How to Solve It

A New Aspect of Mathematical Method

G. PÓLYA

SECOND EDITION

PENGUIN BOOKS

PENGUIN BOOKS

Published by the Penguin Group
27 Wrights Lane, London w8 5TZ, England
Viking Penguin Inc., 40 West 23rd Street, New York, New York 10010, USA
Penguin Books Australia Ltd, Ringwood, Victoria, Australia
Penguin Books Canada Ltd, 2801 John Street, Markham, Ontario, Canada L3R 1B4
Penguin Books (NZ) Ltd, 182–190 Wairau Road, Auckland 10, New Zealand

Penguin Books Ltd, Registered Offices: Harmondsworth, Middlesex, England

First published in the USA by Princeton University Press 1945
Published by arrangement with Princeton University Press in Penguin Books 1990
1 3 5 7 9 10 8 6 4 2

Made and printed in Great Britain by
Cox and Wyman Ltd, Reading, Berks.

Contents

PENGUIN BOOKS

HOW TO SOLVE IT

George Pólya was born in Budapest in 1887. He began to study law at the University of Budapest in 1905 and eventually changed to philosophy, taking mathematics and physics as part of his philosophy course. In 1912 he gained his Ph.D. in mathematics and, after carrying out post-doctoral work at Göttingen and Paris, he took up a teaching position at the Federal Institute of Technology in Zürich in 1914. Four years later he married Stella Weber, a Swiss. After the outbreak of the Second World War the Pólyas moved to the United States in 1940 and settled in California, at Palo Alto. Here George Pólya took a post at Stanford University and in 1945 wrote *How to Solve It*, a book that has sold more than a million copies and has been translated into seventeen languages. He wrote three other books, and his collected papers fill four volumes. He died in 1985.

Ian Stewart was born in Folkestone in 1945. He graduated from Cambridge University in 1966 with a B.A. degree in mathematics and in 1969 obtained a Ph.D. from the University of Warwick, where he is now a Reader in Mathematics. He has held visiting positions in Germany, New Zealand, Connecticut and Texas and is an active research mathematician in nonlinear dynamics and bifurcation theory. He has written popular articles about mathematics for *New Scientist*, the *Economist*, *The Times* and the *Guardian*, and he occasionally contributes to BBC Radio. His books include *Concepts of Modern Mathematics* (Penguin, 1975), *The Problems of Mathematics* and *Does God Play Dice?: The Mathematics of Chaos* (Penguin, 1990). He lives in Coventry with his wife, two sons, two cats and a variable number of goldfish.

Foreword

by Ian Stewart

During my last couple of years at school I used to haunt the public library, looking for mathematics books. I now realize that this was not a normal activity for a 16-year-old, but at the time it seemed entirely natural, for I was in the grip of an irresistible addiction: mathematics. I still am. Among the scores of books that I devoured was one that was, even then, a classic. You hold it in your hands at this moment: *How to Solve It*.

Any red-blooded mathematician would sign a pact with the Devil for that information. Mathematics is *hard*. So are most things that are worth doing, but mathematics demands an unusual mix of intellectual curiosity and nit-picking pedantry. George Pólya knew that mathematics is hard, but unlike most practitioners of the arcane art, he wanted to make it easier. He was a first-class research mathematician, a brilliant teacher and an able expositor. You won't find that combination often.

Pólya noticed that his students didn't know how to solve problems. Countless thousands of mathematics teachers have observed the same thing, but Pólya's thinking went a little deeper. The difficulty was not that his students didn't *know* enough mathematics, or that they didn't understand the mechanics of using what they knew. He came to the conclusion that what they lacked was the ability to direct their thought processes along fruitful channels. They might be able tacticians, but their sense of strategy was faulty. Could this be because they had no idea that there was any such thing as a strategy for solving a mathematical problem?

Pólya's experience in research led him to recognize that there are a number of general problem-solving techniques, which mathematicians use all the time but seldom articulate. He called them *heuristic strategies*. At the beginning of his book you'll find a skeleton outline, dividing the process into four phases:

- understand the problem
- try to use experience from related problems to plan an attack
- carry out the attack

and, finally,

- ask yourself whether you really believe the answer you've got.

There are fashions in the teaching of mathematics. Problem-solving came into vogue in the 1980s, in part as a reaction against the abstraction of "New Mathematics". Pólya became the unwitting guru of the problem-solvers. The 1980 yearbook of the National Council of Teachers of Mathematics in the USA reads as if it has been marinated in Pólya sauce. More recently the idea has resurfaced as the buzzword *investigation*. Children should not learn facts or methods: they should investigate problems for themselves.

It sounds exciting. But what is the evidence? Can people really use Pólya's heuristic strategies to solve problems? On the face of it, the answer seems to be obvious. The anecdotal evidence from practising mathematicians is massive and convincing. *Yes, that's the way mathematicians think.* On the other hand . . . teachers who coached students for the International Mathematics Olympiads came to a unanimous conclusion. Students *don't* learn to solve problems by following Pólya's heuristic strategies. They learn to solve problems by starting with lots of raw talent and honing it razor-sharp on lots of abrasive problems. Programmers trying to develop Artificial Intelligence found that computers couldn't use heuristics either. The future of Pólya's brainchild began to look less bright.

By now you must be wondering whether I'm going to tell you not to buy the book. Not so. Didn't I just say it's a classic? But, in any case, coaches of would-be Olympians and the Artificial Intelligentsia notwithstanding, Pólya was *right*.

Provided you apply his strategies at the right level.

Let me leave that remark hanging, as one might a pheasant, to mature it, while I tell you a little about the man himself.

"Mathematics is in between . . ."

George Pólya was born in Budapest on 13 December 1887. As a child he did not find mathematics especially interesting: he recalled that of his mathematics teachers "two were despicable and one was good". He was very bright: his position in class at the Gymnasium, or secondary school, varied between second and fourth, and apparently he had no trouble maintaining it. At that time Hungary ran the only national mathematics competition in the world for secondary-school pupils, the Eötvös Competition. All students entering college were encouraged to take part. Pólya went to the test centre, but didn't hand in his paper.

In 1905 he began his studies at the University of Budapest. His mother insisted that he should study law. He stood the boredom for one term. He changed to languages and literature, and then to philosophy. As part of his philosophy course he was advised to take mathematics and physics. As a result he came into contact with two outstanding scientists: the physicist Loránd Eötvös, and the mathematician Lipót Fejér. Fejér's lectures were famous, and they attracted many into mathematics. He used to sit in cafés talking to his students about mathematical problems and telling them tales of famous mathematicians. Pólya was among those hooked. "I thought, I am not good enough for physics and I am too good for philosophy. Mathematics is in between."

In 1912 he gained his Ph.D., which was in mathematics with a minor in physics and chemistry. His thesis research was in probability theory. He did post-doctoral work at Göttingen and Paris, and in 1914 took up a teaching position at the Federal Institute of Technology in Zürich. At the outbreak of war he tried to join the Hungarian army but was rejected because of the after-effects of a childhood soccer injury. Later a more desperate Hungary tried to recall him from Switzerland, but by then Pólya had read Bertrand Russell and decided that war was wrong, and he stayed put. In 1918 he married Stella Weber, a Swiss.

In 1940, one World War later, along with thousands of other European intellectuals who found the activities of Adolf Hitler intolerable, the Pólyas arrived in the United States. After a two-year visiting position at Brown University, Pólya settled down in California at Palo Alto and took a post at Stanford University. Here, in 1945, he wrote *How to Solve It.* The book has since sold more than a million copies and has been translated into seventeen languages. He wrote three other books with an educational bent, and four research monographs. His collected papers fill four volumes. The Mathematical Association of America produced a film of his lectures, called *How to Teach Guessing.* It won the "blue ribbon" in the Educational Film Library Association's Film Festival in 1968.

Pólya's research touched many fields of mathematics, among them complex function theory, combinatorics and probability theory. His classification of the seventeen discrete symmetry groups in the plane ("wallpaper patterns") had a significant influence on the artist Maurits Escher, who studied Pólya's paper carefully, transferring it in full to his notebooks.

He was conservative about fashion in mathematics and in education. When he was in Zürich there was a great deal of interest among mathematicians in "intuitionistic logic", which holds that a proposition P and its double negative not-not-P may be different. Hermann Weyl, an enthusiast for intuitionism, bet Pólya that within fifty years the whole of mathematics would have been rewritten in intuitionistic terms; Pólya begged to differ. The terms of the bet were inscribed on a document, to be opened fifty years later. When it was, Pólya won hands down.

Many of Pólya's sayings have been preserved. Asked which mathematician had influenced him most, he said it was Leonhard Euler (Swiss, 1707–83): "Euler did something that no other great mathematician of his stature did. He explained how he found his results, and I was deeply interested in that. It has to do with my interest in problem-solving." He had many students, among them John von Neumann, one of the fathers of the electronic computer, a man so versatile that his involvement with ENIAC was almost a sideline. Pólya said that von Neumann was "the only student of mine I was ever intimidated by. He was so quick. There was a seminar for advanced students in Zürich that I was teaching and von Neumann was in the class. I came to a certain theorem, and I said it is not proved and it may be difficult. Von Neumann didn't say anything but after five minutes he raised his hand. When I called on him he went to the blackboard and proceeded to write down the proof. After that I was afraid of von Neumann."

Pólya remained interested in mathematics throughout his long life, but he felt his age keenly and often reminded visitors that he was approaching a full century. When computers started to make an impact on the teaching of mathematics, Agnes Wieschenberg discussed them with him. "I am almost 100 years old, too old to learn computers, but if I would live in New York I would listen to your computer classes," said Pólya. Paul Erdös promised him a 100th birthday celebration. He replied, "Maybe 100, but not more."

He died in Palo Alto on 7 September 1985, aged 97.

Guidelines, not rules

Let me return to the question of heuristic strategies. Pólya saw them not as a rigid recipe but as a set of practical guidelines. It is inherent in the nature of guidelines that they don't work if you take them too literally. They are something that you must interpret through the eyes of experience. This explains at once why heuristics alone are of little use in Artificial Intelligence. But if they are embedded in a richer structure of machine inference, it turns out that they perform quite well. Pólya's strategies relate to a much deeper level than the operational surface of mathematics. In the same way, it looks as though the Olympians possessed so much raw talent that they already "knew" the heuristic strategies—and a lot more. Their main problem was to enlarge the background against which those strategies could operate. Educationalists have found that Pólya's basic ideas can be made to work but that the skeleton which he laid down needs to be fleshed out before it leads to a successful teaching procedure. Each of Pólya's general strategies must be expanded into a group of related, but distinct, operational tactics.

For example, one principle is that a general problem can often be illuminated by considering special cases. But

the type of special case will vary from problem to problem. In questions about the summation of n terms of a series, it is often worth working out the first few cases, $n = 1, 2, 3, \ldots$. If the problem is about the divisibility properties of integers, then the "right" special case may be when n is prime or when n has a small number of factors, but this time it seldom helps to look at the cases $n = 1, 2, 3, \ldots$.

Thus another educational problem arises, one that is dealt with implicitly in *How to Solve It* but not perhaps given the emphasis it deserves. If a dozen or so general strategies are replaced by several hundred tactics, how is the student to select which one to use? According to Alan Schoenfeld, "Research now indicates that a large part of what comprises competent problem-solving behaviour consists of the ability to monitor and assess what one does while working problems, and to make the most of the problem-solving resources at one's disposal. It also indicates that students are pretty poor at this, partly because issues of 'resource allocation during thinking' are almost never discussed." In short, the would-be problem-solver needs to develop a feeling for when the attack is making progress or when it's bogged down in a dead end. The buzzword for this is *metacognition*.

Developing this kind of ability requires a mixture of general guidelines, specific methods, plenty of practice on examples and the encouragement of a certain kind of introspection. It's as much an art as a science.

Pólya's four phases

You may feel that the first phase in solving a problem by Pólya's method scarcely needs to be made explicit. *Understand the problem*. Obvious, isn't it?

Well, no. And it's a measure of the man's genius that he recognized the need for it. I never cease to be amazed at

All this from one picture. Inspired by an analogy, by family resemblances in two problems that we *knew* were quite different on a technical, operational level. This is not just problem-solving: it's problem-generation. The point that I want to make clear is that, whatever the role of Pólya-type heuristics, someone has to have the two pieces of the puzzle, Hénon-Heiles and Hopf, in his or her head at the same time, otherwise nobody will spot the possible connection. You can't consider lots of related but solved problems unless your head is full of all sorts of bits of mathematical reasoning.

There's an ironic twist to this particular tale. When we started working on the problem we didn't know much about Hamiltonian dynamics. Some time after we had gained a reasonable understanding of what was going on, and why the Hénon-Heiles system looked like a triangularly symmetric Hopf bifurcation, we were looking in a standard research text and found a simple trick, known to most people in Hamiltonian dynamics, that converts one into the other. Even though periodic solutions to Hamiltonian systems don't arise by Hopf bifurcation, you can still detect them that way: you merely have to tinker with them to make them non-Hamiltonian in just the right manner, by adding a bit of friction at the beginning and then taking it away at the end. If we'd known that to begin with, we would probably have decided that there wasn't anything really interesting involved. But by the time we found out, we'd already pushed the theory well beyond what you could get from Hopf bifurcation, by using quite different ideas to get results that won't come from the simple textbook trick. So in this case some selective ignorance turned out to be crucial. Ignorance is a strategy that, on the whole, I don't think you'll find in Pólya. Fortunately it's one that you don't have to teach: most of us can be pretty

ignorant about almost anything with remarkably little effort.

"Use all the data"

There are differences between the way a student goes about solving problems and the way that a research mathematician does. One of Pólya's pieces of advice to students (and a good one) is: "Did you use all the data?" Does the statement of the problem involve information that you haven't used?

This particular strategy is not so much about the mathematics as about what went on inside the teacher's or examiner's head. I recall W. W. Sawyer's graphic description in *Prelude to Mathematics* of the process of "reconstructing the examiner". The problem must have come from somewhere; somebody thought it up. How? Often that very question suggests a line of attack.

But at research level *you invent the problem yourself.* It doesn't come to you with set hypotheses and a set conclusion, nor is there any guarantee that there is an answer at all. You therefore spend a great deal of time developing a feel for the problem, trying to decide what the essential ideas and concepts should be and how everything fits together. This is, if you like, the planning phase, but it's often not very structured. You can use Pólya's strategy only when you've got a pretty clear idea of what you want to prove. "Hang on, we haven't used the fact that M is seven-dimensional! Something isn't right . . ." You need to develop the ability to select the relevant information.

In order to be able to select the crucial information, and make use of it, mathematicians spend a great deal of their time acquiring both a broad background and a repertoire of more specific tricks. In the education of mathematicians it is important not to concentrate solely on problem-solving methods: the actual mathematical content is important too. And a lot of that content must be taught in a fairly conventional way. Life is too short for students to "discover" by their own "investigations" everything that they ought to know. I'm very happy that they should learn to work out ideas by themselves, but I'm also rather worried that, unless we're careful, the next generation of students will be able to talk the hind leg off a donkey about the thought-processes involved in solving problems but will be able to operate only at a very low level in terms of content. And I also feel there's an inherent contradiction in an "investigation" whose success is measured by whether or not investigators find what the teacher wants them to find. There is a real danger here, and the key to avoiding it is to see problem-solving as a practical tool, part of the mathematician's mental equipment, but not as an end in itself.

Pólya knew all this. He intended the advice in his book to be applied flexibly, informed by intelligence and experience. In a sense it is advice to the teacher, who is assumed to recognize familiar sensations, rather than to the student. It is not, and never was meant to be, an easy method for forcing mathematical ability into human heads. I don't believe there is any such method: the ability comes from within those heads, and it can be nurtured and developed but not forced.

Heuristics in the kitchen

Let me give you an example of a problem that is very easy to solve if the "right" background intuition has been developed but much harder if it has not. Many of you will have seen this particular problem before; if you haven't, you might like to experiment, using Pólya's heuristics, and see whether they work for you.

In 1988 I was involved in the preparation of a series of TV programmes whose aim was to convey the enjoyment of mathematics at an elementary level to a broad audience (it averaged 8 million over the seven-show series). The chosen vehicle was puzzles, solved by individuals or teams in the studio. On the surface it resembled just another game show, but the production team made a big effort to include significant mathematical ideas. Part of my job was to make sure they didn't overstep the bounds of accuracy in the effort to keep the ideas simple. We spent many days thinking about the thought processes involved in solving puzzles: they are very similar to those involved in solving a mathematical problem.

It became clear early on that mathematicians have at their disposal certain reflexes that are not naturally present in most of us—not so much techniques as points of view. A sense of symmetry is one, an ability to discard irrelevant information another. In some ways puzzles are more accurate models for research mathematics than problems in maths exams or textbooks are: in puzzles and research much of the information available has no bearing on the final answer, and the trick is to filter out the noise.

One puzzle was about connecting up items of kitchen equipment to electric plugs: A to a, B to b, C to c. The cables must not cross.

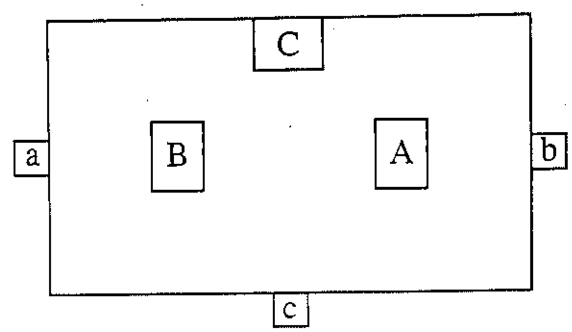

Figure 1

The line of argument—the heuristic—that we extracted from our deliberations is the following. Item C differs from the other two in that it is attached to a wall. So cable Cc cuts the kitchen in two, whereas Aa and Bb don't. If we run cable Cc the wrong way, we can cut off A from a and make our task impossible, like this:

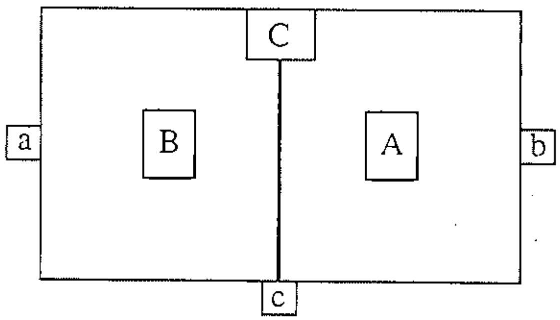

Figure 2

The way the puzzle is set deliberately deceives you into wanting to draw this incorrect connection. That's the art of puzzle-setting. A well-developed mathematical nose can sniff out these red herrings. The key mathematical idea here is *connectedness*. This is fundamentally a topological concept, and at a deeper level the puzzle is really about topology, but we felt that topology is too complicated for prime-time TV. However, I reckon that my audience here can tolerate a lot more, and I'll return to the topological idea below.

Research mathematicians know a very useful heuristic principle: *leave the hard bit to last*. Maybe you can knock off enough easy bits to find out that the problem isn't as hard as you'd thought. There's an equally valid principle that contradicts this completely: *go for the jugular*. It seems not to work as well on this particular puzzle, but sometimes it's better. One of the nice things about heuristics is that

they don't have to be consistent. If one approach doesn't work, try another. You're in trouble only when you've run out of things to try.

Anyway, this suggests connecting up Aa and Bb first. How? However you please. Keep it simple:

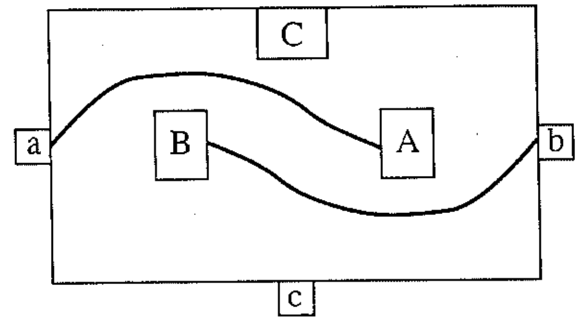

Figure 3

Now you've got a little maze to thread, and the answer's easy:

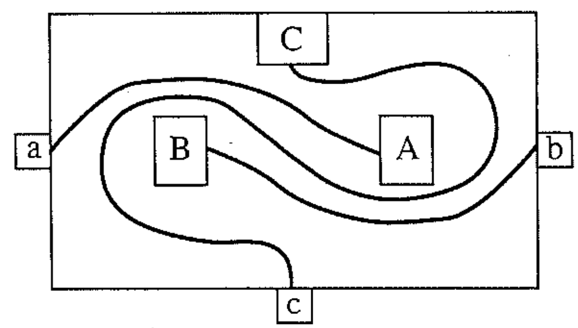

Figure 4

Great!

However, that's not how trained mathematicians will think about it. They have well-developed reflexes that come into play *immediately*. To the professional the puzzle is manifestly an exercise in topology. The shape of the room doesn't matter: a circle or ellipse would do just as well. In fact, you can distort the room any way you choose, as long as you do so by a *continuous* deformation. Not only the outline: *you can distort the floor too.*

In particular you can distort it so that A and B change places:

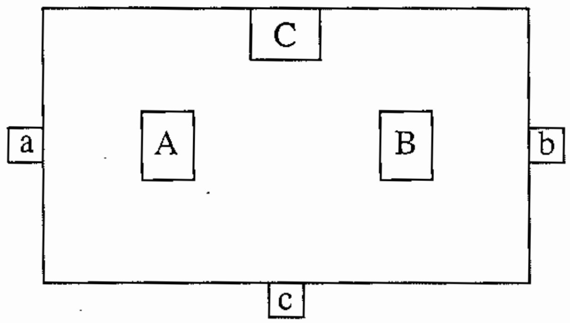

Figure 5

If the problem had been posed in this form to begin with, you'd have considered it a pretty awful puzzle because the answer stares you in the face:

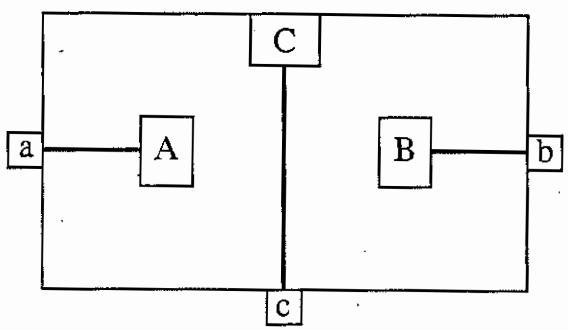

Figure 6

The *coup de grâce*

Except for one final thing: you have to recover the answer to the problem in its original form. Mathematicians are often so pleased that they've "solved it" that they omit to write down this final step. They forget that what they've solved is some other problem that they happen to know is equivalent to the original one. The proof then comes to an abrupt and premature halt. This is a very bad habit because it confuses students no end. It's also a very difficult habit to shed. You stop when you *think* you've reached the end—that is, when the problem rolls over on its back and waves a white flag. When Caesar's given your opponent the thumbs-down, it's easy to forget you still have to administer the *coup de grâce*.

To do that in this case you've got to work out what the deformation that interchanges A and B actually does. Then you can undo it and carry the whole layout of the cables with you. Imagine the kitchen floor is made of putty. Stick two fingers into A and B, and twist your hand through 180 degrees, pulling the putty with you. Everything swirls round, continuously, and A and B change places. To undo this deformation you twist back the other way. You should now be able to see that twisting Figure 6 like this leads directly to Figure 4, the answer we wanted.

This reflex of the professional is not at all obvious. It has to be learnt, and it can be learnt only in a topology course. So what problem-solving resources you have at your disposal depend upon mathematical content as well as problem-solving strategy. New kinds of problem can lead to new strategies; new strategies can solve new kinds of problem. In fact, this is one of the most important ways in which really new mathematics gets developed.

But what, to my mind, is the real lesson that this puzzle teaches us is something that most of you will have to take on trust—though I assure you it's true. Most of us find Figure 1 quite hard and Figure 3 trivial and don't see any connection between the two. Topologists, by contrast, find it very hard to distinguish the two versions at all. *For topologists, the puzzle doesn't exist.*

Does the "swirling" argument flit through their brain so fast that they don't consciously notice it? I don't think so. I think they just know immediately that it can be done. If pushed, they then think for a few seconds and come up with the "swirling" image. I'm no topologist, but that's the order in which I "solved" the puzzle. First I saw that it had to be trivial, then I had to rationalize why.

A student of topology may "know" the theorem that any finite set of points in a rectangle can be moved to any positions you please by a topological transformation—and thus "know" how to solve the puzzle—without being able to lay hands on the actual transformation required. There's a big difference between knowing facts and knowing how those facts fit together. There's also a difficult technical problem for the professional: to prove that what is intuitively obvious here is actually true. A logically rigorous proof that it is possible to interchange A and B by a topological transformation is surprisingly subtle. But all of this is part and parcel of a fully fledged topologist's mode of thought, and it is available immediately and without conscious effort.

If you ever get the chance, listen to mathematicians discussing research over coffee or lunch. Yes, they do that a lot—it's almost impossible to stop them (research, that is, not lunch). Anyway, you'll observe that they jabber incomprehensibly and wave their hands about. The two most important items of mathematical equipment are pen and paper. (And, as the old story has it, a wastepaper basket, thus distinguishing mathematics from philosophy.) It seems to be an almost unfailing rule that in any

gathering of mathematicians nobody has any paper, so they use an old napkin. Nobody has a pen either, but they borrow one from somebody in the business studies department. What gets written on the napkin is cryptic and incoherent. Then suddenly one of them says, "Oh! that's done it!" The others nod sagely. *What's done what?* you think. *How do they know?* But they've all simultaneously come to the same conclusion: they've seen the light at the end of the tunnel. From that moment they all agree that the problem is "solved", even if working out the answer is a twenty-page calculation. The main idea has arrived. The problem has cracked wide open, ripe for plunder.

Weak points

Mathematical problems aren't uniformly impenetrable. They have their weak points, places where you can insert a probe, waggle it, exert some leverage, chip a bit off. Or, sometimes, crack the thing wide open. Mathematicians can sense these things. They don't check the logic of a proof by some laborious calculation with a truth-table: they have the entire plan of campaign mapped out in their heads, and they know when the enemy is on the run.

What we need to do is to equip our students with the ability to sense these weak points, to know whether they're making progress or getting stuck. We need to monitor their thought processes and evaluate them as they go. Pólya's heuristics help: they provide a kind of toolkit for manoeuvring about a problem, seeking an opening. But it's not enough just to cart the toolkit around with you. It's not even enough to be able to name each tool and know how it works. The subtler art is to select the right tool at the right time and to use the tools in such a manner that you get the job finished.

I believe that teachers can make their students aware of this deeper level of subtlety too—at least, to some extent, for those students who are sufficiently receptive. Don't expect silk purses from sows' ears. And be aware that it's a much, much harder task than just following a few rules by rote.

It's clear that Pólya must have known this; indeed, it's what his book is really about. You can gain a great deal of pleasure by reading it and a great deal of insight into the problem-solving mentality. However, Pólya didn't call his book *The Only Way to Solve It*. He didn't envisage it as a "how to" book in the sense of *How to Develop a Super-power Memory*, a recipe book of specific techniques. It is more like *How to Make Friends and Influence People*, general advice to a sensitive and intelligent reader who already has a general comprehension of the nature of the enterprise. It's not some magic prescription or universal nostrum to cure all mathematical ills. But it is the outcome of careful and informed deliberations by one of the great teachers among the ranks of research mathematicians. If you read it with intelligence, sensitivity and plenty of common sense, then within its pages you'll find a great deal of sensible advice and a lot to think about.

Ian Stewart
Warwick University, September 1989

From the Preface to the First Printing

A great discovery solves a great problem but there is a grain of discovery in the solution of any problem. Your problem may be modest; but if it challenges your curiosity and brings into play your inventive faculties, and if you solve it by your own means, you may experience the tension and enjoy the triumph of discovery. Such experiences at a susceptible age may create a taste for mental work and leave their imprint on mind and character for a lifetime.

Thus, a teacher of mathematics has a great opportunity. If he fills his allotted time with drilling his students in routine operations he kills their interest, hampers their intellectual development, and misuses his opportunity. But if he challenges the curiosity of his students by setting them problems proportionate to their knowledge, and helps them to solve their problems with stimulating questions, he may give them a taste for, and some means of, independent thinking.

Also a student whose college curriculum includes some mathematics has a singular opportunity. This opportunity is lost, of course, if he regards mathematics as a subject in which he has to earn so and so much credit and which he should forget after the final examination as quickly as possible. The opportunity may be lost even if the student has some natural talent for mathematics because he, as everybody else, must discover his talents and tastes; he cannot know that he likes raspberry pie if he has never tasted raspberry pie. He may manage to find out, however, that a mathematics problem may be as much fun as a crossword puzzle, or that vigorous mental work may be an exercise as desirable as a fast game of tennis. Having tasted the pleasure in mathematics he will not forget it easily and then there is a good chance that mathematics will become something for him: a hobby, or a tool of his profession, or his profession, or a great ambition.

The author remembers the time when he was a student himself, a somewhat ambitious student, eager to understand a little mathematics and physics. He listened to lectures, read books, tried to take in the solutions and facts presented, but there was a question that disturbed him again and again: "Yes, the solution seems to work, it appears to be correct; but how is it possible to invent such a solution? Yes, this experiment seems to work, this appears to be a fact; but how can people discover such facts? And how could I invent or discover such things by myself?" Today the author is teaching mathematics in a university; he thinks or hopes that some of his more eager students ask similar questions and he tries to satisfy their curiosity. Trying to understand not only the solution of this or that problem but also the motives and procedures of the solution, and trying to explain these motives and procedures to others, he was finally led to write the present book. He hopes that it will be useful to teachers who wish to develop their students' ability to solve problems, and to students who are keen on developing their own abilities.

Although the present book pays special attention to the requirements of students and teachers of mathematics, it should interest anybody concerned with the ways and means of invention and discovery. Such interest may be more widespread than one would assume without reflection. The space devoted by popular newspapers and magazines to crossword puzzles and other riddles seems to show that people spend some time in solving unpractical problems. Behind the desire to solve this or that problem that confers no material advantage, there may be a deeper curiosity, a desire to understand the ways and means, the motives and procedures, of solution.

The following pages are written somewhat concisely, but as simply as possible, and are based on a long and serious study of methods of solution. This sort of study, called *heuristic* by some writers, is not in fashion nowadays but has a long past and, perhaps, some future.

Studying the methods of solving problems, we perceive another face of mathematics. Yes, mathematics has two faces; it is the rigorous science of Euclid but it is also something else. Mathematics presented in the Euclidean way appears as a systematic, deductive science; but mathematics in the making appears as an experimental, inductive science. Both aspects are as old as the science of mathematics itself. But the second aspect is new in one respect; mathematics "in statu nascendi," in the process of being invented, has never before been presented in quite this manner to the student, or to the teacher himself, or to the general public.

The subject of heuristic has manifold connections; mathematicians, logicians, psychologists, educationalists, even philosophers may claim various parts of it as belonging to their special domains. The author, well aware of the possibility of criticism from opposite quarters and keenly conscious of his limitations, has one claim to make: he has some experience in solving problems and in teaching mathematics on various levels.

The subject is more fully dealt with in a more extensive book by the author which is on the way to completion.

Stanford University, August 1, 1944

From the Preface to the Seventh Printing

I am glad to say that I have now succeeded in fulfilling, at least in part, a promise given in the preface to the first printing: The two volumes *Induction and Analogy in Mathematics* and *Patterns of Plausible Inference* which constitute my recent work *Mathematics and Plausible Reasoning* continue the line of thinking begun in *How to Solve It*.

Zurich, August 30, 1954

Preface to the Second Edition

The present second edition adds, besides a few minor improvements, a new fourth part, "Problems, Hints, Solutions."

As this edition was being prepared for print, a study appeared (Educational Testing Service, Princeton, N.J.; *cf. Time,* June 18, 1956) which seems to have formulated a few pertinent observations—they are not new to the people in the know, but it was high time to formulate them for the general public—: " . . . mathematics has the dubious honor of being the least popular subject in the curriculum . . . Future teachers pass through the elementary schools learning to detest mathematics . . . They return to the elementary school to teach a new generation to detest it."

I hope that the present edition, designed for wider diffusion, will convince some of its readers that mathematics, besides being a necessary avenue to engineering jobs and scientific knowledge, may be fun and may also open up a vista of mental activity on the highest level.

Zurich, June 30, 1956

HOW TO SOLVE IT

UNDERSTANDING THE PROBLEM

First.

You have to *understand* the problem.

What is the unknown? What are the data? What is the condition?

Is it possible to satisfy the condition? Is the condition sufficient to determine the unknown? Or is it insufficient? Or redundant? Or contradictory?

Draw a figure. Introduce suitable notation.

Separate the various parts of the condition. Can you write them down?

DEVISING A PLAN

Second.

Find the connection between the data and the unknown. You may be obliged to consider auxiliary problems if an immediate connection cannot be found. You should obtain eventually a *plan* of the solution.

Have you seen it before? Or have you seen the same problem in a slightly different form?

Do you know a related problem? Do you know a theorem that could be useful?

Look at the unknown! And try to think of a familiar problem having the same or a similar unknown.

Here is a problem related to yours and solved before. Could you use it? Could you use its result? Could you use its method? Should you introduce some auxiliary element in order to make its use possible?

Could you restate the problem? Could you restate it still differently? Go back to definitions.

If you cannot solve the proposed problem try to solve first some related problem. Could you imagine a more accessible related problem? A more general problem? A more special problem? An analogous problem? Could you solve a part of the problem? Keep only a part of the condition, drop the other part; how far is the unknown then determined, how can it vary? Could you derive something useful from the data? Could you think of other data appropriate to determine the unknown? Could you change the unknown or the data, or both if necessary, so that the new unknown and the new data are nearer to each other?

Did you use all the data? Did you use the whole condition? Have you taken into account all essential notions involved in the problem?

CARRYING OUT THE PLAN

Third.

Carry out your plan.

Carrying out your plan of the solution, *check each step.* Can you see clearly that the step is correct? Can you prove that it is correct?

LOOKING BACK

Fourth.

Examine the solution obtained.

Can you *check the result?* Can you check the argument?

Can you derive the result differently? Can you see it at a glance?

Can you use the result, or the method, for some other problem?

Introduction

The following considerations are grouped around the preceding list of questions and suggestions entitled "How to Solve It." Any question or suggestion quoted from it will be printed in *italics*, and the whole list will be referred to simply as "the list" or as "our list."

The following pages will discuss the purpose of the list, illustrate its practical use by examples, and explain the underlying notions and mental operations. By way of preliminary explanation, this much may be said: If, using them properly, you address these questions and suggestions to yourself, they may help you to solve your problem. If, using them properly, you address the same questions and suggestions to one of your students, you may help him to solve his problem.

The book is divided into four parts.

The title of the first part is "In the Classroom." It contains twenty sections. Each section will be quoted by its number in heavy type as, for instance, "section **7**." Sections **1** to **5** discuss the "Purpose" of our list in general terms. Sections **6** to **17** explain what are the "Main Divisions, Main Questions" of the list, and discuss a first practical example. Sections **18**, **19**, **20** add "More Examples."

The title of the very short second part is "How to Solve It." It is written in dialogue; a somewhat idealized teacher answers short questions of a somewhat idealized student.

The third and most extensive part is a "Short Dictionary of Heuristic"; we shall refer to it as the "Dictionary."

It contains sixty-seven articles arranged alphabetically. For example, the meaning of the term HEURISTIC (set in small capitals) is explained in an article with this title on page 112. When the title of such an article is referred to within the text it will be set in small capitals. Certain paragraphs of a few articles are more technical; they are enclosed in square brackets. Some articles are fairly closely connected with the first part to which they add further illustrations and more specific comments. Other articles go somewhat beyond the aim of the first part of which they explain the background. There is a key-article on MODERN HEURISTIC. It explains the connection of the main articles and the plan underlying the Dictionary; it contains also directions how to find information about particular items of the list. It must be emphasized that there is a common plan and a certain unity, because the articles of the Dictionary show the greatest outward variety. There are a few longer articles devoted to the systematic though condensed discussion of some general theme; others contain more specific comments, still others cross-references, or historical data, or quotations, or aphorisms, or even jokes.

The Dictionary should not be read too quickly; its text is often condensed, and now and then somewhat subtle. The reader may refer to the Dictionary for information about particular points. If these points come from his experience with his own problems or his own students, the reading has a much better chance to be profitable.

The title of the fourth part is "Problems, Hints, Solutions." It proposes a few problems to the more ambitious reader. Each problem is followed (in proper distance) by a "hint" that may reveal a way to the result which is explained in the "solution."

We have mentioned repeatedly the "student" and the "teacher" and we shall refer to them again and again. It

may be good to observe that the "student" may be a high school student, or a college student, or anyone else who is studying mathematics. Also the "teacher" may be a high school teacher, or a college instructor, or anyone interested in the technique of teaching mathematics. The author looks at the situation sometimes from the point of view of the student and sometimes from that of the teacher (the latter case is preponderant in the first part). Yet most of the time (especially in the third part) the point of view is that of a person who is neither teacher nor student but anxious to solve the problem before him.

PART I. IN THE CLASSROOM

PURPOSE

1. Helping the student. One of the most important tasks of the teacher is to help his students. This task is not quite easy; it demands time, practice, devotion, and sound principles.

The student should acquire as much experience of independent work as possible. But if he is left alone with his problem without any help or with insufficient help, he may make no progress at all. If the teacher helps too much, nothing is left to the student. The teacher should help, but not too much and not too little, so that the student shall have a *reasonable share of the work.*

If the student is not able to do much, the teacher should leave him at least some illusion of independent work. In order to do so, the teacher should help the student discreetly, *unobtrusively.*

The best is, however, to help the student naturally. The teacher should put himself in the student's place, he should see the student's case, he should try to understand what is going on in the student's mind, and ask a question or indicate a step that *could have occurred to the student himself.*

2. Questions, recommendations, mental operations. Trying to help the student effectively but unobtrusively and naturally, the teacher is led to ask the same questions and to indicate the same steps again and again. Thus, in countless problems, we have to ask the question: *What* *is the unknown?* We may vary the words, and ask the same thing in many different ways: What is required? What do you want to find? What are you supposed to seek? The aim of these questions is to focus the student's attention upon the unknown. Sometimes, we obtain the same effect more naturally with a suggestion: *Look at the unknown!* Question and suggestion aim at the same effect; they tend to provoke the same mental operation.

It seemed to the author that it might be worth while to collect and to group questions and suggestions which are typically helpful in discussing problems with students. The list we study contains questions and suggestions of this sort, carefully chosen and arranged; they are equally useful to the problem-solver who works by himself. If the reader is sufficiently acquainted with the list and can see, behind the suggestion, the action suggested, he may realize that the list enumerates, indirectly, *mental operations typically useful for the solution of problems.* These operations are listed in the order in which they are most likely to occur.

3. **Generality** is an important characteristic of the questions and suggestions contained in our list. Take the questions: *What is the unknown? What are the data? What is the condition?* These questions are generally applicable, we can ask them with good effect dealing with all sorts of problems. Their use is not restricted to any subject-matter. Our problem may be algebraic or geometric, mathematical or nonmathematical, theoretical or practical, a serious problem or a mere puzzle; it makes no difference, the questions make sense and might help us to solve the problem.

There is a **restriction,** in fact, but it has nothing to do with the subject-matter. Certain questions and suggestions of the list are applicable to "problems to find" only, not to "problems to prove." If we have a problem of the latter kind we must use different questions; see PROBLEMS TO FIND, PROBLEMS TO PROVE.

4. **Common sense.** The questions and suggestions of our list are general, but, except for their generality, they are natural, simple, obvious, and proceed from plain common sense. Take the suggestion: *Look at the unknown! And try to think of a familiar problem having the same or a similar unknown.* This suggestion advises you to do what you would do anyhow, without any advice, if you were seriously concerned with your problem. Are you hungry? You wish to obtain food and you think of familiar ways of obtaining food. Have you a problem of geometric construction? You wish to construct a triangle and you think of familiar ways of constructing a triangle. Have you a problem of any kind? You wish to find a certain unknown, and you think of familiar ways of finding such an unknown, or some similar unknown. If you do so you follow exactly the suggestion we quoted from our list. And you are on the right track, too; the suggestion is a good one, it suggests to you a procedure which is very frequently successful.

All the questions and suggestions of our list are natural, simple, obvious, just plain common sense; but they state plain common sense in general terms. They suggest a certain conduct which comes naturally to any person who is seriously concerned with his problem and has some common sense. But the person who behaves the right way usually does not care to express his behavior in clear words and, possibly, he cannot express it so; our list tries to express it so.

5. **Teacher and student. Imitation and practice.** There are two aims which the teacher may have in view when addressing to his students a question or a suggestion of the list: First, to help the student to solve the problem

at hand. Second, to develop the student's ability so that he may solve future problems by himself.

Experience shows that the questions and suggestions of our list, appropriately used, very frequently help the student. They have two common characteristics, common sense and generality. As they proceed from plain common sense they very often come naturally; they could have occurred to the student himself. As they are general, they help unobtrusively; they just indicate a general direction and leave plenty for the student to do.

But the two aims we mentioned before are closely connected; if the student succeeds in solving the problem at hand, he adds a little to his ability to solve problems. Then, we should not forget that our questions are general, applicable in many cases. If the same question is repeatedly helpful, the student will scarcely fail to notice it and he will be induced to ask the question by himself in a similar situation. Asking the question repeatedly, he may succeed once in eliciting the right idea. By such a success, he discovers the right way of using the question, and then he has really assimilated it.

The student may absorb a few questions of our list so well that he is finally able to put to himself the right question in the right moment and to perform the corresponding mental operation naturally and vigorously. Such a student has certainly derived the greatest possible profit from our list. What can the teacher do in order to obtain this best possible result?

Solving problems is a practical skill like, let us say, swimming. We acquire any practical skill by imitation and practice. Trying to swim, you imitate what other people do with their hands and feet to keep their heads above water, and, finally, you learn to swim by practicing swimming. Trying to solve problems, you have to observe and to imitate what other people do when solving problems and, finally, you learn to do problems by doing them.

The teacher who wishes to develop his students' ability to do problems must instill some interest for problems into their minds and give them plenty of opportunity for imitation and practice. If the teacher wishes to develop in his students the mental operations which correspond to the questions and suggestions of our list, he puts these questions and suggestions to the students as often as he can do so naturally. Moreover, when the teacher solves a problem before the class, he should dramatize his ideas a little and he should put to himself the same questions which he uses when helping the students. Thanks to such guidance, the student will eventually discover the right use of these questions and suggestions, and doing so he will acquire something that is more important than the knowledge of any particular mathematical fact.

MAIN DIVISIONS, MAIN QUESTIONS

6. Four phases. Trying to find the solution, we may repeatedly change our point of view, our way of looking at the problem. We have to shift our position again and again. Our conception of the problem is likely to be rather incomplete when we start the work; our outlook is different when we have made some progress; it is again different when we have almost obtained the solution.

In order to group conveniently the questions and suggestions of our list, we shall distinguish four phases of the work. First, we have to *understand* the problem; we have to see clearly what is required. Second, we have to see how the various items are connected, how the unknown is linked to the data, in order to obtain the idea of the solution, to make a *plan*. Third, we *carry out* our plan. Fourth, we *look back* at the completed solution, we review and discuss it.

Each of these phases has its importance. It may happen that a student hits upon an exceptionally bright idea and jumping all preparations blurts out with the solution. Such lucky ideas, of course, are most desirable, but something very undesirable and unfortunate may result if the student leaves out any of the four phases without having a good idea. The worst may happen if the student embarks upon computations or constructions without having *understood* the problem. It is generally useless to carry out details without having seen the main connection, or having made a sort of *plan*. Many mistakes can be avoided if, carrying out his plan, the student *checks each step*. Some of the best effects may be lost if the student fails to reexamine and to *reconsider* the completed solution.

7. Understanding the problem. It is foolish to answer a question that you do not understand. It is sad to work for an end that you do not desire. Such foolish and sad things often happen, in and out of school, but the teacher should try to prevent them from happening in his class. The student should understand the problem. But he should not only understand it, he should also desire its solution. If the student is lacking in understanding or in interest, it is not always his fault; the problem should be well chosen, not too difficult and not too easy, natural and interesting, and some time should be allowed for natural and interesting presentation.

First of all, the verbal statement of the problem must be understood. The teacher can check this, up to a certain extent; he asks the student to repeat the statement, and the student should be able to state the problem fluently. The student should also be able to point out the principal parts of the problem, the unknown, the

data, the condition. Hence, the teacher can seldom afford to miss the questions: *What is the unknown? What are the data? What is the condition?*

The student should consider the principal parts of the problem attentively, repeatedly, and from various sides. If there is a figure connected with the problem he should *draw a figure* and point out on it the unknown and the data. If it is necessary to give names to these objects he should *introduce suitable notation;* devoting some attention to the appropriate choice of signs, he is obliged to consider the objects for which the signs have to be chosen. There is another question which may be useful in this preparatory stage provided that we do not expect a definitive answer but just a provisional answer, a guess: *Is it possible to satisfy the condition?*

(In the exposition of Part II [p. 33] "Understanding the problem" is subdivided into two stages: "Getting acquainted" and "Working for better understanding.")

8. Example. Let us illustrate some of the points explained in the foregoing section. We take the following simple problem: *Find the diagonal of a rectangular parallelepiped of which the length, the width, and the height are known.*

In order to discuss this problem profitably, the students must be familiar with the theorem of Pythagoras, and with some of its applications in plane geometry, but they may have very little systematic knowledge in solid geometry. The teacher may rely here upon the student's unsophisticated familiarity with spatial relations.

The teacher can make the problem interesting by making it concrete. The classroom is a rectangular parallelepiped whose dimensions could be measured, and can be estimated; the students have to find, to "measure indirectly," the diagonal of the classroom. The teacher points out the length, the width, and the height of the classroom, indicates the diagonal with a gesture, and enlivens his figure, drawn on the blackboard, by referring repeatedly to the classroom.

The dialogue between the teacher and the students may start as follows:

"What is the unknown?"

"The length of the diagonal of a parallelepiped."

"What are the data?"

"The length, the width, and the height of the parallelepiped."

"Introduce suitable notation. Which letter should denote the unknown?"

"x."

"Which letters would you choose for the length, the width, and the height?"

"a, b, c."

"What is the condition, linking a, b, c, and x?"

"x is the diagonal of the parallelepiped of which a, b, and c are the length, the width, and the height."

"Is it a reasonable problem? I mean, *is the condition sufficient to determine the unknown?"*

"Yes, it is. If we know a, b, c, we know the parallelepiped. If the parallelepiped is determined, the diagonal is determined."

9. Devising a plan. We have a plan when we know, or know at least in outline, which calculations, computations, or constructions we have to perform in order to obtain the unknown. The way from understanding the problem to conceiving a plan may be long and tortuous. In fact, the main achievement in the solution of a problem is to conceive the idea of a plan. This idea may emerge gradually. Or, after apparently unsuccessful trials and a period of hesitation, it may occur suddenly, in a flash, as a "bright idea." The best that the teacher can do for the student is to procure for him, by unobtrusive

help, a bright idea. The questions and suggestions we are going to discuss tend to provoke such an idea.

In order to be able to see the student's position, the teacher should think of his own experience, of his difficulties and successes in solving problems.

We know, of course, that it is hard to have a good idea if we have little knowledge of the subject, and impossible to have it if we have no knowledge. Good ideas are based on past experience and formerly acquired knowledge. Mere remembering is not enough for a good idea, but we cannot have any good idea without recollecting some pertinent facts; materials alone are not enough for constructing a house but we cannot construct a house without collecting the necessary materials. The materials necessary for solving a mathematical problem are certain relevant items of our formerly acquired mathematical knowledge, as formerly solved problems, or formerly proved theorems. Thus, it is often appropriate to start the work with the question: *Do you know a related problem?*

The difficulty is that there are usually too many problems which are somewhat related to our present problem, that is, have some point in common with it. How can we choose the one, or the few, which are really useful? There is a suggestion that puts our finger on an essential common point: *Look at the unknown! And try to think of a familiar problem having the same or a similar unknown.*

If we succeed in recalling a formerly solved problem which is closely related to our present problem, we are lucky. We should try to deserve such luck; we may deserve it by exploiting it. *Here is a problem related to yours and solved before. Could you use it?*

The foregoing questions, well understood and seriously considered, very often help to start the right train of ideas; but they cannot help always, they cannot work magic. If they do not work, we must look around for some other appropriate point of contact, and explore the various aspects of our problem; we have to vary, to transform, to modify the problem. *Could you restate the problem?* Some of the questions of our list hint specific means to vary the problem, as generalization, specialization, use of analogy, dropping a part of the condition, and so on; the details are important but we cannot go into them now. Variation of the problem may lead to some appropriate auxiliary problem: *If you cannot solve the proposed problem try to solve first some related problem.*

Trying to apply various known problems or theorems, considering various modifications, experimenting with various auxiliary problems, we may stray so far from our original problem that we are in danger of losing it altogether. Yet there is a good question that may bring us back to it: *Did you use all the data? Did you use the whole condition?*

10. Example. We return to the example considered in section 8. As we left it, the students just succeeded in understanding the problem and showed some mild interest in it. They could now have some ideas of their own, some initiative. If the teacher, having watched sharply, cannot detect any sign of such initiative he has to resume carefully his dialogue with the students. He must be prepared to repeat with some modification the questions which the students do not answer. He must be prepared to meet often with the disconcerting silence of the students (which will be indicated by dots . . . . .).

"Do you know a related problem?"

. . . . .

"Look at the unknown! Do you know a problem having the same unknown?"

. . . . .

"Well, *what is the unknown?*"

"The diagonal of a parallelepiped."

"Do you know any *problem with the same unknown?*"

"No. We have not had any problem yet about the diagonal of a parallelepiped."

"Do you know any *problem with a similar unknown?*"

. . . . .

"You see, the diagonal is a segment, the segment of a straight line. Did you never solve a problem whose unknown was the length of a line?"

"Of course, we have solved such problems. For instance, to find a side of a right triangle."

"Good! *Here is a problem related to yours and solved before. Could you use it?*"

. . . . .

"You were lucky enough to remember a problem which is related to your present one and which you solved

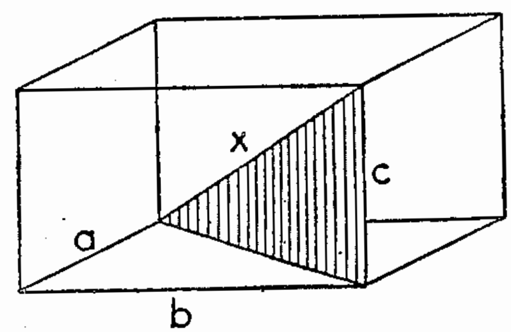

FIG. 1

before. Would you like to use it? *Could you introduce some auxiliary element in order to make its use possible?*"

. . . . .

"Look here, the problem you remembered is about a triangle. Have you any triangle in your figure?"

Let us hope that the last hint was explicit enough to provoke the idea of the solution which is to introduce a right triangle, (emphasized in Fig. 1) of which the required diagonal is the hypotenuse. Yet the teacher should be prepared for the case that even this fairly explicit hint is insufficient to shake the torpor of the students; and so he should be prepared to use a whole gamut of more and more explicit hints.

"Would you like to have a triangle in the figure?"

"What sort of triangle would you like to have in the figure?"

"You cannot find yet the diagonal; but you said that you could find the side of a triangle. Now, what will you do?"

"Could you find the diagonal, if it were a side of a triangle?"

When, eventually, with more or less help, the students succeed in introducing the decisive auxiliary element, the right triangle emphasized in Fig. 1, the teacher should convince himself that the students see sufficiently far ahead before encouraging them to go into actual calculations.

"I think that it was a good idea to draw that triangle. You have now a triangle; but have you the unknown?"

"The unknown is the hypotenuse of the triangle; we can calculate it by the theorem of Pythagoras."

"You can, if both legs are known; but are they?"

"One leg is given, it is c. And the other, I think, is not difficult to find. Yes, the other leg is the hypotenuse of another right triangle."

"Very good! Now I see that you have a plan."

11. Carrying out the plan. To devise a plan, to conceive the idea of the solution is not easy. It takes so much to succeed; formerly acquired knowledge, good mental habits, concentration upon the purpose, and one more thing: good luck. To carry out the plan is much easier; what we need is mainly patience.

The plan gives a general outline; we have to convince

ourselves that the details fit into the outline, and so we have to examine the details one after the other, patiently, till everything is perfectly clear, and no obscure corner remains in which an error could be hidden.

If the student has really conceived a plan, the teacher has now a relatively peaceful time. The main danger is that the student forgets his plan. This may easily happen if the student received his plan from outside, and accepted it on the authority of the teacher; but if he worked for it himself, even with some help, and conceived the final idea with satisfaction, he will not lose this idea easily. Yet the teacher must insist that the student should *check each step.*

We may convince ourselves of the correctness of a step in our reasoning either "intuitively" or "formally." We may concentrate upon the point in question till we see it so clearly and distinctly that we have no doubt that the step is correct; or we may derive the point in question according to formal rules. (The difference between "insight" and "formal proof" is clear enough in many important cases; we may leave further discussion to philosophers.)

The main point is that the student should be honestly convinced of the correctness of each step. In certain cases, the teacher may emphasize the difference between "seeing" and "proving": *Can you see clearly that the step is correct?* But can you also *prove that the step is correct?*

12. Example. Let us resume our work at the point where we left it at the end of section **10**. The student, at last, has got the idea of the solution. He sees the right triangle of which the unknown x is the hypotenuse and the given height c is one of the legs; the other leg is the diagonal of a face. The student must, possibly, be urged to introduce suitable notation. He should choose y to denote that other leg, the diagonal of the face whose sides are a and b. Thus, he may see more clearly the idea of the solution which is to introduce an auxiliary problem whose unknown is y. Finally, working at one right triangle after the other, he may obtain (see Fig. 1)

$$x^2 = y^2 + c^2$$
$$y^2 = a^2 + b^2$$

and hence, eliminating the auxiliary unknown y,

$$x^2 = a^2 + b^2 + c^2$$
$$x = \sqrt{a^2 + b^2 + c^2}\,.$$

The teacher has no reason to interrupt the student if he carries out these details correctly except, possibly, to warn him that he should *check each step*. Thus, the teacher may ask:

"Can you *see clearly* that the triangle with sides x, y, c is a right triangle?"

To this question the student may answer honestly "Yes" but he could be much embarrassed if the teacher, not satisfied with the intuitive conviction of the student, should go on asking:

"But can you *prove* that this triangle is a right triangle?"

Thus, the teacher should rather suppress this question unless the class has had a good initiation in solid geometry. Even in the latter case, there is some danger that the answer to an incidental question may become the main difficulty for the majority of the students.

13. Looking back. Even fairly good students, when they have obtained the solution of the problem and written down neatly the argument, shut their books and look for something else. Doing so, they miss an important and instructive phase of the work. By looking back at the completed solution, by reconsidering and reexamining the result and the path that led to it, they could consolidate their knowledge and develop their ability to solve problems. A good teacher should understand and impress on his students the view that no problem whatever is completely exhausted. There remains always something to do; with sufficient study and penetration, we could improve any solution, and, in any case, we can always improve our understanding of the solution.

The student has now carried through his plan. He has written down the solution, checking each step. Thus, he should have good reasons to believe that his solution is correct. Nevertheless, errors are always possible, especially if the argument is long and involved. Hence, verifications are desirable. Especially, if there is some rapid and intuitive procedure to test either the result or the argument, it should not be overlooked. *Can you check the result? Can you check the argument?*

In order to convince ourselves of the presence or of the quality of an object, we like to see and to touch it. And as we prefer perception through two different senses, so we prefer conviction by two different proofs: *Can you derive the result differently?* We prefer, of course, a short and intuitive argument to a long and heavy one: *Can you see it at a glance?*

One of the first and foremost duties of the teacher is not to give his students the impression that mathematical problems have little connection with each other, and no connection at all with anything else. We have a natural opportunity to investigate the connections of a problem when looking back at its solution. The students will find looking back at the solution really interesting if they have made an honest effort, and have the consciousness of having done well. Then they are eager to see what else they could accomplish with that effort, and how they could do equally well another time. The teacher should encourage the students to imagine cases in which they could utilize again the procedure used, or apply the result obtained. *Can you use the result, or the method, for some other problem?*

14. **Example.** In section 12, the students finally obtained the solution: If the three edges of a rectangular parallelogram, issued from the same corner, are a, b, c, the diagonal is

$$\sqrt{a^2 + b^2 + c^2}.$$

Can you check the result? The teacher cannot expect a good answer to this question from inexperienced students. The students, however, should acquire fairly early the experience that problems "in letters" have a great advantage over purely numerical problems; if the problem is given "in letters" its result is accessible to several tests to which a problem "in numbers" is not susceptible at all. Our example, although fairly simple, is sufficient to show this. The teacher can ask several questions about the result which the students may readily answer with "Yes"; but an answer "No" would show a serious flaw in the result.

"*Did you use all the data?* Do all the data a, b, c appear in your formula for the diagonal?"

"Length, width, and height play the same role in our question; our problem is symmetric with respect to a, b, c. Is the expression you obtained for the diagonal symmetric in a, b, c? Does it remain unchanged when a, b, c are interchanged?"

"Our problem is a problem of solid geometry: to find the diagonal of a parallelepiped with given dimensions a, b, c. Our problem is analogous to a problem of plane geometry: to find the diagonal of a rectangle with given dimensions a, b. Is the result of our 'solid' problem analogous to the result of the 'plane' problem?"

"If the height c decreases, and finally vanishes, the parallelepiped becomes a parallelogram. If you put $c = 0$ in your formula, do you obtain the correct formula for the diagonal of the rectangular parallelogram?"

"If the height c increases, the diagonal increases. Does your formula show this?"

"If all three measures a, b, c of the parallelepiped increase in the same proportion, the diagonal also increases in the same proportion. If, in your formula, you substitute $12a, 12b, 12c$ for a, b, c respectively, the expression of the diagonal, owing to this substitution, should also be multiplied by 12. Is that so?"

"If a, b, c are measured in feet, your formula gives the diagonal measured in feet too; but if you change all measures into inches, the formula should remain correct. Is that so?"

(The two last questions are essentially equivalent; see TEST BY DIMENSION.)

These questions have several good effects. First, an intelligent student cannot help being impressed by the fact that the formula passes so many tests. He was convinced before that the formula is correct because he derived it carefully. But now he is more convinced, and his gain in confidence comes from a different source; it is due to a sort of "experimental evidence." Then, thanks to the foregoing questions, the details of the formula acquire new significance, and are linked up with various facts. The formula has therefore a better chance of being remembered, the knowledge of the student is consolidated. Finally, these questions can be easily transferred to similar problems. After some experience with similar problems, an intelligent student may perceive the underlying general ideas: use of all relevant data, variation of the data, symmetry, analogy. If he gets into the habit of directing his attention to such points, his ability to solve problems may definitely profit.

Can you check the argument? To recheck the argument step by step may be necessary in difficult and important cases. Usually, it is enough to pick out "touchy" points for rechecking. In our case, it may be advisable to discuss retrospectively the question which was less advisable to discuss as the solution was not yet attained: Can you *prove* that the triangle with sides x, y, c is a right triangle? (See the end of section **12.**)

Can you use the result or the method for some other problem? With a little encouragement, and after one or two examples, the students easily find applications which consist essentially in giving some *concrete interpretation* to the abstract mathematical elements of the problem. The teacher himself used such a concrete interpretation as he took the room in which the discussion takes place for the parallelepiped of the problem. A dull student may propose, as application, to calculate the diagonal of the cafeteria instead of the diagonal of the classroom. If the students do not volunteer more imaginative remarks, the teacher himself may put a slightly different problem, for instance: "Being given the length, the width, and the height of a rectangular parallelepiped, find the distance of the center from one of the corners."

The students may use the *result* of the problem they just solved, observing that the distance required is one half of the diagonal they just calculated. Or they may use the *method,* introducing suitable right triangles (the latter alternative is less obvious and somewhat more clumsy in the present case).

After this application, the teacher may discuss the configuration of the four diagonals of the parallelepiped, and the six pyramids of which the six faces are the bases, the center the common vertex, and the semidiagonals the lateral edges. When the geometric imagination of the students is sufficiently enlivened, the teacher should come

back to his question: *Can you use the result, or the method, for some other problem?* Now there is a better chance that the students may find some more interesting concrete interpretation, for instance, the following:

"In the center of the flat rectangular top of a building which is 21 yards long and 16 yards wide, a flagpole is to be erected, 8 yards high. To support the pole, we need four equal cables. The cables should start from the same point, 2 yards under the top of the pole, and end at the four corners of the top of the building. How long is each cable?"

The students may use the *method* of the problem they solved in detail introducing a right triangle in a vertical plane, and another one in a horizontal plane. Or they may use the *result,* imagining a rectangular parallelepiped of which the diagonal, x, is one of the four cables and the edges are

$$a = 10.5 \qquad b = 8 \qquad c = 6.$$

By straightforward application of the formula, $x = 14.5$.

For more examples, see CAN YOU USE THE RESULT?

15. Various approaches. Let us still retain, for a while, the problem we considered in the foregoing sections **8, 10, 12, 14.** The main work, the discovery of the plan, was described in section **10.** Let us observe that the teacher could have proceeded differently. Starting from the same point as in section **10,** he could have followed a somewhat different line, asking the following questions:

"Do you know any related problem?"

"Do you know an *analogous* problem?"

"You see, the proposed problem is a problem of solid geometry. Could you think of a simpler analogous problem of plane geometry?"

"You see, the proposed problem is about a figure in space, it is concerned with the diagonal of a rectangular parallelepiped. What might be an analogous problem about a figure in the plane? It should be concerned with —the diagonal—of—a rectangular—"

"Parallelogram."

The students, even if they are very slow and indifferent, and were not able to guess anything before, are obliged finally to contribute at least a minute part of the idea. Besides, if the students are so slow, the teacher should not take up the present problem about the parallelepiped without having discussed before, in order to prepare the students, the analogous problem about the parallelogram. Then, he can go on now as follows:

"Here is a problem related to yours and solved before. Can you use it?"

"Should you introduce some auxiliary element in order to make its use possible?"

Eventually, the teacher may succeed in suggesting to the students the desirable idea. It consists in conceiving the diagonal of the given parallelepiped as the diagonal of a suitable parallelogram which must be introduced into the figure (as intersection of the parallelepiped with a plane passing through two opposite edges). The idea is essentially the same as before (section **10**) but the approach is different. In section **10**, the contact with the available knowledge of the students was established through the unknown; a formerly solved problem was recollected because its unknown was the same as that of the proposed problem. In the present section analogy provides the contact with the idea of the solution.

16. The teacher's method of questioning shown in the foregoing sections **8, 10, 12, 14, 15** is essentially this: Begin with a general question or suggestion of our list, and, if necessary, come down gradually to more specific and concrete questions or suggestions till you reach one which elicits a response in the student's mind. If you have to help the student exploit his idea, start again, if possible, from a general question or suggestion contained in the list, and return again to some more special one if necessary; and so on.

Of course, our list is just a first list of this kind; it seems to be sufficient for the majority of simple cases, but there is no doubt that it could be perfected. It is important, however, that the suggestions from which we start should be simple, natural, and general, and that their list should be short.

The suggestions must be simple and natural because otherwise they cannot be *unobtrusive.*

The suggestions must be general, applicable not only to the present problem but to problems of all sorts, if they are to help develop the *ability* of the student and not just a special technique.

The list must be short in order that the questions may be often repeated, unartificially, and under varying circumstances; thus, there is a chance that they will be eventually assimilated by the student and will contribute to the development of a *mental habit.*

It is necessary to come down gradually to specific suggestions, in order that the student may have as great a *share of the work* as possible.

This method of questioning is not a rigid one; fortunately so, because, in these matters, any rigid, mechanical, pedantical procedure is necessarily bad. Our method admits a certain elasticity and variation, it admits various approaches (section **15**), it can be and should be so applied that questions asked by the teacher *could have occurred to the student himself.*

If a reader wishes to try the method here proposed in his class he should, of course, proceed with caution. He should study carefully the example introduced in section **8**, and the following examples in sections **18, 19, 20**. He should prepare carefully the examples which he intends to discuss, considering also various approaches. He should start with a few trials and find out gradually how he can manage the method, how the students take it, and how much time it takes.

17. Good questions and bad questions. If the method of questioning formulated in the foregoing section is well understood it helps to judge, by comparison, the quality of certain suggestions which may be offered with the intention of helping the students.

Let us go back to the situation as it presented itself at the beginning of section **10** when the question was asked: *Do you know a related problem?* Instead of this, with the best intention to help the students, the question may be offered: *Could you apply the theorem of Pythagoras?*

The intention may be the best, but the question is about the worst. We must realize in what situation it was offered; then we shall see that there is a long sequence of objections against that sort of "help."

(1) If the student is near to the solution, he may understand the suggestion implied by the question; but if he is not, he quite possibly will not see at all the point at which the question is driving. Thus the question fails to help where help is most needed.

(2) If the suggestion is understood, it gives the whole secret away, very little remains for the student to do.

(3) The suggestion is of too special a nature. Even if the student can make use of it in solving the present problem, nothing is learned for future problems. The question is not instructive.

(4) Even if he understands the suggestion, the student can scarcely understand how the teacher came to the idea of putting such a question. And how could he, the student, find such a question by himself? It appears as an unnatural surprise, as a rabbit pulled out of a hat; it is really not instructive.

None of these objections can be raised against the procedure described in section **10**, or against that in section **15**.

MORE EXAMPLES

18. A problem of construction. *Inscribe a square in a given triangle. Two vertices of the square should be on the base of the triangle, the two other vertices of the square on the two other sides of the triangle, one on each.*

"What is the unknown?"

"A square."

"What are the data?"

"A triangle is given, nothing else."

"What is the condition?"

"The four corners of the square should be on the perimeter of the triangle, two corners on the base, one corner on each of the other two sides."

"Is it possible to satisfy the condition?"

"I think so. I am not so sure."

"You do not seem to find the problem too easy. *If you cannot solve the proposed problem, try to solve first some related problem.* Could you satisfy a *part of the condition?*"

"What do you mean by a part of the condition?"

"You see, the condition is concerned with all the vertices of the square. How many vertices are there?"

"Four."

"A part of the condition would be concerned with less than four vertices. *Keep only a part of the condition, drop the other part.* What part of the condition is easy to satisfy?"

"It is easy to draw a square with two vertices on the perimeter of the triangle—or even one with three vertices on the perimeter!"

"Draw a figure!"

The student draws Fig. 2.

"You *kept only a part of the condition,* and you *dropped the other part. How far is the unknown now determined?*"

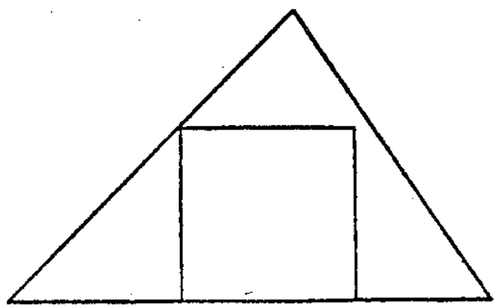

FIG. 2

"The square is not determined if it has only three vertices on the perimeter of the triangle."

"Good! *Draw a figure.*"

The student draws Fig. 3.

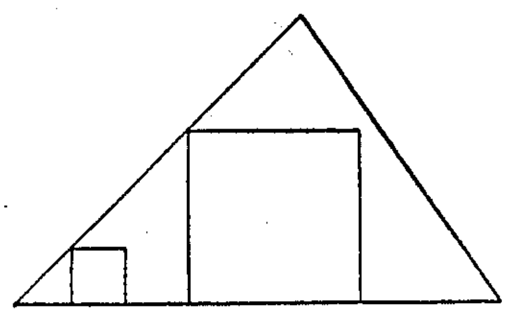

FIG. 3

"The square, as you said, is not determined by the *part of the condition you kept. How can it vary?*"

. . . . .

"Three corners of your square are on the perimeter of the triangle but the fourth corner is not yet there where it should be. Your square, as you said, is undetermined,

it can vary; the same is true of its fourth corner. *How can it vary?*"

. . . . .

"Try it experimentally, if you wish. Draw more squares with three corners on the perimeter in the same way as the two squares already in the figure. Draw small squares and large squares. What seems to be the locus of the fourth corner? *How can it vary?*"

The teacher brought the student very near to the idea of the solution. If the student is able to guess that the locus of the fourth corner is a straight line, he has got it.

19. A problem to prove. *Two angles are in different planes but each side of one is parallel to the corresponding side of the other, and has also the same direction. Prove that such angles are equal.*

What we have to prove is a fundamental theorem of solid geometry. The problem may be proposed to students who are familiar with plane geometry and acquainted with those few facts of solid geometry which prepare the present theorem in Euclid's Elements. (The theorem that we have stated and are going to prove is the proposition 10 of Book XI of Euclid.) Not only questions and suggestions quoted from our list are printed in italics but also others which correspond to them as "problems to prove" correspond to "problems to find." (The correspondence is worked out systematically in PROBLEMS TO FIND, PROBLEMS TO PROVE 5, 6.)

"What is the hypothesis?"

"Two angles are in different planes. Each side of one is parallel to the corresponding side of the other, and has also the same direction.

"What is the conclusion?"

"The angles are equal."

"Draw a figure. Introduce suitable notation."

The student draws the lines of Fig. 4 and chooses, helped more or less by the teacher, the letters as in Fig. 4.

"*What is the hypothesis?* Say it, please, using your notation."

"A, B, C are not in the same plane as A', B', C'. And $AB \parallel A'B'$, $AC \parallel A'C'$. Also AB has the same direction as $A'B'$, and AC the same as $A'C'$."

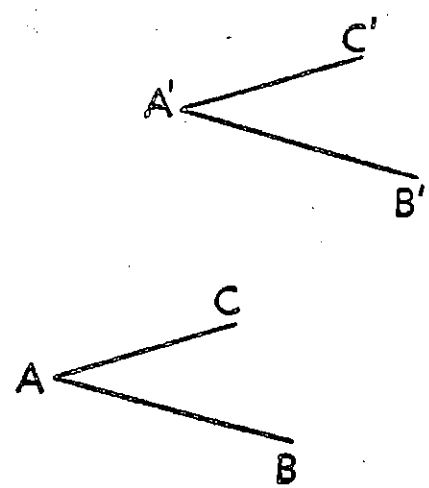

FIG. 4

"*What is the conclusion?*"

"$\angle BAC = \angle B'A'C'$."

"*Look at the conclusion! And try to think of a familiar theorem having the same or a similar conclusion.*"

"If two triangles are congruent, the corresponding angles are equal."

"Very good! Now *here is a theorem related to yours and proved before. Could you use it?*"

"I think so but I do not see yet quite how."

"*Should you introduce some auxiliary element in order to make its use possible?*"

. . . . .

"Well, the theorem which you quoted so well is about triangles, about a pair of congruent triangles. Have you any triangles in your figure?"

"No. But I could introduce some. Let me join B to C, and B' to C'. Then there are two triangles, $\triangle ABC$, $\triangle A'B'C'$."

"Well done. But what are these triangles good for?"

"To prove the conclusion, $\angle BAC = \angle B'A'C'$."

"Good! If you wish to prove this, what kind of triangles do you need?"

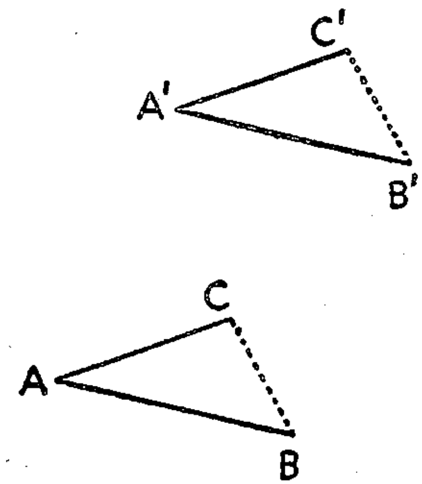

FIG. 5

"Congruent triangles. Yes, of course, I may choose B, C, B', C' so that

$$AB = A'B', AC = A'C'."$$

"Very good! Now, what do you wish to prove?"

"I wish to prove that the triangles are congruent,

$$\triangle ABC = \triangle A'B'C'.$$

If I could prove this, the conclusion $\angle BAC = \angle B'A'C'$ would follow immediately."

"Fine! You have a new aim, you aim at a new conclusion. *Look at the conclusion! And try to think of a* *familiar theorem having the same or a similar conclusion."*

"Two triangles are congruent if—if the three sides of the one are equal respectively to the three sides of the other."

"Well done. You could have chosen a worse one. Now *here is a theorem related to yours and proved before. Could you use it?"*

"I could use it if I knew that $BC = B'C'$."

"That is right! Thus, what is your aim?"

"To prove that $BC = B'C'$."

"Try to think of a familiar theorem having the same or a similar conclusion."

"Yes, I know a theorem finishing: '. . . then the two lines are equal.' But it does not fit in."

"Should you introduce some auxiliary element in order to make its use possible?"

. . . . .

"You see, how could you prove $BC = B'C'$ when there is no connection in the figure between BC and $B'C'$?"

. . . . .

"Did you use the hypothesis? What is the hypothesis?"

"We suppose that $AB \parallel A'B'$, $AC \parallel A'C'$. Yes, of course, I must use that."

"Did you use the whole hypothesis? You say that $AB \parallel A'B'$. Is that all that you know about these lines?"

"No; AB is also equal to $A'B'$, by construction. They are parallel and equal to each other. And so are AC and $A'C'$."

"Two parallel lines of equal length—it is an interesting configuration. *Have you seen it before?"*

"Of course! Yes! Parallelogram! Let me join A to A', B to B', and C to C'."

"The idea is not so bad. How many parallelograms have you now in your figure?"

"Two. No, three. No, two. I mean, there are two of

which you can prove immediately that they are parallelograms. There is a third which seems to be a parallelogram; I hope I can prove that it is one. And then the proof will be finished!"

We could have gathered from his foregoing answers that the student is intelligent. But after this last remark of his, there is no doubt.

This student is able to guess a mathematical result and to distinguish clearly between proof and guess. He knows also that guesses can be more or less plausible. Really, he did profit something from his mathematics classes; he has some real experience in solving problems, he can conceive and exploit a good idea.

20. A rate problem. *Water is flowing into a conical vessel at the rate* r. *The vessel has the shape of a right circular cone, with horizontal base, the vertex pointing downwards; the radius of the base is* a, *the altitude of the*

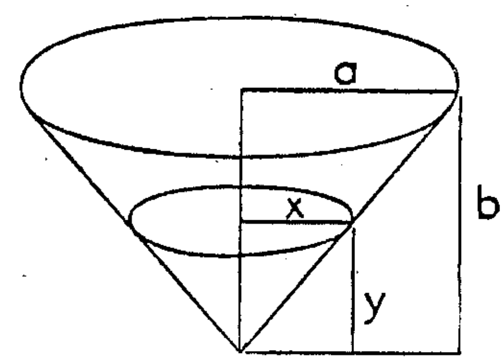

FIG. 6

cone b. *Find the rate at which the surface is rising when the depth of the water is* y. *Finally, obtain the numerical value of the unknown supposing that* $a = 4$ *ft.*, $b = 3$ *ft.*, $r = 2$ *cu. ft. per minute, and* $y = 1$ *ft.*

The students are supposed to know the simplest rules of differentiation and the notion of "rate of change."

"What are the data?"

"The radius of the base of the cone $a = 4$ ft., the altitude of the cone $b = 3$ ft., the rate at which the water is flowing into the vessel $r = 2$ cu. ft. per minute, and the depth of the water at a certain moment, $y = 1$ ft."

"Correct. The statement of the problem seems to suggest that you should disregard, provisionally, the numerical values, work with the letters, express the unknown in terms of a, b, r, y and only finally, after having obtained the expression of the unknown in letters, substitute the numerical values. I would follow this suggestion. Now, *what is the unknown?*"

"The rate at which the surface is rising when the depth of the water is y."

"What is that? Could you say it in other terms?"

"The rate at which the depth of the water is increasing."

"What is that? *Could you restate it still differently?*"

"The rate of change of the depth of the water."

"That is right, the rate of change of y. But what is the rate of change? *Go back to the definition.*"

"The derivative is the rate of change of a function."

"Correct. Now, is y a function? As we said before, we disregard the numerical value of y. Can you imagine that y changes?"

"Yes, y, the depth of the water, increases as the time goes by."

"Thus, y is a function of what?"

"Of the time t."

"Good. *Introduce suitable notation.* How would you write the 'rate of change of y' in mathematical symbols?"

"$\frac{dy}{dt}$"

"Good. Thus, this is your unknown. You have to express it in terms of a, b, r, y. By the way, one of these data is a 'rate.' Which one?"

"r is the rate at which water is flowing into the vessel."

"What is that? Could you say it in other terms?"

"r is the rate of change of the volume of the water in the vessel."

"What is that? *Could you restate it still differently? How would you write it in suitable notation?*"

"$r = \frac{dV}{dt}$."

"What is V?"

"The volume of the water in the vessel at the time t."

"Good. Thus, you have to express $\frac{dy}{dt}$ in terms of a, b, $\frac{dV}{dt}$, y. How will you do it?"

.....

"If you cannot solve the proposed problem try to solve first some related problem. If you do not see yet the connection between $\frac{dy}{dt}$ and the data, try to bring in some simpler connection that could serve as a stepping stone."

.....

"Do you not see that there are other connections? For instance, are y and V independent of each other?"

"No. When y increases, V must increase too."

"Thus, there is a connection. What is the connection?"

"Well, V is the volume of a cone of which the altitude is y. But I do not know yet the radius of the base."

"You may consider it, nevertheless. Call it something, say x."

"$V = \frac{\pi x^2 y}{3}$."

"Correct. Now, what about x? Is it independent of y?"

"No. When the depth of the water, y, increases the radius of the free surface, x, increases too."

"Thus, there is a connection. What is the connection?"

"Of course, similar triangles.

$$x : y = a : b."$$

"One more connection, you see. I would not miss profiting from it. Do not forget, you wished to know the connection between V and y."

"I have

$$x = \frac{ay}{b}$$

$$V = \frac{\pi a^2 y^3}{3b^2}."$$

"Very good. This looks like a stepping stone, does it not? But you should not forget your goal. *What is the unknown?*"

"Well, $\frac{dy}{dt}$."

"You have to find a connection between $\frac{dy}{dt}$, $\frac{dV}{dt}$, and other quantities. And here you have one between y, V, and other quantities. What to do?"

"Differentiate! Of course!

$$\frac{dV}{dt} = \frac{\pi a^2 y^2}{b^2} \frac{dy}{dt}.$$

Here it is."

"Fine! And what about the numerical values?"

"If $a = 4$, $b = 3$, $\frac{dV}{dt} = r = 2$, $y = 1$, then

$$2 = \frac{\pi \times 16 \times 1}{9} \frac{dy}{dt}."$$

PART II. HOW TO SOLVE IT
A DIALOGUE

Getting Acquainted

Where should I start? Start from the statement of the problem.

What can I do? Visualize the problem as a whole as clearly and as vividly as you can. Do not concern yourself with details for the moment.

What can I gain by doing so? You should understand the problem, familiarize yourself with it, impress its purpose on your mind. The attention bestowed on the problem may also stimulate your memory and prepare for the recollection of relevant points.

Working for Better Understanding

Where should I start? Start again from the statement of the problem. Start when this statement is so clear to you and so well impressed on your mind that you may lose sight of it for a while without fear of losing it altogether.

What can I do? Isolate the principal parts of your problem. The hypothesis and the conclusion are the principal parts of a "problem to prove"; the unknown, the data, and the conditions are the principal parts of a "problem to find." Go through the principal parts of your problem, consider them one by one, consider them in turn, consider them in various combinations, relating each detail to other details and each to the whole of the problem.