sn't parallel to
ooking for its
$c_1$ as the 'next
$_0$ as the initial



which is the $x$-coordinate of the point where the tangent line at $(x_1, f(x_1))$ meets the $x$-axis, regarding $x_2$ as the next approximation, and so on. See Fig. 9(*a*). As Fig. 9(*b*), shows, the process may go badly wrong! We shall not attempt very much theory here (see, however, Sections 3.5–3.9 below), but rather illustrate the method with a variety of examples.

Given $f$ and the initial guess $x_0$ we define

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{2}$$

for $k = 0, 1, 2, 3, \ldots$, so long as $f'(x_k) \neq 0$ (and it's very unlikely to be *exactly* 0), and consider the sequence of 'approximations', or *iterates* as we shall call them, $x_0, x_1, x_2, \ldots$.

### 3.1 *Program: Newton–Raphson*

```
1Ø INPUT XØ
2Ø LET X = XØ
3Ø Work out f(x) = a and f'(x) = b from explicit formulae for
   f and its derivative f'
4Ø LET X = X – A/B
5Ø PRINT X
6Ø GOTO 3Ø
```

Ideally insert a line 55 which makes the micro pause before going on to the next iteration. If you can define functions in the program, say FNF(X) for $f$ and FND(X) for $f'$, then you can omit line 3Ø and have

```
 5 DEF FNF(X) = formula for f(x)
 6 DEF FND(X) = formula for f'(x)
4Ø LET X = X – FNF(X)/FND(X)
```

### 3.2 *Exercise*

The graph in Fig. 9(*b*) may remind you of the inverse tangent graph $y = \arctan x$ (also written $\tan^{-1} x$). For $y = \arctan x$ we have $f'(x) = 1/(1 + x^2)$, so line 4Ø of the program becomes

```
4Ø LET X = X – (1 + X*X)*ATN(X)
```

Try letting $x_0$ equal: (*a*) 1; (*b*) 2; (*c*) 1.3; (*d*) 1.4; (*e*) 1.391; (*f*) 1.392. The solution sought is of course $x = 0$. You should find that (*a*), (*c*) and (*e*) are 'good' values for $x_0$, and do indeed give the solution $x = 0$ after a few iterations, but (*b*), (*d*) and (*f*) are 'bad' values and go off to infinity in the manner of Fig. 9(*b*).

What is the 'critical' value of $x_0$ which separates the good from the bad? It is the value of $x_0$ which makes $x_1 = x_0 - (1 + x_0^2)(\arctan x_0)$

precisely equal to $-x_0$, so that a second iteration merely sends us back to $x_0$. (Compare Exercise 4.2 below.) That is, the critical value is a solution of arc tan $x = 2x/(1+x^2)$. Now refer back to Exercise 2.9. (Of course, you could solve *this* equation by Newton–Raphson . . . .)

### 3.3 *Exercise*

What happens if the Newton–Raphson method is used to seek the root $x = 0$ of $f(x) = 0$ where (*a*) $f(x) = x^{1/3}$, (*b*) $f(x) = x^{3/5}$, (*c*) $f(x) = x^{2/3}$? In each case $x_{k+1}$ can be expressed very simply in terms of $x_k$, making it relatively easy to spot what will happen. Sketches of the three curves $y = f(x)$ also help.

Note that if, say, $x^{1/3}$ is to be evaluated on the micro we have to do better than writing $X\uparrow(\frac{1}{3})$, for this will be rejected when $x < 0$ (also possibly for $x = 0$). (The micro tries to work out $(\ln x)/3$, and naturally fails because $\ln x$ is undefined for $x \leqslant 0$.) It is necessary to do something like

$$\text{sign}\,(x)\,(\text{abs}\,(x))^{1/3}$$

where sign$(x)$ is $+1$ for $x > 0$, $-1$ for $x < 0$ and 0 for $x = 0$, while abs $(x) = |x|$ is the absolute value of $x$.

Try $f(x) = 1 + x^{2/3}$ (you may be able to write this in BASIC as $1 + (X*X)\uparrow(\frac{1}{3})$, since $x^2 \geqslant 0$), which gives *no* solutions for $f(x) = 0$. You should find that the iterates $x_k$ eventually oscillate between about $x = \pm 5.19615$. Solve $x - f(x)/f'(x) = -x$ to find the exact value!

In this case, $x_{k+1} = -\frac{1}{2}x_k - \frac{3}{2}x_k^{1/3}$. *If* $x_k$ approaches a limit $l$ as $k \to \infty$, *then* this equation implies $l = -\frac{1}{2}l - \frac{3}{2}l^{1/3}$, which has only the solution $l = 0$. However, as you will discover by taking various values for $x_0$, there is no starting value $x_0$ which makes $x_k \to 0$ as $k \to \infty$. (In fact, if $|x_k| < 1$, then $|x_{k+1}| > |x_k|$.)

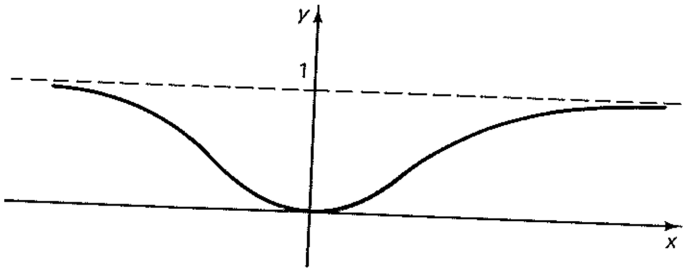
For $f(x) = 1 + x^p$, where $\frac{1}{2} < p < 1$ and $p$ is a fraction with an odd denominator, the $x_k$ eventually oscillate between $\pm(2p-1)^{-1/p}$. Try $p = 18/19$.

### 3.4 *Exercise*

Try seeking the root $x = 0$ of $1 - e^{-x^2} = 0$. Thus $x$ is here replaced by $x - (1 - e^{-x^2})/(2x\,e^{-x^2})$ in line 4Ø of Program 3.1. The shape of the graph $y = 1 - e^{-x^2}$ is suggested in Fig. 10. In BASIC, Y = 1 − EXP( − X*X). Try (*a*) $x_0 = 1$, (*b*) $x_0 = 1.6$, (*c*) $x_0 = 1.528$. Show that the critical value of $x_0$ separating good values of $x_0$ (for which $x_k \to 0$) from bad values (for which $x_k \not\to 0$) is a solution of $4x^2 + 1 = e^{x^2}$. Find this critical value using the method of bisection.

Note that if $x_0$ is much bigger than the critical value (e.g. $x_0 = 2$)

Fig. 10




then the Newton–Raphson method gives numbers $x_k$ which are wildly divergent: the iterates tend to infinity so fast that your computer will probably complain about dividing by zero after a couple of iterations. Also the approximations approach zero painfully slowly for say $x_0 = 1.5$. Why is this?

In the next few items we take the opportunity of making some slightly more technical observations on Newton–Raphson, and the limitations of the method.

### 3.5 *Exercise: Newton–Raphson can deceive us*

Suppose the iterates $x_k$ approach a limit $l$ as $k \to \infty$. Is $f(l) = 0$? If not, we are deceived by the method. Here is a wild example to show that deceit *is* possible. Let

$$f(x) = 1 - 2x \sin(1/x) \quad \text{for } x \neq 0,$$
$$f(0) = 1.$$

See Fig. 11 for a rough sketch. In fact $f$ is continuous for all $x$, and $f'(x)$ exists for all $x \neq 0$ (compare e.g. Spivak (1980) pp. 80, 146).

Take $x_0 = 1/(2\pi)$, so that $f(x_0) = 1$. Verify that $x_1 = 1/(4\pi)$, $x_2 = 1/(8\pi)$, and in general $x_k = 1/(2^{k+1}\pi)$. Thus certainly $x_k \to 0$, but equally certainly $f(0) \neq 0$. Verify also that $f'(x_k) = 2^{k+2}\pi$, so the derivative $\to \infty$ as $k \to \infty$. Try programming this, and starting with $x_0 = 1/(2\pi)$. Try also $x_0 = 0.5$. Explain what you see.

### 3.6 *A sufficient condition for Newton–Raphson to work*

Suppose $x_k \to l$ and $|f'(x)| \leqslant M$ for some constant $M$ and all $x$ sufficiently close to $l$. Then from

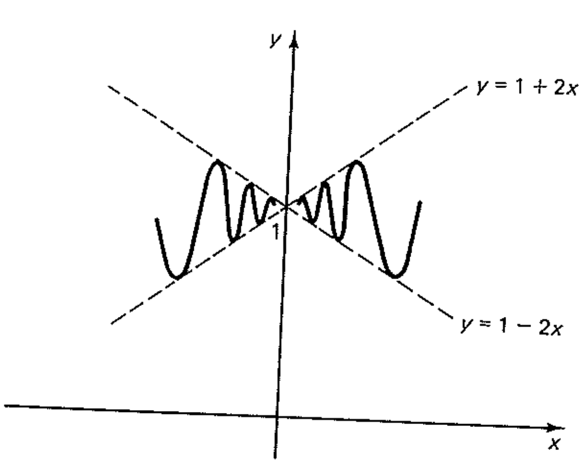
$$(x_{k+1} - x_k)f'(x_k) = f(x_k)$$

(which is the definition of $x_{k+1}$ in terms of $x_k$), it follows that $f(x_k) \to 0$, since the left-hand side $\to 0$. Thus $f(x_k) \to f(l) = 0$ so $l$ *is* a solution to $f(x) = 0$. (Why does this not contradict Exercise 3.5?)

### 3.7 *Do the $x_k$ approach a limit?*

How can we tell for sure that the iterates $x_k$ are approaching a limit $l$? The various numbers $x_k$ appearing on the screen stop changing after a while; what guarantee is there that they won't *start* changing again if we wait

Fig. 11




patiently? Actually none at all, though for 'reasonable' functions $f$ this won't happen. What you *can* do is to check on the value of $f(x_k)$ when $x_k$ has stopped changing, to see if it is nearly zero. Or you can evaluate $f(x_k - \epsilon)$ and $f(x_k + \epsilon)$ for $\epsilon$ around $10^{-7}$ to see if there is a change of sign, indicating a solution which will be $x_k$ to six decimal places or so.

Note the contrast with the bisection method, where we can be *certain*, after a definite number of iterations, that we have located a solution to say six decimal places.

## 3.8 *Another wild example*

This is based on Exercise 3.5, and shows that we can be deceived into believing that $x_k \to l$. We attach a parabola to part of the curve in Exercise 3.5 (Fig. 12). Fix a whole number $n$ and define $\epsilon = 1/(2n+1)\pi$. Let

$$f(x) = \begin{cases} 1 - 2x \sin(1/x) & \text{for } x \geqslant \epsilon \\ a(x^2 + 3x + 2) & \text{for } x < \epsilon \end{cases}$$

where $a = 1/(\epsilon^2 + 3\epsilon + 2)$. Then $f$ is continuous for all $x$; with more trouble we could make a curve that joined up more smoothly at $x = \epsilon$, but we do not need to do that.

Try $x_0 = 1/2\pi$ ( = 0.15915494 . . . ), $n = 100\,000$. (You will need to split line 4Ø of the Program 3.1 into two alternatives, according to whether $x < \epsilon$ or $x > \epsilon$.) You should find that the $x_k$ appear to approach 0 but then change their minds and arrive at $-1$ instead. If $n$ is very large then $\epsilon$ will be recorded as 0 by the micro, of course, so we shall never know in that case that the solution is $x = -1$ rather than $x = 0$.

## 3.9 *When will Newton–Raphson locate a solution?*

Suppose that $f(l) = 0$. What conditions on $x_0$, $l$ and $f$ will guarantee that $x_k \to l$ as $k \to \infty$? A thorough treatment of this question is to be found in books on Numerical Analysis (for example, Young and Gregory (1972), pp. 132, 146); here we indicate a proof of the following fact: *Suppose that $f$, $f'$ and $f''$* ( = second derivative of $f$) *are continuous for all $x$ near $l$, and $f(l) = 0$, $f'(l) \neq 0$. Then, provided $x_0$ is sufficiently close to $l$,* we have $x_k \to l$ as $k \to \infty$.

In fact, let $g(x) = x - f(x)/f'(x)$, so $g(x_k) = x_{k+1}$ and $g(l) = l$. We have

$g'(x) = f(x)f''(x)/(f'(x))^2$. Since $f'(l) \neq 0$ and $f'$ is continuous, there exists $c > 0$ with $|f'(x)| \geq c$ for all $x$ sufficiently close to $l$. Further, for $x$ sufficiently close to $l$, $|f(x)f''(x)| \leq \frac{1}{2}c^2$, since $f(l) = 0$ and $f$ and $f''$ are continuous. So $|g'(x)| \leq \frac{1}{2}$ for all $x$ close enough to $l$. By the Mean Value Theorem (see any book on calculus, e.g. Spivak (1980), p. 179),

$$x_{k+1} - l = g(x_k) - g(l) = (x_k - l)g'(x)$$

for some $x$ between $x_k$ and $l$. Hence, provided $x_0$ is close enough to $l$, $|x_1 - l| \leq \frac{1}{2}|x_0 - l|$, $|x_2 - l| \leq \frac{1}{2}|x_1 - l| \leq \frac{1}{4}|x_0 - l|$, and so on. Generally, $|x_k - l| \leq (1/2^k)|x_0 - l|$, so as $k \to \infty$, we have $x_k \to l$.

The example in Exercise 3.5 above fails because $f''(0)$ does not even exist; also $f'$ is not continuous at $x = 0$.

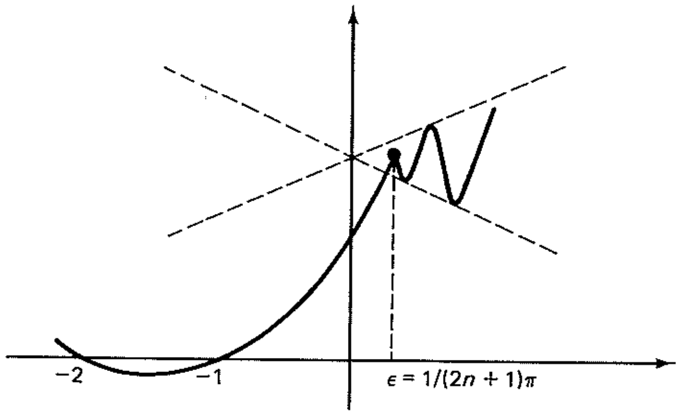
The above result assumes $f'(l) \neq 0$, i.e. $l$ is a *simple* solution of $f(x) = 0$. See Young and Gregory (1972), p. 146 for further details.

We now turn to examples where the Newton–Raphson method can be (more or less) successfully applied to the solution of equations. For the most part, the examples involve *polynomial* functions $f$, and the reader may now wish to proceed straight to Exercise 3.12, using a program such as Program 3.1 above. We pause, however, to consider the problem of *efficiently* evaluating a given polynomial.

## 3.10 *Evaluation of polynomials*

Let $f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ $(n \geq 1)$ be a polynomial, which we want to evaluate for a given value of $x$. Thus we want a program which accepts $x$, $n$, $a_0$, $\ldots$, $a_n$ as inputs and which calculates $f(x)$. The numbers $a_0, \ldots, a_n$ will be stored as an *array*, and because some micros only allow arrays to be indexed from 1 onwards (rather than 0), we write $a_j$ as an array element AA(J + 1) for $j = 0, 1, \ldots, n$. The array AA has dimension $n + 1$, that is contains $n + 1$ entries.

Fig. 12



Brute force will evaluate $f(x)$ by $1 + 2 + \ldots + (n-1)$ multiplications followed by $n$ additions, a total of $\frac{1}{2}n(n+1)$ operations. Suppose, however, that we work out in succession

$$b_1 = a_n x,\ b_2 = x(b_1 + a_{n-1}),\ b_3 = x(b_2 + a_{n-2}),$$

and so on, down to $b_n = x(b_{n-1} + a_1)$ and finally $b_{n+1} = b_n + a_0$. Then a little thought will convince you that $b_{n+1} = f(x)$, and working out $b_{n+1}$ involves only $n$ additions and $n$ multiplications, which is a great improvement if $n$ is large.

### 3.11 *Program: Evaluation*

```
1Ø INPUT N
2Ø DIM AA = N + 1
3Ø FOR J = 1 TO N + 1
4Ø INPUT AA(J)
5Ø NEXT J
6Ø INPUT X
7Ø LET F = AA(N + 1)*X
8Ø IF N = 1 THEN GOTO 12Ø
9Ø FOR J = 1 TO N - 1
1ØØ LET F = X*(F + AA(N - J + 1))
11Ø NEXT J
12Ø LET F = F + AA(1)
13Ø PRINT F
```

The final value of F in the program is $f(x)$.

When both $f(x)$ and $f'(x)$ are needed where

$$f'(x) = a_1 + 2a_2 x + \ldots + na_n x^{n-1},$$

then we could continue:

```
14Ø LET F1 = Ø
15Ø IF N = 1 THEN GOTO 21Ø
16Ø LET F1 = N*AA(N + 1)*X
17Ø IF N = 2 THEN GOTO 21Ø
18Ø FOR J = 1 TO N - 2
19Ø LET F1 = X*(F1 + (N - J)*AA(N - J + 1))
2ØØ NEXT J
21Ø LET F1 = F1 + AA(2)
```

The final value of F1 in the program is then $f'(x)$.

Lastly, we can go on to perform the Newton–Raphson iteration, supposing the $x$ in line 6Ø is the starting value $x_0$.

```
300 X=X-F/F1
310 PRINT X
320 GOTO 70
```

As usual, a line 315 causing the micro to pause until a key is pressed will make the iteration easier to follow.

At any stage it is possible to check the value of $f(x)$, to see whether it is small, or to check $f(x-\epsilon)$ and $f(x+\epsilon)$ for small $\epsilon$ (around $10^{-7}$) to see whether there is a change of sign. General theory predicts that once we are close enough to a root iteration will (subject to some mild conditions such as those in Section 3.9) bring us even closer.

### 3.12 *Exercise: Quadratic equations*

Take $f(x) = x^2 - y$, for various values $y > 0$, to find approximations to $\sqrt{y}$.

Take $f(x) = x^2 - 3x + 2$, with solutions to $f(x) = 0$ given by $x = 1$ and $x = 2$. Check experimentally that, if $x_0 > \frac{3}{2}$ (the half-way point between the roots), then $x_k \to 2$ and, if $x_0 < \frac{3}{2}$, then $x_k \to 1$. What happens if $x_0 = \frac{3}{2}$?

Try some other quadratic equations with real roots and check by experiment that, if $x_0$ is greater than the average of the two roots (i.e. $x_0 > -b/2$ for $x^2 + bx + c = 0$), then $x_k$ converges to the larger root; if less than the average then $x_k$ converges to the smaller root. Sketch diagrams of quadratic curves and the Newton–Raphson iteration to make this plausible.

### 3.13 *Exercise: Quadratic equations with complex roots*

Take $x^2 + x + 1 = 0$, which has complex roots. It is hardly likely that $x_k$ will get close to *them*, but try various values of $x_0$ (not equal to $-\frac{1}{2}$, which makes $f'(x_0) = 0$), and verify experimentally that the numbers $x_1, x_2, x_3, \ldots$ always jump about in a rather random-looking fashion. More of this strange phenomenon later – see §6. Try also $40x^2 - 40x + 11 = 0$.

### 3.14 *Exercise: Sensitivity to the starting value*

A good example with which to illustrate the sensitivity of the Newton–Raphson method to the starting value $x_0$ is the equation $2 \sin x - x = 0$, which has two non-zero solutions and the solution $x = 0$. Try various $x_0$ very close to 1 and on each side of 1 and notice how the iterates may diverge, or converge to any of the three solutions. According to Mackie and Scott (1985) the behaviour of the iterates depends not only on the value of $x_0$ but also on the machine being used!

§4
*Some cubic equations to solve*

**4.1** *Exercise: A simple cubic*
Let $f(x) = x(x-1)(x-2) = x^3 - 3x^2 + 2x$. Start with various values for $x_0$ and see when the corresponding approximations $x_k$ approach 0, when 1 and when 2. Is your result reasonable from a diagram?

**4.2** *Exercise: Alternating iterates*
Let $f(x) = x^3 - 2x + 2$ and $x_0 = 0$. You should find that the iterates $x_k$ alternate between 0 and 1. Find other cubic equations where the iterates alternate $a, b, a, b, \ldots$. (Hint: you need $f(a)/f'(a) = -f(b)/f'(b)$ where $f = 0$ is the cubic equation.) Try taking $x_0$ just $> 0$ or just $< 0$ in the above equation.
Try $2x^3 - 9x^2 + 11x - 3 = 0, \quad x_0 = 1$.

**4.3** *Exercise: Jumping iterates*
Let $f(x) = x^3 - 9x^2 + x - 1$. With $x_0 = 1$, you will find the iterates jump about near 0. In fact there are two complex roots quite close to 0; there is also a real root around 8.9.

**4.4** *A volume problem*
Consider a solid hemisphere, of radius $r$, and a plane parallel to the base of the hemisphere, at height $h$ above the base (Fig. 13). The volume below the plane is then $\pi r^2 h - \frac{1}{3}\pi h^3$. (If you know how to find volumes by integration, then you can check this.) What must $h$ be in order to cut the hemisphere into two parts of *equal* volume? This requires

$$\pi r^2 h - \tfrac{1}{3}\pi h^3 = \tfrac{1}{4}\,(\text{volume of sphere, radius } r)$$
$$= \tfrac{1}{3}\pi r^3.$$

Hence $x = h/r$ satisfies $x^3 - 3x + 1 = 0$. Of course, $0 < x < 1$. In fact, this equation has three real roots; find approximate values for all of them by the Newton–Raphson method.

**4.5** *More volume problems*
There are many variants of Section 4.4. For example, if a sphere of radius $r$ stands on a horizontal surface then a horizontal plane at a height $h$ above the surface cuts off a volume $\pi h^2(r - \frac{1}{3}h)$ below itself. (Deduce this from Section 4.4.) Here $0 \leqslant h \leqslant 2r$, of course. What should $h$ be so that, for example, this volume is $\frac{3}{4}$ that of the sphere? This gives $x = h/r$ satisfying $x^3 - 3x^2 + 3 = 0$. Again this equation has three real roots; the answer to the problem is the one satisfying $1 < x < 2$. Find it by means of Program 3.1.

This also answers the question of the depth $h$ to which a spherical