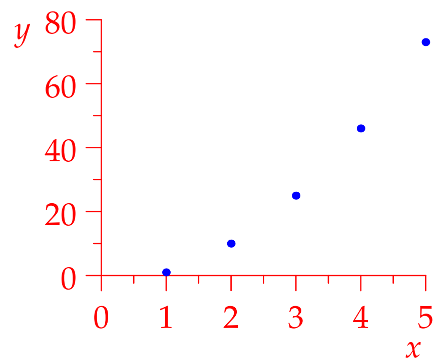
4 Sequences as Functions

We've seen several types of function in this course and used them to model various situations. In practice, one is often faced with the opposite problem: given experimental data, what type of function should you try?

4.1 Polynomial Sequences: First, Second, and Higher Differences

To begin to answer this, it is helpful to think of a very basic type of function: a *sequence.* Loosely speaking, a sequence a function whose domain is a set like the natural numbers, for example

$$f : \mathbb{N} \to \mathbb{R} : n \mapsto 3n^2 - 2$$



defines the sequence

$$\big(f(1), f(2), f(3), \dots\big) = \big(1, 10, 25, 46, 73, \dots\big)$$

Sequences are an intuitive way to expose even elementary school students to the idea of a function without requiring any formal acknowledgement of a formula/rule or even a domain: the *first* term is 1, the *second* is 10, the *third* is 25, etc.

Such a function might arise from an experiment; indeed it is particularly common to take measurements at fixed time intervals. Suppose this has been done, and that what you therefore have is the data set

x	1	2	3	4	5
y	1	10	25	46	73

Could you recover the function $y = f(x) = 3x^2 - 2$ directly from this data? Plotting the data (as above) is a good place to start, though it is hard to decide directly from the plot whether we should be hunting for a quadratic function, or some other power function/polynomial, or even an exponential. Of course, the physical source of real-world data might also provide clues.

A more mathematical approach involves considering how data values *change*:

$$\begin{array}{cccccccc} & & +1 & & +1 & & +1 & \\ x & 1 & \to & 2 & \to & 3 & \to & 4 \\ y & 1 & \to & 10 & \to & 25 & \to & 36 \\ & & +9 & & +15 & & +21 & \\ & & & +6 & & +6 & & \end{array}$$

The first-differences in the x-values are constant whereas those for the y-values are increasing

$$(y_{n+1} - y_n) = (9, 15, 21, 27, \dots)$$

You likely already notice the pattern: the sequence of first-differences is increasing *linearly* as the *arithmetic sequence*

$$y_{n+1} - y_n = 3 + 6n$$

To make this even clearer, note that the sequence of second-differences in the y-values is *constant* $(+6)$. These facts are huge clues that we expect a quadratic function.

But why? Well we can certainly check the following directly:

Linear Model If $f(n) = an + b$, then the sequence of first-differences is constant

$$f(n+1) - f(n) = a$$

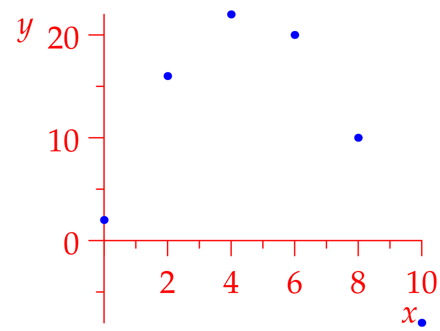
Quadratic Model If $f(n) = an^2 + bn + c$, then the sequence of first-differences is linear and the second-differences are constant:

$$g(n) := f(n+1) - f(n) = 2an + a + b, \qquad g(n+1) - g(n) = 2a \tag{$*$}$$

The relationship between these results and the *derivative*(s) of the original function $f(x)$ should feel intuitive: what happens if you differentiate a quadratic twice? The sequential approach in this chapter is a form of *discrete calculus*: using a pattern of *differences* to predict the original function is similar to how we use knowledge of a derivative $f'(x)$ to find $f(x)$.

Example 4.1. You are given the following data set

x	0	2	4	6	8	10
y	2	16	22	20	10	-8



The x-values have constant first-differences while the y-values have constant second-differences

First-differences: $14, 6, -2, -10, -18$

Second-differences: $-8, -8, -8, -8$

We therefore *suspect* a quadratic model $y = f(x) = ax^2 + bx + c$. But how to find it?

We *could* try using the above formulæ $(*)$, but this is ugly, not least because the constant x-differences are not 1. It is easier just to substitute in the first three values:

$$\begin{cases} 2 = f(0) = c \\ 16 = f(2) = 4a + 2b + c \\ 22 = f(4) = 16a + 4b + c \end{cases} \implies \begin{cases} c = 2 \\ 4a + 2b = 14 \\ 8a + 2b = 10 \end{cases} \implies 4a = -4$$

whence $a = -1$, $b = 9$ and $c = 2$. A quadratic model is therefore

$$y = f(x) = -x^2 + 9x + 2$$

It is easily verified that the remaining data values satisfy this relationship. Indeed, we could have used substituted any three data pairs from the table and found the same model.

Limitations of the Sequential Approach

Consider the question we're answering in Example 4.1, "Find a model satisfying given data." There are at least two issues with our method:

1. Constant second-differences don't guarantee that *only* a quadratic model is suitable. For example,

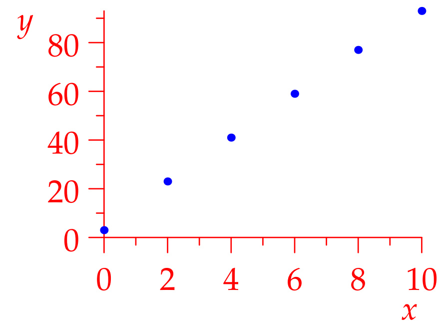
$$y = -n^2 + 9n + 2 + n(n-2)(n-4)(n-6)(n-8)(n-10)$$

   is a very complicated model satisfying the same given data set! Our old model would have predicted $y(1) = 10$, but the new model gives $y(1) = -935$ (!). Since the usefulness of a model often lies in its ability to *predict* as yet unseen results, this is potentially a big problem. Of course simplicity is also a desired property of a model, and our original quadratic certainly provides that.

2. Experimental data is very unlikely to fit such precise patterns: the real-world is noisy with all sorts of error, not least that of the observer. However, if the differences are *close* to satisfying such patterns, then you should feel confident that a linear/quadratic model is a good choice.

Example 4.2. Given the data set

x	0	2	4	6	8	10
y	3	23	41	59	77	93



with sequences of first- and second-differences

First-differences: 20, 18, 18, 18, 16

Second-differences: $-2, 0, 0, -2$

do you think a linear or quadratic model would be superior?

If you wanted a linear model, you'd likely be inclined to try $f(x) = 9x + b$ for some constant b. Here are two options:

1. $f(x) = 9x + 5$ fits the middle four data values perfectly, but as a predictor is too large at the endpoints: $f(0) = 5 > 3$ and $f(10) = 95 > 93$.

2. $f(x) = 9x + 5 - \frac{2}{3}$ doesn't pass through any of the data values but seems to reduce the net error to zero:

x	0	2	4	6	8	10
$f(x) - y$	$-\frac{4}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$-\frac{4}{3}$

$$\implies \sum_x f(x) - y = 0$$

Neither model is perfect, but then this is what you expect with real-world data!

Exercise 4 suggests an approach for finding an approximate quadratic model.

Exercises 4.1. *Key concepts: Sequences are simple functions, Recognizing linear and quadratic models from constant first- and second-differences*

1. For each data set, find a function $y = f(x)$ modelling the data.

   (a)

x	2	4	6	8
y	-1	2	7	14

   (b)

x	2	5	8	11	14
y	-6	-15	-6	21	66

   (c)

x	0	6	9	15
y	3	15	21	33

   (*Be careful here: the x-differences aren't constant!*)

2. Suppose a table of data values containing (x_0, y_0) has constant first-differences in both variables

   $$\Delta x = x_{n+1} - x_n = a, \quad \Delta y = b$$

   Find the equation of the linear function $y = f(x)$ through the data.

3. What relationship do you expect to find with the sequential differences of a cubic function $f(n) = an^3 + bn^2 + cn + d$? What about a degree-$m$ polynomial $f(n) = an^m + bn^{m-1} + \cdots$? Why?

4. Suppose $f(n) = an^2 + bn + c$ is an approximate quadratic model for the data in Example 4.2 that has constant second-differences equalling -1.

   (a) By computing the first- and second-differences for the model f, show that $a = -\frac{1}{8}$.

   (b) Now compute the sequence $y + \frac{1}{8}x^2$. What *might* be reasonable values for b, c?

5. Suppose $f(x)$ is a twice-differentiable function and $h > 0$ is constant.

   (a) Use the mean value theorem from calculus to explain the following.

      i. First-differences satisfy $f(x+h) - f(x) = f'(\xi)h$ for some $\xi \in (x, x+h)$.

      ii. (Hard) Second-differences satisfy

      $$\big(f(x+2h) - f(x+h)\big) - \big(f(x+h) - f(x)\big) = f''(\xi)h(\xi_1 - \xi_2)$$

      for some ξ, ξ_1, ξ_2 satisfying

      $$x < \xi_2 < x+h, \qquad x+h < \xi_2 < x+2h, \qquad \xi_2 < \xi < \xi_1$$

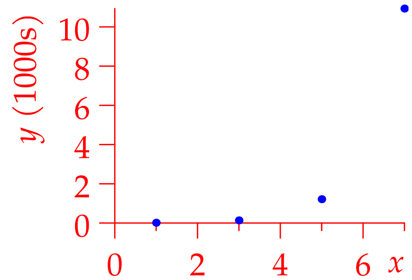
   (b) Find the ξ terms in the above formulæ when $f(x) = ax^2 + bx + c$ is quadratic.

4.2 Exponential, Logarithmic & Power Sequences

Instaed of differences, you might also have to consider *ratios* between successive terms.

Example 4.3. Consider the data below. At first glance, you might believe that the data to be approximately quadratic, but how can we tell? Well, *if* the data were quadratic, *then* we should get constant second-differences. However, applying the constant-difference method doesn't seem to yield anything helpful:

$$\begin{array}{c|ccccccc} x & 1 & \xrightarrow{+2} & 3 & \xrightarrow{+2} & 5 & \xrightarrow{+2} & 7 \\ \hline y & 15 & \xrightarrow[+120]{\times 9} & 135 & \xrightarrow[+1080]{\times 9} & 1215 & \xrightarrow[+9720]{\times 9} & 10935 \\ & & & \xrightarrow{+960} & & \xrightarrow{+8640} & & \end{array}$$



Regardless, since we only have two second-differences, any conclusion relying on this would be very weak. If instead we think about successive ratios of y-values, then a different pattern emerges: the output scales by 9 when 2 is added to the input, $f(x+2) = 9f(x)$. *Exponential* functions convert addition to multiplication,[15] so we try such a function. If $y = f(x) = ba^x$ for some constants a, b, then

$$9f(x) = f(x+2) = ba^{x+2} = ba^2b^x = a^2 f(x) \implies a^2 = 9$$

A suitable model is therefore $y = 5 \cdot 3^x$.

We can see the pattern in the example more generally.

Exponential Model If $f(x) = ba^x$, then adding a constant to x results in

$$f(x+k) = ba^{x+k} = a^k f(x)$$

If x-values have constant differences ($\Delta x = k$), then y-values will be related by a constant *ratio* ($r_y = a^k$). You might remember this as 'addition–product' or 'arithmetic–geometric.'

Such a simple pattern is often disguised:

- Complete data might not be given, so you might have to skip some data values to see a pattern. For example, if our original data was

x	1	3	4	5	7
y	15	135	405	1215	10935

  then the x-values are not in a strictly arithmetic sequence. Ignoring the third pair produces the exponential model above, which is then seen to satisfy $f(4) = 5 \cdot 3^4 = 405$.

- As in Example 4.2, experimental data will likely only *approximately* exhibit such patterns.

[15]This is exactly what the exponential law says: $a^{x+y} = a^x \cdot a^y$.

Example 4.4. A population P of rabbits is measured every two months, producing the data set

t	0	2	4	6	8	10
P	5	7	10	14	19	28

This data seems very close to being quadratic. Consider the first and second sequences of P-differences

$$\Delta P = (2,3,4,5,9), \qquad \Delta\Delta P = (1,1,1,4)$$

The last difference doesn't fit the pattern. Should we look for an exponential model instead? Indeed we *expect* such a model precisely because the data is measuring population growth! We therefore consider the ratios of P-values:

$$(r_y) = (1.4, 1.43, 1.4, 1.36, 1.47)$$

The ratios are *very close* to being constant $r_y \approx 1.4 = \frac{28}{5}$, strongly suggesting an exponential model! To exactly match the first and last data values, we could take the model

$$P(t) \approx 5\left(\frac{28}{5}\right)^{\frac{t}{10}}$$

t	0	2	4	6	8	10
P	5	7.06	9.96	14.06	19.84	28

where predictions are given to 2 decimal places. Only $P(8)$ doesn't match the original data set when rounding to the nearest integer.

Logarithmic and Power Models

We've seen that addition-addition corresponds to a linear model and that addition-multiplication to an exponential. There are two other natural combinations.

Logarithms These operate exactly as exponential models but in reverse. If $f(n) = \log_a x + b$, then

$$f(kx) = \log_a(kx) + b = \log_a k + \log_a x + b = \log_a k + f(x)$$

Otherwise said, *multiplying* the input by a constant $r_x = k$ *adds* a constant $\Delta y = \log_a k$ to the output. This could be summarized as 'product–addition.' Since it is typically easier to work with exponential models, the suggested approach to such cases is simply to switch the roles of x and y: find an exponential model $x = pq^x$.

Power Functions If $f(x) = ax^m$, then

$$f(kx) = a(kx)^m = ak^m x^m = k^m f(x)$$

That is, *multiplying* the input by a constant $r_x = k$ also *multiplies* the output $r_y = k^m$. We have a 'product–product' relationship between successive terms.

Examples 4.5. Find the patterns in the following data and suggest a model $y = f(x)$ in each case.

x	6	18	54	162
y	1	2	3	4

x	3	6	9	12
y	135	1080	3645	8640

Start by computing the x- and y-differences and ratios...

Predicting Values Given Limited Data

Suppose you are given two data points. Can you predict the output for another given input? The answer depends on what type of model you choose. Here is a typical problem.

Example 4.6. Suppose a function $y = g(x)$ satisfies $g(2) = 3$ and $g(4) = 9$. What do you think should be the value of $g(8)$?

Everything depends on the type of model you try...

1. For a linear (addition-addition) model we know that $\Delta x = 4 - 2 = 2$ corresponds to $\Delta y = 9 - 3 = 6$. It follows that

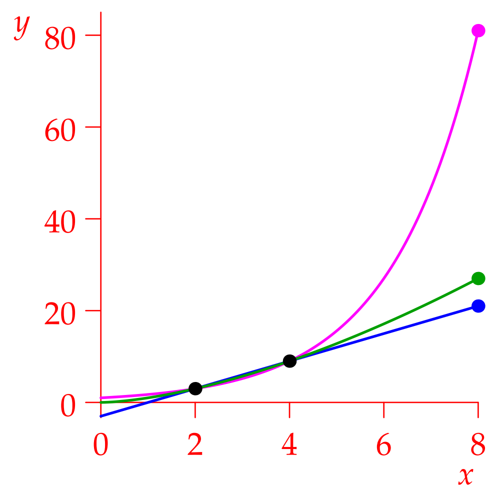
$$\begin{aligned} g(8) &= g(4 + 2\Delta x) = g(4) + 2\Delta y = 9 + 12 \\ &= 21 \end{aligned}$$

2. For an exponential (addition-product) model, the constant x-differences $\Delta x = 2$ correspond to a constant *y-ratio* $r_y = \frac{9}{3} = 3$. We therefore have

$$\begin{aligned} g(8) &= r_y g(6) = r_y^2 g(4) = 9 \cdot 9 \\ &= 81 \end{aligned}$$

3. For a power (product-product) model, $r_x = \frac{4}{2} = 2$ corresponds to $r_y = 3$, whence

$$\begin{aligned} g(8) &= g(4 \cdot 2) = g(4r_x) = r_y g(4) = 3 \cdot 9 \\ &= 27 \end{aligned}$$



1. $g(x) = 3x - 3$
2. $g(x) = 3^{x/2}$
3. $g(x) = x^{\log_2 3}$

Notice how we didn't need to calculate the models explicitly, though they are stated below the graph for convenience. We could also use a logarithmic model to predict $g(8)$. The details of all this are left as an exercise.

Exercises 4.2. *Key concepts: Exponential sequences (addition-product), Logarithmic sequences (product-addition), Power sequences (product-product), Predicting the next term in a sequence*

1. Find the patterns in the following data sets and use them to find a model $y = f(x)$.

   (a)

x	0	1	2	3	4
y	80	120	180	270	405

   (b)

x	2	4	8	10
y	1	16	256	625

   (c)

x	1	3	5	7	9
y	15	5	19	57	119

   (d)

x	1	3	4	6
y	1	36	216	7776

   (e)

x	20	60	180	540
y	2	4	6	8

   (f)

x	2	6	54	486	4374
y	2	4	8	12	16

2. Take logarithms of the power relationship $y = ax^m$. What is the relationship between $\ln y$ and $\ln x$? Use this to give another reason why the inputs and outputs of power functions satisfy a 'product–product' relationship.

3. How does our analysis of exponential functions change if we add a constant to the model? That is, how might you recognize a sequence arising from a function $f(x) = ba^x + c$?

4. Suppose $f(5) = 12$ and $f(10) = 18$. Find the value of $f(20)$ supposing $f(x)$ is a:

   (a) Linear function;

   (b) Exponential function;

   (c) Power function;

   (d) Logarithmic function.

   If an experiment returns $f(20) = 39$, which of the four models do you think would be more appropriate?

5. Recall Example 4.6.

   (a) Derive the equations for the linear, exponential and power models.

   (b) Find the predicted value of $g(8)$ if we use a logarithmic model. Find the model's equation.

4.3 Newton's Method

Sequences are often used as a way to *approximate* something useful. In this section we discuss a general approach for approximating solutions to equations.

Example 4.7. Suppose $x_n > \sqrt{2}$. Then $\frac{2}{x_n} < \frac{2}{\sqrt{2}} = \sqrt{2}$. It seems reasonable that the average

$$x_{n+1} := \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \tag{$*$}$$

should be closer to $\sqrt{2}$. Starting with an initial guess, say $x_0 = 2$, we obtain the sequence

$$x_1 = \frac{1}{2}\left(2 + \frac{2}{2}\right) = \frac{3}{2}, \quad x_2 = \frac{17}{12} = 1.4166\ldots, \quad x_3 = \frac{577}{408} = 1.4142\ldots, \quad \ldots$$

This sequence certainly *appears* to be converging to $\sqrt{2}$...

A rigorous proof that the sequence converges to $\sqrt{2}$ requires more detail than is appropriate for this discussion (see Exercise 7). Two observations should make it seem more believable:

1. *If* the sequence defined by $(*)$ has a limit L, then it must satisfy

$$L = \frac{1}{2}\left(L + \frac{2}{L}\right) \implies 2L^2 = L^2 + 2 \implies L^2 = 2 \implies L = \sqrt{2}$$

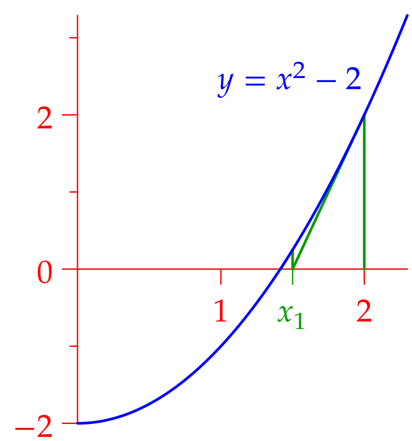
   where we take the positive root since all terms x_n are plainly positive.

2. The algorithm can be interpreted *geometrically*.
   The tangent line to the curve $y = f(x) = x^2 - 2$ at $(x_n, f(x_n))$ has equation

$$\begin{aligned} y &= f(x_n) + f'(x_n)(x - x_n) \\ &= x_n^2 - 2 + 2x_n(x - x_n) \\ &= 2x_n x - x_n^2 - 2 \end{aligned}$$

   Its intersection with the x-axis yields the next iterate:

$$y = 0 \implies x = \frac{x_n^2 + 2}{2x_n} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) = x_{n+1}$$



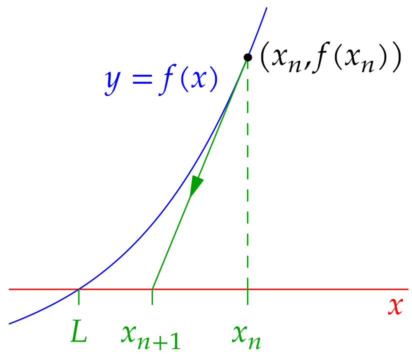
Since it makes use of the average, this approach is sometimes called the *method of the mean*. It may be applied to approximate any square root $\sqrt{a}$ where $a > 0$ (Exercise 4). Variations of this approach have been in use for thousands of years.

More importantly, this geometric interpretation generalizes...

Definition 4.8. Given a differentiable function $f(x)$ with non-zero derivative, the *Newton–Raphson iterates* of an initial value x_0 are defined by the recurrence formula

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

Our two previous observations still hold:



1. If $L = \lim_{n\to\infty} x_n$ exists and $f'(L) \neq 0$, then

$$L = L - \frac{f(L)}{f'(L)} \implies f(L) = 0$$

That is, the limit L is a root of the function $f(x)$.

2. The tangent line at $\big(x_n, f(x_n)\big)$ forms a right-triangle with base $x_n - x_{n+1}$ and height $f(x_n)$. Its slope is

$$f'(x_n) = \frac{f(x_n)}{x_n - x_{n+1}}$$

which rearranges to give the formula $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

For polynomials with integer coefficients, the iterates form a sequence of *rational numbers*.

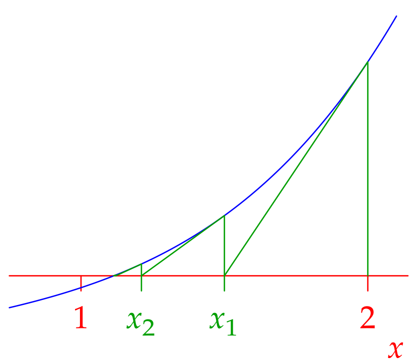
Examples 4.9. We use Newton's method to find approximations to two irrational numbers.

1. To find a root of $f(x) = x^4 + 4x - 6$, start with $x_0 = 2$ and iterate

$$x_{n+1} = x_n - \frac{x_n^4 + 4x_n - 6}{4x_n^3 + 4} = \frac{3(x_n^4 + 2)}{4(x_n^3 + 1)}$$

which yields the sequence (to 3 d.p.)

$$\left(2, \frac{3}{2}, \frac{339}{280}, \ldots\right) = (2, 1.5, 1.211, 1.121, 1.114, 1.114, \ldots)$$



A calculator shows that 1.114 is approximately a root.

2. The irrational number $x = \sqrt{2} + \sqrt{3}$ is a root of the polynomial

$$f(x) = x^4 - 10x^2 + 1$$

Applying Newton's method with $x_0 = 3$ results in the sequence (to 3 d.p.)

$$x_{n+1} = x_n - \frac{x_n^4 - 10x_n^2 + 1}{4x_n^3 - 20x_n} \implies (x_n) = \left(3,\ \frac{19}{6} = 3.167,\ 3.147,\ \ldots\right)$$

Newton's method can be attempted for any differentiable function, though the sequence isn't guaranteed to converge: see for instance Exercise 6. You can find graphical interfaces online for playing with this. Newton's method is one of the most important algorithms in numerical analysis, with many applications and generalizations to other contexts.

Exercises 4.3. *Key concepts: Visualizing and computing with Newton's method*

1. Use Newton's method to find a root of the given function to 4 decimal places.

   (*Use a calculator, but explain what you are doing!*)

   (a) $f(x) = x^3 - 4$
   (b) $f(x) = 2x^3 + x - 1$
   (c) $f(x) = e^x - \sqrt{x} - 2$
   (d) $f(x) = \cos x - x$ (*remember to use radians*)

2. In Example 4.9.1, we claimed that the sequence of iterates approaches an *irrational* number. Why is this obvious?

3. Use Newton's method to find a rational number approximation to $\sqrt[3]{2}$ in lowest terms $\frac{p}{q}$ where $10 < q < 100$.

4. Here is the general *method of the mean*. To approximate $\sqrt{a}$, choose any positive x_0 and define

$$x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

   (a) If the sequence (x_n) has a limit L, prove that $L = \sqrt{a}$.
   (b) Show that this is Newton's method in disguise. What is the relevant function $f(x)$?

5. We might consider a *method of the mean* for approximating $\sqrt[3]{2}$: given x_0, define

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n^2}\right)$$

   (a) If the sequence (x_n) converges, show that its limit is $\sqrt[3]{2}$.
   (b) If $x_n > \sqrt[3]{2}$, show that $\frac{2}{x_n^2} < \sqrt[3]{2}$.
   (c) Let $x_0 = 1$. Compute x_1 and x_2. Compare these with the values obtained using Newton's method for the function $f(x) = x^3 - 2$ with the same initial condition $x_0 = 1$ (e.g., Exercise 3).

6. Let $f(x) = x^3 - 5x$.

   (a) What happens if you apply Newton's method to this function with initial condition $x_0 = 1$? Draw a picture to illustrate.

   (b) (Just for fun!) Investigate what happens for other values of x_0. Can you make any conjectures? Is is possible for x_0 to be *positive* and yet for $x_n \to -\sqrt{5}$? Can you make any sense of what happens if $1 < x_0 < \sqrt{\frac{5}{3}}$?

7. Suppose you perform Newton's method for $f(x) = x^2 - 2$ starting with some $x_0 > 0$.

   (a) Check that

   $$x_{n+1} - \sqrt{2} = \frac{1}{2x_n}(x_n - \sqrt{2})^2 \tag{$*$}$$

   so that all terms of the sequence (except perhaps the first) are larger than $\sqrt{2}$.

   (b) If $n \geq 1$, use $(*)$ to verify that

   $$x_{n+1} - \sqrt{2} \leq \frac{1}{2}(x_n - \sqrt{2})$$

   Hence argue that $\lim_{n\to\infty} x_n = \sqrt{2}$.