vectors lie in the plane orthogonal to the unit diagonal vector $\mathbf{I}$.

The other possibility, how to interpret oriented graphs is the difference inside a string itself. For example, a string **abcda** contains transitions **a** into **b**, **b** into **c**, **c** into **d** and **d** into **a**. The difference is thus:

$$\begin{array}{ccccccc} \mathbf{a} & \rightarrow & \mathbf{b} & & & & \\ & \mathbf{b} & \rightarrow & \mathbf{c} & & & \\ & & \mathbf{c} & \rightarrow & \mathbf{d} & & \\ & & & \mathbf{d} & \rightarrow & \mathbf{a} & \end{array} \ .$$

Similarly differences at higher distances could be compared.

The oriented complete graphs K_n form 2 dimensional edges (known as arcs) of the plane simplices. Unoriented graphs are strings of vectors orthogonal to surfaces of corresponding oriented graphs. The unoriented complete graphs K_n are vector strings going from the coordinate center to the farthest end of the unit cube which sides are diagonals of the n dimensional unit cube, or otherwise, to the farthest end of the cube with the side 0-2, as it is shown on Fig. 12.1. Another graphs correspond to multisets

from these bases, defined as differences or sums of naive matrices. Edges and arcs of graphs form a space which symmetry is more complicated that the symmetry of naive matrices.

The recursive definition of the canonical form of the incidence matrix $\mathbf{S}$ of the complete oriented graph K_n is

$$\begin{pmatrix} \mathbf{S}_{n-1} & \mathbf{0}_{n-1} \\ -\mathbf{I}_{n-1} & \mathbf{J}_{n-1} \end{pmatrix} , \tag{12.3}$$

where $\mathbf{0}_{n-1}$ is the zero vector-column. Similarly, the canonical form of the complete unoriented graph K_n is

$$\begin{pmatrix} \mathbf{G}_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{I}_{n-1} & \mathbf{J}_{n-1} \end{pmatrix} . \tag{12.4}$$

12.7 Quadratic Forms of the Incidence Matrices.

A simple exercise in matrix multiplication shows that the quadratic forms of the incidence matrices of unoriented and oriented graphs have the form

$$(\mathbf{N}_a^{\mathrm{T}} + \mathbf{N})_b^{\mathrm{T}}(\mathbf{N}_a + \mathbf{N}_b) = (\mathbf{N}_a^{\mathrm{T}}\mathbf{N}_a + \mathbf{N}_b^{\mathrm{T}}\mathbf{N}_b) + (\mathbf{N}_a^{\mathrm{T}}\mathbf{N}_b + \mathbf{N}_b^{\mathrm{T}}\mathbf{N}_a) \tag{12.5}$$

$$(\mathbf{N}_a^{\mathrm{T}} - \mathbf{N})_b^{\mathrm{T}}(\mathbf{N}_a - \mathbf{N}_b) = (\mathbf{N}_a^{\mathrm{T}}\mathbf{N}_a + \mathbf{N}_b^{\mathrm{T}}\mathbf{N}_b) - (\mathbf{N}_a^{\mathrm{T}}\mathbf{N}_b + \mathbf{N}_b^{\mathrm{T}}\mathbf{N}_a) \tag{12.6}$$

The quadratical forms are composed from two parts: The diagonal matrix $\mathbf{V}$ formed by the sum of quadratic forms of the two naive matrices $\mathbf{N}_a$ and $\mathbf{N}_b$. The diagonal elements v_j are known as *degrees* of the corresponding vertices.

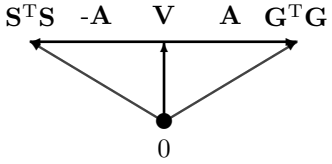
The sum of scalar products

$$(\mathbf{N}_a^{\mathrm{T}}\mathbf{N}_b + \mathbf{N}_b^{\mathrm{T}}\mathbf{N}_a)$$

forms the off-diagonal elements. It is known as the adjacency matrix $\mathbf{A}$ of a graph. Its elements a_{ij} show which vertices are adjacent and in multigraphs how many lines connect both vertices. For it is necessary to have in the incidence matrix identical unit rows or one row with square root of the multiplicity of the line.

The diagonal matrix $\mathbf{V}$ and the adjacency matrix $\mathbf{A}$ can be obtained as the sum or the difference, respectively, of quadratic forms of unoriented and oriented graph

Figure 12.2: Decomposition of quadratic forms $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ into the diagonal vector $\mathbf{V}$ and the adjacency matrix vector $\mathbf{A}$. $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ are orthogonal.



$$\mathbf{V} = 1/2(\mathbf{G}^{\mathrm{T}}\mathbf{G} + \mathbf{S}^{\mathrm{T}}\mathbf{S}) \tag{12.7}$$

$$\mathbf{A} = 1/2(\mathbf{G}^{\mathrm{T}}\mathbf{G} - \mathbf{S}^{\mathrm{T}}\mathbf{S}) \tag{12.8}$$

The relation of both quadratic forms is shown schematically on Fig 12.2. The Hilbert length of the diagonal vector $\mathbf{V}$ is 2m, twice the number of rows in the incidence matrices. The adjacency matrix vector $\mathbf{A}$ has the same length and it is opposite oriented in both quadratic forms, thus $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ end on different planes. If the graph is *regular*, $v_j = const$, then the diagonal matrix $\mathbf{V}$ is collinear with the unit diagonal vector $\mathbf{I}$ and the adjacency matrix $\mathbf{A}$ has the same direction, too.

The diagonal elements of the adjacency matrix $\mathbf{A}$ are zeroes. It is therefore possible to use them inconsistently for noting loops of a graph with loops. At oriented graphs rows corresponding loops are zeroes. But at unoriented graphs, the row corresponding to a loop has value 2, which gives as the quadratic 4 and using formulas 12.7 and 12.8 the loop value 2 appears automatically.

The other quadratic forms $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ and $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ have on the diagonal 2, the number of unit vectors in the rows of the incidence matrices. This is in accord with the fact that each line is registered twice in matrix $\mathbf{V}$ as well as in matrix $\mathbf{A}$. Off diagonal elements are ± 1, if two lines are adjacent having a common vertex. The off-diagonal elements form in such a way the adjacency matrices of line graphs. But at oriented graphs this explanation is complicated by signs which signs can be positive and negative. This sign pattern depends on mutual orientation of arcs. It is unpredictable and must be determined separately.

12.8 Incidence Matrices of Complete Graphs K_n as Operators

The unit matrices $\mathbf{J}$ ($\mathbf{J}^{\mathrm{T}}$) are operators which sum row (or column) elements of the matrix they are acting on, or transfer them into the resulting vector-row (or vector-column). In canonical form of incidence matrices of complete graphs K_n the unit matrices $\mathbf{J}$ are combined with the unit matrices $\mathbf{I}$ with the negative signs. The incidence matrices of complete graphs K_n are frame operators[1]. The framing operation is applied to quadratic forms of coordinate matrices twice. At first $\mathbf{C}\mathbf{C}^{\mathrm{T}}$ is framed

$$\mathbf{S}(*)\mathbf{S}^{\mathrm{T}} \tag{12.9}$$

or

$$\mathbf{G}(*)\mathbf{G}^{\mathrm{T}} \ . \tag{12.10}$$

The result of this operation is the larger matrix with $\binom{n}{2}$ rows and columns. The elements in the product are differences (sums) of all pairs of the elements of the framed matrix. The product is split into the diagonal and off-diagonal parts. The diagonal part is again framed, now in the frame collapsing diagonal elements back into n dimensional symmetrical matrix

$$\mathbf{S}^{\mathrm{T}}(*)\mathbf{S} \tag{12.11}$$

or

$$\mathbf{G}^{\mathrm{T}}(*)\mathbf{G} \ . \tag{12.12}$$

This operation forms the second difference (sum) of $\binom{n}{2}$ of the first differences (sums).

The unit diagonal matrix $\mathbf{I}$ gives $\mathbf{S}(\mathbf{I})\mathbf{S}^{\mathrm{T}}$. This is matrix $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ of the complete graph K_4. Four diagonal elements of $\mathbf{I}$ exploded into six diagonal elements of the product. The diagonal elements (2) are differences of the coordinates (or squared distances, since $\mathbf{I} = \mathbf{I}^2$) of the four vertices of the regular tetrahedron. The diagonal elements are rearranged back into four dimensions as in 12.11 or 12.12.

[1]It is curious that such elementary things can be discovered at the end of the twenties century. Maybe they were just forgotten.

12.9 Blocs Schemes

As we said it is possible to study systematically matrices with an arbitrary number of unit elements in a row. For practical reasons, this number must be constant, otherwise only special configurations were accessible for calculations. From matrices having k unit elements in each row, only matrices having specific properties corresponding to properties of complete graphs were studied. Such matrices are called *block schemes* **B** and give the quadratic forms

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = (l-r)\mathbf{I} + r\mathbf{J}\mathbf{J}^{\mathrm{T}} \ , \tag{12.13}$$

where r is the connectivity of the block. Sometimes there is posed a stronger condition on block schemes, their matrices must be the squared ones and their both quadratic form equivalent

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{B}\mathbf{B}^{\mathrm{T}} \ . \tag{12.14}$$

The unoriented complete graph K_3 is the block with $l = 3$, $r = 1$. The other K_n are not blocs, since in their $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ appear zero elements.

The equation 12.9 shows that each unit vector $\mathbf{e}_j$ must appear in the scheme l-times and each pair of elements r-times. The numbers m, n, k, l ,r are limited by following conditions

$$mk = nl \tag{12.15}$$

$$l(k-1) = r(n-1) \tag{12.16}$$

(12.15) counts the number of units in rows and columns, (12.16) the pairs in rows $mk(k-1)/2$ and in the quadratic form $rn(n-1)/2$. Dividing both sides by l/2, the result is simplified to the final form. The simplest example of a block scheme is the matrix with $m = n = 4, k = l = 3, r = 2$:

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

Block schemes with $k = 3$ are known as the Steiner's 3-tuples. It is clear that the construction of block schemes and finding their numbers is not a simple task. If you are interested, the book [8] is recommended .

12.10 Hadamard Matrices

Another special class of matrices are the *Hadamard matrices* $\mathbf{H}$ with elements $h_{ij} = \pm 1$ and quadratic forms

$$\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{H}\mathbf{H}^{\mathrm{T}} = n\mathbf{I}\,. \tag{12.17}$$

It means that all rows and columns of the Hadamard matrices are orthogonal. The examples of two lowest Hadamard matrices are:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\,.$$

The Hadamard matrices can be symmetrical as well as asymmetrical. There exist some rules how it is possible to construct Hadamard matrices of higher orders. The construction is easy at the $2n$ dimensional matrices, where the blocks of the lower matrices can be used as the building stones

$$\begin{pmatrix} \mathbf{H}_n & \mathbf{H}_n \\ \mathbf{H}_n & -\mathbf{H}_n \end{pmatrix}\,.$$

Chapter 13

Graphs

13.1 Historical Notes

The theory of graphs was formulated, similarly as many other notions in this book, by Euler. Before the World War II, all graph theory could be cumulated in only one book. Today, there are numerous specialized journals dealing with the theory of graphs and its applications.

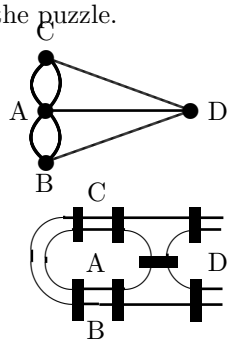
Euler formulated the basic idea of the graph theory when solving the puzzle of the seven bridges in Königsberg (Fig. 13.1). Is it possible to take a walk over all the bridges, and returning back to the starting place, crossing each bridge only once? Euler has shown that the demanded path exists only if in all the crossing points of the roads even number of the roads meet. Three roads intersected in some crossing points of the roads in the Euler's graph. Thus in Königsberg a simple path was impossible.

One wonders if such a configuration of bridges were in Athens, its philosophers on their promenades were interested in such trivial problems and if they have solved it similarly as Euler did, for all similar configurations of ways? Or, was the 7 bridge puzzle not as silly as for children? Some maturity is needed to be interested in relations which can not be seen but only imagined?

Till now all problems in this book were solved by multiplication of different possibilities and summing them. Essentially, old Greeks would have been able to solve them, but they were interested in the geometrical problems, where the problem and its solution can be seen. The possible answer to the above question is that the multidimensional spaces are too abstract to start with.

An advantage of the graph theory is that the graphs connect abstract

Figure 13.1: Seven bridges in Königsberg and the Euler's graph solution of the puzzle.

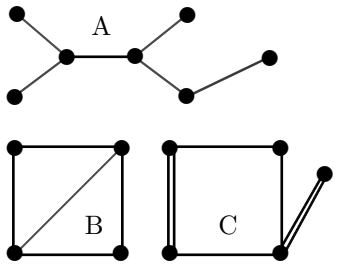


notions with concreteness. They can be drawn on paper and inspected consecutively as a system of points and lines. But this simplicity is deceiving.

Graphs are usually considered to be a binary relation of two sets, the *vertices* and the *edges* or *arcs*, see Fig. 3.2. It is possible to define a theory of anything and there appeared very interesting problems suitable to be studied by young adepts of academic degrees, as For example: the game theory. But some graph problems found very soon practical applications or analogies in physical sciences. Especially chemistry gave many impetuses for utilization of graph theory because the graphs were found to be adequate models of connectivities of atoms in molecules. It seems to be unsubstantial to study walks between vertices of graphs but when these walks are connected directly with complicated measurable physical properties of chemical compounds, as the boiling point is, then such theoretical studies become pragmatical, and give us a deep insight into how our world is constructed.

Graphs were connected with many different matrices: *Incidence matrices* **S** and **G**, *adjacency matrices* **A**, *distance matrices* **D** and other kinds of matrices. All these matrices were used for calculations of eigenvalues and eigenvectors, but the different matrices were not connected into an unified system. Mathematicians were satisfied with the fact, that all graphs can be squeezed into three dimensional space and mapped onto two dimensional paper surface. They ignored the problem of dimensionality of graphs. Different authors considered them to be dimensionless objects, one dimensional objects, two dimensional objects. According to the Occam's razor, there should not be introduced more factors than necessary to explain observed facts. But treating graphs as multidimensional vectors with special configurations unifies the theory, graphs are just a special class of vectors,

Figure 13.2: Examples of unoriented graphs. A – a tree, B – a cycle graph, C – a multigraph.



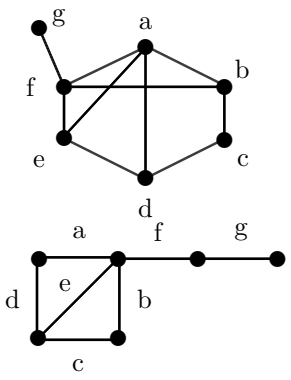
sums, or differences of two vector strings. These vectors belong into the vector space. Properties of sums or differences of two vector strings can be studied conveniently if they are imagined as graphs, compared with existing objects, or at least with small samples of larger structures.

13.2 Some Basic Notions of the Graph Theory

The graph theory has two basic notions. The first one is the *vertex* which is usually depicted as a point, but a vertex can be identified with anything, even with a surface comprising many vertices, if the graph theory is applied to practical problems. The second notion is the *line* representing a relation between two vertices. Lines can be *oriented*, as vectors are, going from a vertex into another, then they are called *arcs*, and/or *unoriented*, just connecting two vertices without any preference of the direction. Then they are called *edges* (Fig. 3.2).

An arc is represented by a row of the incidence matrix **S** formed by the difference of two unit vectors $(\mathbf{e}_i - \mathbf{e}_j)$. According to our convention, both vectors act simultaneously and the beginning of the vector can be placed on the vertex j. The resulting arc vector goes directly from the vertex j into the vertex i. An edge is depicted as a simple line connecting two vertices. Actually the sum of two unit vectors is orthogonal to the line connecting both vertices. It is more instructive to draw an unoriented graph with connecting lines. Nevertheless, for formal reasons we can consider an unoriented graph as a string vectors where each member is orthogonal to its oriented matching element. When the oriented graph is a vector, then the unoriented graph must be a vector, too.

Figure 13.3: Graph and its line graph.



A special line in graphs is the *loop* which connects a vertex with itself. Formal difficulties appear, how to connect oriented loops with matrices, because corresponding rows are zero $(\mathbf{e}_j - \mathbf{e}_j) = \mathbf{0}$. These complications are resulting from the higher order symmetries. An unoriented loop has a double intensity

$$(\mathbf{e}_j + \mathbf{e}_j) = 2\mathbf{e}_j \ , \tag{13.1}$$

and we will see later, how this fact can be exploited.

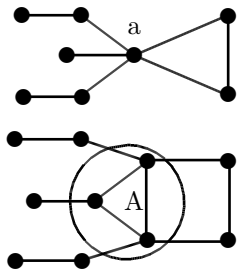
Relations between things can be things, too. For example: in chemistry, if we identify atoms in a molecule with vertices, then bonds between atoms, keeping the molecule together, and determining the structure of the molecule, are bonding electrons. The forces between the nuclei and the electrons are modeled by graphs if into each connecting line a new vertex is inserted and so a *subdivision graph* is formed. Each line in the graph is split into a pair of lines. The generated subdivision graph has (n + m) vertices and 2m lines.

We can construct a *line graph* 13.3, changing lines into vertices, and introducing new incidences defined now by the common vertices of two original lines. If the parent graph had m edges, the sum of its vertex degrees v_j was $2m$. Its line graph has m vertices and the sum of its vertex degrees v_i is

$$\Sigma(v_j^2 - v_j) \ . \tag{13.2}$$

A pair of vertices can be connected by more lines simultaneously. Then we speak about *multigraphs* (13.2, C). Next step is to consider the parallel

Figure 13.4: Restriction of a graph. Vertices in the circle A are joined into one vertex a.



lines as one line with the weight k. It is obvious that the lines need not to be weighted by whole numbers but any weights w_{ij} can be used. From calculations emerge even graphs with imaginary lines.

It is also possible to restrict the graphs by grouping sets of vertices into new vertices and leaving only the lines connecting the new set vertices (Fig. 13.4). This operation simplifies the graph.

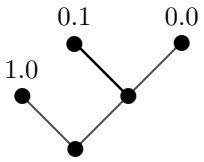
Both elements of the graphs can be indexed (labeled) and unindexed (unlabelled). Usually only vertex labeled graphs are considered. Labeled graphs are sometimes only partially indexed graphs, when only some of their vertices are indexed, or equivalently, several vertices have equal indices. When one vertex is specially labeled, we speak about the *root*.

A special labeling of graphs is their coloring. A task can be formulated to color the vertices in such a way that no incident vertices had the same color. The number of colors indicates the parts of the graph where all vertices are disconnected. No line exists between them. The least number of colors which are necessary to color a connected graph is 2. Then we speak about *bipartite graphs*. For coloring of the planar graphs (cards), which lines do not intersect, we need at least four colors.

Bipartite graphs have an important property, their incidence matrices can be separated into two blocks and their quadratic forms split into two separate blocks.

Graphs are *connected*, if there exists at least one path or walk between all pairs of vertices. It is uninterrupted string of lines connecting given pair of vertices. Mutually unconnected parts of a graph are known as its *components*. At least $(n-1)$ lines are needed to connect all n vertices of a graph and n lines to form a cycle. Connected graphs with $(n-1)$ lines are known as *trees* (13.2), and they are acyclic. A graph formed from more trees is the *forest*.

Figure 13.5: Decision tree. The left branch means 1, the right branch means 0. The root is taken as the decimal point and the consecutive decisions model the more valued logic.



We can find the *center* of a graph, determined as its innermost vertex, or the *diameter* of a graph, as if they were some solid objects. But there appears some difficulties. When we define the center of a graph as the vertex which has the same distance from the most distant vertices, then in linear chains with even number of vertices For example: in the linear chain L_6

we have two candidates for the nomination. It is better to speak about the *centroid* or the *central edge*. Some graphs have no center at all.

To introduce all notions of the graph theory consecutively in a short survey is perplexing. But it is necessary to know some terms.

The *linear chains* L_n are a special class of trees which all vertices except two endings have the degree $v_j = 2$. The vertex degree counts lines incident to the vertex. Linear chains have the longest distance between their extremal vertices and the greatest diameters from all graphs. Another extremal trees are the *stars* S_n. All $(n-1)$ of their vertices are connected to the central vertex directly. The diameter of stars is always 2. The *decisive trees* are trees with one vertex of the degree 2 and all other vertices with degrees 3 or 1. If the vertex of the degree 2 is chosen as the root (Fig. 19.3) then on a walk it is necessary to make a binary decision on each step which side to go. The vertices with degrees 1 are known as the *leaves*. They are connected by the *branches* to the *stem* of the tree. We already know the decision trees as strings in the unit cubes. In a tree, they are joined into bifurcating branches. The indexing of the leaves is known as the binary coding.

The *complete graph* K_n has n(n-1)/2 lines which connect mutually all its vertices. Its diameter is 1 and it has no center. The *complement* $\overline{G}$ of a graph G is defined as the set of lines of the graph G missing in the complete graph K_n on the same vertices, or by the sum

$$K_n = G + \overline{G}\ . \tag{13.3}$$

It follows, that the complementary graph of the complementary graph $\overline{\overline{G}}$ is the initial graph G and that the complementary graph $\overline{K_n}$ of the complete graph K_n is the empty graph G_n with no lines.

13.3 Petrie Matrices

The arcs of the oriented graphs were defined as the differences of two unit vectors $(\mathbf{e}_j - \mathbf{e}_i)$. There is another possibility of mapping arcs and edges on matrices with the unit elements.

An arc is identified directly with the unit vector ($\mathbf{e}_j$ or with a continuous string of the unit vectors $\mathbf{e}_j$. Such matrices are known as the *Petrie matrices*[1] $\mathbf{P}e$.

The Petrie matrices are equivalent to the incidence matrices. A row containing a continuous string of unit symbols corresponds to each an arc of the incidence matrix without an interruption. The string of unit symbols in a Petrie matrix $\mathbf{P}e$ going from i to (p-1) corresponds to the arc between vertices i and p. The arc 1-2 is represented in a Petrie matrix by one unit symbol, the arc 1-6 needs 5 unit symbols.

The canonical forms $\mathbf{P}e$ and $\mathbf{S}$ of K_4 are

$$\begin{array}{cc} \mathbf{P}\mathrm{e} & \mathbf{S} \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \end{array} .$$

Petrie matrices have two important properties:

1. A Petrie matrix $\mathbf{P}e$ of a graph G multiplied by the incidence matrix $\mathbf{S}$ of the linear chain L gives the incidence matrix of the given graph:

$$\mathbf{S}(G) = \mathbf{P}e(G)\mathbf{S}(L)\ . \tag{13.4}$$

From the consecutive units in a row of the Petrie matrix only the first and the last are mapped in the product, all intermediate pairs are annihilated by consecutive pairs of unit symbols with opposite signs from the

[1] There is not enough of simple symbols for all different matrices.

incidence matrix of the linear chain, which vertices are indexed consecutively: $1-2-3-\ldots-n$. For example

			1	1	0	0				−1	1	0	0
			0	−1	1	0				0	−1	1	0
			0	0	−1	1				0	0	−1	1
1	0	0	−1	1	0	0	1	0	0	−1	1	0	0
0	1	0	0	−1	1	0	1	1	0	−1	0	1	0
0	0	1	0	0	−1	1	1	1	1	−1	0	0	1

2. Only the Petrie matrices of trees are nonsingular. The trees have $(n-1)$ arcs. Therefore their Petrie matrices are square matrices and because trees are connected graphs, their Petrie matrices are without empty columns. The importance of this property will be clear in the Chapt. 15.

13.4 Matrices Coding Trees

The Petrie matrices define the trees in space of arcs. The another possibility of coding trees is in the space of their vertices. There exist the *descendant code matrices* and their inverses, showing the relation of vertices as the relation of children to parents. In the descendant code both ends of arcs are used, but the vertices on the path only once. Moreover, the root itself is induced as the element $\mathbf{e}_{11}$ in the first row. The convention is, that the arcs are always going from the root. The resulting code[2] has the matrix $\mathbf{C}$ the lower triangular form and on the diagonal is the unit matrix $\mathbf{I}$. At trees, the first column is at trees the unit matrix $\mathbf{J}$, but the code allows forests, too.

The inverses of code matrices $\mathbf{C}^{-1}$ are in the lower triangular form, and the unit matrices $\mathbf{I}$ are on the diagonal. The off diagonal elements are -1 when the vertex j is the child of the vertex i, and 0 otherwise. Since each child has only one parent, two nonzero elements are in each row, except the first one, and this part of the matrix is the incidence matrix $\mathbf{S}$ of the given tree. Therefore

$$(\mathbf{S} + \mathbf{e}_{11}) = \mathbf{C}^{-1} \; . \tag{13.5}$$

The element $\mathbf{e}_{11}$ is the vector going from the origin of the coordinate system to the vertex 1, or using the graph convention, the arc going from the vertex 0 to the vertex 1. In this case, the zero column containing one -1 element is deleted.

[2]The code matrix $\mathbf{C}$ is simultaneously the coordinate matrix.

For our purposes it is necessary to allow any vertex to become the root without changing the indexes. For this reason, we define the path matrix as vertices on the path between the vertex i to the root j. This is only a permutation of the lower triangular form. For example:

$$\begin{array}{cc} \mathbf{C} & \mathbf{C}^{-1} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} & \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 \end{pmatrix} . \end{array}$$

The permutation of the columns is (3,1,4,2). The row with the unit element is inserted into the incidence matrix as the second one and all arcs are going from the vertex 2.

We already applied the code matrix of the linear chain L_n and its inverse as the operators $\mathbf{T}^{\mathrm{T}}$ and $\mathbf{C}^{-1}$ in Sect. 4.3. Recall that

$$\begin{array}{cccc|cccc} & & & & 1 & 0 & 0 & 0 \\ & & & & 1 & 1 & 0 & 0 \\ & & & & 1 & 1 & 1 & 0 \\ & & & & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1. \end{array}$$

Inspecting it, we see, that $\mathbf{C}^{-1}$ is incidence matrix of the linear chain L_4 which singularity was removed by adding the row with one unit element 1_{11}. For such rooted incidence matrices we will use the starred symbol $\mathbf{S}*$. Similarly the incidence matrices of all trees can be adjusted. The code matrices $\mathbf{C}$ are just their inverses $(\mathbf{S}*)^{-1}$.

It seems that the distinction between the Petrie matrices $\mathbf{P}e$ and the code matrices $\mathbf{C}$ is due to the unit column $\mathbf{J}$ which transforms $(n-1)$ square matrix to n dimensional square matrix. But both sets are different.

The incidence matrices of trees $\mathbf{G}*$ rooted by the unit column $\mathbf{J}$ are nonsingular and they have inverses $\mathbf{G}^{-1}$ which in their turn are code matrices $\mathbf{C}$ of unoriented trees. These code matrices $\mathbf{C}$ must contain negative elements.

For example, for the star we get using the principle of inclusion and exclusion

$$
\begin{array}{cccc|cccc}
 & & & & 1 & 0 & 0 & 0 \\
 & & & & -1 & 1 & 0 & 0 \\
 & & & & -1 & 0 & 1 & 0 \\
 & & & & -1 & 0 & 0 & 1 \\
\hline
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1\,.
\end{array}
$$

The incidence matrices of the unoriented stars $\mathbf{S}*$ and the oriented stars $\mathbf{G}*$ are selfinverse.

Chapter 14

Enumeration of Graphs

14.1 Introduction

We treated enumeration of the naive matrices **N** in detail. To enumerate their sums and differences, known as unoriented and oriented graphs, respectively, is more complicated problem. Therefore only some problems of graph enumeration will be discussed.

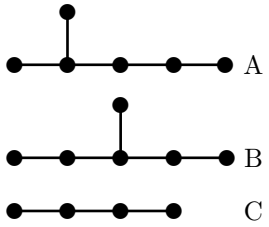
14.2 Enumeration of Trees

Acyclic connected graphs, known as the trees, form the base of the graph space. We explain later why, now we only show some complications of enumeration of graphs on them, as compared to the naive matrices.

Every tree which vertices are labeled can be connected with a string of symbols using the Prüfer algorithm: We choose the pending vertex with the lowest label, mark its neighbor and prune it from the tree (its branch is cut and discarded). This pruning is repeated till from the original tree only $K_2 = L_2$ remains. In such a way we obtain a string of $(n-2)$ symbols. If all n vertices of the original tree had a specific label, then there is obviously n^{n-2} strings corresponding to all possible labeling of trees. For example: L_5 1-5-4-3-2 gives 5,3,4, L_5 2-1-4-3-5 gives 1,4,3. The sequence 4,4, 4 is obtained by pruning the star S_5 rooted in 4.

These strings can be counted by a modified Equation 10.2. A tree has $(n-1)$ edges and the sum of vertex degrees v_j is $\sum v_j = 2(n-1)$. The least possible vertex degree of pending vertices is 1. n vertex degrees are bounded, therefore only $(n-2)$ units can be partitioned in trees. Therefore, we get

Figure 14.1: The smallest pair of graphs on the same partition orbit (A and B) and the graph with a central edge (C).



$$\text{Number of trees } = n^{n-2} = \sum (n!/\prod_k n_k!)([n-2]!/\prod_k (v_k-1)^{n_k} \ . \quad (14.1)$$

The sum is made over all partitions of $(n-2)$ into n parts and v_k replaces m_k.

The equation 14.1 counts trees successfully, but there appears one inconvenience: Different types of trees are counted together when they have the same partition structure. The partition orbits split in the graphs into the suborbits. The smallest pair of trees split in such a way, two different trees on the orbit 322111 are on Fig. 14.1.

The partition orbits are split into the graph orbits with different structures. Similar vertices of graphs are known as the *orbits of a graph*. This use of one notion on different levels is somewhat confusing[1].

The tree A on Fig. 14.1 has 5 different orbits and B only 4. The number of different edges, connecting vertices on different orbits, is lesser than the number of vertex orbits, except for the symmetric edges connecting vertices on the same orbit, as the central edge is in C on Fig. 14.1.

We must explain why the partition orbits are important and find techniques how to count the number of unlabelled trees. But before that we mention another two problems connected with the labeling of trees.

Trees, similarly as other graphs, can be built from trees of lower dimensions. If we use the technique of the Young tables, that is inscribing indices into the Ferrers graphs, we obtain the Young labeled trees. Starting from K_2, there are always $(n-1)$ opportunities how to attach the n-th vertex to $(n-1)$ vertices of trees of the lower level and the number of the

[1]The Sun has its planets and the planets have in their turn their trabants all with their own orbits.

Table 14.1: Trees generated by the polynomial $(x(x+m)^{m-1})$ and the inverse matrix

	1	2	3	4	5	$\sum$	1	2	3	4	5
m=1	1					1	1				
2	2	1				3	-2	1			
3	9	6	1			16	3	-6	1		
4	64	48	12	1		125	-4	24	-12	1	
5	625	500	150	20	1	1296	5	-80	90	-20	1

Young labeled trees must be $(n-1)!$. These trees can be compared to the convolutions.

All trees are generated by the polynomial

$$x(x+m)^{m-1}\ , \tag{14.2}$$

where m is the number of edges in a tree with $(m+1)$ vertices. The powers of x can be interpreted as the number of edges connected to the added vertex forming the root and the terms of the polynomial at x^k can be interpreted as the number of trees rooted in the n-th vertex having the corresponding vertex degree k. For example: for $m=4$ we get:

$$64x^1 + 48x^2 + 12x^3 + 1x^4 = 125\ .$$

16 trees with 4 vertices are attached to the fifth vertex at 4 different places This gives the first coefficient. The second coefficient is obtained by rooting $(L_3 + K_1) = 3 \times 12$ and $2K_2 = 3 \times 4$. The last term corresponds to the star rooted in the fifth vertex.

So we got the new combinatorial identity, which can be tabulated in Table 14.1 together with its inverse matrix

The elements of the inverse matrix can be decomposed into binomial coefficients $\binom{m}{j}$ and elements $j^{(i-j)}$. The next row of the inverse matrix is $-6 \times 1 + 15 \times 16 - 20 \times 27 + 15 \times 16 - 6 \times 5 + 1 \times 1$.

For counting unlabelled trees it is necessary to find the number of rooted trees orbit and the number of the rooted trees with symmetrical edges.

14.3 Symmetry Group of Unoriented Graphs

The incidence matrix $\mathbf{G}$ of the complete unoriented graph K_n has n columns and $n(n-1)/2$ rows. In each column there are $(n-1)$ unit elements and in each row there are two unit elements. Different combinations of pairs

Table 14.2: Relation between S_n and G_n groups

S_n group		s_1^4	$s_1^2 s_2^1$	$s_1^1 s_3^1$	s_2^2	s_4^1
P		1 0 0 0	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0
		0 1 0 0	1 0 0 0	0 0 1 0	1 0 0 0	0 0 1 0
		0 0 1 0	0 0 1 0	1 0 0 0	0 0 0 1	0 0 0 1
		0 0 0 1	0 0 0 1	0 0 0 1	0 0 1 0	1 0 0 0
Initial row	$\mathbf{G}_{K_4}$	Permuted rows (the index of the original row)				
1	1 1 0 0	1	1	2	1	2
2	1 0 1 0	2	3	3	6	4
3	0 1 1 0	3	2	1	1	5
4	1 0 0 1	4	4	6	4	6
5	0 1 0 1	5	6	4	3	3
6	0 0 1 1	6	5	5	2	1
G_4 group		s_1^6	$s_1^2 s_2^2$	s_3^2	$s_1^2 s_2^2$	$s_2^1 s_4^1$.

of the unit vectors correspond to different edges of the graph and can be indexed consecutively by the index i, going from 1 to $n(n-1)/2$.

The incidence matrix $\mathbf{G}$ can be permuted from the left by permutation matrices $\mathbf{P}_{n(n-1)/2}$ forming the group of cyclic permutation $S_{n(n-1)/2}$ and from the right by permutation matrices $\mathbf{P}_n$. These permutations of n columns form the group S_n of cyclic permutations which changes permutations of the larger left hand group $S_{n(n-1)/2}$. This group of graph edges can not be complete, because it is induced by the lesser cyclic group S_n. We will use for the graph group induced by permutations of columns of the incidence matrix $\mathbf{G}_n$ simple notation G_n. In mathematical literature different names are used, as the "wreath product" or the "step group".

In Table 14.2, effects of cyclic permutations on the incidence matrix of the complete graph K_n are shown.

The indexing of the graph edges is done recursively. To the graph with n edges a new vertex is added and to the incidence matrix $\mathbf{G}_n$ a new block having the block form of two unit matrices $(\mathbf{I}_n|\mathbf{J}_n)$. The subgroup $s_1^1 S_n$ of the group S_{n-1}, which leaves the last column in its place, permutes only the elements of the matrix $\mathbf{G}$, but its effect transforms the unit cycle s_1 with one element into n elements of the acting permutation matrix and transforms its cycle structure which adds new cycles to the existing structure of the graph group.

Of course, the group S_{n+1} contains also other subgroups than $s_1^1 S_n$. One of them is the subgroup of the simple cycles s_{n+1}. Each cycle with uneven length k transforms the $(n+1)$-th unit cycle into a new cycle of the same length. In our example $(s_1^1 + s_1^3)$ transforms into

$$(s_1^3 + s_1^3) = s_1^6 \text{ and } (s_1^1 + s_3^1) \text{ into } (s_3^1 + s_3^1) = s_3^2 \,. \tag{14.3}$$

Cycles with even length transform the added unit cycle into two cycles, one having the same length as the original cycle and the other with half length. For this case, we have in our example the cycles of the length 2:

$$[s_1^1 + (s_1^1 s_2^1)] = (s_1^1 + s_1^1 s_2^1) = s_1^2 s_2^2 \,.$$

Actually, each element of a cycle of the length n acts on $(n-1)/2$ induced elements of the group G_n. If n is odd, $(n-1)/2$ is a whole number and if n is even, there remain $n/2$ edges which are permuted and form a new cycle. In our example s_4 generated the new cycle s_2 because the complete graph K_4 has 6 edges. In K_6 with 15 edges, s_6 produces the cyclic structure $s_3^1 s_6^2$.

When there are two cycles of different length, which have not a common divisor, they induce as many cycles as their common divisor is of the length which is equal to their lowest multiple. For example: at $n = 5 : 2 \times 3 = 6$, and there remain 4 elements to be permuted by smaller cycles. This is possible as $s_1^1 s_3^1$. The cycle s_1 is induced by the cycle s_2 which permutes two vectors of only one edge and leaves the identity. The cycle s_3 permutes columns of three edges only and reproduces itself. Some examples of induced subgroups of $S(a)$ and the corresponding graph cycles:

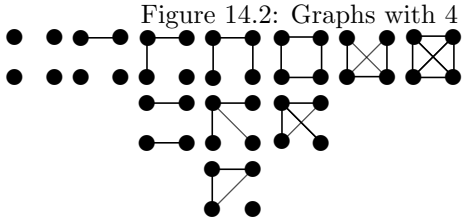
$$\begin{array}{llllll} S_6 & s_6^1\,; & S_7 & s_1^1 s_6^1\,; & S_8 & s_2^1 s_6^1\,; \\ G_6 & s_3 s_6^2\,; & G_7 & s_3 s_6^3\,; & G & s_1^1 s_3^1 s_6^4\,. \end{array}$$

It is possible to generate any graph group either by computing results of multiplication of the incidence matrices by different permutation matrices, or by deducing effects of different cycle structures. Both ways are tedious tasks demanding patience and/or computers. If we remind yourselves that we dealt with sums of only two naive matrices, where all operations seemed easy, we wonder how complicated must be groups of matrices having in each row three or more unit symbols, or groups of matrices of different kinds.

The graph groups G_n can be used for determining the number of all simple graphs with n vertices, similarly as the cycle indices were used. An edge can be present in a graph or not. In a simple graph multiple edges are not allowed and we can imagine that the graphs are placed on vertices of $n(n-1)/2$ dimensional unit cubes which sides are formed by diagonals as on Fig. 12.1, where two diagonal strings in the 3 dimensional cubes were shown.

To represent both possibilities, we insert into the cycle indices the polynomial $(1 + x^k)$ instead of cycles s_k and calculate for all subgroups. The G_4 graph index is

Figure 14.2: Graphs with 4 vertices and k edges



$$G_4 = 1/24\ (s_1^6 + 9s_1^2s_2^2 + 8s_3^2 + 6s_2^1s_4^1)\ . \tag{14.4}$$

It gives

$$Z(G_4,\ 1+x) = 1 + x^1 + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6\ , \tag{14.5}$$

where coefficients at x^k determine the number of different graphs with 4 vertices and k edges. They are shown on Fig. 14.2.

14.4 Symmetries of Unoriented Graphs

We explained the graph group by permutations of columns of the incidence matrix **G** of the complete graph. Now we use this technique to explain the symmetry of other unoriented graphs, which are essentially subsets of k elements of the complete graph.

There are only two possibilities of what a permutation columns of the incidence matrix can do with the rows. A row can be permuted in itself or it can be changed into a row corresponding to another edge. The group of the single edge has two elements: (1)(2) and (12). If the edge is defined on the set of 4 vertices, then there are 4 permutations which leave it unchanged: (1)(2)(3)(4), (1)(2)(34), (12) (3)(4), and (12)(34). We can choose 6 different edges but some will have equal groups, as the edge 3-4 with the edge 1-2.

With k rows we have always three possibilities: Permutations acting on vertices change only the ordering of rows, that is their indexing. Or they change them completely (or at least partially) into rows corresponding to other edges. The result is that one labeled graph is changed into another labeled graph which must have the same number of edges and must belong to the same graph type.

The count of the number b of the permutations which only permute the rows of the incidence matrix **G** of a graph determines the symmetry

of the graph. When we divide the number of all permutations $n!$ by the *symmetry number* b, we obtain the number of the different labeled graphs of the given type. b of a single edge on 4 vertices is 4, and there are indeed $24/4 = 6$ different edges on the set of 4 vertices. The symmetry number of this graph K_4 is 24 and therefore there is only one distinguishable labeling of this graph. The relation of the number b to distinguishable labeling is known as the *Burnside lemma*.

Now we inspect the calculations according to (14.3). The formula

$$(1+x)^6 + 9(1+x)^2(1+x^2)^2 + 8(1+x^3)^2 + 6(1+x^2)(1+x^4) \qquad (14.6)$$

is divided according to its increments in the final result as:

Powers of x	0	1	2	3	4	5	6
s_1^6	1	6	15	20	15	6	1
$9s_1^2s_2^2$	9	18	27	36	27	18	9
$8s_3^2$	8			16			8
$6s_2^1s_4^1$	6		6		6		6
Σ	24	24	48	72	48	24	24
Number of graphs	1	1	2	3	2	1	1

All permutation matrices of the group S_4 transform the empty or complete graph into itself. Therefore their b = 24. When we divide the column sums by 24, we obtain the number of different graphs with k vertices. The number of the distinguishable labeled graphs is given in the first row, where the identity permutations are counted. For a single edge, it gives six different graphs. The number b is produced by three permutations of rows of the type $s_1^2s_2^2$ and by one permutation s_1^6.

At graphs with two edges, 15, 27, and 6 permutations belong to two different graphs, either L_3 and one isolated vertex or two L_2. When we try to divide the permutations into the graph orbits, we can use the fact that both b and the number of different labeling of a graph must be divisors of $n!$. 15 can then be split only as $12 + 3$. Then 27 can be divided as $12 + 12 + 3$. We can use also another criterion, to decide which from both possibilities is right. We exploit possible partitions of vertex degrees. Graphs with two edges have the sum of vertex degrees 4 and for 4 vertices two partitions: 2110 and 1111. There are 12 distinguishable permutations of the first partition and only 1 permutation of the second one. This partition is stable at all permutations, including the cycle of the length 4, therefore the group structure is s_4^1. Both criterions leave as the only possible splitting $12 + 12 + 3$. There are 12 linear chains L_4 with $b = 2$ and the group structure $(s_1^4 + s_2^2)$, and 3 graphs $2K_2$ with $b = 8$. Their group

structure is $s_1^4 + 2s^2 s_2^1 + 3s_2^2 + 2s_4^1$. The graphs with five and six edges are complementary the graphs with none and one edge.

14.5 Oriented graphs

In an simple oriented graph, two arcs between each pair of vertices can exist. The symmetry of the oriented graphs is complicated by this fact. This can be documented on the relation between the number of selfcomplementary unoriented graphs with 4k vertices and the number of selfcomplementary tournaments with 2k vertices. A *tournament* is a connected oriented graph which can have only one from both orientations of arcs.

The complete tournament with 2k vertices has $(4k^2 - 2k)$ arcs, the complete oriented graph with 4k vertices has $(8k^2 - 2k)$ arcs. It is necessary to complete a graph corresponding to a selfcomplementary tournament with 2k vertices, and to generate from each arc two arcs. It can be done as follows: We generate 2k new vertices indexed by dashed indices of the tournament and we connect all odd dashed and undashed vertices having equal index by k arcs. If in the tournament the arc i-j exists, we induce arcs i-j and i-j' in the complementary graph, if there is the arc j-i, we introduce arcs i'-j and i'-j'. The arcs missing in the induced graph are present in the selfcomplementary graph, they correspond to the arcs in the complementary tournament or connect even dashed and undashed vertices. The difference is formed by $4k^2$ arcs and $2k$ vertices.

The difference between the oriented and unoriented graphs can be explained also in a other way. We can use two separate rows for both orientations of the arcs i-j and j-i, respectively. In an simple oriented graph, n(n-1) arcs can be, that is twice the number of the edges. The incidence matrix **S** has twice the number of rows of the incidence matrix **G** and permutations of its columns induce another kind of column permutations changing the signs. The permutation (12)(3)(4) induces the graph permutations (12)(23)(45)(6) (89)(10,11)(12) which leaves only two rows unchanged.

We already mentioned that the number of labeled unoriented graphs is $2^{n(n-1)/2}$. This can be written as the polynomial

$$G(t) = (1+t)^{\binom{n}{2}} \text{ with } t = 1 \, . \tag{14.7}$$

This fact is deduced, from the possibilities of how to fill the adjacency matrix **A** with unit symbols. There are $\binom{n}{2}$ possibilities which are independent. The adjacency matrix is symmetrical. The filling goes simultaneously in the lower and upper off-diagonal positions. The polynomial $G(2)$ gives the number of labeled oriented graphs with only one arc between a pair of

vertices. This corresponds to an adjacency matrix which has only one element in each pair of positions i-j and j-i showing the orientation of the arc. The polynomial $G(3)$ gives the number of oriented graphs with both orientations of arcs, or the number of asymmetrical adjacency matrices which can have a pair of unit symbols in each pair of the corresponding places.

14.6 Connected Unoriented Graphs

A graph is connected if it has only one component. The number of unoriented connected graphs can be determined if we count all graphs rooted in one component C_k with k vertices. Their number is equal to the number of all rooted labeled graphs

$$n2^{n(n-1)/2} = \sum_{k=1}^{n} \binom{n}{k} C_k G_{n-k} \ , \tag{14.8}$$

where G_{n-k} is the number of all rooted graphs with $(n-k)$ vertices, that means $n2^{(n-k)(n-k-1)/2}$ with $G_0 = 1$. The meaning of the left hand side of the identity is clear: Each graph has n possible roots. The right hand side counts each graph according to the number of its components. If it has two components, then it is counted twice, once with k roots, then with $(n-k)$ roots. The empty graph is counted n times due to the binomial coefficient on the right hand side.

When we separate the number of connected graphs C_n, we can determine the number of all rooted labeled graphs recursively. Here a general relation between two generating functions is applied, the normal and the exponential ones. The numbers of the connected labeled graphs G_n are the coefficients of the exponential generating function of the labeled graphs

$$G(x) = \sum_{n=1}^{\infty} C^n(x)/n! = \exp(\sum_{n=1}^{\infty} a^n x^n) \ . \tag{14.9}$$

Because there exists also the normal generating function of the labeled graphs

$$G(x) = \sum_{n=1}^{\infty} A^n x^n) \ , \tag{14.10}$$

both functions can be compared. Inserting $a_0 = 1$, we can logarithm both sides with the result

$$a_n = A_n - 1/n \sum_{n=1}^{\infty} k a_k A_{n-k}) \; . \tag{14.11}$$

(14.8) appears to be just a special case of this identity.
The number of connected graphs C_n is a fast growing function

n	1	2	3	4	5	6
C_n	1	1	4	38	728	26704

.

Chapter 15

Eigenvalues and Eigenvectors

15.1 Interpretation of Eigenvalues

The quadratic forms of the naive matrices $\mathbf{N}^{\mathrm{T}}\mathbf{N}$ are diagonal matrices. Also squares of Hadamard matrices are diagonal matrices. But the second quadratic forms of naive matrices $\mathbf{N}\mathbf{N}^{\mathrm{T}}$ and the quadratic forms of the incidence matrices of graphs **G** and **S** have off-diagonal elements. We interpreted the diagonal and the off-diagonal elements as two orthogonal matrix vectors, giving the unit projections of any matrix vector **M** into the space of rows and columns (see Fig.1.6). In this chapter we will show conditions when a matrix vector can be represented by an equivalent diagonal matrix of *eigenvalues* introduced in Sect. 3.5 and the properties which such a substitute has.

When compared with the naive matrices, one property is clear: The diagonal matrix must have the same length as the matrix vector **M** itself. From this property follows that at the diagonalization, the matrix vector **M** is rotated to decrease the importance of the off-diagonal elements. Alternatively, the vector position is stable and we move the coordinate system, exactly as if we were going around the matrix till a place is found, from where it is possible to see through. Such a point of view has its own set of coordinates.

The going around the matrix is similar with the function of the polarizing filters rotating the light (the set of the eigenvalues is known as the *spectrum* of the matrix) has a pair of matrices known as matrices of *eigenvectors*. The matrix **M** is put between a pair of eigenvector matrices $\mathbf{Z}^{\mathrm{T}}$

and $\mathbf{Z}$ and the resulting product is the equivalent diagonal matrix $\Delta(\mathbf{M})$:

$$\mathbf{Z}^{\mathrm{T}}\mathbf{M}\mathbf{Z} = \Delta(\mathbf{M}) \,. \tag{15.1}$$

In Sect. 3.5, symbols $\mathbf{L}$ and $\mathbf{R}$ were used for both diagonalizing matrices. The difference between these matrices and the eigenvectors is due to the additional demand on the eigenvectors. The eigenvectors are the diagonalizing vectors which are normalized as in the following examples

		$1/\sqrt{2}$	$1/\sqrt{2}$
		$1/\sqrt{2}$	$-1/\sqrt{2}$
0	1	$1/\sqrt{2}$	$-1/\sqrt{2}$
1	0	$1/\sqrt{2}$	$1/\sqrt{2}$
$1/\sqrt{2}$	$1/\sqrt{2}$	1	0
$1/\sqrt{2}$	$-1/\sqrt{2}$	0	-1

		$1/\sqrt{2}$	$1/\sqrt{2}$
		$1/\sqrt{2}$	$-1/\sqrt{2}$
2	1	$3/\sqrt{2}$	$1/\sqrt{2}$
1	2	$3/\sqrt{2}$	$-1/\sqrt{2}$
$1/\sqrt{2}$	$1/\sqrt{2}$	3	0
$1/\sqrt{2}$	$-1/\sqrt{2}$	0	1

The situation is complicated when more eigenvalues are equal and the corresponding values are multiple.

Notice two important properties of eigenvector matrices:

- 1. Their column vectors should be orthogonal and normalized

$$\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{I} \,. \tag{15.2}$$

For example:

		$2^{-1/2}$	$2^{-1/2}$
		$2^{-1/2}$	$-2^{-1/2}$
$2^{-1/2}$	$2^{-1/2}$	1	0
$2^{-1/2}$	$2^{-1/2}$	0	1

Sometimes it is difficult to find the orthogonal eigenvectors, if more eigenvalues are equal (or one eigenvalue is multiple).

- 2. When eigenvectors multiply the matrix $\mathbf{M}$, all its elements are multiplied by the factor corresponding to the eigenvalue λ_j. In other words, the matrix $\mathbf{M}$ behaves to its eigenvector matrices $\mathbf{Z}^{\mathrm{T}}$ and $\mathbf{Z}$ as a diagonal matrix of the eigenvalues

$$\mathbf{MZ} = \lambda_j \mathbf{Z} \, . \tag{15.3}$$

15.2 Eigenvalues and Singular Values

All the above equations were written for the quadratic matrices $\mathbf{M}$, representing quadratic forms. For rectangular matrices we can fill their missing row or column elements by zeroes and for any vector taken as the eigenvector, we obtain a zero eigenvalue. We will not be interested in the eigenvalues of the rectangular matrices, but in the eigenvalues of their quadratic forms, which are known as the *singular values* of the rectangular matrices and of the asymmetric square matrices.

The incidence matrix $\mathbf{S}$ of a tree is $(n-1) \times n$-dimensional. $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ is n-dimensional matrix, $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ is $(n-1)$-dimensional matrix. Both products have the same sets of singular values. In this case $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ must have one zero λ_j. This is true for all connected graphs. The square unit matrix $\mathbf{J}\mathbf{J}^{\mathrm{T}}$ has only one nonzero eigenvalue which is identical with the eigenvalue of $\mathbf{J}^{\mathrm{T}}\mathbf{J}$. This is the sum of n units.

We repeat once again the important fact that on the diagonals of both quadratic forms as well ason the diagonals of squared symmetric matrices appear the squared elements m_{ij}. If a matrix is symmetric, both quadratic form coincide with the square of the matrix $\mathbf{M}^{\mathrm{T}}\mathbf{M} = \mathbf{M}^2$, therefore the singular values of the symmetric matrices coincide with their squared eigenvalues.

15.3 Characteristic Polynomials

Now we approach to the problem of the eigenvalues in another aspect. A matrix and the matrix of its eigenvectors form a system of linear equations which solutions are found when consecutively the diagonal matrix of eigenvalues $\Delta(\lambda)$ is subtracted from the matrix $\mathbf{M}$ and the resulting matrix is multiplied by the eigenvector $\mathbf{z}$:

$$(\mathbf{M} - \lambda\mathbf{I})\mathbf{z} = \mathbf{0} \, . \tag{15.4}$$

For example: the matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

corresponds to the equations

$$(2-\lambda)x + y = 0$$

and

$$x + (2-\lambda)y = 0\ .$$

If we insert the eigenvector $x = 1,\ y = 1$, we get as the solution $\lambda = 3$, for $x = 1,\ y = -1$, the eigenvalue is $\lambda = 1$. We already know the eigenvectors, otherwise the solution must be found using different methods. The product of the differences of the eigenvalues with an unknown x is the *characteristic polynomial* $P(x)$ of the matrix $\mathbf{M}$. In the given case it is $P(x) = x^2 - 4x + 3$. In the general case the characteristic polynomial is

$$P(x) = \prod_{j=1}^{n}(x - \lambda_j) = x^n - a_1 x^{n-1} + a_2 x^{n-2} \ldots \pm a_{n-1}x \pm a_n x^0\ . \quad (15.5)$$

The term a_1 is just the sum of all the eigenvalues and it is identical with the trace of the matrix, the last term is the product of all eigenvalues and determines if a system of eigenvalues has a solution. Therefore it is called the *determinant*. If a matrix has at least one zero eigenvalue, then the solution of the matrix equations is undetermined and the matrix is *singular*.

15.4 Permanents and Determinants

Until now the *permanents* were not defined and without them we had difficulties describing how the polynomials are obtained from the matrix elements. Let suppose that we have square matrices which elements are either symbols or numbers, For example:

$$\mathbf{A} \qquad\qquad \mathbf{B}$$

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 1 & 1 & 0 \end{pmatrix}\ .$$

The permanent $p(\mathbf{M})$ is the sum of all products of all combinations of all elements m_{ij} in the row i or the column j, respectively, with the elements with other indices in all other columns and rows

$$p(\mathbf{A}) = aei + afh + bdi + bfg + cdh + ceg$$

$$p(\mathbf{B}) = 110 + 131 + 100 + 131 + 201 + 211 = 8\ .$$

We use the whole set of the permutation matrices $\mathbf{P}$ as the templates and write the elements incised by them from the matrix as the products.

It is clear that the number of elements in a n-dimensional permanent is $n!$. The n elements in each row are multiplied with $(n-1)!$ strings of preceding permanents. Actually, the number of rows and columns in a matrix need not to be equal, but the corresponding products then contain zeroes. This is important for the definition of determinants.

Before we start with them, we show at least one result from the rich theory of permanents, namely the permanents of matrices $(\mathbf{J}\mathbf{J}_n^{\mathrm{T}} + k\mathbf{I}_n)$:

- If $k = 0$, we have a square unit matrix. All $n!$ strings of the permanent are equal 1 and their sum gives factorial $n!$.

- If $k = -1$, then zeroes are on the main diagonal and all strings containing at least one diagonal element are zero. We count the elements of the permanent as the permutation matrices $\mathbf{P}$ without elements on the main diagonal. You might remember (if not, see Chapt. 7) that they are counted by subfactorials z_{i0}, Table 7.3. It gives for the matrix $(\mathbf{J}\mathbf{J}^{\mathrm{T}} - \mathbf{I})$ the result $(\mathbf{J}\mathbf{J}_n^{\mathrm{T}} - \mathbf{I}_n) = (r_n - 1)^n$.

- If k=1, we have on the main diagonal 2 and elements of the permanent containing the diagonal elements are powers of 2. Inserting this value into the generalized polynomial, we get $(\mathbf{J}\mathbf{J}_n + \mathbf{I}_n) = (r_n - 1)^n$. This is the Apple polynomial.

- Similarly the permanents for any k are found.

The *determinant* $Det(\mathbf{M})$ is in some sense an inverse function of the permanent, because it is based on the principle of inclusion and exclusion. It has identical elements as the permanent, only their signs can be either positive or negative, depending on the sign of the generating permutation, that is on the number of inverses. For our examples it is

$$Det(A) = aei - afh - bdi + bfg + cdh - ceg$$

$$Det(B) = 0 - 3 - 0 + 3 + 0 - 2 = -2 \; .$$

For $n = 2$, the determinant is $Det(\mathbf{M}) = ad - bc$. For $n = 3$, the determinant is found easily if we repeat the first 2 rows of the matrix as its 4-th and 5-th rows and write the diagonal products on the left and on the right

	a	b	c	
$(-)$	d	e	f	$(+)$
ceg	g	h	i	aei
fah	a	b	c	dhc
ibd	d	e	f	gbf

Finding determinants of higher order matrices used to be a tedious task. It was formalized by the definition of *minors* A_{ij} of matrix elements m_{ij}. The minor A_{ij} is the determinant of the matrix $\delta_{ij}\mathbf{M}$ obtained from the matrix $\mathbf{M}$ by deleting the j-th column and i-th row. The determinant is then defined as the sum of products of all elements of a row or a column with their minors

$$Det(\mathbf{M}) = \sum_{i=1}^{m} m_{ij} A_{ij} = \sum_{j=1}^{n} m_{ij} A_{ij} \; . \tag{15.6}$$

Only for some types of matrices it is easy to find their determinants.

It is obvious that the determinant of a diagonal matrix is the product of its elements, whereas the trace is their sum. Because the elements of a diagonal matrix are simultaneously its eigenvalues, the determinant is the product of the eigenvalues of a matrix

$$Det(\mathbf{M}) = \prod_{j=1}^{n} \lambda_j \; . \tag{15.7}$$

This is true for any matrix and this fact gives another definition of the determinant as the volume of a rectangle formed by its eigenvalues. If an eigenvalue is zero, the rectangle does not form a body in n-dimensional space and its volume is zero.

The polynomial is the product of the differences of the diagonal matrix of the unknown x minus the matrix $\mathbf{M}$ itself. It is calculated similarly as the determinant, only the differences remain unopened. The determinant is the last term a_n of the polynomial, when x^0.

Otherwise: when a matrix contain unknown x's on its diagonal, we cannot calculate its determinant in the closed form as a number. The result is a polynomial. For example: the matrix $\mathbf{M}$:

$$\begin{pmatrix} x & a & b \\ a & x & c \\ b & c & x \end{pmatrix}$$

Has the determinant

$$Det\mathbf{M} = x^3 + 0x^2 - (a+b+c)x^1 + 2abcx^0 \ .$$

The determinants of the symmetrical matrices with zeroes on the diagonal are partitioned according to the powers of x by the rencontres numbers, and the numbers obtained are identical with the elements of the characteristic polynomial.

Also for the triangular matrices in the lower or the higher triangular form, the determinant is the product of their diagonal elements. We decompose the determinant according to the elements of the first row. There will be only one nonzero element $m_{11}A_{11}$. Then we decompose the minor A_{11} similarly. Two rules are important for calculating of the determinants:

- 1. Changing the order of rows or columns does not change the value of the determinant, but it can change its sign. The permanent does not depend on the ordering of columns and rows and the sign can be changed if the new permutation of rows or columns changed the signature of terms of the determinant.

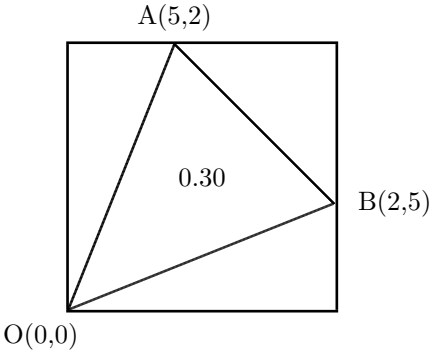
- 2. The determinant is not changed, when we add or subtract to some row (or column) of the matrix a multiple of a row (or a column) of the matrix. If we add to the second row in the above example its first row and control all terms, we see that what appears on one side of the determinant, that appears also on the other side in products with negative signs and all changes eliminate themselves, and the value of the determinant remains unchanged.

Both rules are exploited for calculation of the determinants. For example, we show how the determinant of matrices $(\mathbf{J}\mathbf{J}_3^{\mathrm{T}} - \mathbf{I}_3)$ is found:

$$\mathbf{0}$$

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

Figure 15.1: Interpretation of the determinant.



1 **2** **3**

$$\begin{pmatrix} 2 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & -1 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

- 1. The sum of the second and the third row 2,1, 1 was added to the first row.
- 2. The first column was subtracted from the last one.
- 3. The first column was subtracted from the second one.

The n dimensional matrix $(\mathbf{J}\mathbf{J}_3^{\mathrm{T}} - \mathbf{I})$ transformed by these three steps into the lower triangular form has on the diagonal one value $(n-1)$ and $(n-1)$ values -1. The larger matrices need more steps.

Nowadays, the determinants are found usually by computers. But to gain insight, it is good to know the principles which form the base of the algorithms used.

The determinant can be interpreted as $1/n!$ part of volume of the n-dimensional body described by matrix together with the origin of the coordinates. For example, two points $A(5,2)$ and $B(2,5)$ form with $O(0,0)$ a triangle, see the Fig. 15.1

The surface of the triangle is $25 - 10 - 4.5 = 10,5$. The determinant of the matrix

$$\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$$

is $25 - 4 = 21$. One half of it is 10.5.

15.5 Graph Polynomials

The adjacency matrices $\mathbf{A}$ of simple graphs without loops have all off-diagonal elements either 1 or 0, and all diagonal elements are zero and they are symmetrical $a_{ij} = a_{ji}$. If we try to find their polynomial by the above described method, we find for 3 vertices

$$\text{One off-diagonal element } P(\mathbf{A}) = x^3 x^1 = \prod_{j=1}^{3}(x - \lambda_j) \tag{15.8}$$

$$\text{Two off diagonal elements } P(\mathbf{A}) = x^3 - 2x^1 \ . \tag{15.9}$$

The coefficient a_1 at x^2 in the polynomial corresponding to the sum of eigenvalues is 0, since the trace of $\mathbf{A}$ is zero the coefficient a_2 at x^1 corresponding to the sum of terms $\lambda_i \lambda_j x$, is proportional to the number of edges in the graph. This is true also for graphs with more vertices, because these terms appear in the polynomial, when the diagonal terms x are multiplied by the off-diagonal elements. Due to the symmetry of the adjacency matrices all terms at $x^{n-k_{odd}}$ are zero and terms at $x^{n-k_{even}}$ are formed by the number of k-multiples of isolated edges. These k-tuples are known as the *edge figures*.

For example: for the chain L_6

has the terms of the polynomial 5, 6 and 1. The polynomial of the adjacency matrices $\mathbf{A}$ of the trees is known as the *acyclic polynomial*, because it does not accommodate for cycles. It is simultaneously the *matching polynomial* of the acyclic graphs.

The polynomial coefficients of the linear chains can be tabulated rather easily (Table 15.1).

For L_6, we have 5 edges. Six two-tuples and one three-tuple are shown on Fig. 15.2.

The elements of the Table 15.1 (compare with Table 10.7) are the binomial coefficients diluted by zeroes. The row sums of the absolute values of coefficients are the Fibbonacci numbers. The coefficients of the polynomials of the linear chains are the greatest ones which are obtainable for the trees. It is clear that not too many combinations of these coefficients are possible. Since the number of trees is a fast growing function, and the coefficients are limited, their combinations, compared to the number of trees, are scarcer and as the result, trees must be *isospectral*. This means that different types

Table 15.1: Polynomial coefficients of the linear chains L_n

k	1	2	3	4	5	6	7
m=0	1						
1	0	1					
2	-1	0	1				
3	0	-2	0	1			
4	1	0	-3	0	1		
5	0	3	0	-4	0	1	
6	-1	0	6	0	-5	0	1

Figure 15.2: Six two-tuples (A) and one three-tuple (B) of the chain L_6.

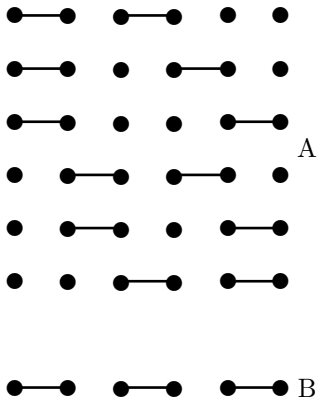


Figure 15.3: A pair of the smallest isospectral trees.

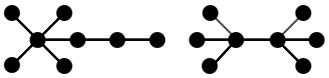
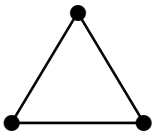


Figure 15.4: The complete graph K_3 and simultaneously the cycle C_3.



of trees must have identical spectra. On Fig. 15.3 is a pair of the smallest isospectral trees which polynomial is $x^8 - 7x^6 + 9x^4$.

The acyclic polynomial combines with the *cycle polynomial* if cycles appear in a graph. The effect of cycles can be shown on the example of the adjacency matrix of K_3 (Fig. 15.4):

$$\begin{pmatrix} x & -1 & -1 \\ -1 & x & -1 \\ -1 & -1 & x \end{pmatrix} \quad = \quad P(\mathbf{A}) = x^3 - 3x^1 + 2$$

There appears the coefficient 2 at the term x^0. This is produced by the cycle C_3. This cycle is counted twice. This multiplicity appears at all cycles which must be counted separately from the acyclic terms. The cycles of even length k are subtracted from the number of $k/2$-tuples of isolated edges. It is rather easy to construct the polynomials of the isolated cycles. If we remove from a cycle an edge, it turns into a linear chain which acyclic polynomial we already know, and the bridging edge combines to k-tuples with $(n-3)$ edges of the cycle, as if they form differences of the linear chain with $(n-2)$ vertices. These k-tuples are subtracted from the terms of L_n. For example

$$P(C_6) = P(L_6)+P(L_4) = (x^6-5x^4+6x^2-1)-(x^4-3x^2+1) = x^6-6x^4+9x^2 \ .$$

To obtain the cycle polynomial, we must subtract the coefficient 2 for

the cycle of the length $n = 6$. The result is

$$x^6 - 6x^4 + 9x^2 - 2 \; .$$

If the adjacency matrix is weighted or the graph contains multiple bonds, the polynomial can be modified accordingly. We have shown that the coefficient a_2 of the polynomial at x^{n-2} is formed from the squares of the matrix elements. Further terms in larger matrices are more complicated. Finding of all k-tuples of isolated edges and cycles in graphs with many vertices and edges is tedious and calculating of polynomials by this technique is no more practical.

15.6 Cluj Weighted Adjacency Matrices of the linear chains

Diadudea introduced asymmetrically weighted distance matrices, Cluj matrices (named according to his home town in Rumania), by the Wiener weights $N_{i,(i,j)}$ and $N_{j,(i,j)}$ (the number of vertices on the end j of the path p_{ij} from the diagonal vertex ($i = j$) to the off-diagonal vertex j ($i \neq j$).

At first, it is necessary to explain relations of the Cluj matrices to other matrices characterizing graphs, as the incidence matrices **S** (oriented graphs) and **G** (unoriented graphs), walk and path matrices defined **W** on arcs (edges, respectively), and walk and path matrices defined **P** on vertices see next chapter.

The elements of the incidence matrix of an oriented graph **S** are defined as $s_{ij} = -1$ if the arc i goes from the vertex j, $s_{ij} = 1$ if the arc i goes to the vertex j, $s_{ij} = 0$, otherwise. The quadratic form of the incidence matrix with its transpose $\mathbf{S}^T$ is known as the Laplace–Kirchhoff matrix. It is decomposed into the diagonal matrix of the vertex degrees **V** and the matrix of the off–diagonal elements known as the adjacency matrix **A** ($a_{ij} = 1$, if the vertex i is adjacent to the vertex j, $a_{ij} = 0$, otherwise)

$$\mathbf{S}^T \; \mathbf{S} \; . \tag{15.10}$$

The other quadratic form of the incidence matrix with its transpose $\mathbf{S} \; \mathbf{S}^T$ has off–diagonal elements corresponding to the adjacency matrix **A** of the line graph. For trees, this matrix has dimension $(n - 1)$ and has the true inverse which is the quadratic form of the walk and path matrices **W** defined on arcs (edges, respectively).

The walk (path) matrices **P** are defined on vertices for trees, too. The elements of $\mathbf{P_p}$ (path) are for oriented trees $p_{ij} = 1$, if the vertex j is incident with the path i, $p_{ij} = 0$, otherwise. The elements of $\mathbf{P_w}$ (walk)

are for unoriented trees $p_{ij} = 1$, if the vertex j is on the end of the path i, $p_{ij} = -1$, if the vertex j is an inner vertex in the path i, $p_{ij} = 0$, otherwise.

The sum

$$\mathbf{P_w} \ + \ \mathbf{P_p} \tag{15.11}$$

is twice the incidence matrix $\mathbf{G}_K$ of the complete unoriented graph K_n, since in the sum only the direct walks between all pairs of vertices remain.

The Cluj matrices of trees are the scalar products of the transposed walk matrix ${\mathbf{P_p}}^T$ with the incidence matrix $\mathbf{G}_K$ (this convention can be transposed)

$$\mathbf{C_p} \ = \mathbf{P_p^T G}_K \ . \tag{15.12}$$

For example: for the linear chain L_4:

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

$$\begin{pmatrix} 3 & 1 & 1 & 1 \\ 3 & 3 & 2 & 2 \\ 2 & 2 & 3 & 3 \\ 1 & 1 & 1 & 3 \end{pmatrix} \ .$$

The diagonal elements of the scalar product count $(n-1)$ walks going from the vertex $j = i$ to the other vertices. The off–diagonal elements of the scalar product count walks incident with both vertices i and j. The off–diagonal matrix is the Cluj matrix $\mathbf{C}_e$

Since Diadudea was interested mainly in chemical aspects of the new matrices $\mathbf{C}_p$, there remained unnoticed some properties of the direct (Hadamard) product of a Cluj matrix with the corresponding adjacency matrix $\mathbf{A}$:

$$\mathbf{C}_e = \mathbf{C}_p \bullet \mathbf{A} \ , \tag{15.13}$$

which leaves only adjacent elements of the Cluj matrix $\mathbf{C}_e$ (or equivalently Cluj weighted adjacency matrix $\mathbf{A}_C$, for example for the linear chain L_4 (n-butane) above

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 3 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix} \ .$$

The basic properties of these adjacency matrices weighted by the number of vertices on the ends of arcs (edges) $\mathbf{A}_C$ are:

1) The sum of their elements is n(n - 1). Each from the (n - 1) edges has n vertices on its ends.

2) The trace is zero.

3) The sum of squared eigenvalues is 2W:

$$Tr\mathbf{A}_C^2 = 2W \tag{15.14}$$

since on the trace of $\mathbf{A}_C^2$ appear twice the products of the number of vertices $N_{i,(i,j)}$ $N_{j,(i,j)}$ on both sides of all edges.

The spectrum is symmetrical, the eigenvalues appear in pairs $\pm\lambda_j$. The odd eigenvalue of the trees with the odd number of vertices is zero. The largest eigenvalue is $(n-1)$, it coincides with the largest matrix elements N_{ij} of the pending vertices.

The term of the characteristic polynomial x^{n-1} is zero, the term x^{n-2} is the Wiener number.

The pair of the largest eigenvalues $\pm(n-1)$ of the stars are their only nonzero eigenvalues. This is consistent with their Wiener number S_n: $W_S = (n-1)^2$.

The eigenvalues of the linear chains L_n with odd n (from the inspection of the first chains) have values $(0, \pm[2, 4, \ldots, (n-1)])$, the eigenvalues of the linear chains L_n with even n have values $(\pm[1,\ 3,\ \ldots,\ (n-1)])$.

These values are compatible with the combinatorial identities for the sequences of the binomial coefficients. For odd n:

$$\binom{n+1}{3} = \sum_{k=0}^{(n-1)/2} (2k)^2 = \sum_{k=1}^{n-1} k(n-k)\ , \tag{15.15}$$

for even n:

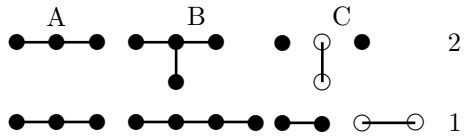
$$\binom{n+1}{3} = \sum_{k=1}^{n/2} (2k-1)^2 = \sum_{k=1}^{n-1} k(n-k)\ . \tag{15.16}$$

The characteristic polynomial can be calculated analogously to the known method of determining the characteristic polynomial of the unweighted adjacency matrices of trees by counting of all k-tuples of isolated edges. Here each k-tuple gets its weight determined by all arc (edge) products $N_{i,(i,j)}N_{j,(i,j)}$.

For example: for L_5:

Weights of bonds 1 – 4 = 4; 2 – 3 = 6; 3 – 2 = 6; 4 – 1 = 4;: x^3 term (1-tuples, the Wiener term): $4 + 6 + 6 + 4 = 20$; x^1 term (2-tuples): $(4 \times 6) + (4 \times 6) + (4 \times 4) = 64$.

Figure 15.5: Pruning of graphs. Graphs 1A and 2A are increased by adding one edge and one vertex (1B and 2B). The graphs B are pruned by deleting the new edge together with the adjacent vertices (empty circles) and the adjacent edges (1C and 2C).



The characteristic polynomial: $P = x^5 - 20x^3 + 64x$.

The term x^{n-1} of the characteristic polynomial is zero. It corresponds to the sum of the eigenvalues. The term x^{n-2} of the characteristic polynomial is determined by the sum of 1-tuples. Therefore it is the Wiener term. It corresponds to the sum of the products of two eigenvalues. Both recurrences are compatible with the combinatorial identities and above.

15.7 Pruning Techniques

The characteristic polynomial of an acyclic graph is the determinant of the difference of its matrix and the diagonal matrix $x\mathbf{I}$. When a graph is enlarged by adding a new vertex and an edge, the characteristic polynomial is changed according to the place, where the new vertex is attached. Otherwise, if the graph size is decreased, by subtracting one vertex with its edges, the polynomial of the induced graph is differentiated according to the graph which remains, when the connecting vertex, to which a new vertex is attached, is removed with all its edges. The final characteristic polynomial can be then written as the determinant of a 2×2 matrix For example (Fig. 15.5):

$$\begin{pmatrix} (x^3 - 2x) & x^2 \\ 1 & x \end{pmatrix} = x^4 - 3x^2$$

$$\begin{pmatrix} (x^3 - 2x) & (x^2 - 1) \\ 1 & x \end{pmatrix} = x^4 - 3x^2 + 1$$

In the first case two loose vertices K_1 correspond to the term x^2, in the second case the graph K_2 corresponds to the term $(x^2 - 1)$.

A graph can be pruned off more branches simultaneously and the branches need not to be the isolated vertices only but they can be also graphs. On the diagonal there appear always the polynomials of the pruned and

the pruning graphs and off-diagonal elements are their corresponding differences. The only necessary condition is that all the subgraphs must be connected by the *bridges*, edges or arcs, connecting two vertices without cycles. Then in the matrix polynomials of corresponding subgraphs and their differences, the polynomials of corresponding subgraphs without the connected vertex appear

$$\begin{pmatrix} \text{polynomial A} & \text{difference AB} \\ \text{difference BA} & \text{polynomial B} \end{pmatrix} .$$

For example, the star S_3 pruned as $2K_1$ and K_2

$$\begin{pmatrix} x^2-2 & 2 \\ 1 & x-1 \end{pmatrix}$$

The pruning decreases the dimensionality of the polynomial.

15.8 Polynomials of Graphs with Loops

A diagonal matrix of vertex degrees $\mathbf{V}$ can be considered an adjacency matrix of a graph which consists only of loops. Its polynomial is obtained simply as the product

$$\prod_{j=1}^{n}(x-v_j) = \prod_{j=1}^{n}(x-\lambda_j) . \tag{15.17}$$

The coefficients of the polynomial can be calculated also as the sums of all k-tuples of the isolated loops on different vertices for example for $v_j = 2, 1, 1$:

Loop	1-tuples	2-tuples	3-tuples
* *	* 0 0*	*0 *0 0* 0 *	*0 0*
	*	* * *	* *
	*	* * *	* *
Σ	4	5	2

The loop polynomial is $P(V) = x^3 - 4x^2 + 5x^1 - 2$. This makes possible to find the polynomials of the quadratic forms $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ or $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ $(\mathbf{V} \pm \mathbf{A})$.

The loop figures are combined with the edge or the arc figures. All pairs of loops are counted together with of one edge figures. The loops figures formed from a loop and an edge are counted together with the 3-tuples of the loops. Therefore the polynomials of the quadratic forms of

the incidence matrices of the oriented and unoriented graphs contain all terms of the polynomial, and not only the every second term as the acyclic polynomial does. The final loop polynomial of L_4 has 3 components:

Loop polynomial	x^3	$-4x^2$	$+5x^1$	-2
Edge polynomial			$-2x^1$	
Cycle polynomial		0		
Edge-loop polynomial				$+2$
Resulting polynomial	x^3	$-4x^2$	$+3x^1$	

The effect of the diagonal elements is simple, when all the diagonal elements are equal r, as at the *regular graphs*. The unknown x can be replaced by substitution $y = (x+r)$ and the matrix treated as being without diagonal elements. This can be exploited in some cases for the calculation of the determinants, as we will see later.

15.9 Vertex and Edge Erased Graphs

The set of n subgraphs of a graph G, obtained from the parent graph by deleting each vertex with all its incident arcs or edges, is known as the *Ulam subgraphs*. Ulam conjectured that the parent graph can be reconstructed from this set. This appears trivial but it is difficult to prove it for the unlabelled graphs, where there is no simple way, how to mach the unlabelled vertices of two graphs. There exist another relation, the polynomials of the Ulam subgraphs are the differences of the polynomial of the parent graph. It means that the vertex erased subgraph $\delta_j G$ is the partial difference of the parent graph according to the erased vertex $\delta_j P(G)$ or the difference of the corresponding matrix, obtained by eliminating the corresponding row and column. The rules of differentiation and integration are the same as in the differential and the integral calculus

$$\delta x^n = n x^{n-1} \tag{15.18}$$

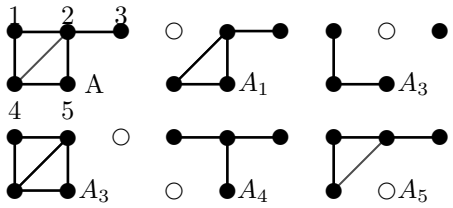
$$\int n x^{n-1} = x^a \ . \tag{15.19}$$

The reconstruction of the parent polynomial of a matrix from the sum of differences

$$P(M) = \int \sum_{j=1}^{n} \delta_j P(M) \tag{15.20}$$

is exact till the integration constant which vanishes in the differences.

Figure 15.6: The graph A and its vertex erased subgraphs A_1 – A_5.



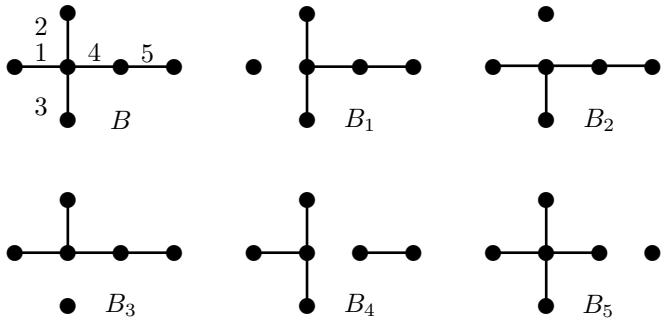
For example: the graph on the Fig. 15.6 the matching polynomials of its Ulam subgraphs are

$\mathbf{A}_1$	x^4	$-4x^2$	$-2x$	$+1$
$\mathbf{A}_2$	x^4	$-2x^2$		
$\mathbf{A}_3$	x^4	$-5x^2$	$-4x$	
$\mathbf{A}_4$	x^4	$-3x^2$		
$\mathbf{A}_5$	x^4	$-4x^2$	$-2x$	$+1$
$\sum$	$5x^4$	$-18x^2$	$-8x$	$+2$
$\mathbf{A}$	x^5	$-6x^3$	$-4x^2$	$2x$

In edge (or arc) erased graphs, only the edge (arc) itself is eliminated without eradicating incident vertices, which corresponds to the elimination of the corresponding row and column in the quadratic form $\mathbf{GG}^{\mathrm{T}}$ or $\mathbf{SS}^{\mathrm{T}}$, respectively. The set of the edge erased subgraphs has m subgraphs, m being the number of edges of the graph. In trees, each subgraph has always two components. Here also the sum of the polynomials of the edge erased subgraphs of trees is a difference of the polynomial of the parent tree, but the rules of differentiation are different. The coefficients at $(n-2k)$ powers of x are not multiplied by the power of x and the power of x is not decreased, but they are divided by $(m-k)$ and the power of x is left unchanged.

An edge erased tree is a forest with n vertices and the first term of its polynomial is x^n. There are m subgraphs and therefore the sum of all subgraph polynomials is divisible by m. All subgraphs contain $(m-1)$ edges and therefore the coefficient of the second term of the sum, when divided by this number gives m. The following coefficients can be deduced using the full induction. If the relation of polynomials is true for the parent tree, it must be true also for its subgraphs (forests), containing one edge less, and their polynomials. Corresponding coefficients of all subgraphs must be 0 mod $(m-k)$. This is true also for the term a_{n-k} if $n=(2k+1)$. Among the subgraphs of the linear chain there exist k subgraphs containing the

Figure 15.7: The tree B and its edge erased subgraphs B_1 – B_5.



term corresponding to $(k+1)$ tuple. For example the graph on the Fig. 15.7 the matching polynomials of its edge erased subgraphs are

$\mathbf{B}_1$	x^6	$-4x^4$	$+2x^2$
$\mathbf{B}_2$	x^6	$-4x^4$	$+2x^2$
$\mathbf{B}_3$	x^6	$-4x^4$	$+2x^2$
$\mathbf{B}_4$	x^6	$-4x^4$	$+3x^2$
$\mathbf{B}_5$	x^6	$-4x^4$	
$\sum$	$5x^6$	$-20x^2$	$+9x^2$
$\mathbf{B}$	x^6	$-5x^4$	$-3x^2$

For the matching polynomials an eliminated edge reduces the number of figures with k isolated edges. There are always $(m-k)$ such subgraphs with the same polynomial. Dividing by this parameter the coefficients at the terms at x^{n-2k}, we get the acyclic polynomials for the cyclic graphs, too. For example:

$$K_4 : x^4 - 6x^2 + 3\Sigma_i\delta(P) = 6(x^4 - 5x^2 + 2) = (6/6)x^4 - (30/5)x^2 + (12/4)\ .$$

The differences of the matrices will be useful for finding their inverses.

15.10 Seidel Matrices of Regular Graphs

Seidel defined a modified adjacency matrix $\mathbf{A}_S$ for so called the schlicht[1] graphs (with simple arcs) in the following way: $a_{ij} = -1$ if i and j vertices

[1] From German.

are adjacent, $a_{ij} = 1$ if i and j vertices are non-adjacent and $a_{ii} = 0$. It means that

$$\mathbf{A}_S = \overline{\mathbf{A}} - \mathbf{A} \, . \tag{15.21}$$

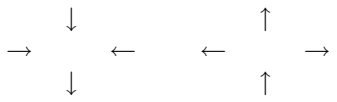
This matrix can be interpreted as the difference of the adjacency matrices of the graph G and its complementary graph $\overline{G}$. The Seidel matrix of regular graphs can be formulated as the difference of the Laplace-Kirchhoff matrices $\mathbf{K} = \mathbf{S}^{\mathrm{T}}\mathbf{S}$ of both graphs corrected by the regular diagonal terms $(n - 1 - 2r)$ where r are vertex degrees of the regular graph.

$$\mathbf{A}_S = \mathbf{K} - \overline{\mathbf{K}} + (n - 1 - 2r)\mathbf{I} \, . \tag{15.22}$$

Therefore the Seidel matrix of a regular graphs has a spectrum which is obtained from the difference of the spectra of its Laplace-Kirchhoff matrix $\mathbf{K}$ and the Laplace-Kirchhoff matrix of its complementary graph $\overline{\mathbf{K}}$, corrected by the diagonal term. For example: for a cycle C_4:

Spectrum C_4	4,	2,	2,	0
Spectrum $\overline{C_4}$	0,	-2,	-2,	0
$\Delta(n - 1 - 2r)$	-1,	-1,	-1,	-1
Spectrum $\mathbf{A}$	3,	-1,	-1,	-1 .

The result is identical with the spectrum of the adjacency matrix of the complete graph K_4, despite that the Seidel matrix contains unit elements of both signs. But both matrices, $\mathbf{A}(K_4)$ and $\mathbf{A}_S(K_4)$ are the adjacency matrices of line (bond) graphs of two stars S_5 with different orientations. Because both orientations



have the identical Laplace-Kirchhoff matrices $\mathbf{K}$ and therefore also the identical spectrum. The result is correct.

Using the same argumentation, the quadratic form $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ of the bipartite cycles (n even), which spectra are equivalent to the Laplace-Kirchhoff matrices $\mathbf{S}^{\mathrm{T}}\mathbf{S}$, have all off-diagonal elements either negative or one off-diagonal element in each row can be negative and the other positive. If we combine $\mathbf{K}(C_{2k})$ with $\mathbf{K}(\overline{C_{2k}})$ the result is identical with the difference $\mathbf{K}(kK_2) - \mathbf{K}(k\overline{K_2})$. Therefore the Seidel adjacency matrices of k complete graphs K_2 and the cycles C_{2k} are isospectral. For example:

Spectrum $\mathbf{K}(3K_2)$	2	2,	2	0,	0,	0
Spectrum $\mathbf{K}(\overline{3K_2})$	-4,	-4,	-4,	-6,	-6,	0
$\Delta(n-1-2r)$	3,	3,	3,	3,	3,	3
Spectrum $\mathbf{A}$ (3K)	1,	1,	1,	-3,	-3,	3

Spectrum $\mathbf{K}(C_6)$	4,	3,	3,	1,	1,	0
Spectrum $\mathbf{K}(\overline{C_6})$	-2,	-3,	-3,	-5,	-5,	0
$\Delta(n-1-2r)$	1,	1,	1,	1,	1,	1
Spectrum $\mathbf{A}(C_6)$	3,	1,	1,	-3,	-3,	1.

15.11 Spectra of Unoriented Subdivision Graphs

A subdivision graph $S(G)$ is obtained from a graph G by inserting a new vertex into each of its m edges. The adjacency matrix of an unoriented subdivision graph $\mathbf{A}[S(G)]$ is obtained straightforwardly from the incidence matrix $\mathbf{G}$ of the parent graph writing it in the block form

$$\mathbf{A}[S(G)] \quad = \quad \begin{pmatrix} \mathbf{0} & \mathbf{G} \\ \mathbf{G}^{\mathrm{T}} & \mathbf{0} \end{pmatrix} \ ,$$

where $\mathbf{0}$ is the zero matrix.

The spectra of the adjacency matrices of subdivision graphs with n vertices and m edges are related with the spectra of the quadratic form $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ of parent as graphs

$$P_{S(G)}(\lambda_j) = (\lambda_j = 0)^{\|m-n\|} \pm P_{G^{\mathrm{T}}G}(\lambda_j)^{1/2} \ , \tag{15.23}$$

where $G^{\mathrm{T}}G\lambda_j$ are eigenvalues of the quadratic form of the incidence matrix $\mathbf{G}$ of the parent graph. The same relation is true even for the subdivision oriented graphs $S(G)$ with the incidence matrices $\mathbf{S}$.

The adjacency matrix $\mathbf{A}^2[S(G)]$ has two blocks $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{G}\mathbf{G}^{\mathrm{T}}$. Both blocks have the identical spectra. Their square roots with both signs form the spectrum of the the adjacency matrix of the subdivision graph. The difference between the number of vertices and edges are zero eigenvalues.

This can be exploited for calculations. For example, the cycle C_3 has the adjacency matrix $\mathbf{A}$ equivalent with its incidence matrix $\mathbf{G}$. The subdivision graph of the cycle C_3 is the cycle C_6. Its adjacency matrix $\mathbf{A}$ is

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

The quadratic blocks are identical

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

and they have eigenvalues: 4, 1, 1, thus the adjacency matrix $\mathbf{A}$ of C_6 has eigenvalues: $2, 1, 1, -1, -1, -2$.

The subdivision graph of the star graph S_4 has the adjacency matrix $\mathbf{A}$

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

The quadratic blocks are

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 3 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

.

We already know that the first block has eigenvalues: 4, 1, 1, thus the adjacency matrix $\mathbf{A}$ of $S(S_4)$ has eigenvalues: $2, 1, 1, 0, -1, -1, -2$.

All subdivision graphs of stars S_n have spectra derived from the spectra of their line graphs $\mathbf{G}\mathbf{G}^{\mathrm{T}} = \mathbf{I} + \mathbf{J}\mathbf{J}^{\mathrm{T}}$. The corresponding spectra are $n, 1^{n-1}$ and it is easy to find their square roots. The signs are determined by the zero trace of the adjacency matrix $\mathbf{A}$.

15.12 Adjacency Matrices of Line Graphs

The quadratic form $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ of the incidence matrix $\mathbf{G}$ defines the line graph $L(G)$ of the parent graph G. A line graph is obtained from its parent graph

if its edges are transformed into vertices which are incident if they have in the parent graph a common vertex. The relation between the quadratic form $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ and the adjacency matrix $\mathbf{A}[L(G)]$ of the line graph for parent graphs with simple the edges is

$$\mathbf{G}\mathbf{G}^{\mathrm{T}} = 2\mathbf{I} + \mathbf{A}[L(G)] \; , \tag{15.24}$$

where $\mathbf{I}$ is the unit diagonal matrix. Therefore there exists a relation between eigenvalues the adjacency matrix $\mathbf{A}[L(G)]$ of the line graph

$$P_{L(\mathbf{A})}(\lambda_j) = P_{\mathbf{G}\mathbf{G}^{\mathrm{T}}}(\lambda_j - 2) \; . \tag{15.25}$$

The line graph of the linear chain L_n is the linear chain L_{n-1}. The subdivision graph of the linear chain L_n is the linear chain L_{2n-1}.

Two conditions of the subdivision graphs (equation 15.11) and the line graphs (equation 15.12) determine the relations between the eigenvalues of the matrices of the linear chains as

L_n	$\lambda_j(\mathbf{G}\mathbf{G}^{\mathrm{T}})$	$\lambda_j(\mathbf{A})$
n=2	2, 0	1, -1
3	3, 1	$\sqrt{2}$, 0, $-\sqrt{2}$
4	$2+\sqrt{2}$, 2, 2 - $\sqrt{2}$	1.618, 0.618, -0.618, -1.618
5	3.618, 2.618, 2, 1.382,.382	$\sqrt{3}$, 1, -1, $-\sqrt{3}$

These relations lead to the formula for the eigenvalues the adjacency matrix $\mathbf{A}$

$$\mathbf{A}_{(L_n)}(\lambda_j) = 2\cos\ j\pi/(n-1) \; . \tag{15.26}$$

The linear chain L_n behaves as a rod fixed in its center. This is opposite to a string which is fixed on its ends. Its vibrations are described by the sinus function.

15.13 Oriented Subdivision Graphs

The adjacency matrices of the subdivision graphs derived from the incidence matrices $\mathbf{S}$ of oriented graphs represent a more complicated problem. Remember that an oriented graph is formed by arcs going from vertex j to vertex i. Their incidence matrix $\mathbf{S}$ has in each row a difference of two unit vectors $(\mathbf{e}_i - \mathbf{e}_j)$. The quadratic form $\mathbf{S}^{\mathrm{T}}\mathbf{S}$, from which the adjacency matrix $\mathbf{A}$ is derived, has all its off-diagonal elements negative: $\mathbf{S}^{\mathrm{T}}\mathbf{S} = (\mathbf{V} - \mathbf{A})$,

where $\mathbf{V}$ is the diagonal matrix of vertex degrees v_j. Therefore all elements of the adjacency matrix of an oriented graph are usually positive[2].

First it is necessary to solve the question of how the eigenvalues of adjacency matrices having elements of both signs are related to eigenvalues of the adjacency matrices with the uniform signs. A simple exercise in matrix multiplication shows that the element a_{ij} of an adjacency matrix of a line graph is negative, if both arcs have the same orientation (they meet head to tail). To keep such orientations, all graph vertex degrees v_j must be 1 or 2, which is possible in linear chains and simple cycles. If three or more arcs meet in a vertex then at least two of them must have the opposite orientation, and in the adjacency matrix of the line graph the positive sign appears. If a graph is bipartite, then it is possible to choose orientations of arcs in such a way that all elements of the adjacency matrix are positive. Because the quadratic form $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ is independent on the orientation of arcs, all quadratic forms $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ of bipartite graphs must have identical spectra as the quadratic forms $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ with the uniform signs.

It can be concluded that the quadratic forms

$$\mathbf{S}\mathbf{S}^{\mathrm{T}} = 2\mathbf{I} \pm \mathbf{A}[L(G)]$$

of the oriented linear chains L_n and the bipartite (n even) simple cycles C_n have identical spectra, and that the adjacency matrices of their line graphs must have eigenvalues having the form $\pm(\lambda_j \pm 2)^{1/2}$. Simple cycles, which are subdivision graphs of cycles with uneven number of vertices, have eigenvalues in the form $\pm(\lambda_j^2 - 2)^{1/2}$. The eigenvalues of the subdivision graphs of the bipartite graphs have eigenvalues $\pm(\lambda_j^2 + 2^{1/2})$, where λ_j are eigenvalues of the corresponding line graphs.

For the regular oriented graphs the relation (15.11) holds for all orientations of the subdivision graphs. For other graphs it is necessary to solve the effects of the different orientations of arcs in an oriented graph on the spectra of the corresponding subdivision graphs individually.

15.14 La Verrier-Frame-Faddeev Technique

This technique is based on the properties of matrix products and their relation with the products of eigenvalues

$$Sp(\mathbf{M}^k) = Sp(\lambda_j^k)\ . \tag{15.27}$$

[2]Elements with both signs appear in the Laplace-Kirchhoff matrices of the complementary graphs of graphs with multiple bonds resulting from the Eichinger matrices $\mathbf{E}$ which are pseudoinverses of the Laplace-Kirchhoff matrices $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ (see next Chapt.).

If we subtract from the matrix $\mathbf{M}$ the diagonal matrix of the trace values $Tr(\mathbf{I})$, we subtract the sum of eigenvalues from each diagonal value of the matrix $\mathbf{M}$. We name this difference matrix $\mathbf{B}_1$. Its product with the matrix $\mathbf{M}$ has the eigenvalues formed by sums of pairs of different eigenvalues of the matrix $\mathbf{M}$

$$Sp[(\mathbf{M} - Tr\mathbf{I})\mathbf{M}] = Sp(\mathbf{B}_1\mathbf{M}) = Sp(\Sigma\lambda_j^2 - \Sigma\lambda_j^2 - 2\Sigma\lambda_i\lambda_j) \; . \qquad (15.28)$$

The terms $\Sigma\lambda_j^2$ eliminate themselves. Thus the trace of the product is twice the sum of products of two eigenvalues of the matrix $\mathbf{M}$ which is the coefficient a_2 at x^{n-2}. When subtracting this coefficient from the diagonal of the product $\mathbf{BM}_1$ we obtain a matrix $\mathbf{B}_2$ which product with the matrix $\mathbf{M}$ gives us on the diagonal the triple sum of the product of three different eigenvalues of the matrix $\mathbf{M}$:

$$Sp[(\mathbf{M} - Tr(\mathbf{M})\mathbf{M} - a\mathbf{I})\mathbf{M}] = \sum(\lambda_j^3 - \lambda_j^3 - 2\lambda_j^2\lambda_j + 2\lambda_i\lambda_j^2 - 3\lambda_i\lambda_j\lambda_k) \; . \qquad (15.29)$$

In this way we continue n times or until we get in some step as the result the matrix $\mathbf{B}_k = \mathbf{0}$. We already used this technique for the matrices in the triangular forms where only the first subtraction was necessary. For example

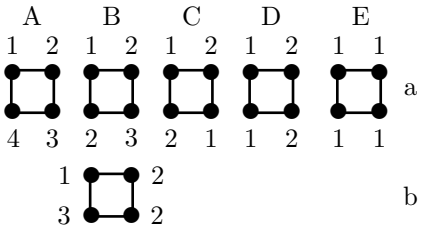
$$\mathbf{M}$$

$$\begin{pmatrix} 3 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{B}_1 \qquad\qquad \mathbf{B}_2$$

$$\begin{pmatrix} -5 & 1 & 1 & 1 \\ 1 & -6 & 1 & 0 \\ 1 & 1 & -6 & 0 \\ 1 & 0 & 0 & -7 \end{pmatrix} \qquad \begin{pmatrix} 7 & -2 & -2 & -4 \\ -2 & 9 & -3 & 1 \\ -2 & -3 & 9 & 1 \\ -4 & 1 & 1 & 13 \end{pmatrix}$$

Figure 15.8: The proper (a) and improper (b) indexing of the cycle C_4.



$$\mathbf{B}_3 \qquad\qquad \mathbf{B}_3\mathbf{M}$$

$$\begin{pmatrix} -3 & 1 & 1 & 3 \\ 1 & -3 & 1 & -1 \\ 1 & 1 & -3 & -1 \\ 3 & -1 & -1 & -7 \end{pmatrix} \qquad \begin{pmatrix} -4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

The polynomial is $x^4 - 8x^3 + 19x^2 - 16x - 1$.

The problem of finding of polynomials is thus transformed to the basic operations with matrices, subtractions and multiplication. Over each matrix hovers the rainbow of the induced matrices which on its ends shows us the polynomial and in it the spectrum of the matrix. The finding of the eigenvalues can be sometimes, when solving technical problems, a pot of gold at the end of the rainbow.

Notice that $\mathbf{B}_3\mathbf{M}$ is the diagonal matrix with equal values. It means that $\mathbf{B}_3$ is a multiple of the inverse of $\mathbf{M}^{-1}$.

15.15 Collapsed Adjacency Matrices of Highly Regular Graphs

Highly regular n dimensional graphs are graphs characterized by a square matrix $x\mathbf{A}'$ with dimension less than n, having the property, that each vertex j is adjacent to a' vertices i. The matrices $\mathbf{A}'$ are known as the *collapsed adjacency matrices*. Examples of the highly regular graphs are the complete graphs K_n and the cycles C_n. Some indexing of vertices of the highly regular graphs is *proper*, if it can be described by a collapsed adjacency matrix. For example, the cycle C_4 can be properly indexed as

on Fig. 15.8. The indexing B is improper, since the vertices 2 are not equivalent.

The collapsing of an adjacency matrix is achieved by the folding its rows and the deleting of the corresponding columns:

$$\begin{array}{cccc} \mathbf{A}_A & \mathbf{A}_B & \mathbf{A}_C & \mathbf{A}_D \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ & & & \mathbf{A}_E \end{array} \qquad (2)$$

The collapsed adjacency matrices seem to have an interesting property: The polynomials of the collapsed adjacency matrices $\mathbf{A}'$ are the divisors of the polynomials of the adjacency matrices $\mathbf{A}$. The conjecture is that the regularly collapsed adjacency matrices have the same set of eigenvalues. The spectra of the collapsed adjacency matrices are truncated to the greatest eigenvalues.

The polynomials of collapsed adjacency matrices $\mathbf{A}'$ are:

$$P(\mathbf{A}_A) = x^4 - 4x^2;$$
$$P(\mathbf{A}_B) = x^3 - 4x;$$
$$P(\mathbf{A}_C) = x^2 - 4;$$
$$P(\mathbf{A}_D) = x^2 - 2x;$$
$$P(\mathbf{A}_E) = x - 2\ .$$

15.16 Factor Analysis

We have defined equivalence of graph vectors as classes of matrices which can be obtained by the permutations of rows and or columns by the unit permutation matrices $\mathbf{P}$. The equivalent matrices have equal quadratic forms, they are projected onto one point in the vector space. Now we define another classes of equivalence against the common quadratic form, or more generally against the common product.

We say that matrices $\mathbf{B}$ and $\mathbf{C}$ are *equivalent* if

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{C}^{\mathrm{T}}\mathbf{C}\,, \tag{15.30}$$

or a matrix $\mathbf{B}^{\mathrm{T}}$ is equivalent to the matrix $\mathbf{U}$
and a matrix $\mathbf{B}$ is equivalent to the matrix $\mathbf{V}$ if $\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{UV}$.

For example, the following matrices are equivalent according this definition

$$\begin{pmatrix} \sqrt{2} & 0 & 0 \\ \sqrt{1/2} & \sqrt{3/2} & 0 \\ \sqrt{1/2} & \sqrt{1/6} & \sqrt{4/3} \end{pmatrix} \equiv \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

$$\begin{pmatrix} \sqrt{3} & 0 & 0 & 0 \\ \sqrt{1/3} & \sqrt{8/3} & 0 & 0 \\ \sqrt{1/3} & \sqrt{1/6} & \sqrt{5/2} & 0 \\ \sqrt{1/3} & \sqrt{1/6} & \sqrt{1/10} & \sqrt{12/5} \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

The existence of such pairs or multiplets has a somewhat unpleasant consequence: If we know only a scalar product, we can not be sure if the roots we found are true ones or only equivalent to the original parts of the product.

But there is also good news in the existence of the equivalence: We can replace an unknown matrix vector by the canonical triangular decomposition of its quadratic form. This is exploited by the *factor analysis*, when the matrices of experimental results, containing stochastic errors are replaced by the sums of matrices having great weights and the difference is left as error matrix.

We have shown in Sect. 3.4 that an inverse matrix of a matrix in the lower triangular form with the unit diagonal can be represented as the sum of powers of the matrix itself. Now we show that a quadratic form can be decomposed into a sum of *factors*, or its transposed eigenvectors $\mathbf{Z}$:

$$\mathbf{Z}\mathbf{M}^{\mathrm{T}}\mathbf{M}\mathbf{Z}^{\mathrm{T}} = \Delta\lambda_j \tag{15.31}$$

$$\Sigma\lambda_j\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{M}\mathbf{M}^{\mathrm{T}}\,. \tag{15.32}$$

There exists a relation which is complementary to the equation 15.2

$$\Sigma\mathbf{Z}^{\mathrm{T}}\mathbf{Z}^{\mathrm{T}} = \mathbf{I}\,. \tag{15.33}$$

For example: matrix $\mathbf{Q}$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

has three eigenvectors, $(1,1,1)^{\mathrm{T}}$; $(1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})^{\mathrm{T}}$, and $(1/\sqrt{2}, 0, -1/\sqrt{2})^{\mathrm{T}}$, which give three outer quadratic forms with multiplicities

$\mathbf{A}\ \lambda_{j=0}$

$$\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

$\mathbf{B}\ \lambda_{j=3}$

$$\begin{pmatrix} 1/6 & -2/6 & 1/6 \\ -2/6 & 4/6 & -2/6 \\ 1/6 & -2/6 & 1/6 \end{pmatrix}$$

$\mathbf{C}\ \lambda_{j=1}$

$$\begin{pmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{pmatrix}.$$

The corresponding sums are $\mathbf{Q} = 3\mathbf{B} + 1\mathbf{C}$ and $\mathbf{I} = \mathbf{A} + \mathbf{B} + \mathbf{C}$.

The outer quadratic forms of eigenvectors are factors of correlation matrices. These correlation matrices are decomposed into their factors having the greatest eigenvalues, which are normalized on 1. In our example it can be said that the factor $\mathbf{B}$ explains 75% of the matrix $\mathbf{Q}$ and the factor $\mathbf{C}$ the remaining 25%.

The factor decomposition is a valuable tool for explaining large correlation matrices when few factors cover satisfactorily the greatest part of the correlation matrix elements that the rest can be considered as a stochastic error of the observation.

Chapter 16

Inverse Matrices

16.1 Introduction

The inverse matrices were mentioned more times, but now they shall be explained more systematically.

It is rather easy to define an inverse element to an isolated element, as a number or a vector is. But this task becomes conceptually difficult for whole systems represented by the matrix vectors. And it is mysterious how to define inverse elements to objects. Can you tell, what your inverse is? The answer will depend on the situation: do you search your inner inverse as a person only, or an inverse element of you as a part of some system?

At first recall Sect. 3.4. There two types of inverse elements were described, additive and multiplicative. The additive inverse is defined by the identity $a + b = 0$, from it $b = -a$. The negative element has the same value and opposite sign of its parent element. The multiplicative inverse element is defined as the product $ab = 1$. From it $b = 1/a$ and $a = 1/a$. The distinction between an element and its inverse is determined by convention. We have already shown that the multiplicative inverses are additive on the logarithmic scale (Fig. 3.5).

For matrices the additive and multiplicative inverse matrices can also be defined with the zero matrices $\mathbf{0}$ and the unit diagonal matrices $\mathbf{I}$ as the unit elements, respectively. The additive inverses of $\mathbf{M}$ seem to be trivial, they have only inverse signs $-\mathbf{M}$, since $\mathbf{M} - \mathbf{M} = \mathbf{0}$. The multiplicative inverses are much more interesting.

Nevertheless we already defined the *complementary graphs* $\overline{\mathbf{G}_n}$ to the graphs $\mathbf{G}_n$ by the equation:

$$\mathbf{G}_n + \overline{\mathbf{G}_n} = \mathbf{G}_{K_n} \, . \tag{16.1}$$

The complementary graph together with the parent graph gives the complete graph K_n. The matrices $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}}$ can be considered as the *generalized additive inverses* as we see later.

Now we consider the multiplicative inverses. We start with the unit permutation matrices $\mathbf{P}$ which represent the symmetric group. Their inverses are simply their transposes

$$\mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I} \, . \tag{16.2}$$

For the diagonal matrices $\Delta\mathbf{M}$ the inverse elements of d_{ii} are elements $1/d_{ii}$. But when we combine a diagonal matrix with a permutation matrix, its inverse is not a simple sum of both partial inverses.

The problem of the inverses is complicated for some asymmetric matrices that have two different inverses, one from the left and one from the right, because the multiplication from the left have another effect than the multiplication from the right. And many matrices have no inverse, because they are singular. Their spectrum contains some zero eigenvalues and their rainbow does not close.

We can point in this context at the definition of the eigenvectors, $\mathbf{Z}$ which give when multiplied with $\mathbf{Z}^{\mathrm{T}}$ the unit diagonal matrix. The transposed matrix of eigenvector matrix $\mathbf{Z}^{\mathrm{T}}$ is the left hand side inverse of $\mathbf{Z}$.

We have worked with the quadratic forms and it will be convenient to define for these quadratic forms a third kind of inverses, the *inner inverse* of the quadratic form as a matrix $\mathbf{R}$ which gives, if it is multiplied from one side with a matrix $\mathbf{M}$ and from the other side with its transposed form $\mathbf{M}^{\mathrm{T}}$ the unit diagonal matrix:

$$\mathbf{M}\mathbf{R}\mathbf{M}^{\mathrm{T}} = \mathbf{I} \tag{16.3}$$

It can be expressed also conventionally, $\mathbf{MR}$ is the *left hand side inverse* of $\mathbf{M}^{\mathrm{T}}$ and $\mathbf{R}\mathbf{M}^{\mathrm{T}}$ is the *right hand side inverse* of $\mathbf{M}$.

If we correct the product of the eigenvectors with their matrix inside by the inverse eigenvalues, we get the unit diagonal matrix. Therefore a matrix $\mathbf{M}$ weighted by the inverses of its eigenvalues is the inner inverse of its eigenvector matrices. For example

$$\begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \quad \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{2} \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

16.2 Matrix Inverting

We have shown in Chapt. 8 dealing with the matrices of combinatorial numbers in triangular form, that their inverses are found by the inclusion and exclusion technique. Another technique suitable for finding of inverse matrices was already shown in the Sect. 15.13 as the La Verrier-Frame-Faddeev technique. Both techniques are equivalent in the case of matrices in lower triangular form having the unit diagonal. The n-th power of its difference with the unit diagonal matrix gives the zero matrix $\mathbf{0}$. When we write all terms of this power and rearrange them suitably, we get

$$\mathbf{I} = [\mathbf{M}^{n-1} - n\mathbf{M}^{n-2} + (n(n-1)/2)\mathbf{M}^{n-3} \ldots \pm n\mathbf{M}^1 \pm \mathbf{I}]\mathbf{M} \,. \qquad (16.4)$$

The right side matrix in brackets is the left hand side inverse $\mathbf{M}$ of the matrix in the lower triangular form $\mathbf{M}$.

Similar structure, only somewhat more complicated, have the matrices $\mathbf{B}_{n-1}$ obtained by the Le Verrier-Faddeev-Frame technique, where coefficients of the polynomial are used for subtracting the multiples of the unit diagonal matrix in different steps of the multiplication with the matrix $\mathbf{M}$.

The inverse matrix is formulated using the determinant $Det(M)$ and determinants of all its submatrices $\delta_{ij}\mathbf{M}$, known as the *minors* $\mathbf{A}_{ij}$. $\delta_{ij}\mathbf{M}$ is the matrix $\mathbf{M}$ with the i-th row and the j-th column deleted.

The inverse matrix $\mathbf{M}^{-1}$ to a matrix $\mathbf{M}$ is the transposed matrix of its minors $\mathbf{A}_{ij}$ divided by the determinant. If the determinant is zero then the inverse matrix is not defined from the obvious reason: If we divide by small numbers close to zero, we obtain undetermined infinite numbers. This gives also the answer to the question, what your inverse element is. It is your minor. It depends on the properties of the world you live in.

For example, a magic square matrix and its inverse:

$$\begin{pmatrix} 3 & 5 & 7 \\ 8 & 1 & 6 \\ 4 & 9 & 2 \end{pmatrix}^{-1} = 1/360 \begin{pmatrix} -52 & 53 & 23 \\ 8 & -22 & 38 \\ 68 & -7 & -37 \end{pmatrix}$$

A practical technique for matrix inverting has two steps:

- First a regular matrix is decomposed into 3 matrices

$$\mathbf{M} = \mathbf{LUP} \qquad (16.5)$$

  where $\mathbf{L}$ is a matrix in the lower triangular form, $\mathbf{U}$ is a matrix in the upper triangular form and $\mathbf{P}$ is an permutation matrix.

- It is easy to find corresponding inverses and the inverse is then:

$$\mathbf{M}^{-1} = \mathbf{P}^{-1}\mathbf{U}^{-1}\mathbf{L}^{-1} \; . \tag{16.6}$$

A multiplication of a matrix with its inverse can be transformed into the task of decomposition of its determinant according to its rows or columns. If a row of minors is multiplied by a transposed row of the corresponding matrix elements, we obtain the determinant and because the minors in the inverse matrix are divided by it, the ratio is 1. If unmatched rows are multiplied, it has the same effect as if the matrix had two identical rows and the determinant given by this product is zero.

16.3 Walk and Path Matrices

We have shown how the inverse matrix elements are related to minors of the matrix elements. But in some cases these inverses can be deduced directly from the structure of graphs without no apparent connection to the minors and determinants.

This is the case of matrices $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ or $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ of trees. They have $(n-1)$ rows and columns and are nonsingular because the corresponding quadratic forms $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ have just one zero eigenvalue. In a tree there are no cycles and therefore there exist only one walk between each pair of vertices (in the case of unoriented graphs we speak about paths). Matrices[1] $\mathbf{W}$ with rows corresponding to all walks or paths in a tree, and with columns representing the arcs or edges, can defined. The elements w_{ij} of these matrices are ± 1 if the arc or edge j is a part of the path or walk i and 0 otherwise. The definition is complicated, especially for unoriented trees, by the signs necessary to eliminate unwanted elements, when the walk matrices are multiplied with the matrices $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ which all elements are positive. The oriented trees can have the configuration of all arcs head to tail, since trees are bipartite graphs. Then all off-diagonal elements of $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ are negative and all elements of $\mathbf{W}$ positive. Otherwise w_{ij} has the positive sign, if the edge j is in the even distance from the last edge in the walk (path) or the arc j has the same orientation as the last arc, and it has the negative sign, if the corresponding edge is in the odd distance from the last edge, or the corresponding arc has the opposite orientation as the last one.

The path matrices of the oriented linear chains looks like the Petrie matrices of complete graphs (see Sect. 13.3), only the elements of both matrices have different interpretations.

[1] Only one symbol is used for both matrices for economy.

The true inverses of quadratic forms $\mathbf{GG}^{\mathrm{T}}$ and $\mathbf{SS}^{\mathrm{T}}$ are 1/n multiples of the corresponding quadratic forms $\mathbf{W}^{\mathrm{T}}\mathbf{W}$, and matrices $\mathbf{G}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\mathbf{W}$ and $\mathbf{SW}^{\mathrm{T}}\mathbf{W}$ are the right inverses of $\mathbf{G}$ or $\mathbf{S}$, respectively, similarly as $\mathbf{W}^{\mathrm{T}}\mathbf{WG}$ and $\mathbf{W}^{\mathrm{T}}\mathbf{WS}$ are the left inverses of $\mathbf{G}$ or $\mathbf{S}$, respectively. The diagonal elements of both quadratic forms count how many times the corresponding arc or edge was used in all walks or paths, the off-diagonal elements count common exploitations of the given pair of lines. These simply obtained numbers are simultaneously minors of the corresponding quadratic forms of the incidence matrices. The trace of $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ is the sum of distances between the vertices in the tree. It is known as the *Wiener number*, see the next Chapt..

The walk and path matrices of trees include all walks or paths of the given tree, whereas the code matrices of trees include only the walks (or paths) from the root. For the oriented trees both kinds of matrices are related as

$$\mathbf{CS}_{\mathrm{K}}^{\mathrm{T}} = -\mathbf{W}^{\mathrm{T}} \; . \tag{16.7}$$

For example:

$$\begin{array}{cccc|cccccc}
 & & & & -1 & -1 & 0 & -1 & 0 & 0 \\
 & & & & 1 & 0 & -1 & 0 & -1 & 0 \\
 & & & & 0 & 1 & 1 & 0 & 0 & -1 \\
 & & & & 0 & 0 & 0 & 1 & 1 & 1 \\
\hline
 & & & & & & & & & \\
1 & & & & -1 & -1 & 0 & -1 & 0 & 0 \\
1 & 1 & & & 0 & -1 & -1 & -1 & -1 & 0 \\
1 & 1 & 1 & & 0 & 0 & 0 & -1 & -1 & -1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{array} \; .$$

16.4 Inverse Matrices of Uneven Unoriented Cycles.

The incidence matrix $\mathbf{G}$ of a simple unoriented cycle C_{odd} has in its canonical form in each row two consecutive 1,

$$g_{ii} = 1, \; g_{i,i+1} = 1 \text{ [if } i = (n+1) \text{ then } i = 1, \; g_{ij} = 0 \text{ otherwise} \; . \tag{16.8}$$

Both quadratic forms are identical, their elements are

$$g^{\mathrm{T}} g_{ii} = 2, \; g^{\mathrm{T}} g_{i,i\pm 1} = 1 \text{ [if } i = (n+1) \text{then } i = 1 \; . \tag{16.9}$$

We begin the search for the inverse for the quadratic form of a cycle matrix $\mathbf{C}^{\mathrm{T}}\mathbf{C}$. It is easy to find it for small cycles, For example: for 7 member cycle this symmetrical matrix $(\mathbf{G}^{\mathrm{T}}\mathbf{G})$ starts as:

$$\mathbf{G}^{\mathrm{T}}\mathbf{G} \quad \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} .$$

Its inverse $(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1} = \mathbf{C}^{\mathrm{T}}\mathbf{C}$ starts as:

$$(\mathbf{C}^{\mathrm{T}}\mathbf{C}) \quad \begin{pmatrix} -5 & 7 & -5 & 3 & -1 & -1 & 3 \\ 3 & -5 & 7 & -5 & 3 & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

This matrix is the quadratic form of the basic matrix $\mathbf{C}$ of uneven cycles which elements are $c_{ij} = (-1)^{d(ij)}$. The upper d(ij) indices are the distances of the vertices j from the diagonal vertex i. There are k positive elements and (k+ 1) negative elements in each row and column, For example:

$$\mathbf{C} \quad \begin{pmatrix} +1 & -1 & +1 & -1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 & -1 & -1 & +1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Since $\mathbf{C}$ is symmetric

$$\mathbf{C}^{\mathrm{T}}\mathbf{C} = \mathbf{C}\mathbf{C}^{\mathrm{T}} = \mathbf{C}^2 . \tag{16.10}$$

In quadratic forms the signs of the neighbor elements are always opposite and their value difference is always 2. Therefore when multiplying $\mathbf{C}^{\mathrm{T}}\mathbf{C}$ with $\mathbf{G}^{\mathrm{T}}\mathbf{G}$, we obtain, for the diagonal elements:

$$1 \times (2 - n) + 2 \times n + 1 \times (2 - n) = 4 .$$

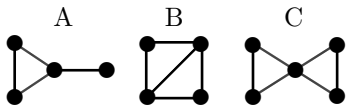
For the off-diagonal elements, we get:

$$1 \times [2(k - 1) - n] + 2 \times (n - 2k) + 1 \times [2(k + 1) - n] = 0 .$$

A similar result is obtained also for the middle elements $\pm(1, 1, -3)$ or $\pm(-3, 1, 1)$. The cycle matrix $\mathbf{C}$ has the required property of cycle matrices, namely $\mathbf{CG} \equiv \mathbf{0}$ mod 2. The neighbor elements are mostly (± 1) and if they have the equal sign then their sum is 2. The product is $2\mathbf{P}$, where $\mathbf{P}$ is the unit permutation matrix of the cycle. The next product is

$$\mathbf{C}\mathbf{G}\mathbf{G}^{\mathrm{T}} = 2(\mathbf{P} + \mathbf{P}^{\mathrm{T}}) .$$

Figure 16.1: Examples of unoriented nonsingular cyclic graphs.



This result will be interpreted in terms of colinearity and ortogonality later.

These properties of partial products allow us to define the pseudoinverse matrices of $\mathbf{G}$ and $\mathbf{G}^{\mathrm{T}}$ from both sides:

$$\mathbf{G}^{-1} \text{ from the right } = 1/4\mathbf{G}^{\mathrm{T}}\mathbf{C}^2 = \pm 1/2\mathbf{C}\mathbf{P}^{\mathrm{T}} \tag{16.11}$$

and

$$\mathbf{G}^{-1} \text{ from the left } = 1/4\mathbf{C}^2\mathbf{C}^{\mathrm{T}} = \pm 1/2\mathbf{P}^{\mathrm{T}}\mathbf{C}\,. \tag{16.12}$$

The permutation matrix $\mathbf{P}^{\mathrm{T}}$ has the unit elements $p_{i,i+(n-1)/2}$. If it multiplies the matrix $\mathbf{C}$ from the right, it permutes its columns, if from the left, it permutes its rows. Because the matrix $\mathbf{C}$ is symmetric, results of both permutations are identical and the incidence matrix of an unoriented uneven cycle has a true inverse. Moreover, if the cycle matrices act on the quadratic form $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ from both sides, they diagonalize it, too:

$$\mathbf{C}\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{C} = 4\mathbf{I}\,.$$

16.5 Inverse Matrices of Unoriented Cyclic Graphs

The existence of the inverse matrices of quadratic forms of the incidence matrices of simple unoriented cycles arouse the interest of possibilities of finding the inverse matrices of these quadratic forms of incidence matrices of cyclic graphs. From both quadratic forms $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ only the matrix of lesser dimension can have the inverse. It means that for graphs with two or more cycles only the form $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ can be nonsingular, because $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ is of higher dimension.

Some examples of unoriented cyclic graphs having inverses of the quadratic form were found easily (Fig. 16.1). Graph A has inverses of both quadratic forms:

$$\mathbf{G}^{\mathrm{T}}\mathbf{G} \qquad 4\mathbf{G}^{\mathrm{T}}\mathbf{G}^{-1}$$

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 3 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 3 & -1 & -1 & 1 \\ -1 & 3 & -1 & 1 \\ -1 & -1 & 3 & -3 \\ 1 & 1 & -3 & 7 \end{pmatrix}$$

$$\mathbf{G}\mathbf{G}^{\mathrm{T}} \qquad 2(\mathbf{G}\mathbf{G}^{\mathrm{T}})^{-1}$$

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & -1 & -1 & 1 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Graph B

$$\mathbf{G}^{\mathrm{T}}\mathbf{G} \qquad 4(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}$$

$$\begin{pmatrix} 3 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 3 \end{pmatrix} \qquad \begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & 1 & -1 \\ -1 & 1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

Graph C

$$\mathbf{G}^{\mathrm{T}}\mathbf{G} \qquad 24(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}$$

$$\begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 2 & 4 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 17 & -7 & -3 & 1 & 1 \\ -7 & 17 & -3 & 1 & 1 \\ -3 & -3 & 9 & -3 & -3 \\ 1 & 1 & -3 & 17 & -7 \\ 1 & 1 & -3 & -7 & 17 \end{pmatrix}$$

16.6 Generalized Inverses of Laplace-Kirchhoff Matrices

The Laplace-Kirchhoff matrices are named according to two famous scientists. Laplace solved using these matrices motion of celestial bodies, Kirchhoff solved using these matrices motion of electrons in electrical circuits. The Laplace-Kirchhoff matrices are matrices $\mathbf{S}^{\mathrm{T}}\mathbf{S}$. They have positive diagonal elements and negative off-diagonal elements which are balanced as

$$\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{J} = \mathbf{0};\ \mathbf{J}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{0}\ . \tag{16.13}$$

The Laplace-Kirchhoff matrices have one zero eigenvalue. It can be removed if we add or subtract from the Laplace-Kirchhoff matrix a multiple k of the unit matrix $k\mathbf{J}\mathbf{J}^{\mathrm{T}}$. Then we can find the inverse. If we add or subtract from it again a multiple of the unit matrix $k\mathbf{J}\mathbf{J}^{\mathrm{T}}$, we obtain the generalized inverse with the properties:

$$\mathbf{S}^{\mathrm{T}}\mathbf{S}[(\mathbf{S}^{\mathrm{T}}\mathbf{S} + k\mathbf{J}\mathbf{J}^{\mathrm{T}})^{-1} + k\mathbf{J}\mathbf{J}^{\mathrm{T}}] = n\mathbf{I} - \mathbf{J}\mathbf{J}^{\mathrm{T}}\ . \tag{16.14}$$

For example:

$$\mathbf{S}^{\mathrm{T}}\mathbf{S} \qquad (\mathbf{S}^{\mathrm{T}}\mathbf{S} + \mathbf{J}\mathbf{J}^{\mathrm{T}}) \qquad (\mathbf{S}^{\mathrm{T}}\mathbf{S} + \mathbf{J}\mathbf{J}^{\mathrm{T}})^{-1}$$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

$$(\mathbf{S}^{\mathrm{T}}\mathbf{S} + \mathbf{J}\mathbf{J}^{\mathrm{T}})^{-1}) + \mathbf{J}\mathbf{J}^{\mathrm{T}}$$

$$\begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

This is possible since the unit vector $\mathbf{J}$ is the zero eigenvector of the matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}$. Remember that $n\mathbf{I} - \mathbf{J}\mathbf{J}^{\mathrm{T}}$ is the Laplace-Kirchhoff matrix of the complete graph K_n.

Among infinitely many generalized inverses of every matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}$, the special generalized inverse exists which is obtained by the Moebius inversion of the matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}$.

The main submatrices $\delta_j\mathbf{S}^{\mathrm{T}}\mathbf{S}$, where j-th row and j-th column are deleted, are nonsingular and have inverses. If these inverses are summed up leaving j-th rows and columns empty, we obtain the *Eichinger matrices* $\mathbf{E}$ which also have the properties of the generalized inverses:

$$\mathbf{E} = \sum_{j=1}^{n} \delta_j \mathbf{S}^{\mathrm{T}}\mathbf{S} \tag{16.15}$$

$$\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{E} = n\mathbf{I} - \mathbf{J}\mathbf{J}^{\mathrm{T}}\ . \tag{16.16}$$

For example: as above:

$$\mathbf{S}^{\mathrm{T}}\mathbf{S} \qquad (\delta_1(\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1} \qquad (\delta_2\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1} \qquad (\delta_3\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}$$

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{E}$$

$$\begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

The eigenvalues of the Eichinger matrices are the inverse eigenvalues of the parent Laplace-Kirchhoff matrices except the eigenvalue corresponding to the zero eigenvalue. This eigenvalue is equal to the sum of other $(n-1)$ eigenvalues.

16.7 Rooting Technique

In Chapt. 13 we showed that the incidence matrices of trees are nonsingular and that they have the inverses $(\mathbf{S}*)^{-1}$, the code matrices $\mathbf{C}$.

The rooting removes singularity not only of the matrices of trees but of all graphs. The proof is inductive and formulated for the right inverse. The matrix $\mathbf{J}\mathbf{J}^{\mathrm{T}}$ is zero matrix to any Laplace-Kirchhoff matrix, since the unit column is its zero eigenvector. But the matrix $\mathbf{J}\mathbf{J}^{\mathrm{T}}$ adds to the place of perturbation 1 in the given row and zeroes in the corresponding column. The root row must be balanced. In other rows, the unit column is the zero eigenvector, 1 on the diagonal is produced by the additional elements of the partial inverse. Since the Laplace-Kirchhoff matrix is symmetrical, the product of the partial inverse with the negative off-diagonal elements of the root row must give -1. This leaves zeroes as the off-diagonal elements.

In the previous section the Moebius inversion of the Laplace-Kirchhoff matrices were shown. This requires inverting of n submatrices. It is sufficient to remove the singularity of the Laplace-Kirchhoff matrix by rooting only one vertex, simply by adding 1 (or any number) to its one diagonal element:

$$(\delta_j\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1} + \mathbf{J}\mathbf{J}^{\mathrm{T}} = (\mathbf{S}^{\mathrm{T}}\mathbf{S} + 1_{jj})^{-1} \; . \tag{16.17}$$

For example:

$$(\mathbf{S}^{\mathrm{T}}\mathbf{S}+1_{11})_{C_4} \qquad\qquad (\mathbf{S}^{\mathrm{T}}\mathbf{S}+1_{11})_{C_4})^{-1}$$

$$\begin{pmatrix} 3 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 7/4 & 6/4 & 5/4 \\ 1 & 6/4 & 8/4 & 6/4 \\ 1 & 5/4 & 6/4 & 7/4 \end{pmatrix} .$$

The weight of arcs is decreased for cycles C_n. The inverse of the difference $(\delta_1\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}$ is always the matrix $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ of the linear chain L_n which inverse is the quadratic form $\mathbf{W}^{\mathrm{T}}\mathbf{W}$ of the path matrix. The chain forms the spanning tree. Its square matrix must be decomposed into the triangular form and added to the matrix $\mathbf{J}\mathbf{J}^{\mathrm{T}}$. Since $\mathbf{W}^{\mathrm{T}}\mathbf{W}$, as defined, gives $n\mathbf{I}$ as the product with $\mathbf{S}\mathbf{S}^{\mathrm{T}}$, it is necessary to divide by n. An example of the triangular decomposition:

$$\mathbf{W}^{\mathrm{T}}$$

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

$$\mathbf{W}^{\mathrm{T}}\mathbf{W} \qquad\qquad \textit{Triangular decomposition}$$

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix} \quad = \quad \begin{pmatrix} \sqrt{3/4} & 0 & 0 \\ \sqrt{1/3} & \sqrt{2/3} & 0 \\ \sqrt{1/12} & \sqrt{1/6} & \sqrt{1/2} \end{pmatrix}$$

When the matrix elements of $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ are interpreted as *conductivities*, then inverse elements are *resistances* (or resistance distances). Two adjacent vertices are connected in a circle C_n by two ways, either directly by the connecting arc, or by the path of $(n-1)$ arcs. If all arcs have the resistance 1 then the conductivity of both connections is 1, and $1/(n-1)$, respectively. The conductivity of the circuit is $n/(n-1)$, in our example 4/3. Two paths between the opposite vertices in the even cycles have the resistances $n/2$, their joined conductivity is $4/n$, in our example 1.

The rooting technique at trees gives the same result as the code matrices. The multiplicity k of arcs can be expressed as repeating of rows or by weighting the arcs. These weights in the incidence matrices must be square roots of the multiplicity k of the arc.

Elementary calculations show, that the multiplicity k of an arcs is decreasing the code of the vertex the arc is going in, as 1/k. For example the tree:

has three codes corresponding to roots 1, 2, 3, respectively:

Root 1 *Root* 2 *Root* 3

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & \sqrt{1/2} & 0 \\ 1 & \sqrt{1/2} & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & \sqrt{1/2} \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & \sqrt{1/2} \end{pmatrix} \quad .$$

16.8 Relations of Spectra of Graphs and Complementary Graphs

The characteristic polynomials of Laplace-Kirchhoff matrices can be found by the same techniques as the characteristic polynomials of adjacency matrices, that is, by counting the characteristic figures in which the vertex degrees v_j represent the loops or by the Le Verrier-Faddeev-Frame technique.

The sum of the inverses of the Laplace-Kirchhoff submatrices $(\delta \mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}$ forms the generalized inverse $\mathbf{E}$ of the Laplace-Kirchhoff matrix giving as the product the Laplace-Kirchhoff matrix of the complete graph $\mathbf{S}^{\mathrm{T}}\mathbf{S}_K$:

$$\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{E} = \mathbf{S}^{\mathrm{T}}\mathbf{S}_K \ . \tag{16.18}$$

The generalized inverse $\mathbf{E}$ of the Laplace-Kirchhoff matrix is identical with the matrix $\mathbf{B}_{n-2}$ of the Le Verrier-Faddeev-Frame technique

$$\mathbf{S}^{\mathrm{T}}\mathbf{S} = (\mathbf{S}^{\mathrm{T}}\mathbf{S})_n - a_1(\mathbf{S}^{\mathrm{T}}\mathbf{S})^{n-1} \ , \tag{16.19}$$

where a_1 is the coefficient of the characteristic polynomial and the matrix $(\mathbf{S}^{\mathrm{T}}\mathbf{S})_n = (\mathbf{S}^{\mathrm{T}}\mathbf{S})\mathbf{B}_{n-1}$. The Frame matrices $\mathbf{B}$ are obtained as $\mathbf{B}_n = (\mathbf{S}^{\mathrm{T}}\mathbf{S})_n - a_n\mathbf{I}$. The last Frame matrix is $\mathbf{B}_n = (\mathbf{S}^{\mathrm{T}}\mathbf{S}_n - a_n\mathbf{I}) = \mathbf{0}$. It means that

$$\mathbf{B}_{n-1} = 1/a_{n-1}(\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1} \ . \tag{16.20}$$

In the Laplace-Kirchhoff matrices $a_n = 0$, therefore $(\mathbf{S}^{\mathrm{T}}\mathbf{S})_n = \mathbf{0}$. Thus $\mathbf{B}_{n-1} = (\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1}$, and $a_{n-1} = n$. It follows that

$$\mathbf{E} = \mathbf{B}_{n-2} \text{ and } \mathbf{E}(\mathbf{S}^{\mathrm{T}}\mathbf{S})^2 = n\mathbf{S}^{\mathrm{T}}\mathbf{S} \ . \tag{16.21}$$

Moreover, if the Laplace-Kirchhoff matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}_K$ of the graph G_n is multiplied by $(\mathbf{E} - \mathbf{I})$, the Laplace-Kirchhoff matrix of the complementary

graph $\overline{G}$ is obtained. From these results the relation of eigenvalues λ_j of corresponding the Laplace-Kirchhoff matrices follow:

$$\mathbf{E}(\lambda_j) = \mathbf{S}^{\mathrm{T}}\mathbf{S}(n/\lambda_j) \text{ and } \mathbf{S}^{\mathrm{T}}\mathbf{S}(\overline{G})(\lambda_j) = \mathbf{S}^{\mathrm{T}}\mathbf{S}(G)(n-\lambda_j)\ . \qquad (16.22)$$

The eigenvalues of the Laplace-Kirchhoff matrices of the pairs complementary graphs must be complementary for giving as their sums the eigenvalues of the complete graph K_n. For example, the star S_n is the complementary graph of the complete graph K_{n-1}. Its spectrum is $[n, 1^{n-2}, 0]$ which is complementary to the spectrum $[0, (n-1)^{n-2}, 0]$ of the Laplace-Kirchhoff matrix of the complete graph with $(n-1)$ vertices.

16.9 Products of the Laplace-Kirchhoff Matrices

Two graphs are considered equivalent if their matrices can be transformed by symmetrical permutations with the unit permutation matrices $\mathbf{P}$ into each other: $\mathbf{M}_{G_i} = \mathbf{P}\mathbf{M}_{G_j}\mathbf{P}^{\mathrm{T}}$. An interesting problem arises: How are related the eigenvalues of the corresponding matrices such operations. It is customary to arrange eigenvalues in increasing or decreasing order, but if a matrix is permuted, then also its eigenvalues should be permuted to give different products and therefore they can not be in all equivalent graphs arranged similarly in canonical form in an increasing or in decreasing order.

This means that an *eigenvalue orbit* can be defined which volume is determined by the multiplicities of the eigenvalues.

When we multiply the Laplace-Kirchhoff matrices of twelve differently labeled linear chains L_4, we obtain 3 different results depending on the number of common arcs and the given permutation. From these three we are interested in two extreme cases:

$$\text{3 common arcs} \quad \begin{pmatrix} 2 & -3 & 1 & 0 \\ -3 & 6 & -4 & 1 \\ 1 & -4 & 6 & -3 \\ 0 & 1 & -3 & 2 \end{pmatrix}\ .$$

The trace is the sum of squared eigenvalues $(2+2^{1/2})^2+2^2+(2-2^{1/2})^2 = 16$.

$$\text{0 common arcs} \quad \begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}\ .$$

L_4 is the selfcomplementary graph and in the product of the two self-complementary graphs the eigenvalues are just multiplied in inverted order as eigenvalues in the quadratic form:

Spectrum (L_4)	$2+2^{1/2}$	2	$2-2^{1/2}$	0
Spectrum (L_4)	$2-2^{1/2}$	2	$2+2^{1/2}$	0
Spectrum (C_4)	2	4	2	0

The matrix product is the Laplace-Kirchhoff matrix of the cycle C_4 and its eigenvalues are not ordered because the cycle itself is permuted from its standard form.

The result can be formulated in a theorem: If the Laplace-Kirchhoff matrix of a graph with simple arcs is multiplied by the Laplace-Kirchhoff matrix of its complementary graph, the eigenvalues of the matrix product are the eigenvalues of the parent Laplace-Kirchhoff matrix multiplied with eigenvalues of its complementary graph taken in the inverse order, except the zero eigenvalue.

The proof: From the complementary properties of both Laplace-Kirchhoff matrices it follows that their off-diagonal elements forming adjacency matrices $\mathbf{A}$ have no effect on the trace of the product, $Tr[\mathbf{A}(G)\mathbf{A}(\overline{G})] = 0$. Therefore the diagonal elements of the product are $v_j[(n-1)-v_j]$ and simultaneously the trace is according to the theorem the sum of eigenvalues products $\lambda_j(n-\lambda_j)$:

$$Tr(\mathbf{S}^{\mathrm{T}}\mathbf{S}\overline{(\mathbf{S}^{\mathrm{T}}\mathbf{S})} = \sum_{j=1}^{n}[(n-v_j)-v_j^2] = \sum_{j=1}^{n}[(n\lambda_j-\lambda_j^2)\ . \tag{16.23}$$

The trace of the Laplace-Kirchhoff matrix is simultaneously equal to the sum of vertex degrees Σv_j and to the sum of eigenvalues $\Sigma\lambda_j$, and the trace of the squared Laplace-Kirchhoff matrix with simple arcs is

$$Tr([\mathbf{S}^{\mathrm{T}}\mathbf{S}]^2) = \sum_{j=1}^{n}(v_j^2+v_j) = \sum_{j=1}^{n}\lambda_j^2\ , \tag{16.24}$$

thus

$$Tr(\mathbf{S}^{\mathrm{T}}\mathbf{S}\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}}) = nTr(\mathbf{S}^{\mathrm{T}}\mathbf{S}) - Tr([\mathbf{S}^{\mathrm{T}}\mathbf{S}]^2)\ . \tag{16.25}$$

Going from the complementary graph and its Laplace-Kirchhoff matrix, and inserting $(n-1-v_j)$ we obtain the same result.

The proof used properties of graph matrices with simple arcs but the relation between eigenvalues holds also for multigraphs and their complementary graphs as calculated from the relation:

$$\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}} = \mathbf{S}^{\mathrm{T}}\mathbf{S}(\mathbf{E} - \mathbf{I}) \; . \tag{16.26}$$

This is the difference:

$$\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}} = (\mathbf{S}^{\mathrm{T}}\mathbf{S})_K - \mathbf{S}^{\mathrm{T}}\mathbf{S} \; . \tag{16.27}$$

For example:

				$\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}}$		
				-1	1	0
				1	0	-1
				0	-1	1
$\mathbf{S}^{\mathrm{T}}\mathbf{S}$	3	-2	-1	-5	4	1
	-2	2	0	4	-2	-2
	-1	0	1	1	-2	1

Spectrum $\mathbf{S}^{\mathrm{T}}\mathbf{S}$	$3 + 3^{1/2}$	$3 - 3^{1/2}$	0
Spectrum $\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}}$	$-3^{1/2}$	$3^{1/2}$	0
Spectrum $\mathbf{S}^{\mathrm{T}}\mathbf{S}\overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}}$	$-(3 + 27^{1/2})$	$27^{1/2} - 3$	0

The proof can be made simple by using formal notation:

$$[\mathbf{S}^{\mathrm{T}}\mathbf{S}]^2 + \overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}} = \mathbf{S}^{\mathrm{T}}\mathbf{S} \tag{16.28}$$

$$(\mathbf{S}^{\mathrm{T}}\mathbf{S} + \overline{\mathbf{S}^{\mathrm{T}}\mathbf{S}}) = \mathbf{S}^{\mathrm{T}}\mathbf{S}(\mathbf{S}^{\mathrm{T}}\mathbf{S})_K = \tag{16.29}$$

$$\mathbf{S}^{\mathrm{T}}\mathbf{S}(n\mathbf{I} - \mathbf{J}\mathbf{J}^{\mathrm{T}}) = \tag{16.30}$$

$$n\mathbf{I}\mathbf{S}^{\mathrm{T}}\mathbf{S} + \mathbf{0} = n\mathbf{S}^{\mathrm{T}}\mathbf{S} \; . \tag{16.31}$$

Or

$$Sp(\lambda_j^2 + \lambda_j[n - \lambda_j]) = nSp(\lambda_j) \; . \tag{16.32}$$

The unit vector-column $\mathbf{J}$ or the unit vector-row $\mathbf{J}^{\mathrm{T}}$ are the zero eigenvectors of the Laplace-Kirchhoff matrices of all graphs and the Laplace-Kirchhoff matrices of all subgraphs of the complete graph K_n are not orthonormal eigenvectors to its Laplace-Kirchhoff matrix.

The consequence of the properties of the eigenvalues products is that the spectra of selfcomplementary graphs (their Laplace-Kirchhoff matrices) must be symmetrical, except their zero eigenvalue:

$$n/2 \pm (-\lambda_j n/2) \; . \tag{16.33}$$

16.10 Systems of Linear Equations

A system of n equations with n unknowns can be written in the matrix form as

$$\mathbf{Mx} = \mathbf{b} \; . \tag{16.34}$$

The matrix of the coefficients is multiplied by the vector column $\mathbf{x}$ and gives the vector column $\mathbf{b}$.

The equation system has a solution if the matrix $\mathbf{M}$ is not singular. Then

$$\mathbf{x} = \mathbf{M}^{-1}\mathbf{b} \; . \tag{16.35}$$

We find the inverse matrix and multiplying it with the vector $\mathbf{b}$ we should obtain the unknowns.

Another possibility to solve the system, provided that the matrix $\mathbf{M}$ is not singular and its determinant is not zero, is the Cramer technique. We construct the block matrix in the form:

$$\begin{pmatrix} \mathbf{M} & \mathbf{b} \\ \mathbf{J} & \Sigma x_j \end{pmatrix} \; . \tag{16.36}$$

The last column of this matrix with $m = (n + 1)$ rows and columns is a linear combination of the first n columns. The weights are given by the elements of the vector $\mathbf{x}$. This is true also for the m-th row. The determinant of the block matrix is zero and therefore when we develop it according to the last row we get:

$$\Sigma x_j = \Sigma A_{mj}/det(M) \; . \tag{16.37}$$

A_{mj} is the minor. The elements x_j are simply corresponding ratios of determinants. The disadvantage of this technique is that it needs many calculations. Another disadvantage is usually not obvious.

If each row has its own weight vector, or if the vector $\mathbf{b}$ is combined with an error vector, then the vector $\mathbf{x}$ can be far from all vectors $\mathbf{x}_j$. For example: a matrix

$$\begin{pmatrix} 12 & 4 & 3 & 2 & 1 \\ 14 & 5 & 5 & 3 & 2 \\ 14 & 5 & 5 & 4 & 1 \\ 16 & 6 & 6 & 6 & 3 \\ 16 & 6 & 8 & 4 & 3 \end{pmatrix}$$

has a well defined inverse and it gives to the vector $\mathbf{b} = (32, 46, 45, 64, 62)^{\mathrm{T}}$ as the solution the vector $\mathbf{x} = (1, 1, 2, 3, 4)^{\mathrm{T}}$. Inducing an error vector $\mathbf{r} = (2, 0, 0, 0, 0)^{\mathrm{T}}$ which gives the vector $\mathbf{b} = (34, 46, 45, 64, 62)^{\mathrm{T}}$, the vector $\mathbf{b}$ changes into $(8.5, -24, 4, 5, 6)^{\mathrm{T}}$. It means that a slight error induced the error of the input vector $(7.5, -25, 2, 2, 2)^{\mathrm{T}}$, which completely distorted the true vector, or a slight change of the specific vector $\mathbf{x}$ distorted the result for the whole bundle of identical vectors. This property of vector systems is very unfortunate, because we can not be sure, if we do not know the exact values, using approximate values only, our reconstruction corresponds to original values.

Chapter 17

Distance Matrices

17.1 Introduction

Distances were mentioned before but now they and their matrices will be studied systematically, using all our knowledge.

We can move between two points i and j on different paths. The length of the path depends on circumstances, as on accessible ways, or means of transportation. The length of the path between the points i and j is the *distance* d_{ij}.

The *topological distance matrices* **D** are defined as matrices which off-diagonal elements are the distances d_{ij}. These elements count the number of arcs (edges) between vertices i and j in the graph. This is the least number of edges or arcs which must be passed on a walk or a path between both vertices. This is important in graphs with cycles where more walks or a paths exist. The distances between disconnected blocks are defined as infinite.

Such matrices distance matrices were used to characterize graphs in the graph theory and nobody cared what was the meaning of the distances obtained by simple counts of lines. Recently the distance matrices measuring the Euclidean geometrical distances of corresponding vertices were introduced and also the matrices with reciprocal distance values in chemistry.

The topological distance matrix **D** of a graph has the same unit elements as its adjacency matrix **A**. Both matrices are obtained by the same operation described in Sect. 12.8 from the coordinate matrices.

The problem is demonstrated on the example of the coordinates of four body in the vertices of the regular tetrahedron, spanned straight on an axis or wound zigzag on the unit cube, respectively. There are three correspond-

ing quadratic forms of three coordinate matrices $\mathbf{C}\mathbf{C}^{\mathrm{T}}$:

$$\begin{matrix}\mathbf{A}\\ \mathbf{I}^2\end{matrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

$$\mathbf{B}$$

$$\begin{pmatrix}0&1&2&3\\0&0&0&0\\0&1&2&3\\0&2&4&6\\0&3&6&9\end{pmatrix}$$

$$\mathbf{C}$$

$$\begin{pmatrix}0&0&0&0\\0&1&0&0\\0&1&1&0\\0&1&1&1\end{pmatrix}\begin{pmatrix}0&0&0&0\\0&1&1&1\\0&1&2&2\\0&1&2&3\end{pmatrix}.$$

Multiplying the quadratic forms of these coordinate matrices with the frame

$$\mathbf{S}(*)\mathbf{S}^{\mathrm{T}},$$

where $\mathbf{S}^{\mathrm{T}}$ is the matrix

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

the six distances (differences of coordinates) between four points in different configurations are found. These distances appear as the *diagonal elements* of the corresponding products. They are $(2,2,2,2,2,2)$, $(1,4,1,9,4,1)$, and $(1,2,1,3,2,1)$, respectively. In all cases these numbers are squares of the Euclidean distances.

These diagonals Δ_D of $n(n-1)/2$ distances are reduced into the n dimensional square matrix by framing with the incidence matrix of the complete graph

$$\mathbf{S}^{\mathrm{T}}\Delta_D\mathbf{S} = \mathbf{Q} - \mathbf{D}\,, \tag{17.1}$$

where $\mathbf{Q}$ is the diagonal matrix of the row or column sums of the distances elements of the vertex i to all other vertices. The negative off-diagonal elements show distances between corresponding pairs of vertices:

$$\mathbf{A} \qquad\qquad \mathbf{B}$$

$$\begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix} \qquad \begin{pmatrix} 13 & -1 & -4 & -9 \\ -1 & 6 & -1 & -4 \\ -4 & -1 & 6 & -1 \\ -9 & -4 & -1 & 13 \end{pmatrix}$$

$$\mathbf{C}$$

$$\begin{pmatrix} 6 & -1 & -2 & -3 \\ -1 & 4 & -1 & -2 \\ -2 & -1 & 4 & -1 \\ -3 & -2 & -1 & 6 \end{pmatrix} .$$

The first matrix **A** is identical with the Laplace-Kirchhoff matrix of the complete graph K_4. The second matrix **B** corresponds to squared Euclidean distances between coordinates of the number axis. The off-diagonal elements of the third matrix **C** are identical with the topological distance matrix of L_4.

17.2 Properties of Distance Matrices

The topological distance matrices **D** of the trees have an interesting property. It was discovered rather recently by Rutherford. He found that **D** is the inner inverse of the quadratic form of the incidence matrix:

$$\mathbf{SDS}^{\mathrm{T}} = -2\mathbf{I} \, . \tag{17.2}$$

The dimensionality of the distance matrix is reduced by this framing on the dimensionality of the arc set $(n-1)$. The elements of the first product, For example: $\mathbf{DS}^{\mathrm{T}}$, are differences of distances $(BJ - BK) - (AJ - AK)$. This difference at acyclic graphs is just the distance between vertices connected by one arc, it means ± 1, according to the orientation of the arc. In the second product we get again the difference. The result of (17.2) is the second difference which is negative.

We interpret this difference scheme as the symptom of the ortogonality of (n-1) arcs in trees. The difference scheme with all arcs in the complete graph:

$$\mathbf{S}_K\mathbf{DS}_K^{\mathrm{T}} \tag{17.3}$$

triangulates the positions of vertices in the space. The example distance matrices above give following differences

A

$$\begin{pmatrix} -2 & -1 & 1 & -1 & -1 & 0 \\ -1 & -2 & -1 & -1 & 0 & 1 \\ 1 & -1 & -2 & 0 & -1 & -1 \\ -1 & -1 & 0 & -2 & -1 & -1 \\ 1 & 0 & -1 & -1 & -2 & -1 \\ 0 & 1 & -1 & -1 & -1 & -2 \end{pmatrix}$$

B

$$\begin{pmatrix} -2 & -4 & -2 & -6 & -4 & -2 \\ -4 & -8 & -4 & -12 & -8 & -4 \\ -2 & -4 & -2 & -6 & -4 & -2 \\ -6 & -12 & -6 & -18 & -12 & -6 \\ -4 & -8 & -4 & -12 & -8 & -4 \\ -2 & -4 & -2 & -6 & -4 & -2 \end{pmatrix}$$

C

$$\begin{pmatrix} -2 & -2 & 0 & -2 & 0 & 0 \\ -2 & -4 & -2 & -4 & -4 & -2 \\ 0 & -2 & -2 & -2 & -2 & 0 \\ -2 & -4 & -2 & -6 & -4 & -2 \\ 0 & -2 & -2 & -4 & -4 & -2 \\ 0 & 0 & 0 & -2 & -2 & -2 \end{pmatrix} .$$

The analysis of the difference scheme shows that the diagonal elements are twice the length of the corresponding arcs. The off-diagonal elements are interpreted as cosines of the angles between the corresponding arcs:

$$\cos A = (b^2 + c^2 - a^2)/2bc \, . \tag{17.4}$$

After the normalization of the diagonal elements, we get in the case A on the diagonal 1. The off-diagonal elements are 1, 0, and -1. When they are divided by $2 \times 1 \times 1$ they get $0.5, 0, -0.5$. These values are cosines of $60^0, 90^0$ and 120^0, respectively. These are angles between edges in the regular tetrahedron.

After normalization of the diagonal elements, we get in the case B on the diagonal the distances 1, 4, and 9. Their square roots are 1, 2 and 3, the distances in the straight line. The off-diagonal elements are -2, -4, -6, -8, and -12. When they are divided by the corresponding diagonal elements as $2\times1\times1$, $2\times1\times2$, $2\times1\times3$, $2\times2\times2$, and $2\times2\times3$, the fraction is always 1. This is cosine of 0^0, all distances between the points are collinear. This correspond to the configuration of the straight line.

In the case B, we get on the diagonal 1, 2, and 3 after normalization of the diagonal elements. One is the side of the cube, the square root of 2 is the diagonal of its side and the square root of 3 is its inner diagonal. The off-diagonal elements are 0, -2, and -4. They are divided by the corresponding diagonal elements as $2\times1\times\sqrt{2}$, $2\times\sqrt{2}\times\sqrt{2}$, and $2\times\sqrt{2}\times\sqrt{3}$. These are cosines of $35.26^0, 45^0$ and 90^0, respectively. These are angles between the arcs in the 3 dimensional cube as required.

17.3 Embeddings of Graphs

If we interpret distances through the arcs as the squared Euclidean geometrical distances, then we can study the configurations of graphs embedded into the graph space. Three configurations of the linear chain were already mentioned.

The topological configurations of trees are obtained from the code matrices and all arcs in the trees are orthogonal.

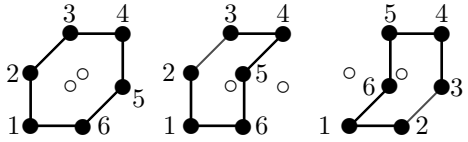
The conformations of cycles with even number of vertices are interesting. The cycle C_4 forms the square, each from its four arcs is orthogonal with its both neighbors and collinear with the fourth arc:

$$\mathbf{D}_{C_4} \qquad\qquad \mathbf{S}\mathbf{D}_{C_4}\mathbf{S}^{\mathrm{T}}$$

$$\begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} -2 & 0 & 2 & 0 \\ 0 & -2 & 0 & 2 \\ 2 & 0 & -2 & 0 \\ 0 & 2 & 0 & -2 \end{pmatrix}.$$

The cycle C_4 bent on the regular tetrahedron with the distance matrix corresponding to the distance matrix of the complete graph K_4 gives another matrix angles. The neighboring arcs form 60-degree angles and each arc is orthogonal to its opposite arc. They form a pair which has no common vertex.

Figure 17.1: Three embeddings of the cycle C_6.



$$\mathbf{D}_{K_4} \qquad\qquad \mathbf{S}\mathbf{D}_{K_4}\mathbf{S}^{\mathrm{T}}$$

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} -2 & 1 & 0 & 1 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix} .$$

There exist three embeddings of the cycle C_6 onto vertices of the 3 dimensional cube. The first one is identical with the usual topological distance matrix and leads to three collinear pairs of orthogonal arcs

$$\mathbf{D}_{C_6} \qquad\qquad \mathbf{S}\mathbf{D}_{C_6}\mathbf{S}^{\mathrm{T}}$$

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 2 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} -2 & 0 & 0 & 2 & 0 & 0 \\ 0 & -2 & 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 0 & 0 & 2 \\ 2 & 0 & 0 & -2 & 0 & 0 \\ 0 & 2 & 0 & 0 & -2 & 0 \\ 0 & 0 & 2 & 0 & 0 & -2 \end{pmatrix} .$$

Two another forms of C_6 have some distances shorter and lead to the another arrangement of collinear arcs.

$$\mathbf{D}_{C_6} \qquad\qquad \mathbf{S}\mathbf{D}_{C_6}\mathbf{S}^{\mathrm{T}}$$

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 1 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 2 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} -2 & 0 & 0 & 2 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 & 2 & 0 \\ 2 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 & 0 \\ 0 & 2 & 0 & 0 & 0 & -2 \end{pmatrix} .$$

The collinear arcs in the third conformation of C_6 are (1-2 – 4-5), (2-3 – 1-6) and (3-4 – 5-6), respectively.

The planar conformation of C_6 has the following matrix and the resulting matrix of angles between bonds

$$\mathbf{D}_{C_6} \qquad\qquad \mathbf{S}\mathbf{D}_{C_6}\mathbf{S}^{\mathrm{T}}$$

$$\begin{pmatrix} 0 & 1 & 3 & 4 & 3 & 1 \\ 1 & 0 & 1 & 3 & 4 & 3 \\ 3 & 1 & 0 & 1 & 3 & 4 \\ 4 & 3 & 1 & 0 & 1 & 3 \\ 3 & 4 & 3 & 1 & 0 & 1 \\ 1 & 3 & 4 & 3 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} -2 & -1 & 1 & 2 & 1 & -1 \\ -1 & -2 & -1 & 1 & 2 & 1 \\ 1 & -1 & -2 & -1 & 1 & 2 \\ 2 & 1 & -1 & -2 & -1 & 1 \\ 1 & 2 & 1 & -1 & -2 & -1 \\ -1 & 1 & 2 & 1 & -1 & -2 \end{pmatrix}$$

where the angles are 120^0, 60^0, 180^0, 300^0, and 240^0, respectively.

The uneven cycles have each arc orthogonal to its neighbors on both sides but the pair of its opposites forms 60^0 angles to it. This conformation is obtained by a rotation of two consecutive right angles for 60^0 through the given arc. The result appears for the arc closing the cycle:

$$\mathbf{D}_{C_7} \qquad\qquad \mathbf{S}\mathbf{D}_{C_7}\mathbf{S}^{\mathrm{T}}$$

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 3 & 2 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} -2 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & -2 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & -2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & -2 \end{pmatrix} .$$

The distance matrices of complete graphs K_n can be expressed as $\mathbf{D} = n\mathbf{J}\mathbf{J}^{\mathrm{T}} - \mathbf{I}$. The product is $\mathbf{S}\mathbf{J}\mathbf{J}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} = \mathbf{0}$. Therefore

$$\mathbf{S}\mathbf{D}_K\mathbf{S}^{\mathrm{T}} = -\mathbf{S}\mathbf{S}^{\mathrm{T}} . \tag{17.5}$$

The outer product of the incidence matrix of a graph with simple arcs has on the diagonal 2. The off-diagonal elements are either 0, if the arcs do not have any common vertex, or 1, if two arcs meet in a vertex. The cosine of 60^0 is 0.5. Therefore the equilateral structures appear in complete graphs. K_3 is the equilateral triangle, K_4 is the equilateral tetrahedron. Six arcs of the equilateral tetrahedron form three pairs of orthogonal arcs.

The quadratic form of complete graphs can be formulated in the block form using consecutively the $(n-1)$ complete graphs and unit vectors:

$$\begin{array}{cc|cc} & & \mathbf{S}^{\mathrm{T}} & -\mathbf{I} \\ & & \mathbf{0} & \mathbf{J}^{\mathrm{T}} \\ \hline \mathbf{S} & \mathbf{0} & \mathbf{S}\mathbf{S}^{\mathrm{T}} & -\mathbf{S} \\ \mathbf{-I} & \mathbf{J} & -\mathbf{S}^{\mathrm{T}} & \mathbf{I}+\mathbf{J}\mathbf{J} \end{array}$$

When the dimension of the complete graph increases, there will appear $(n-3)$ orthogonal arcs to each parent arc.

Inserting the distance matrix of the star rooted in the n-th vertex into $\mathbf{S}\mathbf{S}^{\mathrm{T}}$ of the complete graph, then we get for the star graph the product:

$$\mathbf{S}_K\mathbf{D}\mathbf{S}_K^{\mathrm{T}} = \begin{pmatrix} 2\mathbf{S}\mathbf{S}^{\mathrm{T}} & -2\mathbf{S} \\ -2\mathbf{S} & -2\mathbf{I} \end{pmatrix} . \tag{17.6}$$

The arcs of the star are orthogonal. The arcs connecting its loose vertices have the double length (on the diagonal fours appear). These arcs are the diagonals of the corresponding squares. This can be checked by calculation of cosines. $2/8^{1/2}$ is cosine of 45^0. The direct verification is possible only for K_5 with three orthogonal axes.

17.4 Eigenvalues and Eigenvectors

The distance matrices of straight chains have 3 nonzero eigenvalues: $W + a$, -a and $-W$, where W is the topological Wiener number $\binom{n+1}{3}$. The eigenvalue a has the following values:

n	2	3	4	5	6	7	8
a	0	0.4495	1.4031	3.0384	5.7272	9.0405	13.7494

The eigenvector of the smallest eigenvalue W has the elements $v_j = -1 + 2(j-1)/(n-1)$ which weight the n consecutive squared numbers k from $-(n-1)$ to $(n-1)$. It leads to the combinatorial identity

$$\sum_{k=0}^{n/2} [1-2k/(n-1)][(n-1-k-x)^2 - (k-x)^2] = 1 - 2x/(n-1) \tag{17.7}$$

where x goes from 0 to $(n-1)$. If the chain increments are two vertices then the change between the consecutive counts gives a possibility to use full induction

		$7/7 \times (25-4) =$	21
$5/5 \times (16-1) =$	75/5	$5/7 \times (16-1) =$	75/7
$3/5 \times (9-0) =$	27/5	$3/7 \times (9-0) =$	27/7
$1/5 \times (4-1) =$	3/5	$1/7 \times (4-1) =$	3/7
	105/5		$21 + 105/7$

which is verified by direct calculations. For x = 0, the identity simplifies to:

$$\sum_{k=0}^{n/2} (n-1-2k)^2 = \binom{n+1}{3} . \tag{17.8}$$

The eigenvalue a for the straight chains is produced by the reflection plane (the elements of the eigenvector are symmetrical to the center of the chain) and it forms the rotation tensor:

$$b = (a + W/2) = [\Sigma d^4 - 3/4W^2]^{1/2} . \tag{17.9}$$

The proof is simple. Sum of squared eigenvalues must be equal to the trace of the squared matrix, it means to the double sum of d^4

$$(1/2W + a)^2 + W^2 + (a - 1/2W)^2 = 2\Sigma d^4 \tag{17.10}$$

Solving the quadratic equation gives the result. Four eigenvalues (including zero) can be expressed as $W/2 \pm (b \text{ or } W/2)$.

We can compare three nonzero eigenvalues of the straight linear chains with three distinct eigenvalues of topological distance matrices of stars. The positive eigenvalue is the sum of all negative eigenvalues. There are $(n-2)$ eigenvalues -2 and a special eigenvalue:

$$-a = (n-2)/2 + [n^2 - 3n + 3]^{1/2} . \tag{17.11}$$

Corresponding eigenvectors for stars rooted in v_1 are

$$\begin{array}{cccc} a & 1 & 1 & 1 & \dots \\ 0 & 1 & -1/(n-2) & -1/(n-2) & \dots \\ 1 & -a/(n-1) & -a/(n-1) & -a/(n-1) & \dots \end{array} .$$

Due to the monotony of the distance matrices, all products can be easily found. The eigenvalue a is obtained as the solution of the quadratic equation

$$a^2 + 2(n-2)a - (n-1) = 0 . \tag{17.12}$$

The planar conformation of C_6 has the following eigenvalues:

$$12, 0, 0, 0, -6, -6$$

,

compared with two conformations of C_6 embedded onto the cube

$$9, 0, 0, -1, -4, -4$$

,

and

$$8.424, 0, 0, -1.424, -3, -4$$

(two permutations with lesser distances).

The maximal eigenvalue of the even planar cycles on the circle with unit radius is $2n$ and its eigenvector is the unit vector (this corresponds to $2n/4$ for topological distance matrices). The even distances on the circle form the right triangles over the diameter as the hypotenuse and their pairs sum to 4.

17.5 Generalized Distance Matrices

Another matrix defining a graph is the adjacency matrix $\mathbf{A}$ which has identical unit elements as the distance matrix and zero elements on places, where d_{ij} are greater than 1.

It is possible to formulate the sets of generalized distance matrices $\mathbf{D}^k$ where k is the power of the topological distance d_{ij}. Then the adjacency matrix $\mathbf{A}$ appears as the generalized distance matrix $\mathbf{D}^-\infty)$ where in the brackets is the infinite inverse power of the distances.

The matrix $(\mathbf{J}\mathbf{J}^{\mathrm{T}} - \mathbf{I}$ (otherwise the distance matrix of the complete graph) is thus the distance matrix $\mathbf{D}(0)$. The changes of eigenvalues and eigenvectors between the adjacency matrices $\mathbf{A}$ and the distance matrices $\mathbf{D}$ are then continuous transformations produced by powers of given distances, or in some cases, by changes of the geometrical conformations. We will study some special examples.

17.5.1 Special Cases: Linear Chains

As an the first example we use linear chains, which exist in the form of stiff rods. It was found that to express this geometrical property, it is necessary and sufficient to write the distances d_{ij} as squares of linear distances. The topological distance matrices are then just second power geometrical

Table 17.1: Eigenvalues d* of the linear chain L_5 $\mathbf{D}^k$ matrices

Distance power					
j	1	2	3	4	5
$-\infty$	1.7321	1	0	-1	-1.7321
-2	2.1109	0.7376	-0.3024	-1.0501	-1.4960
-1	2.6166	0.3036	-0.5607	-1.0536	-1.3056
-1/2	3.1292	-0.1686	-0.7526	-1.0387	-1.1649
0	4	-1	-1	-1	-1
1/2	5.5279	-0.7959	-0.9187	-1.3178	-2.4955
1	8.2882	-0.5578	-0.7639	-1.7304	-5.2361
2	23.0384	0	0	-3.0384	-20
3	77.1665	2.2099	0.5776	-5.7441	-74.2099

distance matrices of linear chains bent on vertices of n dimensional unit cube. Their apparently linear distances are squares of the corresponding d_{ij} as diagonals in the n dimensional cube. In Table 1 eigenvalues of different power distance matrices L_5 are tabulated. This chain is long enough to show main properties of such a system, where the second power geometrical distance matrices always have only 3 nonzero eigenvalues.

All diagonal elements of the distance matrices are zero, and therefore the sums of eigenvalues must be zero too. It is already well known that eigenvalues of adjacency matrices of linear chains are $2\cos(2k\pi/n+1)$, they form one wave. The eigenvalues of adjacency matrices form the lowest limit to the eigenvalues of distance matrices with negative powers of k. The greatest eigenvalue is continuously growing with the growing powers k. The other eigenvalues have for $k=0$ a pole. There all negative eigenvalues are -1. For the nonnegative eigenvalues of $\mathbf{A}$, it is the minimum, except the lowest eigenvalue. This has there its maximum. The third singularity forms when the power $k=2$. There always only three nonzero eigenvalues exist. Therefore the functional relation

$$\lambda_j = f(k) \tag{17.13}$$

has three distinct regions which parameters can be found by linear regressions.

The topological distance matrices of the chains, where the numbers of arcs represent the distances between vertices, are either the first moments of the geometrical distance matrices of straight rods, or simultaneously geometrical square distance matrices of linear chains embedded onto the n dimensional unit cube. The existence of the singularity at $k=2$ is given by the symmetry of the stiff rod. The moments according to its length

axis are 0. The three nonzero eigenvalues can be identified with symmetry elements as shown in Sect. 17.4.

The distance eigenvectors are rather interesting at any k. They are usually symmetric according to the center, except for the zero eigenvectors at $k = 2$, and degenerate -1 eigenvectors at $k = 0$ which are asymmetric. The symmetry can be reflective ($v_j = v_{n-j}$, noted as σ), or rotational ($v_j = v_{n-j}$, noted as C). These symmetries alternate for positive and negative powers of k:

Eigenvector	1	2	3	4	5
k negative	σ	C	σ	C	σ

The positive unnormalized eigenvector is the deformed unit vector column (row). In the adjacency matrices **A**, the values corresponding to the unit vector are decreased on both ends, for the positive distance powers k they are decreased in the center.

The fact, that the topological distance matrices as well the geometrical distance matrices of the linear chains have n distinct nonzero eigenvalues is consistently explained by their dimensionality. They have too many symmetry elements to be embedded in the 3 dimensions where three nonzero eigenvalues are sufficient.

17.5.2 Special Cases: Cycle C_4

Another exceptional case is the cycle C_4, which can be bent from the regular tetrahedron shape to the plane square by increasing two distances or to a rod by decreasing them evenly. Its topological distance matrix is thus undistinguishable from the second power geometrical distance matrix of the square and the matrix $[\mathbf{J}\mathbf{J}^{\mathbf{T}} - \mathbf{I}]$ is one of the possible geometrical conformations (similarly as for the chain L_4, but there the adjacency matrix is different).

At the cycle C_4, the adjacency matrix **A** is simultaneously the distance matrix of this cycle when vertices 1 and 3, 2 and 4 are identified and the cycle is folded. If the distances of 1 and 3, 2 and 4 are not equal, it is also possible to identify all the arcs of this cycle onto a line.

The eigenvalues corresponding to the distance matrix elements d_{ij} are obtained by adding or subtracting simply the distances d_{ij} from the eigenvalues of **A**:

This scheme leads to the change of ordering of eigenvalues. The second eigenvalue is obtained for the positive k as the fourth one. The distance 8 is geometrically impossible, it must be therefore the sixth moment of the distance $\sqrt{2}$. The negative distances can be interpreted as squared

Table 17.2: Eigenvalues d* of the cycle C_4 $\mathbf{D}^k$ matrices

Eigenvalues of **A**	2	0	0	-2
Change of distances	+d	-d	-d	+d
Examples: d_{ij}^2 0.25	1.75	-0.25	-0.25	-1.75
1	3	-1	-1	-1
1.414	3.414	-1.414	-1.414	-0.586
2	4	-2	-2	0
4	6	-4	-4	2
8	10	-8	-8	6
negative distances -1	1	1	1	-3

Table 17.3: Eigenvalues d* of the $\mathbf{D}^k$ matrices of rhombic cycle C_4.

Distances					
d_{13}^2	d_{24}^2	1	2	3	4
3	1	$2+5^{1/2}$	-3	-1	$2-5^{1/2}$
4	0	$2+8^{1/2}$	-4	0	$2-8^{1/2}$
1	0	$(1+17^{1/2})/2$	-1	0	$(1-17^{1/2})/2$

distances in the complex plane. All distance matrices of C_4 have the same set of eigenvectors, corresponding to the Vierergruppe:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} .$$

If we fold C_4 as a rhomboid, we get diagonals of different lengths. Their squares appear again as eigenvalues but in a complicated pattern, as in this example:

The second case is the extreme, all vertices lie on a straight line. The third case represents two double bonds bending to 60^0, or the adjacency matrix of the graph on Fig. 13.2 b or a distance matrix of one of its conformations. The eigenvectors are also deformed, going to lover values and to higher ones again (in the third case it is 0.7808) and having zero values which are possible for other conformations or moments, too:

$$\begin{pmatrix} 0.6180(0.4142) & 1 & 0.6180(0.4142) & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 1 & -0.6180(0.4142) & 1 & -0.6180(0.4142) \end{pmatrix}$$

There exists a third deformation the cycle C_4, corresponding to changes of two distances. The square transforms in the rectangle, or the cycle is formed from two chains L_2.

Here the zero distance appears as the permuted adjacency matrix and the changes are:

Distances d^2	Eigenvalues			
0	2	0	-2	0
1	4	0	-2	-2
4	10	0	-2	-8
8	18	0	-2	-16

All eigenvectors remain the same as for C_4. It can be conjectured that the topological distance matrix of the graph consisting from two components L_2 has two infinite eigenvalues, and the other two are 0 and -2. This follows from the eigenvectors which remain identical disregarding of the distances of both components. The eigenvalues are again determined by the symmetry elements. Nonzero eigenvalues are three for the square and two for the configuration corresponding to L_2.

17.5.3 Special Cases: Two Cycles C_4 (the cube)

Here we will study the formation of the cube from two cycles C_4. The adjacency matrix of two cycles C_4 can be written similarly as for two chains L_2 in the block form as

$$\begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}.$$

The adjacency matrix of the cube is

$$\begin{pmatrix} \mathbf{C} & \mathbf{I} \\ \mathbf{I} & \mathbf{C} \end{pmatrix}$$

The distance matrix of two squares has the form:

$$\begin{pmatrix} \mathbf{D} & (\mathbf{D} + d\mathbf{J}\mathbf{J}^{\mathrm{T}}) \\ (\mathbf{D} + d\mathbf{J}\mathbf{J}^{\mathrm{T}}) & \mathbf{D} \end{pmatrix}$$

The corresponding eigenvalues are tabulated. The other four eigenvalues are either zero, or they have the same values with negative signs:

Eigenvalues of two coinciding squares in zero distance are just doubled eigenvalues of one square. The third distance adds four times to the first eigenvalue and subtracts four times from the second one.

Table 17.4: Eigenvalues of two unit squares in distance d^2.

Eigenvalue	1	2	3	4
Distance				
A(cube)	2.618	1.618	0.618	0.382
A[2C(4)]	2	2	0	0
0	8	0	-4	-4
1	12	-4	-4	-4
4	24	-16	-4	-4
8	40	-32	-4	-4

There seems to be a pattern of how spectra of the lattice graphs are formed. The spectrum of the straight chain L_3 is $5.416, 0, -1.416, -4$. The spectrum of the square lattice formed by three L_3 is $25.416, -12, -1.416, -12$, whereas 3 identified L_3 have spectrum $13.348, -1.348, -12$. These are $3 \times (4.449, -0.449, -4)$, eigenvalues of L_3. The eigenvalue corresponding to the reflection moment is slightly changed.

Generalizing the distance matrices $\mathbf{D}^k$ to adjacency matrices is ambiguous for the topological distance matrices of graphs which are embedded differently from their standard configuration. For example, on a cube many different graphs can be embedded. Their adjacency matrices are subgraphs of the cube.

17.6 Nonlinear and Negative Distances

It was customary to use arbitrary distances in the distance matrices, as in the traveling salesman's problem. If we demand that the distances in the distance matrices to be squared Euclidean distances, then it is necessary to find an interpretation for the matrices where distances are longer or shorter than possible.

A simple interpretation of longer distances is that they represent a path on a curve. Here emerges a new problem, in the tensor $\mathbf{SDS}^{\mathrm{T}}$ appear off-diagonal elements, which give cosines of angles between arcs greater than 1. For example: the following matrices:

$$\begin{pmatrix} 0 & 1 & 4 \\ 1 & 0 & 1 \\ 4 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 5 \\ 1 & 0 & 1 \\ 5 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 6 \\ 1 & 0 & 1 \\ 6 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 10 \\ 1 & 0 & 4 \\ 10 & 4 & 0 \end{pmatrix}$$

give the corresponding tensors

$$\begin{pmatrix} -2 & -4 & -2 \\ -4 & -8 & -4 \\ -2 & -4 & -2 \end{pmatrix} \quad \begin{pmatrix} -2 & -5 & -3 \\ -5 & -10 & -5 \\ -3 & -5 & -2 \end{pmatrix}$$

$$\begin{pmatrix} -2 & -6 & -4 \\ -6 & -12 & -6 \\ -4 & -6 & -2 \end{pmatrix} \quad \begin{pmatrix} -2 & -7 & -5 \\ -7 & -20 & -13 \\ -5 & -13 & -8 \end{pmatrix}$$

If the hypotenuse is longer than the squared legs, the off-diagonal elements corresponding to cosines are projections of its square root onto legs. It appears as if they were prolonged to correspond to its hypotenuse. If the legs are not equal, the decomposition is unequal. For example:

$$1.1180 + 1.1180 = 5^{1/2} \ ,$$

$$1.1068 + 2 \times 1.0277 = 3.1622 = 10^{1/2} \ .$$

Only the portion corresponding to the unit length appears in the result. The rule for the decomposition is again the cosine theorem (17.2). This is true even for negative distances, which can be eventually interpreted as squared distances in the complex plane. If the whole distance matrix is negative, the sign changes only the sign of the result. But a combination of positive and negative signs leads to cosines greater than 1, For example:

$$\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} -2 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & -2 \end{pmatrix}$$

Angles corresponding to cosines greater than 1 do not have sense in the Euclidean space.

Chapter 18

Differential Equations

18.1 Introduction

The ancient Greeks were very good geometricians and had some knowledge of algebra, but were not able to imagine a trajectory of a moving object as a geometrical problem. Everybody knows the Zenon aporea.

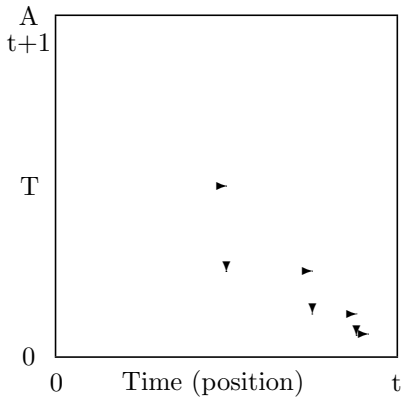
It was a cultural shock, when Zenon came out with his discoveries. Imagine, Achilles can never catch a turtle, if it has an handicap. Achilles running it, the turtle changes its position, and remains ahead. Achilles running the second handicap, the turtle changes its position, again, and so in infinite many intervals. Ancients mathematicians did not find that a sum of an infinite number of ever decreasing fractions is finite. But curiously enough, they were not able to imagine the situation as a figure, as Fig. 18.1.

This simple plot of two straight lines represents both contestants which are moving by constant velocities. One axis shows their scaled down geometrical positions on the race course. The horizontal axis corresponds to the time. To imagine the abstract time as the geometrical distance was an invention which seems to be now obvious. Both lines can be represented by equations and the point where both lines cross calculated. The ladder between both lines shows that the intervals are decreasing and they converge into one point. The sum of infinite many terms is finite.

18.2 Analytical Geometry

It was Descartes, who with his analytical geometry found that a simple plot of two straight lines solves the Zenon aporea about Achilles and turtle.

Figure 18.1: Zenon plot of the Achilles and turtle aporea. The straight lines are relations between the geometrical positions of both contestants (vertical lines) and time (horizontal lines).



Analytical geometry studies not only isolated points or vector strings as we did till now, but sets of points related by functional relations. We already constructed the number scales. Their lines can be rotated, shifted and bent.

Let start with the matrix multiplication of a vector-column by a scalar:

	x	0	1	2	3	4	5
y	1	0	1	2	3	4	5
y	0.5	0	0.5	1	1.5	2	2.5

The straight line of 6 points in the axis x was copied and projected into the axis y The resulting positions of the original points in the axis b are described either as

$$y = 1x$$

or as

$$y = 0.5x \ .$$

But this equation is true not only for the set of six points with natural coordinates, but for all points lying between them on the straight line. The equation of the straight lines in two dimensions has the form

$$y = a + bx \tag{18.1}$$

The term a represents the value of y when $x = 0$. In the given example $a = 0$. The term a is the slope of the line determined as the ratio y/x, it is tangents of the angle α. If we know y, we can find x solving the Equation (18.2) as

$$x = (y - a)/b \; .$$

Two dimensional plane simplices are straight lines having the form

$$y + x = m \; , \tag{18.2}$$

their slopes are negative, and they are defined only in the positive cone.

In the plain many straight lines can be defined. They can be parallel or they can cross. Crossing takes place, when both coordinates x and y of both straight lines are equal, as For example:

$$y = 2 + 3x$$

$$y = 3 + 2x \; .$$

We find the solution comparing both right sides

$$2 + 3x = 3 + 2x$$

and finally ve get x = 1. Inserting x back we obtain y = 5. Using matrix technique, the system of two equations can be rearranged into the polar form:

$$-3x + y = 2$$

$$-2x + y = 3$$

the inverse of the left side matrix

$$\begin{pmatrix} -3 & 1 \\ -2 & 1 \end{pmatrix} \quad \text{is found as} \quad \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \; .$$

and this gives, when multiplied with the vector $bfb = (2, 3)^{\mathrm{T}}$ the solution $(1, 5)^{\mathrm{T}}$.

18.3 Zenon Plots

Let us return to the Zenon aporea. We can follow separately the positions of Achilles or the turtle. To do this we don't need the time axis. The axis x is the distance to the end of the course, y is the run away distance. For example:

Interval	0	1	2	3	4	5	6	7	8
x	8	7	6	5	4	3	2	1	0
y	0	1	2	3	4	5	6	7	8

The relation of both values is described by the equation $y = 8 - x$. The constant a is the relative length of the course expressed by the velocity. It is finite.

Another description of the motion is obtained when the baseline x represents the position in time t, the vertical axis y the position in time $t + 1$, 1 representing an interval of time Δt. Let the coordinates of the measured points to be for simplicity:

Interval	0	1	2	3	4	5	6	7	8
x	0	1	2	3	4	5	6	7	8
y	1	2	3	4	5	6	7	8	9

The motion is described by the equation $y = 1 + x$.

Now let the coordinates to change as follows:

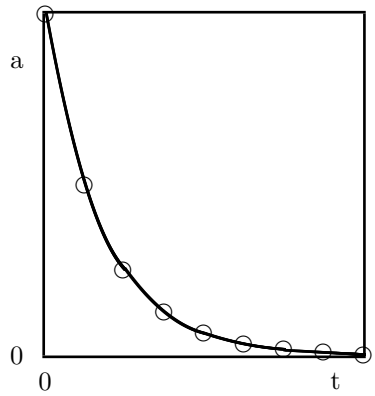
Interval	0	1	2	3	4	5	6	7	8
x	256	128	64	32	16	8	4	2	1
y	0	128	192	224	240	248	252	254	255

The velocity of the motion is not constant, but it is decreasing exponentially. The line depicting values x in different time intervals t on the graph is bent (Fig. 18.2). To straighten it, we must use an logarithmic transformation $y = \log x$. Using binary base, we get the same values as in the first example, the axis x represents the distance to the end of the course, y is the run away distance.

Now again let the baseline x to represent the position in time t, the vertical axis y the position in time $t + 1$, 1 representing an interval of time Δt. The coordinates of the exponential curve are

Interval	1	2	3	4	5	6	7	8	9
x	0	128	192	224	240	248	252	254	255
y	128	64	32	16	8	4	2	1	?

Figure 18.2: Exponential curve. The decreasing distance intervals from Zenon plot of the Achilles and turtle aporea are on the vertical axis, the horizontal axis is the time.



The x values are growing, the y values are decreasing. Both changes are not linear. Nevertheless, if the values x are plotted against the corresponding values y, the plot is linear, see Fig. 18.3.

The plot represents the exponential changes, For example: the radioactive decay or monomolecular chemical reactions if y is the starting substance, x is the product. The corresponding equation is

$$y = 2^{8-t} \,. \tag{18.3}$$

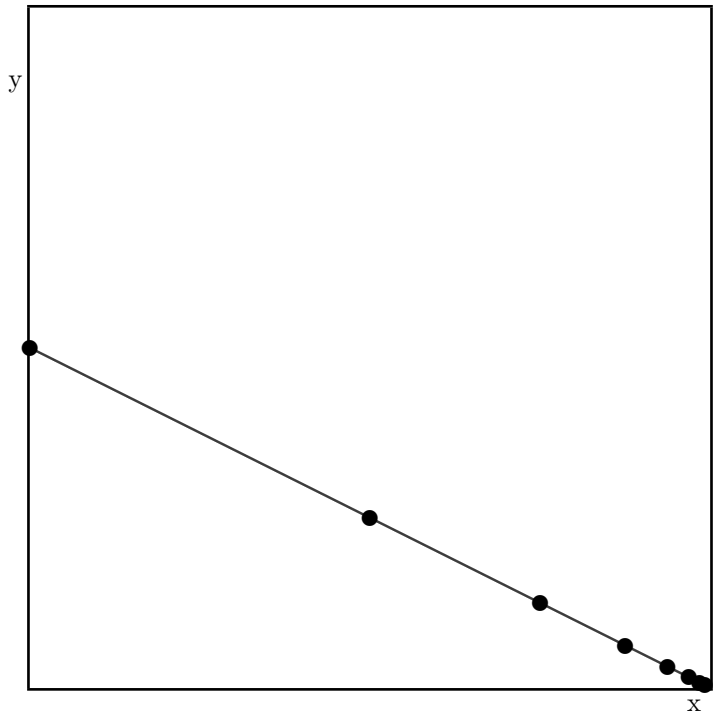
The Zenon aporea is now transformed into its modern form, the question, when the last radioactive atom will decay, and when their starting number is $x = 256$.

We are now in a similar situation as the Greeks were. The decay of radioactive elements is governed by an exponential law. The ratio of decaying atoms in equal time intervals Δ_t is constant. To be sure that all atoms decayed, we need infinitely many such intervals.

Essentially, the infinitely many intervals are needed only for the last atom, if we demand certainty of its decay.

The graph of the process is the same as in the case of the runners, if both axes, time and position, are replaced by positions (concentrations) in consecutive time intervals, t and $(t + 1)$ as if both positions were on two different orthogonal scales. By doing so, these positions are considered to be orthogonal, the exponential movement is transformed into linear, as if

Figure 18.3: Linearization of the exponential curve. The decreasing distances between points correspond to the constant time intervals.



we used the logarithmic scale[1].

18.4 Markov Matrices

Markov was a Russian mathematician who got the somewhat childish idea to study the order in which the consonants follow the vowels in a Pushkin's poem. After a consonant another consonant or a vowel can follow with some statistical probability which is determined by the structure of the language and its use by the author.

Markov studied probabilities of transitions of consecutive phonemes, as consonants c and vowels v in the example

A vv A vc M cv A vc R cc K cv O vc V

Probabilities vv, vc, cc and cv are obtained from the direct counts by dividing them with all possibilities of the transitions (here 7 transitions of 8 letters). When arranged among into matrices, they form the *stochastic matrices* **M** which row sums are 1. The theory of processes connected with these matrices forms a part of the theory of *stochastic processes*.

Each phoneme in a text is considered to be a state of the system which fluctuates constantly between its possible states, forming a chain of consecutive events.

There is another possibility to interpret the phenomenon. A text can be considered as a whole and all observed differences can form one transition into the next state. Or two distinct objects, represented by strings of symbols, can be compared. The differences can be thus expressed as arcs of a graph, for example

	? A	A A	A M	M A	A R	R K	K O	O V	V ?
c	*	0	1	-1	1	0	-1	1	*
v	*	0	-1	1	-1	0	1	-1	*

The two rows with the numbers form the transposed incidence matrix $\mathbf{S}^{\mathrm{T}}$ of the multigraph with loops, zeroes are on places of loops, arcs beginning and ending on the same site, the asterisks * mark the undetermined start and end terms. It is possible to connect the last letter with the first one for removing these loose ends.

[1]The linear movement is the limit of the exponential movement when the constant $k = 0$

Figure 18.4: Transitions of 2 letter strings. The direct transition $cc \leftrightarrow vv$ is impossible.

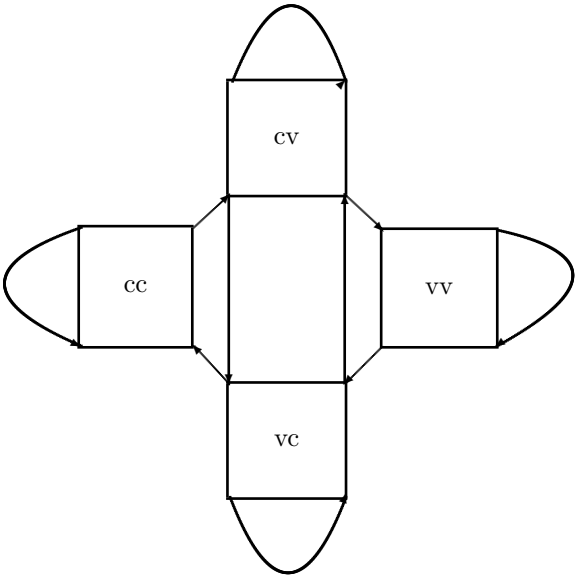
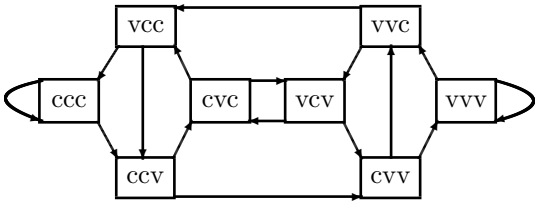


Figure 18.5: Transitions of 3 letter strings.



The string is formed by differences $(\mathbf{e}_i - \mathbf{e}_j)$ and it is clear that we can write it as the incidence matrix of an oriented multigraph with loops. On Fig. 18.4 the possible transitions of 2 letter strings are shown, on Fig. 18.5 the possible transitions of 3 letter strings are shown.

Such transitions are not limited to language. If we followed atoms of the radioactive elements for some periods of time, then each atom either remained unchanged, or it emitted a quantum of radiation, and changed into an atom of another element. Here we do not know the indexing of individual atoms, we can determine only their amount. The amount δx of atoms, which decay in a time interval, is proportional to the number of atoms x, the constant of proportionality being k, and to the length of the time interval δt. The equation describing this process is

$$\delta x/\delta t = -kx \tag{18.4}$$

The solution of this equation is found by separating of the variables in the differential form (very short δt):

$$\delta x/x = \delta(logx) = -k\delta t \tag{18.5}$$

and integrating both sides and delogarithming the result

$$x = Aexp(-kt)\ , \tag{18.6}$$

where A is the initial value of x as the integration constant. This solution has the above mentioned hook: We cannot be ever sure about the time when the last atom in the system decays, there exist only probabilities. This is the discrepancy between differential and integral calculus and finite mathematics.

The process can be visualized by two different plots, either we plot concentrations against elapsed time as on Fig. 18.2, which is the traditional technique, or we plot the concentrations of the changing substance x_t eventually the concentrations of the product $(1-x)_t$ against these concentrations x_{t+1} or $(1-x_{t+1})$, respectively after the constant time interval Δt as on Fig. 18.3 The concentrations points on this plot form straight lines which slopes depend on the velocity constants k.

Once again: The values of a function in two different time intervals were treated as orthogonal vectors. In this way we obtained a plot of a linear function from an exponential function, as if we found a logarithm of the exponential function. The orthogonal projection gave the logarithmic transformation of the exponential velocity of transformation of n atoms of two kinds.

18.5 Multidimensional Systems

According of our definition, matrices of oriented graphs describe motions on planes orthogonal to the unit vectors **I**. We are able to follow conveniently the changes of concentrations of 3 components, which can be drawn on equilateral triangles.

What is easy for two components becomes complicated for systems containing n different components which can each transform into another with different velocities k_{ij}. Nevertheless, the basics remain and such systems are described by generalized Markov matrices **M** which off-diagonal elements k_{ij}. are the rate constants of a system of equations 18.4 and the diagonal elements are the sums of rate constants with negative signs $-\Sigma k_{ij}$. The diagonal elements are either the column sums if the matrix **M** acts on the concentration vector column **c** from the left, or the row sums if the matrix **P** acts on the concentration vector row $\mathbf{c}^{\mathrm{T}}$ from the right.

18.6 Transition Matrices

A transition matrix **P** is formed from two parts, the Markov matrix **M** and the identity matrix **I**

$$\mathbf{P} = (\mathbf{I} + \mathbf{M}) \, . \tag{18.7}$$

M is the asymmetrically split Laplace-Kirchhoff matrix $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ with the negative signs on the diagonal which is normalized to the unit concentrations. The transition matrices **P** have two limits: Either the identity matrix

I, if no change occurs in the given time interval, or the permutation matrices **P**, if all species are transformed into other one within one time interval. We can suppose that each reaction (transition) in which an object is transformed into another species, say $a \rightarrow b$ in a time interval δt is registered in the incidence matrix **S** as the difference of two unit vectors $(\mathbf{e}_i - \mathbf{e}_j)$. These additive operators are transformed in the quadratic form $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ into the multiplicative operators which are normalized, it means the operator k_{ij} is the ratio of transformed objects to all present objects, and the normalized symmetrical quadratic form $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ is split into the row operator $\mathbf{P}_r$ and the column operator $\mathbf{P_c}$

$$-\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{P}_r + \mathbf{P}_c \, . \tag{18.8}$$

The adjacency matrices **A** which we used till now were the symmetrical. They were obtained as the off-diagonal elements of quadratic forms of incidence matrices of either an oriented graph **S**, or an unoriented graph **G** (see Sect. 12.7).

Since the asymmetric adjacency matrices are used as the operators, it is necessary to determine, how they are produced formally. When the vectors-rows **c** are multiplied from the right, then $a_{ij} = k$, when k arcs go from the vertex j in the vertex i, when vectors-columns **c** are multiplied from the left, then $aij = k$, when k arcs go from the vertex i to the vertex j. We will use subscripts r and l for the both kinds of the adjacency matrices **A**.

The orientation of arcs can be expressed by signs, where $a_{ij} = +k$, when k arcs go from the vertex i in the vertex j, or where $a_{ij} = -k$, when k arcs go in the vertex i from the vertex j, or opposite.

If each arc represents one transformation of the object j into the object i, and the counts k_{ij} are normalized, k_{ij}'s become the rates of reactions known in chemistry as the monomolecular reactions, velocity together with the corresponding sums Σk_{ij} on the diagonal with negative signs. When the concentration (or coordinate) vectors **c** are multiplied by these operators, the changes of concentrations are obtained, when the concentration vectors **c** are multiplied by $(\mathbf{I} - \mathbf{P})$, the new concentration vectors are obtained. We suppose, that concentration vectors are rows and the multiplication is from right

$$\mathbf{c}_{t+1}^{\mathrm{T}} = \mathbf{c}_t^{\mathrm{T}}\mathbf{M} \, , \tag{18.9}$$

therefore the sums Σk_{ij} on the diagonal are the column sums.

Let **S** and **G** be the incidence matrices of the same oriented multigraph, where **S** and **G** are identical matrices except for the signs. An unoriented edge corresponds to each arc. The rows of **S** and **G** are the mutually orthogonal vectors.

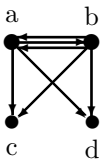


Figure 18.6: Reaction multigraph.

The corresponding scalar products $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{S}$ are the asymmetric matrices showing differences in the orientation of arcs. As an example we use the multigraph defined by the transposed incidence matrix $\mathbf{S}^{\mathrm{T}}$ (see Fig. 18.6)

$$\mathbf{S}^{\mathrm{T}} \begin{pmatrix} -1 & -1 & 0 & -1 & 0 & 1 & -1 \\ 1 & 0 & -1 & 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix} .$$

The elements of the matrix $\mathbf{G}^{\mathrm{T}}\mathbf{S}$ are

$$\begin{pmatrix} -3 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix} .$$

They can be interpreted as
v_{ii} = (arcs in - arcs out)
a_{ij} = (arcs out i in j - arcs out j in i),
then $a_{ij} = 0$ no arc.
The elements of the matrix $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ are

$$\begin{pmatrix} -3 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$

They can be interpreted as
v_{ii} = (arcs in - arcs out)
a_{ij} = (arcs out i in j - arcs out j in i),
then $a_{ij} = 0$ no arc.
The off-diagonal elements of the matrix $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ differ from the off-diagonal elements of the matrix $\mathbf{G}^{\mathrm{T}}\mathbf{S}$ only by signs. The scalar products $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{S}$ can be combined with the quadratic forms of incidence matrices. There are four additive combinations

$$\mathbf{S}^{\mathrm{T}}\mathbf{S} \qquad \mathbf{G}^{\mathrm{T}}\mathbf{G}$$

$$\begin{pmatrix} 5 & -3 & -1 & -1 \\ -3 & 5 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 5 & 3 & 1 & 1 \\ 3 & 5 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{S}^{\mathrm{T}}\mathbf{S} \qquad \mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{S}^{\mathrm{T}}\mathbf{S}$$

$$\begin{pmatrix} 2 & -2 & 0 & 0 \\ -4 & 6 & 0 & -2 \\ -2 & -2 & 4 & 0 \\ -2 & 0 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} -8 & 4 & 2 & 2 \\ 2 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 \end{pmatrix}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{G}^{\mathrm{T}}\mathbf{G} \qquad \mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{G}^{\mathrm{T}}\mathbf{G}$$

$$\begin{pmatrix} 2 & 4 & 2 & 2 \\ 2 & 6 & 2 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} -8 & -2 & 0 & 0 \\ -4 & -4 & 0 & -2 \\ -2 & -2 & 0 & 0 \\ -2 & 0 & 0 & -2 \end{pmatrix} .$$

This gives this pattern

$$\begin{aligned} \mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{S}^{\mathrm{T}}\mathbf{S} &= 2(\mathbf{V}_{in} - \mathbf{A}_r) \\ \mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{S}^{\mathrm{T}}\mathbf{S} &= 2(\mathbf{A}_l - \mathbf{V}_{out}) \\ \mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{G}^{\mathrm{T}}\mathbf{G} &= 2(\mathbf{V}_{in} + \mathbf{A}_l) \\ \mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{G}^{\mathrm{T}}\mathbf{G} &= -2(\mathbf{A}_r + \mathbf{V}_{out}) \, . \end{aligned}$$

The scalar product $(\mathbf{G} - \mathbf{S})^{\mathrm{T}}\mathbf{S}$ can be normalized into the left hand side operator $\mathbf{M}$. The diagonal matrices of vertex degrees (arcs in and out), as well as the asymmetric adjacency matrices can be separated by the transposing sums or differences $\mathbf{G}^{\mathrm{T}}\mathbf{S}$ with $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and combining them with the sums or differences $\mathbf{G}^{\mathrm{T}}\mathbf{S}$ with $\mathbf{S}^{\mathrm{T}}\mathbf{S}$:

$$4\mathbf{V}_{in} = (\mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{S}^{\mathrm{T}}\mathbf{S}) + (\mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{G}^{\mathrm{T}}\mathbf{G})^{\mathrm{T}}$$

$$-4\mathbf{V}_{out} = (\mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{S}^{\mathrm{T}}\mathbf{S}) + (\mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{G}^{\mathrm{T}}\mathbf{G})^{\mathrm{T}}$$

$$-4\mathbf{A}_l = (\mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{S}^{\mathrm{T}}\mathbf{S}) - (\mathbf{G}^{\mathrm{T}}\mathbf{S} + \mathbf{G}^{\mathrm{T}}\mathbf{G})^{\mathrm{T}}$$

$$4\mathbf{A}_r = (\mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{S}^{\mathrm{T}}\mathbf{S}) - (\mathbf{G}^{\mathrm{T}}\mathbf{S} - \mathbf{G}^{\mathrm{T}}\mathbf{G})^{\mathrm{T}} .$$

The same operation with $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ gives the pattern:

$$\mathbf{S}^{\mathrm{T}}\mathbf{G} + \mathbf{S}^{\mathrm{T}}\mathbf{S} = 2(\mathbf{V}_{in} - \mathbf{A}_l)$$

$$\mathbf{S}^{\mathrm{T}}\mathbf{G} - \mathbf{S}^{\mathrm{T}}\mathbf{S} = 2(\mathbf{A}_r - \mathbf{V}_{out})$$

$$\mathbf{S}^{\mathrm{T}}\mathbf{G} + \mathbf{G}^{\mathrm{T}}\mathbf{G} = 2(\mathbf{V}_{in} + \mathbf{A}_r)$$

$$\mathbf{S}^{\mathrm{T}}\mathbf{G} - \mathbf{G}^{\mathrm{T}}\mathbf{G} = -2(\mathbf{A}_l + \mathbf{V}_{out}) .$$

The scalar product $\mathbf{S}^{\mathrm{T}}(\mathbf{G} - \mathbf{S})$ can be normalized into the right hand side operator $\mathbf{M}$. The diagonal matrices of vertex degrees (arcs in and out), as well as the asymmetric adjacency matrices can be separated by transposing the sums or differences $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ with $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and combining them with sums or differences $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ with $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ as above. These transposes are identical with sums or differences of $\mathbf{G}^{\mathrm{T}}\mathbf{S}$, because the transposing changes the ordering of matrices in the product.

The incidence matrices $\mathbf{S}$ and $\mathbf{G}$, or their transposes, used as the multiplication operators, transfer each element of the multiplied matrix vector twice, once on the diagonal, once as off-diagonal element. The sums or differences of these matrices $\mathbf{S}$ and $\mathbf{G}$, which should be transformed into the quadratic matrices, have in each row exactly one element 2 in the ending or starting column, respectively. The results are thus elementary. But these facts are not explained in textbooks or in current literature. If they were studied earlier, they were forgotten.

The double entry accounting of the arcs using the orthogonal vector strings, their sums and differences, quadratic forms, scalar products and transposes, gives a web of related matrices describing the graphs and to them isomorphic objects and their transformations.

The Laplace-Kirchhoff matrix, identical with $\mathbf{S}^{\mathrm{T}}\mathbf{S}$ and used for solving electrical circuits, is symmetrical. It actually describes only the properties of the circuit, the resistances of lines (conductors) connecting the vertices of the net. The direction of the flow is induced by the applied tension. The matrix of currents corresponds to one from the matrices $\mathbf{S}^{\mathrm{T}}\mathbf{G}$ or $\mathbf{G}^{\mathrm{T}}\mathbf{S}$, currents k in the branches always have the opposite signs

$$(\mathbf{S}^{\mathrm{T}}\mathbf{G})_{ij} = -(\mathbf{S}^{\mathrm{T}}\mathbf{G})_{ij} . \tag{18.10}$$

Moreover, flows in and out from all vertices must be balanced, $\Sigma k_{ij} = 0$. Since the resistance can be expressed as the length of a conductor, the inverse problem appears as the resistance distances.

18.7 Equilibrium Concentrations

Finding the diagonal matrix $\mathbf{C}$ of the equilibrium concentrations c_j* for large systems is not a simple task. It requires calculations of the determinants of all submatrices of the matrix product $\delta_j\mathbf{MC}$, obtained by deleting the j-th row and column. Many variants of the Kirchhoff technique of spanning trees were elaborated for this purpose.

Today the technical difficulties are removed by the use of computers but a basic question remains open: is the product $\mathbf{MC}$ a symmetrical matrix or not? Wei and Prater [**?**], who elaborated the matrix technique for solving of systems of exponential equations, argued by the principle of microscopic reversibility according to which the equivalence should be true:

$$c_i^* k_{ij} = c_j^* k_{ji} \; . \tag{18.11}$$

The properties of the essentially positive matrices make the validity of this principle doubtful. We use the properties of the eigenvalues of the Markov matrices and will study the operator $\mathbf{P} = (\mathbf{I} + \mathbf{M})$. This operator transforms the concentration vector $\mathbf{c}_t$ in time t into the concentration vector $\mathbf{c}_{t+1}$ in time $(t + \delta)$.

18.8 Properties of Matrix Sums (I + M)

The matrices $(\mathbf{I} + \mathbf{M})$ have one eigenvalue exactly 1, the other eigenvalues are in the circle $0 < \lambda_j < 1$. The matrix $\mathbf{M}$ has exactly one eigenvalue equal to zero and the remaining $(n - 1)$ eigenvalues in the range limited by the circle given by the rate sums $\Sigma - k_{ij}$. Because a transformation of any species can not be greater that its concentration, the sum of the rate constants must be lesser than 1. If the regular unit matrix $\mathbf{I}$ is added to $\mathbf{M}$, all eigenvalues are increased evenly by 1. This has an important consequence which remained unnoticed: The equilibrium state of the operator $\mathbf{P}$ has one eigenvalue exactly 1, all other eigenvalues are 0. The product of any concentration vector $\mathbf{c}$ with the equilibrium operator $(\mathbf{I} + \mathbf{M})^\infty$ must give the equilibrium concentration vector $\mathbf{c}^*$. Therefore $(1/n)\mathbf{I}(\mathbf{I} + \mathbf{M})^\infty$ has the form of n identical columns of the equilibrium concentration vectors $\mathbf{c}^{\mathrm{T}}$. Because the sum of concentrations is always $\Sigma_{j=1}^{n} = 1$ this result conforms with the condition $\mathbf{c}(\mathbf{I} + \mathbf{M})^\infty = \mathbf{c}^{*\mathrm{T}}$.

The other important property of the equilibrium operator is that its product with the Markov matrix $\mathbf{M}$ must give the zero matrix $\mathbf{0}$: $\mathbf{M}(\mathbf{I}+\mathbf{M})^{\infty}=\mathbf{0}$. To show some consequences, we separate the equilibrium matrix operator into the diagonal matrix $\mathbf{C}$ which elements are equilibrium concentrations c_j^* and the matrix of off-diagonal elements $[\mathbf{M}(\mathbf{I}+\mathbf{M})^{\infty}-\mathbf{C}]$. The products with the Markov matrix have the following forms:

$$\mathbf{M} \;=\; \begin{pmatrix} -c_1^*\Sigma k_{i1} & c_2^* k_{12} & \dots & c_n^* k_{1n} \\ c_1^* k_{21} & -c_2^*\Sigma k_{i2} & \dots & c_n^* k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_1^* k_{n1} & c_2^* k_{n2} & \dots & -c_n^*\Sigma k_{in} \end{pmatrix} .$$

$$\mathbf{M}[(\mathbf{I}+\mathbf{M})^{\infty}-\mathbf{C}]$$

$$\begin{pmatrix} \Sigma_{i=1} c_i k_{1i} & \Sigma_{i\neq 2}(c_i^* k_{i1} - c_1^* k_{i1}) & \dots & \Sigma_{i\neq n}(c_i^* k_{1n} - c_1^* k_{i1}) \\ \Sigma_{i\neq 1}(c_i^* k_{2i} - c_2^* k_{2i}) & \Sigma_{i=2} c_i k_{2i} & \dots & \Sigma_{i\neq n}(c_i^* k_{2n} - c_1^* k_{i2}) \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{i\neq 1}(c_i^* k_{ni} - c_n^* k_{ni}) & \Sigma_{i\neq 2}(c_i^* k_{ni} - c_n^* k_{ni}) & \dots & \Sigma_{i=n} c_i k_{ni} \end{pmatrix} .$$

The equilibrium condition is fulfilled if

$$\sum_{j=n}^{n} c_j^* k_{ji} - \sum_{i=n}^{n} c_i^* k_{ij} = 0 \,. \tag{18.12}$$

All flows in each position in the matrix must be balanced by all outflows to keep equilibrium. For this the principle of microscopic reversibility is not a necessary condition, but it is only a special case from all possibilities, how the equilibrium can be reached.

Because any equilibrium state of the operator $\mathbf{P}$ has exactly one eigenvalue 1 and other $(n-1)$ eigenvalues are 0, it is easy to

find the corresponding eigenvectors. The unit eigenvector is the unit row $\mathbf{J}^{\mathrm{T}}$ or the unit column $\mathbf{J}$, respectively. The zero eigenvectors can be chosen as any $(n-1)$ rows or columns of the Markov matrix. Any Markov matrix is therefore a system of eigenvectors of its equilibrium state.

18.9 Classification of Markov Matrices

A Markov matrix describes its own equilibrium state and all the paths to the equilibrium from any point of the n dimensional concentration simplex. This simplex is a plane orthogonal to the unit vector $\mathbf{I}$, For example: for

3 substances it is an equilateral triangle. Each point of the simplex can be the equilibrium point of the system and to each equilibrium point there go infinitely many paths. Therefore it is necessary to classify the Markov matrices according to the character of paths the matrix produces. If we exclude matrices going to concentrations outside the simplex, there are three possibilities. Easily they can be found for the two dimensional case:

A	**B**	**C**
$p, q < 0.5$	$p = q = 0.5$	$p, q > 0.5$
$\begin{pmatrix} (1-p) & p \\ q & (1-q) \end{pmatrix}$	$\begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$	$\begin{pmatrix} (1-p) & p \\ q & (1-q) \end{pmatrix}$

- **A**: Smooth approach. The transformation lines are inside the frame formed by the diagonal and the axis x. The determinant of **P** is greater than 1. The first step can lead immediately to the equilibrium concentration.

- **B**. Oscillating approach. This can be recognized simply by the reaction constants. If $k_{ij} > c_j^*$, then the system oscillates when the reaction starts from the vertex of the reaction simplex $c^i = 1$. In the first step the concentration c_j jumps over the equilibrium concentration. Here the time conditions should be studied, that is the relations between the time intervals needed for transformation of an object into another one. These intervals are surely different for n different objects and whole reaction intervals. We can not suppose that all objects react simultaneously and therefore the reaction intervals can be much longer than the transformation intervals of individual objects. But this difference induces lability and can lead to oscillations of other kinds.

- **C**. The steepest approach. The reaction path should be a straight line going from any concentration point to the equilibrium. This requires that the reaction constants of each substance must be proportional to the equilibrium concentrations of the target substances. For example: for 3 substances: $c_1 k_{12} = a c_2^*$ and $c_1 k_{13} = a c_3^*$. From the microscopic reversibility conditions $c_2^* k_{23} = c_3^* k_{32}$ we obtain the relation of reaction constants $k_{23}/k_{13} = k_{23}/k_{12}$. For other two substances we obtain similarly for c_2: $k_{21}/k_{31} = k_{23}/k_{12}$ and for c_3: $k_{31}/k_{21} = k_{32}/k_{12}$. Comparing all three results, we see that such approach is possible only for $c_j^* = 1/3$, that is for the center of the simplex.

The principle of microscopic reversibility assures the steepest approach only on straight lines connecting the equilibrium state with vertices of the simplex, one pure substance reacts or one substance is depleted from the equilibrium state. It is a special path and it is questionable. It is much easier to allow the existence of cyclic flows which must be balanced in equilibrium by the condition for species in a cycle

$$k_{ij} = (k + k')/c_i^* \ , \ \textit{where } k' = c_j^* k_{ij} \ . \tag{18.13}$$

The steepest descent to the equilibrium might be the optimal path in the concentration simplex, but it is not possible to prove that it is the only possible path for all reaction systems and conditions. It is not possible to prove that the matrix product **MC** is a symmetrical matrix. On the other side, it is rather easy to find the conditions for the oscillating reaction systems. A sufficient condition is when k_{ij} are relatively great numbers. Of course, such values violate conditions of differential reactions, it is assumed that the increments $\delta x/\delta t$ are infinitesimally small but the matrix multiplication shows why oscillations emerge: in one time interval there are not sufficiently great concentrations of the backfeed products to balance the loss $c_j \Sigma k_{ij}$ if both values c_j and Σk_{ij} are great. Because $(\mathbf{I} + \mathbf{M})^b \neq (\mathbf{I} + b\mathbf{M})$, we cannot choose time intervals Δ_t freely. They should be comparable with intervals needed for reactions. If some reactions need substantially longer times, oscillations emerge as in the Lotka-Woltera cycle.

18.10 Jakobi Approximations

We have shown the exact methods for solving the equation $\mathbf{Mx} = \mathbf{b}$ in Chapt. 16, based on the inverting of the matrix **M** or finding its eigenvalues. In case, when we are not able to do such sophisticated mathematical operations, we can try to guess the right answer. We have counted the matrices and we know, that if we limit ourselves to natural numbers, their number is not infinite. Therefore, using computers, it is possible to find the solution by the trial and error methods, especially, if the results are compared with the target values and impossible combinations excluded. This technique of fluctuation can be compared with the process by which a system seeks its equilibrium.

Let us start with the guess vector **y**. After multiplication with the matrix **M** we get the guess vector **g**. Comparing it with the target vector **b** we obtain the difference $d_{\mathbf{g}-\mathbf{b}}$. If it is zero, our guess coincides with the searched vector and we can end our search. Similarly if the difference $d\mathbf{g} - \mathbf{b}$ is negligible we can stop our search. Otherwise we must correct the original guess vector using $d\mathbf{g} - \mathbf{b}$. But we cannot apply the whole

difference, because the next guess could be as a pendulum on the other side of the true values. We must lessen the fluctuations. The correction must be smaller than the difference, which is achieved by using a multiplication constant c: $0<c<1$. If we choose the constant too low, we need too many steps to find an acceptable value of $\mathbf{g}$, if c is too close to 1, the results could fluctuate, similarly as was shown for the Markov matrices.

Chapter 19

Entropic Measures and Information

19.1 Distances and Logarithms

Maybe you know that information can be measured by its entropy

$$H = -\Sigma p_j \log_2 p_j \tag{19.1}$$

where the sum is made over all probabilities p_j of objects (symbols). These probabilities are unknown and we leave them at first undefined.

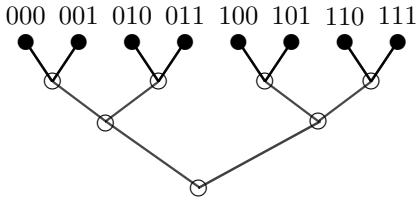
Nobody cared to explain, why this function is suitable as the measure, it was just introduced as an axiom. We now define this function as a simple result of mapping of m objects on vertices of a multidimensional unit cube, or equivalently, indexing these objects by a regular code consisting from 0 and 1 symbols or simply by using binary number scale having equal number of digits:

Decimal	0	1	2	3	4	5	6	7
Binary	000	001	010	011	100	101	110	111

The least necessary number of digits for each object from m objects is close to $\log_2 m$. These digits count edges of a binary decision graph on which leaves the counted objects are placed (Fig. 19.1) [1].

[1] Please arrange the leaves onto the vertices of the cube and draw the decision tree yourselves. I tried it but my figure was too ugly. The cube as well as the decision tree must be deformed.

Figure 19.1: Binary decision tree is isomorphic with indexing of m objects by binary digits.



For all m objects we need at least $m \log_2 m$ digits (in our example 24 digits). This limit is obtainable only if m is a power of 2. Nevertheless it can be used for elementary calculations of logarithms with a satisfactory precision. The number of digits m_j is the distance of the leave j from the root in the decision tree. Therefore the logarithms are related to the distances.

Knowing that $3^5 = 243$, we construct a binary decision tree with 1937 edges

- 128 * 8 = 1024
- 64 * 8 = 512
- 32 * 8 = 256
- 16 * 8 = 128

  Till now 15 branches each with 16 leaves from 16 stems of the fourth degree were used fully for indexing 240 leaves (objects) by 1920 digits. The shorter tree budding from the last stem is used for the last three leaves

- 2 * 6 = 12
- 1 * 5 = 5

The sum of the distances of the leaves from the root is 1937. Thus 1937 : 243 = 7.971. The result of the division is the mean distance which equals to $\log_2 3^4$. The estimation of the binary logarithm of 3 is 7.971 : 5 = 1.597. Since $\lg_2 3 = 1.585$, the precision for such simple calculation is good and could be improved using higher powers of the searched number close to the power of the base number.

The calculations can be done for any natural number of branches. As an example: $5^{10} = 9765625$. The corresponding rooted tree with 10 branches has the length somewhat lesser than 7, which is simply the number of digits. Accepting this rough estimate, and dividing by 10, we get as the estimate 0.70000. The value obtained by the calculator is $\log_{10} 5 = 0.69897$.

After this excursion we return to the entropy function. If we have some information about the counted objects, the necessary number of digits can be decreased. Suppose, that the objects are already indexed by n symbols of an alphabet. The new indexing can be composed from two parts, the symbol j and the binary code proper for each specific symbol. Now we need only $\Sigma m_j \log_2 m_j$ symbols. The difference

$$H = m \log_2 m - \sum_{j=1}^{n} m_j \log_2 m_j = \sum m_j/m \log(m_j/m) \tag{19.2}$$

will be the measure of the information gained by dividing the set of m objects into n labeled subsets. Introducing $p_j = m_j/m$ and dividing the result by the number m, we obtain the entropy H_m relative to 1 object.

For example: the string *aaaabbcd* and its permutations need only 10 digits:

Decimal	0	1	2	3	4	5	6	7
Binary	a00	a01	a10	a11	b0	b1	c	d

The normalized difference against the full tree $H = (24 - 10)/8 = 1.75$ is the information entropy of the string[2].

Unfortunately, this simple explanation does not explain the entropy function H. This is only an approximation of its one form, based on the binary logarithms.

19.2 Boltzmann's Entropy Function H_n

On the Boltzmann's tomb the formula

$$S = -k \ln W \ , \tag{19.3}$$

is engraved, where S stands for the thermodynamic entropy, W to Wahrscheinlichkeit, that means probability, and k is a constant named in honor of Boltzmann. This formula was the cause of his death. He died

[2] A reviewer of a prestigious education journal did not believed it and rejected my paper.

exhausted by vain efforts to prove it. Even his friends concocted aporea to disprove Boltzmann's ideas. His tragedy was, that nobody understood his proof which I try to explain by this book.

The entropy was defined by Clausius by its difference. The entropy difference is the ratio between the specific heat Q needed to increase temperature T of some substance and the given temperature T: dS = dQ/T. If the specific heat was constant, the integrated form would be

$$S = C \log T + S_0 \ . \tag{19.4}$$

It is accepted that the entropy at absolute zero is zero. Therefore the integration constant S_0 must be $-C \log 0$. But the entropy is much more complicated function, because the specific heat Q depends on temperature and has singularities, as the melting and evaporation heats are. We concentrate on the fact that the entropy is a logarithmic function of temperature. What is the temperature? This is a measure of the thermal motion of molecules [3]. In a system of ideal gas, the molecules represented by points move haphazardly and if they collide, they exchange their kinetic energy, but the total amount of the energy at the constant temperature remains constant. Moreover, if the system remains isolated, the distribution of energies of molecules reaches spontaneously an equilibrium. This is the largest orbit, where the system is stable for long periods of time.

The entropy function is considered to be mysterious. Not only for its abstract form (we do not feel it directly as the temperature, pressure and volume) but for its property. It is increasing spontaneously. To decrease the entropy needs an outside action.

We have shown that the surfaces of constant energy in the phase space are planes orthogonal to the unit vector **I**. The system of the ideal gas moves on this plane and for most of the time it remains on each orbit proportionally to its volume. Therefore the system exists in the largest orbit or orbits nearest to it for most of time. We already know the formula for the evaluation of volumes of individual orbits. This is the polynomial coefficient for n permutations

$$n!/\Pi \ n_k! \ . \tag{19.5}$$

The logarithm of this coefficient was proposed by Boltzmann as a mathematical equivalent of entropy, the H function. If n and n_k are large numbers, and in the case of the ideal gas they certainly are (the Avogadro number, determining the number of molecules in 1 mole of gas, is of order 10^{23}), the Stirling approximation of n! can be used and the result is

[3] According to a more sophisticated definition, T is an integrating factor.

$$H_n = -\Sigma(n_k/n)\log(n_k/n)\ . \tag{19.6}$$

This result can be obtained only with natural logarithms, unlike in the information entropy. Usually, the ratios n_k/n are replaced by a symbol p_k, where p should be the probability.

Boltzmann's problem was that he only conjectured the existence of the quanta of energy (they were discovered in time of the Boltzmann's death by Planck) and that he, instead of speaking about symmetry of the partition orbits, introduced ill defined probabilities p_k which replaced the true ratios n_k/n.

One paradox against Boltzmann was connected with the time inversion. The classical mechanics supposed that the time can be inverted. But such time inversion should lead to the decrease of entropy. This could be taken as an evidence against the H theorem. We have shown that space is not insensitive to changes of signs, the negative cone has quite different properties than the positive one. Nevertheless the sign of the entropy changes only classifies the natural processes. We can say that if a time inversion led to the decrease of the entropy of a system then this time inversion is not a spontaneous phenomenon, because its cause lies outside the system.

19.3 Maximal H_n Entropy

Searching the maximal value of the function 19.6 seems to be an easy task. The entropy H_n is maximal when all values $n_j = 1$. This monotone solution has a fault: It can be achieved only at a special value of the arithmetical mean m/n. The sum of the arithmetical progression 1 to n is $\binom{n+1}{2}$, therefore the arithmetical mean of values m_j necessary for the linear distribution is $(n-1)/2$, one half of the number of the objects. This value is acceptable only at small systems.

At large systems as gas molecules are, the monotone distribution is unachievable. The Avogadro number N is 6.023×10^{23} (one mole of hydrogen weights about two grams), the Boltzmann constant k (k= R/N) is 1.38×10^{-23} Joule/grad and the gas constant R is 8.314 Joule/grad. The monotone distribution would require temperatures in Kelvin's (centigrade degrees) in the range of the Avogadro number.

The distribution of gas molecules can not be monotone. Nevertheless, it must be as flat as possible.

We investigate at first relations of the means of some skewed distributions.

The straight slopes

n_k	6	5	4	3	2	1	$\sum 21 = \binom{k+1}{2}$
m_k	0	1	2	3	4	5	
$n_k \times m_k$	0	5	8	9	8	5	$\sum 35 = \binom{k+1}{3}$

give the arithmetical mean $(k-1)/3$, approximately $\sqrt{2n}/3$.

The exponential slopes

n_k	32	16	8	4	2	1	$\sum 63 = 2^6 - 1 = 2^{k+1} - 1$
m_k	0	1	2	3	4	5	
$n_k \times m_k$	0	16	16	12	8	5	$\sum 57 = 2^6 - 7 = 2^{k+1} - 2k + 1$

have the arithmetical mean for all sizes somewhat lesser than 1. Starting m_k values from the lowest value r, the arithmetical mean will be always $r+1$, since we add to the basic distribution $r \times 2^{k+1} - 1$ units. The exponential slopes can be flattened by combining several such distributions:

n_k	8	8	4	4	2	2	1	1	$\sum 30 = 2 \times (2^4 - 1)$
m_k	0	1	2	3	4	5	6	7	
$n_k \times m_k$	0	8	8	12	8	10	6	7	$\sum 59$

The arithmetical mean grows slowly and the slopes can be flattened by equilibrating neighbor values.

A distribution can be symmetrical.

A straight distribution in the form of a ridge roof gives a somewhat better result then the monotone distribution: Its arithmetical mean is in the range of the square root of n:

n_k	1	2	3	4	3	2	1	$\sum 16 = 4^2$
m_k	0	1	2	3	4	5	6	
$n_k \times m_k$	0	2	6	12	12	10	6	$\sum 48 = 3 \times 4^2$

The binomial distribution gives this result

n_k	1	6	15	20	15	6	1	$\sum 64 = 2^6$
m_k	0	1	2	3	4	5	6	
$n_k \times m_k$	0	6	30	60	60	30	6	$\sum 192 = 3 \times 2^6$

If $n = 2^k$ then the arithmetical mean of the binomial distribution is $k/2$. For the Avogadro number $k \simeq 79$ ($2^79 = 6.045 \times 10^23$). The arithmetical

mean is very low. This means that the distribution can be flatter and contain more values than 80.

The flatter binomial distribution can be modeled as

n_k	1	1	4	4	6	6	4	4	1	1	$\sum 32 = 2 \times 2^4$
m_k	0	1	2	3	4	5	6	7	8	9	
$n_k \times m_k$	0	1	8	12	24	30	24	28	8	9	$\sum 144 = 9 \times 2^3$

The entropy can be again increased by leveling the slope as $1, 2, 3, 5, 5 \ldots$. Try the triple binomial distribution and find the corresponding equations.

The increasing and decreasing exponential slopes:

n_k	1	2	4	8	4	2	1	$\sum 22 = 2^3 - 1 + 2^4 - 1$
m_k	0	1	2	3	4	5	6	
$n_k \times m_k$	0	2	8	24	16	10	12	$\sum 72$

The distribution is composed from two components. The decreasing exponential slope with $n = 2^{k+1} - 1$ parts has the mean value $k + 1$. The increasing exponential slope with $n = 2^k - 1$ parts has the sum $n_k \times m_k = \sum_{k=0}^{k-1} k2^k$. Its mean value is somewhat greater than $(k-2)$ but lesser than k, since the last term in the sum is decisive. The arithmetical mean is approximately k. The exponential slopes can be again flattened as before.

The entropy H_n would be maximal when the distribution would be as flat as possible and approaching to the monotone distribution. If there is room enough for all parts, the distribution will be symmetrical one, otherwise it can be skewed one.

19.4 Shannon's Entropy Function H_m

A statement from a recent abstract in Chemical Abstracts [15]: "Boltzmann entropy is an information entropy", is typical for the state of art. It is generally believed, that Shannon entropy function H_m is more sophisticated and therefore better defined than Boltzmann entropy function H_n. But both functions measure related but nevertheless different properties. They can even be additive.

One can speculate, who was Jack with a Lantern, who changed the great enigma connected with entropy into a greater error. Its consequences are spread from mathematics, over physics, biology, social sciences to philosophy.

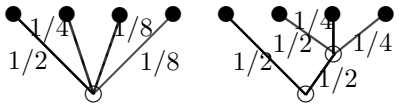
J. Von Neumann gave this advice to Shannon [16]:

Figure 19.2: Decisions from four possibilities.

> "You should call it entropy, for two reasons. In the first place, your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage."

The basic idea of Boltzmann's proof of the H theorem was not understood and remained obscure (Kac [17] "a demonstration").

We have shown the derivation of the equation 19.1 and what it measures. Shannon chose the function H deliberately from somewhat other reason. He was interested in frequencies of symbols in messages (or in the ratios of the individual frequencies m_j of the individual symbols j to the total number m of all symbols m_j/m). The function H is additive when the decisions are split as on Fig. 19.2



The most important difference of 19.2 against 19.6 is the maximal value of both functions. H_m is maximal when all symbols have the same frequencies which are equal to the arithmetical mean $\overline{m} = m/n$. Then $n_{m/n} = n$ (other $n_k = 0$) and the entropy H_n is minimal, zero. The entropy H_m has a cumulative effect on the distribution. It decreases its spread.

The fact of existence of two entropy functions explains the so called redundancy of the information, since H_m in texts is not maximal. When m entropy is maximal, n entropy is minimal and their sum is not optimal. If all symbols appeared in our speech with equal frequencies, differences between words were negligible and difficult to be noticed. There are 6 permutations of *aabb* and only 4 permutations of *aaab*. But there exists 4 strings *abbb* on the same partition and together 8 string.

It is better to explain it on words as basic vectors of information. We must repeat words connected with the subject we are speaking about. These key words which are necessary for understanding are more frequent. Changing word frequencies in messages according their subjects gives us opportunity to formulate more different messages than if all words were used evenly and to recognize immediately what is spoken about.

We have shown the simple interpretation of the information entropy. Now we introduce this function as the analogy of the Boltzmann's entropy

Function H_n. This is the logarithmic measure of the polynomial coefficient for n permutations $n!/Pi\ n_k!$. There exists the polynomial coefficient for m permutations

$$m!/Pi\ m_j! = m!/Pi_k!^{n_k}\ . \tag{19.7}$$

There exist two polynomial coefficients, one for the n permutations, the other for m permutations. What are the properties of the polynomial coefficient for m permutations?

This coefficient determines how many strings can be formed from m symbols on the alphabet of n symbols. In other words, how many different messages have place.

The coefficient

$$m!/\prod_{j=1}^{n}\ n_j = \prod_{k\geq 1} n_k!^{m_k} \tag{19.8}$$

can be modified similarly as in the case of 19.6 using the Stirling approximation of m factorials. Of course, the problem is that the numbers are rather small and the aproximation is worse. The result has the same form as 19.6, except that p_k are the relative frequencies of individual symbols.

There exists a decisive difference, the function H_m has maximum, when all symbol are used evenly.

19.5 Distances and Entropy

To answer a question how many angels can be placed on a point of a needle is not a task of mathematics, but to analyze the work of Maxwell's demon is, since this creature is still with us not only in physics but also in theory of information.

The demon transformed a mixed string of cool molecules c and hot molecules h

chchchchchchchchchchchchchchchchch

into a string in the form

ccccccccccccccccchhhhhhhhhhhhhhhhhhh

Till now we considered both strings as equivalent, since both strings are on the same orbit. When we imagine them in the two dimensional space, both strings are distinguishable. Let fill a long string two volumes of a book. We observe then both strings as two distinct states, one volume with the hot molecules h has a higher temperature than the other one with cool molecules c. The mixed strings (states corresponding to them) have an intermediate temperature and higher physical entropy.

Table 19.1: Logical functions

conjunction:		if p and q,		then (p and q)
alternative:		if p and q,		then (p or q)
implication:		if p and q,		then (p is q)
p	q	conjunction	alternative	implication
1	1	1	1	1
1	0	0	1	0
0	1	0	1	0
0	0	0	0	1

The problem is to find a way how to measure their difference. One possibility is to express it using distances between symbols of one kind. For such short strings it is necessary to close them to a loop, to avoid truncation problems connected with both ends.

The distances between symbols c are then
2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,
and
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,20, respectively.

The distances between symbols h are here the same.

The distribution of distances in both cases is quite different and the effect of mixing can be measured exactly as for the original strings by the polynomial coefficients.

The distribution of distances in the binomial distribution is known as the negative binomial distribution. For more symbols we can speak about the negative polynomial distribution.

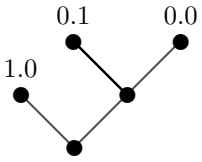
19.6 Logical functions

Our thinking is governed by logical laws, as conjunction, alternative, implication or other logical functions are. Some predicate can be true or false. The true predicate has value 1, the false predicate has value 0. Now there are known many valued logic or the fuzzy logic, when the false predicate can have any value between 1 and 0. Two predicates are combined and the result depends on the law which must be applied.

The logical decision p can be represented by a tree with two branches. The left one means true, its value is 1. The right branch means zero. On the corresponding branch is grafted the tree for the second predicate q. To the ends of its branches the new logical values are attributed according to tables of logical functions.

The conjunction function is obtained by usual multiplication of $p \times q$.

Figure 19.3: Decision tree. The left branch means 1, the right branch means 0. The root is taken as the decimal point.



The tree valued logic, allowing values 1, 0.5 and 0 can be represented by a decision tree with more branches, when the binary point is placed after the first value (Fig. 19.3). The value 0.1 means zero and 2 , it is 0.5, since 1 is 2. 0.0 is rounded to 0. The right branch could have values 1.1 and 1.0, but the values greater than 1 are truncated.

The logical operations can be viewed as operations of symmetry, attributing to different points of logical space given values.

Bibliography

[1] J. Riordan, An Introduction to Combinatorial Analysis, John Wiley, New York, 1958.

[2] L. Boltzmann, Über die Beziehung zwischen dem zweiten Hauptsatze der mechanishen Wärmetheorie und die Wahrscheinlichkeitsrechnung, *Wiener Berichte* **1877**, 76, 373.

[3] C. E. Shannon, The Mathematical Theory of Communication, *Bell System Technical Journal*, **1948**, 27, 379, 623.

[4] J. Hašek, The Brave Soldier Švejk.

[5] M. Kunz, What is Entropy (in Czech), *Věda a technika mládeži*, **1979**, 33, 552, , 616.

[6] W. Feller, An Introduction to Probability Theory and its Applications, J. Willey, New York, 1970, Chapter 10.4.

[7] W. Heisenberg in The Physicist's Conception of Nature, Ed. J. Mehra, D. Reidel, Dortrecht, **1968**, p. 267.

[8] M. Hall Jr., Combinatorial Theory, Blaisdell Publ. Comp., Waltham, 1967.

[9] F. Harary, Graph Theory, Addison-Wesley, Reading, 1969.

[10] F. Harary, E. M. Palmer, Graphical Enumeration, Academic Press, New York, 1973

[11] D. Cvetkovic, M. Doob, H. Sachs, Spectra of Graphs, Deutcher Verlag der Wissenshaften, Berlin, 1980.

[12] G. E. Andrews, The Theory of Partitions, Addison-Wesley Publ. Comp., Reading, MA, 1976.

[13] S. Weinberg, Mathematics, the Unifying Thread in Science, Notices AMS, 1986, 716.

[14] J. Wei, C. D. Prater, Structure and analysis of complex reaction systems. In D.D. Eley, P. W. Selwood, P. B. Weisz Eds., Advances in Catalysis, Vol. XIII, 203-392, Academic Press, New York, 1962.

[15] E. B. Chen, Boltzmann Entropy, Relative Entropy and Related Quantities in Thermodynamic Space, *J. Chem. Phys.*, **1995**, 102, 7169-79; CA 122: 299958.

[16] M. Tribus, E. C. McIrvine, Energy and Information, *Scientific American*, **1971**, 225, 3, 179.

[17] M. Kac in J. Mehra, Ed. The Physicist's Conception of Nature, Reidel, Dordrecht, 1973, p.560.

M. Kunz, A Note about the Negentropy Principle, *MATCH*, **88**, 23, 3.