$$5x^4 - 10x^2 + 2x + 1 = 0$$

in the given intervals:

(a) $-\infty < x < -1$
(b) $-1 < x < 0$
(c) $0 < x < 1$
(d) $1 < x < \infty$.

11. Find the number of positive roots and the number of negative roots of the equation

$$x^4 + x^3 - 2x - 3 = 0.$$

5.9.4 Pointers to Further Study

→ Theory of equations

5.10 Convex Functions

We first give a geometric definition of convex function based on the graph of the function, viewed as a curve. The line segment joining two points on a curve is called a chord, this being the standard usage in the case of a circle.

Definition Let A be an interval. A function $f : A \to \mathbb{R}$ is said to be *strictly convex*, if, for each pair of points a and b in the interval A, with $a < b$, the graph of f for $a < x < b$ lies strictly below the chord joining $(a, f(a))$ and $(b, f(b))$.

Plain convexity is a slightly, but significantly, weaker notion.

Definition Let A be an interval. A function $f : A \to \mathbb{R}$ is said to be *convex*, if, for each pair of points a and b in the interval A, with $a < b$, the graph of f between a and b does not go above the chord joining $(a, f(a))$ and $(b, f(b))$.

Our focus is entirely on strict convexity.[2] At the level of single-variable calculus it is strict convexity that has all the interesting applications. In some calculus texts a strictly convex function is called concave-up, a term that describes it admirably. Its uses explored here (some of them in the exercises) include some interesting deductions about solutions of equations, minimisation problems, the Legendre transform, inflection points and (in the next section) a sharp form of Jensen's inequality. Last

[2] This focus produces a tiresome need to repeat the words "strict" and "strictly". An alternative would have been to use the term "convex" instead of "strictly convex" and in the few places where convexity of the not necessarily strict kind is mentioned, to use "weakly convex". There is a precedent in some of the sources and it is consistent with the rule that the more useful version should have the simpler name. But it is not consistent with multivariate calculus where the greater usefulness of strict convexity compared to convexity is not so apparent.

but not least, an understanding about where a function is strictly convex and where strictly concave is a great aid to sketching its graph, still a useful mathematical skill.

Now we translate strict convexity into algebra. One way to write the equation of the chord is to "proceed from the point $(a, f(a))$" thus

$$y = \left(\frac{f(b) - f(a)}{b - a}\right)(x - a) + f(a).$$

Using this we can write the condition that f is strictly convex as follows. For all a, b and x in A such that $a < x < b$ we require

$$f(x) < \left(\frac{f(b) - f(a)}{b - a}\right)(x - a) + f(a),$$

or equivalently

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a}. \tag{5.3}$$

This inequality asserts that the slope of the chord is an increasing function of its right endpoint (just think of b as variable).

The inequality (5.3) is algebraically equivalent to each of two others; like it they each compare the slope of two chords. They are

$$\frac{f(b) - f(a)}{b - a} < \frac{f(b) - f(x)}{b - x}, \tag{5.4}$$

which asserts that the slope of the chord is an increasing function of its left endpoint, and

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(x)}{b - x}. \tag{5.5}$$

It is a nice exercise for the reader to show that all three inequalities are algebraically equivalent. Any one of them implies the other two. Geometrically this is obvious, as the three quantities being compared are the slopes of three chords forming the sides of a triangle whose vertices are the points $(a, f(a))$, $(x, f(x))$ and $(b, f(b))$ on the curve $y = f(x)$. A picture makes this rather obvious.

There is even a fourth version of the same inequality, also easy to obtain, that rather obviously expresses the claim that the graph is below the chord, namely

$$f(x) < \left(\frac{b - x}{b - a}\right) f(a) + \left(\frac{x - a}{b - a}\right) f(b). \tag{5.6}$$

Exercise Prove that the inequalities (5.3)–(5.6) are algebraically equivalent.

Putting this together we can set out a rather wordy necessary and sufficient condition for strict convexity of the function f; that for every three points a, x and b in the interval of definition, such that $a < x < b$, at least one of the above four inequalities is verified (and if one is true then all are true).

If, however, f is differentiable there is a much simpler criterion.

Proposition 5.18 *A differentiable function f is strictly convex if and only if f' is strictly increasing.*

Proof Suppose that f is strictly convex and differentiable. Let $a < b$. It follows from the inequalities that the quotient $(f(x) - f(a))/(x - a)$ is a strictly increasing function of x for $x > a$, and the quotient $(f(b) - f(x))/(b - x)$ is a strictly increasing function of x for $x < b$. Hence

$$f'(a) = \lim_{x \to a+} \frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a} < \lim_{x \to b-} \frac{f(b) - f(x)}{b - x} = f'(b)$$

giving $f'(a) < f'(b)$.

Conversely suppose that f' is strictly increasing. Let a, x, b be in the interval of definition of f and suppose that $a < x < b$. By the mean value theorem there are points y between a and x, and z between x and b, such that

$$\frac{f(x) - f(a)}{x - a} = f'(y)$$

and

$$\frac{f(b) - f(x)}{b - x} = f'(z).$$

But $f'(y) < f'(z)$ so we find

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(x)}{b - x}.$$

This is inequality (5.5) and shows that f is strictly convex. □

As an immediate consequence we have the most useful test for strict convexity; it is based on calculus rather than geometry, but requires second derivatives.

Proposition 5.19 *A sufficient condition for a twice differentiable function f to be strictly convex is that $f''(x) > 0$ for all x in the interval of definition.*

5.10.1 Tangent Lines and Convexity

Another useful conclusion, and a fifth necessary and sufficient, purely geometric condition for strict convexity, but based on the assumption that the function is differentiable, is the following.

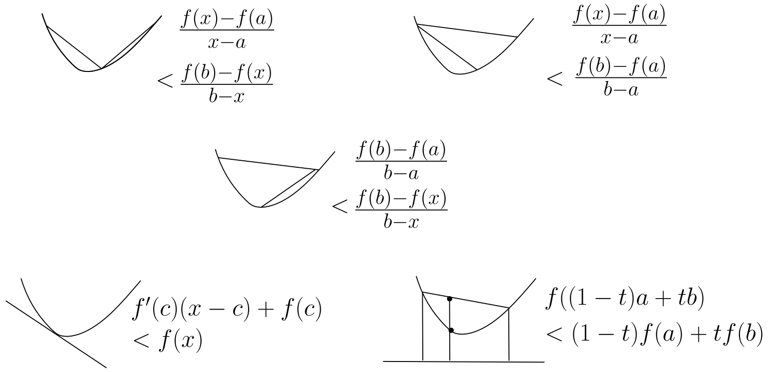

Fig. 5.6 Five views of strict convexity

Proposition 5.20 *Let f be differentiable in the open interval A. A necessary and sufficient condition for f to be strictly convex is that for every c in A, the tangent line to the curve $y = f(x)$ at the point $(c, f(c))$ lies wholly below the curve itself, except that they both contain the point $(c, f(c))$.*

Proof Suppose that f is strictly convex. We know that $(f(c) - f(x))/(c - x)$ is a strictly increasing function of x for $x < c$; and that $(f(x) - f(c))/(x - c)$ is a strictly increasing function of x for $x > c$. Hence if $x < c$ we find

$$\frac{f(c) - f(x)}{c - x} < \lim_{t \to c-} \frac{f(c) - f(t)}{c - t} = f'(c)$$

which implies

$$f(x) > f(c) + f'(c)(x - c)$$

and if $c < x$ we find

$$f'(c) = \lim_{t \to c+} \frac{f(t) - f(c)}{t - c} < \frac{f(x) - f(c)}{x - c}$$

which implies

$$f(x) > f(c) + f'(c)(x - c).$$

This shows that the condition is necessary.

The reader is invited to finish the proof by showing that the condition is sufficient for strict convexity given that f is differentiable. □

The five geometrical conditions for strict convexity are illustrated in Fig. 5.6.

5.10.2 Inflection Points

A function f such that $-f$ is strictly convex is called strictly concave (in some calculus texts it is called concave-down). Let f be differentiable in the interval A. A point $(a, f(a))$ on the curve $y = f(x)$ is called an *inflection point* of the curve if there exists $h > 0$, such that f is strictly convex [respectively, strictly concave] in the interval $]a - h, a[$, and strictly concave [respectively, strictly convex] in the interval $]a, a + h[$.

In other words the function switches from strictly convex to strictly concave, or from strictly concave to strictly convex, at the point a. We say loosely that f has an inflection point at a. Inflection points are illustrated in Fig. 5.7.

For some reason the notion of inflection point is only applied to differentiable functions; there has to be a tangent. Properly an inflection point is a property possessed by a plane curve and not just a graph; it is a point where the curvature changes sign. The concept of curvature really belongs to the study of the differential geometry of plane curves.

A necessary condition for an inflection point at a is that f' has either a local strict maximum or a local strict minimum at a. This is not sufficient. Again if f is twice differentiable it is necessary that $f''(a) = 0$, but still not sufficient. We have to force f'' to change sign, to be strictly positive on one side of a and strictly negative on the other.

A problem left to the exercises is to find a sufficient condition that f has an inflection point at a that builds on higher derivatives of f at a alone.

We often want to sketch the graph of a given function. Nowadays there are many good software packages that do this. A good sketch prepared without the help of a computer should show roughly where the function is strictly convex and where strictly concave. This means having some idea of where f'' is positive, where negative and where the inflection points are that separate these regions.

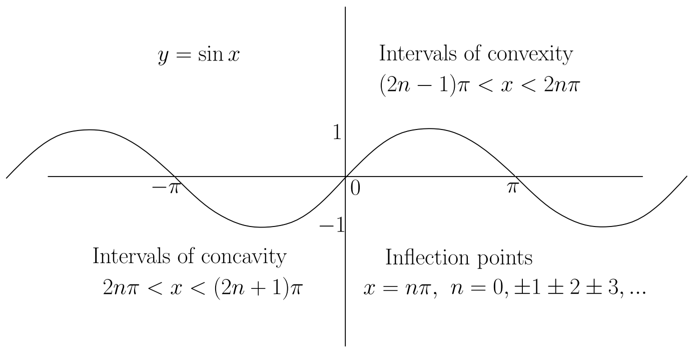

Fig. 5.7 Inflection points of $y = \sin x$

5.10.3 Exercises

1. Let f be a strictly convex function defined in an open interval A and let c be a point in A. Show that the limits

$$\lim_{t\to c-}\frac{f(c)-f(t)}{c-t} \quad \text{and} \quad \lim_{t\to c+}\frac{f(t)-f(c)}{t-c}$$

both exist and that the first is less than or equal to the second. These limits are the left and right derivatives, $D_l f(c)$ and $D_r f(c)$. Give an example to show that they do not have to be equal.

2. Show that a strictly convex function, defined in an interval A, is continuous if A is open, but that continuity may fail if A is not open.
Hint. One way is to use the previous exercise.

3. The function in Proposition 5.20 was assumed to be differentiable. Without assuming differentiability it is possible to say something similar, and obtain a sixth necessary and sufficient, purely geometric condition for strict convexity. Prove the following:

A function f, defined in an open interval A, is strictly convex if and only if it satisfies the following condition: for every c in A there exists a straight line through the point $(c, f(c))$ that lies wholly below the graph of f, except that the line and graph both contain the point $(c, f(c))$.

4. Let f be a convex function and suppose that there exist points $a < x < b$, such that the point $(x, f(x))$ lies on the chord joining $(a, f(a))$ and $(b, f(b))$. Show that the whole of the chord lies on the graph of f. So the graph of a non-strictly convex function differs from that of a strictly convex one by including some straight line segments.
Hint. Consider how the inequalities (5.3)–(5.6) should be modified for a function that is convex but not necessarily strictly convex.

5. Let f be a strictly convex function on the interval $[0, \infty[$ and suppose that $f(0) = 0$. Show that f satisfies

$$f(a+b) > f(a) + f(b)$$

for all positive a and b.

6. Show that if f is a strictly convex function and a and b are constants, then the function $f(x) + ax + b$ is also strictly convex.

7. Suppose that a function f satisfies

$$f\left(\frac{a+b}{2}\right) < \frac{1}{2}f(a) + \frac{1}{2}f(b)$$

for all a and b in its interval of definition.

(a) Show that

$$f\big(ta + (1-t)b\big) < tf(a) + (1-t)f(b)$$

for all a and b and for all *dyadic fractions* t in the interval $0 < t < 1$; that is, for all t of the form $t = k/2^n$ where n is a positive integer and k is an integer in the range $1 \le k \le 2^n - 1$.

(b) Show that if f is continuous then f is strictly convex.
Hint. For item (a) use induction with respect to n. To get started figure out how to handle the case $t = \frac{1}{4}$. For item (b) use the fact that every real number in the interval $[0, 1]$ can be approximated arbitrarily closely by dyadic fractions, as is shown by the binary representation of real numbers, analogous to the decimal representation but using 2 as a base instead of 10. You will need to figure out why the inequality remains strict when t is an arbitrary real number in the interval $0 < t < 1$.

8. Show that a twice differentiable function is convex (of the not necessarily strict kind) if and only if its second derivative is non-negative.
Note. This is a case where convexity is simpler than strict convexity. The counterpart of Proposition 5.19 is a necessary and sufficient condition for convexity, whereas Proposition 5.19 is only a sufficient condition for strict convexity.
9. Let f be a strictly convex function. Show that a straight line intersects the graph of f in at most two points. In other words, given constants a and b, the equation $f(x) = ax + b$ has at most two roots.
10. Let f be a strictly convex function defined in an interval A.

(a) Show that if f attains a minimum it does so at a unique point.
(b) Suppose that there exist distinct points a and b in A such that $f(a) = f(b)$. Show that f attains a minimum (which by (a) occurs at a unique point).
(c) Let c be an interior point of A (that is, c is not an endpoint). Show that f attains a minimum at c if and only if $D_l f(c) \le 0$ and $D_r f(c) \ge 0$ (see Exercise 1).
(d) Suppose that c is an interior point of A and that f is differentiable at c. Show that f attains a minimum at c if and only if $f'(c) = 0$.

11. For the purposes of this exercise we shall call a line that cuts a curve $y = f(x)$ a secant line. A secant line meets the curve and crosses it; it contains points (x_1, y_1) and (x_2, y_2), such that $y_1 < f(x_1)$ and $y_2 > f(x_2)$. Note that this is slightly different from the common usage, which requires a secant to meet the curve in two points, an assumption not made here.
Let f be strictly convex in the whole real line. Show that a secant line that is parallel to some chord of the curve $y = f(x)$ cuts the curve in two points.
12. Let f be strictly convex and defined in the whole real line. Suppose that f attains a minimum. Prove that $\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(x) = \infty$.
13. Let f be defined in an open interval A and let c be a point in A. Show that the following is sufficient for f to have an inflection point at c: the derivatives

$f^{(j)}(c)$ exist up to $j = m$, $f^{(j)}(c) = 0$ for $j = 2, ..., m-1$, $f^{(m)}(c) \neq 0$ and m is odd.

In the following series of exercises we study the Legendre transform. This is an important construction associated with convex functions that has many applications, both theoretical and practical.

14. Let f be strictly convex and differentiable in the open interval A. Let $c = \inf f'$ and $d = \sup f'$. For each p in the interval $B :=]c, d[$ let g_p be the function $g_p(x) = px - f(x)$.
    (a) Show that g_p attains a maximum value at a unique point x_p in A.
    (b) For each p in B we let
    $$f_*(p) = px_p - f(x_p).$$
    The function f_* is called the Legendre transform of f. Now suppose that f is twice differentiable and that $f'' > 0$. Show that
    $$(f_*)' = (f')^{-1}$$
    and deduce that f_* is strictly convex.
    (c) What if the second derivative does not exist? Can you prove the formula in item (b) from first principles, that is, by arguing from the difference quotient?
    (d) Show that $f_{**} = f$. Algebraically, the operation of passing from f to f_* is an *involution*. The same operation applied to f_* brings one back to f.
15. Prove Young's inequality. Given that f is strictly convex and differentiable, then
    $$px \leq f_*(p) + f(x)$$
    for all x in A and p in B (where A and B are the domains of f and f_* respectively).
16. Show that the power function x^a, with $a > 1$, is strictly convex in its interval of definition $0 < x < \infty$. It therefore has a Legendre transform. Obtain nice results by computing the Legendre transform of x^a/a and writing down the result of Young's inequality.
17. Try the previous exercise for the function e^x. You will need some knowledge of the exponential function and natural logarithm.
18. Let f be a strictly convex function defined in an open interval A. Let B be the set of all real numbers p, such that the graph $y = f(x)$ has a chord with slope p.
    (a) Show that B is an interval.
    (b) Show that for each p in B the function $px - f(x)$ attains a maximum at a unique point x_p in A.

This shows how the Legendre transform can be defined for strictly convex functions that are not everywhere differentiable. We set $f_*(p) = px_p - f(x_p)$ for each p in B. However, f_*, though convex, may fail to be strictly convex.

(c) Let f be defined by

$$f(x) = \begin{cases} (x-1)^2 \text{ if } x < 0 \\ (x+1)^2 \text{ if } x \geq 0. \end{cases}$$

Show that f is strictly convex and compute f_*. Show that the latter is convex but not strictly convex.

5.11 (◇) Jensen's Inequality

Jensen claimed that his inequality implied almost all known inequalities as special cases.[3] If this was only partially true it would make it a remarkable object of study. Actually Jensen's inequality is a natural enough extension of the fourth inequality characterising strictly convex functions, inequality (5.6).

Proposition 5.21 *Let f be a strictly convex function with domain A, let x_j, $(j = 1, 2, \ldots, n)$, be points in A, and let t_j, $(j = 1, 2, \ldots, n)$, be positive numbers such that $\sum_{j=1}^{n} t_j = 1$. Then*

$$f\left(\sum_{j=1}^{n} t_j x_j\right) \leq \sum_{j=1}^{n} t_j f(x_j).$$

Equality holds if and only if the numbers x_j are all equal.

Proof We set $c = \sum_{j=1}^{n} t_j x_j$. Because the numbers t_j are positive and sum to 1, it follows that c belongs to the interval A. By the result of Sect. 5.10, Exercise 3, there exists a line through $(c, f(c))$, that lies wholly below the graph $y = f(x)$, except that both the line and the graph contain the point $(c, f(c))$. Let the line have the equation $y = f(c) + m(x - c)$. Then for all $x \neq c$ we have

$$f(c) + m(x - c) < f(x)$$

whilst for $x = c$ we have equality. We now find

$$\sum_{j=1}^{n} t_j \big(f(c) + m(x_j - c)\big) \leq \sum_{j=1}^{n} t_j f(x_j)$$

[3]This is stated in the book "A Course of Analysis" by E. G. Phillips, originally published in 1930. I don't know what the author's source was; maybe he knew Jensen.