Quiz 1 Solutions

Professor Greenside

The following solutions are more detailed than what was required to get full credit on the problems.

1. (a) (3 points) For one point each, describe three ways that the Earth is not in thermodynamic equilibrium in addition to the fact that the temperature inside the Earth is not uniform.

Answer: To get credit for this problem, it was not sufficient to list a criterion, you had to be specific about how the criterion was or was not satisfied in the context of Earth. Some possible answers:

- i. The Earth is not time independent: the air, oceans, and molten rock and iron (deep in the Earth) all show macroscopic motion, and there are chemical processes such as weathering of rock that are not in steady state.
- ii. There is relative motion: the air in the atmosphere does not move rigidly with the air near the ground moving more slowly than the air higher in the sky, the water in the oceans does not move rigidly, the molten iron in the iron core is convecting and so does not move rigidly.
- iii. The state of the Earth is not independent of its history. This can be seen in many ways: there are unmixed minerals and rocks, there are details of the Earth such as the shapes and positions of the continents that are specific to Earth's history, the composition of the air continues to change over time (partly because of human technology).
- (b) (5 points) Estimate to the nearest power of ten the ratio of "the thermal relaxation time of the Earth" to "the age of the solar system" which is about 5 billion years. The radius of the Earth is 6,400 km and the thermal diffusivity of rock and iron is about $10^{-6} \text{ m}^2/\text{s}$.

Hint: Round all numbers to the nearest power of ten as soon as possible to minimize the arithmetic.

Answer: I accepted answers of 100 or 1,000, or 10,000, which are all reasonable based on different ways of rounding. This answer is surprising: it will take a cosmological amount of time, longer than the current age of the universe, for the Earth to reach thermal equilibrium. Of course, there may be other time scales such as a relaxation time to equilibrate with respect to concentrations and chemical reactions so reaching thermodynamic equilibrium might take even longer than reaching thermal equilibrium.

Let's answer this problem two ways, first by rounding to the nearest power of ten so that we avoid working with any numerical digits, second by retaining just one significant digit throughout.

First way: estimate by replacing all numbers by the nearest power of ten. The relaxation time is L^2/κ where L is the largest distance that separates one part of a system from another part except if the system is coming into equilibrium with a constant temperature surface, in which case L is the biggest distance from an interior point to the surface. So for this problem, we should choose L = R where R is the radius of the sphere since Earth is coming into equilibrium with a constant-temperature surface that is basically zero degrees K (the temperature of outer space). It is also ok to choose L = 2R for this problem since you get the same order of magnitude.

We are given R = 6,400 km and $\kappa = 10^{-6} \text{ m}^2/\text{s}$ so we need to convert km to m or vice versa to have consistent length units. We are also given the age of the solar system in years which means, when we take a ratio of L^2/κ to the age of the solar system, we have to convert years to seconds or vice versa, to have consistent units of time.

In general, you should try to combine all numbers together in one calculation rather than calculate the answer in several stages since the former reduces possible errors and helps one to identify cancellations that can reduce the amount of work. Thus the calculation would start like this:

$$\frac{L^2/\kappa}{\text{age of solar system}} = \frac{\left[\left(6,400\,\text{km} \times 10^3\,\frac{\text{m}}{\text{km}} \right)^2 / 10^{-6}\,\text{m}^2/\text{s} \right]}{5 \times 10^9\,\text{years} \times 365\,\frac{\text{days}}{\text{year}} \times 24\,\frac{\text{hours}}{\text{day}} \times 3,600\,\frac{\text{seconds}}{\text{hour}}.$$
(1)

I strongly recommend that, in the future, you begin all calculations in the same way: gather all numbers together with all their known digits, and explicitly give units.

At this point, we could calculate the answer using a calculator or replace each number with a leading significant digit, or replace numbers with the nearest power of ten. Doing the latter, we make the substitutions:

$$6,400 \to 10^4, \qquad 5 \to 10, \qquad 365 \to 10^2, \qquad 24 \to 10^1, \qquad 3600 \to 10^3.$$
 (2)

The reason why I replace the 5 in the denominator with a 10 and not with a 1 is because there are several other numbers in the denominator that I am replacing with smaller power of ten, and I want to balance the errors to improve the overall accuracy of the estimate. Eq. (1) now becomes

$$\frac{L^2/\kappa}{\text{age of solar system}} \approx \frac{\left\lfloor \left(10^4 \times 10^3\right)^2 / 10^{-6} \right\rfloor}{\left(10 \times 10^9\right) \times 10^2 \times 10^1 \times 10^3}$$
(3)

$$=\frac{\left(10^7\right)^2 \times 10^6}{10^{1+9+2+1+3}}\tag{4}$$

$$=\frac{10^{20}}{10^{16}}=10^4,\tag{5}$$

so the relaxation time of the Earth compared to the age of the solar system is about 10,000. Let's redo the calculation by trying to get a better estimate in which retain one significant digit at all the various steps. We again start with Eq. (1) where all numbers with all significant digits are listed. But now we round each number to the leading digit so the substitutions Eq. (2) become

$$6,400 \to 6 \times 10^3, \qquad 365 \to 4 \times 10^2, \qquad 24 \to 2 \times 10^1, \qquad 3,600 \to 4 \times 10^3.$$
 (6)

Making these substitutions, Eq. (1) becomes:

$$\frac{L^2/\kappa}{\text{age of solar system}} \approx \frac{\left[\left(\left(6 \times 10^3\right) \times 10^3\right)^2 / 10^{-6}\right]}{\left(5 \times 10^9\right) \times \left(4 \times 10^2\right) \times \left(2 \times 10^1\right) \times \left(4 \times 10^3\right)}$$
(7)

$$= \frac{36}{5 \times 4 \times 2 \times 4} \times \frac{\left[(10^3 \times 10^3)^2 \times 10^6 \right]}{10^{9+2+1+3}} \tag{8}$$

$$=\frac{36}{5\times4\times2\times4}\times10^{3}$$
(9)

$$\approx 300,$$
 (10)

to one significant digit. Notice how I separated all the leading digits from the powers of ten so I can concentrate on simplifying the combinations of digits.

The combination of digits $36/(5 \times 4 \times 2 \times 4)$ can be simplified in several ways. One is to divide out factors as much as possible, e.g., replace 36/4 with 9 to get $9/(5 \times 4 \times 2)$. Next would be to observe that 9 is nearly 10 and so replace $9/(5 \times 2)$ with 1, leaving $1/4 = 0.25 \approx 3 \times 10^{-1}$ so the answer to one digit would be 300, quite a bit smaller than the first estimate of 10^4 . Alternatively, we could replace 36 with 40 right away (trying always to retain just one significant digit), and then observe that $40/(5 \times 4 \times 2) = 1$ in which case the combination of digits reduces to $1/4 \approx 0.25 = 3 \times 10^{-1}$ and we again get 300. A calculator applied to Eq. (1) gives 260 to two significant digits, so our estimate to one significant digit is a good one.

Please appreciate that a precise value is not needed here since the relaxation time L^2/κ is a rough guideline for how long it takes to reach equilibrium; the estimate by replacing all numbers with nearest powers of ten provided a quick useful way to get the ballpark of what to expect, and we are off by a factor of 30 compared to the answer with one significant digit. The key scientific insight to obtain from this problem is that, no matter how we make the estimate, the relaxation time is really long, billions of years.

A few more comments to keep in mind for future quizzes, exams, and when you do calculations of your own as scientists or engineers:

- i. There is one exception to when it is ok to replace a number with the nearest power of ten right away and that is when a number is being raised to a power. Thus replacing $(4 \times 10^{-4})^3$ with $(10^{-4})^3 = 10^{-12}$ is not wise since $4^3 = 64 \approx 10^2$ and we have lost a factor of 100. So just for those numbers being raised to powers, you should retain one significant digit and raise that digit by itself to the power and then replace by the nearest power of ten.
- ii. Too many students made errors with units and forgot to convert kilometers to meters or years to seconds and ended up with wildly wrong answers. To avoid these kinds of errors, *please list the units with your numbers as you start your calculation* as I did in Eq. (1). Once you know that all units are consistent, it is ok to drop units as you proceed with your calculation.
- iii. You should *always* use scientific notation. Many students wrote out numbers with many zeros, e.g., 6, 400, 000 m for the radius of the Earth. It is hard to count many zeros quickly and so working with numbers like these increases the chance that you will lose or gain an extra zero, which can be a big error. It also takes too much time to write out all these zeros, and especially to multiply or divide numbers with many zeros.
- 2. (5 points) A molecule undergoes a random walk in which it successively travels a distance d, collides with another molecule, and then moves off in a random new direction. Show that the ensemble average of the final position $\mathbf{X}(N)$ of the molecule after N successive random steps $\Delta \mathbf{X}_i$

$$\mathbf{X}(N) = \sum_{i=1}^{N} \Delta \mathbf{X}_{i},\tag{11}$$

is the zero vector $\mathbf{0}$, i.e., the origin of the coordinate system where the molecule started off.

Answer: The problem requires thinking about the quantity $\langle \mathbf{X}(N) \rangle$. Using the additivity of averages, that $\langle f + g \rangle = \langle f \rangle + \langle g \rangle$, Eq. (11) gives

$$\langle \mathbf{X}(N) \rangle = \sum_{i=1}^{N} \langle \Delta \mathbf{X}_i \rangle.$$
(12)

The result will be true if the ensemble average of each separate vector step $\langle \Delta \mathbf{X}_i \rangle$ is the zero vector. But this actually is the case. The ensemble average of the *i*th step can be written as

$$\langle \Delta \mathbf{X}_i \rangle = \frac{1}{K} \sum_{k=1}^K \Delta \mathbf{X}_i^{(k)},\tag{13}$$

where we imagine K identical experiments being performed, all starting off in an identical way, and we record the value of the *i*th step $\Delta \mathbf{X}_i^{(k)}$ for each of the K experiments and then average the resulting set of K vectors. But these K vectors all have the same length d (by assumption of the random walk) but point in randomly different directions in three dimensions. Provided K is large enough, there will always be one vector that points nearly opposite to another vector and so the sum will be zero since each vector in the sum is canceled by another vector in the sum.

Although there is no need to do so, this argument can be made a bit more precise by observing that each step vector $\Delta \mathbf{X}_i^{(k)} = d \, \hat{\mathbf{n}}_i^{(k)}$ can be written in terms of a unit vector $\hat{\mathbf{n}}_i^{(k)}$ whose tip lies on the unit sphere. The ensemble average of a step vector is then equivalent to summing all the vectors $\hat{\mathbf{u}}_i^{(k)}$ over the surface of a unit sphere, at least for large enough K (and the ensemble average can be thought of as the value obtained in the limit $K \to \infty$), and it is pretty obvious this sum has to be zero by symmetry, for any unit vector pointing in one direction, there is a unit vector pointing in an opposite direction and their sum will be the zero vector.

It is a useful exercise with spherical coordinates to verify this obvious conclusion. In spherical coordinates, a unit vector pointing in the direction (θ, ϕ) from the origin can be written as

$$\hat{\mathbf{n}}(\theta,\phi) = \left(\sin(\theta)\cos(\phi),\sin(\theta)\sin(\phi),\cos(\theta)\right),\tag{14}$$

so the ensemble average Eq. (13) becomes an integration over the the surface of the unit sphere:

$$\langle \Delta \mathbf{X}_i \rangle = \frac{1}{K} \sum_{k=1}^K \Delta \mathbf{X}_i^{(k)} \tag{15}$$

$$= d \times \frac{1}{K} \sum_{k=1}^{K} \hat{\mathbf{n}}_i(\theta_k, \phi_k)$$
(16)

$$= d \times \int_0^{2\pi} \int_0^{\pi} \left(\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta)\right) \times \frac{\sin(\theta)\,d\theta\,d\phi}{4\pi},\tag{17}$$

where the first integral is with respect to ϕ and the second integral is with respect to θ . Notice a subtle point, that because we are using spherical coordinates, we need to include the factor $d\Omega/(4\pi) = \sin(\theta) d\theta d\phi/(4\pi)$ to weight each vector $\hat{\mathbf{n}}(\theta, \phi)$ by the fraction of the unit sphere pointing in the direction (θ, ϕ) . We can see that this integral must give zero without doing any integrations since each component has odd symmetry with respect to θ or ϕ and so gives zero when integrated over a domain symmetric with respect to ϕ or θ .

Note: some students tried mightily to prove the above result by relating it to the discussion in class, in which we took the ensemble average of the square of the vector $\mathbf{X}(N)$ and then argued that various dot products would average to zero. This argument does not work here and that was actually the point of this question: to see if you could understand conceptually how an ensemble average applies to a step vector $\Delta \mathbf{X}_i$ directly.

3. (5 points) Given that

$$e^{\pm x} \approx 1 \pm x + \frac{1}{2}x^2 \pm \frac{1}{6}x^3, \quad \text{for } |x| \ll 1,$$
 (18)

find the cubic polynomial that most accurately approximates the hyperbolic tangent function

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

near x = 0. (Hint: At some point, use $1/(1-x) \approx 1 + x + x^2 + \dots$ for $|x| \ll 1$.)

Answer: The goal of this problem was to determine a Taylor series approximation up to cubic terms for the hyperbolic tanh function (pronounced in English as "tanch"), i.e., to find a third-order polynomial $c_0+c_1x+c_2x^2+c_3x^3$ that best approximates this function in the vicinity of the location x = 0. The wrong way to do this would have been to start differentiating tanh(x) up to third order, the algebra is excessive and unnecessary. Instead, as you learned in the first homework assignment, the idea was for you to use the given Taylor series of the exponential function to generate a ratio of polynomials, which you could then simplify using the identity $1/(1-x) \approx 1 + x + x^2$ to convert a reciprocal to a product.

Before grinding through some algebra, note that you can use parity symmetry of the function, whether it is even or odd, to get a quick idea of what terms will appear in the Taylor series. Thus you should be able to see that the tanh(x) function is an odd function so that f(-x) = -f(x). This implies that only odd powers can appear in the Taylor series about x = 0 and so the polynomial approximation to cubic order has only two terms of the form $c_1x + c_3x^3$; whatever your calculation, the constant term and x^2 terms better be zero. Similarly, a function like tan(cos(x)) is an even function, f(-x) = f(x)and so all odd powers of x must be missing in the Taylor series. If a function has neither even or odd parity such as sin(x) + cos(x) or e^x , then you can generally expect all powers of x to appear.

The steps go like this. First, the numerator of the tanh function can be written as:

$$e^{x} - e^{-x} = \left(1 + x + \frac{1}{2}x^{2} + \frac{1}{6}x^{3}\right) - \left(1 - x + \frac{1}{2}x^{2} - \frac{1}{6}x^{3}\right) = 2\left(x + \frac{1}{6}x^{3}\right)$$
(19)

while the denominator can be written as

$$e^{x} + e^{-x} = \left(1 + x + \frac{1}{2}x^{2} + \frac{1}{6}x^{3}\right) + \left(1 - x + \frac{1}{2}x^{2} - \frac{1}{6}x^{3}\right) = 2\left(1 + \frac{1}{2}x^{2}\right)$$
(20)

up to cubic terms. Substituting into $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$, we get

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$
(21)

$$\approx \frac{x + \frac{1}{6}x^3}{1 + \frac{1}{2}x^2} \tag{22}$$

$$\approx \left(x + \frac{1}{6}x^3\right) \times \left(1 - \frac{1}{2}x^2\right) \tag{23}$$

$$= x - \frac{1}{2}x^3 + \frac{1}{6}x^3 + \text{h.o.t.s}$$
(24)

$$= x - \frac{1}{3}x^3.$$
 (25)

This last line is the answer. Here is a line-by-line justification of the above algebra:

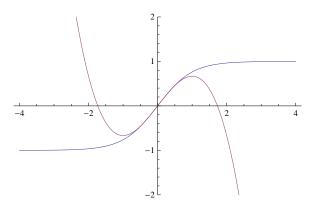
- (a) Eq. (22) was obtained by substituting the Taylor series approximations up to third-order terms from Eq. (19) and Eq. (20) into Eq. (21).
- (b) Eq. (23) was obtained by using the two lowest-order terms of the geometric series $1/(1-x) \approx 1+x$ with x replaced by $-(1/2)x^2$. This trick replaces division by the polynomial $1 + (1/2)x^2$ with multiplication by the polynomial $1 (1/2)x^2$:

$$\frac{1}{1+\frac{1}{2}x^2} \approx 1 - \frac{1}{2}x^2.$$
(26)

I ignored all higher-order terms in the geometric series because the next term in the geometric series involves squaring $(1/2)x^2$ which gives a 4th-order power of x, and this can be thrown out since we are only retaining terms to 3rd-order.

(c) Eq. (24) was obtained by multiplying out the two polynomials in Eq. (23), and by immediately setting to zero any power of x higher than 3. The abbreviation **h.o.t.s** means "higher order terms" and is a common way to refer to infinitely many terms that can be ignored because they are sufficiently small in the limit $x \to 0$.

The following Mathematica plot shows that the Taylor series Eq. (25) in purple accurately approximates the tanh function in blue (relative error smaller than 10%) over the range [-1,1] and then deviates substantially, if anything because tanh(x) is bounded between -1 and 1 while the cubic polynomial is unbounded.



We will use the answer Eq. (25) later in the course, it shows up when calculating properties of a paramagnet and some polymers.

Note: some students didn't remember or notice that, when using the approximation $1/(1-x) \approx 1+x+x^2$, that |x| must be small compared to one, else you can't ignore the higher powers like x^3 , x^4 , and so on. Thus the following step:

$$\frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}} = \frac{1 - e^{-2x}}{1 - [-e^{-2x}]} \approx \left(1 - e^{-2x}\right) \times \left(1 + \left[-e^{-2x}\right] + \left[-e^{-2x}\right]^2\right),\tag{27}$$

is not valid because, for x near zero, $e^{-2x} \approx 1$ is not small compared to one. Similarly, you can not write

$$\frac{1}{2+x} = \frac{1}{1-(-1-x)},\tag{28}$$

and then use the geometric series with $x \to -(1+x)$ because, for x near zero, -1-x is close to -1 and so is not small compared to one. The trick is instead to pull out a factor of 2 in the denominator to get a "one plus something small compared to one":

$$\frac{1}{2+x} = \frac{1}{2} \times \frac{1}{1+\frac{1}{2}x},\tag{29}$$

and then use the geometric series with $x \to -x/2$ since x/2 is small compared to 1 if x is small compared to 1.

True or False Questions (2 points each)

For each of the following statements, please circle \mathbf{T} or \mathbf{F} to indicate respectively whether the statement is true or false. No justification is needed.

1. **T** / **F** Water vapor with a temperature of 200° C is twice as hot as boiling water with a temperature of 100° C.

Answer: F This statement would be true if the temperature scale was kelvin, not Celsius, one has to use an absolute temperature scale to compare "hotness". The ratio of hotness of the vapor to boiling water is $(273 + 200)/(273 + 100) = 473/373 \approx 1.3$, which is not close to being twice as hot.

2. **T** / **F** The time scale for a system to approach thermal equilibrium does not depend on how big the temperature differences are within the system.

Answer: F This problem was not well posed so I discounted it from the grade. The answer is "true" in that the relaxation time L^2/κ does not depend on details of the temperature difference, which is the answer I had in mind when creating this question. But as one student correctly pointed out to me, it is possible for a temperature difference in a system to be so great that the system changes its character, e.g., a solid could melt, a liquid could vaporize, or the substance could transform chemically into a new substance. In these cases, the mechanism of heat transport (κ) can change substantially, in which case the answer would be "false".

3. **T** / **F** In order for a macroscopic system to be in thermodynamic equilibrium, the number of atoms in that system can not change over time.

Answer: F The surface of some material that is in contact with a surrounding gas would be an example. The surface and surrounding gas can reach thermodynamic equilibrium but there is no reason why the number of atoms on the surface must be constant in time. Also, equilibrium is defined by macroscopic observables such as pressure, temperature, and chemical potential and these macroscopic observables allow microscopic variations in the number of atoms.

4. **T** / **F** An aluminum cylindrical rod has a length L that is much greater than its radius r and one end of the rod has a higher temperature than the other end. If the rod is sitting in outer space (i.e., in a vacuum), then the thermal relaxation time for the rod is r^2/κ .

Answer: T This problem was also not well posed and discounted from the grade.

What I originally had in mind was a rod that was thermally isolated from the world, e.g., by surrounding the rod with a thick layer of a material like Styrofoam that conducts heat slowly (a small thermal diffusivity κ). Then heat can not enter or leave the rod (at least over a time scale that can be made quite long) and the only way for the rod to reach thermal equilibrium would be along its length. Then the relaxation time would be determined by the longest distance over which the information has to travel for equilibration to occur giving $\tau = L^2/\kappa$.

But a vacuum is not the same as an ideal insulator (I did not think carefully enough here), the rod can lose energy by radiating it way in the form of blackbody radiation. The answer then depends on the relative rate at which energy is lost by radiation versus the relaxation time L^2/κ , and there is no way for the class to answer this question at this point in the course. (They will be able to answer it later in the semester.) So the answer could go either way depending on what assumptions one makes about the mean temperature of the rod (which controls the rate of blackbody radiation) and the thermal diffusivity of the rod.

5. **T** / **F** An aluminum cylindrical rod has a length L that is much greater than its radius r and one end of the rod has a higher temperature than the other end. If the rod is sitting in air with a constant temperature T, then the thermal relaxation time for the rod is r^2/κ .

Answer: T Here there is no ambiguity: the rod will eventually reach thermodynamic equilibrium with the surrounding air so the question is: what is the biggest distance over which information has to travel from the air (surface of the rod) to any point in the rod. But no point in the rod is further away from the surface than the radius r so the relaxation time is r^2/κ .

The point of this and the previous question is that identifying the relaxation time requires stating some context. For a rod surrounded by an ideal insulator, the relaxation time is L^2/κ , for the exact same rod surrounded by a constant temperature medium, the relaxation time is much faster, r^2/κ .

Note that to maintain a constant temperature environment of the rod, it is not sufficient to place the rod in air since the air near the rod will warm up or cool down in response to the temperature of the rod and then you have a complicated situation of a rod that is out of thermal equilibrium and air near the rod that is also out of thermal equilibrium. To make this problem well posed scientifically, you would have to blow air of constant temperature past the rod at a rate fast enough that the air near the surface of the rod always has temperature T.

6. **T** / **F** For a gas of volume V containing N identical molecules, the quantity $(V/N)^{1/3}$ is approximately equal to the mean free path of the molecules.

Answer: F By definition, the mean free path is the average distance a molecule travels before undergoing a collision. I mentioned in lecture and in my lecture notes that, that for air at room temperature, the mean free path is quite a bit bigger than the average spacing between molecules, 300 nm versus about 5 nm.

The difference between mean free path and average distance between molecules is important for our kinetic theory. When asking "how many molecules with average speed v collide with the wall within a short time Δt ?", one has to choose the distance $v\Delta t$ to be comparable to or smaller than a mean free path, otherwise molecules in the little virtual cylinder will collide with one another before hitting the wall, typically knocking a molecule out of the cylinder.

7. **T** / **F** If $a \propto b$ and $c \propto d$ then $a + c \propto b + d$.

Answer: F This problem tested whether you understood the concept of proportionality. $a \propto b$ means that there is a nonzero constant k_1 such that $a = k_1 b$, and $c \propto d$ means there is a constant k_2 such that $c = k_2 d$. Then $a + c = k_1 b + k_2 d$ and generally $k_1 \neq k_2$ so $k_1 b + k_2 d \neq k_3 (b + d)$ for some constant k_3 . So the statement is false.

Proportionality is mainly useful when multiplying quantities or raising quantities to powers. So it is true that $a \propto b$ and $c \propto d$ implies $ac \propto bd$ and that $a/c \propto b/d$ provided c and d are nonzero. The statement $a \propto b$ also implies that $a^f \propto b^f$ for any real number f. These observations allow one to make useful manipulations. For example, if the mass $M \propto V$ of a mammoth is proportional to its volume $V = L^3$, then the size L of a roughly cubical or spherical mammoth would satisfy $L \propto V^{1/3} \propto M^{1/3}$ and so the thermal relaxation time of the mammoth would be $\tau \propto L^2 \propto M^{2/3}$.

8. \mathbf{T} / \mathbf{F} The result of executing the Mathematica code

will be the list $\{ -2, -2 \}$.

Answer: T This tested whether you understand some basic notation that you learned from the Mathematical tutorial of the first homework assignment. The **Table** command generates a list, here a list of two numbers $x_i - y_i$ for i = 1 and i = 2. The notation x[[i]] is the notation for accessing the *i*the component or member of the list x.