Module 2: Problem Solving Strategies and Estimation

2.1 Problem Solving

Solving problems is the most common task used to measure understanding in technical and scientific courses, and in many aspects of life as well. In general, problem solving requires factual and procedural knowledge in the area of the problem, plus knowledge of numerous schema, plus skill in overall problem solving. Schema is loosely defined as a "specific type of problem" such as principal, rate, and interest problems, one-dimensional kinematic problems with constant acceleration, etc. In most introductory university courses, improving problem solving relies on three things:

1. increasing domain knowledge, particularly definitions and procedures
2. learning schema for various types of problems and how to recognize that a particular problem belongs to a known schema
3. becoming more conscious of and insightful about the process of problem solving.

To improve your problem solving ability in a course, the most essential change of attitude is to focus more on the *process of solution* rather than on *obtaining the answer.* For homework problems there is frequently a simple way to obtain the answer, often involving some specific insight. This will quickly get you the answer, but you will not build schema that will help solve related problems further down the road. Moreover, if you rely on insight, when you get stuck on a problem, you're stuck with no plan or fallback position. At MIT you will see very few exam problems that are exactly the same as problems you have seen before, but most will use the same schema.

General Approach to Problem Solving

A great many physics textbook authors (e.g. Young and Freedman, Knight, Halliday, Resnick and Cartwright…) recommend overall problem solving strategies. These are typically four-step procedures that descend from George Polya's influential book, *How to Solve It*, on problem solving[1]. Here are his four steps:

I. Understand – get a conceptual grasp of the problem

What is the problem asking? What are the given conditions and assumptions? What domain of knowledge is involved? What is to be found and how is this determined or constrained by the given conditions?

What knowledge is relevant? E.g. in physics, does this problem involve kinematics, forces, energy, momentum, angular momentum, equilibrium? If the problem involves two different areas of knowledge, try to separate the problem into parts. Is there motion or is it static? If the problem involves vector quantities such as velocity or momentum, think of these geometrically

[1] G. Polya, *How to Solve It*, *2nd ed.,* Princeton University Press, 1957.

(as arrows that add vectorially). Get conceptual understanding: is some physical quantity (energy, momentum, angular momentum, etc.) conserved? Have you done problems that involve the same concepts in roughly the same way?

Model: Real life contains great complexity, so in physics (chemistry, economics…) you actually solve a model problem that contains the essential elements of the real problem. The bike and rider become a point mass (unless angular momentum is involved), the ladder's mass is regarded as being uniformly distributed along its length, the car is assumed to have constant acceleration or constant power (obviously not true when it shifts gears), etc. Become sensitive to information that is implicitly assumed (Presence of gravity? No friction? That the collision is of short duration relative to the timescale of the subsequent motion? …).

Advice: Write *your own* representation of the problem's stated data; *redraw* the picture with your labeling and comments. Get the problem into your brain! Go systematically down the list of topics in the course or for that week if you are stuck.

II. Devise a Plan - set up a procedure to obtain the desired solution

General - Have you seen a problem like this – i.e., does the problem fit in a schema you already know? Is a part of this problem a known schema? Could you simplify this problem so that it is? Can you find *any* useful results for the given problem and data even if it is not the solution (e.g. in the special case of motion on an incline when the plane is at $q = 0$)? Can you imagine a route to the solution if only you knew some apparently not given information? If your solution plan involves equations, count the unknowns and check that you have that many independent equations.

In Physics, exploit the freedoms you have: use a particular type of coordinate system (e.g. polar) to simplify the problem, pick the orientation of a coordinate system to get the unknowns in one equation only (e.g. only the x-direction), pick the position of the origin to eliminate torques from forces you don't know, pick a system so that an unknown force acts entirely within it and hence does not change the system's momentum… Given that the problem involves some particular thing (constant acceleration, momentum) think over *all* the equations that involve this concept.

III. Carry our your plan – solve the problem!

This generally involves mathematical manipulations. Try to keep them as simple as possible by not substituting in lengthy algebraic expressions until the end is in sight, make your work as neat as you can to ease checking and reduce careless mistakes. Keep a clear idea of where you are going and have been (label the equations and what you have now found), if possible, check each step as you proceed.

IV. Look Back – check your solution and method of solution

Can you see that the answer is correct now that you have it – often simply by retrospective inspection? Can you solve it a different way? Is the problem equivalent to one you've solved before if the variables have some specific values?

In physics: Check dimensions if analytic, units if numerical. Check special cases (for instance, for a problem involving two massive objects moving on an inclined plane, if $m_1 = m_2$ or $q = 0$ does the solution reduce to a simple expression that you can easily derive by inspection or a simple argument?) Is the scaling what you'd expect (an energy should vary as the velocity squared, or linearly with the height). Does it depend sensibly on the various quantities (e.g. is the acceleration less if the masses are larger, more if the spring has a larger k)? Is the answer physically reasonable (especially if numbers are given or reasonable ones substituted).

Review the schema of your solution: Review and try to remember the outline of the solution – what is the model, the physical approximations, the concepts needed, and any tricky math manipulation.

2.2 Significant Figures, Scientific Notation, and Rounding

Significant Figures

We shall define significant figures by the following rules.[2]

1. The leftmost nonzero digit is the most significant digit.

2. If there is no decimal place, the rightmost nonzero digit is the least significant digit.

3. If there is a decimal point then the right most digit is the least significant digit even if it is a zero.

4. All digits between the least and most significant digits are counted as significant digits.

[2] Philip R Bevington and D. Keith Robinson, Data Reduction and Error Analysis for the Physical Sciences, 2nd Edition, McGraw-Hill, Inc., New York, 1992.

When reporting the results of an experiment, the number of significant digits used in reporting the result is the number of digits needed to state the result of that measurement (or a calculation based on that measurement) without any loss of precision.

There are exceptions to these rules, so you may want to carry around one extra significant digit until you report your result. For example if you multiply $2 \times 0.56 = 1.12$, not 1.1.

There is some ambiguity about the number of significant figure when the rightmost digit is 0, for example 1050, with no terminal decimal point. This has only three significant digits. If all the digits are significant the number should be written as 1050., with a terminal decimal point. To avoid this ambiguity it is wiser to use scientific notation.

Scientific Notation

Careless use of significant digits can be easily avoided by the use of decimal notation times the appropriate power of ten for the number. Then all the significant digits are manifestly evident in the decimal number. So the number $1050 = 1.05 \times 10^3$ while the number $1050. = 1.050 \times 10^3$.

Rounding

To round off a number by eliminating insignificant digits we have three rules. For practical purposes, rounding will be done automatically by a calculator or computer, and all we need do is set the desired number of significant figures for whichever tool is used.

1. If the fraction is greater than 1/2, increment the new least significant digit.

2. If the fraction is less than 1/2, do not increment.

3. If the fraction equals 1/2, increment the least significant digit only if it is odd.

The reason for Rule 3 is that a fractional value of 1/2 may result from a previous rounding up of a fraction that was slightly less than 1/2 or a rounding down of a fraction that was slightly greater than 1/2. For example, 1.249 and 1.251 both round to three significant digits 1.25. If we were to round again to two significant digits, both would yield the same value, either 1.2 or 1.3 depending on our convention in Rule 3. Choosing to round up if the resulting last digit is odd and to round down if the resulting last digit is even reduces the systematic errors that would otherwise be introduced into the average of a group of such numbers.

2.3 Order of Magnitude Estimates - Fermi Problems

Counting is the first mathematical skill we learn. We came to use this skill by distinguishing elements into groups of similar objects, but we run into problems when our desired objects are not easily identified, or there are too many to count.

Rather than spending a huge amount of effort to attempt an exact count, we can try to estimate the number of objects in a collection. For example, we can try to estimate the total number of grains of sand contained in a bucket of sand. Since we can see individual grains of sand, we expect the number to be very large but finite. Sometimes we can try to estimate a number which we are fairly sure but not certain is finite, such as the number of particles in the universe (See Chapter 20).

We can also assign numbers to quantities that carry dimensions, such as mass, length, time, or charge, which may be difficult to measure exactly. We may be interested in estimating the mass of the air inside a room, or the length of telephone wire in the United States, or the amount of time that we have slept in our lives, or the number of electrons inside our body. So we choose some set of units, such as kilograms, miles, hours, and coulombs, and then we can attempt to estimate the number with respect to our standard quantity.

Often we are interested in estimating quantities such as speed, force, energy, or power. We may want to estimate our natural walking speed, or the force of wind acting against a bicycle rider, or the total energy consumption of a country, or the electrical power necessary to operate this institute. All of these quantities have no exact, well-defined value; they instead lie within some range of values.

When we make these types of estimates, we should be satisfied if our estimate is reasonably close to the middle of the range of possible values. But what does "reasonably close" mean? Once again, this depends on what quantities we are estimating. If we are describing a quantity that has a very large number associated with it, then an estimate within an order of magnitude should be satisfactory. The number of molecules in a breath of air is close to 10^{22}; an estimate anywhere between 10^{21} and 10^{23} molecules is close enough. If we are trying to win a contest by estimating the number of marbles in a glass container, we cannot be so imprecise; we must hope that our estimate is within 1% of the real quantity.

These types of estimations are called *Fermi Problems*. The technique is named after the physicist Enrico Fermi, who was famous for making these sorts of "back of the envelope" calculations.

Methodology for Estimation Problems

Estimating is a skill that improves with practice. Here are two guiding principles that may help you get started.

(1) You must identify a set of quantities that can be estimated or calculated.

(2) You must establish an approximate or exact relationship between these quantities and the quantity to be estimated in the problem.

Estimations may be characterized by a precise relationship between an estimated quantity and the quantity of interest in the problem. When we estimate, we are drawing upon what we know. But different people are more familiar with certain things than others. If you are basing your estimate on a fact that you already know, the accuracy of your estimate will depend on the accuracy of your previous knowledge. When there is no precise relationship between estimated quantities and the quantity to be estimated in the problem, then the accuracy of the result will depend on the type of relationships you decide upon. There are often many approaches to an estimation problem leading to a reasonably accurate estimate. So use your creativity and imagination!

Example: Lining Up Pennies

Suppose you want to line pennies up, diameter to diameter, until the total length is 1 kilometer . How many pennies will you need? How accurate is this estimation?

Solution: The first step is to consider what type of quantity is being estimated. In this example we are estimating a dimensionless scalar quantity, the number of pennies. We can now give a precise relationship for the number of pennies needed to mark off 1 kilometer

$$\#\ \text{of pennies} = \frac{\text{total distance}}{\text{diameter of penny}}\,. \tag{2.3.1}$$

We can estimate a penny to be approximately 2 centimeters wide. Therefore the number of pennies is

$$\#\ \text{of pennies} = \frac{\text{total distance}}{\text{length of a penny}} = \frac{(1\ \text{km})}{(2\ \text{cm})(1\ \text{km}/10^5\ \text{cm})} = 5\times 10^4\ \text{pennies}\,. \tag{2.3.2}$$

When applying numbers to relationships we must be careful to convert units whenever necessary.

How accurate is this estimation? If you measure the size of a penny, you will find out that the width is $1.9\ \text{cm}$, so our estimate was accurate to within 5%. This accuracy was fortuitous. Suppose we estimated the length of a penny to be 1 cm. Then our estimate for the total number of pennies would be within a factor of 2, a margin of error we can live with for this type of problem.

Example: Estimate the total mass of all the water in the earth's oceans.

Solution: In this example we are estimating mass, a quantity that is a fundamental in SI units, and is measured in kg. Initially we will try to estimate two quantities: the density of water and the volume of water contained in the oceans. Then the relationship we want is

$$(\text{mass})_{\text{ocean}}=(\text{density})_{\text{water}}(\text{volume})_{\text{ocean}}\,. \tag{2.3.3}$$

One of the hardest aspects of estimation problems is to decide which relationship applies. One way to check your work is to check dimensions. Density has dimensions of mass/volume, so our relationship is

$$\left(\text{mass}_{\text{ocean}}\right)=\left(\frac{\text{mass}}{\text{volume}}\right)\left(\text{volume}_{\text{ocean}}\right). \tag{2.3.4}$$

The density of fresh water is $\rho_{\text{water}} = 1.0\ \text{g}\cdot\text{cm}^{-3}$; the density of seawater is slightly higher, but the difference won't matter for this estimate. You could estimate this density by envisioning how much mass is contained in a one-liter bottle of water. (The density of water is a point of reference for all density problems. Suppose we need to estimate the density of iron. If we compare iron to water, we estimate that iron is 5 to 10 times denser than water. The actual density of iron is $\rho_{\text{iron}} = 7.8\ \text{g}\cdot\text{cm}^{-3}$).

Since there is no precise relationship, estimating the volume of water in the oceans is much harder. Let's model the volume occupied by the oceans as if they completely cover the earth, forming a spherical shell (Figure 1.5, which is decidedly not to scale). The volume of a spherical shell of radius R_{earth} and thickness d is

$$\text{volume}_{\text{shell}} \cong \left(4\ \ R_{\text{earth}}^2 d\right), \tag{2.3.5}$$

where R_{earth} is the radius of the earth and d is the average depth of the ocean.

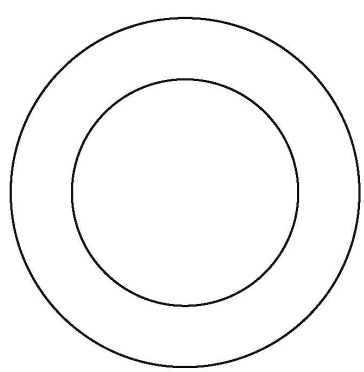

Figure 1.5 A model for estimating the mass of the oceans.

We first estimate that the oceans cover about 75% of the surface of the earth. So the volume of the oceans is

$$\text{volume}_{\text{ocean}} \cong (0.75)(4\pi R_{\text{earth}}^2 d). \tag{2.3.6}$$

We therefore have two more quantities to estimate, the average depth of the ocean, which we can estimate the order of magnitude as $d \cong 1\,\text{km}$, and the radius of the earth, which is approximately $R_{\text{earth}} \cong 6 \times 10^3\,\text{km}$. (The quantity that you may remember is the circumference of the earth, about $25{,}000\,\text{miles}$. Historically the circumference of the earth was defined to be $4 \times 10^7\,\text{m}$). The radius R_{earth} and the circumference s are exactly related by

$$s = 2\pi R_{\text{earth}}\,. \tag{2.3.7}$$

Thus

$$R_{\text{earth}} = \frac{s}{2\pi} = \frac{(2.5 \times 10^4\ \text{mi})(1.6\ \text{km} \cdot \text{mi}^{-1})}{2\pi} = 6.4 \times 10^3\ \text{km} \tag{2.3.8}$$

We will use $R_{\text{earth}} \cong 6 \times 10^3\,\text{km}$; additional accuracy is not necessary for this problem, since the ocean depth estimate is clearly less accurate. In fact, the factor of 75% is not needed, but included more or less from habit.

Altogether, our estimate for the mass of the oceans is

$$(\text{mass})_{\text{ocean}} = (\text{density})_{\text{water}} (\text{volume})_{\text{ocean}} \cong \rho_{\text{water}} (0.75)(4\pi R_{\text{earth}}^2 d), \tag{2.3.9}$$

$$(\text{mass})_{\text{ocean}} \cong \left(\frac{1\text{g}}{\text{cm}^3}\right)\left(\frac{1\,\text{kg}}{10^3\ \text{g}}\right)\left(\frac{(10^5\ \text{cm})^3}{(1\,\text{km})^3}\right)(0.75)(4\ \ )(6 \times 10^3\ \text{km})^2(1\text{km}), \tag{2.3.10}$$

$$(\text{mass})_{\text{ocean}} \cong 3 \times 10^{20}\ \text{kg} \cong 10^{20}\ \text{kg}$$