RANDOM WALKS

E. B. DYNKIN AND V. A. USPENSKII

D.C. HEATH AND COMPANY

Random Walks

PART THREE OF *MATHEMATICAL CONVERSATIONS*

E. B. Dynkin and V. A. Uspenskii

Translated and adapted from the first Russian edition (1952) by

NORMAN D. WHALAND, JR., *and* OLGA A. TITELBAUM

SURVEY OF
RECENT EAST EUROPEAN MATHEMATICAL LITERATURE

A project conducted by

ALFRED L. PUTNAM *and* IZAAK WIRSZUP

Department of Mathematics,
The University of Chicago, under a
grant from the National Science Foundation

D. C. HEATH AND COMPANY BOSTON

Library of Congress Catalog Card Number: 63-19838

Printed in the United States of America
(6 F 3)

PREFACE TO THE AMERICAN EDITION

THIS BOOKLET is a translation of Part Three of *Mathematical Conversations* by E. B. Dynkin and V. A. Uspenskii, published as Number 6 in the Russian series, *Library of the Mathematics Circle*. The book is based on the material covered during the academic years 1945–46 and 1946–47 in one of the sections of the School Mathematics Circle at Moscow State University. One of the authors was the instructor of the section, and the other was a participant.

The primary aim was not so much to impart new information as to teach an active, creative attitude toward mathematics. The most successful topics took shape only as the work progressed. A series of consecutive meetings was devoted to each topic. A meeting would usually begin with problems whose formulation required no new concepts, but whose solution led the students directly into a new area of inquiry. These problems would sometimes be solved during the meeting, but more often they were left as homework. At the next meeting, the instructor would discuss the solutions of the problems and then use them as a basis for generalizations. Whenever possible, the material was presented in sequences of related problems.

This booklet presents one of the topics, *Random Walks,* in a considerably revised and expanded form. Like the discussions on which it is based, it retains the practice of interrupting the presentation with problems whose solutions are essential to what follows. To understand this material, the reader should be familiar with high school algebra.

INSTRUCTIONS FOR THE USE OF THIS BOOKLET

This booklet is devoted to a single topic, and it should therefore be read in order. Moreover, it is designed for the reader's active participation, and the problems form an organic part of the text. Most of the problems are grouped in sequences, each sequence forming a unit and building up to a final result contained in the last problem of the sequence. Sometimes the aim of a sequence of problems is not some definite result, but rather mastery of a new method.

Finally, a few of the problems are practice exercises, designed to help the reader master new concepts (for example, Problems 1–3).

Before attempting to solve a problem, the reader should examine all the problems in the given sequence. Solutions are provided following the Concluding Remarks, but it is recommended that the reader look at them *only* after he has tried to solve *all* the problems of a sequence. If he looks at the solutions too soon, they may set his mind working in a certain direction, but with independent thought he may arrive at new and original methods. The experience of the *School Mathematics Circles* has shown that sometimes simpler and more elegant solutions are found than those expected by the authors of the problems.

The reader may not always be able to solve all the problems of a sequence independently. If, after solving the first few problems, he should run into difficulties, he may find it helpful to read the solutions of the problems he has already solved. If these do not suggest an approach to the next problem, he should look at its solution, and then proceed to try the rest by himself. Eventually, he should read all the solutions, whether or not he has succeeded in solving the problems independently, as they have been carefully prepared, and many of them are accompanied by conclusions and remarks of a fundamental nature.

Although the problems here are basic, this is by no means merely an exercise book. The text is also important. The relation between problems and text differs in the various chapters. In some the essential ideas are set forth in the text, but in others they are in the problems, and the text merely introduces concepts and states results. The text and problems are always closely related and must be read in the order in which they appear in the book.

In conclusion, we advise the reader not to begrudge the time spent on solving the problems. Each sequence, indeed each problem, solved independently enlarges the arsenal of resources at his disposal. One idea arrived at independently is worth a dozen borrowed ones. Even if persistent attempts to solve a problem do not lead to success, the time is not spent in vain, as he will then see its solution in a new light. He can look for the reason for his failure and can discover the fundamental idea that leads to success.

CONTENTS

Random Walks

Introduction

There is a well-known board game called *Circus*. Two players take turns rolling a die, and each in turn moves his piece forward on a square board which is divided into 100 numbered squares. The piece is moved as many squares forward as the spots on the die indicate. At the beginning of the game, both pieces are placed on the first square. The winner is the one who first reaches the square numbered 100. There is one further rule: If a piece reaches a square with a red number, it moves to another square (either forward or backward) whose number is blue; this move is specified by the "circus act" marked on the board. Obviously, in this game the motion of the pieces depends not on the skill of the players, but rather on chance (assuming the die is rolled "fairly"). The motion of these pieces is one simple example of a *random walk*.

We give another example of a random walk. Two friends live in a city whose map is shown in Figure 1. They leave their house,

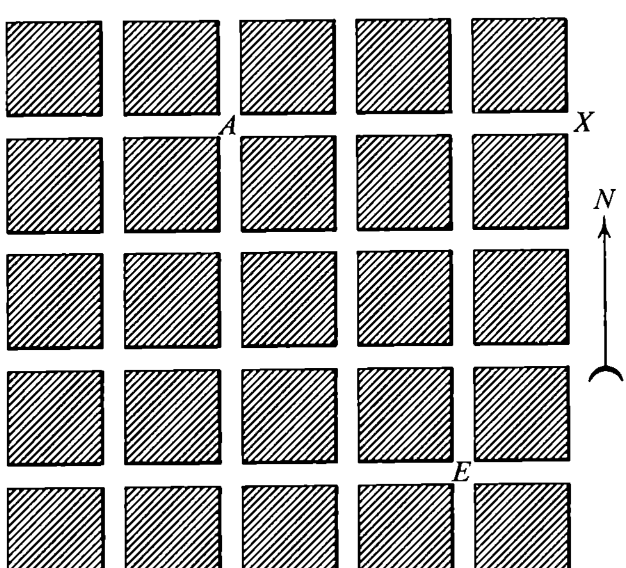

Fig. 1

which stands at the intersection A, and set out to go for a walk. However, they disagree as to the route they are to take; they agree that at each intersection, beginning with A, they will toss two coins and proceed north, east, south, or west, depending on which of the four possible tosses (heads-heads, heads-tails, tails-heads, tails-tails) comes up. Thus, they toss the coins and begin their walk at the intersection A; when they reach the following intersection, they again toss the coins and choose their further path according to the result. If they come to the edge of town (for example, to the point X), they will turn around and come back.

Cases similar to the examples just given (but much more complicated) are encountered in nature. Brownian motion will serve as an example: "If a light powder is suspended in water, and if a drop of this water, together with the particles contained therein, is placed under the microscope, one can observe that the particles appear to be alive, because they are in continuous zigzag motion." Here, the random path (Fig. 2) has an essentially more complicated character than in the previous example: in the first place, the particles can change the direction of their motion at any moment, while in the example of the walk in the city, this was possible only at intersections; in the second place, the previous example permitted only four possible directions of motion, whereas the particles can move in any direction.

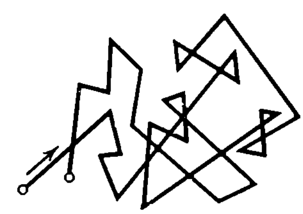

Fig. 2

In this section, we shall consider a few simple examples of random walks. We return to the board game described above and investigate the duration of the game. To solve this problem we consider a less complicated board (Fig. 3), one that has only 25 squares, and add the rule that a piece landing on square 24 must automatically return to the starting point. There are no other rules for moving from square to square, and the squares are all colored with the same color. How long does it take to play a complete game on this board? It is possible for a player to reach square 25 as early as the fourth move; this occurs if he rolls four successive sixes. On the other hand, if each player's piece repeatedly lands on square

21	22	23	24	25
20	19	18	17	16
11	12	13	14	15
10	9	8	7	6
1	2	3	4	5

Fig. 3

24, the game will not have ended after 1,000 rolls of the die. It can even happen that the game never ends at all; this occurs, for example, if both players roll the following sequence:

6, 6, 6, 5, 6, 6, 6, 5, 6, 6, 6, 5, 6, 6, 6, 5, 6, 6, 6, 5, ...

In this case, no bound can be given for the duration of the game. But try playing one game after another. You will be convinced that the game does end and, in fact, rather quickly. Just what is the trouble here?

We have already determined that it is impossible to give a definite bound for the duration of the game. The situation is basically altered, however, when an absolutely certain answer is not required. To clarify this last remark, let us consider some more examples.

EXAMPLE 1. Suppose we have an urn which contains 1,000 balls; suppose, also, that one of the balls is black, the others white. A ball is taken from this urn at random. Can it happen that the black ball is chosen?

Answer. It is certainly possible, but it is not at all likely.

EXAMPLE 2. Suppose a man who does not know Russian starts to type on a Russian typewriter. Is it possible that he thereby writes Pushkin's story, *The Captain's Daughter?*

Answer. Obviously, this is not absolutely impossible. However, it hardly appears to be a real possibility to anyone. (Indeed, the essays of two school children may also coincide word for word. But, as a rule, one quite correctly regards this as evidence that one of the essays was copied.)

EXAMPLE 3. According to the kinetic theory of gases, air is composed of a great number of molecules that are moving randomly. There is almost no interaction between the individual molecules; hence, the position of one molecule in space does not influence the positions of the other molecules. We imagine the space in the room we are in to be divided into an upper and a lower half. It is not impossible (this follows from the kinetic theory of gases) for all of the molecules of air in the room to go into the upper half, suffocating everyone in the room. This conclusion appears far-fetched to us. But does this mean that the kinetic theory is false?

Answer. It is to be hoped that the reader does not draw this conclusion, in view of the examples discussed above. The phenomenon that we have just presented is indeed not absolutely impossible; one can, however, view it as practically impossible. The basis for this conclusion is even more apparent in this example than in the previous example: there are incomparably more molecules in a room than there are letters in the story *The Captain's Daughter.*

Thus, we may quite correctly regard a highly unlikely occurrence as impossible for all practical purposes. Moreover, there are different degrees of unlikelihood. The unlikelihood of the occurrence of the event in the second example is much greater than that in the first, and the unlikelihood in the third example is much greater than that in the second.

We return again to our problem, which we now reformulate as follows: What *limit* can be estimated for the length of the game, so that a game whose length exceeds the estimated limit is practically impossible in the same degree as are the events of our examples?

How one solves this and similar problems will become clear to the reader after working through this book. We shall give only the answer here. If we estimate that the game ends after at most 200 throws, an error will be about as unlikely as the event in the first example. If the estimate is increased to 30 million, an error will be about as unlikely as the event in the second example. If the estimate is raised to 10^{25}, an error will be as unlikely as the event in the third example.

1. Probability

1. FUNDAMENTAL PROPERTIES OF PROBABILITY

First we shall learn how to calculate probabilities. Consider two urns with 100 balls in each. In the first urn, one ball is white, the other 99 black. In the second urn, 10 balls are white and 90 black. From which urn is one more likely to draw a white ball?[1] The reader will answer without hesitation: from the second. If we ask how many times greater this probability is, the reader will certainly answer that it is ten times as great. Suppose, now, that we have a third urn, in which all 100 balls are white. As before, we conclude that a white ball can be drawn from the third urn with 100 times the probability as from the first urn. The ball taken from the third urn will surely be a white ball. If we define the probability of this last, inevitable event to be the number 1, it follows from what has been said that the probability of drawing a white ball from the first urn is equal to $\frac{1}{100}$, and the probability of drawing a white ball from the second urn is $\frac{10}{100}$.

We consider the general case, in which the urn contains n balls, of which m are white. From the same considerations as above, we can conclude that the probability of a white ball being drawn from the urn is $\frac{m}{n}$. The urn problem is extraordinarily useful because many problems can be reduced to this form.

EXAMPLE 1. What is the probability that heads will come up on the toss of a coin? We consider an urn with one white and one black ball. Let the white ball correspond to heads, the black to tails.

Answer. Obviously, the probability that heads will come up is equal to the probability that one draws a white ball from our urn, and this equals $\frac{1}{2}$.

[1] Here it is assumed that the balls in the urns are completely uniform and well mixed and that one does not look when drawing; then, one has the same probability of drawing any ball.

EXAMPLE 2. What is the probability that one rolls a five with a die? The problem may be thought of as an urn with six balls, one of which is white. What is the probability that the white ball will be drawn?

Answer. $\frac{1}{6}$.

EXAMPLE 3. A domino is drawn at random from a box of dominos. What is the probability that there is a six on one end of this domino? We consider an urn with 28 balls, of which 7 (those corresponding to the 7 dominos that have a six on one of their halves) are white.

Answer. The probability that a domino of the desired kind is drawn is equal to the probability that a white ball is drawn from the urn, that is, $\frac{7}{28}$.

EXAMPLE 4. There are 5 red, 7 blue, and 13 black balls in an urn. What is the probability that either a red or a blue ball is drawn? There are 12 white and 13 black balls in a second urn. What is the probability that a white ball is drawn?

Answer. The probability that either a red or a blue ball is drawn from the first urn is equal to the probability that a white ball is drawn from the second urn, that is, $\frac{12}{25}$.

In general, a trial may have n equally probable outcomes, of which m yield a desired event A, the others yielding an undesired event. Each such trial is equivalent to drawing a ball from an urn containing n balls, of which m are white and the rest black. The occurrence of the event A has exactly the same probability as the drawing of a white ball from the urn, that is,

$$\frac{m}{n}.$$

DEFINITION. *The* probability *of an event A is equal to the number of possible favorable outcomes divided by the total number of possible outcomes.*

We denote the *probability* of the event A by

$$\mathbf{P}(A).$$

We now formulate the following properties of probability:

Property 1. If the event A *implies* the event B, that is, if each occurrence of the event A is followed by an occurrence of event B (or, the event B always occurs when the event A does), then

$$\mathbf{P}(A) \leq \mathbf{P}(B).$$

Property 2. If the events A and B are *mutually exclusive* (that is, it is impossible that both A and B occur), then

$$\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B), \tag{1}$$

where $A + B$ is understood to mean the event that consists of the occurrence of either A or B.

Property 3. If the events A and B are *exact opposites of each other,* that is, if the occurrence of A is the same as the nonoccurrence of B, then

$$\mathbf{P}(A) + \mathbf{P}(B) = 1. \tag{2}$$

Property 4. If the event E is *certain,* that is, if E must occur, then

$$\mathbf{P}(E) = 1.$$

Property 5. If the event O is *impossible,* that is, if O cannot occur, then

$$\mathbf{P}(O) = 0.$$

It is easy to obtain these properties from a consideraton of the urn problem, and the reader may derive them for himself.[1] We shall only clarify a few definitions: an example of *mutually exclusive* events is the drawing of a blue ball and the drawing of a red ball (when only one ball can be drawn from the urn); an example of *opposite results* is the tossing of a coin, where either heads or tails must come up. The drawing of a white ball from an urn that contains only white balls is a certain event; the drawing of a black ball from this urn would then be an impossible event.

Although many problems can be reduced to the urn problem, there are many (and among them the most interesting) that cannot be reduced to this problem. However, Properties 1–5 of probability are always true.

[1] We have already derived a special case of Property 2 in Example 4 on page 6. Here we restate the Example: A is the event of a red ball being drawn; B is the event of a blue ball being drawn; and $A + B$ becomes the event of either a red or a blue ball being drawn.

2. CONDITIONAL PROBABILITY

We now wish to become acquainted with so-called *conditional probability;* first, a few examples.

At recess, students of the first and second grades gather in the playground to play. Eleven pupils of the first grade take part, 8 boys and 3 girls, together with 6 pupils of the second grade, 2 boys and 4 girls. It is decided by lot who is to begin.[1] What is the probability that the lot falls on a first-grade student? To calculate this probability, the number of students of the first grade must be divided by the number of all participants. The result equals $\frac{11}{17}$.

We now assume that we know that the game will be started by a boy, and ask what influence this has on the probability of interest to us. We want to know what the probability is now for a member of the first grade to start the game. All 10 boys that participate in the game are equally likely to begin. Of these 10 boys, 8 are pupils in the first grade. Hence, the probability that a member of the first grade begins is, in this case, equal to $\frac{8}{10}$. We see that the probability has changed.

We have obtained the *conditional probability* that a member of the first grade begins, *on condition* that the game is begun by a boy.

DEFINITION. *The* conditional probability *of an event B given an event A is the probability that the event B will occur, if it is known that a previous event A is certain to occur.* It is denoted by

$$\mathbf{P}(B|A).$$

Problem 1. In the example above calculate the conditional probability that a member of the first grade begins, on condition that the game is begun by a girl.

In our example, let B denote "the game is begun by a member of the first grade" and A denote "a boy begins." As we calculated, $\mathbf{P}(B) = \frac{11}{17}$ and $\mathbf{P}(B|A) = \frac{8}{10}$. Hence, in this case $\mathbf{P}(B|A) \neq \mathbf{P}(B)$. The occurrence of the event A thus has a significant influence on the probability of the event B.

We add yet another property of probability to the five already enumerated in section 1.

[1] The drawing must guarantee that the probability of any participant beginning the game is the same. Among the possible forms of the drawing, that one is best which achieves this equality of probability most completely.

Property 6.

$$\mathbf{P}(AB) = \mathbf{P}(A)\,\mathbf{P}(B|A). \tag{3}$$

By AB, we are to understand the event in which both A and B occur.

We shall now indicate how one can derive this property. We shall verify Property 6 by means of our previous example of the students in the playground. The person who begins the game is determined by lot. As before, let A be the event "A boy begins" and B the event "A member of the first grade begins." Then, AB is the event "A boy of the first grade begins." Of the 17 possible outcomes, the event AB can occur in 8 ways (8 boys are pupils in the first grade); the event A can occur in 10 ways. Hence, $\mathbf{P}(AB) = \frac{8}{17}$ and $\mathbf{P}(A) = \frac{10}{17}$. We have already found that $\mathbf{P}(B|A) = \frac{8}{10}$; thus we have

$$\frac{8}{17} = \frac{10}{17} \cdot \frac{8}{10},$$

i.e.,

$$\mathbf{P}(AB) = \mathbf{P}(A)\,\mathbf{P}(B|A).$$

If A and B are independent events, neither the occurrence nor nonoccurrence of the event A has any effect on the probability of the event B; hence, the conditional probability $\mathbf{P}(B|A)$ of the event B on the condition A is equal to the unconditional probability $\mathbf{P}(B)$:

$$\mathbf{P}(B|A) = \mathbf{P}(B).$$

In this case, formula (3) takes the form

$$\mathbf{P}(AB) = \mathbf{P}(A)\,\mathbf{P}(B),$$

and we obtain:

Property 6*a*. If A and B are independent events, then

$$\mathbf{P}(AB) = \mathbf{P}(A)\,\mathbf{P}(B). \tag{4}$$

EXAMPLE. Successive flips of a coin are independent events. Hence, the probability of obtaining heads twice in a row is $\mathbf{P}$("heads come up on the first toss" and "heads come up on the second toss") $=$ $\mathbf{P}$("heads come up on the first toss") $\cdot$ $\mathbf{P}$("heads come up on the second toss") $= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Formulas (1) and (4), the formulas for addition and multiplication of probabilities, can easily be generalized to the case of more than two mutually exclusive or independent events. Let n events $A_1, A_2, \ldots, A_n$, any two of which are mutually exclusive, be given. Then

$$\mathbf{P}(A_1 + A_2 + A_3 + \cdots + A_n) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) + \cdots + \mathbf{P}(A_n). \quad (5)$$

We prove this formula for $n = 3$. The event A_3 is mutually exclusive of both A_1 and A_2; from this it clearly follows that A_3 and $A_1 + A_2$ are mutually exclusive. Then, by Property 2, we have the following formula

$$\mathbf{P}(A_1 + A_2 + A_3) = \mathbf{P}(A_1 + A_2) + \mathbf{P}(A_3). \quad (5')$$

But A_1 and A_2 are also mutually exclusive, so that

$$\mathbf{P}(A_1 + A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2). \quad (5'')$$

Formula (5) follows, for the case $n = 3$, from (5′) and (5″). Formula (5) can be proved analogously for arbitrary n.

If A_1 and A_2 are independent, and if A_1A_2 and A_3 are also independent, then

$$\mathbf{P}(A_1A_2A_3) = \mathbf{P}(A_1A_2)\mathbf{P}(A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3).$$

More generally, if A_1 and A_2 are independent, and, also, each of the pairs of events A_1A_2 and A_3, $A_1A_2A_3$ and $A_4, \ldots, A_1A_2A_3 \cdots A_{n-1}$ and A_n are independent, then

$$\mathbf{P}(A_1A_2A_3 \cdots A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3) \cdots \mathbf{P}(A_n). \quad (6)$$

For example, let A_k be the event that the kth toss of a coin is heads. Then, the probability that the coin comes up heads n times is given by formula (6).

$$\mathbf{P}(A_1A_2 \cdots A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2) \cdots \mathbf{P}(A_n) = \underbrace{\tfrac{1}{2}\cdot\tfrac{1}{2}\cdot \cdots \cdot\tfrac{1}{2}}_{n \text{ times}} = \frac{1}{2^n}.$$

This same number gives the probability that the coin comes up tails n times and, in general, that on n tosses any previously specified sequence of heads and tails comes up.

Problem 2. What is the probability that no six will come up on six rolls of a die?

Problem 3. We understand by p the probability that the target is hit on one shot. Calculate the probability that in n shots, one hits the target.

We mentioned in the introduction that events of small probability can, with good grounds, be regarded as practically impossible, and that there are varying degrees of unlikelihood. We shall make this more precise.

If one is to apply the methods of probability theory to the study of a phenomenon in nature, he must each time choose an arbitrarily small number ε, *the permissible probability of a deviation* (*an error*). If we have predicted a course of events by arguments based on probability theory, we must admit the possibility of error in our prediction and demand only that the probability of this error is not greater than ε. We start with the assumption that all events whose probability is smaller than ε are to be regarded as *practically impossible,* and that all events whose probability is greater than $1 - \varepsilon$ can be assumed to be *practically certain.*

DEFINITION. *The number ε is the* permissible degree (*or the* magnitude) of uncertainty *and the number* $1 - \varepsilon$ *is the* required degree (*or* magnitude) of certainty.

Obviously, the value of ε must be chosen for each individual problem according to the practical requirements on the correctness of the conclusions. Frequently used values of ε are 0.01, 0.005, 0.001, and 0.0001.

Let us clarify our definition and discussion by considering the example of the number of shots that are necessary for a single hit on a target (see Problem 3 above). Suppose that for each shot the probability of a hit is 0.2. How many shots must be made to hit the target once? It is clear that the number of shots cannot be given with absolute certainty. For it can happen that the target is hit on the first shot; at the same time, one cannot exclude the possibility that after 100 or 200 shots none have yet hit the target. Hence, we shall not seek after absolute certainty, but rather introduce a *permissible degree of uncertainty* ε. It appears entirely acceptable, for example, to set the value ε equal to 0.001. The statement that the target will have been hit after n shots is false with probability $(1 - 0.2)^n = (0.8)^n$.

We now choose n so that

$$(0.8)^n < 0.001.$$

The smallest value of n that satisfies this inequality is 31. (It is easy to calculate this with the aid of a table of logarithms.) Hence, the statement that the target will be hit at least once after 31 shots is false with a probability that does not exceed the permissible bound 0.001. Hence, under our requirements for the degree of certainty, we can say that it is practically certain that the target will be hit after 31 shots.

The number n varies with the required degree of certainty ε. The values of n for different degrees of certainty are compiled in the following table:

ε	n
0.01	21
0.005	24
0.001	31
0.0001	42

Example. Calculate the probability that no six comes up on infinitely many rolls of a die.

Solution. We first calculate the probability of the event B_n that no six has come up after n rolls. In the solution of Problem 2, page 10, we found $\mathbf{P}(B_6) = \left(\frac{5}{6}\right)^6$. Analogously, we find (by formula (6)),

$$\mathbf{P}(B_n) = \left(\frac{5}{6}\right)^n,$$

for arbitrary n. We denote by B the event of interest to us, that no six appears in an infinite sequence of rolls. If the event B occurs, then all of the events $B_1, B_2, \ldots, B_n, \ldots$ must occur. Hence, on the basis of the first property of probability:

$$\mathbf{P}(B) \leq \mathbf{P}(B_1) = \frac{5}{6},$$

$$\mathbf{P}(B) \leq \mathbf{P}(B_2) = \left(\frac{5}{6}\right)^2,$$

$$\cdots\cdots\cdots\cdots$$

$$\mathbf{P}(B) \leq \mathbf{P}(B_n) = \left(\frac{5}{6}\right)^n,$$

. . . .

The numbers $\frac{5}{6}, \left(\frac{5}{6}\right)^2, \ldots, \left(\frac{5}{6}\right)^n, \ldots$ are the terms of an infinite decreasing geometric progression; these terms will eventually be smaller than any predetermined positive number, provided n is sufficiently large.[1] Hence, $\mathbf{P}(B)$ also becomes smaller than any positive number, that is, $\mathbf{P}(B) = 0$. Hence, the probability that no six appears in an infinite sequence of rolls of a die is equal to zero.

The events that we have dealt with up to this time have either been impossible, or, if possible, had a probability greater than zero. Here we encounter for the first time an event whose probability is equal to zero and which, nevertheless, appears to be logically possible. We could not obtain this result if the probability of our event were calculated by the rule set forth at the beginning of this section, namely, as the ratio of the number of favorable events to the total number of all possible events.

The result that we have obtained as our solution to this exercise can be interpreted in the following way. However high we may wish the degree of certainty to be, a number of rolls can be given for which the six must come up at least once with this certainty.[2] This is the precise meaning of the statement that the six comes up with a certainty of 1 in infinitely many rolls.

3. THE FORMULA FOR COMPLETE PROBABILITY

DEFINITION. *A system of events* $A_1, A_2, A_3, \ldots, A_n$ *is called* complete *if at least one of the events must occur (in other words, if the event* $A_1 + A_2 + A_3 + \cdots + A_n$ *is certain).*

If the events $A_1, A_2, \ldots, A_n$ form a *complete* system and if they are mutually exclusive, then

$$\mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_n) = 1 \tag{7}$$

(this follows from formula (5) and Property 4).

[1] Proofs of this are found in the section on limits in many calculus textbooks.

[2] In fact, if we say that a six occurs in n rolls, we can err with a probability of $(\frac{5}{6})^n$. For sufficiently great n, the probability of an error can be made arbitrarily small.

Property 7. (The formula for complete probability.) Let a complete system of mutually exclusive events $A_1, A_2, A_3, \ldots, A_n$ be given. Then the probability of an arbitrary event B can be calculated by the formula

$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B|A_1) + \mathbf{P}(A_2)\mathbf{P}(B|A_2) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B|A_n). \qquad (8)$$

To prove this, note that since one of the events $A_1, A_2, \ldots, A_n$ occurs, the occurrence of the event B is equivalent to the appearance of one of the events $BA_1, BA_2, \ldots, BA_n$. Hence,

$$\mathbf{P}(B) = \mathbf{P}((A_1B) + (A_2B) + \cdots + (A_nB)).$$

Since the events $A_1B, A_2B, \ldots, A_nB$ are mutually exclusive (as $A_1, A_2, \ldots, A_n$ are mutually exclusive), we have, by formula (5),

$$\mathbf{P}(B) = \mathbf{P}(A_1B) + \mathbf{P}(A_2B) + \cdots + \mathbf{P}(A_nB).$$

Applying Property 6, we obtain

$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B|A_1) + \mathbf{P}(A_2)\mathbf{P}(B|A_2) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B|A_n).$$

Let us use this formula to calculate the probability that in the game we described on page 8 a member of the first grade begins. Here, B means that a member of the first grade begins, A_1 that a boy begins, and A_2 that a girl begins. We obtain

$$\mathbf{P}(A_1) = \frac{10}{17}, \ \mathbf{P}(A_2) = \frac{7}{17}, \ \mathbf{P}(B|A_1) = \frac{8}{10}, \ \mathbf{P}(B|A_2) = \frac{3}{7},$$

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(A_1)\mathbf{P}(B|A_1) + \mathbf{P}(A_2)\mathbf{P}(B|A_2) \\ &= \frac{10}{17}\cdot\frac{8}{10} + \frac{7}{17}\cdot\frac{3}{7} = \frac{8}{17} + \frac{3}{17} = \frac{11}{17}.\end{aligned}$$

This result agrees with the previous calculation.

Problem 4. Two players alternately toss a coin, and the one that first tosses heads wins. What is the probability that the game never ends? What is the probability that the first player wins? What is the probability that the other player wins?

Problem 5. A particle at point A (Fig. 4) can, in the next unit of time, remain at A with the probability p_{11}, or move to point B with the probability p_{12}. If it is at B, it can, in the next unit of time, remain at B with the probability p_{22} or move to point A with the probability p_{21}. What is the probability that the particle is at point A after n units of time, if at the initial moment it is:

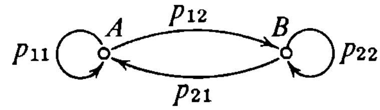

Fig. 4

(a) At point A? (b) At point B?

Many different phenomena can be reduced to the diagram in Figure 4. Following the example of the distinguished Russian mathematician A. A. Markov, we consider the first 20,000 letters of the poem *Eugene Onegin* (except *ъ* and *ь*). Our particle is at point A if the letter is a vowel and at point B if the letter is a consonant. The succession of vowels and consonants is represented by the motion of the particle on the diagram shown in Figure 4. Obviously, the probability that the letter following a vowel is a consonant is greater than that the letter is once more a vowel. In fact, Markov's calculations show that the probability p_{12} for the appearance of a consonant under the condition that the previous letter was a vowel is approximately equal to 0.872, while the probability p_{11} for the appearance of a vowel under the same condition is equal to about 0.128. It was shown in the same way that $p_{21} \approx 0.663$ and $p_{22} \approx 0.337$.

Similar counts made of the first 100,000 letters of Aksakov's story *The Childhood of Bagrov's Grandson* give a somewhat different result:

$$p_{11} \approx 0.147; \qquad p_{21} \approx 0.695;$$
$$p_{12} \approx 0.853; \qquad p_{22} \approx 0.305.$$

With certain limitations, a number of meteorological phenomena can also be handled in a similar fashion, for example, the sequence of clear and cloudy days. The probability that a cloudy day will follow a cloudy day is greater than the probability that the following day will be clear; the probability that a clear day will follow a clear day is greater than the probability that the following day will be cloudy. The probability of a change from a clear day to a cloudy one and from a cloudy to a clear one, etc. (a movement of the particle in Problem 5 corresponds to this change), proves to be approximately constant for a definite place and season and can be calculated from observations.

Problem 6. Two identical-looking urns stand in a room. Suppose there are a balls in the left urn and b in the right. Several people come into the room, one after the other, and either transfer a ball from the right urn to the left or from the left urn to the right. It is assumed that the probability that a ball is transferred from the right urn to the left is equal to the probability that a ball is transferred from the left urn to the right, that is, $\frac{1}{2}$. The experiment goes on until one of the urns is empty. What is the probability that the left urn becomes empty? What is the probability that the right urn becomes empty? What is the probability that the experiment does not end?

Problem 7. A caterpillar crawls along the edges of a wire cube (Fig. 5). On reaching a corner, the probability that it will crawl onto any particular edge that leads out from this corner is $\frac{1}{3}$. The points A and B are daubed with glue. The caterpillar starts out from the point 0. What is the probability that it sticks to the point A? What is the probability that it sticks to the point B?

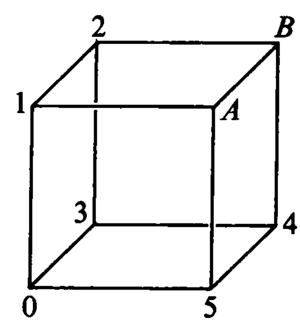

Fig. 5

2. Problems Concerning a Random Walk on an Infinite Line

4. GRAPH OF COIN TOSSES

We mark the points 0, ± 1, ± 2, ± 3, ... on a straight line and carry out the following experiment: We place a marker on the point 0 and toss a coin. If heads comes up, we move the marker one place to the left, and if tails comes up, we move the marker one place to the right. Now we toss the coin for a second time, for a third time, etc., and each time move the piece according to the result of the toss. We can assume that the two possible outcomes of a toss have equal probabilities, so that for each toss the probability that the marker is moved to the left is exactly as great as the probability that the marker is moved to the right; that is, the probability is $\frac{1}{2}$.

Clearly, after the first toss, the marker is on the point -1 or 1; after the second toss, on one of the points -2, 0, or 2; after the third, on -3, -1, 1, 3, etc. The diagram shown in Figure 6 gives a graphical picture of the possible positions of the marker at each moment.

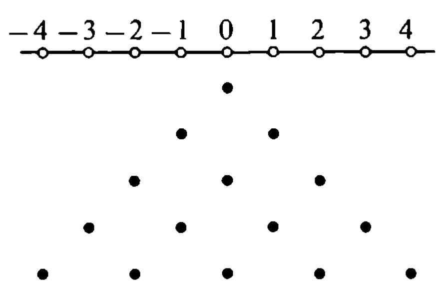

Fig. 6

The points shown in this diagram form a triangle. (This triangle can be continued downwards indefinitely.) The vertex of the triangle lies under the number 0, corresponding to the fact that 0 was the initial position of our marker. The first row of the triangle consists of two points. The numbers lying over these points, -1 and 1, show where the marker can stand after the first toss. The following row shows where the marker can stand after the second toss, and so on.

5. THE TRIANGLE OF PROBABILITIES

The diagram shown in Figure 6 exhibits all of the possible positions of the marker, but of these some are more, others less probable. We seek to calculate these probabilities. At the beginning, the marker is on the point 0 with probability 1. After the coin has been tossed for the first time, the marker is found with a probability of $\frac{1}{2}$ on each of the points -1 and 1. After the second toss, one can have obtained the following results:

heads–heads, heads–tails, tails–heads, tails–tails.

These four results are all equally probable, and, consequently, each has the probability $\frac{1}{4}$. After the first result, the marker is on point -2; after the second and third, on 0; and after the fourth, on $+2$. Hence, after the first two tosses the marker is found on the point -2 with a probability of $\frac{1}{4}$; on the point 0 with a probability of $\frac{2}{4}$; and on the point $+2$ with a probability of $\frac{1}{4}$. Similarly, the probability of each possible position of the marker after the third, fourth, . . . toss can be calculated. If one replaces every point of the diagram by the corresponding probability, one obtains the triangle of numbers shown in Figure 7. This triangle (we shall call it the *triangle of probabilities*) has a noteworthy property: *each of its numbers is equal to half of the sum of the two numbers standing above it.* This property can be easily verified for all of the numbers shown in Figure 7. It is clear, however, that this check still does not prove that this property continues to hold for an arbitrary continuation of the triangle.

-4	-3	-2	-1	0	1	2	3	4
				1				
			$\frac{1}{2}$		$\frac{1}{2}$			
		$\frac{1}{4}$		$\frac{2}{4}$		$\frac{1}{4}$		
	$\frac{1}{8}$		$\frac{3}{8}$		$\frac{3}{8}$		$\frac{1}{8}$	
$\frac{1}{16}$		$\frac{4}{16}$		$\frac{6}{16}$		$\frac{4}{16}$		$\frac{1}{16}$

Fig. 7

Problem 8. Prove, with the help of the formula for complete probability (Property 7, page 14), that

$$Z_n{}^k = \frac{1}{2}(Z_{n-1}{}^{k-1} + Z_{n-1}{}^{k+1}), \tag{1}$$

where $Z_n{}^k$ denotes the probability that the marker is at the point k at time n.

With the help of the *rule of the half-sum*, the triangle of probabilities can be easily continued. The first nine rows of the triangle of probabilities are written out in Figure 8 (not counting the zeroth

$$
\begin{array}{ccccccccccccccccccccl}
 & & & & & & & & & 1 & & & & & & & & & & & \text{0th row} \\
 & & & & & & & & \frac{1}{2} & & \frac{1}{2} & & & & & & & & & & \text{1st row} \\
 & & & & & & & \frac{1}{4} & & \frac{2}{4} & & \frac{1}{4} & & & & & & & & & \text{2nd row} \\
 & & & & & & \frac{1}{8} & & \frac{3}{8} & & \frac{3}{8} & & \frac{1}{8} & & & & & & & & \text{3rd row} \\
 & & & & & \frac{1}{16} & & \frac{4}{16} & & \frac{6}{16} & & \frac{4}{16} & & \frac{1}{16} & & & & & & & \text{4th row} \\
 & & & & \frac{1}{32} & & \frac{5}{32} & & \frac{10}{32} & & \frac{10}{32} & & \frac{5}{32} & & \frac{1}{32} & & & & & & \text{5th row} \\
 & & & \frac{1}{64} & & \frac{6}{64} & & \frac{15}{64} & & \frac{20}{64} & & \frac{15}{64} & & \frac{6}{64} & & \frac{1}{64} & & & & & \text{6th row} \\
 & & \frac{1}{128} & & \frac{7}{128} & & \frac{21}{128} & & \frac{35}{128} & & \frac{35}{128} & & \frac{21}{128} & & \frac{7}{128} & & \frac{1}{128} & & & & \text{7th row} \\
 & \frac{1}{256} & & \frac{8}{256} & & \frac{28}{256} & & \frac{56}{256} & & \frac{70}{256} & & \frac{56}{256} & & \frac{28}{256} & & \frac{8}{256} & & \frac{1}{256} & & & \text{8th row} \\
\frac{1}{512} & & \frac{9}{512} & & \frac{36}{512} & & \frac{84}{512} & & \frac{126}{512} & & \frac{126}{512} & & \frac{84}{512} & & \frac{36}{512} & & \frac{9}{512} & & \frac{1}{512} & & \text{9th row} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & &
\end{array}
$$

Fig. 8

row, the vertex of the triangle). Their direct calculation (by the method used for the first four rows) would not be easy.

Problem 9. Prove that the sum of the elements of each row of the triangle of probabilities is equal to one.

We remark that the *rule of the half-sum* is equivalent to the following *halving-rule*. We consider an arbitrary row of the triangle, halve every number of this row, and place one half below and to the right, the other half below and to the left (see Figure 9, in

$$
\begin{tikzcd}
 & \frac{1}{16} \arrow[dl] \arrow[dr] & & \frac{4}{16} \arrow[dl] \arrow[dr] & & \frac{6}{16} \arrow[dl] \arrow[dr] & & \frac{4}{16} \arrow[dl] \arrow[dr] & & \frac{1}{16} \arrow[dl] \arrow[dr] & \\
\frac{1}{32} & & \frac{1}{32}+\frac{4}{32} & & \frac{4}{32}+\frac{6}{32} & & \frac{6}{32}+\frac{4}{32} & & \frac{4}{32}+\frac{1}{32} & & \frac{1}{32}
\end{tikzcd}
$$

Fig. 9

which this is carried out for the fourth row). If we add the numbers that stand on the same point, we obtain the next row of the triangle.

Let us now imagine that at the initial time there is a unit mass at the point O, and that in the course of a second it splits into two equal parts, one half moving to the left and the other half to the right. During the second second these halves again divide into two equal parts, one of which moves to the right, the other to the left, etc.

It is clear that the sum of the masses that are found at the point k after n seconds is equal to the number in our triangle that stands in the kth place of the nth row. This connection between the problem of the random motion of a marker and the problem of the shifting of a dividing mass is very useful in the solution of a number of problems.[1]

If the first row of the triangle of probabilities is multiplied by 2, the second by 2^2, the third by $2^3, \ldots$, the nth by 2^n, we obtain a triangle that consists only of whole numbers. The reader can easily verify that, in this new triangle, every number is equal to the sum of the two numbers standing above it. This triangle is called *Pascal's triangle;* see E. B. Dynkin and V. A. Uspenskii, *Problems in the Theory of Numbers* (Boston: D. C. Heath and Company, 1963), Chapter 3, section 18.

Let us divide every element of the triangle of probabilities by its left neighbor. Of course, the elements of the left edge of the triangle have no left neighbors. We strike out these elements and what remains is the quotient triangle in Figure 10. The law by

$$\frac{1}{1}$$
$$\frac{2}{1} \quad \frac{1}{2}$$
$$\frac{3}{1} \quad \frac{2}{2} \quad \frac{1}{3}$$
$$\frac{4}{1} \quad \frac{3}{2} \quad \frac{2}{3} \quad \frac{1}{4}$$
$$\frac{5}{1} \quad \frac{4}{2} \quad \frac{3}{3} \quad \frac{2}{4} \quad \frac{1}{5}$$
$$\frac{6}{1} \quad \frac{5}{2} \quad \frac{4}{3} \quad \frac{3}{4} \quad \frac{2}{5} \quad \frac{1}{6}$$
$$\frac{7}{1} \quad \frac{6}{2} \quad \frac{5}{3} \quad \frac{4}{4} \quad \frac{3}{5} \quad \frac{2}{6} \quad \frac{1}{7}$$
$$\frac{8}{1} \quad \frac{7}{2} \quad \frac{6}{3} \quad \frac{5}{4} \quad \frac{4}{5} \quad \frac{3}{6} \quad \frac{2}{7} \quad \frac{1}{8}$$
$$\frac{9}{1} \quad \frac{8}{2} \quad \frac{7}{3} \quad \frac{6}{4} \quad \frac{5}{5} \quad \frac{4}{6} \quad \frac{3}{7} \quad \frac{2}{8} \quad \frac{1}{9}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Fig. 10

which this triangle is constructed is easy to discover. In the nth row, the denominators of the fractions run through all the numbers from 1 to n, while the numerators run from n to 1. We leave it to the reader to verify this law with the aid of formula (1).

[1] An analogous scheme was considered in one of the problems of the Eighth Moscow Mathematical Olympiad. This exercise dealt with 2^n men who start out from the vertex of the triangle in Figure 6, half of them going down and to the left, and half going down and to the right. At each point, they continue according to the *halving-rule.* It was asked how many people there are at each point of the nth row.

The triangle of probabilities can easily be recovered from the quotient triangle. We begin the nth row with the number $\frac{1}{2^n}$ and then reconstruct all of the elements of this row, one after the other, by multiplying the element already obtained by the corresponding element of the nth row of the quotient triangle. We thus obtain the triangle of probabilities in the form given in Figure 11.

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & \frac{1}{2} & & \frac{1}{2}\cdot\frac{1}{1} & & & \\
 & & \frac{1}{4} & & \frac{1}{4}\cdot\frac{2}{1} & & \frac{1}{4}\cdot\frac{2}{1}\cdot\frac{1}{2} & & \\
 & \frac{1}{8} & & \frac{1}{8}\cdot\frac{3}{1} & & \frac{1}{8}\cdot\frac{3}{1}\cdot\frac{2}{2} & & \frac{1}{8}\cdot\frac{3}{1}\cdot\frac{2}{2}\cdot\frac{1}{3} & \\
\frac{1}{16} & & \frac{1}{16}\cdot\frac{4}{1} & & \frac{1}{16}\cdot\frac{4}{1}\cdot\frac{3}{2} & & \frac{1}{16}\cdot\frac{4}{1}\cdot\frac{3}{2}\cdot\frac{2}{3} & & \frac{1}{16}\cdot\frac{4}{1}\cdot\frac{3}{2}\cdot\frac{2}{3}\cdot\frac{1}{4} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}
$$

Fig. 11

From this, it can be seen that the $(k + 1)$st element of the nth row of the triangle of probabilities is equal to

$$\frac{1}{2^n}\cdot\frac{n}{1}\cdot\frac{n-1}{2}\cdot\frac{n-2}{3}\cdot\ldots\cdot\frac{n-k+1}{k}. \tag{2}$$

6. CENTRAL ELEMENTS OF THE TRIANGLE OF PROBABILITIES

We are particularly interested in the central elements of the triangle of probabilities, that is, the elements that lie on its axis of symmetry. These elements are found only on even rows. The central element $Z_{2k}{}^0$, which lies on the $2k$th row, is the probability that the marker will have returned to the initial position after $2k$ moves. Denoting the probability of this event by w_{2k}, we have

$$w_0 = 1,\ w_2 = \frac{2}{4},\ w_4 = \frac{6}{16},\ w_6 = \frac{20}{64},\ w_8 = \frac{70}{256},\ \ldots,$$

or, by Figure 11,

$$w_0 = 1,\ w_2 = \frac{1}{4}\cdot\frac{2}{1},\ w_4 = \frac{1}{16}\cdot\frac{4}{1}\cdot\frac{3}{2},\ \ldots.$$

If we use the general expression (2) and the fact that the $2k$th row consists of $2k + 1$ numbers, of which k are to the right and k to the left of the middle term, we obtain the formula

$$w_{2k} = \frac{1}{2^{2k}} \cdot \frac{2k}{1} \cdot \frac{2k-1}{2} \cdot \frac{2k-2}{3} \cdot \dots \cdot \frac{k+1}{k}. \tag{3}$$

The value of w_{2k} can be easily calculated using this formula, provided k is not too large. If k is very large, it is extremely difficult to calculate this fraction. (Try, for example, to calculate $w_{10,000}$!) We can estimate w_{2k} by making use of the following remarkable inequality:

$$\frac{1}{\sqrt{4k}} \leq w_{2k} < \frac{1}{\sqrt{2k}} \quad (k = 1, 2, 3, \dots). \tag{4}$$

To prove this inequality, we first transform formula (3):

$$\begin{aligned} w_{2k} &= \frac{1}{2^{2k}} \cdot \frac{2k}{1} \cdot \frac{2k-1}{2} \cdot \frac{2k-2}{3} \cdot \dots \cdot \frac{k+1}{k} \\ &= \frac{1}{2^{2k}} \cdot \frac{2k}{1} \cdot \frac{2k-1}{2} \cdot \frac{2k-2}{3} \cdot \dots \cdot \frac{k+1}{k} \cdot \frac{k}{k} \cdot \frac{k-1}{k-1} \cdot \dots \cdot \frac{1}{1} \\ &= \frac{1}{2^{2k}} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2k-1) \cdot 2 \cdot 4 \cdot 6 \cdot \dots \cdot 2k}{1 \cdot 2 \cdot 3 \cdot \dots \cdot k \cdot 1 \cdot 2 \cdot 3 \cdot \dots \cdot k} \\ &= \frac{1}{2^{2k}} \cdot \frac{1}{1} \cdot \frac{3}{2} \cdot \frac{5}{3} \cdot \dots \cdot \frac{2k-1}{k} \cdot \frac{2}{1} \cdot \frac{4}{2} \cdot \frac{6}{3} \cdot \dots \cdot \frac{2k}{k} \\ &= \frac{1}{2^k} \cdot \frac{1}{1} \cdot \frac{3}{2} \cdot \frac{5}{3} \cdot \dots \cdot \frac{2k-1}{k} = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \dots \cdot \frac{2k-1}{2k}. \end{aligned}$$

We now write the three products

$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{5}{6} \cdot \frac{6}{7} \cdot \frac{7}{8} \cdot \frac{8}{9} \cdot \frac{9}{10} \cdot \dots \cdot \frac{2k-2}{2k-1} \cdot \frac{2k-1}{2k}, \tag{5}$$

$$\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{7}{8} \cdot \frac{7}{8} \cdot \frac{9}{10} \cdot \frac{9}{10} \cdot \dots \cdot \frac{2k-1}{2k} \cdot \frac{2k-1}{2k}, \tag{6}$$

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{5}{6} \cdot \frac{6}{7} \cdot \frac{7}{8} \cdot \frac{8}{9} \cdot \frac{9}{10} \cdot \frac{10}{11} \cdot \dots \cdot \frac{2k-1}{2k} \cdot \frac{2k}{2k+1} \tag{7}$$

under one another. It is easily seen that of the three numbers in any column, the second is at least equal to the first, and the third is at least equal to the second. Hence, (5) has the smallest and (7) the greatest value. The middle product is equal to

$$\left(\frac{1}{2}\right)^2 \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{5}{6}\right)^2 \cdot \dots \cdot \left(\frac{2k-1}{2k}\right)^2 = w_{2k}{}^2,$$

the upper, after simplification, is equal to $\frac{1}{4k}$, and the lower is equal to $\frac{1}{2k+1}$. Hence,

$$\frac{1}{4k} \leq w_{2k}{}^2 \leq \frac{1}{2k+1}.$$

Then, certainly,

$$\frac{1}{4k} \leq w_{2k}{}^2 < \frac{1}{2k},$$

or, if we take the square root,

$$\frac{1}{\sqrt{4k}} \leq w_{2k} < \frac{1}{\sqrt{2k}},$$

which was to be proved.

By the same method, still more exact approximations for w_{2k} can be found.

Problem 10. Prove that, for all $k \geq 2$, we have

$$\sqrt{3 \cdot \left(\frac{1}{2}\right)^2 \cdot \frac{3}{4} \cdot \frac{1}{2k}} \leq w_{2k} < \sqrt{3 \cdot \left(\frac{1}{2}\right)^2 \cdot \frac{1}{2k}}; \tag{8}$$

for all $k \geq 3$, we have

$$\sqrt{5 \cdot \left(\frac{1}{2} \cdot \frac{3}{4}\right)^2 \cdot \frac{5}{6} \cdot \frac{1}{2k}} \leq w_{2k} < \sqrt{5 \cdot \left(\frac{1}{2} \cdot \frac{3}{4}\right)^2 \cdot \frac{1}{2k}}; \tag{9}$$

and, in general, for all $k \geq a$

$$\sqrt{(2a-1) \cdot \left(\frac{1}{2} \cdot \frac{3}{4} \cdot \dots \cdot \frac{2a-3}{2a-2}\right)^2 \cdot \frac{2a-1}{2a} \cdot \frac{1}{2k}}$$

$$\leq w_{2k} < \sqrt{(2a-1) \cdot \left(\frac{1}{2} \cdot \frac{3}{4} \cdot \dots \cdot \frac{2a-3}{2a-2}\right)^2 \cdot \frac{1}{2k}}. \tag{10}$$

Hint. In each of the products (5) and (7), change the initial terms in such a way that they coincide with the initial terms of the product (6).

Problem 10 yields a sequence of increasingly accurate approximations for w_{2k}; the ratio of the lower to the upper bound serves as a measure of the accuracy. For approximation (4), this is equal to $\sqrt{\frac{1}{2}}$; for approximation (8), it is equal to $\sqrt{\frac{3}{4}}$; for approximation (9), it is equal to $\sqrt{\frac{5}{6}}$; and for approximation (10), it is equal to $\sqrt{1-\frac{1}{2a}}$, and, thus, approaches 1 as a limit. If the products in the inequalities (8) and (9) are calculated out, we obtain

$$\frac{1}{\sqrt{3.56k}} \leq w_{2k} < \frac{1}{\sqrt{2.66k}} \quad (\text{for } k \geq 2), \tag{11}$$

$$\frac{1}{\sqrt{3.42k}} \leq w_{2k} < \frac{1}{\sqrt{2.84k}} \quad (\text{for } k \geq 3). \tag{12}$$

On substituting the values 4, 5, 60, 150 for a in the general inequality (10), we obtain

$$\frac{1}{\sqrt{3.35k}} \leq w_{2k} < \frac{1}{\sqrt{2.92k}} \quad (\text{for } k \geq 4), \tag{13}$$

$$\frac{1}{\sqrt{3.31k}} \leq w_{2k} < \frac{1}{\sqrt{2.98k}} \quad (\text{for } k \geq 5), \tag{14}$$

$$\frac{1}{\sqrt{3.18k}} \leq w_{2k} < \frac{1}{\sqrt{3.10k}} \quad (\text{for } k \geq 60), \tag{15}$$

$$\frac{1}{\sqrt{3.15k}} \leq w_{2k} < \frac{1}{\sqrt{3.14k}} \quad (\text{for } k \geq 150). \tag{16}$$

The coefficients of k constantly decrease on the left side of the inequality, while they constantly increase on the right side. One can prove rigorously (we shall not do so here) that both of these sequences converge to the same limit, and that this limit is the well-known π (the ratio of the circumference of a circle to its diameter). Hence, w_{2k} can be calculated for large values of $2k$ by the following approximation formula:

$$w_{2k} \approx \frac{1}{\sqrt{\pi k}}. \tag{17}$$

One can prove (and the reader can verify this for himself) that for $k = 25$ this formula yields an approximation that is correct to two significant figures; the larger the value of k, the more accurate the approximation.

Problem 11. Calculate $w_{10,000}$ to two significant figures.

In the triangle of probabilities, the numbers increase as we move from the edges to the middle. Hence, by using the inequality (4) it is not difficult to find the upper bound of the numbers in the nth row.

Problem 12. Prove that all elements of the nth row of the triangle of probabilities are less than or equal to $\frac{1}{\sqrt{n}}$.

7. ESTIMATION OF ARBITRARY ELEMENTS OF THE TRIANGLE

Since we have found an approximation formula for the terms lying in the middle of the triangle of probabilities, it is only natural to seek an equally convenient formula for the other terms of the triangle.

Consider the $2k$th row. We denote the middle term of this row by v_0 (above, we called this term w_{2k}) and number all of the terms to its right in this row

$$v_0, v_1, v_2, \ldots, v_{k-2}, v_{k-1}, v_k.$$

Now, the elements on the right side of the $2k$th row of the quotient triangle are

$$\frac{k}{k+1}, \frac{k-1}{k+2}, \ldots, \frac{2}{2k-1}, \frac{1}{2k}.$$

By the definition of the quotient triangle, we have

$$\frac{v_1}{v_0} = \frac{k}{k+1}; \quad \frac{v_2}{v_1} = \frac{k-1}{k+2}; \quad \ldots; \quad \frac{v_{k-1}}{v_{k-2}} = \frac{2}{2k-1}; \quad \frac{v_k}{v_{k-1}} = \frac{1}{2k}.$$

If we choose a number s between 0 and k, and estimate the ratio $\frac{v_s}{v_0}$, we obtain

$$\frac{v_s}{v_0} = \frac{v_1}{v_0} \cdot \frac{v_2}{v_1} \cdot \ldots \cdot \frac{v_{s-1}}{v_{s-2}} \cdot \frac{v_s}{v_{s-1}}$$

$$= \frac{k}{k+1} \cdot \frac{k-1}{k+2} \cdot \ldots \cdot \frac{k-s+2}{k+s-1} \cdot \frac{k-s+1}{k+s}.$$

If we reverse the order of the denominators appearing in these factors, we get

$$\frac{v_s}{v_0} = \frac{k}{k+s} \cdot \frac{k-1}{k+s-1} \cdot \ldots \cdot \frac{k-s+2}{k+2} \cdot \frac{k-s+1}{k+1}$$

$$= \left(1 - \frac{s}{k+s}\right)\left(1 - \frac{s}{k+s-1}\right) \cdots \left(1 - \frac{s}{k+2}\right)\left(1 - \frac{s}{k+1}\right).$$

It is easy to see that the first factor is greater than each of those following it, while the last factor is less than each of those preceding it. It follows that

$$\left(1 - \frac{s}{k+1}\right)^s \leq \frac{v_s}{v_0} \leq \left(1 - \frac{s}{k+s}\right)^s, \tag{18}$$

or,

$$\left(\frac{k+1-s}{k+1}\right)^s \leq \frac{v_s}{v_0} \leq \left(\frac{k}{k+s}\right)^s. \tag{19}$$

Problem 13. Between what limits does the probability $Z_{120}{}^{20}$ lie?

One can find analogous estimates for the elements in the odd-numbered rows; however, we shall not concern ourselves with this.

8. THE LAW OF THE SQUARE ROOT OF *n*

Assume that our experiment of tossing the coin and moving the marker is continued for sufficiently long, say for 1,000 moves. How far is the marker then from the starting point? In any case, not farther than 1,000 steps to the right or left. And this is the only thing that we can state with absolute certainty. If we seek to assert that the marker moves less than 1,000 steps, that is, not more than 998 steps, then an error is not impossible, for the same side of the coin can come up every time in 1,000 tosses (heads or tails), and the marker will then be found at $-1{,}000$ or at $+1{,}000$. The probability that this happens is, however, so small $\left(\text{it is equal to } 2\cdot\frac{1}{2^{1{,}000}}\right)$ that we can regard such an outcome as practically impossible. Likewise, the probability that the marker passes the nine hundred ninetieth mark is small. For this to happen, the marker must reach one of the points

$$-1{,}000,\ -998,\ -996,\ -994,\ -992,$$
$$+992,\ +994,\ +996,\ +998,\ +1{,}000.$$

To calculate the probability of this, we must consider the thousandth row of the triangle of probabilities and form the sum of the five outer left and the five outer right terms. Because of the symmetry of the triangle, this sum is equal to twice the sum of the five outer left or right terms:

$$2\left(\frac{1}{2^{1{,}000}}+\frac{1}{2^{1{,}000}}\cdot\frac{1{,}000}{1}+\frac{1}{2^{1{,}000}}\cdot\frac{1{,}000}{1}\cdot\frac{999}{2}\right.$$
$$+\frac{1}{2^{1{,}000}}\cdot\frac{1{,}000}{1}\cdot\frac{999}{2}\cdot\frac{998}{3}$$
$$\left.+\frac{1}{2^{1{,}000}}\cdot\frac{1{,}000}{1}\cdot\frac{999}{2}\cdot\frac{998}{3}\cdot\frac{997}{4}\right)\approx 0.\underbrace{00\ldots0\ldots}_{290\text{ zeros}}8.$$

Hence, we can regard an error here as practically impossible. This is not surprising, for when it is assumed that after one thousand moves the marker has not passed the 990th mark, of the 1,001 possible positions we neglect the ten most improbable ones. The statement that the marker remains under the hundredth mark (that is, that it does not pass the ninety-eighth mark) is much more risky. Here, we neglect the greater part of the possible positions, namely, 902 out of 1,001. How justified are we in doing this? How small is the probability of an error in this statement or, on the other hand, how close to one is it? This question cannot be answered without calculation. To carry it through, we must find the sum of the 451 left outer and the 451 right outer terms in the thousandth row of the triangle of probabilities. While no knowledge is necessary for this calculation except that of the four fundamental arithmetic operations, few of our readers could carry it through to the end. It would take too much time and energy.[1] We shall therefore seek to estimate the probability, and forego an exact calculation. Our considerations will be altogether general. To simplify the calculations, however, we shall assume from now on that the number n of steps that the marker makes is even.

Let us consider the $2k$th row of the triangle of probabilities and number its terms from the middle term to the right edge

$$v_0, v_1, v_2, \ldots, v_k.$$

To estimate the sum

$$S = v_r + v_{r+1} + \cdots + v_k,$$

we use inequality (19) of section 7.

By this inequality, we have

$$\frac{v_r}{v_0} \leq \left(\frac{k}{k+r}\right)^r,$$

$$\frac{v_{r+1}}{v_0} \leq \left(\frac{k}{k+r+1}\right)^{r+1},$$

$$\frac{v_{r+2}}{v_0} \leq \left(\frac{k}{k+r+2}\right)^{r+2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{v_k}{v_0} \leq \left(\frac{k}{k+k}\right)^k.$$

For brevity, we denote $\dfrac{k}{k+r}$ by g. One sees immediately that no fraction enclosed in parentheses on the right-hand sides of our inequalities exceeds g. Hence,

$$\frac{v_r}{v_0} \leq g^r, \quad \frac{v_{r+1}}{v_0} \leq g^{r+1}, \quad \frac{v_{r+2}}{v_0} \leq g^{r+2}, \quad \ldots, \quad \frac{v_k}{v_0} \leq g^k.$$

[1] Naturally, we can calculate this sum in the following way: We subtract from one the sum of the 99 middle terms. The number of summands in this sum is significantly smaller (99 instead of 902), but these summands themselves are considerably more difficult to calculate.

Adding these inequalities, we obtain

$$\frac{S}{v_0} \leq g^r + g^{r+1} + g^{r+2} + \cdots + g^k.$$

The right-hand side forms a geometric series, whose sum is $\frac{g^r - g^{k+1}}{1 - g}$; thus,

$$\frac{S}{v_0} \leq \frac{g^r - g^{k+1}}{1 - g}.$$

But, $g = \frac{k}{k + r} < 1$, so that $1 - g > 0$, and the inequality is strengthened when we omit the negative number $\frac{-g^{k+1}}{1 - g}$ from the right-hand side. Hence,

$$\frac{S}{v_0} < \frac{g^r}{1 - g}.$$

If we multiply this inequality by v_0 and substitute the original value for g, we finally obtain

$$S < v_0 \frac{k + r}{r} \left(\frac{k}{k + r}\right)^r. \tag{20}$$

Now we are in a position to estimate the probability P that after $n = 2k$ moves, the marker is not less than $m = 2r$ steps from the starting point. This probability is equal to twice the probability that after $n = 2k$ moves, the marker is not less than $m = 2r$ steps to the right of the starting point. The probability of this last event is precisely the sum S, which we have already estimated. In formula (20) we replace v_0 (for the middle term of the $2k$th row) by the more convenient symbol w_{2k}, used previously; this reminds us that the middle term depends on the number of the row. We multiply both sides of inequality (20) by 2, replace k by $\frac{n}{2}$ and r by $\frac{m}{2}$, and thus obtain

$$P \leq 2w_n \cdot \frac{n + m}{m} \left(\frac{n}{n + m}\right)^{\frac{m}{2}}. \tag{21}$$

The middle term

$$w_n = w_{2k}$$

can be estimated by either formula (4) or (11)–(16) of section 6. By these formulas, we have

$$w_{2k} < \frac{1}{\sqrt{Bk}}, \tag{22}$$

where B is a number which can be chosen as near to π as desired, provided k is large enough. (In (4) $B = 2$; in (11) $B = 2.66$, etc.)

We replace k by $\frac{n}{2}$ in the inequality (22). Comparing this inequality with (21), we obtain

$$P \leq 2\sqrt{\frac{2}{Bn}} \cdot \frac{n+m}{m} \cdot \left(\frac{n}{n+m}\right)^{\frac{m}{2}}, \tag{23}$$

and can use this formula to calculate the numerical example discussed at the beginning of this section. Suppose the marker makes 1,000 moves. What is the probability that it is not less than 100 steps from the starting point? We substitute $n = 1{,}000$ and $m = 100$ in formula (23):

$$P \leq 2\sqrt{\frac{2}{1{,}000B}} \cdot 11 \cdot \left(\frac{10}{11}\right)^{50}.$$

Since $k = \frac{n}{2} > 150$, we can set $B = 3.14$ according to formula (16) of section 6. We then obtain

$$P < 0.0048.$$

The event that the marker is at least 100 steps from the starting point after 1,000 moves is thus highly improbable. If a not too high degree of certainty is demanded and, say, 0.005 is permitted as the degree of uncertainty, one can say that it is practically impossible that the marker has moved more than 98 steps after 1,000 moves; that is, it is practically certain that the marker has moved less than 100 steps.

For the further study of the motion of the marker, it is convenient to replace the approximation (23) by another that is not so precise but is simpler and more convenient to calculate. We require a preliminary inequality, the proof of which is left to the reader.

Problem 14. Prove that for arbitrary positive p and integral positive r,

$$(1 + p)^r \geq 1 + rp. \tag{24}$$

Let m and n be positive even numbers. We substitute $r = \frac{m}{2}$ and $p = \frac{m}{n}$ in the inequality (24) and obtain

$$\left(1 + \frac{m}{n}\right)^{\frac{m}{2}} \geq 1 + \frac{m^2}{2n} > \frac{m^2}{2n}.$$

Hence,

$$\left(\frac{n}{n + m}\right)^{\frac{m}{2}} = \frac{1}{\left(1 + \frac{m}{n}\right)^{\frac{m}{2}}} < \frac{2n}{m^2}. \tag{25}$$

From inequalities (21) and (25) it follows that

$$P < 2w_n(n + m)\frac{2n}{m^3}. \tag{26}$$

According to inequality (4) in section 6, $w_n < \frac{1}{\sqrt{n}}$; furthermore, $m \leq n$. Hence, the estimate

$$P < \frac{2}{\sqrt{n}} \cdot 2n \cdot \frac{2n}{m^3} = \left(\frac{2\sqrt{n}}{m}\right)^3$$

follows from inequality (26).

Thus, if it is asserted that the marker is less than m steps distant from the starting point after n moves, the probability of error is less than $\left(\frac{2\sqrt{n}}{m}\right)^3$.

We choose an arbitrary positive number t and estimate the probability of an error in the following statement:

(A) After n moves the marker is less than $t\sqrt{n}$ steps away from the starting point.

We denote by m the smallest even number that satisfies the condition

$$m \geq t\sqrt{n}.$$

Since the distance of the marker from the starting point after an even number of moves is an even number, the statement (A) is equivalent to the following statement:

(B) After n moves the marker is less than m steps away from the starting point.

Consequently, the probability of an error in the statement (A) is equal to the probability of an error in the statement (B). This probability is smaller than

$$\left(\frac{2\sqrt{n}}{m}\right)^3 \leq \left(\frac{2\sqrt{n}}{t\sqrt{n}}\right)^3 = \left(\frac{2}{t}\right)^3.$$

Hence, we have proved the following important law:

LAW (The Law of the Square Root of n). *With probability of error less than* $\left(\frac{2}{t}\right)^3$, *one can assert that after n moves, the distance of the marker from the starting point is less than* $t\sqrt{n}$ *(that is, the marker is situated between* $-t\sqrt{n}$ *and* $+t\sqrt{n}$*).*

We choose a certain degree of uncertainty, for example 0.005, and determine t so that

$$\left(\frac{2}{t}\right)^3 = 0.005.$$

As a solution of this equation for t, we find[1]

$$t = \frac{2}{\sqrt[3]{0.005}} \approx 12.$$

It follows from the *law of the square root of n* that, for every value n, the statement that *the marker has moved less than* $12\sqrt{n}$ *steps from the starting point in n moves* is practically certain.

We compile the following table:

n	$n' = 12\sqrt{n}$	$\frac{n'}{n} = \frac{12}{\sqrt{n}}$
2,500	600	0.24
10,000	1,200	0.12
250,000	6,000	0.024
1,000,000	12,000	0.012

[1] Here the approximation is greater than the actual value.

The second column of the table gives almost certain bounds on the distance of the marker from the starting point for various values of n. The ratio $\frac{n'}{n} = \frac{12}{\sqrt{n}}$ approaches zero as n increases without bound.

Let us now assume that the marker is m steps distant from the starting point at the end of n moves. We call the ratio $\frac{m}{n}$ the *reduced velocity* of the marker. If a particle starts out from the point 0 with this velocity and does not change direction, then at time n, its displacement amounts to m steps.

For example: If the marker is at the point -20 after 100 steps, its *reduced velocity* amounts to $\frac{1}{5}$. The reduced velocity varies between 1 (when the marker always moves in one direction) and 0 (when it returns to the starting point, that is, when it makes exactly as many moves to the left as to the right). From the *law of the square root of n* one easily deduces:

THEOREM. *When the motion of the marker is continued sufficiently long, it is practically certain that the reduced velocity is close to zero.*

In fact, it is practically certain that the displacement of the marker is less than $12\sqrt{n}$, and, hence, that its reduced velocity is less than $\frac{12\sqrt{n}}{n} = \frac{12}{\sqrt{n}}$. If n is sufficiently large, this bound will be arbitrarily close to zero.

Until now, we have taken 0.005 as the permissible probability of error. However, we can repeat our considerations without significant alteration for an arbitrary value ε of this tolerance. As a result, we come to the following conclusion:

One can say with probability of error less than ε that:

(*a*) The displacement of the marker after n moves is less than $\frac{2}{\sqrt[3]{\varepsilon}}\sqrt{n}$;

(*b*) The absolute value of its reduced velocity after n moves is less than $\frac{2}{\sqrt[3]{\varepsilon}} \Big/ \sqrt{n}$.

Problem 15. Let the permissible probability of error be 0.05. Give a practically certain bound for the displacement and for the reduced velocity of the marker after 1,000 moves.

Problem 16. Let α be an arbitrary positive number. Prove that it can be stated with the probability of error less than

$$\left(\frac{2}{\alpha\sqrt{n}}\right)^3$$

that the reduced velocity of the marker after n moves is less than α.

Problem 17. Determine the number of moves that are sufficient for the reduced velocity of the marker to be smaller than 0.01, with the probability of error not greater than 0.001.

9. THE LAW OF LARGE NUMBERS

We now recall that the marker moves in accordance with the outcome of the toss of a coin. If on n tosses of a coin there are l tails and $n - l$ heads, the marker moves l steps to the right and $n - l$ steps to the left, finally reaching the point

$$l - (n - l) = 2l - n.$$

The reduced velocity of the marker after n moves is given by the absolute value of

$$\frac{2l - n}{n} = 2\frac{l}{n} - 1. \tag{27}$$

The fraction $\frac{l}{n}$ characterizes the *relative frequency* with which tails comes up.

Let a permissible probability of error be given. We know that for large values of n it can be asserted with practical certainty that the reduced velocity is close to zero. It is clear from equality (27) that for a small reduced velocity $2\frac{l}{n}$ is approximately equal to 1, and the relative frequency $\frac{l}{n}$ is consequently near $\frac{1}{2}$. In other words:

If a coin is tossed very often, it is practically certain that the frequency with which heads comes up is approximately equal to $\frac{1}{2}$.

Roughly speaking, it is practically certain that heads comes up exactly as often as tails. A more exact formulation would be:

Choose an arbitrarily permissible probability of error ε and an arbitrarily small number α. If the number of tosses of the coin exceeds

$$N = \frac{1}{\alpha^2 \sqrt[3]{\varepsilon^2}},$$

it can be asserted with a probability of error less than ε that the frequency with which tails comes up differs from $\frac{1}{2}$ by less than α.

The proof for this exact formulation is easily obtained from part (*b*) on page 33. For $n > \frac{1}{\alpha^2 \sqrt[3]{\varepsilon^2}}$ we have

$$\frac{1}{\sqrt{n}} < \alpha \sqrt[3]{\varepsilon},$$

and

$$\frac{\frac{2}{\sqrt[3]{\varepsilon}}}{\sqrt{n}} < 2\alpha.$$

Hence, the absolute value of the reduced velocity of the marker is less than 2α, with the probability of error less than ε. In this case, the reduced velocity is equal to the absolute value of $\frac{2l - n}{n} = 2\frac{l}{n} - 1$. Hence, it can be stated with the probability of error less than ε that $2\frac{l}{n}$ does not differ from 1 by more than 2α, or in other words, that $\frac{l}{n}$ differs from $\frac{1}{2}$ by less than α.

Problem 18. How often must a coin be tossed so that it can be asserted with the probability of error less than 0.01 that the frequency with which tails comes up lies between 0.4 and 0.6?

Suppose now that a die is tossed instead of a coin. How often does the six come up? If the same calculations and arguments are carried through for this new case as for the example of the coin, we obtain the following result: for a great number of tosses, the frequency with which a six comes up lies near $\frac{1}{6}$ with practical certainty.

We consider yet another experiment. An urn contains a balls, of which b are white and the rest black. A ball is drawn n times from this urn, and is returned to the urn each time. How often will a white ball be drawn? One can prove that it is practically certain that for sufficiently many trials, the frequency with which a white ball is drawn lies near $\frac{b}{a}$.

We now formulate a general result including all of the above formulations as special cases.

Suppose that an experiment is carried out in which an event A can either occur or not occur (a toss of tails, the roll of a six, drawing a white ball out of the urn, etc.), *and let the probability of the occurrence of the event A be p.* (In our examples p was equal to $\frac{1}{2}$, $\frac{1}{6}$, and $\frac{b}{a}$, respectively.) *Suppose this experiment is repeated many times, the result of each trial not influencing the results of the succeeding ones. Then, for a large number of trials, it is practically certain that the frequency of the event A will be approximately equal to the probability p of this event.*

This general formulation can be made more precise, exactly as in the formulation for the case of the coin.

This result is essentially nothing but a restatement of a well-known theorem of Bernoulli,[1] which sets forth the simplest form of one of the fundamental laws of probability theory, *the law of large numbers.* Here we cannot go into the generalizations of Bernoulli's theorem. We only remark that the most important is due to the Russian mathematician P. L. Chebyshev.

The reader will appreciate the great significance of the law of large numbers. By the statement that the frequency of occurrence of an event A approaches the probability of A with practical certainty for a large series of trials, the law of large numbers makes possible the experimental determination of this probability. In many cases, the experimental method for the determination of a probability is the only possible one. Furthermore, the knowledge of the connection between probability and frequency enables one to draw practical conclusions about the frequency of appearance of an event in a long series of experiments from the theoretically calculated probability of this event. The connection between probability and frequency is fundamental in many applications of probability theory to physics, technology, etc.

[1] Jacob Bernoulli (1654–1705), famous Swiss mathematician.

3. Random Walks with Finitely Many States

In the preceding chapter, we considered the simplest example of a random walk, a random walk on a line. The problems posed there were of the following sort: At a given moment, where could the marker be, what is the probability that it is at a given point, and how far from the starting point is it? In this chapter, we consider more complicated schemes for random walks, including, for example, the random stroll through the city and the children's game mentioned in the Introduction. The problems related to these schemes differ somewhat from those investigated in Chapter 2. If an arbitrary point is chosen, we ask whether a particle can ever reach it, and if so, when. The first question can be answered with the aid of a general theorem (p. 46). An exact answer for the second, dealing with the question of the number of moves necessary to reach a given point with a given probability, can be found only in the simplest cases (see Problem 20 to follow). For the general case, we can give only an approximation for the necessary number of moves.

10. RANDOM WALKS ON A FINITE LINE

Let us make a slight change in the scheme for a random walk on a straight line. We place reflecting barriers in the path of the moving particle at the points m_1 and m_2 (see Fig. 12). These barriers

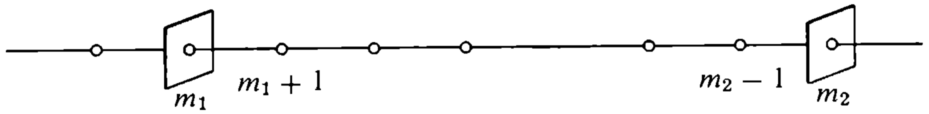

Fig. 12

cause particles that reach m_1 to move to $m_1 + 1$ on the next move, and those that reach m_2 to move to $m_2 - 1$ on the next move. The motion of the particle will thus take place between the points m_1 and m_2.

We could also restrict the motion of the particle by placing absorbing rather than reflecting barriers at the points m_1 and m_2. In this case, the particle upon reaching either the point m_1 or the point m_2 would remain there permanently. (We have already considered this diagram in the solution of Problem 6.)

Finally, we could place a reflecting barrier at one of the two points and an absorbing barrier at the other.

Problem 19. Show that the probability that after n moves the particle has reached the point m at least once does not depend on whether there is a barrier at m, and if there is, whether it is a reflecting or an absorbing barrier.

We place a reflecting barrier at point 0. A particle makes n moves starting from the point 1. What is the probability that it touches the point 3 at least once? To calculate this probability we can, according to Problem 19, place an absorbing barrier at point 3 (Fig. 13). But then, the event that after n moves the particle touches the point 3 at least once is the same as the event that the particle is at point 3 after n moves. (For, when it reaches this point it remains there.) We wish to calculate the probability d_n that the particle is at the point 3 after n moves. By a_n, b_n, c_n, we denote the probability of the events that the particle is at point 0, point 1, point 2, respectively.

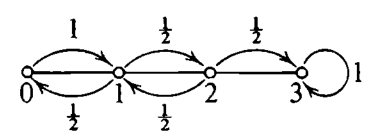

Fig. 13

To be at point 0 after n moves (the probability for this is a_n), the particle must be at point 1 after $n - 1$ moves (the probability for this is b_{n-1}), and then go from 1 to 0 (this occurs with probability $\frac{1}{2}$). By Property 6, we have

$$a_n = b_{n-1} \cdot \frac{1}{2}. \tag{1}$$

Similarly, with the aid of Properties 6 and 7, we obtain the relations

$$b_n = a_{n-1} \cdot 1 + c_{n-1} \cdot \tfrac{1}{2}, \tag{2}$$

$$c_n = b_{n-1} \cdot \tfrac{1}{2}, \tag{3}$$

$$d_n = d_{n-1} \cdot 1 + c_{n-1} \cdot \tfrac{1}{2}. \tag{4}$$

Problem 20. Using relations (1)–(4), show that

$$d_{2k} = d_{2k+1} = 1 - \left(\frac{3}{4}\right)^k. \tag{5}$$

Equality (5) can be interpreted as follows: the statement that after $2k$ moves the particle has touched the point 3 at least once, is false with probability $(\frac{3}{4})^k$. How many moves must be made for this probability to be less than 0.01?

We obtain $k_0 = 16.006$ as a solution of the equation $(\frac{3}{4})^{k_0} = 0.01$. Thus, we conclude that, for

$$k > 16.006, \tag{6}$$

the inequality

$$\left(\frac{3}{4}\right)^k < 0.01 \tag{7}$$

is satisfied. Hence, for every $k \geq 17$, that is, for every number of moves greater than or equal to $2 \cdot 17 = 34$, the particle reaches the point 3 with a probability greater than 0.99. (Compare with the calculations on page 45.)

11. RANDOM WALKS THROUGH A CITY

We return to a consideration of the random walk through a city, which we have described in the Introduction (Fig. 1). We ask whether and when the friends reach the intersection E. We shall prove: If the time of their stroll is without limit, they reach the intersection E, exactly as with all other intersections, with probability 1. Furthermore, we shall estimate the probability that E is reached in a given number of moves.

We state our argument in general form, not assuming that the city must necessarily be of the form shown in Figure 1. Suppose that a traveler goes through the city. If he reaches an intersection from which k streets go out, the probability that he chooses to continue along the first street is p_1, that he chooses the second, $p_2, \ldots,$ and that he chooses the kth street, p_k (the case that he chooses with a certain probability the street by which he arrived is included in this enumeration). We assume that the numbers $p_1, p_2, \ldots, p_k$ are different from zero, and that for a given intersection the probabilities remain constant. This means that the traveler always chooses his further path from a certain intersection with the same probabilities, independent of the direction from which he entered the intersection and the number of times he has traversed it. (In the example given on page 1, we have $k = 4$ and $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$ for every intersection.) If the traveler reaches the edge of the city, he turns around and comes back. We shall assume that he continues in this way indefinitely. We claim that, no matter how the traveler wanders, for each intersection, the probability that he comes to this intersection is 1 (regardless of the location of the intersection).

Proof. Part 1. Let E_0 be *any* intersection; we shall show that the probability that the traveler comes to E_0 is 1. Let the remaining intersections be denoted by $E_1, E_2, \ldots, E_v$.

Suppose the positive integer N and the number $\alpha > 0$ are such that the probability of his arriving at E_0 at least once after completing N moves,[1] regardless of his path, is greater than or equal to α.

We split the totality of moves the traveler makes into groups of N moves each. Let the first subsequence consist of the first to the Nth move, the second, the $(N + 1)$st to the $2N$th, etc. We denote by D_k the event that after making all the moves contained in the first k groups, the traveler has come to E_0 at least once. We denote by $\bar{D}_k$ the contrary event (that the traveler has not reached the point E_0 at any time while making the moves contained in the first k groups).

The aim of the first part of the proof is to establish the inequality

$$\mathbf{P}(\bar{D}_k) \leq \mathbf{P}(\bar{D}_{k-1})(1 - \alpha).$$

To do this, we consider the event $F_{k-1}^{(s)}$: "After making all moves in the first $k - 1$ groups, the traveler has not reached E_0 and is at E_s after $(k - 1)N$ moves."

The events $D_{k-1}, F_{k-1}^{(1)}, F_{k-1}^{(2)}, \ldots, F_{k-1}^{(v)}$ are pairwise mutually exclusive; they also form a complete system. By Property 7,

$$\mathbf{P}(D_k) = \mathbf{P}(D_{k-1})\mathbf{P}(D_k|D_{k-1}) + \mathbf{P}(F_{k-1}^{(1)})\mathbf{P}(D_k|F_{k-1}^{(1)}) + \cdots + \mathbf{P}(F_{k-1}^{(v)})\mathbf{P}(D_k|F_{k-1}^{(v)}). \tag{8}$$

Obviously,

$$\mathbf{P}(D_k|D_{k-1}) = 1. \tag{9}$$

[1] For brevity, we say that the traveler has made a "move" when he goes from one intersection to a neighboring one.

Furthermore, $\mathbf{P}(D_k|F_{k-1}^{(1)})$ is the probability that the traveler reaches the point E_0 with a move in the kth group, upon the condition that he was at E_1 at the beginning of this group. In other words: $\mathbf{P}(D_k|F_{k-1}^{(1)})$ is the probability that E_0 is reached at least once in N moves, if the path begins at E_1. By the choice of α, this probability is greater than or equal to α, that is,

$$\mathbf{P}(D_k|F_{k-1}^{(1)}) \geq \alpha. \tag{10}$$

Likewise,

$$\mathbf{P}(D_k|F_{k-1}^{(2)}) \geq \alpha, \quad \ldots, \quad \mathbf{P}(D_k|F_{k-1}^{(v)}) \geq \alpha. \tag{11}$$

Hence, it follows from Formulas (8)–(11):

$$\mathbf{P}(D_k) \geq \mathbf{P}(D_{k-1}) + \mathbf{P}(F_{k-1}^{(1)})\alpha + \mathbf{P}(F_{k-1}^{(2)})\alpha + \cdots + \mathbf{P}(F_{k-1}^{(v)})\alpha$$

$$= \mathbf{P}(D_{k-1}) + [\mathbf{P}(F_{k-1}^{(1)}) + \mathbf{P}(F_{k-1}^{(2)}) + \cdots + \mathbf{P}(F_{k-1}^{(v)})]\alpha. \tag{12}$$

But the sum of the probabilities of the pairwise mutually exclusive events D_{k-1}, $F_{k-1}^{(1)}$, $F_{k-1}^{(2)}$, $\ldots$, $F_{k-1}^{(v)}$ is equal to 1, since these events form a complete system (Formula (7) of section 3). Hence, the sum in the square brackets is equal to

$$1 - \mathbf{P}(D_{k-1}),$$

and Formula (12) takes the form

$$\mathbf{P}(D_k) \geq \mathbf{P}(D_{k-1}) + [1 - \mathbf{P}(D_{k-1})]\alpha. \tag{13}$$

By Property 3,

$$\mathbf{P}(\bar{D}_k) = 1 - \mathbf{P}(D_k),$$

$$\mathbf{P}(\bar{D}_{k-1}) = 1 - \mathbf{P}(D_{k-1}).$$

From this and (13), we obtain

$$\mathbf{P}(\bar{D}_k) = 1 - \mathbf{P}(D_k) \leq 1 - \mathbf{P}(D_{k-1}) - [1 - \mathbf{P}(D_{k-1})]\alpha$$

$$= \mathbf{P}(\bar{D}_{k-1}) - \mathbf{P}(\bar{D}_{k-1})\alpha = \mathbf{P}(\bar{D}_{k-1})(1 - \alpha).$$

Proof. Part 2. We now seek numbers N and α possessing the desired properties.

We choose an arbitrary intersection E_i. The traveler can reach the intersection E_0 from the intersection E_i by moving along various paths. We choose one of the paths $E_iE_jE_k \ldots E_rE_sE_0$. We denote the number of moves in this path by N_i. Let us calculate the probability α_i that the traveler traverses the path $E_iE_jE_k \ldots E_rE_sE_0$ if he starts at E_i.

We denote the probability that the traveler goes from the intersection E_i along the path E_iE_j by p_{ij}. The probability p_{ij} is the conditional probability that the traveler is at E_j after the nth move, under the condition that he was at E_i after the $(n - 1)$st move. Then the probability that the traveler starting out from E_i takes the path $E_iE_jE_k$ (Fig. 14) is equal to $p_{ij}p_{jk}$. Indeed, by Property 6,

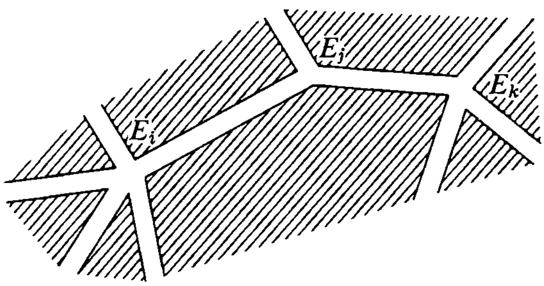

Fig. 14

this probability is the product of the probability that the traveler is at E_j after one move and the probability (under this condition) that he is at E_k after one further move. The first factor is equal to p_{ij}, the second is equal to p_{jk}; hence, the desired probability is equal to $p_{ij}p_{jk}$. In general, the probability that a traveler starting out from E_i chooses the path $E_iE_jE_k \ldots E_rE_sE_0$ is $p_{ij}p_{jk} \ldots p_{rs}p_{s0} = \alpha_i$.

We carry through this calculation for every intersection E_1, E_2, $\ldots E_v$. We thus obtain the numbers $N_1, N_2, \ldots, N_v$; $\alpha_1, \alpha_2, \ldots, \alpha_v$. Let N be the greatest of the numbers $N_1, N_2, \ldots, N_v$, and α be the smallest of the numbers $\alpha_1, \alpha_2, \ldots, \alpha_v$. We must show that N and α have the desired properties.

First of all, the numbers $\alpha_1, \alpha_2, \ldots, \alpha_v$ are all positive; hence, we also have $\alpha > 0$.

Suppose the traveler has traveled N moves from the intersection E_i. He has thereby reached the point E_0 at least once, with a probability greater than or equal to α; the event "the traveler took the path $E_iE_jE_k \ldots E_rE_sE_0$ in the first N_i moves" implies the event "the traveler reached E_0 at least once in N moves" (for if the first event occurs, the traveler reaches E_0 after having made N_i moves). By Property 1, the probability of the second event is greater than or equal to the probability of the first, which is $\alpha_i \geq \alpha$.

Proof. Part 3. We estimate, successively, the probabilities $\mathbf{P}(\bar{D}_1)$, $\mathbf{P}(\bar{D}_2), \ldots, \mathbf{P}(\bar{D}_k), \ldots$.

First of all, $\mathbf{P}(D_1) \geq \alpha$, and hence, by Property 3,

$$\mathbf{P}(\bar{D}_1) = 1 - \mathbf{P}(D_1) \leq 1 - \alpha.$$

Furthermore, we find, with the help of the formula

$$\mathbf{P}(\bar{D}_k) \leq \mathbf{P}(\bar{D}_{k-1})(1 - \alpha),$$

obtained in the first part of the proof, that

$$\left.\begin{array}{l} \mathbf{P}(\bar{D}_2) \leq \mathbf{P}(\bar{D}_1)(1 - \alpha) \leq (1 - \alpha)^2, \\ \mathbf{P}(\bar{D}_3) \leq \mathbf{P}(\bar{D}_2)(1 - \alpha) \leq (1 - \alpha)^3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{P}(\bar{D}_k) \leq \mathbf{P}(\bar{D}_{k-1})(1 - \alpha) \leq (1 - \alpha)^k, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \tag{14}$$

We consider the event $\bar{D}$ (that the traveler never reaches E_0). $\bar{D}$ implies each of the events $\bar{D}_1, \bar{D}_2, \ldots, \bar{D}_k, \ldots$. Hence, $\mathbf{P}(\bar{D})$ is, by Property 1, not greater than the probabilities $\mathbf{P}(\bar{D}_1), \mathbf{P}(\bar{D}_2), \ldots, \mathbf{P}(\bar{D}_k), \ldots$ and, hence, not greater than any of the numbers

$$1 - \alpha, (1 - \alpha)^2, \ldots, (1 - \alpha)^k, \ldots. \tag{15}$$

Since $\alpha > 0$ and $1 - \alpha < 1$, the terms of the sequence (15) form an infinite decreasing geometric progression, which for increasing k becomes smaller than any arbitrary positive number ε. Hence, the probability that E_0 is never reached is smaller than any positive number, and is thus equal to zero. By Property 3, the probability that E_0 will be reached is then equal to 1.

The following theorem has thus been proved:

THEOREM. *The traveler reaches every intersection with a probability of* 1 *regardless of the intersection from which he starts.*

Remark 1. We consider two arbitrary intersections, denote them by E_0 and E_1, and assume that the traveler starts his path at the intersection E_1. Let us calculate the probability of the events that

A_0 : The traveler reaches E_0 at least once;

A_{01} : The traveler reaches E_0 and then the intersection E_1;

A_{010}: The traveler reaches E_0, then reaches E_1, and after that returns to E_0, etc.

By the theorem just proved, we have $\mathbf{P}(A_0) = 1$. By Property 6, the probability of the event A_{01} is equal to the product of $\mathbf{P}(A_0)$ and the probability that the traveler, on going out from E_0, then reaches E_1. By our theorem, both factors are equal to 1; hence, $\mathbf{P}(A_{01}) = 1$. Furthermore, the probability $\mathbf{P}(A_{010})$ is, by the same Property 6, equal to the product of $\mathbf{P}(A_{01})$ and the probability that the traveler, on going out from E_1, finally reaches E_0. From this, it is obvious that

$$\mathbf{P}(A_{010}) = 1.$$

In a similar manner, we show that

$$\mathbf{P}(A_{0101}) = \mathbf{P}(A_{01010}) = \cdots = 1.$$

We pick an arbitrary positive integer k. It follows from what has been proved that:

a) With the probability 1, the traveler returns to the starting point at least k times.

b) With the probability 1, the traveler reaches an arbitrary preassigned intersection E_0 at least k times.

Remark 2. The formula

$$\mathbf{P}(\bar{D}_k) \leq (1 - \alpha)^k$$

(compare with (14)) gives an estimate for the probability of error of the statement that after kN moves the traveler has reached E_0 at least once. We choose an arbitrary probability ε. Then, we can find a number k such that

$$(1 - \alpha)^k < \varepsilon.$$

But then, certainly,

$$\mathbf{P}(\bar{D}_k) < \varepsilon,$$

and the statement that after kN moves the traveler has reached E_0 at least once, has a probability of error less than ε. Thus, even for arbitrarily great demands on the degree of certainty, one can give a number of moves in which the traveler reaches E_0 with practical certainty.

Formula (14) is general; that is, it permits an approximation of the probability $\mathbf{P}(\bar{D}_k)$ for an arbitrary real city. However, it yields only a rough approximation to this probability. We shall apply this approximation to the consideration of an earlier example (see Fig. 13). In this example, one can take $N = 3$ and $\alpha = \frac{1}{4}$. Hence, the probability that after $k \cdot 3$ moves the traveler never reaches the point 3 satisfies the inequality

$$\mathbf{P}(\bar{D}_k) \leq \left(\frac{3}{4}\right)^k.$$

We ask that

$$\left(\frac{3}{4}\right)^k < 0.01.$$

The smallest value of k that satisfies this inequality is equal to 17 (see p. 39). Therefore, the smallest suitable number of subsequences is equal to 17, and we reach the point 3 after $17 \cdot 3 = 51$ moves with a probability of at least 0.99. However, as was shown earlier, even 34 moves are sufficient (not less). Thus, even in this simple case, our approximation yields a result which is less accurate than the exact calculation by a factor of one and one half. In more complicated examples, the result becomes still less precise.

Our traveler need not go on foot; he can use one of the municipal methods of conveyance: streetcar, bus, trolley bus, or subway. At a stop there is a certain probability that he enters, say, a bus. After that, he gets off with a certain probability at each of the following stops and continues with a certain probability.

In this case, the random path of the traveler through the city does not differ in essence from the board game *Circus* mentioned in the Introduction. One can instead name the game *Journey Through a City*. The squares of the board are represented as street intersections, and municipal conveyances substituted for the circus attractions. The sudden motion to another edge of the board corresponds, perhaps, to the journey to the next subway station. We wish to use a device of chance that offers more possibilities than a die. One can, for example, use an urn, and determine the direction of motion by drawing from the urn a piece of paper bearing the designation of an intersection. It is clear, however, that a single urn will not suffice, since the point that we reach depends on the point from which we start. Ideally, one would use as many different urns as there are squares on the board.

12. MARKOV CHAINS

We now consider an arbitrary diagram of points $E_1, E_2, \ldots, E_n$, several of which are joined by arrows that point in the possible directions of motion.

DEFINITION. *A system of n points or states for which we know the possibilities of transitions between them as well as the probabilities that these transitions take place, is called a* Markov chain.

The probability that one goes in one step from E_i to E_j is generally denoted by p_{ij} (in particular, p_{ii} denotes the probability of not moving from E_i on a move).[1] The transition from E_i to E_j is possible if $p_{ij} > 0$ (in this case, we draw an arrow from E_i to E_j).

DEFINITION. *A* Markov chain *is called* irreducible *if one can go from any position E_i to any other position E_j by means of a chain of possible transitions.*

In terms of the arrows, this means that one can go from any point E_i to any other point E_j in the direction of the arrows. Figure 15 gives an example of a chain that is not *irreducible* (it is impossible to go from E_2 to E_1 in the direction of the arrows). Figures 16 to 20 are examples of irreducible chains. We have already considered a *Markov chain* of two states in Problem 5 (Fig. 4).

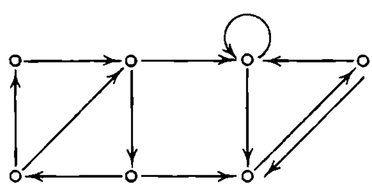

Fig. 15

The following general theorem holds:

THEOREM. *In the motion of a particle through an arbitrary system of states that form an irreducible Markov chain, this particle reaches any state with a probability of* 1 *(independent of its starting point).*

This theorem was already proved, in essence, as a theorem on the random walk of a traveler through a city. The reader can ascertain without difficulty that our proof is also valid, step for step, without an alteration for arbitrary irreducible Markov chains.

The Markov chains are named after the noted Russian mathematician Andrei Andreievich Markov (1856–1922), who discovered

[1] We remark that, for every i,

$$p_{i1} + p_{i2} + \cdots + p_{ii} + \cdots + p_{in} = 1,$$

since $p_{i1}, p_{i2}, \ldots, p_{in}$ are the probabilities of mutually exclusive events (going from E_i to E_1, from E_i to E_2, and finally from E_i to E_n) that form a complete system.

and investigated them. Markov chains are very important because of their applications to science and technology, as well as for probability theory itself.

In the solution of Problem 5 it was remarked that for a chain of two states the probability that after n moves a given position is reached depends less and less on the starting point with increasing n. This fact holds for all irreducible Markov chains (with minor limitations).

13. THE MEETING PROBLEM

We return once more to the city plan in Figure 1. Suppose that, as before, the friends start out on their walk from intersection A, but that this time they go their own ways; that is, each of them tosses two coins independently of the other and chooses his path according to the outcome of his own toss. Will the friends meet again after they have left A? We shall show that with probability 1 this meeting will take place somewhere. Moreover, if any crossing E is fixed beforehand, it can even be asserted that with probability 1 the friends will meet at E.

We restate this problem in general terms as follows:

THEOREM. *Two particles move along an irreducible chain K, beginning their motion at the same time at an arbitrary state. At each move, each particle moves independently of the other, from one state to another. The probability that the two particles meet in an arbitrary preassigned state is equal to* 1.

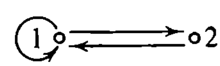

Fig. 16

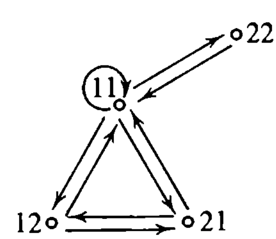

Fig. 17

First, we explain an important relation between Markov chains.

Let the chain K consist of n states E_1, E_2, $\ldots$, E_n. We consider n^2 points, and denote them by $E_{11}, E_{12}, \ldots, E_{1n}, E_{21}, E_{22}, \ldots, E_{2n}, \ldots, E_{n1}, E_{n2}, \ldots, E_{nn}$. We shall denote the state of the pair of particles on the chain K by a marker that is situated on one of the points $E_{11}, \ldots, E_{nn}$. If the first particle is in state E_i, and the second is in state E_k, we place the marker on the point E_{ik}. The passage of the first particle from E_i to E_j occurs with the probability p_{ij}, and the passage of the other particle from E_k to E_l

with the probability p_{kl}. The simultaneous passage of the first particle from E_i to E_j, and the second from E_k to E_l occurs with the probability $p_{ij}p_{kl}$ (by Formula (4), section 2), and, hence, the marker goes from E_{ik} to E_{jl} with the probability $p_{ij}p_{kl}$. We have thus obtained a new Markov chain, which we denote by K^2. The chain K^2 obtained from the chain K shown in Figure 16 can be seen in Figure 17.

According to the general rule, we can draw an arrow from E_{ik} to E_{jl} when the probability of the passage from E_{ik} to E_{jl} is positive. From this it follows that there is an arrow from E_{ik} to E_{jl} if and only if the two probabilities p_{ij} and p_{kl} are positive, that is, if and only if an arrow leads from E_i to E_j and an arrow leads from E_k to E_l. The system of arrows in the chain K^2 is thus formed from the system of arrows in the chain K, the value of p_{ij} being unimportant.

Problem 21. (*a*) The system of arrows of a chain K is shown in Figure 18. Construct the system for the chain K^2;

(*b*) Do the same for the system shown in Figure 19;

(*c*) Do the same for the system shown in Figure 20.

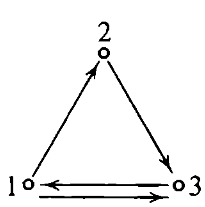

Fig. 18

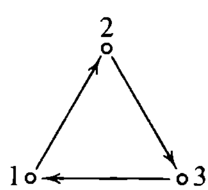

Fig. 19

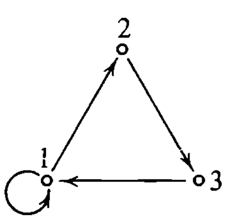

Fig. 20

We consider the totality L of states in K^2 which one can reach from E_{11} by moving along the arrows.

Problem 22. Prove that:

(*a*) The positions $E_{22}, E_{33}, \ldots, E_{nn}$ are contained in L.

(*b*) L is an irreducible chain.

The proof of the meeting theorem can now be given in a few words.

Proof. Let K be an arbitrary irreducible chain, and suppose that the two particles begin their motion at the same time in the state E_{11}. We construct the chain K^2. This may not be irreducible (as was shown, for example, in Problem 21*b*). We form the chain L from K^2. This chain is (as was shown in Problem 22*b*) irreducible, and we can apply the general theorem formulated on page 46. Hence, after leaving the state E_{11}, the marker will reach any state of L with probability 1, including E_{11}, E_{22}, ..., E_{nn}. But this means that our particles meet with probability 1 in any arbitrary preassigned state.

It is proved in exactly the same way that 3 or 4 or, in general, n particles that begin their motion at the same time at a state of an irreducible chain, meet again with probability 1, and furthermore, with probability 1 they reach any arbitrary preassigned state at the same time.

4. Random Walks with Infinitely Many States

The Markov chains that we considered in the previous chapter had finitely many states. We now turn to a consideration of chains with infinitely many states. (We have already met one such chain in the diagram of the random walk on a line.) Exactly as in the previous section, the question of interest to us is whether a particle reaches a given point, and if it does, how fast. The random walk of a particle on an infinite chain differs qualitatively from the motion on a finite chain. For example, one cannot in general assert here that the particle reaches every position with the probability 1 (although this is possible in individual cases).

We limit ourselves to the simplest chains with infinitely many states. First, we investigate the scheme, already familiar to us, of the random walk on a straight line (we shall also call this chain an "infinite path"). Then, we shall consider another simple example of a chain with infinitely many states, an infinitely large city with a checkerboard pattern (Fig. 21). If a traveler reaches any crossing, he continues in any of the four directions with a probability $\frac{1}{4}$.

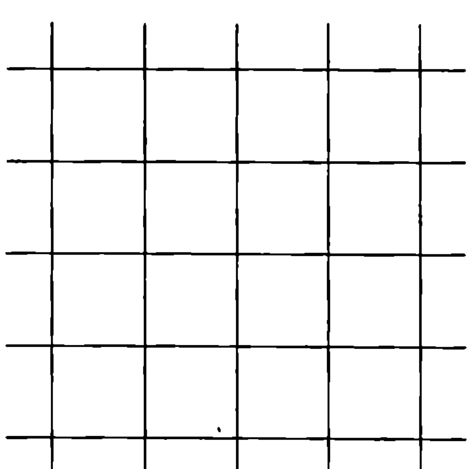

Fig. 21

Problem 23. Let the points 0, ±1, ±2, ±3, ... be marked on a line (Fig. 22). A particle located at a point n might move at the next

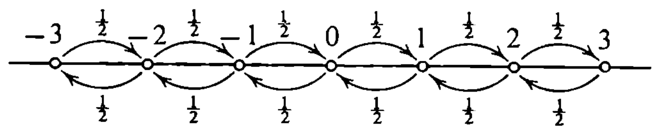

Fig. 22

moment to the point $n + 1$ with the probability $\frac{1}{2}$ and to the point $n - 1$ with the same probability. At the beginning, the particle is situated at the point 0. Find:

(*a*) The probability x that the particle reaches the point 1 at least once.

(*b*) The probability y that the particle reaches the point -1 at least once.

(*c*) The probability z that the particle returns to the point 0 at some time (that is, that it is situated at the zero point at a time other than the beginning of the random walk).

Problem 24. Prove that the particle of Problem 23 reaches every point with the probability 1.

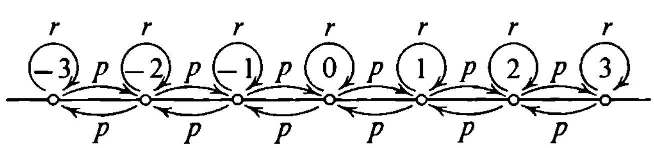

Fig. 23

Problem 25. Suppose that in the diagram in Figure 23 the particle moves from the point n to the point $n + 1$ in a unit of time with probability p, moves to the point $n - 1$ with the same probability p, and remains in the same place with the probability r (in Problem 23, $p = \frac{1}{2}$ and $r = 0$). Prove that the particle reaches every point with the probability 1 for $p > 0$, regardless of the starting point.

It is established for the chains in Problems 24 to 27 that the particle, wherever it might start, reaches every position with probability 1. However, the chain shown in Figure 24 does not have

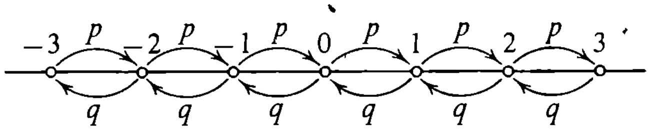

Fig. 24

this property; here, the particle goes from n to $n+1$ with the probability p, and from n to $n-1$ with the probability $q = 1 - p$, where $p > q$. One can prove (we shall refrain from doing so here) that the probability that the particle ever reaches the point -1 after starting out from the point 0 is less than 1.

Problem 26. Using the fact that in Figure 24 the probability of ever reaching the point -1 from the point 0 is smaller than 1, find the probabilities x, y, z (of Problem 23).

In Problem 23, it developed that a particle moving on the line returns to its starting point with a probability 1. How many moves must it make to make this return practically certain? The calculation shows that, for a probability greater than 0.99, several thousand moves are necessary; if this probability is to be greater than 0.999, then several hundred thousand moves are necessary.

One can show that the probability of a return is less than $1 - \varepsilon$ when the particle has made less than $\frac{0.2}{\varepsilon^2}$ moves. When it has made more than $\frac{0.87}{\varepsilon^2}$ moves, this probability is greater than $1 - \varepsilon$.

It was further shown that the particle reaches every point with probability 1. We ask again, how many moves are necessary for the particle to have reached a given point with practical certainty. It is clear that the farther the point of interest to us is from the starting point, the greater the number of moves necessary for the particle to reach it. To reach the point 10 with a probability of 0.99, more than 100,000 moves are necessary; to reach the point 100 with the same probability, tens of millions of moves are necessary.

We set

$$N_1 = \frac{2(\frac{3}{4}k - 1)^2}{\varepsilon^2},$$

$$N_2 = \frac{10k^2}{\varepsilon^2} + \frac{6}{\varepsilon} + 2.$$

One can prove that the probability that the particle reaches the point $2k$ even once is smaller than $1 - \varepsilon$, if it makes less than N_1 moves.[1] If it makes more than N_2 moves, the probability of the same event is greater than $1 - \varepsilon$.

15. THE MEETING PROBLEM

We now turn to the problem of determining the probability that two travelers on an infinite path K will meet (Fig. 22). In the previous chapter, we solved the analogous problem for an arbitrary Markov chain with finitely many states by reducing it to the problem of a single particle moving on a new, more complicated Markov chain with finitely many states. This device was sufficient for the solution of the meeting problem, since we had a general theorem about the random motion of particles that held for all arbitrarily complicated irreducible Markov chains with finitely many states. In the case of infinitely many states, the meeting problem cannot be solved in this fashion, since we lack a corresponding theorem holding for all chains with infinitely many states. We are therefore compelled to solve the meeting problem by a direct method.

In the case of infinitely many states, we cannot reduce the motion of two particles to the motion of one particle on a more complicated chain. On the contrary, we must reduce the problem of motion of a particle on a complicated chain to a problem on the random motion of two (or several) particles on a simpler chain. We shall proceed in this way in the following investigation of an infinitely large city with a checkerboard pattern. We shall reduce the motion of a particle (marker) in this city to the motion of two particles (travelers) on an infinite path K.

We make use of the following auxiliary problem.

[1] This statement holds for $k \geq 2$, $\varepsilon \leq \frac{3}{4} - \frac{1}{k}$.

Problem 27. Prove that the sum

$$S_k = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{k-1} + \frac{1}{k}$$

increases with k and can surpass any arbitrary value.

THEOREM. *Suppose the travelers begin their journey at the same time at point* 0. *Then they meet again at this point with probability* 1.

Proof. Part 1. We consider the events:

A_s: Both travelers reach the point 0 after s moves.

B_s: Both travelers reach the point 0 after s moves, without a previous meeting[1] at this point.

C_s: Up to and including the sth move, the travelers have not met at the point 0.

The events $B_1, B_2, B_3, \ldots, B_{s-1}, B_s, C_s$ are pairwise mutually exclusive and form a complete system. Hence (Property 7),

$$\begin{aligned}\mathbf{P}(A_s) = \mathbf{P}(B_1)\mathbf{P}(A_s|B_1) + \mathbf{P}(B_2)\mathbf{P}(A_s|B_2) + \cdots \\ + \mathbf{P}(B_s)\mathbf{P}(A_s|B_s) + \mathbf{P}(C_s)\mathbf{P}(A_s|C_s). \qquad (1)\end{aligned}$$

Clearly,

$$\mathbf{P}(A_s|C_s) = 0, \quad \mathbf{P}(A_s|B_s) = 1.$$

Furthermore, $\mathbf{P}(A_s|B_i)$ is the probability that both travelers are at point 0 after the sth move, under the condition that they had already met there after the ith move; that is, $\mathbf{P}(A_s|B_i)$ is the probability that the two travelers, starting out from 0, return there after $s - i$ moves. Hence,

$$\mathbf{P}(A_s|B_i) = \mathbf{P}(A_{s-i}).$$

If we now set

$$\mathbf{P}(A_s) = a_s,$$

and

$$\mathbf{P}(B_s) = b_s,$$

equality (1) assumes the form

$$a_s = b_1 a_{s-1} + b_2 a_{s-2} + \cdots + b_{s-1} a_1 + b_s.$$

[1] Except, of course, at the initial moment.

We write this equality out for $s = 1, 2, 3, \ldots, n$ and add (we write out those a_k, b_k that are equal to zero for the sake of clarity):

$$\left.\begin{array}{l}
a_1 = b_1, \\
a_2 = b_2 + a_1 b_1, \\
a_3 = b_3 + a_1 b_2 + a_2 b_1, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_n = b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_{n-1} b_1 \\
\hline
R_n = Q_n + a_1 Q_{n-1} + a_2 Q_{n-2} + \cdots + a_{n-1} Q_1 \,.
\end{array}\right\} \quad (2)$$

Here we have the set

$$R_n = a_1 + a_2 + a_3 + \cdots + a_n,$$
$$Q_n = b_1 + b_2 + b_3 + \cdots + b_n.$$

Clearly, Q_n is the probability of the event D_n: "in the course of the first n moves, the travelers meet at least once at the point 0"; for we have

$$\begin{aligned}
D_n &= B_1 + B_2 + B_3 + \cdots + B_n, \\
\mathbf{P}(D_n) &= \mathbf{P}(B_1) + \mathbf{P}(B_2) + \cdots + \mathbf{P}(B_n) \\
&= b_1 + b_2 + \cdots + b_n \\
&= Q_n.
\end{aligned}$$

The probability Q that the travelers meet at 0 at some time (the probability we are seeking) is, by Property 1, greater than or equal to Q_n; that is,

$$Q \geq Q_n, \quad (3)$$

for every n.

It follows from (2) and (3) that

$$R_n \leq Q + a_1 Q + a_2 Q + \cdots + a_{n-1} Q = (1 + R_{n-1})Q,$$

$$Q \geq \frac{R_n}{1 + R_{n-1}}. \quad (4)$$

The inequality (4) holds for every n.

Proof. Part 2. We shall now show that R_n exceeds every arbitrary number with increasing n; we have $R_n = a_1 + a_2 + a_3 + \cdots + a_n$, where a_n is the probability that the two travelers meet at 0 after n moves. Since the two travelers move independently of each other, a_n can be found by the multiplication formula for probabilities:

$$a_n = w_n^2,$$

where w_n is the probability that one traveler is at point 0 after n moves. The probability w_{2k+1} is, obviously, equal to zero. As was shown in Chapter 2,

$$w_{2k} \geq \frac{1}{\sqrt{4k}}.$$

Hence,

$$a_{2k+1} = 0, \ a_{2k} \geq \frac{1}{4k}.$$

From this it follows that

$$R_{2k+1} = R_{2k} \geq \frac{1}{4} S_k,$$

where $S_k = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k}$. Since, however, S_k increases without bound, R_n also increases without bound.

Proof. Part 3. We now prove that $Q = 1$. First of all, Q, like all probabilities, is either less than or equal to 1. We assume that $Q = 1 - d$, where $d > 0$, and come to a contradiction. By inequality (4),

$$1 - d \geq \frac{R_n}{1 + R_{n-1}} = \frac{R_{n-1} + a_n}{R_{n-1} + 1},$$

$$R_{n-1} + 1 - d(R_{n-1} + 1) \geq R_{n-1} + a_n,$$

$$1 - a_n \geq d(R_{n-1} + 1),$$

$$1 > dR_{n-1},$$

$$R_{n-1} < \frac{1}{d}.$$

But this is false, since R_{n-1} increases without bound for increasing n.

16. THE INFINITELY LARGE CITY WITH A CHECKERBOARD PATTERN

We have already said that we shall reduce the motion through an infinitely large city with a checkerboard pattern to the motion of two travelers on an infinite path K.

Let us imagine that two travelers move on the path K. The chain K^2 is constructed by the process given in the preceding chapter, and consists of the positions (0, 0), (0, 1), (0, -1), (1, 1), (1, -1), (-1, 1), (-1, -1), etc. The chain K^2 is shown in Figure 25.

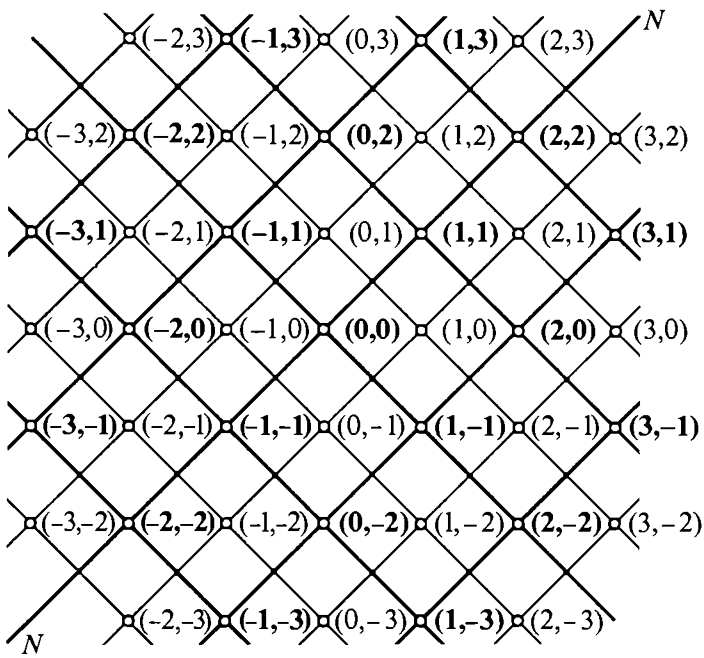

Fig. 25

Suppose that the travelers are at the points m and n of the chain K. The first traveler goes with a probability $\frac{1}{2}$ from m to $m + 1$, and the second traveler goes with the same probability from n to $n + 1$. Hence, the probability that the two travelers go from the point pair (m, n) to the point pair $(m + 1, n + 1)$ is equal to $\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$. This means that the marker that moves on the chain K^2 and reflects the position of both travelers on K goes from (m, n) to $(m + 1, n + 1)$ with the probability $\frac{1}{4}$. The marker goes from (m, n) to each of the other points

$$(m + 1, n - 1), (m - 1, n + 1), (m - 1, n - 1)$$

neighboring (m, n) with the same probability $\frac{1}{4}$.

We remark that at each move the distance between the travelers either changes by two or not at all. Therefore, if at the beginning the travelers are at the points k and l, they can subsequently reach at the same time only such points m and n whose difference $m - n$ is of the same parity as $k - l$. Hence, the chain K^2 is not irreducible; it divides into two irreducible chains: into L, containing the pairs (m, n) with even difference (that is, the points that in Figure 25 are joined by thick lines), and M, containing the pairs with odd difference (that is, the points in Figure 25 that are joined by thin lines).

We assume that the marker starts out from the position (0, 0) of the chain L (corresponding to the fact that both travelers start out from the position 0 of the chain K at the same time). The marker must remain on the boundaries of L in its further motion. The chain L is shown separately in Figure 26, where it is rotated 45° from its

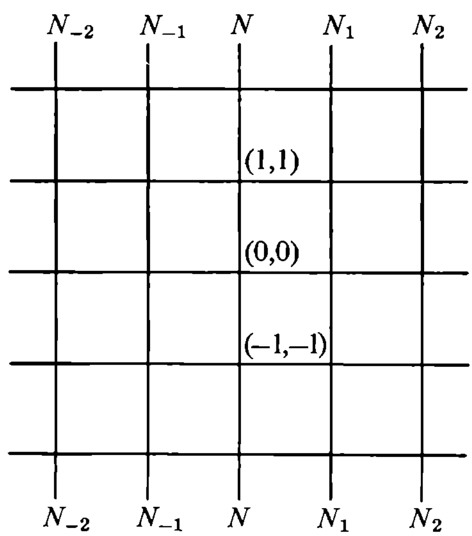

Fig. 26

position in Figure 25. However, Figure 26 coincides with Figure 21. We find, therefore, that the problem of the random walk of two travelers on the path K is equivalent to the problem of the random walk of one traveler through an infinitely large city of checkerboard pattern.

We have proved that two travelers starting out from the point 0 of the path K at the same time, will, with probability 1, again meet at this point. In terms of the city L, this means that a marker that begins its motion at (0, 0) returns there with a probability 1. Thus, simply translating the result of the previous section into this new language, we have proved that *the marker returns to the starting point with probability 1.*

The return of the marker to the starting point takes much longer in the infinitely large city with a checkerboard pattern than on the infinite path; that is, many more moves must be made for this return to be practically certain.

For example, to obtain a probability of return of more than 0.99 the marker must make an astronomical number of moves: more than 10^{88} (in the case of the path K, 10,000 moves sufficed).

In general, if the marker makes less than $10^{\frac{0.9}{\varepsilon}-1.7}$ moves, the probability that it returns at least once to the starting point is less than $1 - \varepsilon$. This probability becomes greater than $1 - \varepsilon$ when the marker makes more than $10^{\frac{2}{\varepsilon}-1.6}$ moves.

We now prove that the marker moving about the infinitely large city with a checkerboard pattern not only returns to the starting point with probability 1, but also reaches any arbitrary preassigned intersection with probability 1.

Let E_0 and E_1 be two neighboring intersections. By x, we denote the probability that the figure reaches the point E_1 at any time, having started out from E_0. Due to the uniform construction of our city, this probability has the same value for any pair of neighboring intersections.

Suppose that the marker starts out from E_0. After the first move it can have reached one of the four intersections 0_1, 0_2, 0_3, 0_4 next to E_0. We consider the events:

A_i: After one move, the marker reaches the point 0_i.

B: The marker returns at any time to E_0.

By Chapter 1, Property 7,

$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B|A_1) + \mathbf{P}(A_2)\mathbf{P}(B|A_2) + \mathbf{P}(A_3)\mathbf{P}(B|A_3) + \mathbf{P}(A_4)\mathbf{P}(B|A_4). \tag{5}$$

From the plan of the city, one sees that

$$\mathbf{P}(A_1) = \mathbf{P}(A_2) = \mathbf{P}(A_3) = \mathbf{P}(A_4) = \frac{1}{4}.$$

The occurrence of B under the condition A_i means that the marker reaches the point E_0 on starting out from the point 0_i (to which it came after the first move).

Hence,

$$\mathbf{P}(B|A_1) = \mathbf{P}(B|A_2) = \mathbf{P}(B|A_3) = \mathbf{P}(B|A_4) = x.$$

Finally, as was proved above, $\mathbf{P}(B) = 1$. If all these values are substituted in Formula (5), we obtain

$$1 = \frac{1}{4}x + \frac{1}{4}x + \frac{1}{4}x + \frac{1}{4}x,$$

and thus, $x = 1$.

Now, let E be an arbitrary intersection of the checkerboard city. We consider one of the paths that lead from E_0 to E, and number in order all of the intersections that belong to this path: E_0, $E_1, E_2, \ldots, E_n = E$.

By D we mean the event that the marker, on starting out from E_0, reaches E_1, then E_2, etc., and finally, $E_n = E$.

The probability that the marker reaches E_{i+1} at some time after starting out from E_i is, as was calculated, equal to 1. If we multiply these probabilities for $i = 0, 1, 2, \ldots, n - 1$ according to the multiplication formula for probabilities, we obtain $\mathbf{P}(D)$. Clearly, the occurrence of the event D implies the occurrence of the event C, that the marker reaches E at some time.

By Chapter 1, Property 1,

$$\mathbf{P}(C) \geq \mathbf{P}(D) = 1,$$

and, consequently, $\mathbf{P}(C) = 1$.

The consideration of the random walk through a city of the kind given makes it possible to sharpen still more the formulation dealing with the infinite path. It was shown that the travelers moving on the infinite path meet at the starting point with a probability 1. We can now assert that the travelers meet at any arbitrary preassigned point n with the probability 1. For this is equivalent to the statement that a marker moving in the city with a checkerboard pattern reaches the intersection (n, n) with probability 1.

This result, however, cannot be generalized to arbitrarily many travelers as in the case of a finite Markov chain. For four or more travelers moving on the path K, the probability that they all meet is smaller than 1. For three travelers, the probability that they all meet somewhere is equal to 1, but the probability that they meet in a preassigned point is less than 1.

We can summarize the results of our investigations of the problems on motion and meeting in a city of checkerboard design in the following theorems:

THEOREM 1. *Let A and B be two arbitrary points on the path K. Two travelers that start out from the point A at the same time meet at the point B with probability* 1.

THEOREM 2. *Let A and B be two arbitrary intersections in the city L. A traveler starting from the intersection A reaches the intersection B with probability* 1.

The proofs of these theorems are intimately connected. The proof of the second theorem arises from the special case of the first, in which the points A and B coincide. The general formulation of Theorem 1 arises, in turn, from Theorem 2.

The special case of Theorem 1 that constitutes the first link in our chain of reasoning demands for its proof some calculations. These calculations can be avoided if one contents himself with a weaker assertion than that of Theorem 1.

THEOREM 1*a*. *Two travelers starting out at the same time from an arbitrary point n of the path K meet again with probability* 1.

If we carry over this statement to the context of our city, it assumes the following form:

THEOREM 2*a*. *A traveler starting out from the intersection (n, n) of the city L reaches, with probability* 1, *a point of the line NN* (Fig. 26).

The proof of this last theorem can be carried through with the aid of Problem 25 alone. Hence, we pose the following problem:

Problem 28. Deduce Theorem 2*a* with the aid of Problem 25.

Concluding Remarks

In this booklet we have considered problems in probability theory. It was our aim to acquaint the reader with the concepts and methods of this unique science by means of examples that were sufficiently clear and which, at the same time, required for their consideration more complicated processes than the simple calculation of the number of desired outcomes. The fact is that the overwhelming majority of the problems that are dealt with by modern probability theory cannot be solved by such simple calculations.

On the other hand, we could not consider more complicated and interesting examples, since they require for their solution tools that lie outside the realm of elementary mathematics. However, we do not wish to leave the reader with the impression that probability theory is the science of children's games and strolls through a city. In reality, the domain of application of probability theory is very great. Probability theory is widely used in technology (radio engineering, the design of telephone networks, quality control of production, etc.), ballistics (the investigation of the scattering of shots), the evaluation of experimental results (the theory of errors). Probability theory also finds important and diversified applications in physics. We have already spoken of some of them (Brownian motion).

Beginning with the work of the great mathematician P. L. Chebyshev (1821–1894), Russian scientists have taken a leading role in the theory of probability. Chebyshev's work has been continued by his students, A. A. Markov (1856–1922) and A. M. Lyapunov (1857–1918). The Soviet school of probability theory has produced such distinguished scholars as S. N. Bernstein (1880–), A. N. Kolmogorov (1903–), and A. Ya. Khinchin (1894–1959).

The following literature is recommended:

Feller, William. *An Introduction to Probability Theory and Its Applications*. Vol. I. 2d ed. (New York: John Wiley & Sons, Inc., 1957.)

Gnedenko, B. V., and Khinchin, A. Ya. *Elementary Introduction to the Theory of Probability*. Translated by W. R. Stahl. (San Francisco and London: W. H. Freeman and Company, 1961.)

Goldberg, Samuel. *Probability, an Introduction*. (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1960.)

Wolf, Frank L. *Elements of Probability and Statistics*. (New York: McGraw-Hill Book Company, Inc., 1962.)

Solutions to Problems

PROBLEM 1. There are 7 possible events, of which 3 satisfy the condition. Hence, the desired conditional probability is equal to $\frac{3}{7}$.

PROBLEM 2. The probability that a six does not come up on one toss is equal to $\frac{5}{6}$. The probability that no six appears on six tosses is, by formula (6) on page 10,

$$\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = \left(\frac{5}{6}\right)^6 = \frac{15{,}625}{46{,}656} \approx 0.34.$$

PROBLEM 3. We denote the desired probability by x. Then, the probability of no hits in n shots is equal to $1 - x$, of no hits on one shot is equal to $1 - p$ (Property 3). Exactly as in the previous exercise we have, by formula (6),

$$1 - x = (1 - p)^n,$$

and thus,

$$x = 1 - (1 - p)^n.$$

PROBLEM 4. We denote the probability of the event A, that the first player wins, by x, the probability of the event B, that the second player wins, by y, and the probability of the event C, that the game never ends, by z.

The events A, B, and C are pairwise mutually exclusive and form a complete system; hence,

$$x + y + z = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) = 1. \tag{1}$$

The probability that the second player wins is equal to the probability that tails comes up on the first toss (this probability is equal to $\frac{1}{2}$) multiplied by the probability that this player wins under this condition (Property 6). Then, however, the second player is in the position of the first, and, hence, this conditional probability is equal to x. Thus,

$$y = \frac{1}{2}x. \tag{2}$$

The probability that the first player wins is found by the formula for the complete probability. We consider the events:

D: "Heads comes up on the first toss."

F: "Tails comes up on the first toss."

The events D and F are mutually exclusive and form a complete system; by Property 7, we have

$$x = \mathbf{P}(A) = \mathbf{P}(D)\mathbf{P}(A|D) + \mathbf{P}(F)\mathbf{P}(A|F).$$

Obviously, $\mathbf{P}(D) = \mathbf{P}(F) = \frac{1}{2}$ and $\mathbf{P}(A|D) = 1$. To calculate $\mathbf{P}(A|F)$, we remark that the first player is in the position of the second, and the second player in the position of the first, when tails comes up on the first toss; the probability of his winning is then y. Hence,

$$x = \frac{1}{2} + \frac{1}{2}y. \tag{3}$$

If we solve the equations (2) and (3), we find $x = \frac{2}{3}$ and $y = \frac{1}{3}$. We obtain $z = 0$ from equation (1).

The fact that the probability of the game continuing indefinitely without an outcome is equal to zero can be obtained by direct calculation (exactly as can be done for the probability of the event that no six comes up in rolling a die).

PROBLEM 5. (*a*) We denote by $p_{11}^{(n)}$ the probability that the particle, on starting out from A, is again at A after n moves, and by $p_{12}^{(n)}$ the probability that the particle, on starting out from A, is at B after n moves. If the particle starts out from point A, the event that the particle is at A after $n-1$ moves and the event that this particle is at B after $n-1$ moves form a complete system of mutually exclusive events; hence, on the basis of Property 7,

$$p_{11}^{(n)} = p_{11}^{(n-1)}p_{11} + p_{12}^{(n-1)}p_{21}.$$

On the other hand, $p_{11}^{(n-1)} + p_{12}^{(n-1)} = 1$ (Property 3); from this it follows that

$$\begin{aligned} p_{11}^{(n)} &= p_{11}^{(n-1)}p_{11} + (1 - p_{11}^{(n-1)})p_{21} \\ &= p_{11}^{(n-1)}p_{11} + p_{21} - p_{11}^{(n-1)}p_{21} \\ &= p_{21} + (p_{11} - p_{21})p_{11}^{(n-1)}. \end{aligned}$$

If we set

$$p_{11} - p_{21} = q,$$

we obtain

$$\begin{aligned} p_{11}^{(n)} &= p_{21} + qp_{11}^{(n-1)} \\ &= p_{21} + q(p_{21} + qp_{11}^{(n-2)}) \\ &= p_{21} + qp_{21} + q^2p_{11}^{(n-2)} \\ &= p_{21} + qp_{21} + q^2(p_{21} + qp_{11}^{(n-3)}) \\ &= p_{21} + qp_{21} + q^2p_{21} + q^3p_{11}^{(n-3)} \\ &= p_{21} + qp_{21} + q^2p_{21} + \cdots + q^{n-2}p_{21} + q^{n-1}p_{11}^{[n-(n-1)]}. \end{aligned}$$

But,

$$p_{11}^{[n-(n-1)]} = p_{11}^{1} = p_{11},$$

from which we obtain

$$p_{11}^{(n)} = p_{21}(1 + q + q^2 + \cdots + q^{n-2}) + q^{n-1}p_{11};$$

or, if we sum by the formula for geometric series,[1]

$$p_{11}^{(n)} = p_{21}\frac{1 - q^{n-1}}{1 - q} + q^{n-1}p_{11} = \frac{p_{21}}{1 - q} + q^{n-1}\left(p_{11} - \frac{p_{21}}{1 - q}\right).$$

By taking into consideration that $p_{11} + p_{12} = 1$ (Property 3), we obtain:

$$1 - q = 1 - p_{11} + p_{21} = p_{12} + p_{21},$$

$$\begin{aligned}
p_{11}^{(n)} &= \frac{p_{21}}{p_{12} + p_{21}} + q^{n-1}\left(p_{11} - \frac{p_{21}}{p_{12} + p_{21}}\right) \\
&= \frac{p_{21}}{p_{12} + p_{21}} + q^{n-1}\frac{p_{12}p_{11} + p_{21}p_{11} - p_{21}}{p_{12} + p_{21}} \\
&= \frac{p_{21}}{p_{12} + p_{21}} + q^{n-1}\frac{p_{12}p_{11} - p_{21}(1 - p_{11})}{p_{12} + p_{21}} \\
&= \frac{p_{21}}{p_{12} + p_{21}} + (p_{11} - p_{21})^{n-1}\frac{p_{12}p_{11} - p_{21}p_{12}}{p_{12} + p_{21}} \\
&= \frac{p_{21}}{p_{12} + p_{21}} + \frac{p_{12}}{p_{12} + p_{21}}(p_{11} - p_{21})^{n}.
\end{aligned}$$

(*b*) To find the probability $p_{21}^{(n)}$ that the point A is reached after n moves on starting out from B, we first calculate $p_{12}^{(n)}$. We find this probability by the formula

$$\begin{aligned}
p_{12}^{(n)} &= 1 - p_{11}^{(n)} \\
&= 1 - \frac{p_{21}}{p_{12} + p_{21}} - \frac{p_{12}}{p_{12} + p_{21}}(p_{11} - p_{21})^{n} \\
&= \frac{p_{12}}{p_{12} + p_{21}} - \frac{p_{12}}{p_{12} + p_{21}}(p_{11} - p_{21})^{n}.
\end{aligned}$$

The probability $p_{21}^{(n)}$ arises from the probability $p_{12}^{(n)}$ when we exchange the points A and B and, correspondingly, replace the index 1 by 2 and 2 by 1. Hence,

$$p_{21}^{(n)} = \frac{p_{21}}{p_{21} + p_{12}} - \frac{p_{21}}{p_{21} + p_{12}}(p_{22} - p_{12})^{n}.$$

[1] See, for example, I. S. Sominskii, *The Method of Mathematical Induction* (Boston: D. C. Heath and Company, 1963), pages 14 and 40.

Remark. The probabilities $p_{11}^{(n)}$ and $p_{21}^{(n)}$ differ by the quantity

$$p_{11}^{(n)} - p_{21}^{(n)} = \frac{p_{12}}{p_{12} + p_{21}}(p_{11} - p_{21})^n + \frac{p_{21}}{p_{12} + p_{21}}(p_{22} - p_{12})^n. \quad (1)$$

With the exception of the degenerate cases

$$p_{11} = 1,\ p_{12} = 0,\ p_{21} = 0,\ p_{22} = 1,$$
$$p_{11} = 0,\ p_{12} = 1,\ p_{21} = 1,\ p_{22} = 0,$$

we always have that

$$-1 < p_{11} - p_{21} < 1,$$
$$-1 < p_{22} - p_{12} < 1.$$

Hence, as n increases, the summands of the right-hand side of equality (1) approach zero, since they are terms of a decreasing geometric progression. Hence, the difference $p_{11}^{(n)} - p_{21}^{(n)}$ also becomes ever closer to zero. In other words, with increasing n, the probability that the particle is at the point A after n moves becomes increasingly independent of the point from which it starts out.

PROBLEM 6. We shall represent the number of balls in the left urn by means of a marker that moves on the line of numbers. At the beginning, the marker is at point a (Fig. 27). In each unit of time it moves with probability $\frac{1}{2}$ to the

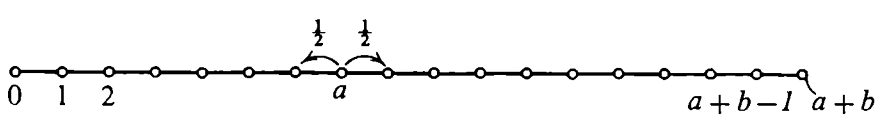

Fig. 27

right (if a ball is moved from the right urn to the left, so that the number of balls in the left urn increases by one), and with the same probability, to the left (if a ball is moved from the left urn to the right, so that the number of balls in the left urn decreases by one). This goes on until the marker reaches the point 0 or the point $(a + b)$ for the first time. If the marker reaches the point 0, the left urn has become empty, and if it reaches the point $a + b$, the right urn has become empty. We denote the probability that the marker reaches the point 0, under the condition that it starts out from the point k, by p_k (p_a is sought). Obviously, $p_0 = 1$ and $p_{a+b} = 0$.

Suppose that the marker is at point k.

We consider the two events:

A: The marker is at $k + 1$ after one move.

B: The marker is at $k - 1$ after one move.

By assumption, $\mathbf{P}(A) = \mathbf{P}(B) = \frac{1}{2}$. If the event A occurs, the probability that the marker reaches the point 0 is equal to p_{k+1}, and if the event B occurs, this probability is equal to p_{k-1}.

By the formula for complete probability, we have

$$p_k = \frac{1}{2}p_{k+1} + \frac{1}{2}p_{k-1},$$

from which we obtain

$$2p_k = p_{k+1} + p_{k-1}, \qquad p_{k+1} - p_k = p_k - p_{k-1}.$$

We denote the constant difference $p_{k+1} - p_k$ by d and write

$$\begin{aligned} p_k - p_{k-1} &= d, \\ p_{k-1} - p_{k-2} &= d, \\ &\cdots\cdots \\ p_2 - p_1 &= d, \\ p_1 - p_0 &= d. \end{aligned}$$

By addition of these equalities we obtain

$$\begin{aligned} p_k - p_0 &= kd, \\ p_k - 1 &= kd, \\ p_k &= 1 + kd, \end{aligned}$$

or, setting $k = a + b$,

$$0 = p_{a+b} = 1 + (a + b)d, \quad d = -\frac{1}{a + b},$$

$$p_k = 1 - \frac{k}{a + b} = \frac{a + b - k}{a + b}.$$

The probability in which we are interested equals

$$p_a = \frac{b}{a + b}.$$

The probability that the right urn becomes empty is then, obviously,

$$p_b = \frac{a}{a + b}.$$

The probability that the experiment ends, that is, that one of the urns becomes empty, is, by Property 2,

$$\frac{b}{a + b} + \frac{a}{a + b} = 1.$$

The probability that the experiment never ends is, by Property 3,

$$1 - 1 = 0$$

Remark. This problem is known in the history of mathematics as the "gambler's ruin problem." Its classical formulation is as follows:

Two people gamble. The probability of a victory for each of them during each game is equal to $\frac{1}{2}$. The first gambler has a rubles, the second b rubles. Play continues until one of the players loses his last ruble. What is the probability of ruin for each of the gamblers?

In Problem 6, the players are replaced by urns and the money by balls.

PROBLEM 7. At each vertex of the cube we write the probability (Fig. 28) that the caterpillar after leaving this vertex gets stuck at the point A. The same probability x is written at the vertices 1 and 5 of Figure 5, since these vertices have the same position with respect to the points A and B. Likewise, the same probability y is written at the vertices 2 and 4. We denote the probability that the caterpillar crawls to A from the points 0 and 3 by z and u, respectively.

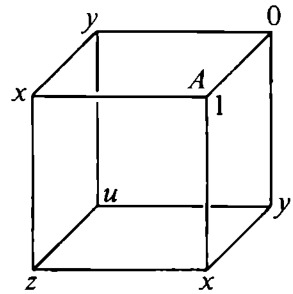

Fig. 28

We now assume that the caterpillar is at the vertex 1 (Fig. 5), and consider the complete system of pairwise mutually exclusive events:

C_1: The caterpillar crawls along the path $1A$,

$$\mathbf{P}(C_1) = \frac{1}{3}.$$

C_2: The caterpillar crawls along the path 12,

$$\mathbf{P}(C_2) = \frac{1}{3}.$$

C_3: The caterpillar crawls along the path 10,

$$\mathbf{P}(C_3) = \frac{1}{3}.$$

The probability that the caterpillar reaches the point A is equal to 1 under the condition C_1, equal to y under the condition C_2, and equal to z under the condition C_3. Hence, (by Property 7),

$$x = \frac{1}{3}\cdot 1 + \frac{1}{3}y + \frac{1}{3}z. \tag{1}$$

Similarly, one finds the three relations

$$y = \frac{1}{3}\cdot 0 + \frac{1}{3}x + \frac{1}{3}u, \tag{2}$$

$$z = \frac{1}{3}x + \frac{1}{3}x + \frac{1}{3}u, \tag{3}$$

$$u = \frac{1}{3}y + \frac{1}{3}y + \frac{1}{3}z. \tag{4}$$

We find

$$x = \frac{9}{14},\ y = \frac{5}{14},\ z = \frac{4}{7},\ u = \frac{3}{7},$$

as solutions of the equations (1), (2), (3), and (4).

The probability z' that the point B is reached if the caterpillar leaves from the point 0 is found by the following considerations. Obviously, the probability that B is reached from the vertex 0 is equal to the probability that the point A is reached from the vertex 3. Hence, $z' = u = \frac{3}{7}$. By the same considerations, we find

$$x' = y = \frac{5}{14},\ y' = x = \frac{9}{14},\ u' = z = \frac{4}{7}.$$

Remark. We remark that

$$x + x' = y + y' = z + z' = u + u' = 1.$$

But, by Property 2, $z + z'$ is the probability that the caterpillar reaches either point starting from the vertex 0, and $1 - (z + z')$ is the probability that the caterpillar reaches neither point starting from the same vertex (Property 3).

Thus, the probability that the caterpillar reaches neither point A nor point B on starting out from the point 0 is equal to zero. One easily sees that it is equally true when any other vertex is chosen as the starting point.

We have already met a similar result in the solution of Problem 6, where it was shown that the probability that the marker reaches neither 0 nor $a + b$ is equal to zero. The probability that the game of Problem 4 never ends is also equal to zero. Finally one can easily ascertain that in the diagram of Problem 5, the probability that the particle never reaches, say, point B is likewise equal to zero (if $p_{12} > 0$).

All these facts are special cases of the general theorem about random walks, which we prove in Chapter 2.

PROBLEM 8. We consider the events:

A: After n moves, the marker is at point k.
B_1: After $n - 1$ moves, the marker is at point $k - 1$.
B_2: After $n - 1$ moves, the marker is at the point $k + 1$.
B_3: After $n - 1$ moves, the marker is neither at $k - 1$ nor at $k + 1$.

By Property 7,

$$\mathbf{P}(A) = \mathbf{P}(B_1)\,\mathbf{P}(A|B_1) + \mathbf{P}(B_2)\,\mathbf{P}(A|B_2) + \mathbf{P}(B_3)\,\mathbf{P}(A|B_3),$$

and we have,

$$\mathbf{P}(A) = Z_n{}^k,\ \mathbf{P}(B_1) = Z_{n-1}{}^{k-1},\ \mathbf{P}(B_2) = Z_{n-1}{}^{k+1},$$

$$\mathbf{P}(A|B_1) = \frac{1}{2},\ \mathbf{P}(A|B_2) = \frac{1}{2},\ \mathbf{P}(A|B_3) = 0.$$

From this, it follows that

$$Z_n{}^k = \frac{Z_{n-1}{}^{k-1} + Z_{n-1}{}^{k+1}}{2}.$$

PROBLEM 9. The elements that stand in the rows of the triangle of probabilities are the probabilities of events that are pairwise mutually exclusive and form a complete system.

PROBLEM 10. For this proof, we write three products, one under the other (compare with (5), (6), and (7) on page 22):

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\cdot\frac{2a-3}{2a-2}\cdot\frac{2a-1}{2a}\cdot\frac{2a-1}{2a}\cdot\frac{2a}{2a+1}$$
$$\times\frac{2a+1}{2a+2}\cdot\ldots\cdot\frac{2k-2}{2k-1}\cdot\frac{2k-1}{2k},$$

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\cdot\frac{2a-3}{2a-2}\cdot\frac{2a-1}{2a}\cdot\frac{2a-1}{2a}\cdot\frac{2a+1}{2a+2}$$
$$\times\frac{2a+1}{2a+2}\cdot\ldots\cdot\frac{2k-1}{2k}\cdot\frac{2k-1}{2k},$$

$$\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\cdot\frac{2a-3}{2a-2}\cdot\frac{2a-1}{2a}\cdot\frac{2a}{2a+1}\cdot\frac{2a+1}{2a+2}$$
$$\times\frac{2a+2}{2a+3}\cdot\ldots\cdot\frac{2k-1}{2k}\cdot\frac{2k}{2k+1}.$$

We see, immediately, that the second product is not smaller than the first, and the third is greater than the second. However, the second product is equal to w_{2k}^2, and, after a few simplifications, the first and third take the form

$$\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\right)^2\cdot\frac{2a-1}{2a}\cdot\frac{2a-1}{2a}\cdot\frac{2a}{2a+1}\cdot\frac{2a+1}{2a+2}\cdot\ldots\cdot\frac{2k-1}{2k}$$
$$=\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\right)^2\cdot\frac{2a-1}{2a}\cdot\frac{2a-1}{2k},$$

$$\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\right)^2\cdot\frac{2a-1}{2a}\cdot\frac{2a}{2a+1}\cdot\frac{2a+1}{2a+2}\cdot\frac{2a+2}{2a+3}\cdot\ldots\cdot\frac{2k}{2k+1}$$
$$=\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\right)^2\cdot\frac{2a-1}{2k+1}.$$

Hence,

$$(2a-1)\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\right)^2\frac{2a-1}{2a}\frac{1}{2k}$$
$$\leq w_{2k}^2<(2a-1)\cdot\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2a-3}{2a-2}\right)^2\cdot\frac{1}{2k},$$

and the desired inequality follows, on taking the square root.

PROBLEM 11. For $k \geq 150$ (inequality (16) on page 24),

$$\frac{1}{\sqrt{3.15k}} \leq w_{2k} < \frac{1}{\sqrt{3.14k}},$$

from which it follows that

$$\frac{1}{1.7749\sqrt{k}} < w_{2k} < \frac{1}{1.7719\sqrt{k}},$$

$$\frac{0.5633}{\sqrt{k}} < w_{2k} < \frac{0.5644}{\sqrt{k}}.$$

In our case, $k = \frac{10{,}000}{2} = 5{,}000$, and, hence,

$$70.710 < \sqrt{k} < 70.711; \; 0.01414 < \frac{1}{\sqrt{k}} < 0.01415.$$

Finally, we obtain

$$0.5633 \times 0.01414 < w_{2k} < 0.5644 \times 0.01415,$$
$$0.007965 < w_{2k} < 0.007987.$$

On rounding off this value to two significant digits, we have

$$w_{2k} = 0.0080.$$

PROBLEM 12. Let $n = 2k$. All terms of the nth row are smaller than the middle term, which satisfies the inequality

$$w_n = w_{2k} < \frac{1}{\sqrt{2k}} = \frac{1}{\sqrt{n}}.$$

Now, let $n = 2k - 1$. There is no middle term in odd-numbered rows, and the largest terms in these rows are the equal terms Z_n^{-1} and Z_n^1. But,

$$Z_n^1 = \frac{Z_n^{-1} + Z_n^1}{2} = Z_{n+1}^0.$$

As just shown,

$$Z_{n+1}^0 < \frac{1}{\sqrt{n+1}}.$$

Hence,

$$Z_n^1 < \frac{1}{\sqrt{n+1}} < \frac{1}{\sqrt{n}};$$

the other terms of the nth row are thus certainly smaller than $\frac{1}{\sqrt{n}}$.

PROBLEM 13. Since the sth term from the middle term of the $2k$th row is $Z_{2k}{}^{2s}$, one can rewrite the inequality (19) on page 26 in the following way:

$$\left(\frac{k+1-s}{k+1}\right)^s \leq \frac{Z_{2k}{}^{2s}}{w_{2k}} \leq \left(\frac{k}{k+s}\right)^s,$$

or,

$$w_{2k}\left(\frac{k+1-s}{k+1}\right)^s \leq Z_{2k}{}^{2s} \leq w_{2k}\left(\frac{k}{k+s}\right)^s.$$

For $k \geq 60$ (see inequality (15) on page 24),

$$\frac{1}{\sqrt{3.18k}} \leq w_{2k} < \frac{1}{\sqrt{3.10k}}.$$

Hence, for $k \geq 60$,

$$\frac{1}{\sqrt{3.18k}}\left(\frac{k+1-s}{k+1}\right)^s \leq Z_{2k}{}^{2s} \leq \frac{1}{\sqrt{3.10k}}\left(\frac{k}{k+s}\right)^s.$$

For an approximation to $Z_{120}{}^{20}$, one must set $s = 10$ and $k = 60$. Then,

$$\frac{1}{\sqrt{3.18 \cdot 60}}\left(\frac{51}{61}\right)^{10} \leq Z_{120}{}^{20} \leq \frac{1}{\sqrt{3.1 \cdot 60}}\left(\frac{60}{70}\right)^{10},$$

and, hence,

$$0.012 < Z_{120}{}^{20} < 0.016.$$

PROBLEM 14. We have

$$(1+p)^r = \underbrace{(1+p)(1+p)\cdots(1+p)}_{r \text{ times}}.$$

Multiplying out these factors, we obtain a number of terms. One of the terms is 1, obtained from multiplying together the 1 from each factor. After that, we obtain p by multiplying p from the first factor by 1 from all the other factors; we likewise obtain p on taking p from the second factor and multiplying by 1 from all the other parentheses, etc. Hence, p occurs r times in the expansion and

$$\begin{aligned}(1+p)^r &= 1 + \underbrace{(p+p+\cdots+p)}_{r \text{ terms}} + \cdots \\ &= 1 + rp + \cdots.\end{aligned}$$

The terms we have not written out are all positive; hence,

$$(1+p)^r \geq 1 + rp.$$

PROBLEM 15. One can take $\frac{2}{\sqrt[3]{0.05}} \cdot \sqrt{n} = 5.429\sqrt{n}$ as the practically certain bound for the displacement after n moves. In particular, for $n = 1{,}000$, 171.7 is the practically certain bound of the displacement. The largest even number that does not exceed 171.7 is 170; hence, one can state with practical certainty that the marker is not more than 170 steps from the starting position after 1,000 moves.

To obtain the practically certain bound for the reduced velocity, one must divide the bound for the deviation by the number of moves. We find that, with practical certainty, the marker has a reduced velocity of at most 0.17 after 1,000 moves.

PROBLEM 16. We set

$$\varepsilon = \left(\frac{2}{\alpha\sqrt{n}}\right)^3.$$

One can say with the probability of error less than ε that after n moves the reduced velocity of the marker is less than $\frac{2}{\sqrt[3]{\varepsilon}\sqrt{n}}$. But,

$$\frac{2}{\sqrt[3]{\varepsilon}\sqrt{n}} = \frac{2}{\frac{2}{\alpha\sqrt{n}}\sqrt{n}} = \alpha.$$

PROBLEM 17. One can assert with a probability of error less than 0.001 that after n moves the reduced velocity of the marker is less than $\frac{2}{\sqrt[3]{0.001}\sqrt{n}}$.

We now choose n in such a way that

$$\frac{2}{\sqrt[3]{0.001}\sqrt{n}} \leq 0.01;$$

from this, it follows that $n \geq 4 \cdot 10^6$.

Hence, one can assert with the probability of error less than 0.001 that the reduced velocity after $4 \cdot 10^6$ moves is smaller than

$$\frac{2}{\sqrt[3]{0.001}\sqrt{4 \cdot 10^6}} = 0.01.$$

PROBLEM 18. In order to assert with probability of error less than 0.01 that the frequency with which heads comes up differs from 0.5 by less than 0.1, we can take any number of tosses that is greater than

$$N = \frac{1}{(0.1)^2 \sqrt[3]{0.01}} = 464.17.$$

It is thus sufficient to make 465 tosses.

PROBLEM 19. If the particle is at m after n moves, it could have arrived there first after one move, after two moves, . . . , or after n moves. Of interest to us is the probability that is expressed, with the help of Formula (5) of section 2, as the sum

$$b_1 + b_2 + \cdots + b_n,$$

where b_k is the probability that m is reached for the first time after k moves. Hence, it suffices to show that the probability b_k does not depend on a barrier.

We now prove that b_k does not depend on what kind of barrier is at the point m. We consider all paths of k moves that begin at the starting point, end at the point m, but otherwise do not touch the point m. These paths are denoted by $A_1, A_2, \ldots, A_s$. We denote by a_i the probability that the particle uses the path A_i for its first k moves. The event that the particle reaches the point m for the first time after k moves is the same as the event that it used either the path A_1, the path $A_2, \ldots$, or, finally, the path A_s. Thus,

$$b_k = a_1 + a_2 + \cdots + a_s.$$

However, none of the paths $A_1, A_2, \ldots, A_s$ goes through m; hence, the probabilities $a_1, a_2, \ldots, a_s$ (and with them also b_k) do not depend on whether or not there is a barrier at the point m or what kind of barrier it may be.

Remark. These considerations have a general character and can be applied to every scheme of a random walk. Hence, the statement formulated in the exercise is valid for arbitrary schemes.

PROBLEM 20. From the relations (1), (2), and (3) on page 38, we find that

$$b_n = \frac{1}{2} b_{n-2} + \frac{1}{4} b_{n-2} = \frac{3}{4} b_{n-2},$$

$$b_{2k} = \left(\frac{3}{4}\right)^k b_0,$$

$$b_{2k+1} = \left(\frac{3}{4}\right)^k b_1.$$

But, $b_0 = 1$ (at the beginning the particle is at 1) and $b_1 = 0$ (after one move, the particle is no longer at 1). Hence,

$$\left.\begin{aligned} b_{2k} &= \left(\frac{3}{4}\right)^k, \\ b_{2k+1} &= 0. \end{aligned}\right\} \qquad (1)$$

It follows from the equalities (3) and (4) on page 38 that

$$d_n = d_{n-1} + \left(\frac{1}{4}\right) b_{n-2}.$$

In addition, by taking into consideration equation (1), we obtain

$$d_{2k+1} = d_{2k} + \frac{1}{4} b_{2k-1} = d_{2k},$$

$$\begin{aligned} d_{2k} &= d_{2k-1} + \frac{1}{4} b_{2k-2} = d_{2k-2} + \frac{1}{4}\left(\frac{3}{4}\right)^{k-1} \\ &= d_{2k-4} + \frac{1}{4}\left(\frac{3}{4}\right)^{k-2} + \frac{1}{4}\left(\frac{3}{4}\right)^{k-1} \\ &= d_0 + \frac{1}{4}\left(\frac{3}{4}\right)^{0} + \frac{1}{4}\left(\frac{3}{4}\right)^{1} + \cdots + \frac{1}{4}\left(\frac{3}{4}\right)^{k-1}. \end{aligned}$$

Since $d_0 = 0$, we finally obtain (by using the formula for the sum of a geometric series)

$$d_{2k} = \frac{1}{4}\left[1 + \frac{3}{4} + \cdots + \left(\frac{3}{4}\right)^{k-1}\right] = \frac{1}{4} \cdot \frac{1 - (\frac{3}{4})^k}{1 - \frac{3}{4}} = 1 - \left(\frac{3}{4}\right)^k.$$

PROBLEM 21.

(*a*)

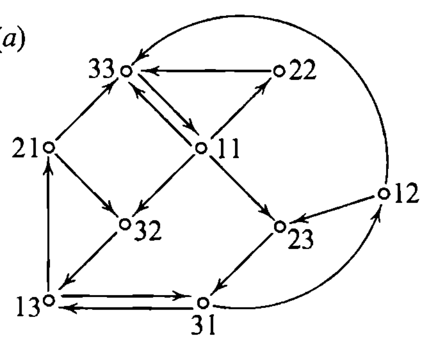

Fig. 29

(*c*)

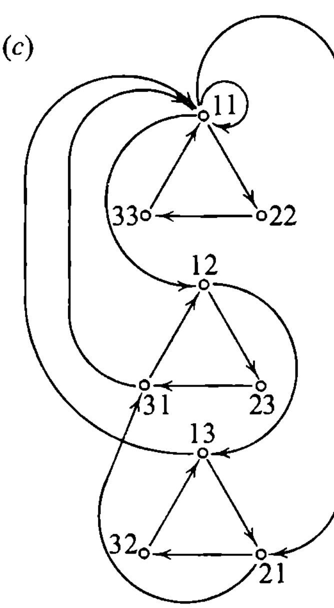

Fig. 31

(*b*)

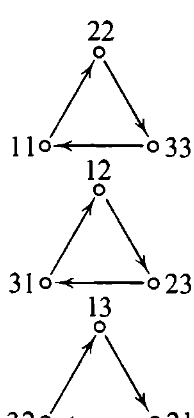

Fig. 30

PROBLEM 22. (*a*) The second particle, just as the first, can also reach E_k from E_1 in z moves.[1] If we relate this path of the two particles to the motion of the marker on K^2, we find that the marker reaches E_{kk} from E_{11} in z moves.

(*b*) It suffices to prove that one can reach E_{11} from an arbitrary position E_{ik} of the chain L. (One can, then, also reach any arbitrary position E_{jl} from it along the path $E_{ik} - E_{11} - E_{jl}$.) Thus, we have to show that one can reach E_{11} from the position E_{ik} belonging to L. Suppose that one reaches the point E_{ik} from E_{11} after s moves. In this case, the first particle goes from E_1 to E_i in s moves, and the second goes from E_1 to E_k (Fig. 32). Now, let E_1

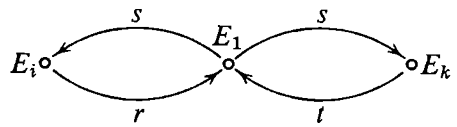

Fig. 32

be reached from E_i in r moves and E_1 from E_k in t moves (the chain K is irreducible!). Then (Fig. 32), the first particle can go from E_i to E_1 in $r + s + t$ moves (on the path $E_i - E_1 - E_k - E_1$) and the second particle from E_k to E_1 in $t + s + r$ moves (on the path $E_k - E_1 - E_i - E_1$); that is, the first particle requires the same number of moves to go from E_i to E_1 as the second does to go from E_k to E_1, namely $z = s + r + t$ moves. If this path of the two particles is carried over to the motion of the marker on K^2, the marker can reach E_{11} from E_{ik} in z moves.

PROBLEM 23. Since all points are fully equivalent, x is really the probability that a particle leaving $n - 1$ reaches n at some time. Likewise, y is the probability that a particle leaving $n + 1$ reaches n.

We assume that the particle starts out from the point 0 and consider the events:

A: The particle returns to the point 0 at some time.

B_1: The particle reaches the point 1 on the first move.

B_2: The particle reaches the point -1 on the first move.

By the formula for complete probability (Property 7, section 3), we have

$$\mathbf{P}(A) = \mathbf{P}(B_1)\mathbf{P}(A|B_1) + \mathbf{P}(B_2)\mathbf{P}(A|B_2). \tag{1}$$

If the condition B_1 is fulfilled, the particle is at the point 1 after one move. For the event A to occur under this condition it is necessary that the particle reach 0 at some time on starting out from 1. The probability for this is equal to y. Hence, $\mathbf{P}(A|B_1) = y$. Likewise, $\mathbf{P}(A|B_2) = x$. Furthermore, $\mathbf{P}(A) = z$ and $\mathbf{P}(B_1) = \mathbf{P}(B_2) = \frac{1}{2}$. If these values are substituted in formula (1), we obtain

$$z = \frac{1}{2}y + \frac{1}{2}x. \tag{2}$$

Furthermore, it is obvious that $x = y$, and, thus, $z = x$.

[1] "Can reach" means "can reach by moving in the direction of the arrows."

To calculate x, we consider the event C (that the particle reaches 1 from 0 at some time). By Property 7, we have

$$x = \mathbf{P}(C) = \mathbf{P}(B_1)\mathbf{P}(C|B_1) + \mathbf{P}(B_2)\mathbf{P}(C|B_2). \tag{3}$$

We have $\mathbf{P}(B_1) = \mathbf{P}(B_2) = \frac{1}{2}$ and $\mathbf{P}(C|B_1) = 1$. Furthermore, $\mathbf{P}(C|B_2)$ is the probability that the particle reaches 1 at some time on starting from -1. We shall calculate this probability. Thus, suppose the particle is located at -1. The probability of the event D (the particle reaches the point 1 at some time) must now be found. For this, we introduce the event F (the particle reaches 0 at some time). Obviously, $FD = D$ (for, the occurrence of D is equivalent to the joint occurrence of the events F and D). Hence, by Property 6,

$$\mathbf{P}(D) = \mathbf{P}(FD) = \mathbf{P}(F)\mathbf{P}(D|F).$$

But, $\mathbf{P}(F) = x$, $\mathbf{P}(D|F) = x$. Hence,

$$\mathbf{P}(D) = x^2. \tag{4}$$

Finally, we find from the formulas (3) and (4) that:

$$x = \frac{1}{2} + \frac{1}{2}x^2, \tag{5}$$

$$x^2 - 2x + 1 = 0,$$

$$(x - 1)^2 = 0,$$

from which it follows that $x = 1$, and, thus, $y = z = 1$.

PROBLEM 24. It must be shown that the particle reaches the point n with probability 1 (for the sake of definiteness, let $n > 0$). For $n = 1$, this was proved in the previous exercise. Let the statement be proved for the point n; we will prove that it then holds for $n + 1$ as well.

Let us consider the event A_{n+1} (the particle reaches $n + 1$ at some time). Obviously, $A_{n+1} = A_nA_{n+1}$; by Property 6, we then have

$$\mathbf{P}(A_{n+1}) = \mathbf{P}(A_n)\mathbf{P}(A_{n+1}|A_n).$$

But, $\mathbf{P}(A_{n+1}|A_n) = x = 1$. Furthermore, $\mathbf{P}(A_n) = 1$, by assumption. Hence, $\mathbf{P}(A_{n+1}) = 1$.

The probability that a preassigned point is reached twice, three times, or, in general, n times, is found exactly as in Remark 1 on page 43; it is equal to 1.

PROBLEM 25. We remark first that $p + p + r = 1$. (See footnote 1 on page 46.) We may denote by x, y, z the probabilities of the same events as in Problem 23. We leave it to the reader to find for himself the equations,

$$z = r + px + py, \tag{1}$$

$$x = p + rx + px^2. \tag{2}$$

(One derives them exactly as (2) and (5) in the solution of Problem 23. One has merely to consider, in addition to B_1 and B_2, the event B_3: the particle remains at the point 0 after the first move). Obviously, $x = y$. Hence,

$$z = r + 2px.$$

We transform (2):

$$\begin{aligned} x &= p + (1 - 2p)x + px^2, \\ 0 &= p - 2px + px^2. \end{aligned}$$

Since $p \neq 0$,

$$1 - 2x + x^2 = 0,$$

from which one obtains $x = 1$ and $z = r + 2p = 1$.

Exactly as in Problem 24, we ascertain that the particle reaches every point with the probability 1.

PROBLEM 26. With the same considerations as were applied in the solution of Problems 23 and 25, we obtain

$$z = py + qx, \tag{1}$$

$$x = p + qx^2, \tag{2}$$

$$y = q + py^2 \tag{3}$$

where, in general, $x \neq y$. The equation (2) yields

$$\begin{aligned} x &= \frac{1 \pm \sqrt{1 - 4pq}}{2q} \\ &= \frac{1 \pm \sqrt{1 - 4p(1 - p)}}{2q} \\ &= \frac{1 \pm \sqrt{(1 - 2p)^2}}{2q} \\ &= \frac{1 \pm (1 - 2p)}{2q} \end{aligned}$$

as solutions; the equation (2) thus has the roots

$$x_1 = \frac{1 + 1 - 2p}{2q} = \frac{2 - 2p}{2q} = \frac{1 - p}{q} = \frac{q}{q} = 1,$$

$$x_2 = \frac{1 - 1 + 2p}{2q} = \frac{2p}{2q} = \frac{p}{q}.$$

Since $\frac{p}{q} > 1$ and x (like every probability) is less than or equal to 1, $x \neq x_2$ and, hence, $x = x_1 = 1$.

We obtain

$$y_1 = 1, \quad y_2 = \frac{q}{p}$$

as a solution of equation (3). By assumption, $y \neq 1$, and, therefore, $y = \frac{q}{p}$. Hence,

$$z = py + qx = 2q.$$

PROBLEM 27. We prove that for $k \geq 2^m$ the inequality $S_k > \frac{m}{2}$ is satisfied. We have

$$S_{2^m} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^m} = 1 + T_1 + T_2 + \cdots + T_m,$$

where

$$T_p = \underbrace{\frac{1}{2^{p-1}+1} + \frac{1}{2^{p-1}+2} + \cdots + \frac{1}{2^p}}_{2^{p-1} \text{ summands}}.$$

Clearly,

$$T_p \geq \underbrace{\frac{1}{2^p} + \frac{1}{2^p} + \cdots + \frac{1}{2^p}}_{2^{p-1} \text{ summands}} = \frac{2^{p-1}}{2^p} = \frac{1}{2}.$$

Hence,

$$S_k \geq S_{2^m} \geq 1 + \underbrace{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}}_{m \text{ summands}} = 1 + \frac{m}{2}.$$

PROBLEM 28. We consider the straight lines N_1N_1, $N_{-1}N_{-1}$, N_2N_2, $N_{-2}N_{-2}$, that run parallel to the line NN (Fig. 26). Let a traveler be at a point of the line N_kN_k. The traveler goes upward on the next move with a probability $\frac{1}{4}$ and downward on the next move with a probability $\frac{1}{4}$; that is, he remains on the line N_kN_k with the probability $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. The traveler goes to the right and reaches the line $N_{k+1}N_{k+1}$ with a probability $\frac{1}{4}$ or goes to the left and reaches the line $N_{k-1}N_{k-1}$ with the probability $\frac{1}{4}$. If we regard the line N_kN_k as one position, we obtain the situation that we have already dealt with in Problem 25. The lines N_kN_k here correspond to the points, with $p = \frac{1}{4}$ and $r = \frac{1}{2}$. By what was proved there, the traveler reaches each line with probability 1; thus, also the line NN.

RANDOM WALKS is a translation of Part Three of *Mathematical Conversations* by E. B. Dynkin and V. A. Uspenskii, which was published in the Russian series, *Library of the Mathematics Circle.* The originality of the exposition and the variety of the problems presented here make this booklet especially useful in stimulating an inventive approach to mathematics. Extensive solutions to all problems are provided.

This booklet deals with some of the more elementary problems in probability theory. The exposition ranges from the simplest examples of a random walk on a line to such more complex examples as random walks through a city, and Markov chains.

The booklet is designed for the reader's active participation, as the problems are carefully integrated with the text and should be solved in sequence. The reader should have a background of high school algebra.

E. B. DYNKIN, a Professor at Moscow State University, is an eminent mathematician and author, whose specialties are higher algebra, topology, and probability theory. V. A. USPENSKII, a Lecturer at Moscow State University, specializes in mathematical logic.