Some discussion and/or hints on assignment 07.
There may be misprints, so use with care.

– – – – – – – – – – – – – – – ⋆ – – – – – – – – – – – – –

1. **The unix/linux command line**

2. **Newton's method for finding a minimum**

   (a) Consider the single-variable function

$$f(x) = 1 + x^{4/3}$$

which has a single minimum at $x = 0$. Write down the iteration equation for Newton' method.
The iterations do not converge. By sketching the derivative $f'(x)$, explain graphically why.

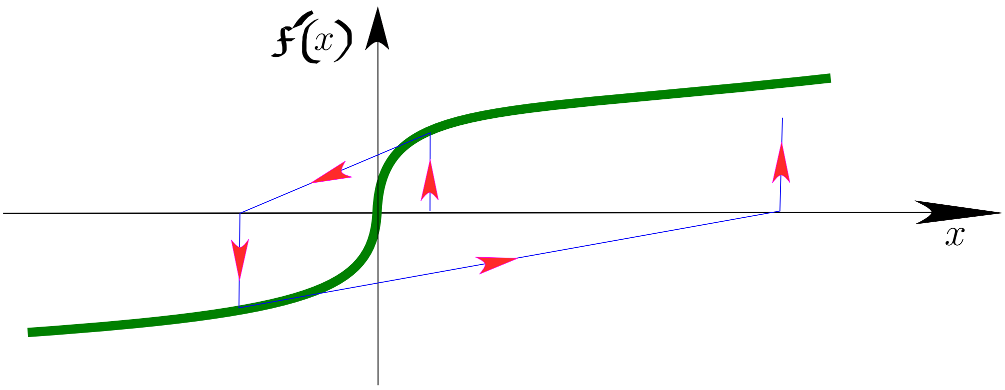
Discussion/Hint $\longrightarrow$

$$f'(x) = (4/3)x^{1/3}; \quad f''(x)(4/9)x^{-2/3}.$$

So the iteration equation is

$$x_{k+1} = x_k - f'(x_k)/f''(x_k) = x_k - 3x_k = -2x_k$$

What a simple iteration equation. If you start from $x = 1$ you get $1, -2, 4, -8, 16, \ldots$. Never converges to $x = 0$.
If you plot $f'(x)$, you should be able to see this graphically, if you know what Newton's iteration is trying to do. For root-finding, the Newton-Raphson algorithm approximates the function $f(x)$ at each step by it's tangent at that point. Please look up how this works graphically, e.g, wikipedia Newton's method.
For minimization, the algorithm approximates the derivative $f'(x)$ at each step by its tangent. So if you have plotted $f'(x)$, you should be able to graphically see how the iteration proceeds:



Finally, another way to look at this: For a single-variable iterative scheme

$$x_{k+1} = g(x_k),$$

you can show graphically or algebraically that, if $g'(x) \in (-1, 1)$, there is convergence, if $g'(x)$ is outside this range, there is no convergence. E.g., see

`http://wwwf.imperial.ac.uk/metric/metric_public/numerical_methods/iteration/fixed_point_iteration.html`

In our case, $g(x) = -2x$, so $g'(x) = -2$ which is outside $(-1, 1)$. Therefore the iteration will not converge.

Comment 1: If $g'(x) \in (-1, 1)$ for some x and not for some other values of x, it's more complicated, but our case is simple.

Comment 2: these considerations are more general than Newton's iteration, and generally valid for any fixed-point iteration.

=-=-=-=-=-=-=-=-=-=

(b) We want to find the location of the minimum of the function of 2 variables

$$f(x, y) = x^2 + xy + y^2 + 3x$$

using Newton's method. Set up the iteration equations. You will have to invert a 2×2 matrix.

Because the function is quadratic, the iteration should converge in a single step. Starting from any initial point (you can choose), show that the iteration converges in a single step.

Discussion/Hint $\longrightarrow$

The Hessian turns out to be a constant matrix, not a function of (x,y), because the function is quadratic. Once you invert the Hessian and calculate

$$H^{-1}\nabla f$$

you get the vector

$$\begin{pmatrix} x+2 \\ y-1 \end{pmatrix}$$

So that the iteration equations are

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \begin{pmatrix} x+2 \\ y-1 \end{pmatrix}$$

which simplifies to

$$x_{k+1} = -2; \qquad y_{k+1} = 1.$$

This does not even depend on what the current value of the iterates are — the next value will always be $(-2, 1)$. In other words, the minimum is reached in a single iteration!

This is not surprising because we are dealing with a quadratic function, for which Newton's iteration is exact.

=-=-=-=-=-=-=-=-=-=

(c) I have a multidimensional minimization problem and calculating the inverse of the Hessian ($\mathbb{H}^{-1}$) is too hard for me. As a very crude approximation, I replace $\mathbb{H}^{-1}$ by $\lambda\mathbb{I}$, where λ is a small number and $\mathbb{I}$ is the unit matrix. The resulting algorithm is then equivalent to a widely used algorithm. Which one?

Discussion/Hint $\longrightarrow$

You should be able to show that you end up with the "gradient descent" algorithm.

=-=-=-=-=-=-=-=-=-=

3. **Statistical Mechanics**

Note: Some of the following is worked out in the student projects linked to on the webpage, under “Ising model”.

Remember $Z = \sum_\alpha e^{-\beta E_\alpha}$, where E_α is the energy of the configuration α. At thermal equilibrium, each configuration appears with probability $P_\alpha = \frac{1}{Z} e^{-\beta E_\alpha}$.

(a) Show that the system energy and its square has expectation values

$$\langle E \rangle = \frac{1}{Z}\frac{\partial Z}{\partial \beta} = \frac{\partial}{\partial \beta} \ln Z \qquad \text{and} \quad \langle E^2 \rangle = \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2}$$

Discussion/Hint $\longrightarrow$

The crucial point for this problem is:
If $e^{-\beta E_a}$ is the probability (up to normalization) of the configuration a, then any quantity x has the expectation value:

$$\langle x \rangle = \frac{1}{Z} \sum_a x_a e^{-\beta E_a}$$

Here $1/Z$ serves as the normalization. Z is called the partition function.
So for example

$$\langle E \rangle = \frac{1}{Z} \sum_a E_a e^{-\beta E_a}$$
$$\langle E^2 \rangle = \frac{1}{Z} \sum_a (E_a)^2 e^{-\beta E_a}$$

Taking derivatives of $Z = \sum_\alpha e^{-\beta E_\alpha}$ with respect to β should give the desired relations.

=-=-=-=-=-=-=-=-=-=

(b) Hence show that the specific heat

$$C_v = \frac{\partial \langle E \rangle}{\partial T}$$

can be written as $C_v = \beta^2 \Big(\langle E^2 \rangle - \langle E \rangle^2 \Big)$. I am probably setting the Boltzmann constant to unity here.

Discussion/Hint $\longrightarrow$

An exercise in taking derivatives. The relevant definitions for $\langle E \rangle$ and $\langle E^2 \rangle$ are above.

=-=-=-=-=-=-=-=-=-=

(c) Perform the corresponding derivation for the magnetic susceptibility. You might have to remember that the configuration energy contains a term $-BM_\alpha$, where B is the magnetic field (often written as H) and M_α is the total magnetization of configuration α.

Discussion/Hint $\longrightarrow$

You could define $E_\alpha = A_\alpha - BM_\alpha$, to help remember that the configuration energy contains a term $-BM_\alpha$. The A_α part of the configuration energy does not depend on B. The definitions of expectation values are

$$\langle M \rangle = \frac{1}{Z} \sum_a M_\alpha e^{-\beta E_\alpha}$$

$$\langle M^2 \rangle = \frac{1}{Z} \sum_a (M_\alpha)^2 e^{-\beta E_\alpha}$$

One can now proceed to calculate

$$\chi = \frac{\partial \langle M \rangle}{\partial B}$$

and express the resulting expression in terms of $\langle M \rangle$ and $\langle M^2 \rangle$.

=-=-=-=-=-=-=-=-=-=