2. Find the stationary values of the following (check whether they are relative maxima or minima or inflection points), assuming the domain to be the interval $[0, \infty)$:
   (*a*) $y = x^3 - 3x + 5$
   (*b*) $y = \frac{1}{3}x^3 - x^2 + x + 10$
   (*c*) $y = -x^3 + 4.5x^2 - 6x + 6$
3. Show that the function $y = x + 1/x$ (with $x \neq 0$) has two relative extrema, one a maximum and the other a minimum. Is the "minimum" larger or smaller than the "maximum"? How is this paradoxical result possible?
4. Let $T = \phi(x)$ be a *total* function (e.g., total product or total cost):
   (*a*) Write out the expressions for the *marginal* function M and the *average* function A.
   (*b*) Show that, when A reaches a relative extremum, M and A must have the same value.
   (*c*) What general principle does this suggest for the drawing of a marginal curve and an average curve in the same diagram?
   (*d*) What can you conclude about the elasticity of the total function T at the point where A reaches an extreme value?

9.3 Second and Higher Derivatives

Hitherto we have considered only the first derivative $f'(x)$ of a function $y = f(x)$; now let us introduce the concept of *second derivative* (short for *second-order derivative*), and derivatives of even higher orders. These will enable us to develop alternative criteria for locating the relative extrema of a function.

Derivative of a Derivative

Since the first derivative $f'(x)$ is itself a function of x, it, too, should be differentiable with respect to x, provided that it is continuous and smooth. The result of this differentiation, known as the second derivative of the function f, is denoted by

$f''(x)$ where the double prime indicates that $f(x)$ has been differentiated with respect to x twice, and where the expression (x) following the double prime suggests that the second derivative is again a function of x

or

$\frac{d^2y}{dx^2}$ where the notation stems from the consideration that the second derivative means, in fact, $\frac{d}{dx}\left(\frac{dy}{dx}\right)$; hence, the d^2 (read: "d-two") in the numerator and dx^2 (read: "dx squared") in the denominator of this symbol.

If the second derivative $f''(x)$ exists for all x values in the domain, the function $f(x)$ is said to be *twice differentiable;* if, in addition, $f''(x)$ is continuous, the function $f(x)$ is said to be *twice continuously differentiable.* Just as the notation $f \in C^{(1)}$ or $f \in C'$ is often used to indicate that the function f is continuously differentiable, an analogous notation

$$f \in C^{(2)} \qquad \text{or} \qquad f \in C''$$

can be used to signify that f is twice continuously differentiable.

As a function of x the second derivative can be differentiated with respect to x again to produce a *third* derivative, which in turn can be the source of a *fourth* derivative, and so on ad infinitum, as long as the differentiability condition is met. These higher-order derivatives are symbolized along the same line as the second derivative:

$$f'''(x), f^{(4)}(x), \dots, f^{(n)}(x) \qquad \text{[with superscripts enclosed in ( )]}$$

or

$$\frac{d^3y}{dx^3}, \frac{d^4y}{dx^4}, \dots, \frac{d^ny}{dx^n}$$

The last of these can also be written as $\frac{d^n}{dx^n}y$, where the $\frac{d^n}{dx^n}$ part serves as an operator symbol instructing us to take the nth derivative of (some function) with respect to x.

Almost all the *specific* functions we shall be working with possess continuous derivatives up to any order we desire; i.e., they are continuously differentiable any number of times. Whenever a *general* function is used, such as $f(x)$, we always assume that it has derivatives up to any order we need.

Example 1 Find the first through the fifth derivatives of the function

$$y = f(x) = 4x^4 - x^3 + 17x^2 + 3x - 1$$

The desired derivatives are as follows:

$$f'(x) = 16x^3 - 3x^2 + 34x + 3$$
$$f''(x) = 48x^2 - 6x + 34$$
$$f'''(x) = 96x - 6$$
$$f^{(4)}(x) = 96$$
$$f^{(5)}(x) = 0$$

In this particular (polynomial) example, we note that each successive derivative function emerges as a lower-order polynomial—from cubic to quadratic, to linear, to constant. We note also that the fifth derivative, being the derivative of a constant, is equal to zero for all values of x; we could therefore have written it as $f^{(5)}(x) \equiv 0$ as well. The equation $f^{(5)}(x) = 0$ should be carefully distinguished from the equation $f^{(5)}(x_0) = 0$ (zero at x_0 only). Also, understand that the statement $f^{(5)}(x) \equiv 0$ does not mean that the fifth derivative does not exist; it indeed exists, and has the value zero.

Example 2 Find the first four derivatives of the rational function

$$y = g(x) = \frac{x}{1+x} \qquad (x \neq -1)$$

These derivatives can be found either by use of the *quotient rule*, or, after rewriting the function as $y = x(1+x)^{-1}$, by the product rule:

$$\left.\begin{aligned} g'(x) &= (1+x)^{-2} \\ g''(x) &= -2(1+x)^{-3} \\ g'''(x) &= 6(1+x)^{-4} \\ g^{(4)}(x) &= -24(1+x)^{-5} \end{aligned}\right\} \qquad (x \neq -1)$$

In this case, repeated derivation evidently does not tend to simplify the subsequent derivative expressions.

Note that, like the primitive function $g(x)$, all the successive derivatives obtained are themselves functions of x. Given specific values of x, however, these derivative functions will then take specific values. When $x = 2$, for instance, the second derivative in Example 2 can be evaluated as

$$g''(2) = -2(3)^{-3} = \frac{-2}{27}$$

and similarly for other values of x. It is of the utmost importance to realize that to evaluate this second derivative $g''(x)$ at $x = 2$, as we did, we must first obtain $g''(x)$ from $g'(x)$ and then substitute $x = 2$ into the equation for $g''(x)$. It is *incorrect* to substitute $x = 2$ into $g(x)$ or $g'(x)$ *prior* to the differentiation process leading to $g''(x)$.

Interpretation of the Second Derivative

The derivative function $f'(x)$ measures the rate of change of the function f. By the same token, the second-derivative function f'' is the measure of the rate of change of the first derivative f'; in other words, the second derivative measures the *rate of change* of the *rate of change* of the original function f. To put it differently, with a given infinitesimal increase in the independent variable x from a point $x = x_0$,

$$\left.\begin{array}{l} f'(x_0) > 0 \\ f'(x_0) < 0 \end{array}\right\} \text{means that the } \textit{value of the function} \text{ tends to } \left\{\begin{array}{l} \text{increase} \\ \text{decrease} \end{array}\right.$$

whereas, with regard to the second derivative,

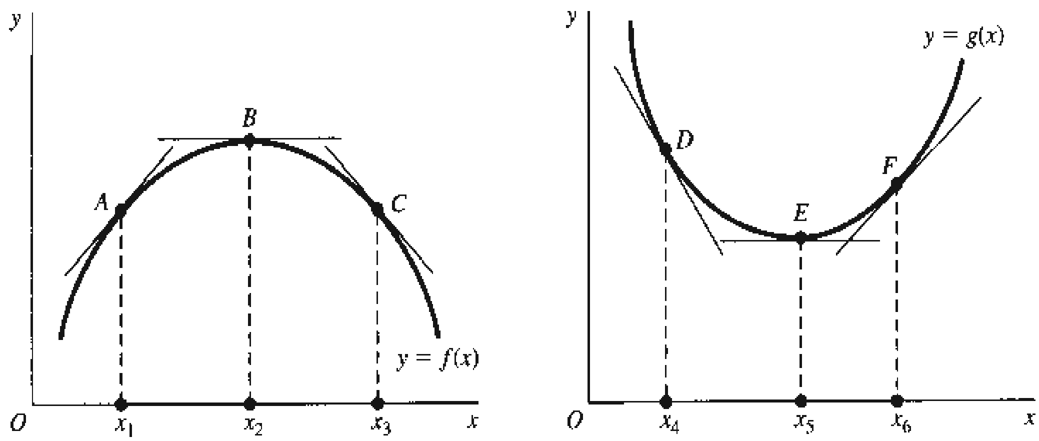
$$\left.\begin{array}{l} f''(x_0) > 0 \\ f''(x_0) < 0 \end{array}\right\} \text{means that the } \textit{slope of the curve} \text{ tends to } \left\{\begin{array}{l} \text{increase} \\ \text{decrease} \end{array}\right.$$

Thus a positive first derivative coupled with a positive second derivative at $x = x_0$ implies that the slope of the curve at that point is *positive and increasing*. In other words, the value of the function is increasing at an increasing rate. Likewise, a positive first derivative with a negative second derivative indicates that the slope of the curve is *positive but decreasing*—the value of the function is increasing at a decreasing rate. The case of a negative first derivative can be interpreted analogously, but a warning is in order in this case: When $f'(x_0) < 0$ and $f''(x_0) > 0$, the slope of the curve is *negative and increasing*, but this does *not* mean that the slope is changing, say, from (-10) to (-11); on the contrary, the change should be from (-11), a smaller number, to (-10), a larger number. In other words, the negative slope must tend to be *less* steep as x increases. Lastly, when $f'(x_0) < 0$ and $f''(x_0) < 0$, the slope of the curve must be *negative and decreasing*. This refers to a negative slope that tends to become *steeper* as x increases.

All of this can be further clarified with a graphical explanation. Figure 9.5*a* illustrates a function with $f''(x) < 0$ throughout. Since the slope must steadily decrease as x increases on the graph, we will, when we move from left to right, pass through a point A with a positive slope, then a point B with zero slope, and then a point C with a negative slope. It may happen, of course, that a function with $f''(x) < 0$ is characterized by $f'(x) > 0$ everywhere, and thus plots only as the rising portion of an inverse U-shaped curve, or, with $f'(x) < 0$ everywhere, plots only as the declining portion of that curve.

The opposite case of a function with $f''(x) > 0$ throughout is illustrated in Fig. 9.5*b*. Here, as we pass through points D to E to F, the slope steadily increases and changes from

FIGURE 9.5



(*a*) (*b*)

negative to zero to positive. Again, we add that a function characterized by $f''(x) > 0$ throughout may, depending on the first-derivative specification, plot only as the declining or the rising portion of a U-shaped curve.

From Fig. 9.5, it is evident that the second derivative $f''(x)$ relates to the *curvature* of a graph; it determines how the curve tends to bend itself. To describe the two types of differing curvatures discussed, we refer to the one in Fig. 9.5*a* as *strictly concave,* and the one in Fig. 9.5*b* as *strictly convex*. And, understandably, a function whose graph is strictly concave (strictly convex) is called a *strictly concave* (*strictly convex*) *function*. The precise geometric characterization of a strictly concave function is as follows. If we pick any pair of points *M* and *N* on its curve and join them by a straight line, the line segment *MN* must lie entirely *below* the curve, except at points *M* and *N*. The characterization of a strictly convex function can be obtained by substituting the word *above* for the word *below* in the last statement. Try this out in Fig. 9.5. If the characterizing condition is relaxed somewhat, so that the line segment *MN* is allowed to lie *either* below the curve, *or* along (coinciding with) the curve, then we will be describing instead a *concave function,* without the adverb *strictly*. Similarly, if the line segment *MN either* lies above, *or* lies along the curve, then the function is *convex,* again without the adverb *strictly*. Note that, since the line segment *MN* may coincide with a (nonstrictly) concave or convex curve, the latter may very well contain a linear segment. In contrast, a *strictly* concave or convex curve can never contain a linear segment anywhere. It follows that while a strictly concave (convex) function is automatically a concave (convex) function, the converse is not true.†

From our earlier discussion of the second derivative, we may now infer that if the second derivative $f''(x)$ is negative for *all* x, then the primitive function $f(x)$ must be a strictly concave function. Similarly, $f(x)$ must be strictly convex, if $f''(x)$ is positive for *all* x. Despite this, it is *not* valid to reverse this inference and say that, if $f(x)$ is strictly concave (strictly convex), then $f''(x)$ must be negative (positive) for all x. This is because, in certain exceptional cases, the second derivative may have a *zero* value at a stationary point on such a curve. An example of this can be found in the function $y = f(x) = x^4$, which plots as a strictly convex curve, but whose derivatives

$$f'(x) = 4x^3 \qquad f''(x) = 12x^2$$

† We shall discuss these concepts further in Sec. 11.5.

indicate that, at the stationary point where $x = 0$, the value of the second derivative is $f''(0) = 0$. Note, however, that at any other point, with $x \neq 0$, the second derivative of this function does have the (expected) positive sign. Aside from the possibility of a zero value at a stationary point, therefore, the second derivative of a strictly concave or convex function may be expected in general to adhere to a single algebraic sign.

For other types of function, the second derivative may take both positive and negative values, depending on the value of x. In Fig. 9.3*a* and *b*, for instance, both $f(x)$ and $g(x)$ undergo a sign change in the second derivative at their respective inflection points J and K. According to Fig. 9.3*a*′, the slope of $f'(x)$—that is, the value of $f''(x)$—changes from negative to positive at $x = j$; the exact opposite occurs with the slope of $g'(x)$—that is, the value of $g''(x)$—on the basis of Fig. 9.3*b*′. Translated into curvature terms, this means that the graph of $f(x)$ turns from strictly concave to strictly convex at point J, whereas the graph of $g(x)$ has the reverse change at point K. Consequently, instead of characterizing an inflection point as a point where the first derivative reaches an extreme value, we may alternatively characterize it as a point where the function undergoes a change in curvature or a change in the sign of its second derivative.

An Application

The two curves in Fig. 9.5 exemplify the graphs of quadratic functions, which may be expressed generally in the form

$$y = ax^2 + bx + c \qquad (a \neq 0)$$

From our discussion of the second derivative, we can now derive a convenient way of determining whether a given quadratic function will have a strictly convex (U-shaped) or a strictly concave (inverse U-shaped) graph.

Since the second derivative of the quadratic function cited is $d^2y/dx^2 = 2a$, this derivative will always have the same algebraic sign as the coefficient a. Recalling that a positive second derivative implies a strictly convex curve, we can infer that a positive coefficient a in the preceding quadratic function gives rise to a U-shaped graph. In contrast, a negative coefficient a leads to a strictly concave curve, shaped like an inverted U.

As intimated at the end of Sec. 9.2, the relative extremum of this function will also prove to be its absolute extremum, because in a quadratic function there can be found only a single valley or peak, evident in a U or inverted U, respectively.

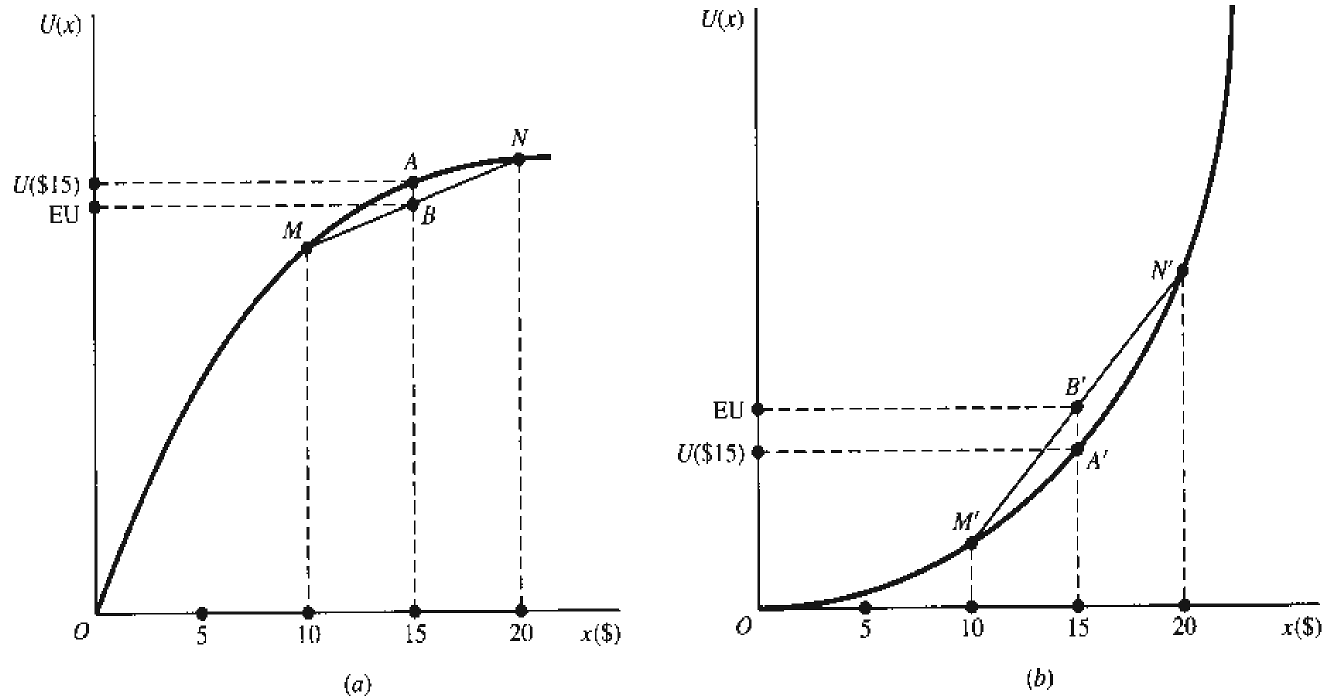
Attitudes toward Risk

The most common application of the concept of marginal utility is to the context of goods consumption. But in another useful application, we consider the marginal utility of *income*, or more to the point of the present discussion, the *payoff* to a betting game, and use this concept to distinguish between different individuals' attitudes toward risk.

Consider the game where, for a fixed sum of money paid in advance (the *cost* of the game), you can throw a die and collect \$10 if an odd number shows up, or \$20 if the number is even. In view of the equal probability of the two outcomes, the mathematically *expected value of payoff* is

$$\text{EV} = 0.5 \times \$10 + 0.5 \times \$20 = \$15$$

FIGURE 9.6



The game is deemed a *fair game,* or *fair bet,* if the cost of the game is exactly \$15. Despite its fairness, playing such a game still involves a risk, for even though the probability distribution of the two possible outcomes is known, the actual result of any individual play is not. Hence, people who are "risk-averse" would consistently decline to play such a game. On the other hand, there are "risk-loving" or "risk-preferring" people who would welcome fair games, or even games with odds set against them (i.e., with the cost of the game exceeding the expected value of payoff).

The explanation for such diverse attitudes toward risk is easily found in the differing utility functions people possess. Assume that a potential player has the strictly concave utility function $U = U(x)$ depicted in Fig. 9.6*a*, where x denotes the payoff, with $U(0) = 0$, $U'(x) > 0$ (positive marginal utility of income or payoff), and $U''(x) < 0$ (diminishing marginal utility) for all x. The economic decision facing this person involves the choice between two courses of action: First, by not playing the game, the person saves the \$15 cost of the game (= EV) and thus enjoys the utility level $U(\$15)$, measured by the height of point A on the curve. Second, by playing, the person has a .5 probability of receiving \$10 and thus enjoying $U(\$10)$ (see point M), plus a .5 probability of receiving \$20 and thus enjoying $U(\$20)$ (see point N). The *expected utility from playing* is, therefore, equal to

$$\text{EU} = 0.5 \times U(\$10) + 0.5 \times U(\$20)$$

which, being the average of the height of M and that of N, is measured by the height of point B, the midpoint on the line segment MN. Since, by the defining property of a strictly concave utility function, line segment MN must lie below arc MN, point B must be lower than point A; that is, EU, the expected utility from playing, falls short of the utility of the cost of the game, and the game should be avoided. For this reason, a strictly concave utility function is associated with risk-averse behavior.

For a risk-loving person, the decision process is analogous, but the opposite choice will be made, because now the relevant utility function is a strictly convex one. In Fig. 9.6*b*,

$U(\$15)$, the utility of keeping the \$15 by not playing the game, is shown by point A' on the curve, and EU, the expected utility from playing, is given by B', the midpoint on the line segment $M'N'$. But this time line segment $M'N'$ lies above arc $M'N'$, and point B' is above point A'. Thus there definitely is a positive incentive to play the game. In contrast to the situation in Fig. 9.6*a*, we can thus associate a strictly convex utility function with risk-loving behavior.

EXERCISE 9.3

1. Find the second and third derivatives of the following functions:
   (*a*) $ax^2 + bx + c$ (*c*) $\frac{3x}{1-x}$ $(x \neq 1)$
   (*b*) $7x^4 - 3x - 4$ (*d*) $\frac{1+x}{1-x}$ $(x \neq 1)$
2. Which of the following quadratic functions are strictly convex?
   (*a*) $y = 9x^2 - 4x + 8$ (*c*) $u = 9 - 2x^2$
   (*b*) $w = -3x^2 + 39$ (*d*) $v = 8 - 5x + x^2$
3. Draw (*a*) a concave curve which is *not* strictly concave, and (*b*) a curve which qualifies simultaneously as a concave curve and a convex curve.
4. Given the function $y = a - \frac{b}{c+x}$ $(a, b, c > 0;\ x \geq 0)$, determine the general shape of its graph by examining (*a*) its first and second derivatives, (*b*) its vertical intercept, and (*c*) the limit of y as x tends to infinity. If this function is to be used as a consumption function, how should the parameters be restricted in order to make it economically sensible?
5. Draw the graph of a function $f(x)$ such that $f'(x) \equiv 0$, and the graph of a function $g(x)$ such that $g'(3) = 0$. Summarize in one sentence the essential difference between $f(x)$ and $g(x)$ in terms of the concept of stationary point.
6. A person who is neither risk-averse nor risk-loving (indifferent toward a fair game) is said to be "risk-neutral."
   (*a*) What kind of utility function would you use to characterize such a person?
   (*b*) Using the die-throwing game detailed in the text, describe the relationship between $U(\$15)$ and EU for the risk-neutral person.

9.4 Second-Derivative Test

Returning to the pair of extreme points B and E in Fig. 9.5 and remembering the newly established relationship between the second derivative and the curvature of a curve, we should be able to see the validity of the following criterion for a relative extremum:

Second-derivative test for relative extremum If the value of the first derivative of a function f at $x = x_0$ is $f'(x_0) = 0$, then the value of the function at x_0, $f(x_0)$, will be

a. A relative *maximum* if the second-derivative value at x_0 is $f''(x_0) < 0$.
b. A relative *minimum* if the second-derivative value at x_0 is $f''(x_0) > 0$.

This test is in general more convenient to use than the first-derivative test, because it does not require us to check the derivative sign to both the left and the right of x_0. But it has the drawback that no unequivocal conclusion can be drawn in the event that $f''(x_0) = 0$. For then the stationary value $f(x_0)$ can be *either* a relative maximum, *or* a relative minimum, *or* even an inflectional value.† When the situation of $f''(x_0) = 0$ is encountered, we must either revert to the first-derivative test, or resort to another test, to be developed in Sec. 9.6, that involves the third or even higher derivatives. For most problems in economics, however, the second-derivative test would usually be adequate for determining a relative maximum or minimum.

Example 1 Find the relative extremum of the function

$$y = f(x) = 4x^2 - x$$

The first and second derivatives are

$$f'(x) = 8x - 1 \qquad \text{and} \qquad f''(x) = 8$$

Setting $f'(x)$ equal to zero and solving the resulting equation, we find the (only) critical value to be $x^* = \frac{1}{8}$, which yields the (only) stationary value $f\left(\frac{1}{8}\right) = -\frac{1}{16}$. Because the second derivative is positive (in this case it is indeed positive for any value of x), the extremum is established as a minimum. Further, since the given function plots as a U-shaped curve, the relative minimum is also the absolute minimum.

Example 2 Find the relative extrema of the function

$$y = g(x) = x^3 - 3x^2 + 2$$

The first two derivatives of this function are

$$g'(x) = 3x^2 - 6x \qquad \text{and} \qquad g''(x) = 6x - 6$$

Setting $g'(x)$ equal to zero and solving the resulting quadratic equation, $3x^2 - 6x = 0$, we obtain the critical values $x_1^* = 2$ and $x_2^* = 0$, which in turn yield the two stationary values:

$$g(2) = -2 \qquad [\text{a minimum because } g''(2) = 6 > 0]$$
$$g(0) = 2 \qquad [\text{a maximum because } g''(0) = -6 < 0]$$

Necessary versus Sufficient Conditions

As was the case with the first-derivative test, the zero-slope condition $f'(x) = 0$ plays the role of a *necessary* condition in the second-derivative test. Since this condition is based on the first-order derivative, it is often referred to as the *first-order condition*. Once we find the first-order condition satisfied at $x = x_0$, the negative (positive) sign of $f''(x_0)$ is *sufficient* to establish the stationary value in question as a relative maximum (minimum). These sufficient conditions, which are based on the second-order derivative, are often referred to as *second-order conditions*.

† To see that an inflection point is possible when $f''(x_0) = 0$, let us refer back to Fig. 9.3*a* and 9.3*a'*. Point J in the upper diagram is an inflection point, with $x = j$ as its critical value. Since the $f'(x)$ curve in the lower diagram attains a minimum at $x = j$, the slope of $f'(x)$ [i.e., $f''(x)$] must be zero at the critical value $x = j$. Thus point J illustrates an inflection point occurring when $f''(x_0) = 0$.

To see that a relative extremum is also consistent with $f''(x_0) = 0$, consider the function $y = x^4$. This function plots as a U-shaped curve and has a minimum, $y = 0$, attained at the critical value $x = 0$. Since the second derivative of this function is $f''(x) = 12x^2$, we again obtain a zero value for this derivative at the critical value $x = 0$. Thus this function illustrates a relative extremum occurring when $f''(x_0) = 0$.

TABLE 9.1 Conditions for a Relative Extremum: $y = f(x)$

Condition	Maximum	Minimum
First-order necessary	$f'(x) = 0$	$f'(x) = 0$
Second-order necessary†	$f''(x) \le 0$	$f''(x) \ge 0$
Second-order sufficient†	$f''(x) < 0$	$f''(x) > 0$

†Applicable only after the first-order necessary condition has been satisfied.

It bears repeating that the first-order condition is *necessary*, but *not sufficient*, for a relative maximum or minimum. (Remember inflection points?) In sharp contrast, the second-order condition that $f''(x)$ be negative (positive) at the critical value x_0 is *sufficient* for a relative maximum (minimum), but it is *not necessary*. [Remember the relative extremum that occurs when $f''(x_0) = 0$?] For this reason, one should carefully guard against the following line of argument: "Since the stationary value $f(x_0)$ is already known to be a minimum, we must have $f''(x_0) > 0$." The reasoning here is faulty because it incorrectly treats the positive sign of $f''(x_0)$ as a necessary condition for $f(x_0)$ to be a minimum.

This is not to say that second-order derivatives can never be used in stating *necessary* conditions for relative extrema. Indeed they can. But care must then be taken to allow for the fact that a relative maximum (minimum) can occur not only when $f''(x_0)$ is negative (positive), but also when $f''(x_0)$ is zero. Consequently, *second-order necessary conditions* must be couched in terms of weak inequalities: for a stationary value $f(x_0)$ to be a relative $\begin{Bmatrix} \text{maximum} \\ \text{minimum} \end{Bmatrix}$, it is necessary that $f''(x_0) \begin{Bmatrix} \le \\ \ge \end{Bmatrix} 0$.

The preceding discussion can be summed up in Table 9.1. All the equations and inequalities in the table are in the nature of conditions (requirements) to be met, rather than descriptive specifications of a given function. In particular, the equation $f'(x) = 0$ does not signify that function f has a zero slope everywhere; rather, it states the stipulation that only those values of x that satisfy this requirement can qualify as critical values.

Conditions for Profit Maximization

We shall now present an economic example of extreme-value problems, i.e., problems of optimization.

One of the first things that a student of economics learns is that, in order to maximize profit, a firm must equate marginal cost and marginal revenue. Let us show the mathematical derivation of this condition. To keep the analysis on a general level, we shall work with the total-revenue function $R = R(Q)$ and total-cost function $C = C(Q)$, both of which are functions of a single variable Q. From these it follows that a profit function (the objective function) may also be formulated in terms of Q (the choice variable):

$$\pi = \pi(Q) = R(Q) - C(Q) \tag{9.1}$$

To find the profit-maximizing output level, we must satisfy the first-order necessary condition for a maximum: $d\pi/dQ = 0$. Accordingly, let us differentiate (9.1) with respect to Q and set the resulting derivative equal to zero: The result is

$$\begin{aligned} \frac{d\pi}{dQ} &\equiv \pi'(Q) = R'(Q) - C'(Q) \\ &= 0 \qquad \text{iff} \qquad R'(Q) = C'(Q) \end{aligned} \tag{9.2}$$

Thus the *optimum* output (*equilibrium* output) Q^* must satisfy the equation $R'(Q^*) = C'(Q^*)$, or MR = MC. This condition constitutes the first-order condition for profit maximization.

However, the first-order condition may lead to a minimum rather than a maximum; thus we must check the second-order condition next. We can obtain the second derivative by differentiating the first derivative in (9.2) with respect to Q:

$$\frac{d^2\pi}{dQ^2} \equiv \pi''(Q) = R''(Q) - C''(Q)$$
$$\leq 0 \qquad \text{iff} \qquad R''(Q) \leq C''(Q)$$

This last inequality is the second-order necessary condition for maximization. If it is not met, then Q^* cannot possibly maximize profit; in fact, it minimizes profit. If $R''(Q^*) = C''(Q^*)$, then we are unable to reach a definite conclusion. The best scenario is to find $R''(Q^*) < C''(Q^*)$, which satisfies the second-order sufficient condition for a maximum. In that case, we can conclusively take Q^* to be a profit-maximizing output. Economically, this would mean that, if the rate of change of MR is less than the rate of change of MC at the output where MC = MR, then that output will maximize profit.

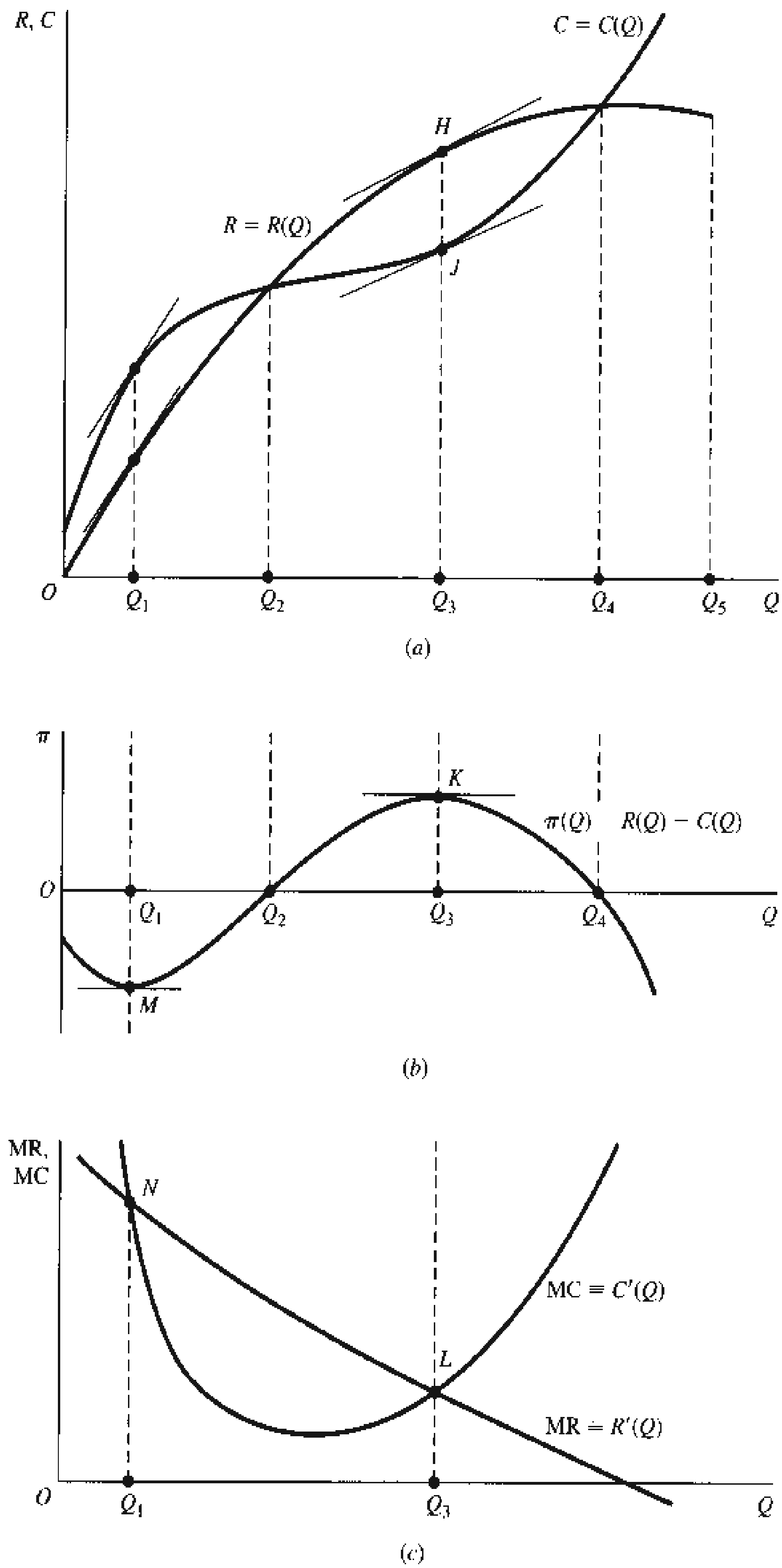
These conditions are illustrated in Fig. 9.7. In Fig. 9.7*a* we have drawn a total-revenue and a total-cost curve, which are seen to intersect twice, at output levels of Q_2 and Q_4. In the open interval (Q_2, Q_4), total revenue R exceeds total cost C, and thus π is positive. But in the intervals $[0, Q_2)$ and $(Q_4, Q_5]$, where Q_5 represents the upper limit of the firm's productive capacity, π is negative. This fact is reflected in Fig. 9.7*b*, where the profit curve—obtained by plotting the vertical distance between the R and C curves for each level of output—lies above the horizontal axis only in the interval (Q_2, Q_4).

When we set $d\pi/dQ = 0$, in line with the first-order condition, it is our intention to locate the peak point K on the profit curve, at output Q_3, where the slope of the curve is zero. However, the relative-minimum point M (output Q_1) will also offer itself as a candidate, because it, too, meets the zero-slope requirement. Below, we shall resort to the second-order condition to eliminate the "wrong" kind of extremum.

The first-order condition $d\pi/dQ = 0$ is equivalent to the condition $R'(Q) = C'(Q)$. In Fig. 9.7*a*, the output level Q_3 satisfies this, because the R and C curves do have the same slope at Q_3 (the tangent lines drawn to the two curves at H and J are parallel to each other). The same is true for output Q_1. Since the equality of the slopes of R and C means the equality of MR and MC, outputs Q_3 and Q_1 must obviously be where the MR and MC curves intersect, as illustrated in Fig. 9.7*c*.

How does the second-order condition enter into the picture? Let us first look at Fig. 9.7*b*. At point K, the second derivative of the π function will (barring the exceptional zero-value case) have a negative value, $\pi''(Q_3) < 0$, because the curve is inverse U-shaped around K; this means that Q_3 will maximize profit. At point M, on the other hand, we would expect that $\pi''(Q_1) > 0$; thus Q_1 provides a relative minimum for π instead. The second-order sufficient condition for a maximum can, of course, be stated alternatively as $R''(Q) < C''(Q)$, that is, that the slope of the MR curve be less than the slope of the MC curve. From Fig. 9.7*c*, it is immediately apparent that output Q_3 satisfies this condition, since the slope of MR is negative while that of MC is positive at point L. But output Q_1 violates this condition because both MC and MR have negative slopes, and that of MR is *numerically smaller* than that of MC at point N, which implies that $R''(Q_1)$ is *greater* than

FIGURE 9.7



$C''(Q_1)$ instead. In fact, therefore, output Q_1 also violates the second-order *necessary* condition for a relative maximum, but satisfies the second-order *sufficient* condition for a relative minimum.

Example 3 Let the $R(Q)$ and $C(Q)$ functions be

$$R(Q) = 1{,}200Q - 2Q^2$$
$$C(Q) = Q^3 - 61.25Q^2 + 1{,}528.5Q + 2{,}000$$

Then the profit function is

$$\pi(Q) = -Q^3 + 59.25Q^2 - 328.5Q - 2{,}000$$

where R, C, and π are all in dollar units and Q is in units of (say) tons per week. This profit function has two critical values, $Q = 3$ and $Q = 36.5$, because

$$\frac{d\pi}{dQ} = -3Q^2 + 118.5Q - 328.5 = 0 \qquad \text{when } Q = \begin{cases} 3 \\ 36.5 \end{cases}$$

But since the second derivative is

$$\frac{d^2\pi}{dQ^2} = -6Q + 118.5 \qquad \begin{cases} > 0 & \text{when } Q = 3 \\ < 0 & \text{when } Q = 36.5 \end{cases}$$

the profit-maximizing output is $Q^* = 36.5$ (tons per week). (The other output minimizes profit.) By substituting Q^* into the profit function, we can find the maximized profit to be $\pi^* = \pi(36.5) = 16{,}318.44$ (dollars per week).

As an alternative approach to the preceding, we can first find the MR and MC functions and then equate the two, i.e., find their intersection. Since

$$R'(Q) = 1{,}200 - 4Q$$
$$C'(Q) = 3Q^2 - 122.5Q + 1{,}528.5$$

equating the two functions will result in a quadratic equation identical with $d\pi/dQ = 0$ which has yielded the two critical values of Q cited previously.

Coefficients of a Cubic Total-Cost Function

In Example 3, a cubic function is used to represent the total-cost function. The traditional total-cost curve $C = C(Q)$, as illustrated in Fig. 9.7*a*, is supposed to contain two wiggles that form a concave segment (decreasing marginal cost) and a subsequent convex segment (increasing marginal cost). Since the graph of a cubic function always contains exactly two wiggles, as illustrated in Fig. 9.4, it should suit that role well. However, Fig. 9.4 immediately alerts us to a problem: the cubic function can possibly produce a downward-sloping segment in its graph, whereas the total-cost function, to make economic sense, should be upward-sloping everywhere (a larger output always entails a higher total cost). If we wish to use a cubic total-cost function such as

$$C = C(Q) = aQ^3 + bQ^2 + cQ + d \qquad \textbf{(9.3)}$$

therefore, it is essential to place appropriate restrictions on the parameters so as to prevent the C curve from ever bending downward.

An equivalent way of stating this requirement is that the MC function should be positive throughout, and this can be ensured only if the *absolute minimum* of the MC function turns out to be positive. Differentiating (9.3) with respect to Q, we obtain the MC function

$$MC = C'(Q) = 3aQ^2 + 2bQ + c \qquad \textbf{(9.4)}$$

which, because it is a quadratic, plots as a parabola as in Fig. 9.7*c*. In order for the MC curve to stay positive (above the horizontal axis) everywhere, it is necessary that the parabola be U-shaped (otherwise, with an inverse U, the curve is bound to extend itself into the second quadrant). Hence the coefficient of the Q^2 term in (9.4) has to be positive; i.e., we must impose the restriction $a > 0$. This restriction, however, is by no means sufficient, because the minimum value of a U-shaped MC curve—call it MC_{min} (a relative minimum which also happens to be an absolute minimum)—may still occur below the horizontal axis. Thus we must next find MC_{min} and ascertain the parameter restrictions that would make it positive.

According to our knowledge of relative extremum, the minimum of MC will occur where

$$\frac{d}{dQ}\text{MC} = 6aQ + 2b = 0$$

The output level that satisfies this first-order condition is

$$Q^* = \frac{-2b}{6a} = \frac{-b}{3a}$$

This minimizes (rather than maximizes) MC because the second derivative $d^2(\text{MC})/dQ^2 = 6a$ is assuredly positive in view of the restriction $a > 0$. The knowledge of Q^* now enables us to calculate MC_{min}, but we may first infer the sign of coefficient b from it. Inasmuch as negative output levels are ruled out, we see that b can never be positive (given $a > 0$). Moreover, since the law of diminishing returns is assumed to set in at a positive output level (that is, MC is assumed to have an initial declining segment), Q^* should be positive (rather than zero). Consequently, we must impose the restriction $b < 0$.

It is a simple matter now to substitute the MC-minimizing output Q^* into (9.4) to find that

$$\text{MC}_{\text{min}} = 3a\left(\frac{-b}{3a}\right)^2 + 2b\frac{-b}{3a} + c = \frac{3ac - b^2}{3a}$$

Thus, to guarantee the positivity of MC_{min}, we must impose the restriction† $b^2 < 3ac$. This last restriction, we may add, in effect also implies the restriction $c > 0$. (Why?)

The preceding discussion has involved the three parameters a, b, and c. What about the other parameter, d? The answer is that there is need for a restriction on d also, but that has nothing to do with the problem of keeping the MC positive. If we let $Q = 0$ in (9.3), we find

† This restriction may also be obtained by the method of *completing the square*. The MC function can be successively transformed as follows:

$$\begin{aligned}\text{MC} &= 3aQ^2 + 2bQ + c \\ &= \left(3aQ^2 + 2bQ + \frac{b^2}{3a}\right) - \frac{b^2}{3a} + c \\ &= \left(\sqrt{3a}Q + \sqrt{\frac{b^2}{3a}}\right)^2 + \frac{-b^2 + 3ac}{3a}\end{aligned}$$

Since the squared expression can possibly be zero, we must, in order to ensure the positivity of MC, require that $b^2 < 3ac$ on the knowledge that $a > 0$.

that $C(0) = d$. The role of d is thus to determine the vertical intercept of the C curve only, with no bearing on its slope. Since the economic meaning of d is the fixed cost of a firm, the appropriate restriction (in the short-run context) would be $d > 0$.

In sum, the coefficients of the total-cost function (9.3) should be restricted as follows (assuming the short-run context):

$$a, c, d > 0 \qquad b < 0 \qquad b^2 < 3ac \tag{9.5}$$

As you can readily verify, the $C(Q)$ function in Example 3 does satisfy (9.5).

Upward-Sloping Marginal-Revenue Curve

The marginal-revenue curve in Fig. 9.7*c* is shown to be downward-sloping throughout. This, of course, is how the MR curve is traditionally drawn for a firm under imperfect competition. However, the possibility of the MR curve being partially, or even wholly, upward-sloping can by no means be ruled out a priori.[†]

Given an average-revenue function $\text{AR} = f(Q)$, the marginal-revenue function can be expressed by

$$\text{MR} = f(Q) + Qf'(Q) \qquad \text{[from (7.7)]}$$

The slope of the MR curve can thus be ascertained from the derivative

$$\frac{d}{dQ}\text{MR} = f'(Q) + f'(Q) + Qf''(Q) = 2f'(Q) + Qf''(Q)$$

As long as the AR curve is downward-sloping (as it would be under imperfect competition), the $2f'(Q)$ term is assuredly negative. But the $Qf''(Q)$ term can be either negative, zero, or positive, depending on the sign of the second derivative of the AR function, i.e., depending on whether the AR curve is strictly concave, linear, or strictly convex. If the AR curve is strictly convex either in its entirety (as illustrated in Fig. 7.2) or along a specific segment, the possibility will exist that the (positive) $Qf''(Q)$ term may dominate the (negative) $2f'(Q)$ term, thereby causing the MR curve to be wholly or partially upward-sloping.

Example 4

Let the average-revenue function be

$$\text{AR} = f(Q) = 8{,}000 - 23Q + 1.1Q^2 - 0.018Q^3$$

As can be verified (see Exercise 9.4-7), this function gives rise to a downward-sloping AR curve, as is appropriate for a firm under imperfect competition. Since

$$\text{MR} = f(Q) + Qf'(Q) = 8{,}000 - 46Q + 3.3Q^2 - 0.072Q^3$$

it follows that the slope of MR is

$$\frac{d}{dQ}\text{MR} = -46 + 6.6Q - 0.216Q^2$$

Because this is a quadratic function and since the coefficient of Q^2 is negative, $d\text{MR}/dQ$ must plot as an inverse-U-shaped curve against Q, such as shown in Fig. 9.5*a*. If a segment of this curve happens to lie above the horizontal axis, the slope of MR will take positive values.

[†] This point is emphatically brought out in John P. Formby, Stephen Layson, and W. James Smith, "The Law of Demand, Positive Sloping Marginal Revenue, and Multiple Profit Equilibria," *Economic Inquiry*, April 1982, pp. 303–311.

Setting $dMR/dQ = 0$, and applying the quadratic formula, we find the two zeros of the quadratic function to be $Q_1 = 10.76$ and $Q_2 = 19.79$ (approximately). This means that, for values of Q in the open interval (Q_1, Q_2), the dMR/dQ curve does lie above the horizontal axis. Thus the marginal-revenue curve indeed is positively sloped for output levels between Q_1 and Q_2.

The presence of a positively sloped segment on the MR curve has interesting implications. Such an MR curve may produce more than one intersection with the MC curve satisfying the second-order sufficient condition for profit maximization. While all such intersections constitute local optima, however, only one of them is the global optimum that the firm is seeking.

EXERCISE 9.4

1. Find the relative maxima and minima of y by the second-derivative test:

   (*a*) $y = -2x^2 + 8x + 25$ (*c*) $y = \frac{1}{3}x^3 - 3x^2 + 5x + 3$

   (*b*) $y = x^3 + 6x^2 + 9$ (*d*) $y = \dfrac{2x}{1-2x} \quad \left(x \neq \dfrac{1}{2}\right)$

2. Mr. Greenthumb wishes to mark out a rectangular flower bed, using a wall of his house as one side of the rectangle. The other three sides are to be marked by wire netting, of which he has only 64 ft available. What are the length L and width W of the rectangle that would give him the largest possible planting area? How do you make sure that your answer gives the largest, not the smallest area?

3. A firm has the following total-cost and demand functions:

   $$C = \frac{1}{3}Q^3 - 7Q^2 + 111Q + 50$$

   $$Q = 100 - P$$

   (*a*) Does the total-cost function satisfy the coefficient restrictions of (9.5)?

   (*b*) Write out the total-revenue function R in terms of Q.

   (*c*) Formulate the total-profit function π in terms of Q.

   (*d*) Find the profit-maximizing level of output Q^*.

   (*e*) What is the maximum profit?

4. If coefficient b in (9.3) were to take a zero value, what would happen to the marginal-cost and total-cost curves?

5. A quadratic profit function $\pi(Q) = hQ^2 + jQ + k$ is to be used to reflect the following assumptions:

   (*a*) If nothing is produced, the profit will be negative (because of fixed costs).

   (*b*) The profit function is strictly concave.

   (*c*) The maximum profit occurs at a positive output level Q^*.

   What parameter restrictions are called for?

6. A purely competitive firm has a single variable input L (labor), with the wage rate W_0 per period. Its fixed inputs cost the firm a total of F dollars per period. The price of the product is P_0.

   (*a*) Write the production function, revenue function, cost function, and profit function of the firm.

(*b*) What is the first-order condition for profit maximization? Give this condition an economic interpretation.

(*c*) What economic circumstances would ensure that profit is maximized rather than minimized?

7. Use the following procedure to verify that the AR curve in Example 4 is negatively sloped:

(*a*) Denote the slope of AR by S. Write an expression for S.

(*b*) Find the maximum value of S, S_{max}, by using the second-derivative test.

(*c*) Then deduce from the value of S_{max} that the AR curve is negatively sloped throughout.

9.5 Maclaurin and Taylor Series

The time has now come for us to develop a test for relative extrema that can apply even when the second derivative turns out to have a zero value at the stationary point. Before we can do that, however, it is first necessary to discuss the so-called expansion of a function $y = f(x)$ into what are known, respectively, as a *Maclaurin series* (expansion around the point $x = 0$) and a *Taylor series* (expansion around any point $x = x_0$).

To *expand* a function $y = f(x)$ around a point x_0 means, in the present context, to transform that function into a *polynomial* form, in which the coefficients of the various terms are expressed in terms of the derivative values $f'(x_0)$, $f''(x_0)$, etc.—all evaluated at the point of expansion x_0. In the Maclaurin series, these will be evaluated at $x = 0$; thus we have $f'(0)$, $f''(0)$, etc., in the coefficients. The result of expansion is a *power series* because, being a polynomial, it consists of a sum of power functions.

Maclaurin Series of a Polynomial Function

Let us consider first the expansion of a *polynomial* function of the nth degree,

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots + a_nx^n \qquad \textbf{(9.6)}$$

into an equivalent nth-degree polynomial where the coefficients (a_0, a_1, etc.) are expressed instead in terms of the derivative values $f'(0)$, $f''(0)$, etc. Since this involves the transformation of one polynomial into another of the same degree, it may seem a sterile and purposeless exercise, but actually it will serve to shed much light on the whole idea of expansion.

Since the power series after expansion will involve the derivatives of various orders of the function f, let us first find these. By successive differentiation of (9.6), we can get the derivatives as follows:

$$\begin{aligned}
f'(x) &= a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \cdots + na_nx^{n-1} \\
f''(x) &= 2a_2 + 3(2)a_3x + 4(3)a_4x^2 + \cdots + n(n-1)a_nx^{n-2} \\
f'''(x) &= 3(2)a_3 + 4(3)(2)a_4x + \cdots + n(n-1)(n-2)a_nx^{n-3} \\
f^{(4)}(x) &= 4(3)(2)a_4 + 5(4)(3)(2)a_5x + \cdots + n(n-1)(n-2)(n-3)a_nx^{n-4} \\
&\vdots \\
f^{(n)}(x) &= n(n-1)(n-2)(n-3)\cdots(3)(2)(1)a_n
\end{aligned}$$