CHAPTER 5

# The methods of dating the ancient events offered by mathematical statistics

In our opinion, the main task of chronology analysis is to create independent statistical methods for the dating of ancient events. Only after that can one proceed to recreate chronology as a whole on the basis of results obtained. A single method – even as efficient as the astronomical one described above – is not enough for a profound study of the problem, because dating is an extremely sophisticated task that requires different methods of cross-verification. Advanced modern methodology of mathematical statistics makes it possible to offer a new approach to the dating of events described in ancient chronicles. This chapter describes new empirico-statistical methods developed by the author and his colleagues, as well as certain ways to apply them in chronological analysis.

This program was implemented in the following way.

1) New empirico-statistical methods of dating ancient events were developed, based on several statistical principles (models) proposed by the author in [884]-[886], [888]-[891], [895]-[905], [1129]-[1132], and [1135]. For a detailed account, see [MET1] and [MET2]. The primary principles, and models based thereupon, were laid out by the author in his report at the 3rd International Conference on Probability Theory and Mathematical Statistics, Vilnius, 1981 ([885]).

We proposed:

- The maxima correlation principle;
- The small distortions principle (for ruler dynasties);
- The frequency damping principle, the frequency duplication principle, and the geographic maps "improvement" principles.

The development of these methods was then related in a report made at the 4th International Conference on Probability Theory and Mathematical Statistics, Vilnius, 1985 ([901]) and the 1st International Congress of the Bernoulli Society for Mathematical Statistics and Probability Theory, 1986 ([1130]). Later on, new empirico-statistical models were proposed and verified by experiments in a series of works by V. V. Fedorov, A. T. Fomenko, V. V. Kalashnikov, G. V. Nosovskiy, and S. T. Rachev ([357], [590]-[613], [723], [1140] and [868]).

2) Those principles and models, as well as their efficiency, were verified by a sufficient amount of authentic material from medieval and contemporary history of the XVI-XX century, proving accuracy of the results obtained by these methods.

3) The same methods were applied to chronological material of ancient history normally dated to periods preceding the X-XIV century A.D. See [884], [886]-[888], [891], [895], [897], [898], [900], [903]

and [905]. Strange "repetitions" and "recurrences" were discovered in the Scaligerian version of the ancient and medieval history, the ones that we shall be referring to as "phantom duplicates".

4) All of these phantom duplicates were ordered into a system on the global chronological map outlined by the author in his articles [886], [888], [894], [896] and [905]. We do not absolutely consider the suggested methods to be universal ones, their applicability limits being clearly defined (see below). The only criterion for the correctness of results obtained is the conformity we discovered between the dates calculated by different methods, including the astronomical dating method described earlier.

5) On the basis of the global chronological map representing "the Scaligerian textbook of ancient history", we managed to restore a tentative origin of the Scaligerian version of the ancient and mediaeval chronology. We shall encapsulate some of those methods below.

## 1. THE LOCAL MAXIMA METHOD

### 1.1. The historical text volume function

The maxima correlation principle, and a method based thereupon, were proposed and developed by the author in [884], [885], [888] and [1129].

Let un assume that we discovered a historical text *X*, e.g., a previously unknown chronicle relating previously unknown events within a significant time interval, from year *A* to year *B*. Moreover, we may know nothing of the chronology in which these years were recorded. We shall hereinafter mark this time interval as (*A, B*). A typical situation: dates of events described in a chronicle are counted down from some event of local importance, such as the foundation of a town, accession of a ruler, etc. In such cases we would say that the chronicle dates the events in a *relative* chronology, which would allow us to distinguish these from the *absolute* dates in terms of B.C. or A.D. A natural question arises, namely: "How does one restore the *absolute* dates of events described in an antique document?" – for instance, the Julian date for the foundation of a town used to calculate the dates of the events?

Certainly, if we already know some of the events described from a dated chronicle, then we can "link" these events to the contemporary time scale. However, if such identification is impossible, the task of dating becomes more complicated. Moreover, the events described in the chronicle discovered may turn out to have already been known to us, though the appearance of their description is still beyond recognition because the chronicle is written in a different language, the chronicler uses completely different names, nicknames, geographic names, etc. Therefore, one might as well use a method of empirico-statistical nature, which makes it possible to sometimes date events on the basis of formal quantitative characteristics of the text under study.

Let us assume that a historical text *X* is broken up into fragments *X*(*t*), each describing a comparatively short time interval, for example, a year (or a decade) number *t*. There exist numerous examples of such texts – e.g., the *per annum* chronicles, or those describing events *year after year*, "per annum": diaries, many historical literary works, history textbooks and monographs. We shall be referring to the fragments *X*(*t*) as "chapters". They line up naturally in a chronological sequence according to the internal relative chronology of the chronicle in question. Many historical texts explicitly feature such "fragmentation into chapters", each describing a single year. Such are, for instance, many Russian chronicles ([671], [672]), as well as the famous ***Radzivillovskaya Letopis' (Povest vremenny'kh let)*** / *The Radzivil Chronicle* (*Story of Years of Time*) [715]. The famous Roman book ***Liber Pontificalis***, (T. Mommsen, ***Gestorum Pontificum Romanorum***, 1898) is of a similar nature.

Various characteristics of the information volume reported by chronicle *X* about year *t* can be measured as:

1) *vol X*(*t*) = number of pages in "chapter" *X*(*t*). Call this number the *volume* of "chapter" *X*(*t*). The volume can be zero if year *t* is not described in chronicle *X*, or missing. Instead of pages, one can count the number of lines, symbols, and so on. That neither affects the idea, nor the application of the method.

2) The total number of times year **t** is mentioned in chronicle *X*.

3) The number of names of all historical characters mentioned in "chapter" *X*(*t*).

4) The number of times a certain specific name (character) is mentioned in "chapter" $X(t)$.

5) The number of references to some other text in "chapter" $X(t)$.

The fund of quantitative characteristics like this is fairly large and important – each one, as we see, assigns a specific number to each year *t* described in the chronicle. In general, different numbers will correspond to different years; therefore, volumes of "chapter" $X(t)$ will largely be changing as the number (year) **t** changes. We shall call the succession of volumes $X(A), \ldots, X(B)$ *the volume function* of the per annum text *X*.

### 1.2. The maxima correlation principle

Thus, we assume a certain historical period from year *A* to year *B* in the history of one state *S* is described in a per annum chronicle *X* exhaustively enough, that is, chronicle *X* has already been, or can be, broken up into pieces – "chapters" $X(t)$, each describing one year *t*. We shall calculate the volume of each such piece – e. g., the number of words or symbols, pages, and so on – and then present the obtained numbers as a graph, with years *t* on the horizontal axis, and volumes of "chapters", or *vol* $X(t)$, on the vertical axis (fig. 5.1). The result shall be a graphic presentation of the volume function for this chronicle *X*.

A respective volume function graph for another per annum chronicle *Y*, describing the year-after-year "flow of events" of the same epoch $(A, B)$, will, as a matter of fact, look different (fig. 5.1). The point is that the personal interests of chroniclers *X* and *Y* play a major part in distribution of volumes – e.g., the information focus and per annum distribution in chronicle *X* on the history of art, and military chronicle *Y* will differ substantially. For example, chronicler *X* of a "defeated party" would describe the defeat of his army in a sparing and reserved manner – a few lines only. On the contrary, chronicler *Y* of a "victorious party" would render the story of the same battle in a great detail, enthusiastically, and eloquently, on several pages.

How vital are those differences? Or, are there characteristics of volume graphs that can only be defined by the time interval $(A, B)$, the history of a state *S*, and unambiguously characterize all, or almost all, chronicles describing this time interval and this state?

Years *t* in which the graph *peaks*, or reaches its *local maxima*, turn out to be a crucial characteristic of volume graph *vol* $X(t)$. The fact that the graph peaks at a given point *t* means that this year is described in the chronicle *in greater detail* – e.g., on more pages than the adjacent ones. Hence, the peaks of the graph, or its local maxima, indicate years a chronicler described in detail on the time interval $(A, B)$. In different chronicles *X* and *Y*, absolutely different years can be "described in detail".

What is the reason for such an uneven description of different years? A possible explanation: a chronicler described an "ancient year" in greater detail because more information on that "ancient year" was available – such as a bulk of old documents larger than that for adjacent years.

The course of our further argumentations is as follows.

1) We shall formulate a *theoretical model*, or *statistical hypothesis,* that will allow us to predict what years from the time interval $(A, B)$ will be reported in detail by a later chronicler, not a contemporary of the ancient events he describes.

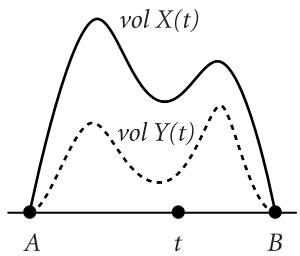


Fig. 5.1. The volume graphs for the two chronicles, *X* and *Y*, relating the events of the same historical epoch.

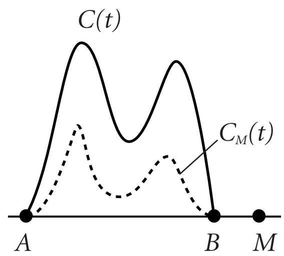


Fig. 5.2. The graph of the "primary information fund" $C(t)$, and the graph of the "remaining information fund" (the texts that survived until the epoch *M*) peak almost simultaneously.

2) Then, we shall mathematically formulate that statistical model, hypothesis.

3) We shall test its correctness on the fairly extensive reliable historical material of the XVI-XX century.

4) Upon discovering experimental proof for the theoretical model, we shall offer a method for dating of ancient events.

Let $C(t)$ be the volume of all texts written about the year $t$ by its contemporaries (fig. 5.2). As done above, we shall construct a numerical volume graph of the time interval $(A, B)$. We certainly are *not aware* today of the precise appearance of this graph $C(t)$. The fact is, the original texts written by contemporaries of the events of the year $t$ became gradually lost over the course of time, and only a certain part has survived. The graph $C(t)$ can be called *the primary information fund graph*. Let us assume that contemporaries described certain years of the epoch $(A, B)$ in greater detail, i.e., recorded an especially large amount of information about these years. We are not discussing reasons for this "original unevenness" as being fairly irrelevant to us now. In the sense of the volume graph $C(t)$ such years – "described by contemporaries in detail" – will be noted for peaks of the graph on these precise years.

A question: 'How does the loss and oblivion of information occur, which in the course of time can distort the graph $C(t)$ and decrease its altitude?' Let us relate *the information loss model.*

Although the altitude of the graph $C(t)$ decreases over the course of time, nonetheless, *from the years in which especially many texts were created by contemporaries, more will survive.*

To restate the model, it is useful to fix a certain moment in time $M$ to the right of point $B$ on fig. 5.2, and construct a graph $C_M(t)$ showing the volume of texts that "survived" until the moment $M$ and describe the events of the year $t$ in the epoch $(A, B)$.

In other words, the number $C_M(t)$ shows the volume of the original ancient texts from the year $t$ that survived until the "fund observation moment" in the year $M$. The graph $C_M(t)$ can be referred to as the graph of the "residual information fund" that survived from the epoch $(A, B)$ until the year $M$. Now our model may be restated in the following way.

*Peaks on both the residual fund volume graph $C_M(t)$ and the original primary information fund graph $C(t)$ must occur approximately in the same years of the time interval* $(A, B)$.

The model is obviously quite difficult to test as it is, because the primary information fund graph $C(t)$ is unknown today. But it is still possible to verify one of the consequences of the theoretical model (hypothesis).

Since later chroniclers $X$ and $Y$ describing the same historical period $(A, B)$ and the "flow of events" are no longer contemporaries of those ancient events, they have to rely on more or less the same set of texts available in their time. Thus, they would describe in greater detail "on the average" the years from which more texts survived, and in less detail the years of which little information was available. In other words, the chroniclers should increase the detail level of their rendition for the years that yielded more old texts.

In the language of volume graphs, the model looks as follows. If chronicler $X$ lives in epoch $M$, then he will rely on the residual fund $C_M(t)$. If the other chronicler $Y$ lives in epoch $N$ that is generally different from epoch $M$, then he relies on the available information fund $C_N(t)$. See fig. 5.3.

It is quite natural to expect the chroniclers $X$ and $Y$ to work "on the average" in good faith, therefore describing in greater detail those years of the ancient (for them) epoch $(A, B)$ from which more information and old texts are available.

In other words, peaks on the volume graph *vol*

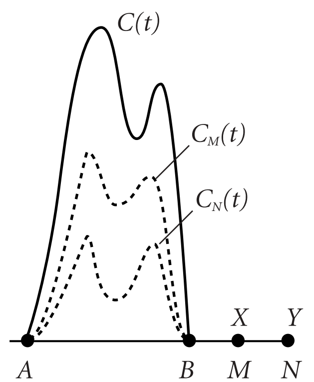


Fig. 5.3. The graphs of the remaining information funds peak around the same period of time as the graph of the primary graph, *C(t)*. The chronicle volume functions *X* and *Y* peak in roughly the same points as the volume graphs of the information that survived until their epoch.

$X(t)$ and the graph $C_M(t)$ will occur in the same years. In their turn, peaks on the graph *vol* $Y(t)$ and the graph $C_N(t)$ will occur approximately over the same years, fig. 5.3.

But the peaks of the residual fund graph $C_M(t)$ *are close* to those of the original, primary graph $C(t)$. Likewise, the splash points of the residual fund graph $C_N(t)$ are close to the splash points of the primary graph $C(t)$. Hence, splashes on the volume graphs for chronicles *X* and *Y*, or the graphs *vol* $X(t)$ and *vol* $Y(t)$, must occur *approximately at the same time*, in "the same" points of the time axis. In other words, their local maxima points must distinctly correlate, fig. 5.1.

In doing so, the *amplitudes* of graphs *vol* $X(t)$ and *vol* $Y(t)$ can certainly differ substantially, fig. 5.4, which does not appear to affect the arguments stated.

The final formula for *the maxima correlation principle* is as follows, preceding the reasoning regarded as the primary consideration.

**The maxima correlation principle**

a) If two chronicles (texts) *X* and *Y* are *a priori dependent*, i.e., describe the same "flow of events" of historical period (*A, B*) of the same state **S**, then *local maxima* (*splashes*) on volume graphs of the chronicles *X* and *Y must occur simultaneously* on the time interval (*A, B*). In other words, the years "described in detail in chronicle *X*" and the years "described in detail in chronicle *Y*" must be close or coincident, fig. 5.4.

b) On the contrary, if chronicles *X* and *Y* are *a priori independent*, i.e., describe either different historical periods (*A, B*) and (*C, D*), or different "flows of events" in different states, then the volume graphs for chronicles *X* and *Y* reach their local maxima *in different points*. In other words, the peaks of the graphs *vol* $X(t)$ and *vol* $Y(t)$ should not correlate, q.v. in fig. 5.5. In doing so, we are supposed to have provisionally combined (identified) segments (*A, B*) and (*C, D*) of the same length before comparing the two graphs.

We shall conditionally call all other pairs of texts, i.e., neither *a priori* dependent nor *a priori* independent, *neutral*, and make no assertions regarding them.

This principle is confirmed if, for the majority of pairs of actual and large enough *dependent* chronicles *X* and *Y*, i. e., those describing the same "flow of events", the peaks on volume graphs for *X* and *Y* do actually occur approximately at the same time, in the same years, while *the magnitude of these peaks can be substantially different.*

On the contrary, for actual *independent* chronicles, the peaks should not correlate in any way. For specific dependent chronicles, the synchronism of volume graph splashes can only be approximate.

## 1.3. Statistical model

The rough idea is as follows. For quantitative evaluation of peak proximity we shall calculate the number $f(X, Y)$ – the sum of numbers $f[k]$ squared, where $f[k]$ is the distance in years between the peak "$k$" of volume graph *X* and the peak "$k$" of volume graph *Y*. If the peaks on both graphs should occur simultaneously, then the peaking moments with identical numbers will coincide, and all numbers $f[k]$ equal zero. Upon reviewing a fairly large fund of authentic

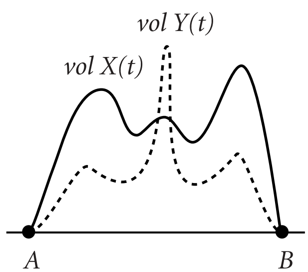


Fig. 5.4. Volume graphs of the dependent chronicles *X* and *Y* which relate the events of roughly the same epoch, peak almost simultaneously. However, the peaks may significantly differ from each other in size.

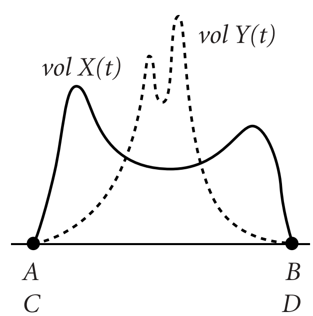


Fig. 5.5. Volume graphs of independent chronicles *X* and *Y* relating to completely different epochs, peak in different points (after the superposition of time intervals (*A, B*) and (*C, D*)).

texts $H$ and calculating the number $f(X, H)$ for each of them, we then choose only those texts $H$ for which this number does not exceed the number $f(X, Y)$. Upon calculating the portion of such texts in the whole fund of texts $H$, we obtain a coefficient that, according to the hypothesis of random vector $H$ distribution, can be interpreted as probability $p(X, Y)$ ([904], [908], [1137] and [884]). If the coefficient $p(X, Y)$ is small, then the chronicles $X$ and $Y$ are dependent, or describe approximately the same "flow of events". If the coefficient is large, then the chronicles $X$ and $Y$ are independent, that is, they report of different "flows of events".

Now we pass on to a more detailed description of the statistical model. Doubtlessly, the peaks on real volume graphs can be only simultaneous approximately. To estimate just how simultaneous the peaks on both graphs are, the mathematical methods of statistics allow us to define a certain number $p(X, Y)$ that measures the mismatch of the years described in detail in the chronicle $X$, and the years described in detail in the chronicle $Y$. It turns out that if the proximity of peaks on both graphs is regarded as random, the number $p(X, Y)$ can be seen as the probability coefficient of this event (which, however, is not at all that key for the efficiency of the method). The smaller this number is, the greater the coincidence of the years described in detail in $X$ with those described in detail in $Y$. We shall formulate a mathematical definition of the coefficient $p(X, Y)$.

Let us examine the time interval $(A, B)$ and the volume graph $vol\ X(t)$ that reaches local maxima in certain points $m_1, \dots, m_{n-1}$. For the purpose of simplicity, we consider each local maximum (peak) to culminate exactly in one point. In general, these points, or years, $m_i$ break up the time interval $(A, B)$ into a number of segments of different length, q.v. in fig. 5.6. Measuring the length of these segments in years, that is, measuring the distance between the points of adjoining local maxima $m_i$ and $m_{i+1}$, we obtain a sequence of integers $a(X)=(x_1, \dots, x_n)$. This means that the number $x_1$ is the distance from the point $A$ to the first local maximum, the number $x_2$ is the distance from the first local maximum to the second one, and so on, the number $x_n$ being the distance from the last local maximum $m_{n-1}$ to the point $B$.

This sequence can be represented by the vector $a(X)$ in Euclidean space $R^n$ of dimension $n$. For instance, in case of two local maxima, i.e., if $n = 3$, we have an integer-valued vector $a(X) = (x_1, x_2, x_3)$ in three-dimensional space. Let the vector $a(X) = (x_1, \dots, x_n)$ be called *the local maxima vector* for the chronicle $X$.

For the other chronicle $Y$ we have, generally speaking, a different vector $a(Y)=(y_1, \dots, y_m)$. We assume that chronicle $Y$ describes events of the time interval $(C, D)$, the length of which is equal to that of the time interval $(A, B)$, i. e., $B - A = D - C$. To compare volume graphs of the chronicles $X$ and $Y$, we shall combine the two previous time segments $(A, B)$ and $(C, D)$ of the same length, and superpose them over each other. Naturally, the number of local maxima of the graphs $vol\ X(t)$ and $vol\ Y(t)$ can be different. However, without rigidly restricting commonness, it is possible

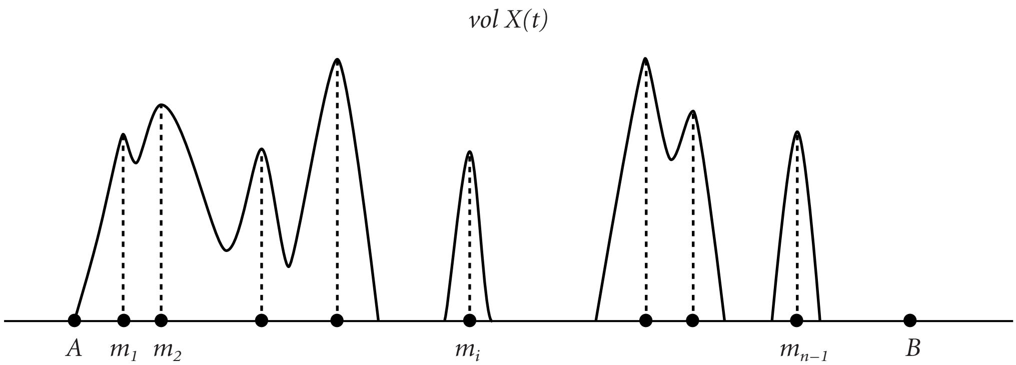


Fig. 5.6. Chronicle volume graph peaks divide the time interval $(A, B)$ into smaller intervals.

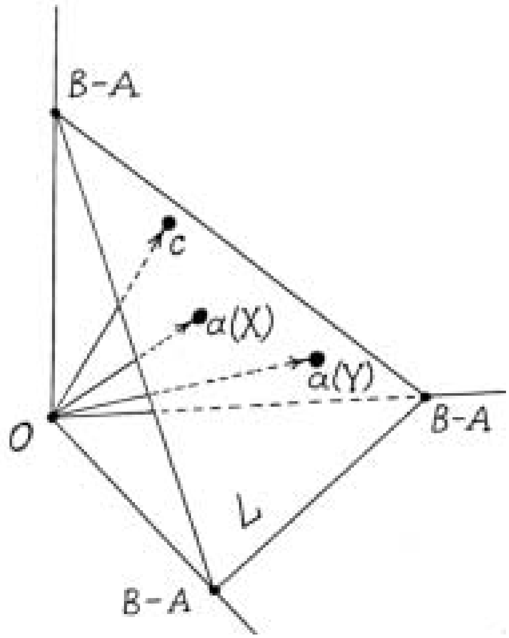


Fig. 5.7. Local maxima vectors for $a(X)$ and $a(Y)$ of the two chronicles compared ($X$ and $Y$) can be conventionalized as two vectors in Euclidean space.

to say that the number of maxima is identical, and thus the vectors $a(X)$ and $a(Y)$ of two comparable chronicles $X$ and $Y$ have the same number of coordinates. Indeed, if the maxima number of two comparable graphs is different, then it is possible to proceed as follows. We shall consider certain maxima *multiple*, i.e., believe several local maxima to have merged at this point. In doing so, lengths of relevant segments corresponding to these multiple maxima can be considered to equal zero. Stipulating this, we can apparently equalize the number of local maxima on the volume graphs of the chronicles $X$ and $Y$. Of course, such an operation – the introduction of multiple maxima – is not unique. We shall settle on a certain variant for the introduction of multiple maxima so far. Later on, we shall get rid of this ambiguity by minimizing all necessary proximity coefficients along all possible variations of multiple maxima introduction. We shall note that the multiple maxima introduction means the appearance of void components, i.e., segments of zero length, in certain places of vector $a(X)$.

Thus, comparing chronicles $X$ and $Y$, we can assume that both vectors $a(X)=(x_1, \dots ,x_n)$ and $a(Y) = (y_1, \dots ,y_n)$ have the same number of coordinates and thus are situated in the same Euclidean space $R^n$. We shall note that the sum of the coordinates of each vector is the same, equalling $B - A = D - C$, or the length of the time interval $(A, B)$. Thus,

$$x_1 + \dots + x_n = y_1 + \dots + y_n = B - A.$$

Now we shall consider the set of all integer-valued vectors $c = (c_1, \dots ,c_n)$, the coordinates of which are non-negative with the sum $c_1 + \dots + c_n$ equalling the same value, namely, $B - A$, or the length of the time interval $(A, B)$. We shall denote the set of all those vectors with the letter $S$. Geometrically, those vectors can be presented as originating from the beginning of coordinates, or from the point $0$ in $R^n$. Let us consider the ends of all such vectors $c = (c_1, \dots , c_n)$, all of them situated on a "multi-dimensional simplex" $L$ defined in the space $R^n$ by one equation

$$c_1 + \dots + c_n = B - A$$

where all coordinates $c_1, \dots , c_n$ are real non-negative numbers. Set $S$ is presented geometrically as a set of "integer points" on simplex $L$, or a set of all points with integer-valued coordinates, from $L$.

It is clear that the ends of the local maxima vectors $a(X)$ and $a(Y)$ for chronicles $X$ and $Y$ belong to the set $S$, fig.5.7.

Now we shall fix the vector $a(X)=(x_1, \dots , x_n)$ and examine all vectors $c = (c_1, \dots , c_n)$ with real coordinates belonging to the simplex $L$ and such as to comply with an additional correlation,

$$(c_1 - x_1)^2 + \dots + (c_n - x_n)^2 \le (y_1 - x_1)^2 + \dots + (y_n - x_n)^2.$$

We shall denote the set of all such vectors $c = (c_1, \dots , c_n)$ as $K$. These vectors are mathematically described as being remote from the fixed vector $a(X)$ on a distance not exceeding the distance $r(X, Y)$ from vector $a(X)$ to vector $a(Y)$. Speaking of the distance between the vectors, we mean the distance between their ends. We shall recall that the value

$$(y_1 - x_1)^2 + \dots + (y_n - x_n)^2$$

is equal to the squared distance $r(X, Y)$ between the vectors $a(X)$ and $a(Y)$. Therefore, set $K$ is a part of simplex $L$, fitting the "$n$-dimensional" ball with the radius of $r(X, Y)$ and the centre in the point $a(X)$.

Let us now calculate how many "integer-valued vectors" set $K$ and set $L$ have each. We shall denote the values obtained as $m(K)$ and $m(L)$, respectively. As a "preliminary coefficient" $p'(X, Y)$ we shall use a ratio of these two values, i. e.,

$$p'(X, Y) = m(K) / m(L),$$

that is,

$$p'(X, Y) = \frac{\textit{number of "integer points" in set K}}{\textit{number of "integer points" in set L}}\ .$$

Since set $K$ is only a part of set $L$, the number $p'(X, Y)$ is enclosed in the segment [0, 1].

If vectors $a(X)$ and $a(Y)$ coincide, then $p'(X,Y) = 0$. If, on the contrary, the vectors are far away from each other, then the value $p'(X, Y)$ is close to, and can even equal one.

We shall note a useful, though not mandatory hereinafter, interpretation of the number $p'(X, Y)$. Let us assume that the vector $c = (c_1, \dots c_n)$ randomly runs across all vectors from the set $S$, and in doing so, it can appear in any point of this set, with an equal probability. In such cases, the random vector $c = (c_1, \dots ,c_n)$ is said to be *uniformly* distributed over the set $S$, i.e., among the set of the "integer points" $(n–1)$-dimensional simplex $L$. Then, the value $p'(X, Y)$ we defined allows for a probability interpretation, as being simply equal to the probability of a random event, when the distance between random vector $c = (c_1, \dots, c_n)$ and the fixed vector $a(X)$ does not exceed the distance between vectors $a(X)$ and $a(Y)$. The smaller this probability, the less accidental is the proximity of vectors $a(X)$ and $a(Y)$. In other words, their proximity in this case indicates a certain *dependence* between them. And the smaller the value $p'(X, Y)$, the stronger this dependence.

The uniformity of distribution of the random vector $c = (c_1, \dots , c_n)$ on simplex $L$, or rather on set $S$ of its "integer points", may be justified by the fact that this vector depicts the distance between adjacent local maxima of the volume function of "chapters" of historical chronicles or other similar texts describing the given time interval $(A, B)$. In considering various chronicles relating the history of different states in different historical epochs, it is quite natural to assume that a local multiple maxima may appear "with equal probability" in any point of the time interval $(A, B)$.

The described construction was completed in assumption that we fixed a certain variant of multiple maxima introduction for volume graphs of chronicles. Variants like that exist in a great number, no doubt. We shall consider all such variants and for each of them, calculate a separate value $p'(X, Y)$, upon which we shall take the least of all obtained values and denote it as $p''(X, Y)$, i.e., minimize the coefficient $p'(X, Y)$ through all possible methods of local multiple maxima introduction of graphs $vol\ X(t)$ and $vol\ Y(t)$.

We shall eventually recall that, upon calculating the coefficient $p''(X, Y)$, the chronicle $X$ and $Y$ appeared to be in unequal positions. The fact is that we were considering an "$n$-dimensional ball" of radius $r(X, Y)$ with its centre in point $a(X)$. In order to eliminate the apparent discrepancy between chronicles $X$ and $Y$, we shall simply swap them and repeat the construction described above, now taking the point $a(Y)$ as the centre of the "$n$-dimensional ball". As a result, a certain value will be obtained, which we denote as $p''(Y, X)$. In the capacity of "symmetrical coefficient" $p(X, Y)$, we shall take a simple average of the values $p''(X, Y)$ and $p''(Y, X)$, i. e.,

$$p(X, Y) = \frac{p''(X, Y) + p''(Y, X)}{2}.$$

For the sake of clarity, we shall explain the meaning of the preliminary coefficient $p'(X, Y)$ on an example of a volume graph with only two local maxima. In this case, both vectors,

$$a(X) = (x_1, x_2, x_3) \text{ and } a(Y) = (y_1, y_2, y_3),$$

are vectors in 3-dimensional Euclidean space, their ends lying on a two-dimensional equilateral triangle $L$ that truncates the same number $B – A$ from the coordinate axes in the space $R^3$. See fig. 5.8. If we mark the distance between points $a(X)$ and $a(Y)$ as $|a(X) – a(Y)|$, then set $K$ is the intersection of the triangle $L$ with the three-dimensional ball, the centre of which is in the point $a(X)$ and the radius equal to $|a(X) – a(Y)|$. After that, we need to calculate the number of "integer points", i. e., points with integer-valued coordinates, in set $K$ and triangle $L$. Taking the ratio of the numbers obtained, we arrive at the coefficient $p'(X, Y)$.

For specific calculations, it is quite convenient to

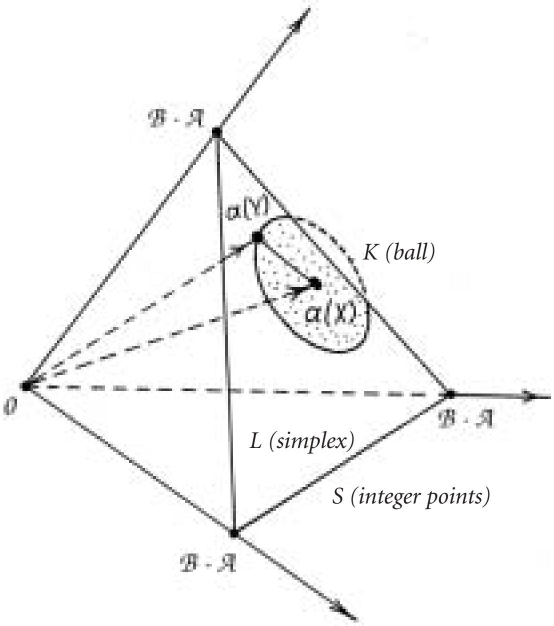


Fig. 5.8. Vectors $a(X)$ and $a(Y)$ define the "ball", part of which becomes included in simplex $L$.

use an approximate method of calculating the coefficient $p(X, Y)$. The fact is that computation of the number of integer points in set $K$ is quite difficult, but appears to be possible to simplify by proceeding from "discrete model" to the "continuous model". It is well known that if $(n-1)$-dimensional set $K$ in $(n-1)$-dimensional simplex $L$ is rather large, then the number of integer points in $K$ is approximately equal to $(n-1)$-dimensional volume of set $K$. Therefore, from the very beginning it is possible to use the ratio of $(n-1)$-dimensional volume $K$ to $(n-1)$-dimensional volume $L$ as the preliminary coefficient $p'(X, Y)$, i. e.,

$$p'(X, Y) = \frac{(n\text{-}1)\text{-}dimensional\ volume\ K}{(n\text{-}1)\text{-}dimensional\ volume\ L}.$$

For instance, in case of two local maxima, ratio

$$\frac{area\ of\ set\ K}{area\ of\ traingle\ L}$$

should be taken as the coefficient $p'(X, Y)$.

When the value of $B - A$ is small, the "discrete coefficient" and the "continuous coefficient" are certainly different. But we in our researches deal with several decades' and even several hundred years' time intervals $B - A$, therefore for our purposes we can, without making a great mistake, use the "continuous model" $p'(X, Y)$ in all confidence. Precise mathematical formulae for the calculation of the "continuous coefficient" $p'(X, Y)$ and for its lower and upper boundaries are presented in the work [884], page 107.

Let us present one more specification of the statistical model described above. When working with specific volume graphs of historical texts, one should "smoothen" those graphs in order to eliminate minute random peaks. We have made our graph even by "proximity averaging", that is, by replacing the value of the volume function at each point $t$ by a simple average of three values of the function, namely, at the points $t-1$, $t$, $t+1$. In the capacity of the "final coefficient" $p(X, Y)$, its value as calculated for such "smoothed graphs" should be taken.

The maxima correlation principle stated above will be confirmed if, for the majority of pairs of *a priori* dependent texts $X$ and $Y$, the coefficient $p(X, Y)$ turns out to be small, and for the majority of the *a priori* independent texts it turns out to be, on the contrary, large.

## 1.4. Experimental test of the maxima correlation principle. Examples of dependent and independent historical texts

In 1978-1985 we conducted the first extensive experiment in the computation of numbers $p(X, Y)$ for several dozen pairs of specific historical texts: chronicles, annals, and so on. See details in [904], [908], [1137] and [884].

The coefficient $p(X, Y)$ turned out to distinguish between *a priori dependent* and *a priori independent* pairs of historical texts well enough. It was discovered that for all examined pairs of actual chronicles $X$, $Y$ describing *obviously different* events (different historical epochs or different states), i.e., for all *independent* texts, the number $p(X, Y)$ fluctuates from *1* to *1/100*, where the number of local maxima ranges from *10* to *15*. On the contrary, when historical chronicles $X$ and $Y$ were *a priori dependent*, that is, described the same events, the number $p(X, Y)$ for the same number of maxima doesn't exceed $10^{-8}$.

Thus, the spread between the coefficient values for dependent and independent texts is approximately 5-6 orders of magnitude. We shall emphasize the fact that it is not the absolute value of the obtained coefficients that is of importance here, but the fact that the "zone of coefficients for *a priori* dependent texts" is separated by *several orders of magnitude* from the "zone of coefficients for *a priori* independent texts". Let us present several examples. Exact values of volume functions for especially interesting chronicles are presented in the Appendix at the end of the book, in order to avoid the overload of current narration.

**EXAMPLE 1.**

Volume graphs for two *a priori* dependent historical texts are presented in fig. 5.9, fig. 5.10 and fig. 5.11.

Namely, in the capacity of text *X* we took a historical monograph ***Essays on the History of Ancient Rome*** by V. S. Sergeyev, a contemporary author. – Vol.1-2, OGIZ, Moscow, 1938.

In the capacity of text *Y* we took the "antique" source, ***The History of Rome*** by Titus Livy. – Vol.1-6, Moscow, 1897-1899.

According to the Scaligerian chronology, these texts describe events in the time interval allegedly of 757-287 B.C. Thus, here *A* = 757 B.C., *B* = 287 B.C. Both texts describe the same historical epoch, approximately the same events. *Primary* peaks of the volume graphs obviously occur at virtually the same time. For quantitative comparison of functions, it is necessary to smoothen "ripples", i. e., secondary peaks that can be superposed over the main, initial oscillations on the

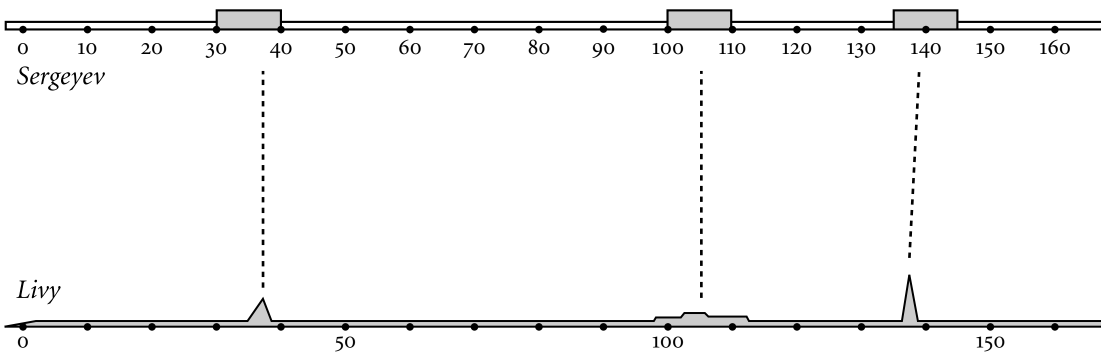


Fig. 5.9. Volume functions of the chronicle of the "ancient" Titus Livy and a modern textbook by Sergeyev. One sees a very explicit correlation. Part one.

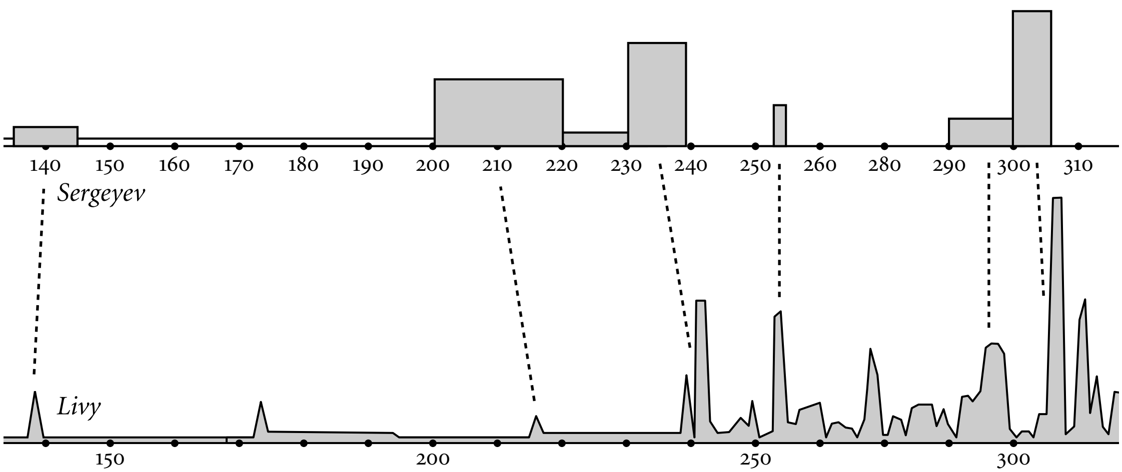


Fig. 5.10. Volume functions of the chronicle of the "ancient" Titus Livy and a modern textbook by Sergeyev. Part two.

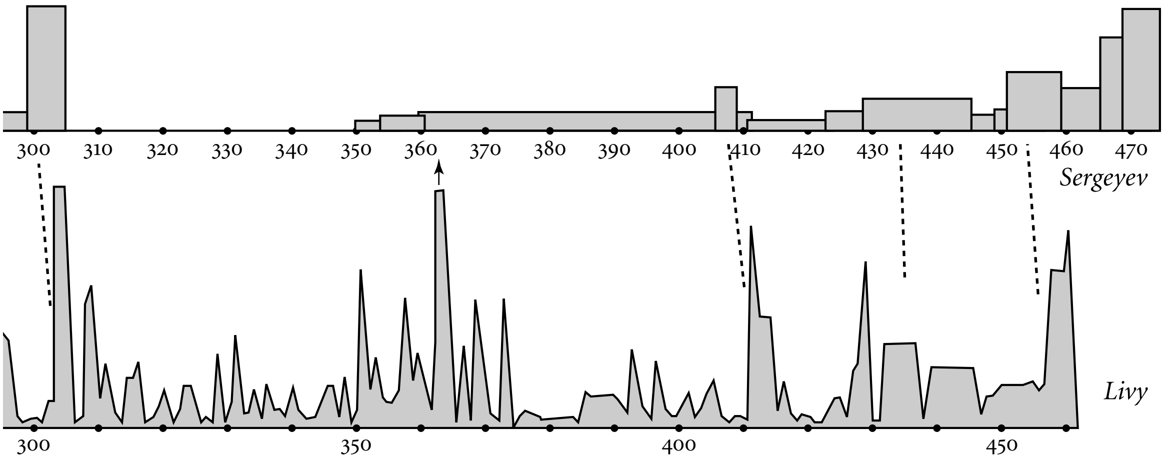


Fig. 5.11. Volume functions of the chronicle of the "ancient" Titus Livy and a modern textbook by Sergeyev. Part three.

graph. When computing the coefficient $p(X, Y)$ we have smoothed these graphs to emphasize only their *main* local maxima, not exceeding 15 in number. It turned out that $p(X, Y) = 2 \times 10^{-12}$. The small value of the coefficient indicates *dependence* between the texts compared, which comes as no surprise in this particular case. As we have already noted, both texts describe the same historical time interval of the "ancient" Rome. The small value of the coefficient $p(X, Y)$ proves the fact that if we consider the observed proximity of the splash points on both graphs as an random event, then its probability is extremely small. As we can see, the contemporary author V. S. Sergeyev reproduced the "ancient" original in his book quite accurately. He certainly supplemented it with his own considerations and commentaries, which, however, turn out to have no influence on the character of dependence between those texts.

Now, we shall use the book by V. S. Sergeyev as the "chronicle" $X'$ once again, and as the "chronicle" $Y'$, the same book, but with the order of the years in the text replaced by the opposite one – in other words, as if we have read the book by Sergeyev "back to front". In this case, $p(X',Y')$ turns out to equal 1/3, a value substantially closer to 1 than the previous one and demonstrating the independence of compared texts – hardly surprising, since the operation of "inverting the chronicle" yields two *a priori* independent texts.

**Example 2.**

We shall regard the following *a priori* dependent historical texts as examples – the two Russian chronicles:

*X* – ***Nikiforovskaya letopis'*** (The Nikiforov Chronicle) [672],

*Y* – ***Suprasl'skaya letopis'*** (The Suprasl' Chronicle) [672].

Both chronicles cover the time interval of allegedly 850-1256 A.D.

Their volume graphs are presented at fig.5.12. Both volume graphs of "chapters" allegedly of 850-1255 A.D. have 31 peaks occurring virtually simultaneously, in the same years. The calculation yields $p(X, Y) = 10^{-24}$, a fairly small value; therefore, dependence between those texts is confirmed. In Chron1, Appendix 5.1, we present precise numerical data for the volume functions of these chronicles.

**Example 3.**

We now shall consider two other Russian chronicles:

*X* – ***Kholmogorskaya letopis'*** (The Kholmogory Chronicle) [672],

*Y* – ***Povest' vremennykh let*** (Story of Years of Time).

Both chronicles cover the time interval of allegedly 850-1000 A.D. Volume graphs of the chronicles reach their local maxima *virtually simultaneously* as well, which is again not by accident but in the order of things – otherwise, the sole chance out of $10^{15}$ would

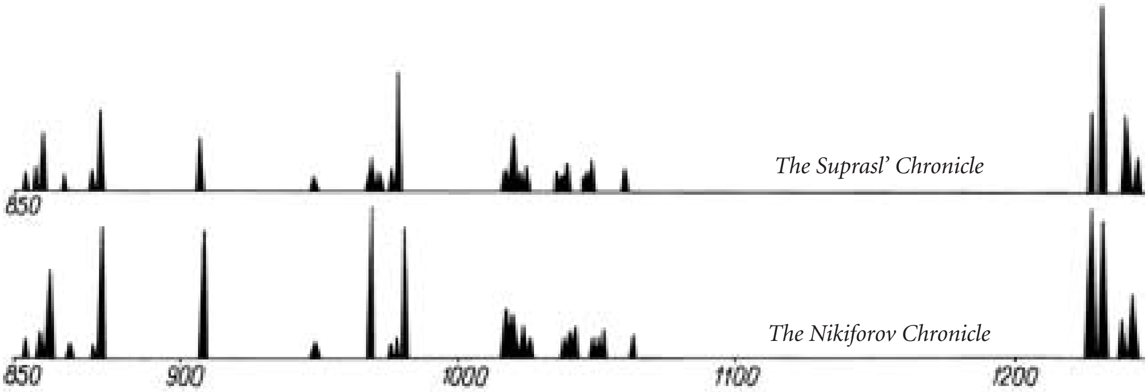


Fig. 5.12. Volume graphs for dependent chronicles: the Suprasl'skaya and the Nikiforskaya. The graph peaks are almost simultaneous.

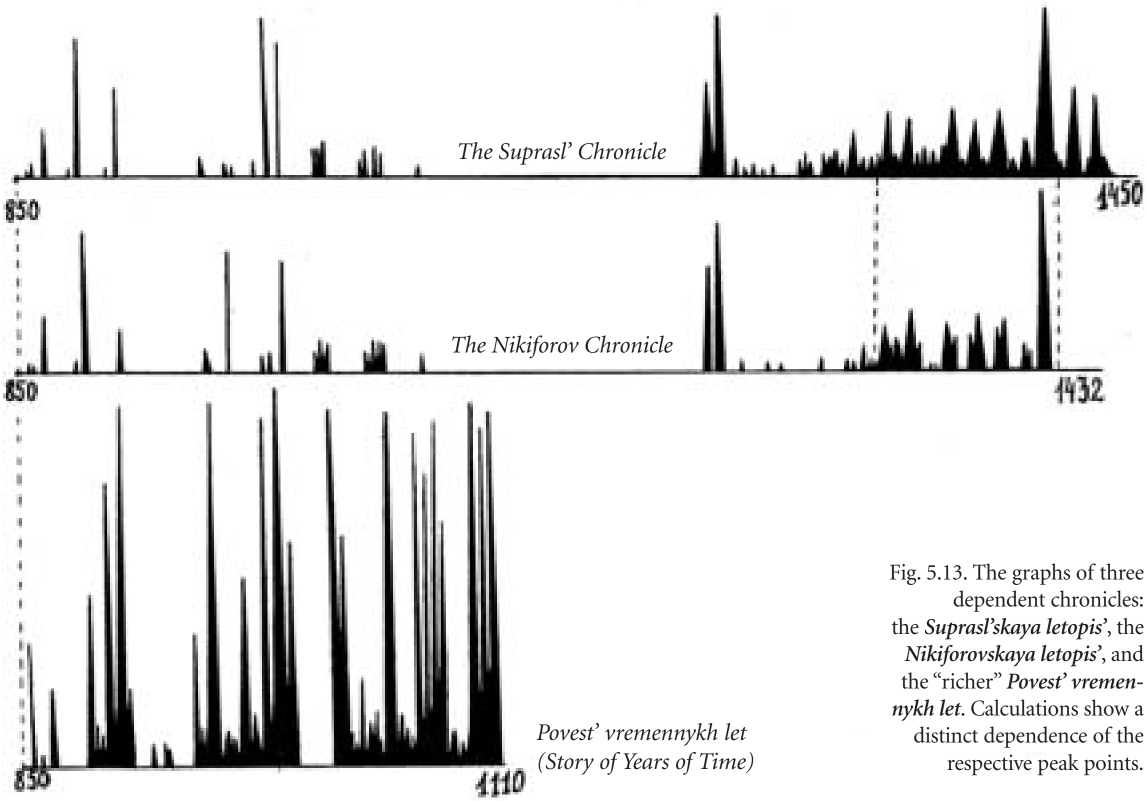


Fig. 5.13. The graphs of three dependent chronicles: the ***Suprasl'skaya letopis'***, the ***Nikiforovskaya letopis'***, and the "richer" ***Povest' vremennykh let***. Calculations show a distinct dependence of the respective peak points.

have been realized. Here, $p(X, Y) = 10^{-15}$. These two chronicles are dependent in the stated time interval. Fig. 5.13 simultaneously presents three volume graphs – for ***Suprasl'skaya letopis'***, ***Nikiforovskaya letopis'***, and ***Povest' vremennykh let***, the latter chronicle being "richer", therefore its graph has more local maxima, and its dependence is not so obvious. Nevertheless, an explicit dependence between those three graphs is as well revealed after smoothing. We shall describe comparison between "rich" and "poor" chronicles in the next chapters. The distribution of volumes of the mentioned chronicles is given in CHRON1, Appendix 5.1.

**Example 4.**

An example from the mediaeval Roman history.

*X* – ***The History of the City of Rome in the Middle Ages***, a fundamental monograph by F. Gregorovius, a German historian, Vols. 1-5 ([196]). This book was written in the XIX century on the basis of a great number of mediaeval secular and ecclesiastic documents.

*Y* – ***Liber Pontificalis*** (T. Mommsen, ***Gestorum Pontificum Romanorum***, 1898). This "Book of Pontiffs", the list and biography of the mediaeval Roman Popes, was restored by Theodor Mommsen, a German historian of the XIX century, from mediaeval Roman texts. Here, $p(X,Y) = 10^{-10}$, which demonstrates an obvious dependence between these two texts. To assume such proximity is accidental, the sole chance out of 10 billion would have been realized.

And so on. The several dozen examples of historical texts we processed, – *a priori dependent* as well as *a priori independent*, – confirmed our theoretical model. Thus, we managed to reveal regularities that allow us to statistically characterize *dependent* historical texts, or those covering the same time interval and the same "flow of events" in the history of the same region or the same state. In the meantime, experiments have demonstrated the following: if two historical texts *X* and *Y* are, on the contrary, *independent*, that is, describe obviously different historical epochs, or different regions, or essentially different "flows of events", then the peaks on volume graphs *vol* $X(t)$ and *vol* $Y(t)$ occur in substantially different years. In the latter case, a typical value of coefficient $p(X, Y)$, the local maxima varying from 10 to 15, fluctuates from 1 to 1/100. Here is a typical example.

**Example 5.**

We now return to the "ancient" history of Rome. In the capacity of compared texts *X* and *Y*, we have taken two other fragments from the book ***Essays on the History of Ancient Rome*** by V. S. Sergeyev ([767]). The first fragment covers the alleged years 520-380 B.C., the second one – the alleged years 380-240 B.C. These periods are considered independent. The computation of the coefficient $p(X, Y)$ yields 1/5, a striking value, different from typical values – $10^{-12}$ – $10^{-6}$ – for *a priori* dependent texts with a similar value of local maxima by several orders of magnitude. Thus, these two texts, "two halves" of the book by V. S. Sergeyev, are truly independent.

Above, we have used a numerical characteristic of volume for the "chapter". However, as our research has demonstrated, a similar statistical regularity becomes apparent for fairly large historical texts when other numerical characteristics are used – for instance, the number of names in each "chapter", the number of references to other chronicles, etc.

In our computational experiment we compared:

a) ancient texts with ancient texts;

b) ancient texts with contemporary texts;

c) contemporary texts with contemporary texts.

As we have already mentioned, other numerical characteristics of texts were analyzed along with volume graphs of "chapters". For instance, graphs for number of names mentioned, numbers of a specific year's mentions in the text, the frequency of references to some other fixed text, and so on ([904], [908], [1137] and [884]). The same *maxima correlation principle* turns out to be true for all of these characteristics – namely, the peaks on graphs for dependent texts occur virtually simultaneously, and as for independent texts, their peaks do not correlate at all.

We shall formulate one more consequence of our basic model, the statistical hypothesis.

If two historical texts are *a priori dependent*, that is, if they describe the same "flow of events" on the same time interval in the history of the same state, then the peaks on corresponding graphs for any pair of numerical characteristics stated above occur approximately in the same years. In other words, if a year is recorded by both chronicles in more detail than the adjoining ones, then the number of mentions of this year, as well as the number of names of characters mentioned in that year, and so on, will increase (locally) in both chronicles. The situation for *a priori independent* texts is directly opposite – no correlation between the stated numerical characteristics is due.

The "secondary maxima correlation principle" proved to be correct when tested on specific, *a priori* dependent, historical texts ([884], pp.110-111).

## 1.5. Method of dating the historical events

Since our theoretical model is supported by the results of experiments, we can now propose a new method of dating the ancient events, – not a universal one, though, – and describe the main idea thereof.

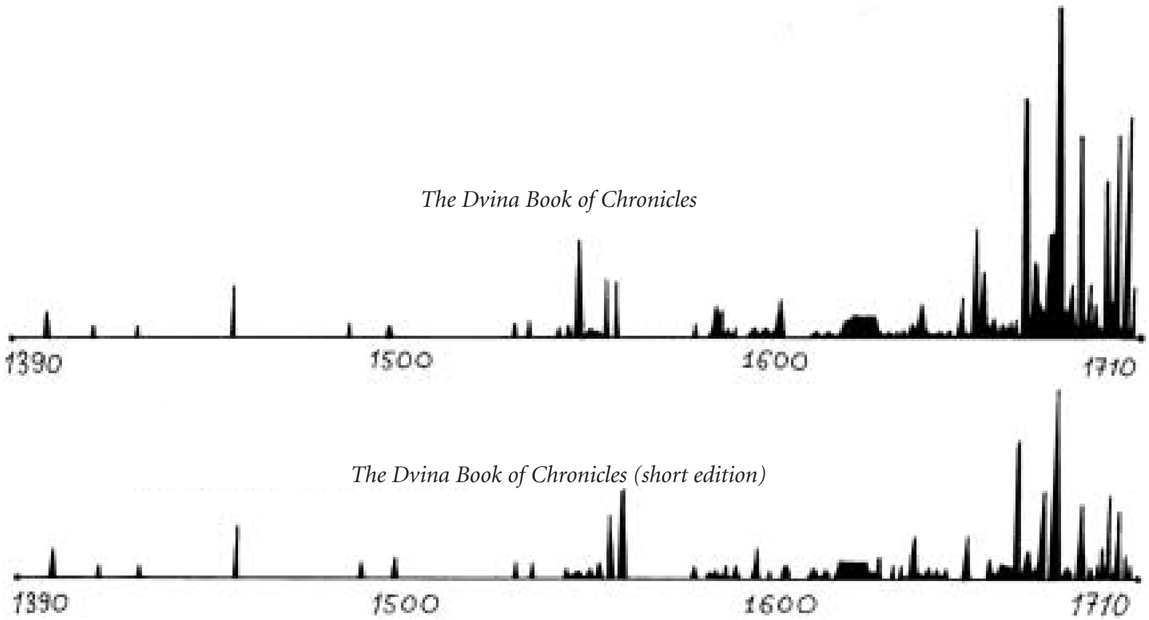


Fig. 5.14. Volume graphs for dependent chronicles: that of the ***Dvina Book of Chronicles***, and its shorter edition. Both graphs peak practically simultaneously.

Let $Y$ be a historical text covering an unknown "flow of events", its absolute dates being lost. Let years $t$ be counted in the text from some event of a local importance, for instance, the foundation of a town, the coronation of a king, whose absolute dates remain unknown to us. We shall calculate the volume graph of "chapters" for text $Y$ and compare it with the volume graphs of other texts, for which we know the absolute dating of events described. If text $X$ is revealed among those texts, and its number $p(X, Y)$ is small – i. e., has the same order of magnitude as pairs of dependent texts (for instance, does not exceed $10^{-8}$ for the corresponding number of local maxima) – then a conclusion can be made, with a sufficient probability, of coincidence or the proximity of the "flows of events" described in those texts. Moreover, the smaller the number $p(X, Y)$, the bigger this chance.

Furthermore, both compared texts may appear completely different – for instance, two versions of the same chronicle written in different countries, by different chroniclers, in different languages.

This method of dating was experimentally tested on mediaeval texts with *a priori* known dates, and the newly acquired dating coincided with those. Now, let us give a few typical examples.

**Example 6.**

In the capacity of the text $Y$, we have chosen a Russian chronicle, the so-called short edition of the ***Dvinskoy Letopisets*** (The Dvina Book of Chronicles), describing the events in the time interval of 320 years ([672]). We shall try and date the events recorded in this chronicle using said method. Looking through all chronicles published in ***The Complete Russian Chronicles***, we shall soon discover text $X$, for which the peaks on volume graph $vol\, X(t)$ occur virtually in the same years that those on graph $vol\, Y(t)$ of the chronicle $Y$, fig. 5.14.

While comparing the graphs, we made sure to have preliminarily superposed time intervals $(A, B)$ and $(C, D)$ one over another. The result of calculation is $p(X, Y) = 2 \times 10^{-25}$. Therefore, these two chronicles most probably describe approximately the same "flows of events". Thus, we manage to date the events recorded in text $Y$ in a fairly formal way, on the basis of the sole comparison of statistical characteristics of texts. The chronicle $X$ turns out to be a lengthy edition of the ***Dvinskoy Letopisets*** ([672]). This chronicle is considered to describe the "flow of events" of 1390-1707 A.D.

As a result, the dating of the text $Y$ we obtained co-

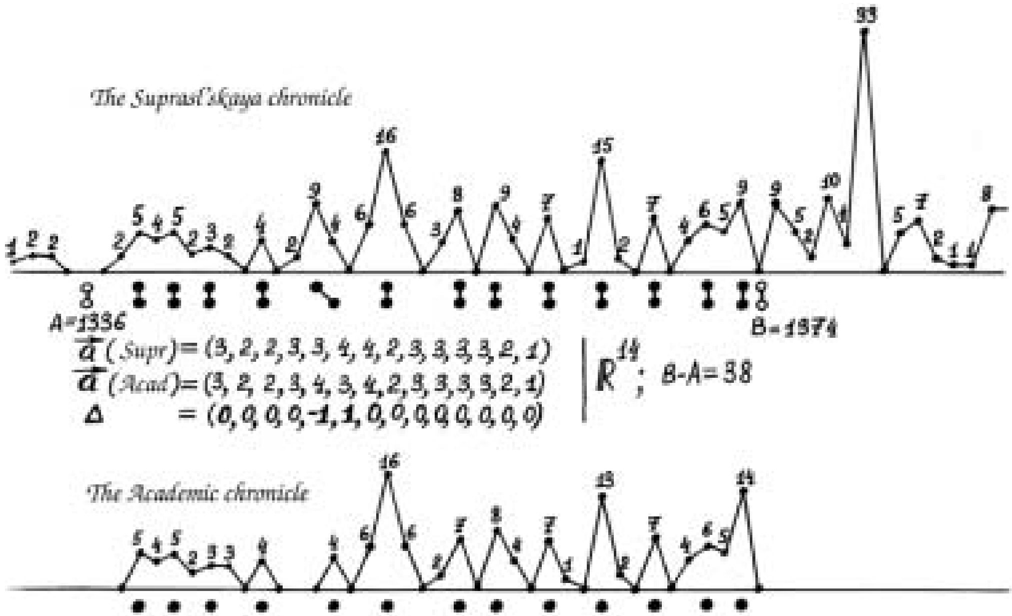


Fig. 5.15. Graphs of dependent chronicles – the ***Suprasl'skaya*** and the ***Akademicheskaya*** on the interval of 1336-1374 A.D. The peaks of the volume graphs occur identically all the time, with just one exception. The locations of local maxima of the graph are marked with thick black dots underneath the graphs – in case of the ***Supraslskaya Chronicle***, these two chains of dots are nearby each other. One sees that the peak points only fail to coincide once. The two chronicles are thus clearly of a dependent nature.

incides with its standard dating, which proves the efficiency of our method.

**Example 7.**

We shall take the Russian ***Akademicheskaya letopis'*** (Academic Chronicle) ([672]) as "text $Y$ with unknown dating". Following the example described above, we soon discover text $X$, namely, a part of the ***Suprasl'skaya letopis'*** ([672]) thought to have described years 1336-1374 A.D. The peaks on the volume graph *vol* $X(t)$ turns out to occur virtually in the same years as those on the volume graph *vol* $Y(t)$, fig. 5.15.

Calculation yields the result $p(X, Y) = 10^{-14}$. Such a small value of the coefficient clearly indicates the dependence of these two texts. Since chronicle $X$ is dated, we can date the chronicle $Y$, too. The obtained dating of text $Y$ coincides with its dating as known before.

Our research was based on several dozens of similar texts of the XVI-XIX century, and in all cases the acquired dating of the "unknown text $Y$" coincided with its usual dating.

In fact, we have learnt nothing new from the examples stated above, because the dating of the short edition of the ***Dvinskoy Letopisets***, for instance, has been known in advance, and we had no reasons to doubt its correctness, since it belongs to the XIV-XVIII century, that is, the epoch when the chronology is more or less dependable. Nevertheless, soon we shall see our method to yield very interesting results regarding chronicles attributed to earlier epochs, that is, those preceding the XIV century A.D.

The maxima correlation principle has been stated above in its rough form, without an attempt to go deep into statistical detail, because we were only after being understood by our readers as fast as possible. Meanwhile, a strict mathematical presentation of the method and its clarifications demand a substantially more detailed study. We would refer our readers wishing to delve into the described method to scientific publications [884] and [892].

The coefficient $p(X, Y)$ can conditionally be called PACY – the Probability of Accidental Coincidence of Years described in detail by chronicles $X$ and $Y$.

A further development and adjustment of the idea is presented in the works by V. V. Fedorov and A. T. Fomenko ([868]), as well as A. T. Fomenko, V. V. Kalashnikov and S. T. Rachev ([357]). It was further revealed that the maxima correlation principle manifests itself most explicitly when comparing historical texts of approximately the same volume and "density of description". Moreover, in some cases not only the local maxima points for *a priori* dependent texts, but also their volume functions, or amplitudes, turned out to correlate! The correlation of volume function amplitudes is especially visible when comparing "fairly poor" texts, or the chronicles with large lacunae – considerable time intervals not reflected in the chronicle. The process of writing "fairly poor" chronicles turns out to be subject to a fairly interesting principle – "respect for information", or "preservation of rarities", a regularity discovered by A. T. Fomenko and S. T. Rachev ([723] and [1140]). For preliminary research in this direction and the formulation of the principle of respect for information, see works [723] and [1140], as well as below in the paragraph written by A. T. Fomenko and S. T. Rachev.

The maxima correlation principle was successfully applied to the analysis of certain Russian chronicles of the period of "strife" at the end of the XVI century – beginning of the XVII century A.D. See related works by A. T. Fomenko and L. E. Morozova ([902] and [548]). N. S. Kellin took a major part in this research as well. See below the part written by A. T. Fomenko, N. S. Kellin, and L. E. Morozova.

## 2. VOLUME FUNCTIONS OF HISTORICAL TEXTS AND THE AMPLITUDE CORRELATION PRINCIPLE

*By A. T. Fomenko and S. T. Rachev*

(S. T. Rachev, doctor of physics and mathematics, Professor, specialist in the field of probability theory and mathematical statistics, Research Fellow of the Institute of Mathematics of the Bulgarian Academy of Sciences; currently works in the USA.)

### 2.1. Dependent and independent chronicles. Volume function maxima correlation

We shall describe the results published by the authors in [723] and [1140]. As above, we shall call two historical chronicles *X* and *Y dependent* if they can be traced back to a common original source and record approximately the same events on the same time interval (*A, B*) in the history of the same region.

On the contrary, we shall consider two chronicles *independent* if they record events of substantially different time intervals (*A, B*) and (*C, D*), or describe events in obviously different geographical regions. We shall consider two time intervals *substantially different* if their intersection on the time axis (i.e., their common part) does not exceed half of their length. Hereinafter, for the sake of simplicity, we shall assume that chronicles compared describe time intervals of the same length, i. e., $B - A = D - C$.

Let chronicle *X* describe events on the time interval (*A, B*), and parameter *t* run through the years from year *A* to year *B*. As above, we shall mark the part of the chronicle that describes the events in the year *t* as *X*(*t*). For the sake of brevity, we shall conventionally call fragments *X*(*t*) *chapters*. Let us calculate the volume of each fragment in certain units, for instance, in quantity of lines, or in pages. In the examples below, the volume of chapters is calculated in lines. However, the choice of measurement unit is not of great importance here. During statistical processing we have normalized the volume of chapters by dividing them by the total volume of the chronicle, thus levelling a possible difference in choice of volume measurement units. So, we obtain the function *vol X*(*t*) that we call the *volume function* of the chronicle.

The correlation principle for local maxima points of the volume graphs was formulated and experimentally tested by A. T. Fomenko in [884]. The main idea in the basis of the principle and the methods pertinent to it is as follows: dependence or independence of chronicles can in certain cases be established by comparing their volume functions. Generally speaking, *local maxima points of volume graphs of dependent chronicles should "correlate"* (in a proper precise sense, see above), *while independent chronicles should not display any "correlation"*, fig.5.1.

In their work [357], A. T. Fomenko, V. V. Kalashnikov and S. T. Rachev, applied the general idea of volume function correlation for dependent chronicles, and the absence of correlation for independent chronicles, to *volume functions themselves*, that is, considering their *amplitudes.* Since the research involved the amplitudes of graphs, this enhanced form of correlation principle should have been tested on specific

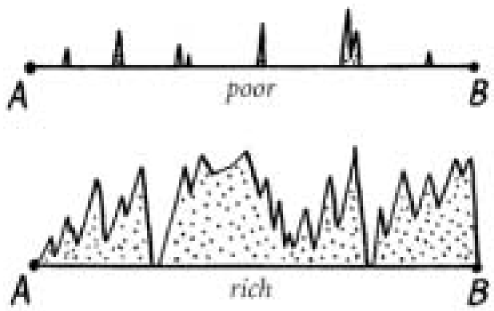


Fig. 5.16. Volume graphs of a rich chronicle and a poor one.

chronicles, which were performed in [357] with participation of N. Y. Rives. Detection methods for dependent and independent chronicles as offered in [357], turned out to be fairly efficient when comparing chronicles of approximately *the same volume*. However, the picture was becoming "smudged" when chronicles of substantially *different volumes* were compared. The current work specifies a new class of chronicles, for which the enhanced form of the local maxima amplitude correlation principle is correct.

The maxima correlation principle discovered by A. T. Fomenko relied upon the fact that different chroniclers, describing the same historical epoch, would generally use *the same volume* or fund of information that survived until their time. That is why, as our statistical experiments prove, they *would describe in greater detail only those years from which many texts survived, and in smaller detail all the rest of them.*

We shall recall the notion of primary information volume for events of epoch $(A, B)$. Let $C(t)$ be the volume of all documents written by the contemporaries of year $t$ about the events of that year, fig. 5.2. Now, let $X$ and $Y$ be chroniclers who are not contemporaries of the epoch $(A, B)$ but willing to write its history. Let $M$ (respectively $N$) be the year in which chronicler $X$ (respectively $Y$) creates the chronicle for the epoch $(A,B)$.

We shall recall that $C_M(t)$ is the volume of documents that survived from the epoch $(A, B)$ till the moment $M$, or the epoch of the chronicler $X$, – in other words, the remainder of primary texts survived till $M$. Graph $C_M(t)$ is the volume graph for the surviving information about the events of the epoch $(A, B)$. $C_N(t)$ is defined similarly.

The maxima correlation principle ensues from the following principle. Each chronicler $X$, describing the epoch $(A, B)$, "on the average" talks in greater detail about years in which the graph $C_M(t)$ peaks – i. e., the more documents from the epoch $(A, B)$ are available to the chronicler $X$, the more detailed is his description of that time, q.v. in fig.5.3.

## 2.2. Rich and poor chronicles and chronicle zones

The definition of a poor chronicle or a rich one is intuitively clear from fig. 5.16. We shall call the chronicle with the "majority" of volumes *vol* $X(t)$ equalling zero *poor*, where most of the years haven't been described by a chronicler. On the contrary, we shall call the chronicle with the "majority" of volumes *vol* $X(t)$ other than zero and fairly large *rich*, where a chronicler reports ample information about the epoch $(A, B)$.

In fact, for actual examples it is sometimes difficult to categorize a chronicle as either poor or rich, therefore, the introduction of new definitions – *poor zone* and *rich zone* of a chronicle – would be practical. Fig. 5.17 presents a relative volume graph of a chronicle with a *poor* beginning and a *rich* ending. Our research experience for specific chronicles makes it clear that *the beginning* of a long chronicle is a *poor zone*,


Fig. 5.17. The poor initial zone of a chronicle, and a richer zone following it.


Fig. 5.18. The rich and the poor zones may alternate within one and the same chronicle.

and its ending is a *rich zone*, typically, although there are chronicles with a poor zone "in the middle", q.v. in fig. 5.18.

### 2.3. Significant and insignificant zeroes of volume functions

In our study of a specific chronicle we shall assume the first year for which *vol X*(*A*) differs from zero as the leftmost point *A* on the time axis, the year is *described* by a chronicler, in other words, we shall call the zero of a volume graph *significant* if it is located *to the right* from the first non-null value, fig.5.19. If the zero is *to the left* from the first non-null value of the graph, then we shall call it *insignificant*. An insignificant zero indicates that not only does the chronicler know nothing about that particular year, but also nothing of preceding years in general. A significant zero indicates that, although the chronicler knows nothing about that particular year, he knows at least something about some of the *previous* years.

From this moment on, we shall not normalize the volume function, since we want to consider the magnitude of amplitudes of local maxima in our research.

### 2.4.The information respect principle

Let us consider a certain historical epoch (*A, B*) and a chronicler *X* who lives in year *M*, where *M* is much bigger than *B*, fig.5.20. Describing the events of the epoch (*A, B*), the chronicler *X* has to rely on the surviving information fund $C_M(t)$, still available in his time. Our idea is that the chronicler *X* treats poor and rich zones of the survived information fund differently.

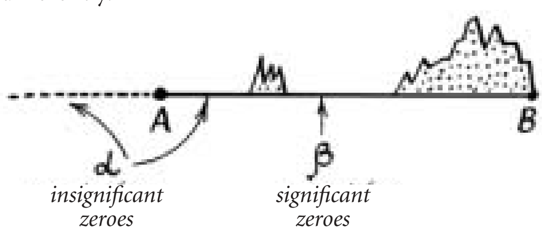


Fig. 5.19. Significant and insignificant zeroes of the chronicle volume function.

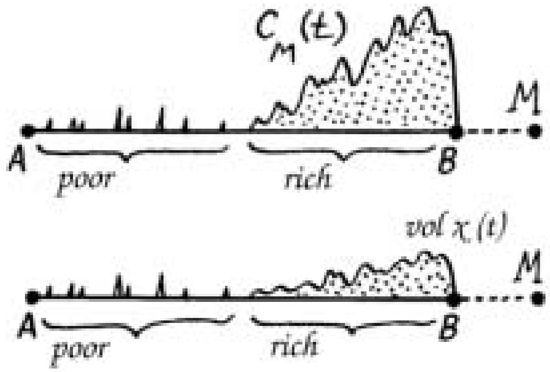


Fig. 5.20. The scribe accurately and scrupulously copies the "poor" zone of the remaining information fund of his time, and treats its richer zones with less reverence, selecting materials the way he sees fit.

We shall briefly formulate the model, the *information respect principle,* in the following way.

*A chronicler's respect for surviving information is in inverse proportion to its volume.*

The intuitive justification of this principle is clear. If some information has survived against a "zero-surrounded background", that is, when to the left and the right of it are the years of which the chronicler knows nothing, then the chronicler has to *highly appreciate* those scarce shreds of information miraculously spared by time. He copies them quite painstakingly, irrespective of his personal attitude towards their contents. Moreover, a chronicler in a *poor zone* of survived information fund has little space. He is limited in his freedom of action by a fairly small volume of surviving information. Therefore, the chronicler reproduces in good faith (by and large), the *amplitudes* of the volume function $C_M(t)$ for the information surviving in its poor zones.

The situation is different in what concerns the *rich zones*. A chronicler faces the necessity to *select* important things from the abundant choice of information. But the larger the volume of surviving information, the less does the chronicler appreciate individual pieces thereof, which often leads to distortions of volume graph amplitudes of the fund surviving in rich zones. Our statistical experiments have proved its veracity. The chronicler is free to be as subjective as he pleases: he can choose one kind of data and be intentionally "indifferent" to other.

## 2.5. The amplitude correlation principle of volume graphs in the poor zones of chronicles

We shall draw consequences from the information respect principle.

Let two chroniclers *X* and *Y* describe the same events on the same time interval (*A, B*). Each of them "copies" the volume graph of *poor zones* of the surviving information fund on the events of epoch (*A, B*) fairly well. Therefore, *the volume graphs of chronicles* X *and* Y *will look alike within poor zones*. Now we can formulate the model – *the amplitude correlation principle in poor zones*.

a) If chronicles *X* and *Y* are *dependent*, i. e., describe approximately the same events and trace back to a common original source, then their volume graphs *vol X*(*t*) and *vol Y*(*t*) should correlate quite well within their poor zones. In the meantime, within their rich zones there may be no amplitude correlation (upon superposition of graphs) at all.

b) If chronicles *X* and *Y* are *independent*, their volume graphs within their poor zones should be also independent, that is, there should be no amplitude correlation (upon the superposition of graphs).

That is, in case of poor dependent chronicles not only do the peaks of comparable graphs correlate, but also *their amplitudes*.

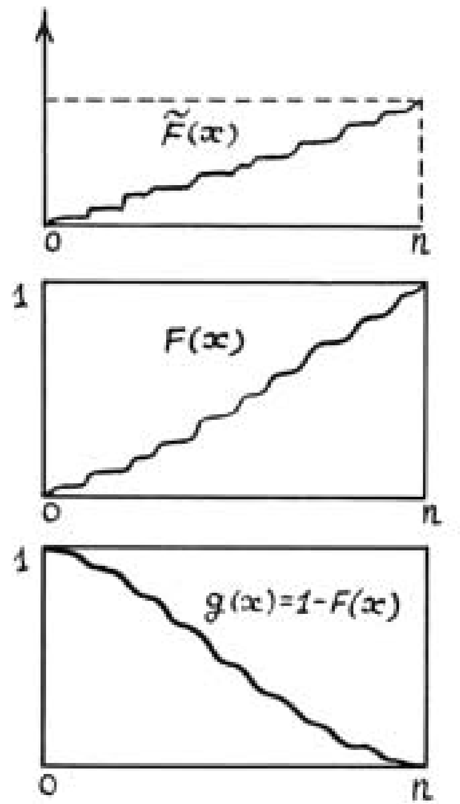


Fig. 5.21. Function graphs $F(x)$ and $g(x) = 1 - F(x)$.

## 2.6. Description of statistical model and formalization

We shall now consider the time period (*A, B*) and introduce the coordinate *x* varying from 0 to $B - A$ thereon, where $B - A$ is the length of the time period that we are interested in. It is clear that $x = t - A$. Let $f(x) = vol\, X(x)$ be the volume function of the chronicle *X*. We shall mark as $G(x)$ the function

$$G(x) = f(0) + f(1) + \ldots + f(x),$$

or, the "integral" of the function *f* from 0 to *x*. We shall call this function *the accumulated sum* of the chronicle *X*, and consider a normalized accumulated sum

$$F(x) = G(x) / vol\, X,$$

where *vol X* is the total volume of the chronicle *X*. The normalized accumulated sum is presented as a non-decreasing graph with values increasing from 0 to 1, character of this increase differing for various chronicles.

Let us consider a new function $g(x) = 1 - F(x)$. See fig. 5.21. Its graph does not increase. Omitting mathematical precision, we shall formulate the next model.

*The function* $g(x) = 1 - F(x)$ *should behave in the poor, early zone of the chronicle as function* $exp(-\lambda x^{\alpha})$.

In mathematical statistics, distributions of such kind are called the Weibull-Gnedenko distributions which are used in mathematical statistics for the description of similar processes.

Therefore, we have two degrees of freedom at our disposal: the parameter $\lambda$ and the parameter $\alpha$, swapping which, we can try to approximate the function $1 - F(x)$. If we manage to do it for specific chronicles, this will prove our theoretical model.

The statistical experiment that we performed with actual chronicles demonstrated that the decrease of

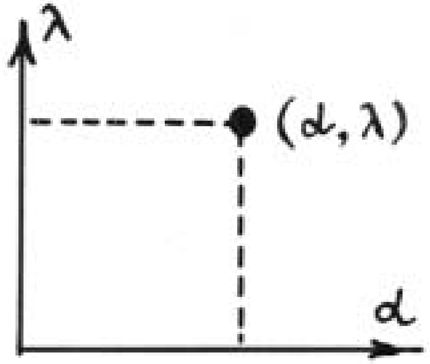


Fig. 5.22. Depiction of the two parameters – the shape and the volume of the chronicle in question – with a point on a plane.

the graph $1 - F(x)$ is indeed fairly well approximated by the function $exp(-\lambda x^{\alpha})$, given a suitable choice of values for $\lambda$ and $\alpha$.

As a result, we can juxtapose over each chronicle – or rather over its beginning – the poor zone thereof, of the two numbers $\lambda$ and $\alpha$ reflecting the character of the chronicle's volume function behaviour. We shall call $\lambda$ the parameter of the chronicle's *volume*, and $\alpha$ the parameter of the chronicle's *form*.

The parameter $\alpha$ turns out to be more important to us since, as statistical experiments have demonstrated, it is this parameter that has a better sense of the distribution character of individual scarce peaks of volume graphs within the poor zone of a chronicle. The parameter $\alpha$ will be the first to indicate whether chronicles are dependent or independent. The parameter $\lambda$ is rather responsible for the chronicle's volume, it demonstrates how rich or poor the chronicle is.

So, our hypothesis, or the statistical model may now be formulated in the following way.

a) If chronicles $X$ and $Y$ are *dependent*, then their pairs of corresponding parameters $(\alpha_X, \lambda_X)$ and $(\alpha_Y, \lambda_Y)$ should be *similar*, stipulating that they are calculated for the poor zones of the chronicles.

b) If the chronicles $X$ and $Y$ are *independent*, then their pairs of corresponding parameters $(\alpha_X, \lambda_X)$ and $(\alpha_Y, \lambda_Y)$ should be at some distance from each other.

It is convenient to picture the pair of numbers $(\alpha, \lambda)$ as a point on an ordinary plane with Cartesian coordinates $\alpha$ and $\lambda$. See fig. 5.22.

## 2.7. The hypothesis about the increase of the "form" parameter of a chronicle in the course of time

We shall now consider two different historical epochs: one with a poor primary information fund, and one with a rich primary fund. In the latter case, we assume the volume of this fund to be more or less constant for each year. Then, it can be demonstrated (omitting mathematical details) that the value $\alpha$ in the first, poor case should be *less* than the value of $\alpha$ in the second, rich case ([723], [1140]). See also articles 2.13 – 2.15. In other words, *poor primary funds* are characterized by *small* values of $\alpha$, and the *rich primary information funds* by *large* values of $\alpha$.

But the closer historical epoch $(A, B)$ is to our time, the better do the primary information funds survive. Today, for instance, written information is, by and large, on the average kept better than in the distant past. Therefore, the value of the parameter $\alpha$ should "on the average" *increase*, as we shift the time period $(A, B)$ under study from left to right on the time axis, i.e., closer to us.

## 2.8. The list and characteristics of the Russian chronicles we investigated

1) ***Povest' vremennykh let*** (*Story of Years of Time*). See ***Literary Memorials of the Ancient Rus'. The Beginning of the Russian Literature.*** Moscow, 1978.

This famous chronicle covers the events in the history of Russia, allegedly between the IX and XII century A.D. The main part of the chronicle describes the epoch of the alleged years 850-1110 A.D. in the consensual chronology. The chronicle begins with a poor zone approximately one hundred years long, starting allegedly in 850 A.D. and ending in the alleged year 940 A.D. The next part of the chronicle, beyond 1050-1110 A.D., is fairly rich.

2) ***Nikiforovskaya letopis'*** (*The Nikiforov Chronicle*), of the Byelorussian-Lithuanian group of chronicles. See ***The Complete Russian Chronicles***, Volume 35, Moscow, 1980. The period of 650 between the alleged years 850 A.D. and 1450 A.D. has been taken for our research work.

3) ***Suprasl'skaya letopis'*** (*The Suprasl' Chronicle*), of the Byelorussian-Lithuanian group of chronicles. See

***The Complete Russian Chronicles*** (***CRC*** for short), volume 35, Moscow, 1980. The period for which this chronicle provides the dates is allegedly 850-1450 A.D. This chronicle, as well as the Nikiforov one, can be rather ranked among *poor* texts in comparison with the richer ***Povest' vremennykh let***.

4) ***Akademicheskaya letopis'*** (*The Academy Chronicle*). See ***CRC***, volume 35, Moscow, 1980. We have researched the period of 1338-1378 A.D. This chronicle is intermediate between poor and rich texts.

5) ***Kholmogorskaya letopis'*** (*The Kholmogory Chronicle*). See ***CRC***, volume 33, St. Petersburg, 1977. It covers the period of the alleged years 850-1560 A.D. This chronicle contains both rich and poor zones.

6) ***Dvinskoy letopisets*** (*The Dvina Book of Chronicles*). Short and full editions. See ***CRC***, volume 33, St. Petersburg, 1977. It covers the period of 1390-1750 A.D. This chronicle contains both rich and poor zones.

All these chronicles begin with *poor zones*, which comes as no surprise. A. T. Fomenko calculated the volume functions. See CHRON1, Appendix 5.1. Among the listed chronicles, there are *a priori dependent* and *a priori independent* ones. For instance, among the *a priori dependent* are:

a) ***Nikiforovskaya letopis'*** and ***Suprasl'skaya letopis'***;

b) ***Povest' vremennykh let*** and ***Nikiforovskaya letopis'***, therefore ***Suprasl'skaya letopis'***, too.

c) Short and full versions of ***Dvinskoy letopisets***.

*A priori independent*, for instance, are the part of ***Dvinskoy letopisets*** covering the XIV century A.D., and the next one covering the XV century A.D.

The fact of dependence or independence of the listed chronicles has been confirmed in [884] and [868] on the basis of the maxima correlation principle, q.v. above.

## 2.9. The final table of the numeric experiment

All listed chronicles were divided into pieces covering approximately 100 years, each one examined with the method stated above. As a result, the parameters $\alpha_X$ and $\lambda_X$, and the correlation coefficient $r$ indicating how well the corresponding graph $exp(-\lambda x^{\alpha})$ approximates the decreasing graph $1 - F(x)$, were calculated (see table 5.1).

TABLE 5.1

| *Symbol* | *Chronicle* | *Epoch (A.D.)* | $\alpha$ | $\lambda$ | $r$ |
|---|---|---|---|---|---|
| P1 | ***Povest' vremennykh let*** | 854-950 | 1.847 | 3.9 × 10 | 0.953 |
| P2 | ***Povest' vremennykh let*** | 918-1018 | 3.003 | 1.6 × 10 | 0.955 |
| P3 | ***Povest' vremennykh let*** | 960-1060 | 2.497 | 4 × 10 | 0.956 |
| P4 | ***Povest' vremennykh let*** | 998-1098 | 2.378 | 1.3 × 10 | 0.954 |
| N1 | ***Nikiforovskaya letopis'*** | 854-960 | 1.511 | 9.3 × 10 | 0.966 |
| N2 | ***Nikiforovskaya letopis'*** | 960-1060 | 2.406 | 5 × 10 | 0.917 |
| N3 | ***Nikiforovskaya letopis'*** | 1110-1210 | 3.685 | 7 × 10 | 0.660 |
| N4 | ***Nikiforovskaya letopis'*** | 1236-1340 | 0.341 | 0.488 | 0.768 |
| N5 | ***Nikiforovskaya letopis'*** | 1330-1432 | 1.390 | 3.9 × 10 | 0.953 |
| S1 | ***Suprasl'skaya letopis'*** | 854-950 | 1.604 | 8.2 × 0 | 0.969 |
| S2 | ***Suprasl'skaya letopis'*** | 960-1060 | 2.584 | 3 × 10 | 0.943 |
| S3 | ***Suprasl'skaya letopis'*** | 1110-1210 | 3.617 | 7.8 × 10 | 0.656 |
| S4 | ***Suprasl'skaya letopis'*** | 1236-1340 | 0.405 | 0.384 | 0.808 |
| S5 | ***Suprasl'skaya letopis'*** | 1330-1432 | 2.354 | 1.6 × 10 | 0.983 |
| S6 | ***Suprasl'skaya letopis'*** | 1432-1450 | 2.089 | 1.3 × 10 | 0.977 |
| A | ***Akademicheskaya letopis'*** | 1336-1374 | 2.185 | 8 × 10 | 0.960 |
| D1 | ***Dvinskoy letopisets*** | 1396-1498 | 0.648 | 0.119 | 0.844 |
| D2 | ***Dvinskoy letopisets*** | 1500-1600 | 4.060 | 2.2 × 10 | 0.875 |
| K | ***Kholmogorskaya letopis'*** | 852-946 | 1.311 | 7.3 × 10 | 0.960 |

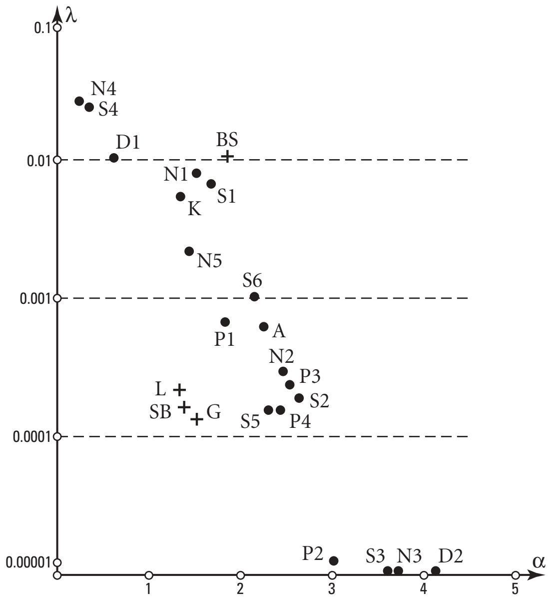


Fig. 5.23. Numeric parameters of shape and volume of the Russian chronicles that we have studied as points on a plane. The points marked by crosses stand for supplementary Russian chronicles which will be covered in more detail later on.

All the obtained value pairs (α, λ) were represented as points on the plane, fig. 5.23, with values of α from 0 to 6 plotted along the horizontal axis. In our experiment, we have not yet encountered values of α exceeding 5. Along the vertical axis we plotted the values of λ, but had to use a shifting, alternating scale. In particular, the first horizontal strip corresponds to the values of λ from 0 to 0.0001, the step size being 0.00001; the next horizontal strip corresponds to the values of λ from 0.0001 to 0.001 (scale factor 0.0001), and so on. Points on fig.5.23 represent pairs of numbers (α, λ) that we have calculated for the chronicles marked with respective abbreviations next to the points.

## 2.10. Interesting consequences of the numeric experiment. The confirmation of the statistical model

As we can see, in all cases considered, the decreasing function $1 - F(x)$ is very well approximated by the function $exp(-\lambda x^{\alpha})$, given suitable choice of parameters α and λ. See the last column of the table 5.1, where the values of the correlation coefficient *r* are apparently extremely close to 1. Thus, our statistical model is confirmed by the Russian chronicles under study – in particular, it turns out that volume functions of large historical chronicles can be modelled using the Weibull-Gnedenko distribution, a fact fairly interesting and useful in itself.

## 2.11. Comparison of a priori dependent Russian chronicles

We must make sure that points representing *a priori dependent* chronicles, or their fragments, must lie close by on the plane (α, λ). For instance, ***Nikiforovskaya letopis'*** and ***Suprasl'skaya letopis'*** were broken up into pieces: 850-950 A.D., 960-1060 A.D., 1110-1310 A.D., 1236-1340 A.D., and 1330-1432 A.D.

EXAMPLE 1. Fig. 5.23 makes it evident that the corresponding points *N1* and *S1*, or the first fragments of ***Nikiforovskaya letopis'*** and ***Suprasl'skaya letopis'*** respectively, *virtually coincide* on the plane (α, λ).

EXAMPLE 2. Points *N2* and *S2* are also *very close.*

EXAMPLE 3. Points *N3* and *S3 virtually coincide.*

EXAMPLE 4. Points *N4* and *S4 virtually coincide.*

EXAMPLE 5. Points *N5* and *S5*, on the contrary, "come apart" on the plane, indicating the absence of amplitude correlation. And indeed we find ourselves in the *rich* zone of the chronicle, for which our rule is not necessarily applicable.

EXAMPLE 6. Volume graphs of ***Nikiforovskaya letopis'*** and ***Suprasl'skaya letopis'*** are presented in fig. 5.24. Amplitude correlation of these chronicles, comparably poor in volume, is quite visible and confirmed by our numerical experiment.

EXAMPLE 7. The following pair of the comparable chronicles is especially interesting, because we compare a *poor* and a *rich* dependent text, – namely, ***Povest' vremennykh let*** and ***Nikiforovskaya letopis'***, or ***Suprasl'skaya letopis'***. The volume graph of ***Povest' vremennykh let*** is presented on fig. 5.24. There is no explicit *visual amplitude* correlation. Only at the beginning of all three chronicles, ***Povest' vremennykh let, Nikiforovskaya letopis'****, and* ***Suprasl'skaya letopis'*** is the *amplitude* correlation present; from about 950 A.D., it gradually becomes diluted.

EXAMPLE 8. ***Povest' vremennykh let*** was broken up into pieces: 854-950 A.D., 918-1018 A.D., 960-1060 A.D. and 998-1098 A.D. The point *P1*, that is, the one

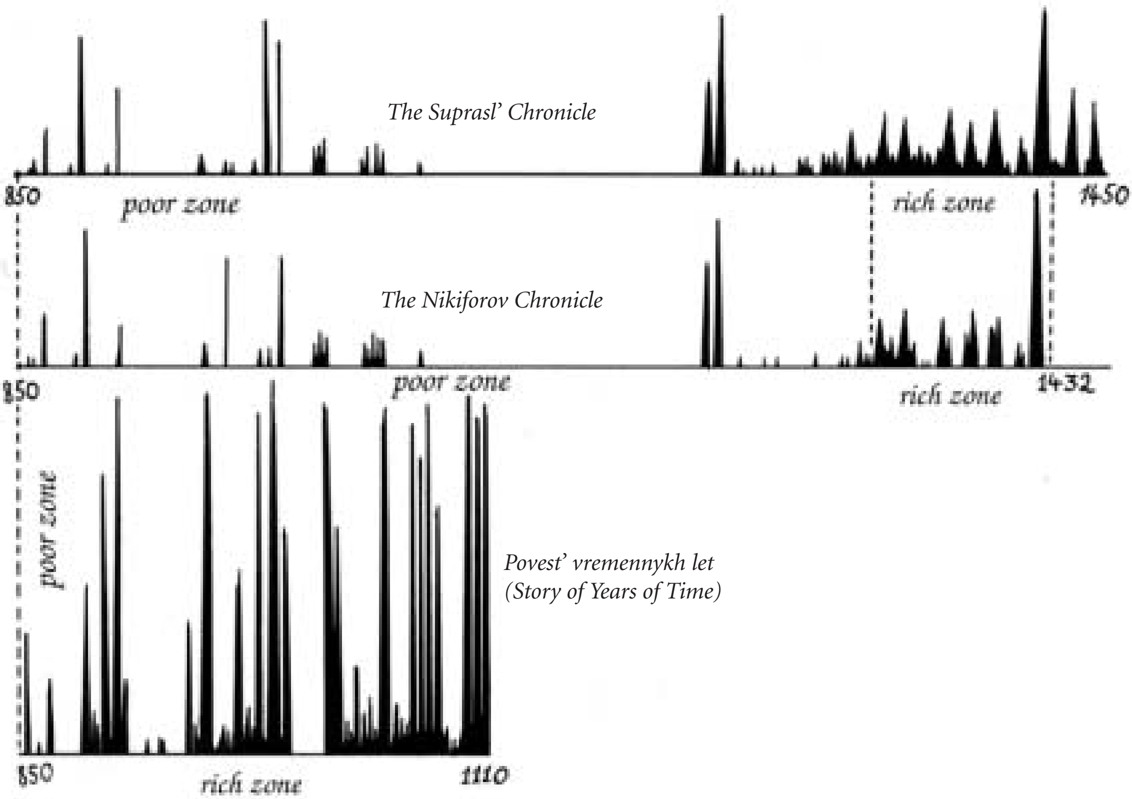


Fig. 5.24. Volume graphs of the Suprasl' chronicle, the Nikiforov chronicle and the ***Story of Years of Time*** chronicle with the rich and the poor zones emphasized.

corresponding to the period 854-950 A.D., seems to be far away on the plane (α, λ) from the virtually coinciding points *N1* and *S1,* which correspond to the pieces of ***Nikiforovskaya letopis'*** *and* ***Suprasl'skaya letopis'*** of 854-950 A.D., q.v. in fig. 5.23. However, we shall recall that the *main* parameter for us is α, or the form parameter. Comparing values of α for points *P1* and the pair of points *N1* and *S1*, that is, simply projecting these points on the horizontal axis, we can see that all three values of α are very close to each other. Therefore, the *rich* chronicle *P1*, i.e., ***Povest' vremennykh let***, is actually *dependent* in relation to the two *poor* chronicles *S1* and *N1*, i.e., ***Suprasl'skaya letopis'*** and ***Nikiforovskaya letopis'***. Thus, our method makes it possible to discover the *dependency* between *poor* and *rich* chronicles with certainty.

EXAMPLE 9. The points *P3, N2* and *S2 virtually coincide*, q.v. in fig. 5.23.

EXAMPLE 10. Finally, let us compare points *P4* and *N2, S2* corresponding to the chronicles describing close historical epochs. We can see that all three points are *very close* to each other on the plane. We have completely exhausted ***Povest' vremennykh let.***

Therefore, our *amplitude correlation* principle *for dependent texts in their poor zones* has been confirmed, – in certain cases, even for the rich zones of chronicles.

## 2.12. Comparison of a priori independent Russian chronicles

To avoid qualms about the obvious independence of compared chronicles, we shall restrict ourselves to the texts recording time periods after 1300 A.D. only, those close to our time.

EXAMPLE 11. Let us break up, for instance, ***Dvinskoy letopisets*** into two parts: 1396-1498 A.D. and 1500-1600 A.D. We have had no reason to doubt their independence. Turning to fig.5.23, we can see the corresponding points *D1* and *D2* to be *far away* from each other indeed – in diametrically opposite ends of the field filled with points representing the results of our experiment.

**Example 12.** Let us review ***Nikiforovskaya letopis'*** of 1110-1210 A.D. and its segment of 1236-1340 A.D. Although, according to the consensual chronology, they refer to *different* historical epochs, one cannot assert obvious independence of the two chronicles *a priori* since they describe events preceding 1300 A.D. Nonetheless, fig. 5.23 makes it clear that their corresponding points *N3* and *N4* are *far away* from each other on the plane (α, λ), which probably indicates their independence.

The experiments we performed with other independent chronicles (tables omitted) demonstrate that *obvious independence* of chronicles manifests itself in a substantial remoteness of points representing them on the plane (α, λ).

## 2.13. Growth of form parameter in the course of time for the Russian chronicles after the XIII century

If we examine the Russian chronicles distributed over the interval between the alleged IX-XVII centuries A.D., we shall see that this effect is not represented in fig. 5.23 with sufficient clarity. However, the situation becomes much clearer if we reduce ourselves to the chronicles beginning approximately from 1200 A.D. and closer to our time – i.e., from the moment when the consensual chronology may be trusted (to some extent, at least). The plane in fig. 5.23 is broken down into segments in accordance with different scales for parameter λ. Let us compare the positions of points found within one strip and describing events superceding the year 1200.

Fig.5.23 distinctly demonstrates that for all of three such points found within the fourth segment, – namely, the points *N4:* 1236-1340, *S4:* 1236-1340, *D1:* 1396-1498, – parameter α does actually *grow* over the course of time.

The third segment contains only two such points: *N5:* 1330-1432, and *S6:* 1432-1450. As we see, parameter alpha *grows over the course of time* as well, since point *S6* is located *to the right* of point *N5*.

The second strip in fig. 5.23 contains only two such points – *S5:* 1330-1432, and *A:* 1336-1374. These values α are very close to each other, virtually coinciding. This is understandable, since the epochs described in texts *A* and *S5* are close by.

The first segment has four points. Only one of them, *D2*, describes the period after 1200; therefore, it is impossible to verify our hypothesis within this segment. Nevertheless, one cannot fail to note that, if we examine all these four points formally, parameter α shall evidently increase in the course of time as well, although we certainly cannot trust the Scaliger-Miller chronology preceding the year 1200.

Let us now compare the positions of points *N4:* 1236-1340, and *N5:* 1330-1432, disregarding the values of λ. Point *N5* is evidently located *to the right* of point *N4*, i. e., parameter α *does actually grow over the course of time*.

The same is also true for points *D1* and *D2*. Point *D2:* 1500-1600 is located *to the right of* point *D1*: 1396-1498, and here parameter α *grows over the course of time* as well.

Finally, the mutual arrangement of points *S4:* 1236-1340, *S5:* 1330-1432, and *S6:* 1432-1450 also confirms our hypothesis of the *growth of parameter* α over the course of time.

The *growth* of the parameter α over the course of time that we discovered assumes a natural explanation: the more recent the chronicle, the "more uniform" its volume function.

And yet it is impossible to make an unambiguous conclusion about the growth of the parameter α over the course of time for individual chronicles on the basis of a small number of experiments. Extra research is necessary.

## 2.14. Growth of the average form parameter over the course of time for groups of Russian chronicles of the XIII-XVI century

In certain cases of the preceding paragraph, we possibly attempted to measure sufficiently rough values "too accurately". Therefore, it is more natural to examine not just various chronicles and their parts, but rather the *groups* of chronicles approximately related to one period of, say, 50 or 100 years. Then, the average values of the parameter for these groups of texts should be compared. Let us examine the texts

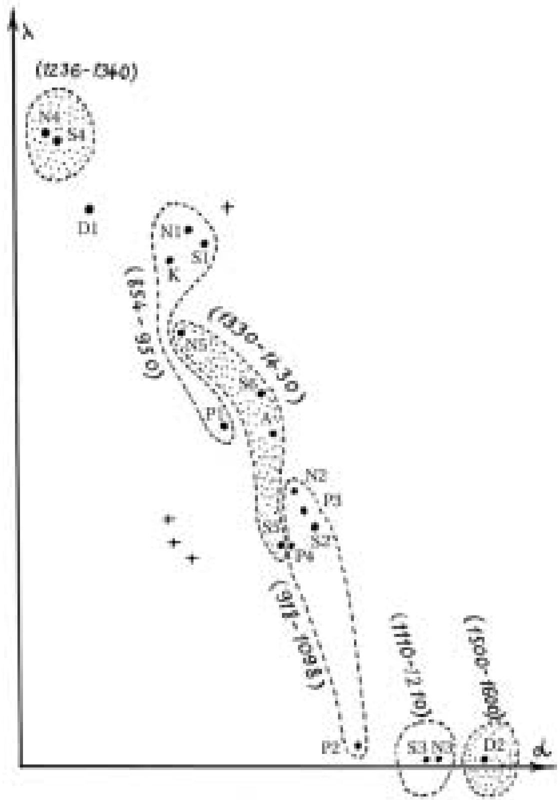


Fig. 5.25. The chronological shift of 300-400 years and its manifestation in the Russian history. One sees a "shaded group" of chronicles next to each "white group". The gap in time between them equals three or four centuries.

beginning with 1200 A.D. and those closer to us. See the result in fig. 5.25. The points, or the chronicles corresponding thereto, are united into several groups corresponding to different periods of history.

Group of years 1236-1340 – two chronicles: *N4* and *S4*.

Group of years 1330-1450 – four chronicles: *N5, S5, S6, A*.

Group of years 1500-1600 – one chronicle *D2*.

In fig. 5.25 it is distinctly evident that each next group is located *to the right of* the preceding one, which matches *the growth* of parameter α over the course of time. The only exception is chronicle *D1:* 1396-1498, found next to the group of chronicles of years 1236-1340. Thus, the "integration of the picture" causes the effect of the growth of parameter α with the flow of time to manifest itself explicitly enough.

## 2.15. Growth of the average parameter of form over the course of time for the groups of Russian chronicles of the alleged IX-XIII century

The Russian chronicles found in this epoch are united into several groups describing close historical periods, – namely:

Group of years 854-950 – four chronicles: *N1, K, S1, P1*.

Group of years 918-1098 – five chronicles: *N2, S2, P2, P3, P4*.

Group of years 1110-1210 – two chronicles: *S3* and *N3*.

In fig. 5.25 it is distinctly evident that each of these groups is located *to the right* of the preceding one, which again indicates the growth of parameter α over the course of time.

**Conclusion.** In the Russian chronicles believed today to date back to the alleged IX-XIII century A.D., and those currently dated back to the XIII-XVI centuries A.D., the parameter α grows evenly over the course of time on the average, which confirms our statistical hypothesis. But the even growth of the parameter α over the course of time discovered by us now makes possible the usage of this effect to establish the correctness or the inaccuracy of the chronology of various chronicles. Let us cite an example.

## 2.16. Chronological shift by 300 or 400 years in Russian history

Fig. 5.25 vividly demonstrates an exceptionally interesting phenomenon.

a) A group of Russian chronicles of the alleged years 918-1098 is characterized with approximately the same values of the parameter α as a group of later Russian chronicles of 1330-1430. Moreover, for both groups of chronicles the growth rate of α over the course of time is more or less the same. In fig. 5.25 these two groups of texts are positioned in such a way that their projections on the horizontal axis are close by. In this case, the Scaliger-Miller dating of these two groups of chronicles differs by approximately 300-400 years. *Thus, we reveal a chronological shift of approximately 300-400 years in the Romanov version of the Russian history*.

b) An absolutely similar effect also appears in the comparison of a group of Russian chronicles allegedly dated to 854-950, and a group of more recent Russian chronicles of 1236-1340 and 1330-1430. The group of 854-950 is located in fig. 5.25 between the groups of 1236-1340 and 1330-1430. Consequently, the values of the parameter $\alpha$ for the two groups of chronicles, which are normally set apart by approximately 300-400 years, once again prove to be very close to each other. *Again a chronological shift by 300-400 years is found in the Romanov version of the Russian history*.

c) We see a perfectly similar effect while comparing the parameters $\alpha$ for a group of Russian chronicles allegedly dated to 1110-1210 and 1500-1600. The values of $\alpha$ prove to be in sufficient propinquity once again. *We see the same chronological shift by approximately 400 years again*.

**An important conclusion.** Comparison of the values of parameter $\alpha$ shows that our statistical experiment with a large group of Russian chronicles *revealed a chronological shift of 300-400 years in the Romanov version of the Russian history*. Apparently, certain Russian chronicles, and therefore the events described therein, were dated incorrectly. As a result, certain actual events of the XIV-XVI century A.D. "slipped backwards in time" by 300-400 years and gave birth to their "phantom reflections" in the epoch of the alleged IX-XIII century A.D. We shall see further on that this 300-400 year shift in the Russian history is also revealed by means of completely independent methods.

## 2.17. Conclusions

1) A new empirico-statistical model that allows us to statistically recognize *dependent* and *independent* chronicles, as well as the statistical principles of *information respect* and *amplitude correlation* for the poor zones of chronicles, have been formulated.

2) Our model and both of the principles, namely, the statistical hypotheses, have been tested by a numeric experiment on the material of Russian chronicles. The model and both of the principles have been confirmed by trustworthy and reliably dated material.

3) It allows us to propose a procedure for the recognition of dependent and independent chronicles.

4) We have obtained the following statistical conclusions as a result of our analysis of several Russian chronicles.

4a. A damping graph $1 - F(x)$, where $F(x)$ is a normalized accumulated sum of the volume function of the chronicle, can be approximated sufficiently well by the function $exp(-\lambda x^{\alpha})$ with a suitable selection of parameters $\alpha$ and $\lambda$.

4b. For the *dependent* chronicles $X$ and $Y$, points $(\alpha_X, \lambda_X)$ and $(\alpha_Y, \lambda_Y)$ corresponding to them on the plane $(\alpha, \lambda)$ are in propinquity.

4c. For the *independent* chronicles $X$ and $Y$, points $(\alpha_X, \lambda_X)$ and $(\alpha_Y, \lambda_Y)$ corresponding to them on the plane $(\alpha, \lambda)$, on the contrary, are distant.

4d. The parameter $\alpha$, and sometimes also parameter $\lambda$, usually characterizes an entire group of chronicles describing events of the specified period. In other words, the parameter $\alpha$ is in a certain sense an "invariant of historical epoch" and its chronicles. This effect may be considered established for the Russian chronicles of the XIV-XVII centuries, i.e. more or less reliably dated texts.

5) Our statistical experiment with a large group of Russian chronicles *revealed a chronological shift by 300-400 years in the Romanov version of the Russian history*.

# 3. THE MAXIMA CORRELATION PRINCIPLE ON THE MATERIAL OF THE SOURCES PERTINENT TO THE EPOCH OF STRIFE IN THE HISTORY OF RUSSIA (1584-1619)

*A. T. Fomenko, N. S. Kellin, L. E. Morozova*

(N. S. Kellin, Candidate of Physical and Mathematical Sciences, senior researcher of the M. V. Keldysh Institute of Applied Mathematics of the Russian Academy of Science, Moscow.
L. E. Morozova, Candidate of Historical Sciences, associate of the Institute of History of the USSR, the USSR Academy of Science.)

We will show now how the maxima correlation principle formulated by A. T. Fomenko manifests itself in a group of dependent historical texts related to the epoch of strife in Russia (the end of the XVI – the beginning of the XVII century A.D.). We have divided each of 20 texts into per annum fragments, or pieces describing the events of separate years, and then N. S. Kellin and L. E. Morozova calculated volumes of all those "chapters" – namely, a number of words in each "chapter". The results obtained were formalized in a united Table 5.2, where the volume of per annum fragments from 1584 to 1619 is indicated for each of the 20 texts.

Here is the list of the investigated texts:

1) ***Povest' o Chestnom Zhitii,*** 2) ***Povest Kako Voskhiti***, 3) ***Povest Kako Otmsti***, 4) ***Zhitie Dmitriya (Touloupova)***, 5) ***Zhitie Dmitriya (Maliutina)***, 6) ***Skazanie O Grishke***, 7) ***Skazanie o Fyodore***, 8) ***Skazanie o Samozvantse***, 9) ***Povest Shakhovskogo***, 10) ***Zhitie Iova***, 11) ***Skazanie Avraamiya*** (1st edition), 13) ***The Chronographer of 1617***, 14) ***Vremennik Timofeyeva***, 15) ***Povest' Katyreva*** (1st edition), 16) ***Povest' Katyreva*** (2nd edition), 17) ***Inoye Skazaniye***, 18) ***Piskaryovskiy Letopisets***, 19) ***Noviy Letopisets.***

Three more texts were added later: 20) ***Izvet Varlaama,*** 21) ***Bel'skiy Letopisets*** and 22) ***Skazaniye O Skopine***.

Below is Table 5.2 of the per annum fragment volumes for the first 19 texts. The years are plotted along the horizontal axis, and the numbers of texts along the vertical. Years are indicated in abbreviated form: 84, 85, 86, etc. instead of 1584, 1585, 1586, etc.

All these historical texts basically describe the same events, therefore they are dependent, based on the same fund of surviving information. Table 5.2 shows that correlation between the peaks, i.e., local maxima of volume functions of these texts, is expressed clearly. It is evident that the peaks on almost all graphs occur virtually simultaneously, in particular, in the years: 1584, 1587, 1591 and 1598.

Now let us consider the result of the second numeric experiment, in which the 19 preceding texts were followed by the three additional texts (see above), with the time limits extended as well – namely, the interval of 1584-1598 A.D. was supplemented with years 1598-1606 – and a table similar to the preceding one was plotted. In Table 5.3, the symbol (•) marks the positions of the local maxima for all 22 historical texts within the range between 1584 and 1606 A.D.

It is distinctly evident that the peaks of all volume functions occur virtually simultaneously, which is explained by the dependence of these texts. *Consequently,*

**Table 5.2**

| | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1: | 432 | 288 | | 200 | 375 | 376 | 1112 | 1632 | | | | | | | 2784 |
| 2: | 140 | 455 | | 458 | | | | 105 | | | | | | | 196 |
| 3: | 230 | | | 800 | | | | 157 | | | | | | | 380 |
| 4: | 120 | | | | | | | 740 | | | | | | | 48 |
| 5: | 180 | | | 500 | 400 | 300 | 306 | 500 | | | | | | | 400 |
| 6: | 152 | | 52 | 180 | | | | 76 | | | | | | | 68 |
| 7: | 240 | 200 | 206 | 240 | 200 | 208 | 210 | 2884 | | | | 20 | 22 | 26 | 756 |
| 8: | 20 | | | | | | | 93 | | | | | | | 128 |
| 9: | 128 | | | | | | | 600 | | | | 20 | 26 | 28 | 360 |
| 10: | 240 | 200 | 100 | 102 | 106 | 450 | | 60 | 56 | 52 | 51 | 50 | 50 | 52 | |
| 11: | 44 | | | 42 | | | | 108 | | | | | | | 306 |
| 12: | 54 | | | 42 | | | | 347 | | | | | | | 112 |
| 13: | 312 | | | 172 | 43 | 42 | | 132 | | | | | | | 324 |
| 14: | 900 | | | 120 | | | | 4420 | 26 | 22 | 20 | 20 | 26 | 28 | 3000 |
| 15: | 150 | | | 120 | | | | 300 | | | | | | | 500 |
| 16: | 152 | | | 86 | | | | 300 | | | | 10 | 10 | 12 | 434 |
| 17: | 264 | | | 675 | | | | 863 | 92 | 90 | | 90 | 92 | 94 | 1034 |
| 18: | 325 | 75 | 50 | 44 | 32 | 46 | 122 | 430 | 86 | 35 | 140 | 20 | 20 | 110 | 1160 |
| 19: | 441 | 99 | 150 | 152 | 54 | 54 | 189 | 1548 | 522 | 36 | 342 | 648 | 50 | 50 | 540 |

**Table 5.3**

| | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 00 | 01 | 02 | 03 | 04 | 05 | 06 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1: | • | | | | | | | • | | | | | | | • | | | | | | | | |
| 2: | | • | | • | | | | • | | | | | | | • | | | | • | | • | | |
| 3: | • | | | • | | | | • | | | | | | | • | | | • | | | | | • |
| 4: | • | | | | | | | • | | | | | | | • | | | • | | | | | • |
| 5: | • | | | • | | | | • | | | | | | | • | | | | | | | | • |
| 6: | • | | | • | | | | • | | | | | | | • | | | | | | | | • |
| 7: | • | | | • | | | | • | | | | | | | • | | | • | | | | | • |
| 8: | • | | | | | | | • | | | | | | | • | | | | • | | | | • |
| 9: | • | | | | | | | • | | | | | | | • | | | | | | | | • |
| 10: | • | | | | | • | | • | | | | | | • | | | | | | • | | | • |
| 11: | • | | | • | | | | • | | | | | | | • | | | | | | • | | • |
| 12: | • | | | • | | | | • | | | | | | | • | | | • | | • | | • | |
| 13: | • | | | • | | | | • | | | | | | | • | | | | | | | | |
| 14: | • | | | • | | | | • | | | | | | | • | | | | | | | | • |
| 15: | • | | | • | | | | • | | | | | | | • | | | • | | | | • | |
| 16: | • | | | • | | | | • | | | | | | | • | | | | | • | | | • |
| 17: | • | | | • | | | | • | | | | | | | • | | | | | | | | • |
| 18: | • | | | | | | | • | | • | | | | | • | | • | | | • | | • | |
| 19: | • | | | • | | | | • | | | • | | | | • | | • | | | • | | | • |
| 20: | • | | | • | | | | • | | | | | | | • | | | | • | | | | • |
| 21: | • | | | • | | | | • | | • | | | | | | • | • | | | | • | | • |
| 22: | | | | | | | | | | | | | | | • | | • | | | • | | • | |

*this confirms the peak correlation principle for the volume functions of dependent texts.*

This dependence of texts can be expressed numerically. Let us introduce the following "distance" between volume functions *vol* $X(t)$ and *vol* $Y(t)$ for the two texts $X$ and $Y$, each divided into clusters of separate per annum fragments $X(t)$ and $Y(t)$, respectively. Let us recall that the fragments $X(t)$ and $Y(t)$ describe the events of just one year $t$.

Let parameter $t$ vary within the time interval from year $A$ to year $B$. Let us designate by $t(X, 1)$, $t(X, 2)$, …, $t(X, N)$ the years in which such peaks, or local maxima, occur on volume graph *vol* $X(t)$. Accordingly, let us designate the peaks of the volume graph *vol* $Y(t)$ by $t(Y, 1)$, $t(Y, 2)$, …, $t(Y, M)$.

For each point $t(X, i)$, let us find *the point nearest to it* in the sequence $t(Y, 1)$, $t(Y, 2)$,…, $t(Y, M)$. Let it be a certain point $t(Y, k)$. Let $p(i)$ designate the distance between them in years, i.e. the absolute difference value $t(X, i) - t(Y, k)$. In other words, we shall find out which local maximum of $Y$ is the nearest to the selected local maximum of $X$.

In a perfectly similar manner, swapping the roles of $X$ and $Y$, for each point $t(Y, j)$ let us find *the nearest point to it* in the sequence $t(X, 1)$, $t(X, 2)$, …, $t(X, N)$. Let it be a certain point $t(X, s)$. Let $q(j)$ designate the distance between them in years, or the absolute value of difference $t(Y, j) - t(X, s)$.

Finally, we assume the following sum as "the distance between $X$ and $Y$":

$$R(X, Y) = p(1)+p(2)+ \dots + p(N)+ q(1)+ q(2)+ \dots + q(M).$$

The meaning of the distance $R(X, Y)$ is completely clear. For each local maximum of function *vol* $X(t)$ we find the nearest local maximum of function *vol* $Y(t)$, determine the distance between them in years, and sum up the numbers obtained. Then we repeat this operation after swapping the positions of chronicles $X$ and $Y$. Summing up the numbers obtained, we obtain $R(X, Y)$. It is clear that $R(X, Y) = R(Y, X)$.

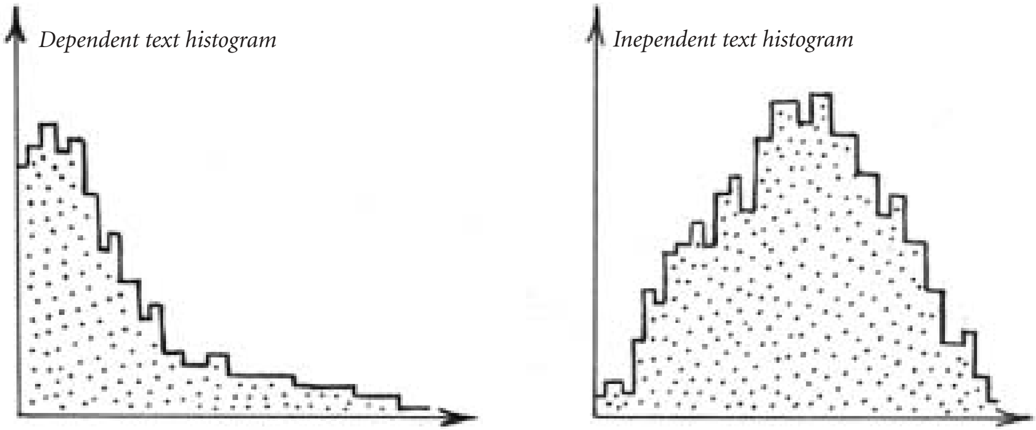


Fig. 5.26. Histograms for dependent and independent historical texts.

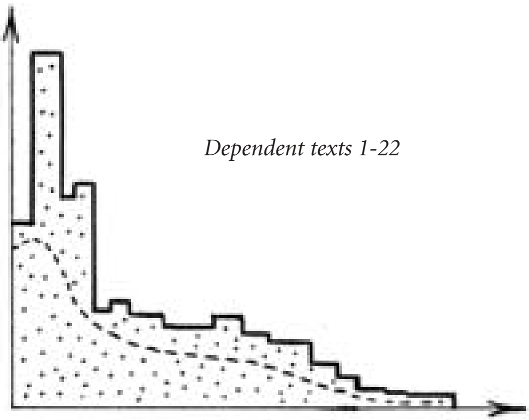


Fig. 5.27. Histogram for the dependent texts 1-22.

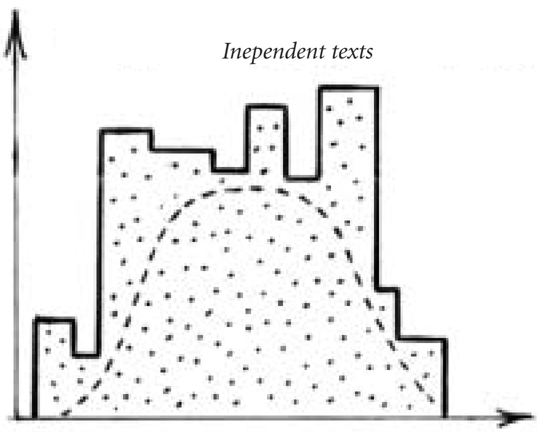


Fig. 5.28. Histogram for independent texts.

If distance $R(X, Y)$ equals zero for a certain pair of texts $X$ and $Y$, consequently, their volume function graphs peak *simultaneously*. The greater this distance, the worse the correlation between their local maxima points. It is also possible to examine the asymmetrical distance from $X$ to $Y$, assuming that

$$p(X, Y) = p(1) + p(2) + \dots + p(N).$$

Likewise, the asymmetrical distance from $Y$ to $X$ is determined, namely,

$$q(Y, X) = q(1)+ q(2)+...+ q(M).$$

Let us numerically estimate a degree of dependence between the historical texts 1-22 listed above, for which we shall calculate a 22×22 square matrix of two-by-two distances $R(X, Y)$, where $X$ and $Y$ pass through all texts 1-22, independently from each other. Let us then calculate a frequency histogram, for which we shall consider the horizontal axis, on which we shall mark the integer points: 0, 1, 2, 3, … and plot the following graph. Let us calculate the number of zeroes in the matrix $\{R(X, Y)\}$ obtained earlier. The number obtained will be plotted on the vertical axis at the point of which horizontal coordinate is equal to zero. Then we shall calculate the number of unities in the matrix $\{R(X, Y)\}$, plot the obtained number on the vertical axis at the point of which horizontal coordinate is equal to 1, and so on. We shall come up with a graph called frequency histogram. What can a study of the obtained histogram tell us?

If the chronicles selected for the analysis *are de-*

*pendent*, then the majority of two-by-two distances between the chronicles must be expressed *in small numbers*, which is to say, the chronicles must be close to each other, meaning that the majority of matrix elements $\{R(X, Y)\}$ must be "small" or close to zero. In that case, however, the absolute maximum of the frequency histogram must be shifted *to the left*, that is, there should be a large set of small frequencies. On the contrary, if there are many *independent* texts among those under investigation, then the maximum of the frequency histogram is shifted to the right, q.v. in fig. 5.26. The share of "large" and "medium" two-by-two distances between the chronicles should therefore increase.

This observation makes it possible to evaluate the degree of dependence or independence for a group of chronicles by plotting an appropriate frequency histogram based on matrix $\{R(X, Y)\}$. Namely, a shift of the maximum *to the left* indicates a possible *dependence* of chronicles, while a shift of the maximum *to the right* indicates a possible *independence*.

This idea was used to evaluate the degree of dependence of historical texts 1-22 enumerated above. Fig. 5.27 shows the experimental histogram of the matrix $\{R(X, Y)\}$ for texts 1-22. This matrix proved to possess many small numbers, therefore the maximum of the histogram is visibly shifted to the left. *This indicates the dependence of historical texts 1-22.*

For comparison, let us plot a histogram for independent texts. To present an example, we decided to compare three chronicles *A, B, C* mentioned below with the preceding texts 1-22. The three additional chronicles are:

A: ***Povest' Vremennykh Let***, allegedly 850-1110 A.D.,
B: ***Akademicheskaya letopis'***, allegedly 1336-1446 A.D.,
C: ***Nikiforovskaya letopis'***, allegedly 850-1430 A.D.

For each of them, a volume function was calculated and all local maxima found. Let us calculate all two-by-two distances of $\{R(X, Y)\}$, where *X* passes through the three chronicles *A, B, C,* and *Y* passes through the historical texts 1-22. As a result, we obtain a rectangular $3\times22$ matrix $\{R(X, Y)\}$. Then a frequency histogram was calculated, with its result shown in fig. 5.28. *A totally different nature* of this histogram is distinctly visible – its maximum moved *to the right*. This indicates *independence* of two groups of texts: $\{A, B, C\}$ and {*texts 1-22*}. Each of these groups can certainly contain dependent texts.

## 4. THE METHOD FOR THE RECOGNITION AND DATING OF THE DYNASTIES OF RULERS
### The small dynastic distortions principle

### 4.1. The formulation of the small dynastic distortions principle

The small dynastic distortions principle, and a method based thereupon, was proposed and developed by the author in [884], [885], [888], [1129], [895] and [1130].

Let us assume a historical text to be found, describing a dynasty of rulers unknown to us, indicating the duration of their rule. The question arises whether this dynasty is a new one, unknown to us, and therefore requiring dating, or is it one of the dynasties we know, but described in the terms we are not used to – for example, the names of rulers are altered, etc.? The answer is in the procedure below ([904], [908], [1137], [885] and [886]).

Let us examine the *k* value of any successive actual rulers or kings in the history of some state or region. We shall agree to name this sequence *an actual dynasty*; its members by no means have to be related, though. Frequently, the same actual dynasty is described in different documents, by different chroniclers, and from different points of view – for example, the activity of rulers, their significance, personal qualities, and so forth, evaluated in a different way. Nevertheless, there are the "invariant" facts, the description of which is less dependent on sympathies or antipathies of chroniclers. These more or less "invariant facts" include, for example, *the duration of the rule of a king*. Usually there are no special reasons for a chronicler to significantly or intentionally distort this figure. However, chroniclers would frequently encounter natural difficulties in calculating reign duration for this or that king.

These natural difficulties are as follows: incompleteness of information, distortion in documents, etc. They sometimes resulted in the fact that chronicles or tables by different chroniclers would report different numbers, which to them seemed to be the reign duration of the same king. Such divergences, sometimes significant, are characteristic, for example, for the pharaohs in the tables by H. Brugsch ([99]) and in the ***Chronological Tables*** by J. Blair ([76]). For example, the tables by J. Blair, going as far as the beginning of the XIX century, collected all basic historical dynasties, with dates of rule, the information about which is available to us. The value of the tables by J. Blair for us lies in the fact that they were compiled in an epoch sufficiently close to the time of the creation of the Scaligerian chronology. Therefore, they contain clearer imprints of the "Scaligerian activity" which were subsequently shaded and plastered by the historians of the XIX-XX century.

Thus, each chronicler describing an actual dynasty $M$ calculates the reign duration of its kings in his own way, to the best of his abilities and possibilities. As a result, he obtains a certain sequence of numbers $a = (a_1, a_2, \dots, a_k)$, where number $a_i$ shows, possibly with an error, the actual reign duration for a king with the value $i$. Let us recall that the value $k$ represents the total number of kings in the dynasty. We agreed to call this sequence of values extracted from the chronicle, *a dynasty of annals*, convenient to be represented as vector $a$ in Euclidean space $R^k$.

Another chronicler describing the same real dynasty $M$ may assign somewhat different reign durations to the same kings. As a result, another dynasty of annals $b = (b_1, b_2, \dots, b_k)$ will appear. Thus, the same actual dynasty $M$, described in different chronicles, may be depicted therein as different dynasties of annals $a$ and $b$. The question is that of how great resulting distortions are? In this case, errors and objective difficulties impeding precise determination of the actual duration of rule play a significant part. We describe the basic types of errors below.

Let us formulate a statistical model, or a hypothesis, which we agree to call *the small distortions principle*.

### The small distortions principle for the reign durations.

If the two dynasties of annals $a$ and $b$ are "slightly" different, they refer to the same actual dynasty $M$, i.e., these are two versions of its descriptions in different chronicles. We call such dynasties of annals *dependent*.

On the contrary, if the two dynasties of annals $a$ and $b$ refer to two different actual dynasties $M$ and $N$, they differ "considerably". We call them *independent*.

We shall call the remaining pairs of dynasties *neutral*.

In other words, according to this hypothetical model, *different chroniclers would distort the same actual dynasty "slightly" when writing their chronicles*. In any case, the resulting differences proved to be smaller "on the average" than those existing between evidently different, or independent, actual dynasties.

The hypothesis or the model formulated above requires an experimental verification. In case of its validity, an important and by no means obvious quality is revealed, one that characterizes the activity of ancient chroniclers. Namely, *the dynasties of annals that appeared in the description of the same actual dynasty differ from one another and from their prototype less than truly different actual dynasties do*.

Is there a natural numerical coefficient, or a measure $c(a, b)$, computed for each pair of dynasties of annals $a$ and $b$ and possessing the quality of being "small" for dependent dynasties and, on the contrary, "large" for independent ones? In other words, this coefficient must distinguish between the dependent and independent dynasties. We have discovered such coefficient.

It turns out that, in order to evaluate the "proximity" of the two dynasties $a$ and $b$, it is possible to introduce the numerical coefficient $c(a, b)$, similar to the coefficient $PACY = p(X, Y)$ as described above. This coefficient $c(a, b)$ also stands for probability. Let us first describe a rough idea of determining the coefficient $c(a, b)$. The dynasty of annals may be conveniently presented in the form of a graph, with the number of kings on the horizontal axis, and the duration of their reigns on the vertical axis. We will say that dynasty $q$ "is similar" to the two dynasties $a$ and $b$, if the graph of dynasty $q$ differs from the graph of dynasty $a$ no more than the graph of dynasty $b$ differs from the graph of dynasty $a$. See details below in [904], [1137], [885], [886]and [884].

The part that dynasties "similar" to dynasties $a$ and $b$ constitute in the set of all dynasties is assumed as $c(a, b)$. In other words, we calculate the ratio:

$$\frac{\text{quantity of dynasties "similar" to a and b}}{\text{total quantity of dynasties described in the chronicles}}$$

Chroniclers may determine the reign durations of kings with an error. We actually extract only their approximate values from the chronicles. It is possible to describe the mechanisms of probability resulting in such errors mathematically. Furthermore, we considered two additional errors that the chroniclers might have possibly made: the permutation of two successive kings and the replacement of these two successive kings by one "king" with a summary duration of rule.

The coefficient $c(a, b)$ may be called *PACD,* i.e., Probability of Accidental Coincidence of Dynasties *a* and *b*.

### 4.2. The statistical model

Let us now provide a formal definition of the coefficient $c(a, b)$, designating the set of all actual dynasties with the length *k*, i.e., consisting of *k* sequential kings, as *D*. We will actually have to denote as set *D* those historical dynasties the information about which is available to us from the preserved historical chronicles. We have compiled an almost complete list of all such dynasties based on a large number of different chronological tables listed below. On the basis of these tables, we composed a list of all groups of 15 successive kings, who, according to the Scaligerian chronology, had ruled within the range of 4000 B.C. – 1900 A.D. in Europe, the Mediterranean, the Middle East, Egypt, and Asia.

Each dynasty of annals may be depicted as a vector in *k*-dimensional Euclidean space $R^k$. In our specific experiment we assumed $k = 15$, q.v. above. We consider two dynasties essentially different if the number of kings, or actual rulers simultaneously listed in both dynasties does not exceed $k/2$, or a half of the members of the entire dynasty. Two randomly chosen real dynasties may intersect, have common members, since we may declare, at our own discretion, one or another king as "the progenitor of a dynasty". Along with dependent and independent dynasties, there also exist "intermediate" or "neutral" pairs of dynasties, in which the number of common kings, or actual rulers, exceeds $k/2$ "(although the dynasties aren't dependent). It is clear that if the total number of dynasties in question is large, the quantity of intermediate or neutral pairs of dynasties is relatively small. Therefore, primary attention should be paid to dependent and independent pairs of dynasties.

The small distortions principle as formulated above means that in practice, "on the average", chroniclers made insignificant mistakes, which means that they would not distort actual numerical data greatly.

Let us now discuss the errors most frequently made by chroniclers in calculating the reign durations of ancient kings. We found these three types of errors while working on a large number of actual historical texts. These particular errors proved to most frequently result in the distortion of actual durations of rules of kings.

*Error one.* The permutation or confusion of two adjacent kings.

*Error two.* The replacement of two kings by one, whose duration of rule equals the sum of durations of both rules.

*Error three.* Inaccuracy in calculating the very reign duration per se. The longer the duration, the greater error the chronicler would usually make in its determination.

These three types of errors may be described and simulated mathematically. Let us begin with errors *(1)* and *(2)*. We shall examine a dynasty $p = (p_1, p_2, \dots, p_k)$ from the set *D*. We shall call vector $q = (q_1, q_2, \dots, q_k)$ *a virtual variation* of vector (dynasty) *p*, and designate it as $q=vir(p)$, if each coordinate $q_i$ of vector *q* is derived from coordinates of vector *p* in one of the two following procedures *(1)* and *(2).*

*(1)* Either $q_i = p_i$ (the coordinate does not change), or $p_i$ coincides with $p_{i-1}$, or $p_i$ coincides with $p_{i+1}$, i.e., with one of the "adjacent coordinates" of vector *p*.

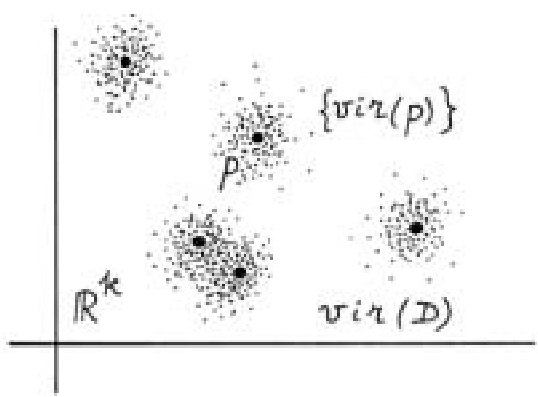


Fig. 5.29. Each *p* dynasty spawns a certain set *vir*(*p*) of virtual dynasties. They are represented geometrically as "clouds", or "globular clusters" surrounding the point *p* in space.

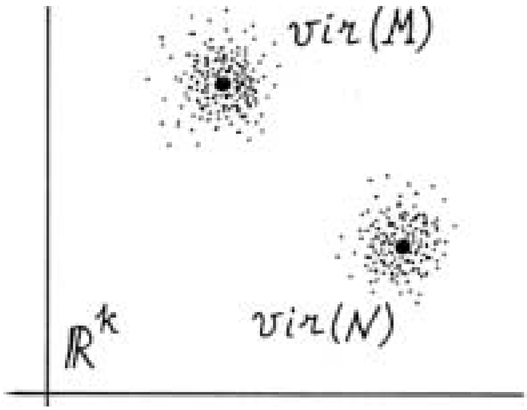


Fig. 5.30. "Globular clusters" *vir*(*M*) and *vir*(*N*) corresponding to two *a priori* independent and different dynasties *M* and *N* that are separated by a considerable distance.

*(2)* Either $q_i = p_i$, or $q_i$ coincides with the number $p_i + p_{i+1}$.

It is clear that each such vector (dynasty) $q$ may be considered as a dynasty of annals, resulting from an actual dynasty $p$ by a "reproduction thereof" due to errors *(1)* and *(2)* made by chroniclers. In other words, we take each real dynasty $p = (p_1, p_2, \dots, p_k)$ from the list $D$ and apply "disturbances" *(1)* and *(2)* to it, which means that we either swap places of two adjacent numbers $p_i$ and $p_{i+1}$, or substitute a certain number $p_i$ by the sum $p_i + p_{i+1}$, or sum $p_{i-1} + p_i$. For each number $i$, we use the above operations just once, that is, we do not consider "long iterations" of operations at the same place $i$. As a result, we obtain a certain number of virtual dynasties $\{q = vir(p)\}$ from one dynasty $p$. The quantity of such virtual dynasties is easy to calculate.

Thus, each "point" from set $D$ is "multiplied" and generates a certain set of "virtual points" surrounding it, a "surrounding cloud", or "globular cluster", fig. 5.29. We may come across some of the obtained virtual dynasties in a certain chronicle (in this case they will be dynasties of annals), while others remain just "theoretically possible", or "virtual".

By uniting all virtual dynasties obtained from all actual dynasties $p$, which compose our list of dynasties $D$, we obtain a certain set $vir(D)$, i. e., "a cloaking cloud" for the initial set of dynasties $D$.

Thus, for each actual dynasty $M$ the set of dynasties of annals describing it can be pictured as a "globular cluster" $vir(M)$. Let us now consider the two actual dynasties $M$ and $N$. If the small distortions principle formulated by us is accurate, then the globular clusters $vir(M)$ and $vir(N)$ corresponding to two *a priori* independent, different actual dynasties $M$ and $N$ do not intersect in space $R^k$, which means that they must be arranged sufficiently far from each other, q.v. in fig. 5.30.

Now let $a$ and $b$ stand for two certain dynasties from set $vir(D)$, for example, two dynasties of annals, q.v. in fig. 5.31. We would like to introduce a certain quantitative measure of proximity between two dynasties, or "measure the distance between them" – estimate how distant they are from each other, in other words, the easiest method would be as follows. Regarding both dynasties as vectors in space $R^k$, it would be possible just to take the Euclidean distance between them, or calculate the number $r(a,b)$, the square of which assumes the form of

$$(a_1 - b_1)^2 + \dots + (a_k - b_k)^2.$$

However, numeric experiments with specific dynasties of annals show that this distance does not make it possible to confidently separate dependent and in-

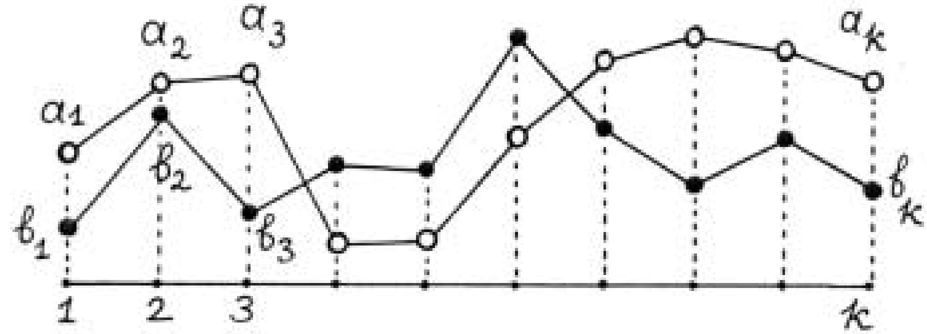


Fig. 5.31. A demonstrative visual representation of the reign lengths of dynasties *a* and *b* as graphs.

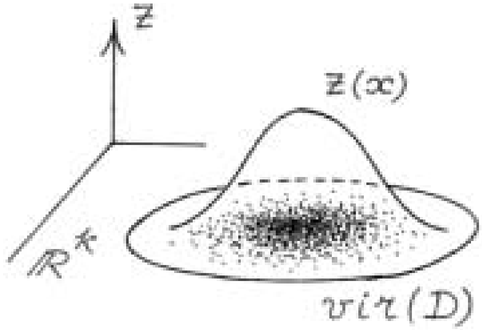


Fig. 5.32. A density function demonstrating the distribution of points pertinent to the set *vir(D)*.

dependent pairs of dynasties. In other words, such distances between *a priori* dependent dynasties of annals, and those between *a priori* independent ones, prove to be comparable to each other. They appear to have "the same order of magnitude".

Moreover, it is impossible to determine the "similarity" or "dissimilarity" of two dynasties, or, to be more precise, graphs of their rule, "at a glance". Visual similarity of two graphs can indicate nothing. It is possible to give examples of *a priori* independent dynasties, the graphs of rule of which prove to be "very similar", although there will be no actual dependence. It turns out that visual proximity can easily lead to confusion in this problem. A reliable quantitative estimation is necessary, one that would eliminate unsteady subjective considerations like "similar" or "not similar".

Thus, the aim is to explain whether such a natural measure of proximity does exist in general within a set of all virtual dynasties, which would make it possible to confidently separate dependent dynasties from independent ones, or make the "distance" between *a priori* dependent dynasties "small", and the "distance" between *a priori* independent dynasties "large". Moreover, these "small" and "large" values should be essentially different from one another, for example, by one or several orders of magnitude.

Such a measure of proximity, or "distance between dynasties", appears to actually exist. We will now turn to the description of this coefficient *c(a, b)*.

Thus, we plotted a set of dynasties *D* in space $R^{15}$. Two most typical errors usually committed by chroniclers were simulated. Each dynasty of the set *D* was subjected to disturbances of types *(1)* and *(2)*. In this case, each point from *D* multiplied into several points, which led to the increase of the set. We designated the set obtained as *vir(D)*. The set *vir(D)* turned out to consist of approximately $15 \times 10^{11}$ points.

We will consider "dynastic vector *a*" to be a random vector in $R^k$, passing through the set *vir(D)*. Then, on the basis of the set *vir(D)* we can create a probability density function *z*. With this aim in mind, the entire space $R^{15}$ was divided into standard cubes of sufficiently small size, so that no point of the set *vir(D)* would fall on the boundary of any cube. If *x* is an internal point of a cube, then we may assume that

$$z(x) = \frac{\text{the number of points from the set vir(D) falling into the cube}}{\text{the total quantity of points in the set vir(D)}}.$$

It is clear that for a point *x*, which lies on a boundary of any cube, it is possible to consider $z(x) = 0$. Function *z(x)* reaches its maximum in the area of especially high concentration of dynasties from the set *vir(D)*, and it drops to zero where there are no points of set *(D)*, fig. 5.32. Thus, the graph of function *z(x)* clearly shows how the set of virtual dynasties *vir(D)* is distributed within space $R^k$, – in other words, where this set is "thick", "dense", and where it is rarefied.

Now we are given two dynasties

$$a = (a_1, \dots a_k) \text{ and } b = (b_1, \dots ,b_k),$$

and we want to estimate how close or distant they are. Let us plot a *k*-dimensional parallelepiped *P'(a, b)* with its center in point *a*, which has as diagonal vector *a-b*, fig. 5.33. If we project the parallelepiped *P'(a, b)* on the *i*-coordinate axis, we will obtain a segment with the ends

$$[a_i - |a_i - b_i|, a_i + |a_i - b_i|].$$

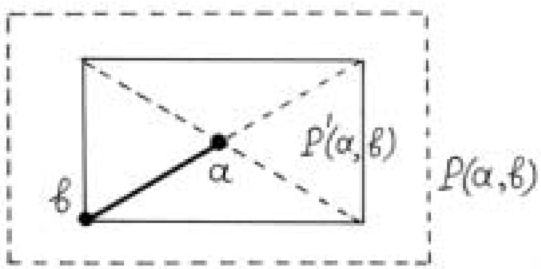


Fig. 5.33. Parallelepipeds *P'*(*a*, *b*) and *P*(*a*, *b*).

As a preliminary coefficient $c'(a, b)$ we will assume the number

$$c'(a, b) = \frac{\text{the number of points of the set } vir(D) \text{ falling in } P'(a, b)}{\text{the total number of points in the set } vir(D)}.$$

It is clear that number $c'(a, b)$ is the integral of density function $z(x)$ along the parallelepiped $P'(a, b)$.

The meaning of this preliminary coefficient $c'(a,b)$ is clear. It is natural to call dynasties, or vectors of $vir(D)$, falling into parallelepiped $P'(a, b)$, "similar" to dynasties $a$ and $b$. In fact, each of such dynasties is located no further from dynasty $a$ than dynasty $b$ is located from dynasty $a$. Consequently, as a measure of proximity of two dynasties $a$ and $b$, we take the part of dynasties "similar" to $a$ and $b$ in the set of all dynasties $vir(D)$.

However, such coefficient $c'(a, b)$ is not sufficiently good yet, since it does not consider the circumstance that the chroniclers could determine certain reign durations with a certain error, – the longer the rule, the larger the error. In other words, we have to take into account the error of chroniclers *(3)* discussed above.

Let us switch to the simulation of error *(3)*. Let $T$ be duration of a reign. It is clear that the duration of rule may be considered a random variable determined for "the set of all kings". Let us designate the number of kings ruling for $T$ years as $g(T)$. In the paper [884] the author of the present book experimentally calculated this frequency histogram $g(T)$ (density of distribution of the indicated random value) on the basis given in ***Chronological Tables*** by J. Blair ([76]). Let us assume $h(T) = 1/g(T)$ and call $h(T)$ a function of the chroniclers' errors. The lower the probability that a random variable, or the duration of reign, assumes the value of $T$, the greater the error $h(T)$ in the determination of duration $T$. In other words, chroniclers calculate small, "short" reign durations better, and in doing so, make insignificant mistakes. On the contrary, a chronicler would calculate long reign durations, those encountered rather rarely, with a significant error. The longer the reign, the greater the possible error.

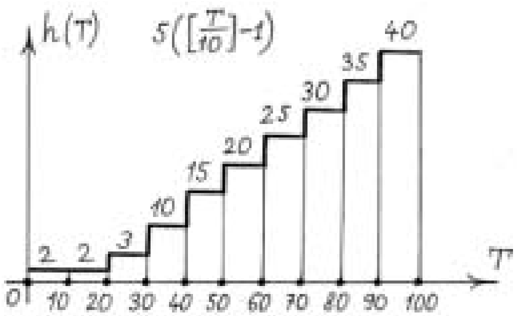


Fig. 5.34. A "scribe error function" calculated experimentally.

The errors function $h(T)$ for indicated probability density of a random value (reign duration) was determined experimentally ([884], p. 115). Let us divide the segment [0, 100] of integer axis $T$ into ten segments of identical length, namely:

[0, 9], [10, 19], [20, 29], [30, 39], … [90, 99].

Then it appears that:

$h(T) = 2$, if $T$ varies from 0 to 19,
$h(T) = 3$, if $T$ varies from 20 to 29,
$h(T) = 5\,([T/10]-1)$, if $T$ varies from 30 to 100.

The integer part of number $s$ is designated as $[s]$, fig. 5.34.

Let us now consider the errors of chroniclers in plotting the "environment" for point $a$. For this end, we expand the parallelepiped $P'(a, b)$, making it a larger parallelepiped $P(a, b)$, where point $a$ is again its centre, and segments with the ends

$$[a_i - |a_i - b_i| - h(a_i),\ a_i + |a_i - b_i| + h(a_i)]$$

are orthogonal projections thereof on the coordinate axes.

It is clear that the parallelepiped $P'(a, b)$ lies entirely within the large parallelepiped $P(a, b)$, q.v. in fig. 5.33. Vector $a - b + h(a)$ is the diagonal of this large parallelepiped, where vector $h(a)$ is

$$h(a) = (h(a_1), \dots , h(a_k)).$$

It is possible to name it *the vector of chroniclers' errors*.

Thus, we simulated all three basic errors that the chroniclers would make while calculating reign durations. As the final coefficient $c(a, b)$ measuring the proximity or distance from each other of two dynasties $a$ and $b$, we assume the following number:

$$c(a, b) = \frac{\text{the number of points from the set } vir(D) \text{ falling in } P(a, b)}{\text{the total number of points in the set } vir(D)}.$$

It is clear that the number $c(a, b)$ is the integral of density function $z(x)$ along the parallelepiped $P(a, b)$. In fig. 5.35, the number $c(a, b)$ is symbolically pre-

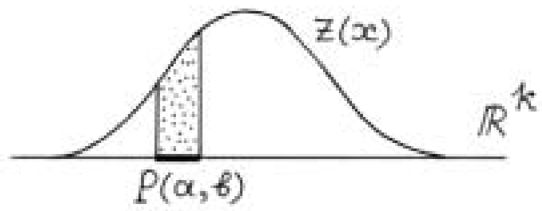


Fig. 5.35. Coefficient *c(a, b)* presented as the volume of a prism, or an integral of the function *z(x)* along the parallelepiped *P(a, b).*

sented as the volume of a prism with parallelepiped *P(a, b)* as its base, and limited on top by the graph of the function *z*. Number *c(a, b)* may, if desired, be interpreted as the probability that a random "dynastic vector" distributed in space $R^k$ with density function *z* proves to be at a distance from point *a*, keeping within the distance between points *a* and *b*, with the error *h*(*a*) taken into account. In other words, the random "dynastic" vector distributed with the density function *z* falls into the environment *P(a, b)* of point *a* with the "radius" $a - b + h(a)$.

It is evident from the above that the role of dynasties *a* and *b* in the calculation of the coefficient *c(a, b)* is not the same. Dynasty *a* was placed into the centre of parallelepiped *P(a, b),* and dynasty *b* determined its diagonal. Certainly, it was possible "to grant equal rights" to dynasties *a* and *b*, likewise the preceding coefficient *p(X, Y)*. In other words, it is possible to change the positions of dynasties *a* and *b*, calculate coefficient *c(b, a),* and then obtain the arithmetic mean value of numbers *c(a, b)* and *c(b, a)*. We refrained from this for two reasons. Firstly, as certain experiments have shown, replacement of coefficient *c(a, b)* by its "symmetric analogue" does not actually change the obtained results. Secondly, in certain cases dynasties *a* and *b* may actually have unequal rights in the sense that one of them may be the original, and the second merely its duplicate, a phantom reflection. In this case it is natural to place dynasty *a*, which claims to be the original, in the centre of the parallelepiped, and consider the "phantom reflection" *b* a "disturbance" of dynasty *a*. The resulting differences between coefficients *c(a, b)* and *c(b, a),* albeit small, may serve as useful material for further, more complex research, which has not been performed yet.

## 4.3. Refinement of the model and the computation experiment

The small distortions principle as formulated above was checked on the basis of coefficient *c(a, b)*.

1) For verification purpose we used ***Chronological Tables*** by J. Blair ([76]) containing virtually all basic chronological data from the Scaligerian version of the history of Europe, the Mediterranean, the Middle East, Egypt, and Asia allegedly from 4000 B.C. to 1800 A.D. This data was then complemented with lists of rulers and their reign durations taken from other tables and monographs, both mediaeval and contemporary. Let us mention the following books here, for example: C. Bemont, G. Monod ([64]), E. Bickerman ([72]), H. Brugsch ([99]), A. A. Vasilyev ([120]), F. Gregorovius ([195] and [196]), J. Assad ([240]), C. Diehl ([247]), F. Kohlrausch ([415]), S. G. Lozinsky ([492]), B. Niese ([579]), V. S. Sergeyev ([766] and ([767]), ***Chronologie égiptienne*** ([1069]), F. K. Ginzel ([1155]), L. Ideler ([1205]), ***L'art de vérifier les dates des faits historiques*** ([1236]), T. Mommsen ([1275]), Isaac Newton ([1298]), D. Petavius ([1337]), J. Scaliger ([1387]).

2) As we have already noted, by dynasty we understand a sequence of actual rulers of the country, irrespectively of their titles and kinship. Subsequently, we will sometimes refer to them as kings for the sake of brevity.

3) The existence of co-rulers sometimes makes it difficult to arrange dynasties into a sequence. We accepted the simplest principle of ordering – by the average reign durations.

4) We will call the sequence of numbers showing the reign durations of all rulers over the course of the entire history of a certain state (where the length of a sequence is not limited *a priori*) *a dynastic current*. Sub-sequences obtained by neglecting some of *co-rulers* will be called *dynastic jets*. Each jet is to be *even*, which means that middles of periods of rule must increase monotonically. A dynastic jet must also be *complete*, or cover the entire historical period included in the given flow without gaps or lapses; reign period superpositions are in order here.

5) In actual situations the above requirements may be somewhat disrupted for natural reasons, – for example, one or several years of interregnum may be

missing in a chronicler's story, – therefore insignificant *gaps* have to be acceptable. We only allowed gaps with durations not exceeding one year. Furthermore, while analysing dynastic currents and jets, the possibility of authentic picture distortion as a result of abovementioned errors *(1)*, *(2)*, and *(3)*, made by chroniclers – must be constantly kept in mind.

6) Another reason for the distortion of a clear formal picture lies in the fact that the beginning of a king's reign is sometimes hard to determine for certain. For example, should we start counting from the moment of actual accession, or from the moment of formal inauguration? Different tables give diverse variants of the beginning of rule of Friedrich II: 1196, 1212, 1215, and 1220 A.D. At the same time, usually there is no problem to determine the end of a rule – in most cases, the death of a king. Thus, a need arises for the "bifurcation", or even a review of the three versions of a king. Fortunately, in practice larger numbers of versions are exceptionally rare. All these versions were included in a general dynastic current. In doing so, not one single jet under research should have contained two different versions of the same reign.

7) A complete list *D* of all dynasties of annals with the length of 15 – i.e., a list of all dynasties of 15 successive kings – was made for all states of the above-indicated geographical regions on the basis of chronological data that we collected from the Scaligerian version. Moreover, every king could appear in several 15-member dynasties, i. e., dynasties may "overlap". Let us enumerate the basic dynastic currents that underwent statistical analysis. They are: the bishops and popes in Rome, patriarchs of Byzantium, Saracens, high priests in Judah, Greek-Bactrians, exarchs in Ravenne, pharaoh dynasties of Egypt, the mediaeval dynasties of Egypt, dynasties of Byzantium, the Roman empire, Spain, Russia, France, Italy, Ottoman = Ataman empire, Scotland, Lacedaemon, Germany, Sweden, Denmark, Israel, Judah, Babylon, Syria, Portugal, Parthia, the kingdom of Bosporus, Macedonia, Poland, England.

8) Having applied disturbances of types *(1)* and *(2)*, see above, to list *D* of dynasties of annals, we turned out to have obtained approximately $15 \times 10^{11}$ virtual dynasties, i.e., the set *vir(D)* appears to contain approximately $15 \times 10^{11}$ points.

## 4.4. Result of the experiment: coefficient *c(a, b)* positively distinguishes between the dependent and independent dynasties of kings

Computation experiment performed in 1977-1979 that M. Zamaletdinov, P. Puchkov, and yours truly performed together confirmed the small distortions principle. Namely, the number *PACD=c(a, b)* turned out to never exceed $10^{-8}$, and usually vary from $10^{-12}$ to $10^{-10}$, for *a priori* dependent dynasties of annals *a* and *b*. In probabilistic interpretation, it means that if we examine the observed proximity of two dependent dynasties of annals as a random event, then its probability is small, such event is exceptionally rare, since only one of hundred billion chances occurs.

It further appeared that if two dynasties of annals *a* and *b* refer to two *a priori* different real dynasties, coefficient *PACD = c(a, b)* "is substantially larger" – namely, never less than $10^{-3}$, or "large". Likewise, in the case of coefficient *p(X, Y)*, we are certainly not interested absolute values of *PACD = c(a, b)* but, rather, the difference of several orders of magnitude between the "dependent zone" and the "independent zone", q.v. in fig. 5.36.

Thus, with the aid of coefficient *PACD* it was possible to discover the essential difference between *a priori* dependent and *a priori* independent dynasties of annals.

## 4.5. The method of dating the royal dynasties and the method detecting the phantom dynastic duplicates

And so, the coefficient *c(a, b)* helps us to distinguish between dependent and independent pairs of dynasties of annals with reasonable certainty. The important experimental condition is that the mistakes of chroniclers are not "too grave". In any case, their errors are substantially less than the value distinguishing between independent dynasties.

This makes it possible to propose a new method of recognizing dependent dynasties of annals and a dating procedure for unknown dynasties within the framework of the experiment performed. Likewise in the paragraph above, for an unknown dynasty *d* we calculate the coefficient *c(a, d)*, where *a* denotes known and already dated dynasties of annals. Let us assume that we have discovered dynasty *a,* for which the coefficient *c(a, d)* is small, that is to say, it does not exceed $10^{-8}$. This allows us to say that dynasties *a* and *d* are dependent with the probability of 1 – *c(a, d)*, – i.e., dynasties of annals *a* and *d* obviously correspond to one actual dynasty *M,* the dating of which is already known to us. Thus, we can date the dynasty of annals *d*.

This procedure was tested on mediaeval dynasties with a known dating. The efficiency of the procedure was completely confirmed ([904] and [908]).

The same method makes it possible to reveal phantom duplicates in the "Scaligerian textbook on history". Namely, if we find two dynasties of annals *a* and *b*, for which coefficient *c(a, b)* does not exceed $10^{-8}$, this allows us to assume having just seen two copies, or two versions describing *the same* actual dynasty *M* multiplied on the pages of different chronicles, and then placed into different parts of the "Scaligerian textbook".

Let us reiterate that any conclusions or hypotheses appealing to "similarities" or on the contrary, "dissimilarities" of dynasties may be considered sensible only when based on extensive numeric experiments, similar to the ones performed by us. Otherwise, vague subjective considerations hardly worthy of being discussed may surface.

## 5. THE FREQUENCY DAMPING PRINCIPLE
### The method of ordering of historical texts in time

The frequency damping principle, and a method based on it, was proposed and developed by the author in [884], [886], [888], [1129], [891], [895], [898], [901] and [1130].

The present method makes it possible to find a chronologically correct order of separate text fragments, reveal duplicates therein on the basis of analysing, or the sum total of proper names mentioned in the text. As in the foregoing procedures, we aim at creating a method of dating based on numeric, or *quantitative* characteristics of texts, not necessarily requiring the analysis of the semantic content of texts, which may be fairly ambiguous and vague. If a document mentions any "famous" characters previously known to us, that are described in other chronicles already dated, it allows us to date the events described therein. However, if such identification does not immediately succeed, and, furthermore, if the events of several generations with a large quantity of previously unknown characters are described, then the task of establishing the identity of characters with the previously known ones becomes more complicated. For the sake of brevity, let us call a text fragment describing events of one generation "a generation chapter".

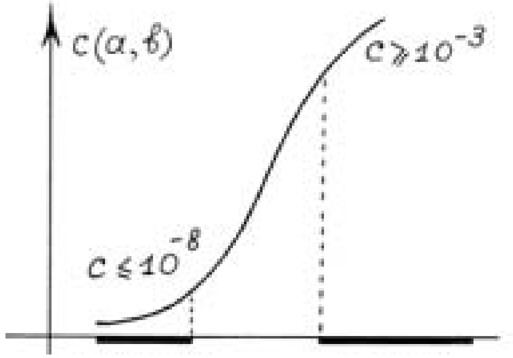


Fig. 5.36. Coefficient *c (a, b)* allows to differentiate between the dependent and the independent dynastic pairs.

We will consider an average duration of one "generation" to be the average reign duration of actual kings reflected in the chronicles available to us. This *average reign duration*, calculated by the author of this book while working on Blair's chronological tables ([76]), proved to be equal 17.1 ([884]).

While working with actual historical texts, one may sometimes come across a problem of separating "generation chapters" contained therein. In such cases we restricted ourselves to an approximate division of a text into successive fragments. Let chronicle *X* describe the events of a sufficiently large time interval *(A, B)*, during which at least several generations of characters have changed. Let chronicle *X* be divided into "generation chapters" *X(T)*, where *T* is the ordinal number of a generation described in fragment *X(T)* in the numeration of "chapters" fixed in the text.

The question arises of whether those "generation chapters" are *correctly* numbered, as ordered in the chronicle. Or, if this numeration is lost or doubtful, *how does one restore it*? In other words, how does one correctly arrange the "chapters" related to each other temporally? For the overwhelming majority of actual historical texts, the following formula appears to apply: *full name = character.* It means the following:

Let a time interval described by a chronicler be sufficiently long – for example, several decades or centuries. Then, as we have tested during the analyses of a large collection of historical documents, in the overwhelming majority of cases, *different characters* have *different full names*. A full name may consist of several words, for example, Charles the Bald. In other words, *the number of different persons with identical full names is negligibly small in comparison with the number of all characters*. This is correct for several hundred historical texts that we investigated, referring to Rome, Greece, Germany, Italy, Russia, England, etc. This is not surprising. In fact, a chronicler is interested in distinguishing between different characters in order to avoid confusion, and the easiest method to attain this is to assign different full names to different persons. This simple psychological circumstance is confirmed by calculations.

Let us now formulate *the frequency damping principle* describing a chronologically correct order of "generation chapters".

With the correct numeration of "generation chapters" in place, a chronicler *passing from descriptions of one generation to the next one changes characters as well.* Namely, describing the generations preceding the generation *Q*, he says nothing about the characters of this generation, since they have not been born yet. Then, in his description of generation *Q*, the chronicler mostly speaks about the characters of this generation, since the events described are directly connected to them. Finally, passing to the description of subsequent generations, the chronicler mentions the previous characters in decreasing frequency, since he describes new events, the characters of which replace the ones departed.

It is important to emphasize here that we do not imply any separate names, but rather a *complete reservoir of all names* used in generation *Q*.

Briefly, our model is formulated as follows. *Every generation gives birth to new historic characters. Upon the change of generations, these characters change, too.*

In spite of its seeming simplicity, this principle proved to be useful in the creation of *the method of dating*. The frequency damping principle has an equivalent re-definition. Since the characters are virtually unambiguously determined by their full names (name = character), we will study the reservoir of all full names contained in the text. We will usually omit the term "full", while constantly implying it. Moreover, an overwhelming majority of historical names proved to be "simple", consisting of one word. Therefore, while processing large historical texts with a significant fund of names, it is possible to consider just the "elementary name units", dividing occasional full names into separate words they consist of.

Let us examine a group of *all* names appearing in the text for the first time in "generation chapter" *Q*. Let us agree to call these names *Q-names,* and corresponding characters *Q-characters.* We will designate the number of *all* references to *all* of these names in this "chapter", with multiplicities, by *K(Q, Q).* Let us then calculate how often the same names are mentioned in "chapter" *T*. Let us designate the resulting number as *K(Q, T)*. If the same name is repeated several times, or with a multiplicity, then *all* those mentions shall be calculated. Let us plot a graph placing the number of "chapters" along the horizontal axis, and numbers *K(Q, T)* along the vertical, where *Q* is a constant, and *T* is a variable, and obtain a separate graph for each *Q*. The frequency damping principle is then formulated as follows.

With the chronologically correct numeration of "generation chapters", every graph *K(Q, T)* has to assume the following form: *to the left of point* Q, *the graph equals zero; point* Q *is the absolute maximum of the graph; then the graph incrementally decreases, fading out more or less evenly*, q.v. in fig. 5.37.

We shall call the graph in fig. 5.37 an ideal one. The formulated principle must be tested experimentally. If it is accurate, and the "chapters" in a chronicle are chronologically correctly streamlined, then all experimental graphs must be close to the ideal one. The undertaken experimental verification has completely confirmed the frequency damping principle ([904] and [908]). Let us give some typical examples.

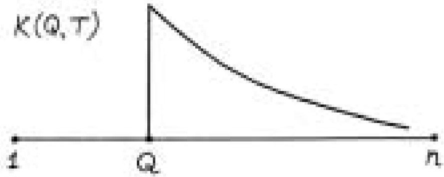


Fig. 5.37. The theoretical "ideal" frequency damping graph.