CHAPTER FOUR

CONDITIONS OF PROJECT OPTIMALITY

Production Theory

Economic forces act within the private sector of the economy to determine the supply curves of individual productive units (firms) and integrate them into a market-supply curve. Production theory attempts to explain the operation of these forces in ordering private sector production and thereby determining:

1 The total expenditure on inputs (raw material, machinery, labor, etc.)
2 The division of this expenditure among individual inputs
3 The way inputs are combined to produce each type of output
4 The amount of each output produced

Water resources development is a production process. In planning a production process for the public sector, many valuable insights can be gained from analyzing how economic forces would act to order production under ideal conditions.

4-1 INPUT AND OUTPUT The basic purpose of production is to convert material (input) into a more useful form (output). A water resources project is constructed to produce such desired output as irrigation water, reduced flood damage, a navigable channel, or electric power from a set of such inputs as earth, concrete, steel, and natural streamflow. As expressions of varied composition, both input and output are vector quantities. Each coordinate of the vector represents a specific input or output item.

The input vector consists of the sum of all the individual inputs. Thus,

$$X = x_1 + x_2 + \cdot \cdot \cdot + x_m \tag{4-1}$$

where X is the total input vector and the x are m individual inputs. The inputs fall by nature into two types. One type is the earth, concrete, and steel which go into a construction project, the required production capital. The other type is the natural streamflow from which the output is produced, the required raw material.

Evaluation of the input vector is complicated by the fact that the timing and magnitude of future streamflows cannot be predicted in advance and vary over wide ranges. Streamflow may be expressed as a continuous hydrograph, a running plot of stream discharge throughout the life of the project, as a probability distribution, or as such distribution moments as the mean and the standard deviation. Furthermore, one must keep in mind that all the other input and output vectors also have probability distributions even though their variance may be much smaller because one cannot predict any future event with absolute certainty.

The concept of the input vector can be simplified by defining its coordinates as the intermediate products of reservoirs, channel improvements, or powerhouses rather than the construction items of earth, concrete, and steel. The optimum combination of construction items in building a given intermediate product (say a dam of specified size) is best found from an engineering economy study seeking to minimize total cost. The first phase or the economy study determines how much earth, concrete, and steel should go into a dam of given size. The second phase or optimality analysis determines how big the dam should be and when and where it should be built.

The output vector consists of the sum of all the individual outputs. Thus,

$$Y = y_a + y_b + \cdot \cdot \cdot + y_n \tag{4-2}$$

where Y is the total output vector and the y are n individual outputs. Again, each output can only be predicted during planning and has a probability distribution expressing the range of possible values. The individual outputs are distinguished by type (irrigation from flood control) and by location (the area served.)

The composition of both vectors varies with time. Input contains investment in original construction, periodic replacement, and regular operation and maintenance. The output occurs in a time stream lasting the life of the project and varying from year to year but normally increasing with the general growth of the economy.

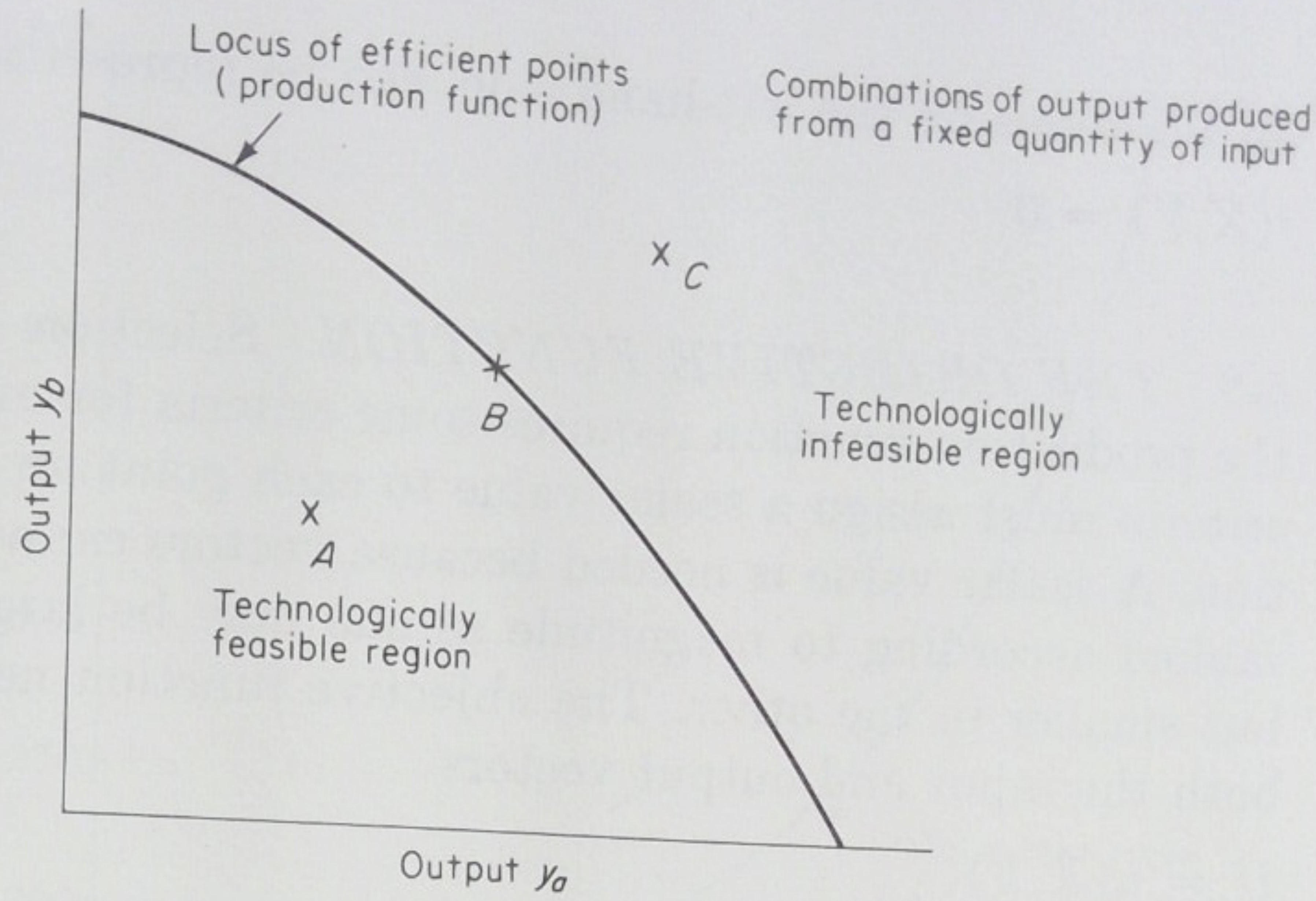


FIGURE 4-1 Production function.

4-2 THE PRODUCTION FUNCTION Economists have traditionally expressed the ability of a production process to produce an output vector from an input vector by a production function. As a simple illustration, one might have a production process in which two outputs are produced (Fig. 4-1). Engineering analysis may show that the combination of outputs represented by point A could physically be produced and the combination of point C would not be physically possible to produce. Continued analysis of alternative production possibilities will show which output vectors can and which output vectors cannot be produced. All those that can be produced are said to fall in the *technologically feasible region.*

Some points in the technologically feasible region are efficient, while others are inefficient. For example, it would be physically possible to dispose of the input vector without producing anything, but this would be wasteful and inefficient. The production represented by point A is inefficient because with the same input, the output vector could be increased to point B. For an inefficient point, the output vector can be unambiguously increased without increasing the input vector or the same output vector can be produced after an unambiguous decrease in the input vector. An unambiguous increase of a vector means some coordinates are increased without any being decreased. An unambiguous decrease means some coordinates are decreased without any being increased. The locus of efficient points is the outer limit of the technologically feasible region. The production function is the mathematical representation of this line. It is related to the entire input and output vectors and, by putting all the terms on the left-hand side, can be represented by the expression

$$f(X,Y) = 0 \tag{4-3}$$

4-3 THE OBJECTIVE FUNCTION Selection of the best point on the production function requires some criteria for evaluating worth. The criteria must assign a scalar value to each point on the production function. A scalar value is needed because vectors cannot be unambiguously ranked according to magnitude as one may be larger in one coordinate but smaller in the other. The objective function necessarily depends on both the input and output vectors.

$$U = u(X,Y) \tag{4-4}$$

Where n outputs are produced from m inputs,

$$U = \sum_{j=a}^{n} B_j y_j - \sum_{i=1}^{m} C_i x_i \tag{4-5}$$

where the B_i refer to the unit benefits associated with the corresponding coordinates of the output vector, and the C_j refer to the unit costs associated with the corresponding coordinates of the input vector. Conceptually, both benefits and costs may be either measured in monetary units or defined with respect to some broader based social welfare function (Sec. 5-2) without affecting the optimality criteria derived below.

4-4 COST AND BENEFIT CURVES Economic evaluation of production alternatives is based on the variation in total production cost with level of production output (called the *total-cost curve*) and the variation in the resulting benefit with level of production output (called the *total-benefit curve*). The total-cost curve is developed by summing the required input costs for a series of levels of output. The total-benefit curve is developed by summing values received by output users.

The total cost includes fixed cost and variable cost. Fixed costs remain constant regardless of output. They include capital recovery charges on the production facilities and other overhead costs. Variable costs vary with level of output. They include the costs of labor and material which can be added or deleted according to the level of production. Variable costs are marginal costs and are used to determine the optimum level of production according to the incremental-cost principle. Fixed costs are not marginal and thus have no influence on the optimum level of production, but they do influence whether or not total benefits exceed total costs or whether the project should be constructed at all.

Average-cost or -benefit curves are developed from total-cost or

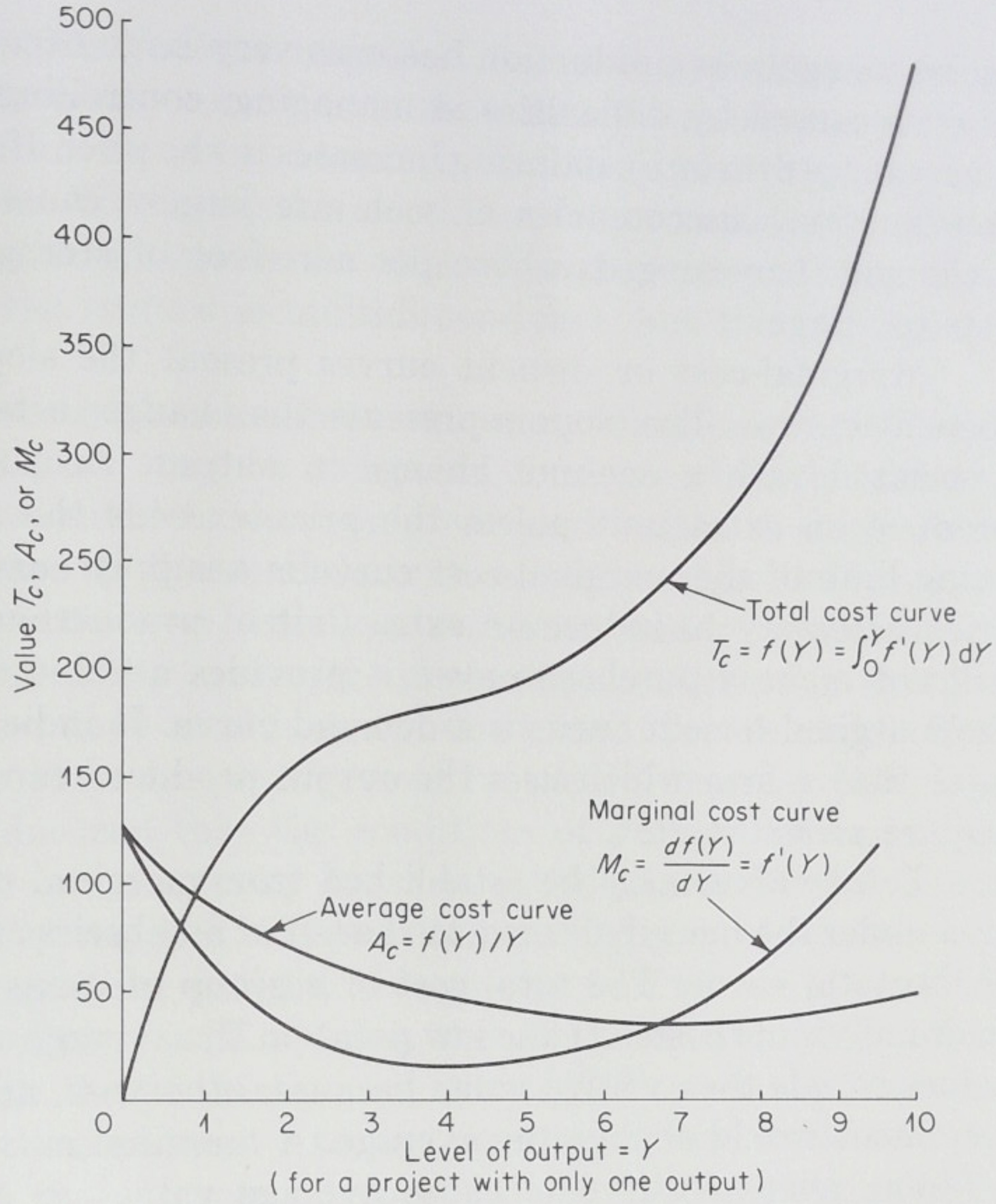


FIGURE 4-2 Representative total-, average-, and marginal-cost curves.

-benefit curves by dividing the total value by the level of output (Fig. 4-2). Average-cost curves are usually U-shaped. They decrease at first because of *economies of scale*, savings in production cost per unit stemming from increases in size of plant and output. Economies of scale result from (1) specialization of labor and (2) advanced technology. Specialization of labor may be impossible in a small plant because no single task requires the full-time effort of any individual worker, but as the size of plant and production increases, assembly line techniques become practical and reduce unit production costs. Advanced technology may be available, but its fixed cost may be too large to be warranted for a small plant. However, increased output spreads fixed cost over more units to make advanced technology economical. In reservoir construction, much of the fixed cost stems from providing sediment storage space, spillways to pass rare floods, access roads, and other minimum structural requirements. Average cost increases again as production becomes very large because of *diseconomies of scale* caused by difficulties in managing, controlling, and coordinating a very large firm or by inducing increases in the price of inputs. In reservoir construction, diseconomies of scale are largely caused by a decreasing yield and increasing dam size per acre-foot of storage as reservoir size gets too large.

Marginal-cost or -benefit curves present the slope of total-cost or -benefit curves. The slope represents the change in total cost or benefit associated with a one-unit change in output. Because a firm will not produce an extra unit unless the price exceeds the marginal cost, the rising limb of the marginal-cost curve is a supply curve. It indicates the price necessary to induce an extra unit of production. Because a buyer will not make a purchase unless it provides a value exceeding the cost, the marginal-benefit curve is a demand curve. It indicates the maximum price that a firm which uses the output produced can afford to spend to acquire an extra unit.

Total values may be established from marginal curves because the area under the marginal curve to the left of an abscissa equals the ordinate of the total curve. The total cost of a group of items equals the sum of their individual costs. At the low point on the average curve, the marginal values equals the average value because, otherwise, adding the marginal increment would change the average. A marginal curve must plot below a falling average curve to cause average values to drop. Obversely, a marginal curve must plot above a rising average curve. Therefore, marginal curves, like average curves, are generally U-shaped, but justified more to the left.

4-5 OPTIMALITY CONDITIONS One approach to determining the optimum production process would be systematic trial-and-error evaluation of the net benefit [Eq. (4-5)] for each point on the production function [Eq. (4-3)], which will lead to the point having maximum value. The search is made easier by being able to recognize the characteristics peculiar to such a point.

A mathematician distinguishes necessary from sufficient conditions when seeking such characteristics. Necessary conditions are characteristics the solution must have, but they do not guarantee that a point having them is the solution. Sufficient conditions guarantee a solution. A simple example applying this distinction to project optimization is found in the production of one output from two inputs. If the two inputs are plotted on a horizontal plane, the maximum output which can be produced from each combination of the two can be plotted vertically to produce what resembles a topographic contour map and is called a *production* or a *response surface*. A necessary condition to show that the maximum output has been produced from these two inputs is that the surface slope downward in all directions. However, this condition is not sufficient because there might be a higher hilltop on the other side of a valley. For greater numbers of inputs the surface is multidimensional, but the principle is the same.

Proof that a given peak is the maximum may follow one of two lines. The height of each peak may be computed to show which is highest. Evidence may be presented to show only one peak exists. A single peak will in fact be the case if all second derivatives of the objective function are continuously negative, a situation occurring under conditions of diminishing marginal utility as supply curves slope continuously upward and demand curves slope continuously downward to the right. The second approach is more frequently used in water resources planning. It needs to be emphasized that the conditions of project optimality to be developed below are necessary but not sufficient conditions. Proof of absolute optimality requires evidence that no higher peak exists.

For a water resources project, the goal is to find an alternative having maximum value of $u(X,Y)$ with the constraint that only alternatives contained on the production function $f(X,Y) = 0$ need be considered. The conditions necessary to having a maximum value may be determined by (1) the geometrical approach or (2) the mathematical approach. The geometrical approach follows immediately, while the mathematical approach is presented later.

Geometric Derivation of Basic Rules

4-6 *OPTIMALITY CONDITION 1: COMBINATION OF INPUTS*

The optimal production process must use the least costly combination of inputs able to produce any given level of output. For example, the sizes of the two dams to provide flood control must be selected to achieve the desired level of flood reduction at minimum cost. The least-cost combination of inputs can be found geometrically by the use of isoquant lines and isocost lines. If the problem is reduced to two dimensions for practical presentation, isoquant lines (Fig. 4-3) show different combinations of two inputs which can produce equal amounts of a single output.

Isoquants are analogous to indifference curves and have analogous characteristics:

1 Two isoquants cannot intersect. An intersection would require the maximum output which could be produced with the same input to be two different amounts.

2 Isoquants slope downward to the right because increased use of one input reduces the quantity of another required to obtain a given level of output. Channel improvement can be substituted for reservoirs to provide flood control.

3 Isoquants are convex to the origin because of the decreasing ability of one input to be substituted for another to obtain a given level of output. As more channel improvement and less reservoir storage are used to produce a given level of flood control, the larger is the incremental channel improvement required to effect a unit reduction in flood storage. This is called the *principle of diminishing marginal rate of substitution.*

In Fig. 4-3, the isoquant for output y_a shows the possible combinations of x_1 and x_2 which could be used in its production. The most efficient combination depends on the unit prices of the inputs, just as the unit prices of goods guide spending to maximize consumer satisfaction (Sec. 3-10).

Isocost lines (Fig. 4-4) indicate the input combinations that can be purchased by a given production budget and are analogous to the line of attainable combinations–used in demand analysis. If the production budget is T, the price of x_1 is P_{x_1}, and the price of x_2 is P_{x_2};

$$T = P_{x_1}x_1 + P_{x_2}x_2 \tag{4-6}$$

which is the equation of a straight line with a slope of P_{x_1}/P_{x_2}.

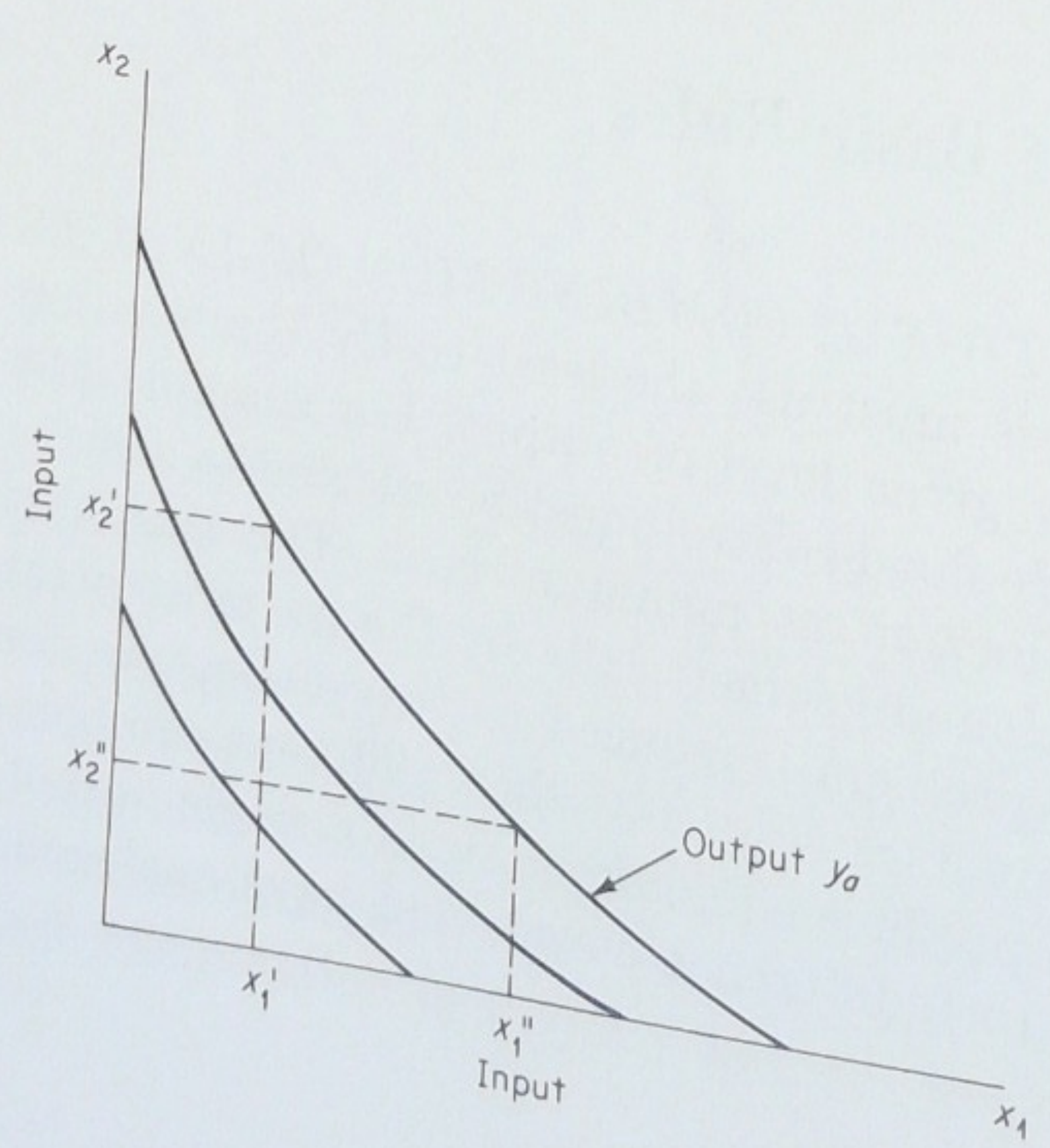


FIGURE 4-3 Isoquants (lines of equal output).

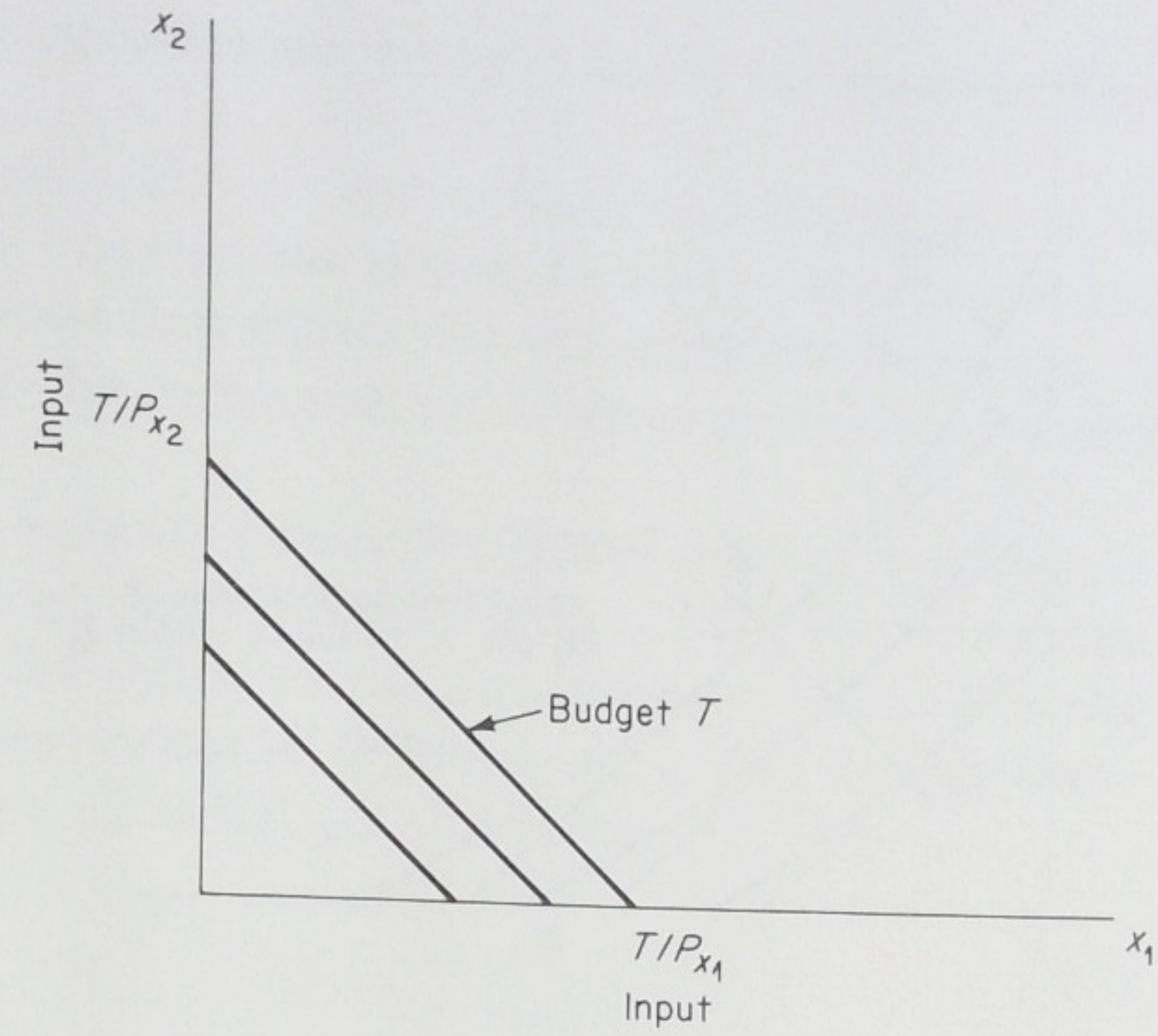


FIGURE 4-4 Isocost lines (lines of equal cost).

Production of a given level of output with the least-cost combination of resources occurs where an isocost line (slope of P_{x_1}/P_{x_2}) is tangent to the isoquant (slope of $MRS_{x_2x_1}$). Therefore,

$$MRS_{x_2x_1} = \frac{P_{x_1}}{P_{x_2}} \tag{4-7}$$

Figure 4-5 shows a number of isoquants with tangent isocosts. The line AB joining the points of tangency is called the *expansion path* and is the locus of the least-cost combinations of inputs for varying levels of total output.

An example will clarify the procedure for combining inputs. Suppose the axes in Fig. 4-5 represent two single-purpose flood control reservoirs. We can calculate combinations of storage capacity in the two reservoirs providing a fixed level of flood peak reduction (isoquant) and the storage combinations that can be constructed with a fixed budget (isocost). As many isoquants and isocosts as needed may be calculated and drawn. Each point of tangency represents the least-cost combination of reservoirs to provide a given level of flood reduction.

4-7 *OPTIMALITY CONDITION 2: COMBINATION OF OUTPUTS*

With two outputs, such as municipal water supply and hydropower, total production must be divided between the two to maximize benefits. One may begin the analysis by plotting on coordinate axes representing two

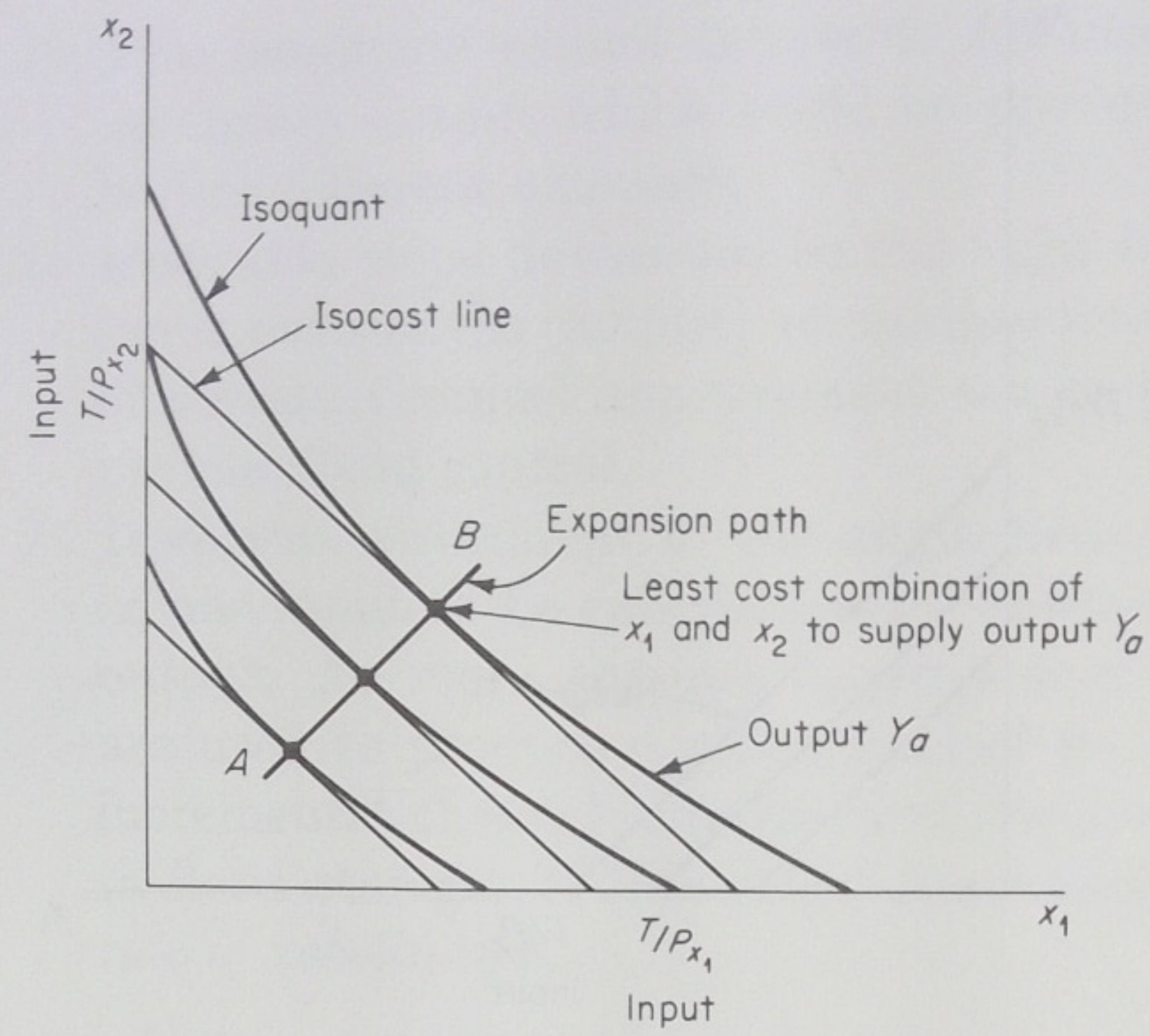


FIGURE 4-5 Determination of least-cost combination of inputs.

outputs y_a and y_b (Fig. 4-6) each of the family of curves showing combinations of outputs that can be produced at a given cost. Each curve indicates all combinations of outputs y_a and y_b that can be produced for the indicated sum and is called a *product-transformation curve* because to move along it, one output must be increased while the other is reduced.

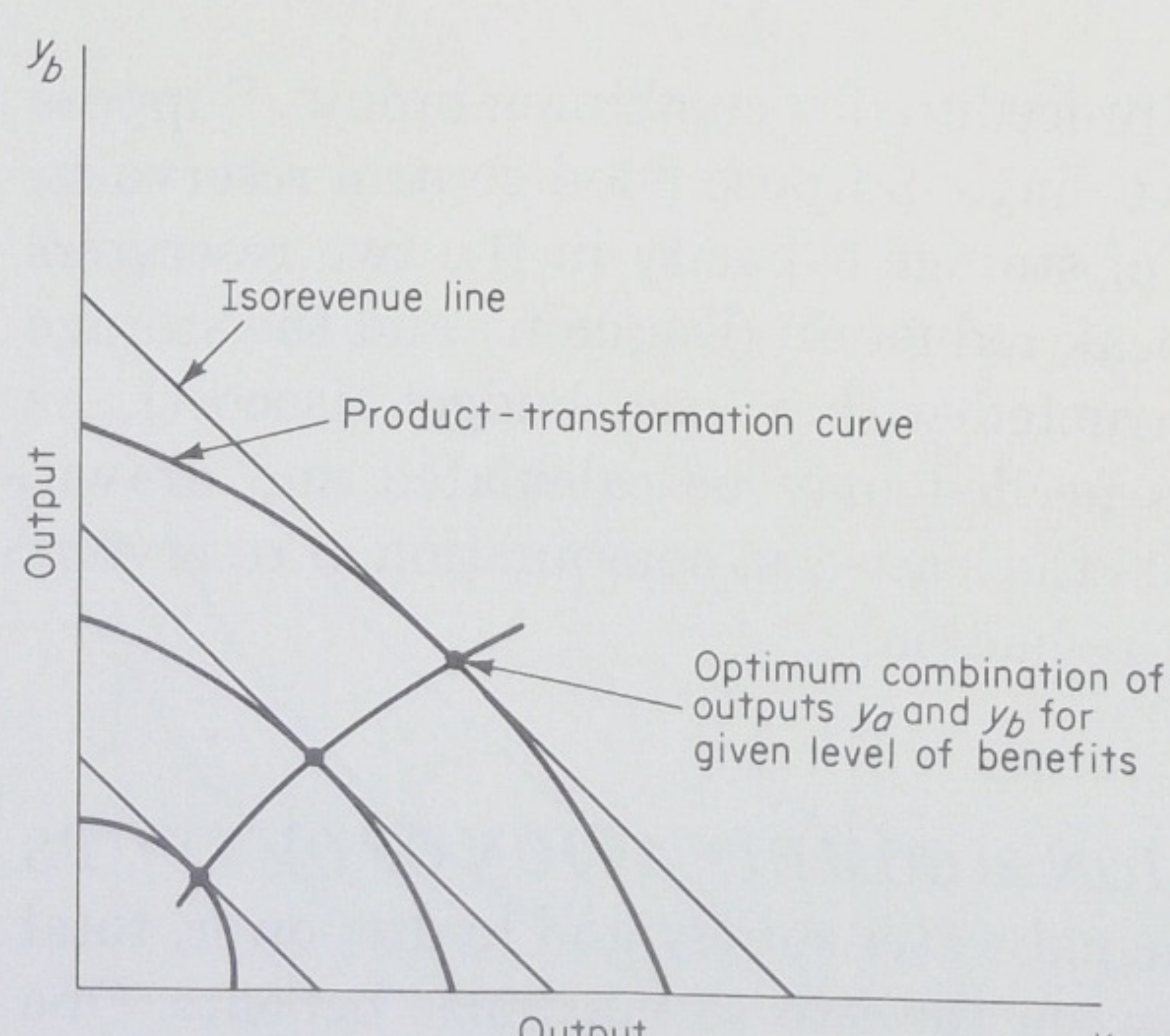


FIGURE 4-6 Optimum combination of outputs.

The slope of the product-transformation curve is called the *marginal rate of transformation.*

A family of parallel lines called *isorevenue lines* may also be drawn in Fig. 4-6. The slopes of these lines are the ratios of the market prices of the two outputs. Each isorevenue line shows the different combinations of outputs that would sell for the same amount of gross revenue or would produce a given benefit.

The optimum mix of outputs achieves a given level of benefits at least cost, or put another way, the maximum level of benefits for a given level of costs. In Fig. 4-6, the optimum combinations are located at the points of tangency of the isorevenue lines (slope of P_{y_a}/P_{y_b}) and product-transformation curves (slope of $MRT_{y_a y_b}$). Therefore,

$$MRT_{y_a y_b} = \frac{P_{y_a}}{P_{y_b}} \tag{4-8}$$

Product-transformation curves are concave to the origin (Fig. 4-6) if the outputs are joint products of the same productive process or if the production of one is facilitated by production of the other. But if the production of one output hinders production of another, the product-transformation curves are convex to the origin. In this case, benefits are maximized by producing only one of the two outputs, a boundary solution for which Eq. (4-8) does not apply (Fig. 4-7). The product-transformation curve reaches the highest isorevenue line on the y_a axis; therefore, only

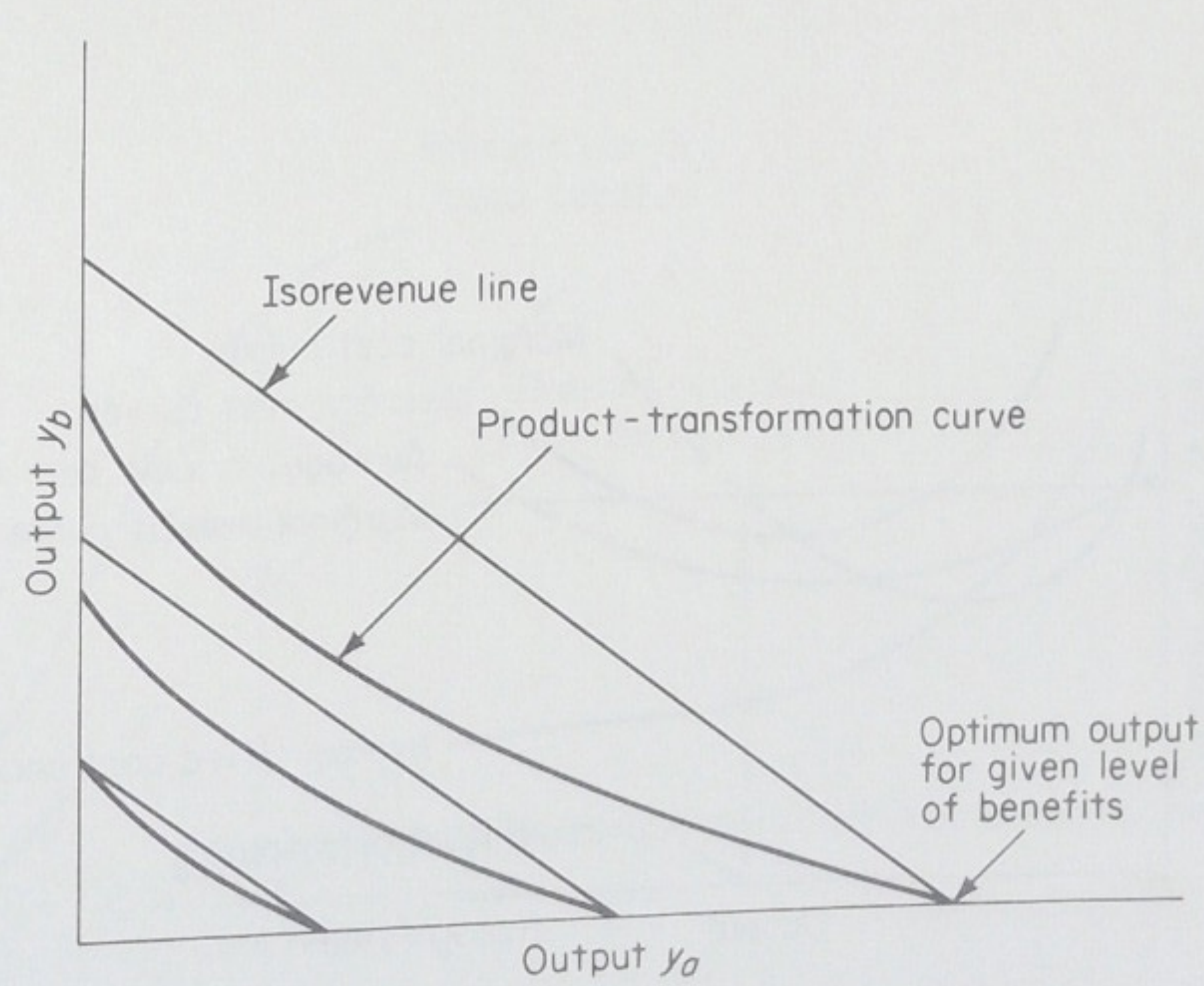


FIGURE 4-7 Optimum output for concave product-transformation curves.

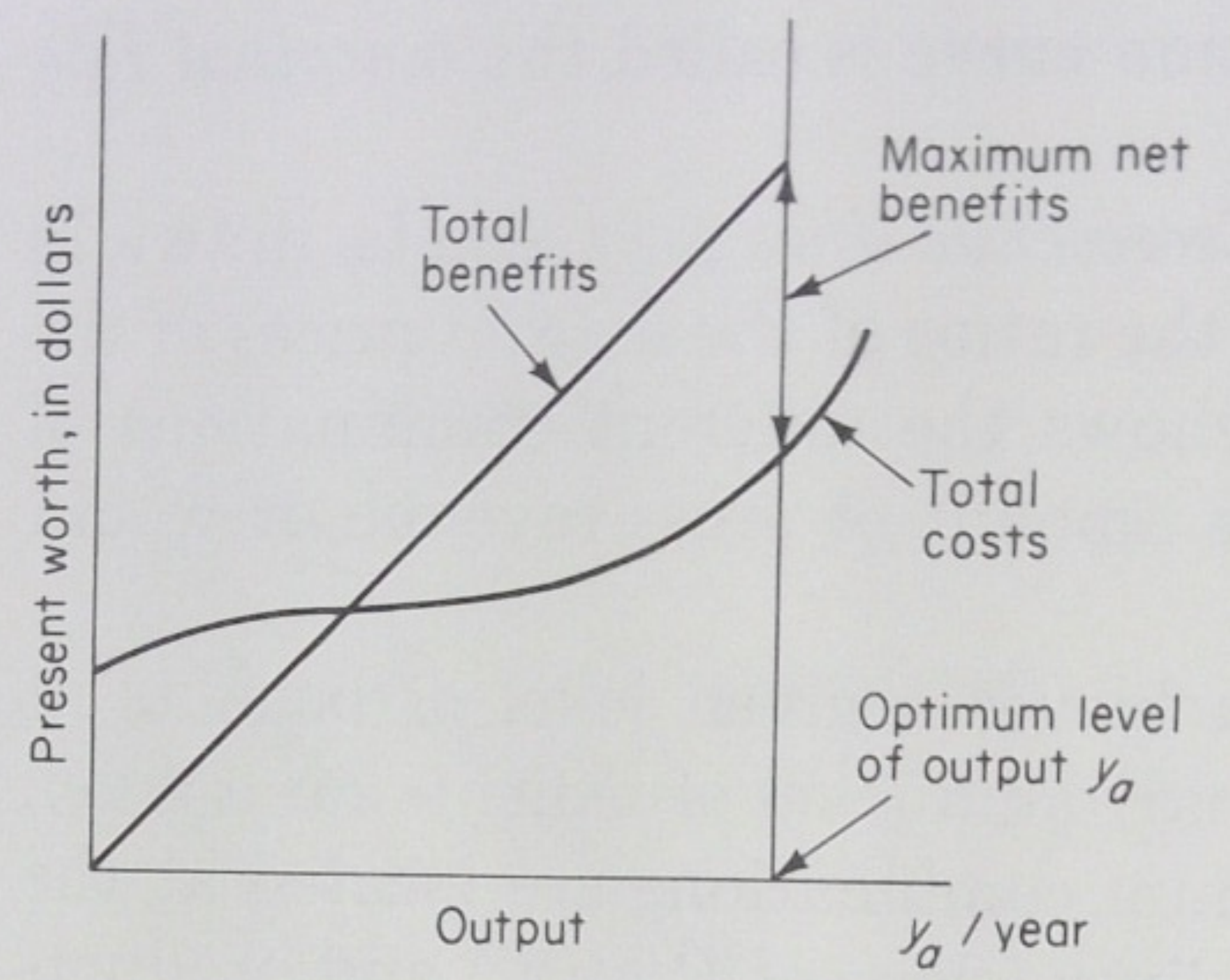


FIGURE 4-8 Determination of optimum level of output.

y_a would be produced. Note the isorevenue line is not tangent to the product-transformation curve at this point.

Optimality condition 2 may be summarized by saying production should be divided between two outputs so that the marginal benefit of any input in the production of one equals the marginal benefit of the input in production of the other. Otherwise production could be shifted between the outputs to increase benefits.

4-8 OPTIMALITY CONDITION 3: LEVEL OF OUTPUT Optimality condition 3 determines the optimum level of output, on the assumption that conditions 1 and 2 have already been met. It states that benefit is maximized if output is increased up to the point where the marginal

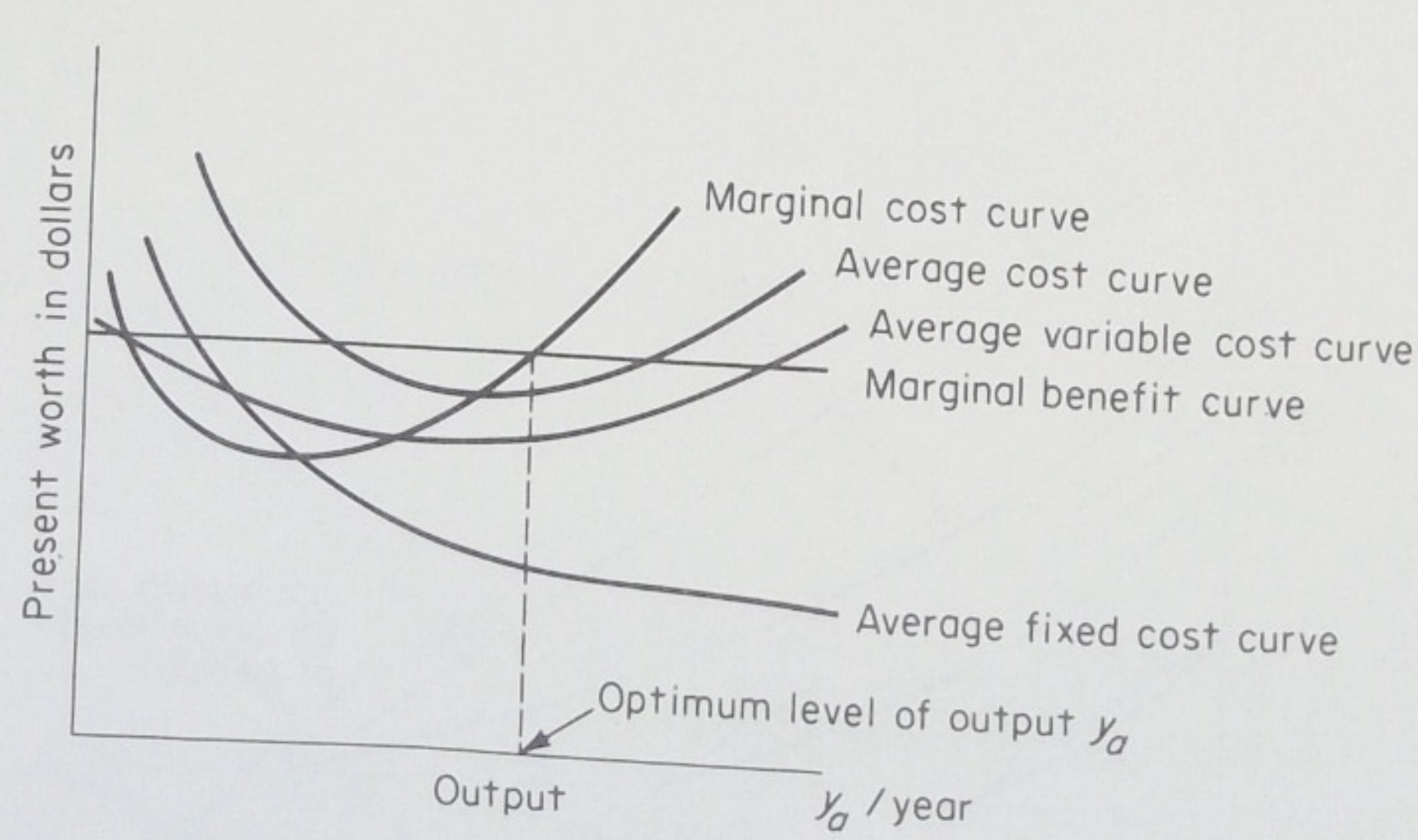


FIGURE 4-9 Optimum level of output by using marginal-cost and marginal-benefit curves.

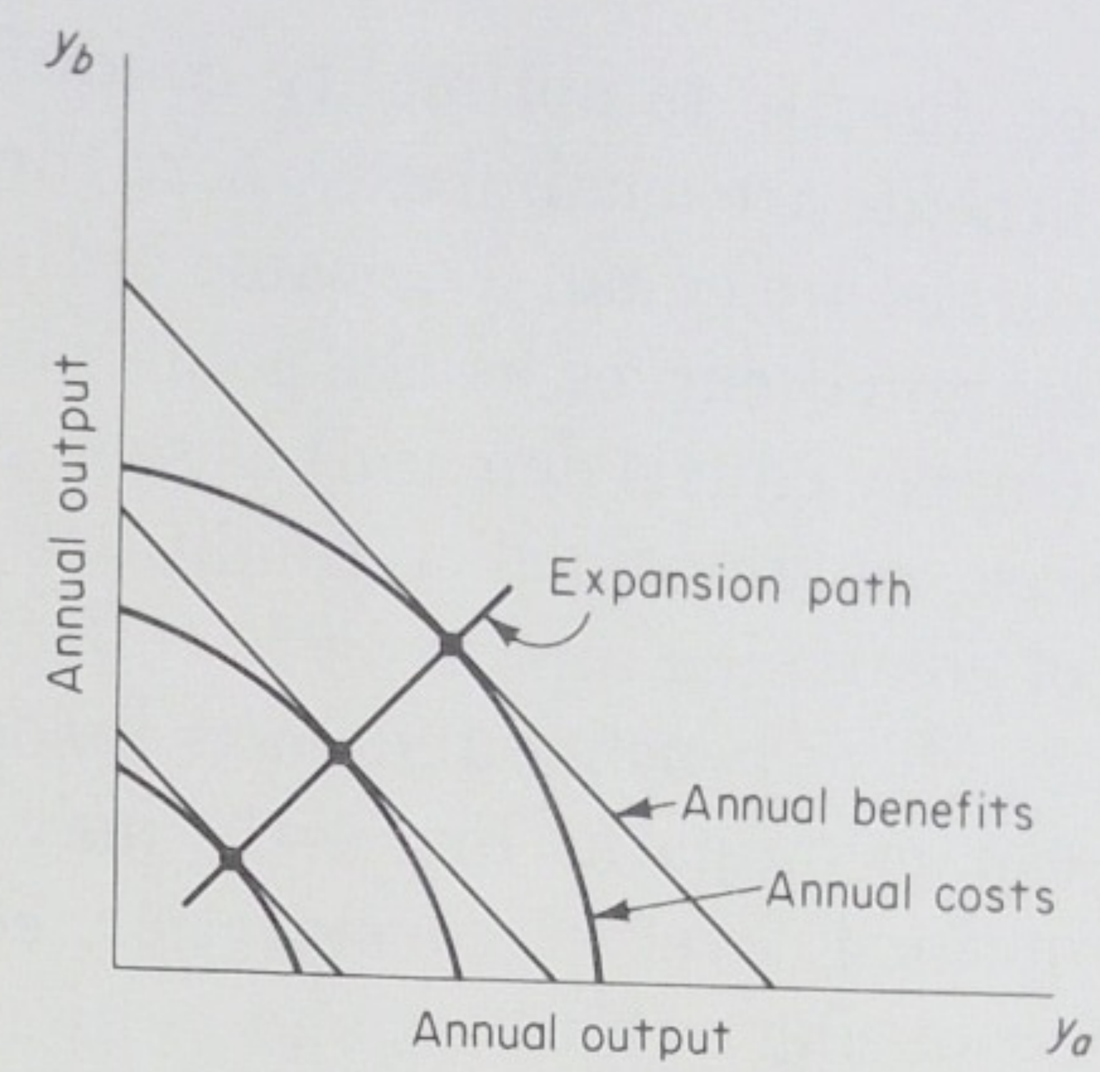


FIGURE 4-10 Optimum construction costs and combination of outputs.

costs equal the marginal benefits, or in engineering-economy terminology, incremental costs equal incremental benefits. The two marginal values are equal where the slopes of the total-value curves are equal or where the distance between them is maximum (see Fig. 4-11).

For the two-input one-output case, optimality condition 1 provides the basis for calculating the minimum cost of attaining different levels of output. The results may be plotted in a total-cost curve (Fig. 4-8.) The total-benefit curve may be plotted by multiplying the unit price of the output times the quantity of output. Under the conditions of pure competition, the unit price is constant; hence the total-benefit line is straight. For only one output, optimality condition 2 does not apply, and we can

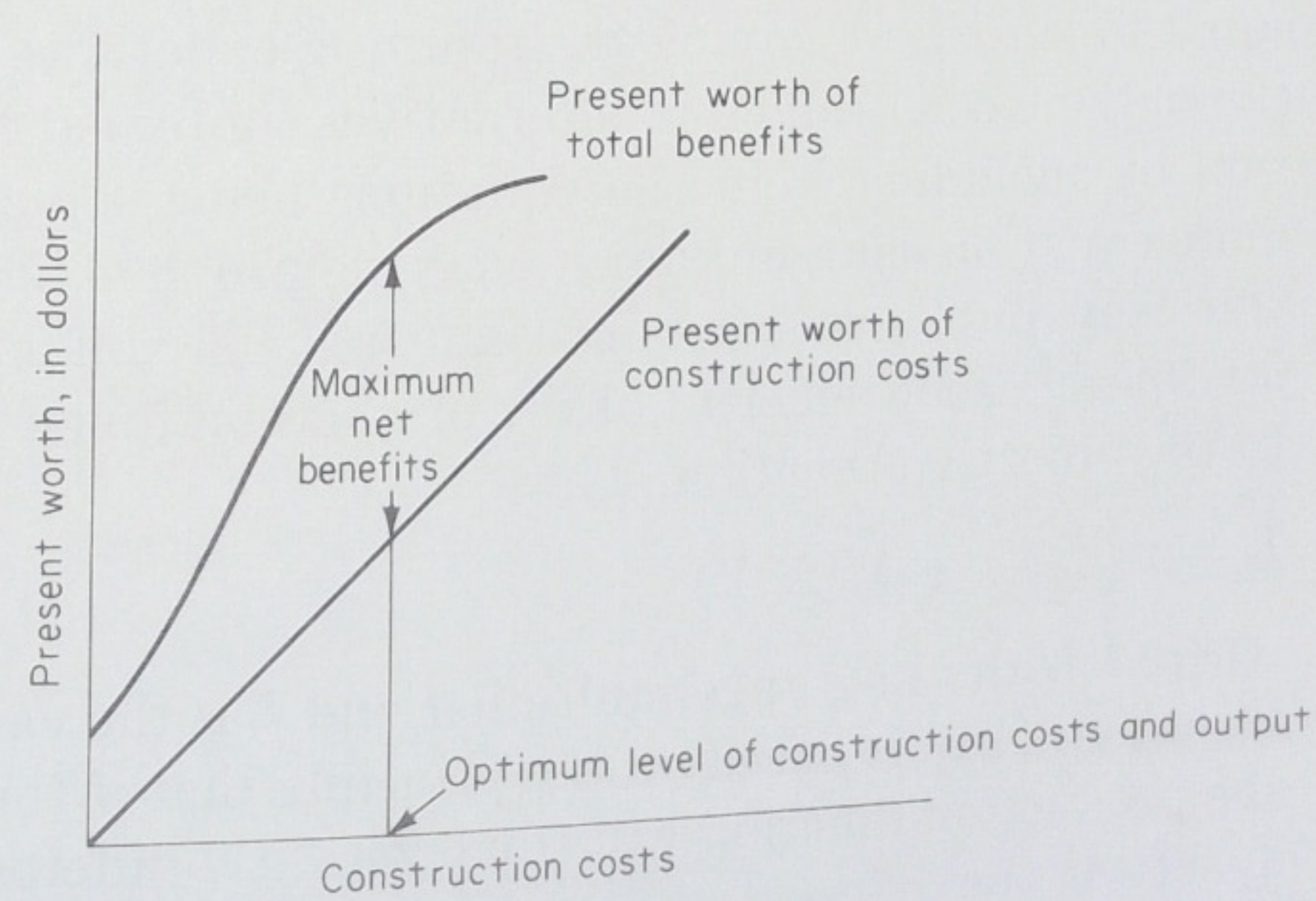


FIGURE 4-11 Determination of optimum level of construction costs and output.

go directly to optimality condition 3. As can be seen from Fig. 4-8, net benefits are maximized where the slopes of the total-benefit and total-cost curves are equal. The same optimum level of output can be expressed on marginal curves as the point where marginal cost equals marginal benefit (supply equals demand) as seen in Fig. 4-9. For the one-input two-output case, only optimality conditions 2 and 3 apply since there is no problem of combining inputs.

The benefits and costs associated with output combinations on the expansion path (Fig. 4-10) may be plotted (Fig. 4-11) to determine the optimal level of construction cost. This cost then is divided between producing outputs y_a and y_b by referring back to the corresponding point on the expansion path. For the multiple-input multiple-output case, total-cost and total-benefit curves must be used as described in the next chapter.

Mathematical Derivation of Basic Rules

The goal of project optimization is to maximize the objective function $u(X,Y)$ by choosing the best alternative on the production function $f(X,Y) = 0$, where X is an m coordinate and Y is an n coordinate vector.

4-9 LAGRANGE MULTIPLIERS Differential calculus can be used to find such a maximum by differentiating the objective function with respect to each of the $(n + m)$ vector components, setting each differential equal to zero, and solving the resulting equations. However, the solution can only be constrained to alternatives contained on the production function by including it in the equations being solved. This introduces one more equation than unknown, and the problem becomes overdetermined. One way out is to introduce an artificial unknown called a *Lagrange multiplier*[1] as a coefficient of the production function and add the product to the objective function. Thus,

$$L = u(X,Y) + \lambda f(X,Y) \tag{4-9}$$

where λ is the Lagrange multiplier and L is the variable to be maximized. Equation (4-9) is based on the principle that if the constraint is satisfied, the production function will equal zero. By differentiating with respect to λ, setting the differential (which will be the production function) equal to

[1] Lagrange multipliers are described in much greater detail in William J. Baumol, "Economic Theory and Operations Analysis" (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1961) pp. 54–59.

zero, and including it in the set of equations to be solved, one incorporates the production function into the solution. The approach is illustrated by Ex. 4-1.

EXAMPLE 4-1

In a constrained maximization problem, the objective function is $Y = 10ab$, and the constraint is $5a + b = 200$. In other words, we are seeking the values for a and b which maximize the first expression without exceeding the upper limit of the second expression. Placed in the format of Eq. (4-9), the constraint expression $5a + b - 200 = 0$ may be added to the objective function without changing its value, to give

$$Y = 10ab + \lambda(5a + b - 200)$$

By partial differentiation with respect to each of the unknowns a, b, and λ and by setting each differential equal to zero,

$$\frac{\partial Y}{\partial a} = 10b + 5\lambda = 0$$

$$\frac{\partial Y}{\partial b} = 10a + \lambda = 0$$

$$\frac{\partial Y}{\partial \lambda} = 5a + b - 200 = 0$$

Solution of the three equations gives $a = 20$, $b = 100$, and $\lambda = -200$. When these values are substituted in the objective function, the maximum value is found to be $Y = 20{,}000$.

The economic significance of λ is that if the number on the right-hand side of the constraint equation had been 201 instead of 200, the optimum value of Y would have been $20{,}000 + \lambda$, or 20,200.

Lagrange multipliers permit constrained maximization by introducing as many artificial unknowns as there are constraints to make the number of unknowns and equations equal. The problem could be solved without using Lagrange multipliers by substituting the constraint expression in the objective function before using the differential calculus approach, but for many expressions the algebra makes this approach difficult if not impossible.

4-10 APPLICATION OF THE LAGRANGE MULTIPLIER In order to find the maximum, Eq. (4-9) must be differentiated individually with respect to each coordinate of the two vectors as well as λ; and each differential must be set equal to zero. Thus,

$$\frac{\partial u(X,Y)}{\partial x_i} = -\lambda \frac{\partial f(X,Y)}{\partial x_i} \qquad i = 1, 2, \ldots, m \tag{4-10}$$

and

$$\frac{\partial u(X,Y)}{\partial y_j} = -\lambda \frac{\partial f(X,Y)}{\partial y_j} \qquad j = a, b, \ldots, n \tag{4-11}$$

and

$$\frac{\partial u(X,Y)}{\partial \lambda} = f(X,Y) \tag{4-12}$$

By dividing Eq. (4-11) into Eq. (4-10) and pairs of Eqs. (4-10) and (4-11) into each other, one obtains

$$\frac{\partial u(X,Y)/\partial x_i}{\partial u(X,Y)/\partial y_j} = \frac{\partial f(X,Y)/\partial x_i}{\partial f(X,Y)/\partial y_j} \tag{4-13}$$

$$\frac{\partial u(X,Y)/\partial x_1}{\partial u(X,Y)/\partial x_2} = \frac{\partial f(X,Y)/\partial x_1}{\partial f(X,Y)/\partial x_2} \tag{4-14}$$

$$\frac{\partial u(X,Y)/\partial y_a}{\partial u(X,Y)/\partial y_b} = \frac{\partial f(X,Y)/\partial y_a}{\partial f(X,Y)/\partial y_b} \tag{4-15}$$

Since $f(X,Y)$ must equal zero, an increase in one element must be offset by a decrease in another. Therefore,

$$\frac{\partial f(X,Y)/\partial x_i}{\partial f(X,Y)/\partial y_j} = -\frac{\partial y_j}{\partial x_i} \tag{4-16}$$

$$\frac{\partial f(X,Y)/\partial x_1}{\partial f(X,Y)/\partial x_2} = -\frac{\partial x_2}{\partial x_1} \tag{4-17}$$

$$\frac{\partial f(X,Y)/\partial y_a}{\partial f(X,Y)/\partial y_b} = -\frac{\partial y_b}{\partial y_a} \tag{4-18}$$

By combining Eqs. (4-13) and (4-16), (4-14) and (4-17), and (4-15) and (4-18), one finally achieves

$$\frac{\partial u(X,Y)/\partial x_i}{\partial u(X,Y)/\partial y_j} = -\frac{\partial y_j}{\partial x_i} \tag{4-19}$$

$$\frac{\partial u(X,Y)/\partial x_1}{\partial u(X,Y)/\partial x_2} = -\frac{\partial x_2}{\partial x_1} \tag{4-20}$$

$$\frac{\partial u(X,Y)/\partial y_a}{\partial u(X,Y)/\partial y_b} = -\frac{\partial y_b}{\partial y_a} \tag{4-21}$$

4-11 THREE BASIC OPTIMALITY CONDITIONS Analysis of the meaning of the terms in Eq. (4-19) reveals $\partial u(X,Y)/\partial x_i$ to equal the marginal cost of input i, or MC_i, and $\partial u(X,Y)/\partial y_j$ to equal the marginal benefit from output j, or MB_j. The term on the right-hand side of the expression, $-\partial y_j/\partial x_i$, is what economists call the *marginal physical product*, or the additional output which can be produced per unit of increase in input. The negative sign results from the opposite nature of inputs and outputs. Thus

$$\frac{MC_i}{MB_j} = MPP_{ij} \tag{4-22}$$

where MPP_{ij} is read as the marginal physical productivity of the ith input when devoted to the jth output.[1] Similar analysis of Eq. (4-20) reveals MC_1 and MC_2. The marginal rate of substitution was defined in Sec. 4-6 as the marginal rate at which quantities of the second input need to be substituted for a unit reduction in the first input while holding the level of production constant, $-\partial x_2/\partial x_1$. Thus

$$\frac{MC_1}{MC_2} = MRS_{21} \tag{4-23}$$

Equation (4-21) contains MB_a, MB_b, and the marginal rate of transformation (Sec. 4-7), or the marginal rate at which production can be shifted from the second output to the first to effect a unit change in the first without changing the input. Thus

$$\frac{MB_a}{MB_b} = MRT_{ba} \tag{4-24}$$

4-12 APPLICATION OF OPTIMALITY CONDITIONS

The three equations of the last section may be used to answer the four questions fundamental to structuring production. The application may be illustrated by a water resources project which produces two outputs, flood control y_a and irrigation y_b, from two inputs, reservoir storage x_1 and channel improvement x_2.

The first fundamental question is: How should the inputs be combined to produce a given output? The answer is found in Eq. (4-23). The marginal cost of an input is its unit price. If unit price varies with amount purchased, the marginal price at the input actually used should be applied. Equation (4-23) says the inputs should be combined in such amounts that the ratio of their prices equals the marginal rate at which one input can be substituted for another with all other components of the production function constant. With the other inputs and outputs constant, and based on a typical production function, one may evaluate x_1 as a function of x_2 to obtain the data in Table 4-1 and the curve in Fig. 4-12. The curve is

[1] This notation is simplified from that previously used, which would be $MPP_{x_i y_j}$.

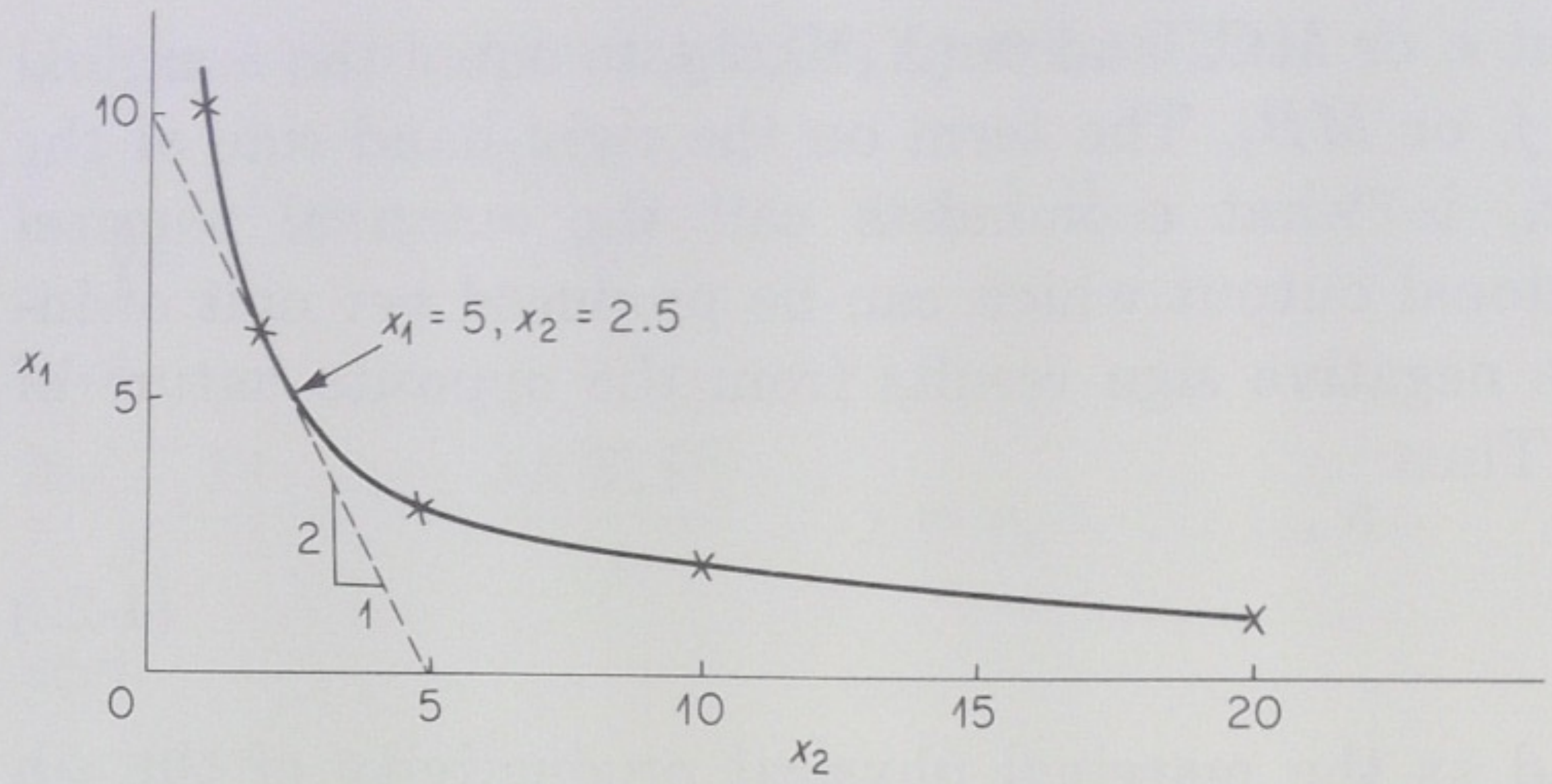


FIGURE 4-12 Optimum combination of inputs.

concave toward the origin because of an increasing inability to substitute a second input for the first as one approaches the point where the first is not used at all. Our rule says to substitute x_2 for x_1 until the marginal rate of substitution MRS of x_2 for x_1 equals the ratio of marginal costs MC. This would give the point $x_1 = 5, x_2 = 2.5$. The total-cost TC column in Table 4-1 and the price line in Fig. 4-12 show how this is also the point of minimum total cost.

The second fundamental question is: How should total production be divided among specific outputs? The answer is found in Eq. (4-24). The rule says that the outputs should be produced in such amounts that the ratio of their unit benefits equals the marginal rate at which production can be shifted from one output to another with all other inputs and outputs

TABLE 4-1 Optimum Rate of Substitution

x_1	x_2	MRS	TC
10	1		60
		0.25	
6	2		50
		0.5	
4	3		50
		2	
3	5		65
		5	
2	10		110
		10	
1	20		205

$MC_1 = \$5$
$MC_2 = \$10$
Optimum $MRS_{21} = 5/10 = 0.5$

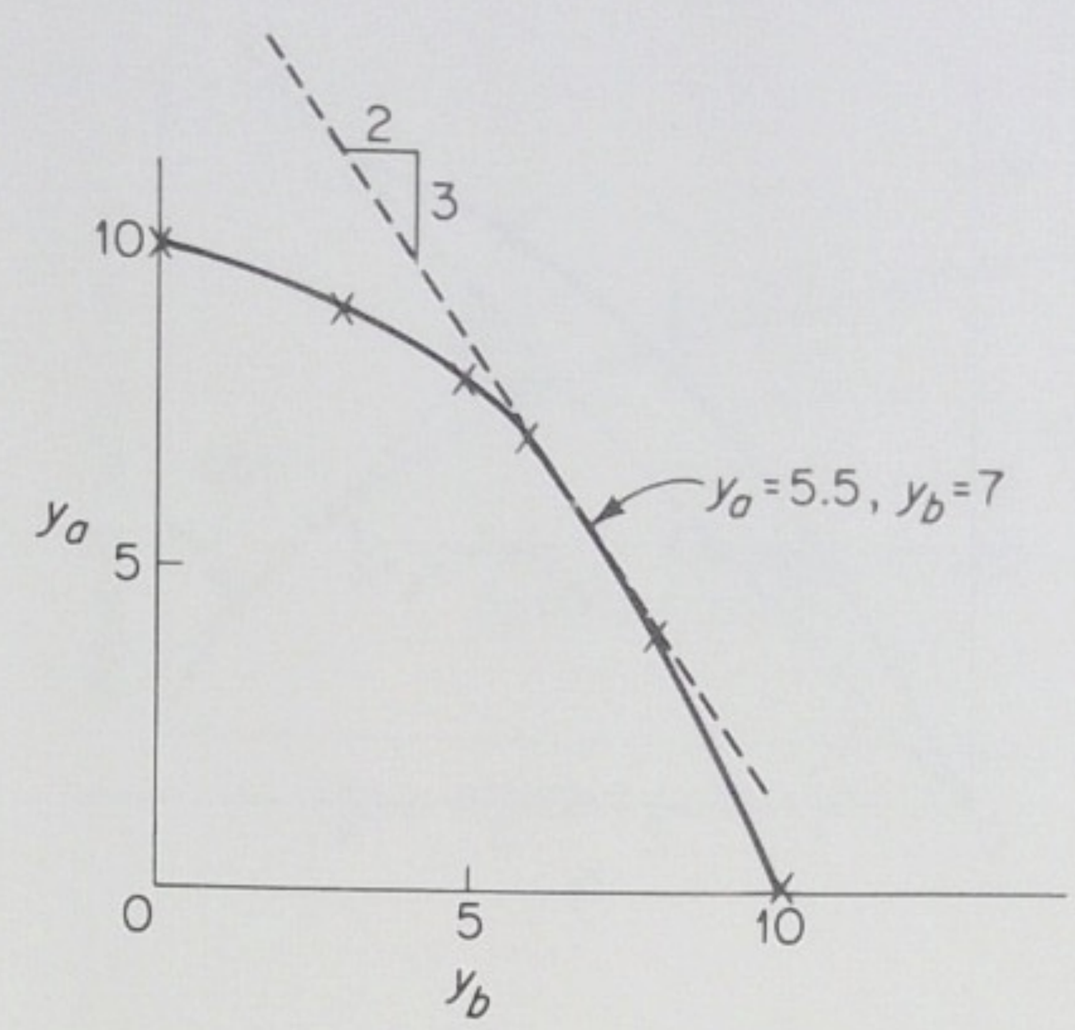


FIGURE 4-13 Optimum division of outputs.

held constant. For this condition, one may evaluate y_a as a function of y_b to obtain the data in Table 4-2 and the curve in Fig. 4-13. The curve is concave away from the origin because the first units of additional output can normally be produced at less marginal cost than later ones. Our rule says to substitute y_b for y_a until the marginal rate of transformation MRT of y_b for y_a equals the ratio of their marginal benefits MB. This corresponds to the point $y_a = 5.5$, $y_b = 7$. The total-benefit TB column in Table 4-2 and the income line in Fig. 4-13 show how this is also the point of maximum total benefit.

The third fundamental question is: How much of a specified input should be devoted to the production of a specified output? The answer is found in Eq. (4-22). The rule says that the input should be utilized in such amount that the ratio of marginal input cost to marginal output

TABLE 4-2 Optimum Rate of Transformation

y_a	y_b	MRT	TB
10	0		20
		3	
9	3		27
		2	
8	5		31
		1	
7	6		32
		0.67	
4	8		32
		0.50	
0	10		30

$MB_a = \$2$
$MB_b = \$3$
Optimum $MRT_{ba} = 2/3 = 0.67$

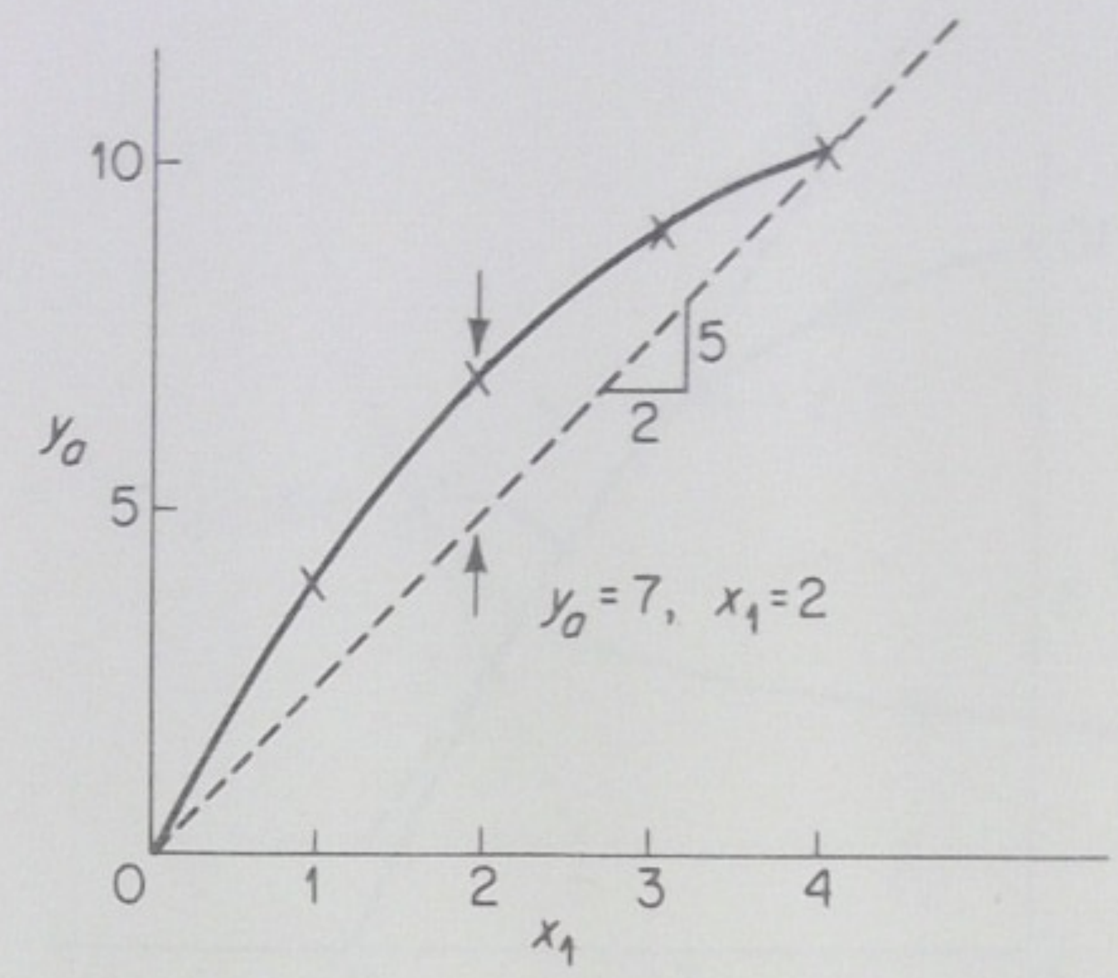


FIGURE 4-14 Optimum input to use in producing a specified output.

benefit equals the ratio of marginal physical output to marginal physical input. For this condition, one may evaluate y_a as a function of x_1 to obtain the data in Table 4-3 and the curve in Fig. 4-14. This is only one of four possible curves of this type for our two-input two-output example, one for each combination of output and input. Our rule says to increase the amount of x_1 used in producing y_a until the marginal increase in y_a for a unit of x_1 equals the ratio of their marginal unit values. This corresponds to the point $y_a = 7$, $x_1 = 2$. The net-benefit $B - C$ column in Table 4-3 and the distance between the two lines in Fig. 4-14 show this to be also the point of maximum excess of benefits over costs.

The fourth fundamental question is: How large should the total project be? The answer is found by shifting the approach demonstrated by Fig. 4-14 from coordinate pairs to the total input and output vectors. However, the analysis requires the selection and evaluation of trial vectors as no straightforward solution of the type used in answering the first three questions is possible. A trial output vector is selected. The first rule is

TABLE 4-3 Optimum Physical Product

x_1	y_a	MPP_{1a}	$B - C$
0	0		0
		4	
1	4		3
		3	
2	7		4
		2	
3	9		3
		1	
4	10		0

$MC_1 = 5$
$MB_a = 2$
Optimum $MPP_{1a} = 2.5$

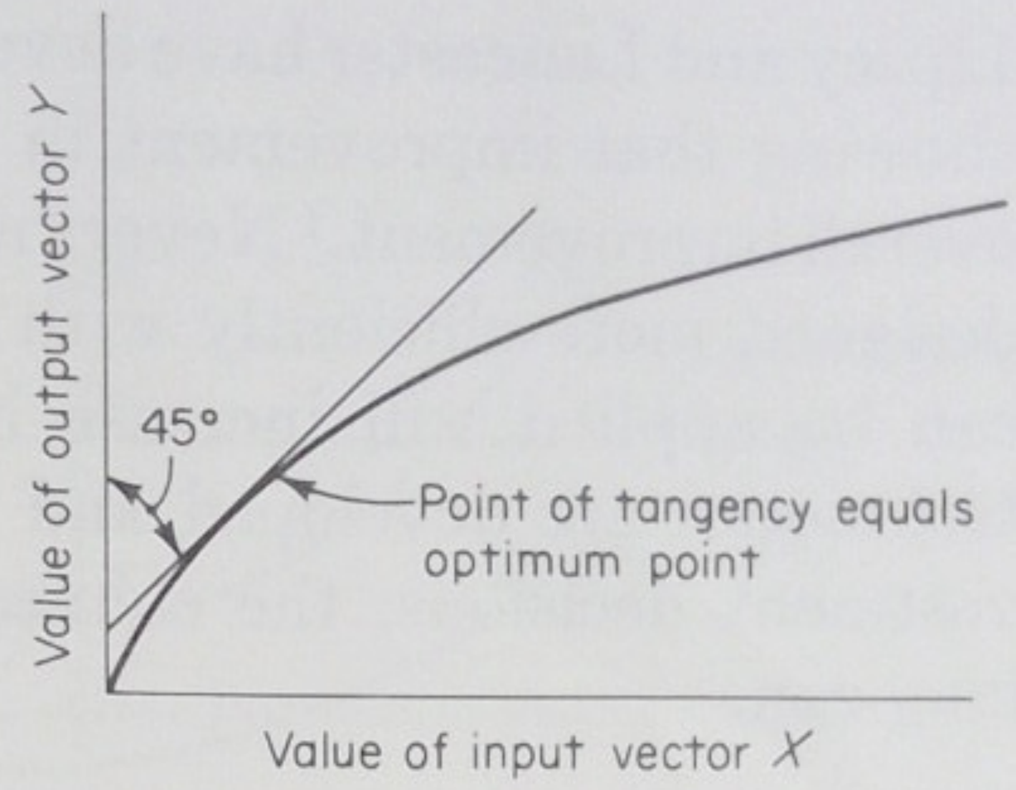


FIGURE 4-15 Selection of the optimum combination of input and output vectors.

applied to determine the optimum input vector for that output vector. The value of the output vector is determined as the sum of the products of the output coordinate magnitudes and the marginal benefits. The value of the input vector is determined as the sum of the products of the input coordinate magnitudes and the marginal costs. Theoretically, an infinite number of trial output vectors would produce points covering the entire area below and to the right of the curve in Fig. 4-15. Because efficient outputs have the maximum value for a given input value, they lie along the locus of points bounding the area on the upper left side. These are also points which meet the conditions of the second and third rules. The optimum input and output vectors are found by the same rule used to determine the optimum level of output in Fig. 4-8.

The complexity of design optimization can be appreciated when one realizes that Eqs. (4-22) to (4-24) only help establish whether a particular project design is in fact optimum. The development of a trial project design requires a long series of economy studies seeking the least-cost means of achieving the desired end, or what economists would call *defining the production function*. Project formulation is a trial-and-error process in which promising designs are tested to determine the resulting net benefits.

While the rules developed in the preceding sections are not even necessary for project optimality if the response surface peaks adjacent to one of the axes (Fig. 4-7) and while they are not in themselves able to guarantee project optimality (Sec. 4-5), neither of these limitations is likely to be encountered in water resources project design. A practical problem of more general consequence is likely to arise in applying optimality criteria where intangible values or outside constraints dictate the project design. Is there any advantage in applying any of the optimality criteria where all of them cannot be satisfied? Is it worthwhile to devote so much effort to optimizing water resources development while optimality criteria are ignored in virtually every other sector of public spending?

Lipsey and Lancaster have developed a "general theory of the second best" showing that improvement in one sector will not invariably produce an overall improvement.[1] Nevertheless, the greater probability is that projects designed more efficiently with respect to as many optimality criteria as can be applied will increase human welfare in the short run. As new techniques are developed and implemented in other kinds of public investment decisions, the chance for improvement is even greater in the long run.

Market Allocation under Pure Competition

The derived optimality conditions would be automatically achieved by a market economy under the ideal conditions of pure competition. This can be shown by analysis of production situations in the very short run, short run, and long run.

4-13 VERY SHORT RUN ANALYSIS In the very short run, the output has already been produced and its amount is fixed. Since no freedom is left to alter design and production decisions, none of the optimality criteria for governing such choices applies. The producer tries to sell at the best price available as long as one can be found in excess of the value of the output as scrap or in excess of the present worth of the net profit expected from storage for later sale. However, a price below the marginal production cost would direct that production be halted as soon as possible.

4-14 SHORT-RUN ANALYSIS In the short run, the firm is free to vary the level of production in response to market conditions. However, an individual firm does not have time to vary the capacity of its production facilities. Industry output can only vary within the capacity of existing firms.

The Firm A firm producing an output in relatively small quantities cannot affect the price by changing the supply and hence faces a horizontal demand curve (Fig. 4-16). The demand curve is the firm's marginal-revenue curve since the incremental revenue from each additional unit of output equals the unit price. Profits are a maximum at the output Y'

[1] Richard Lipsey and Kevin Lancaster, The General Theory of the Second Best, *Rev. Econ. Studies*, vol. 24 (December, 1956).

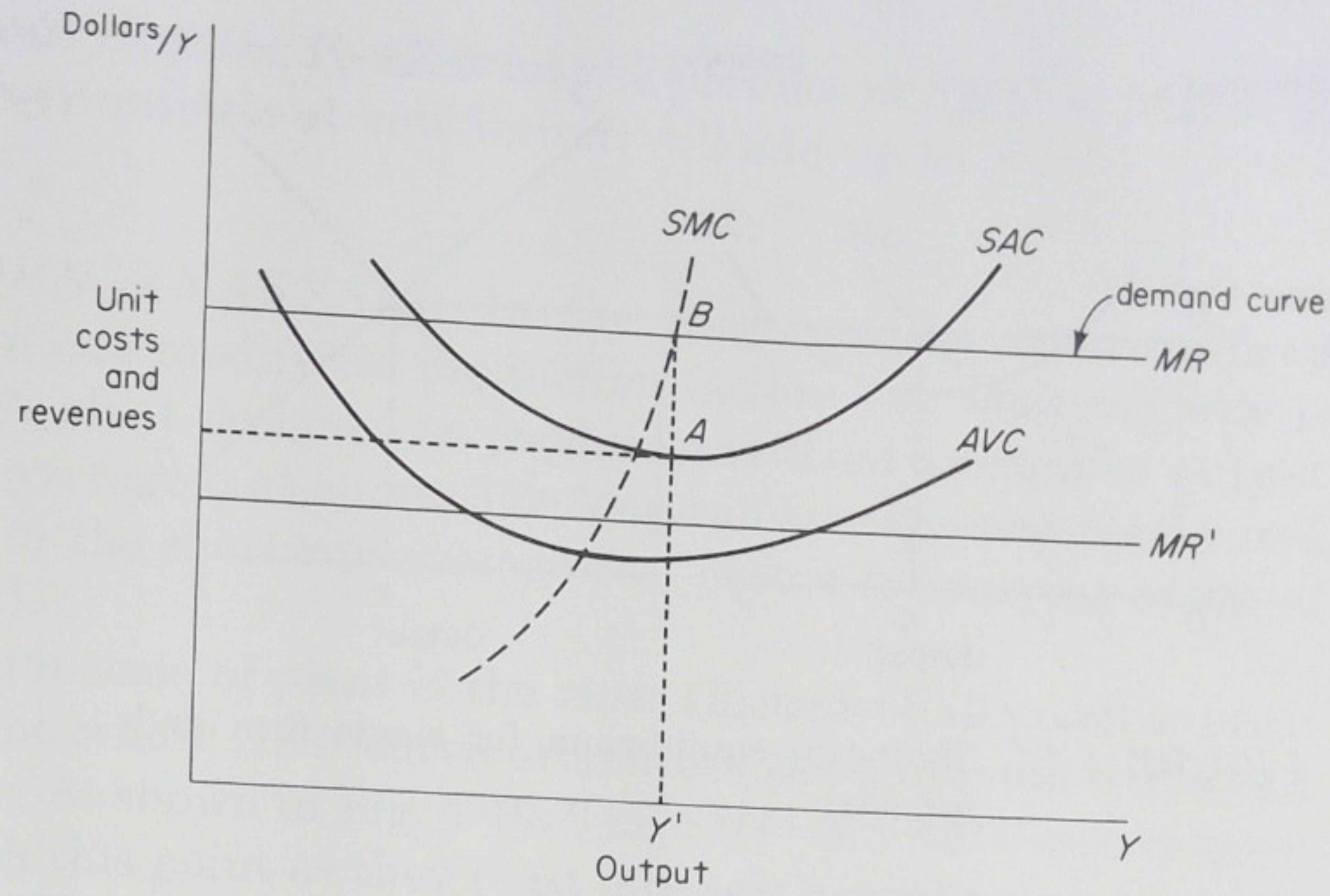


FIGURE 4-16 The firm's short-run unit-revenue and unit-cost curves

where $SMC = MR$. Here SMC is the short-run marginal cost, or the variable cost per marginal unit of production from a fixed plant, and MR is the marginal revenue.

If the market price (MR' in Fig. 4-16) is less than short-run average costs SAC at all possible levels of output, the firm will lose money. However, it should continue to operate at a loss as long as the market price exceeds the average variable costs AVC. This is because revenues exceed the cost of production and can be used to partially defray fixed costs which continue regardless of whether the plant shuts down or not.

The profit-maximizing firm should produce the output for which marginal cost equals market price ($MC = MR$) unless market price falls below the firm's average variable costs. In that event it should shut down. The firm's short-run supply curve is that part of the SMC curve which lies above the AVC curve.

The Industry The price faced by the firm is determined by the composite supply and demand curves faced by the group of firms comprising the industry. The short-run industry-supply curve is the horizontal summation of the firms' short-run-supply curves as long as production input prices are not affected by the industry. If input prices are affected, firms' unit-cost curves will shift and cause some adjustment in the industry supply curve.

Figure 4-17 shows how prices are signaled by the industry to the

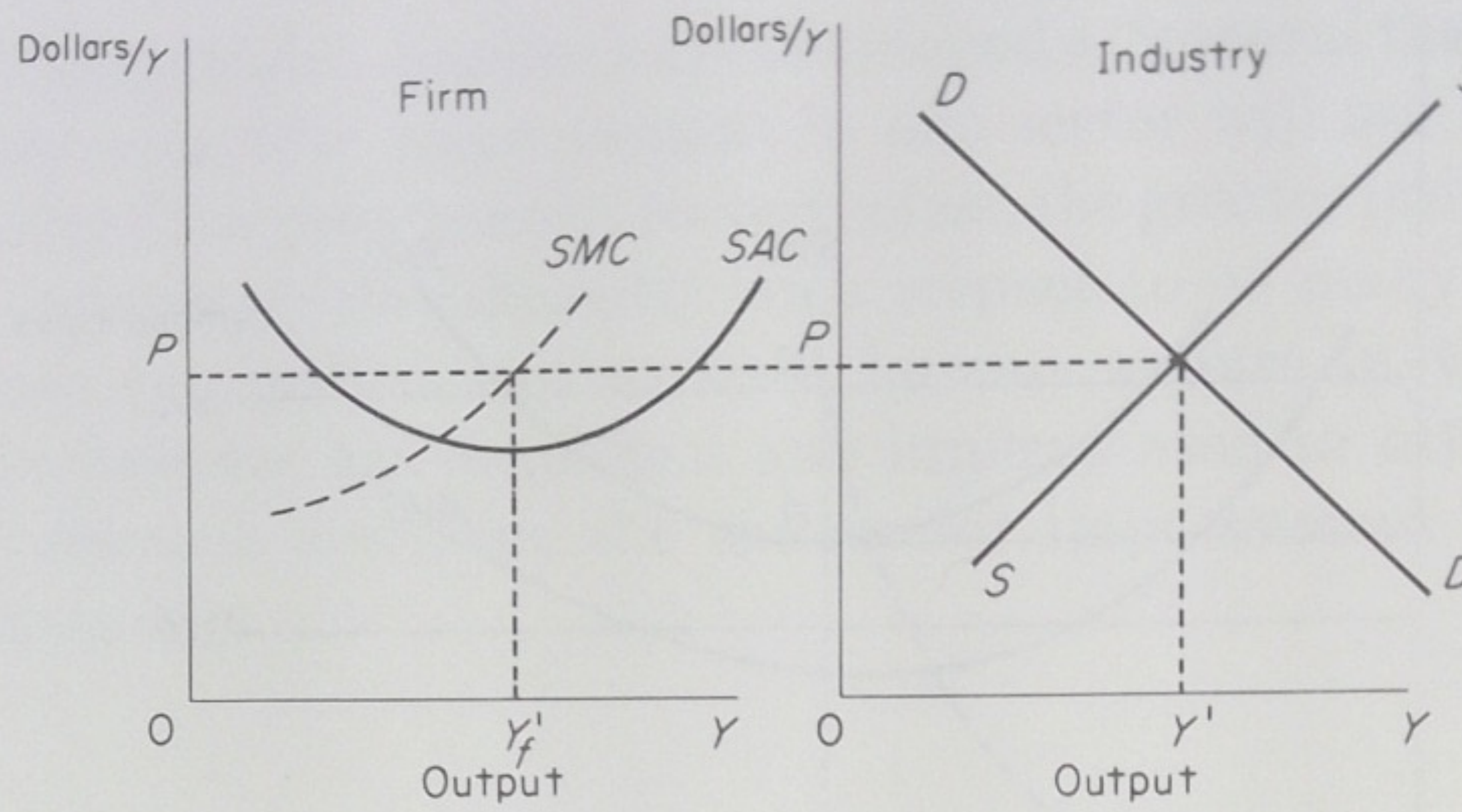


FIGURE 4-17 Short-run equilibrium for single firm and industry

firm. The horizontal axis for the industry is greatly compressed since the output of the firm is very small compared with the output for the industry as a whole. The price axes are the same. Suppose we have the industry demand *DD*, which is the horizontal summation of the consumer-demand curves, and the industry supply *SS*, which is the horizontal summation of the firms' supply curves. The short-run equilibrium market price *P* is established by the interaction of *DD* and *SS*. This price becomes the horizontal demand curve for the firm because it can sell as much as it

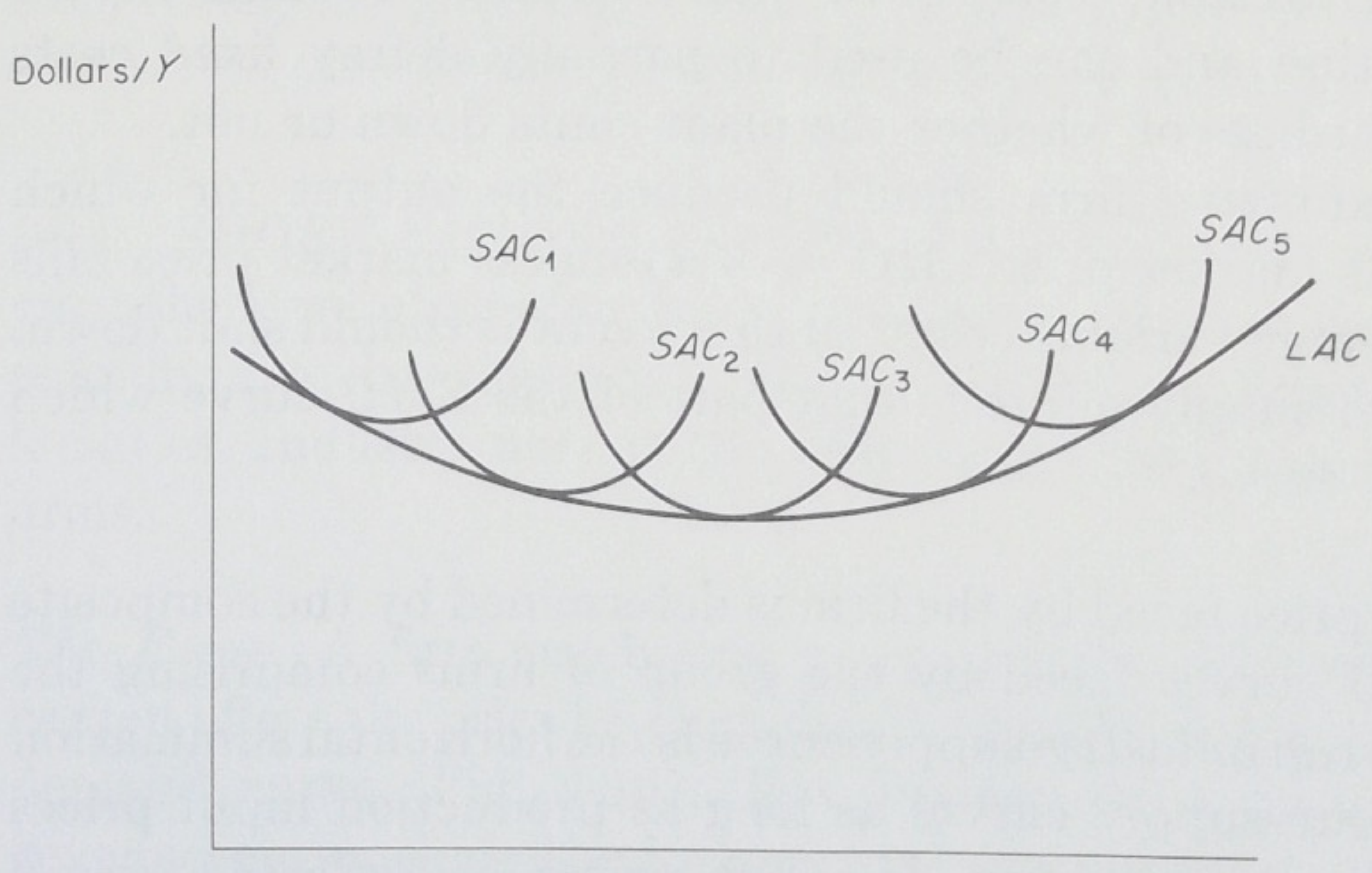


FIGURE 4-18 Long-run average-cost curves.

can make without affecting P. All firms will produce at $SMC = MR = P$, and together their outputs at equilibrium will add up to Y'.

4-15 LONG-RUN ANALYSIS In the long run, no costs are fixed because the firm can modify its production facilities in whatever way is advantageous. A plant designed to produce any fixed amount of output has a short-run average-cost curve. The long-run average-cost LAC curve is the envelope of the short-run average-cost curves for varying scales of plant (Fig. 4-18).

The optimum scale of plant is the most efficient of all possible plant sizes. It is the one whose SAC curve establishes the minimum-cost point of the LAC curve as shown in Fig. 4-19. The SMC and LMC curves must also pass through this point as they must intersect average-cost curves at their low point. The long-run profit-maximizing rule is for a firm to produce the output for which $LMC = LAC = SAC = SMC$. Each firm must operate an optimum scale of plant at the optimum rate of output. Free entry and exit of firms will restrict production to those firms which have their average-cost curves tangent to the demand line DD. Under pure competition, consumers get products at prices equal to long-run average costs.

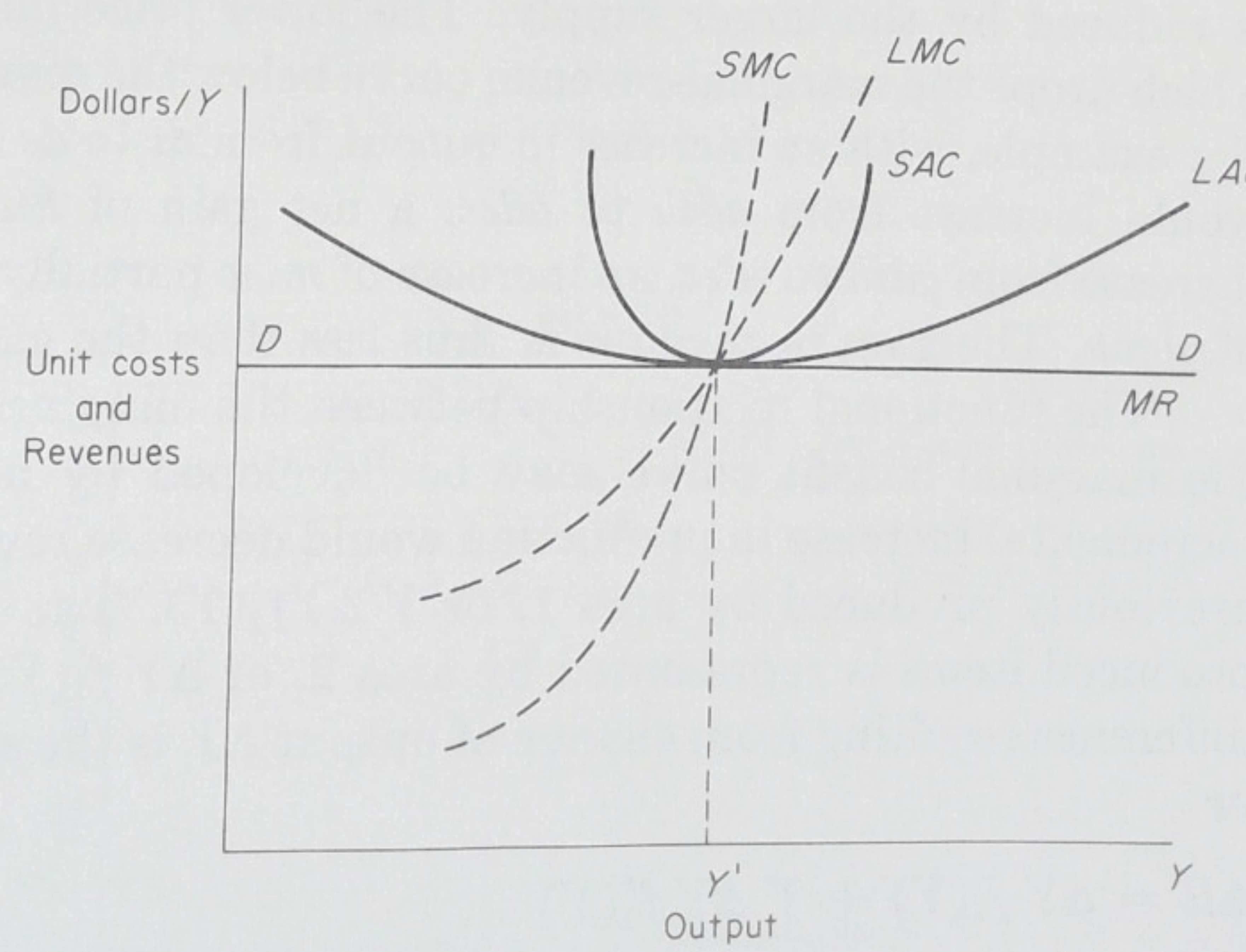


FIGURE 4-19 Long-run cost curves and determination of optimum scale of plant and output.

Applications

4-16 USE OF SUPPLY AND DEMAND CURVES IN PROJECT EVALUATION Application of the principles developed in this chapter to the economic evaluation of engineering projects requires analysis of the demand for project output and of the cost of project production. Long-run values should be used to optimize design in project planning, and short-run values should be used to optimize operation of an existing facility. The demand may be graphically depicted by a marginal-benefit curve, and the supply may be depicted by a marginal-cost curve. Potential-output uses vary greatly in value. Whenever all available output is devoted to the most beneficial use available, marginal benefit monotonically decreases with increasing output. Each increment of output is put to a lower-value use than the previous increment as the higher-value uses are satisfied. Expressed mathematically,

$$MB = f_1(Y) \qquad (4\text{-}25)$$

or marginal benefit is functionally related to the level of output Y.

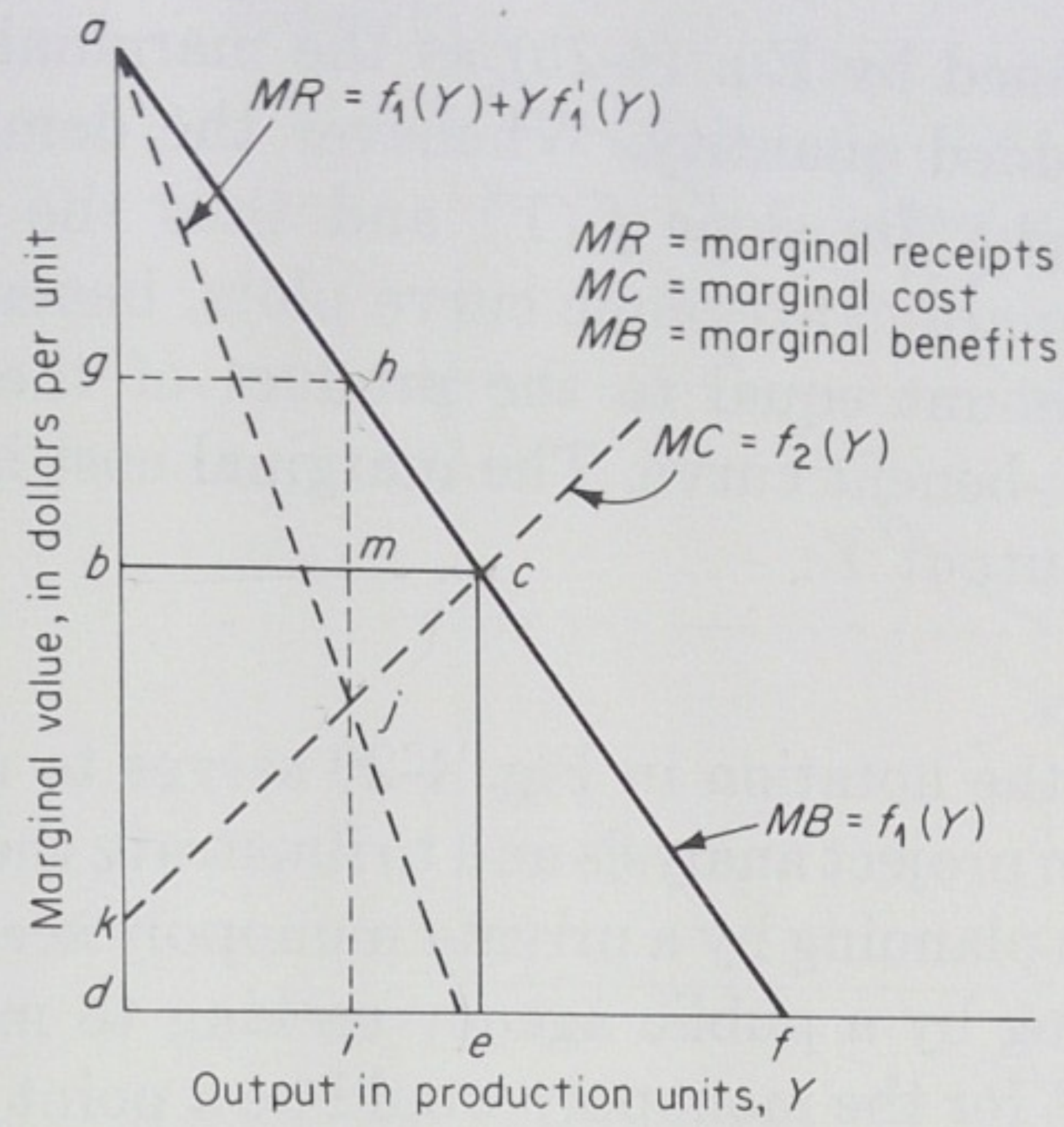
The marginal revenue which could be realized from sale of additional output would be less than the marginal benefit. Revenue depends on the sale price, while benefit is received through use and is thus determined by the nature of the use. When production increases, the total benefit is increased by the area added under the marginal-benefit curve. The price is reduced by the larger supply. The lower price inflicts a revenue loss which drops the marginal-revenue curve below the marginal-benefit curve. For example, with an increase in output from *di* to *de* (Fig. 4-20), benefits would increase from *adih* to *adec*, a net gain of *hiec*. Revenues would increase from *gdih* to *bdec*, an increase of *miec* partially offset by a decrease of *gbmh*. The gain in revenue is thus less than the gain in benefit.

The functional relationship between the marginal-revenue curve and the marginal-benefit curve may be developed by use of Fig. 4-21. An incremental increase in production would decrease revenue from the items previously produced by area 1, or $Y\,\Delta Y f_1'(Y)$. The revenue from newly produced items is represented by area 2, or $\Delta Y\, f_1(Y)$. The total revenue difference resulting from change of output ΔY is the sum of the two areas, or

$$\Delta R = \Delta Y\, f_1(Y) + Y\,\Delta Y\, f_1'(Y) \qquad (4\text{-}26)$$

Marginal revenue is determined by decreasing ΔY to differential size to get

$$MR = \frac{dR}{dY} = f_1(Y) + Yf_1'(Y) \qquad (4\text{-}27)$$



	Private monopoly		Public works projects
Goal	Maximize net receipts		Maximize net benefits
Optimum point	*j*, *MR* = *MC*		*c*, *MB* = *MC*
At optimum point: gross receipts	*gdih*		*bdec*
Production	*di*		*de*
Price	*hi*		*ce*
Production cost	*kdij*		*kdec*
Net receipts (producer's surplus)	*gkjh*	>	*bkc*
Gross benefits	*adih*		*adec*
Net benefits	*akjh*	<	*akc*
Consumer's surplus	*agh*		*abc*

FIGURE 4-20 Definition of terms in benefit computations.

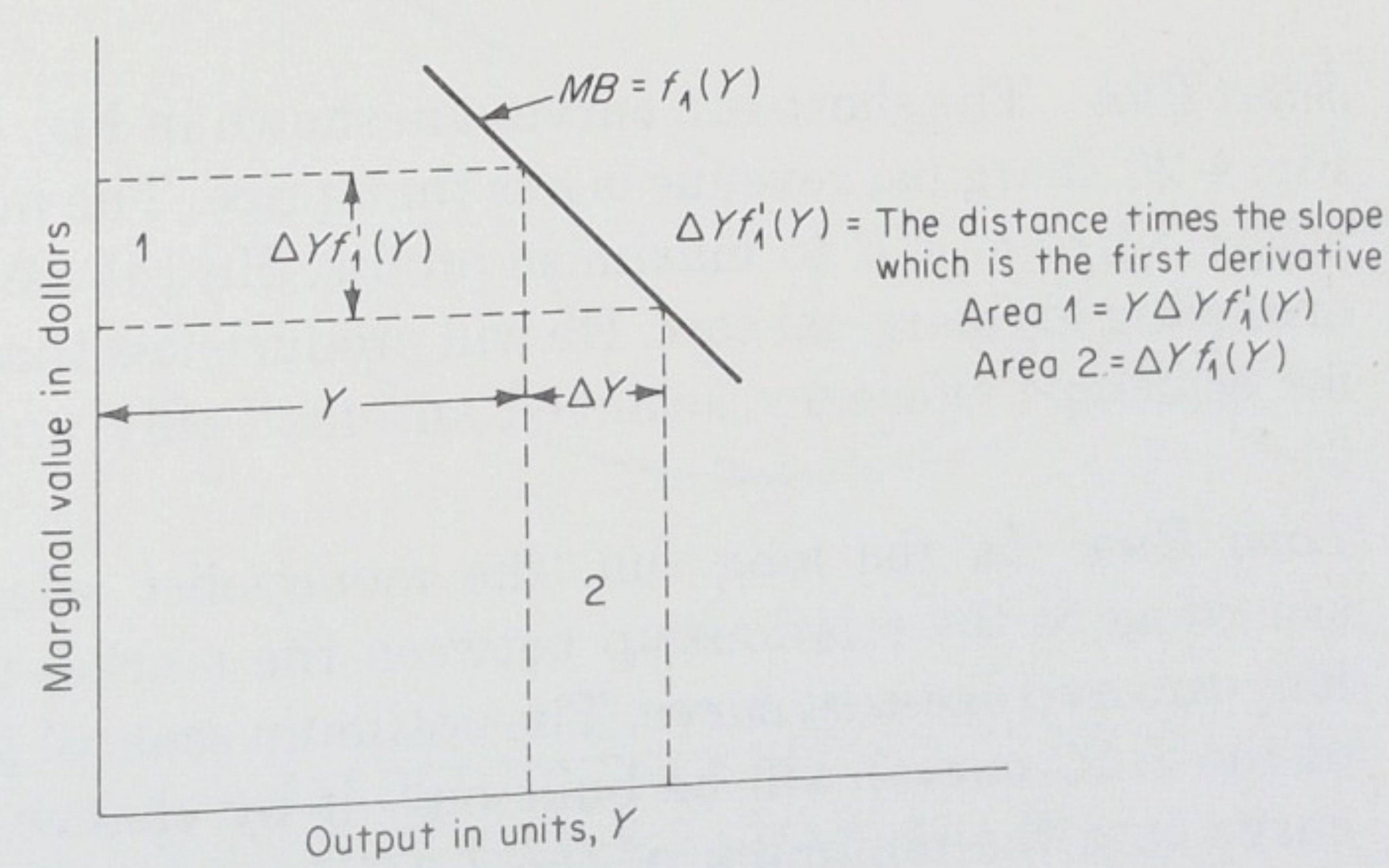


FIGURE 4-21 Effect of incremental change in output on marginal revenue.

The first term is defined by Eq. (4-25) as the marginal benefit, and the second term is an added quantity. Whenever the demand curve slopes downward to the right, its slope $f_1'(Y)$ and thus the second term are negative. Thus, the marginal-revenue curve plots, beneath the marginal-benefit curve, an amount equal to the product of the output and the slope of the marginal-benefit curve. The marginal cost is also some function of the level of output Y:

$$MC = f_2(Y) \tag{4-28}$$

The balance of the notation in Fig. 4-20 serves to define the benefit and cost terms used in project analysis and to illustrate the difference which would result between planning by a private monopoly seeking to maximize revenues and planning by a public agency seeking to maximize benefits. The optimum project for the monopoly would be a point j where marginal revenue equals marginal cost. The optimum project for the public agency would be at point c where marginal benefit equals marginal cost. At each optimum point, the optimum production, gross receipts, production cost, net receipts, producer's surplus, gross benefits, net benefits, and consumer's surplus are defined as shown.

4-17 MARKET ALLOCATION UNDER PURE MONOPOLY Pure monopoly is the market situation where a firm produces the entire supply of an output for which there are no good substitutes, a rare situation because most outputs have substitutes. The monopolist has the same cost curves as the firm in pure competition but faces the entire market-demand curve. His actions affect price as well as the economic equilibrium achieved in the short run as well as the long run.

Short Run The short-run curves are shown in Fig. 4-22. As illustrated by Fig. 4-20, marginal revenue is less than price. The monopolist will produce where $SMC = MR$ to maximize profits. He will be able to sell at a price exceeding his marginal cost. He will produce less than the optimum output for economic efficiency achieved automatically under pure competition.

Long Run In the long run, the monopolist selects his scale of plant according to the relationship between the market-demand curve and his long-run average-cost curve. The optimum scale of plant (at the minimum of the LAC curve) will be built only if by chance the marginal-revenue curve cuts the minimum of the LAC curve as shown in Fig. 4-23. The monopolist will follow the profit-maximizing rule of $LMC = MR$. Since $SMC = LMC = MR = SAC = LAC$ at output Y', the firm is in both

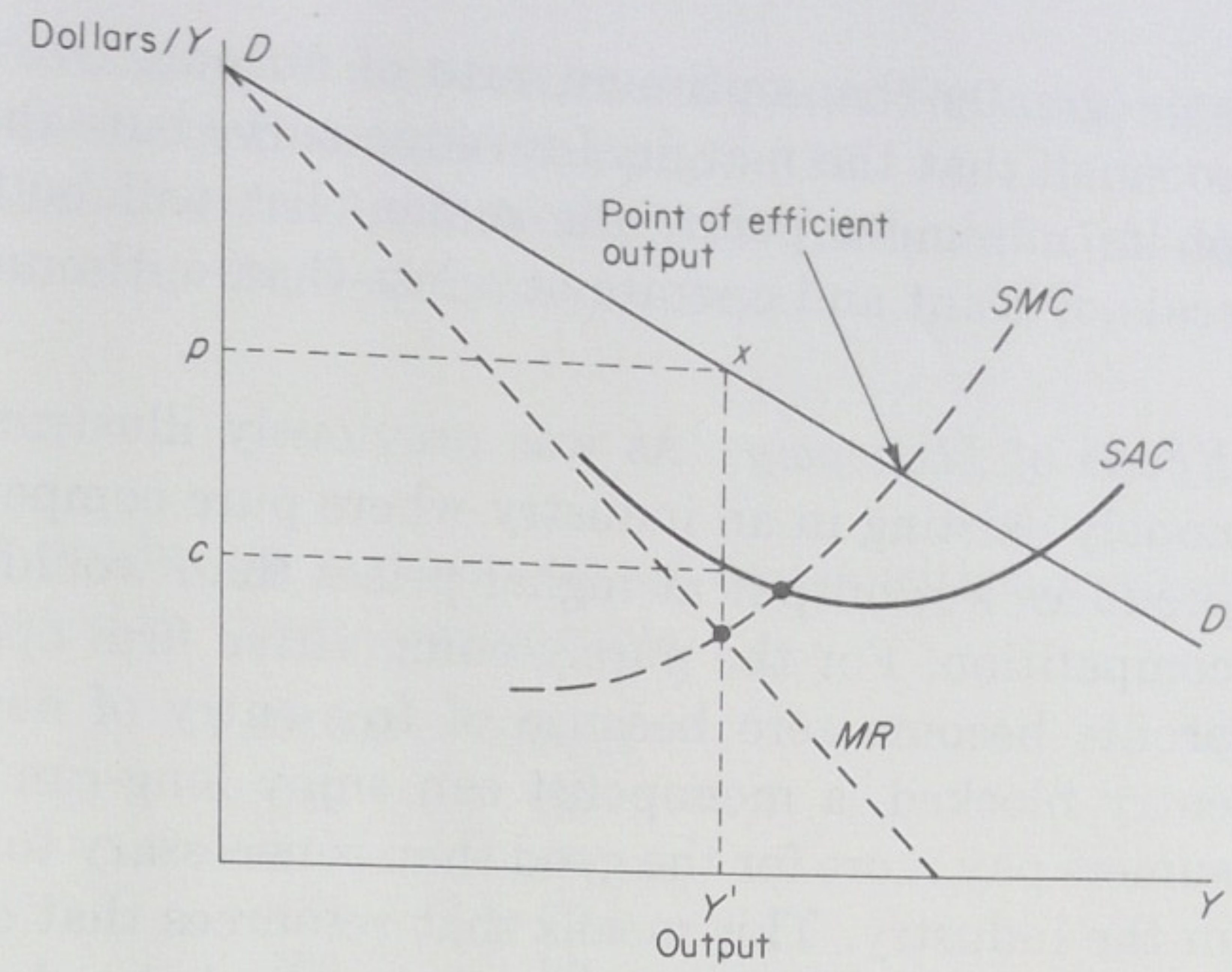


FIGURE 4-22 Unit curves for pure monopoly.

short-run and long-run equilibrium. The monopolist's profit is $(p - c)Y'$. He operates an optimum scale of plant at an optimum rate of output.

However, if the market is so large that the monopolist's marginal-revenue curve cuts the *LAC* curve to the right of its minimum point, the monopolist will build a larger-than-optimum scale of plant and operate

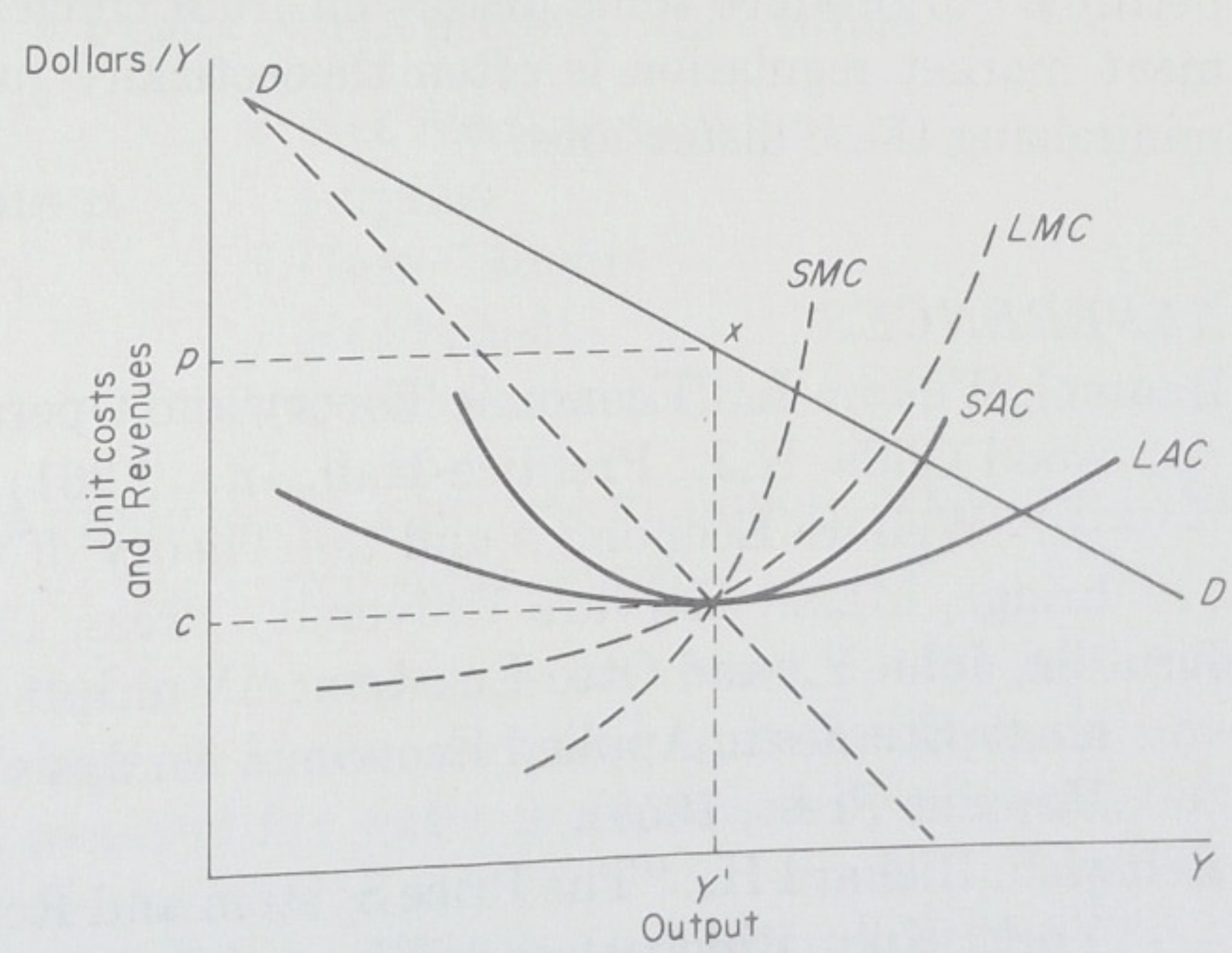


FIGURE 4-23 Monopoly situation for optimum scale of plant.

at a greater-than-optimum rate of output. Obversely, if the market is so small that the marginal-revenue curve cuts the *LAC* curve to the left of its minimum point, the monopolist will build a less-than-optimum scale of plant and operate at a less-than-optimum rate of output.

Effects of Monopoly As was previously illustrated by Fig. 4-20, a monopoly existing in an industry where pure competition would be possible produces less output at higher prices than would be the case under pure competition. For the purely competitive firm operating in the long run, profits become zero because of free entry of new firms. However, with entry blocked, a monopolist can enjoy long-run profits. Therefore, consumers pay more for the good than is necessary to induce resources to stay in the industry. This means that resources that could be used to expand output of the desired good are being used elsewhere in lower-valued uses.

It is also possible in an industry where the market is limited relative to the optimum scale of plant and rate of output for monopoly to result in lower costs per unit than under pure competition. An example would be a large thermoelectric plant in a small country. The unit cost of electricity produced by a monopoly may be considerably less than that for a number of firms producing with less than optimum scales of plant and rates of output.

Either way, monopoly distorts the efficiency equilibrium achieved automatically under pure competition to the degree the monopolist does not operate an optimum scale of plant at the optimum rate of output. A partial monopoly or oligopoly or any other departure from pure competition will produce some distortion from efficiency conditions. Government market regulation is often theoretically justified with the goal of minimizing these distortions.

REFERENCES

Baumol, William J.: "Economic Theory and Operations Analysis" (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1961).

———: "Welfare Economics and the Theory of the State," 2d ed. (Cambridge, Mass.: Harvard University Press, 1967).

Krutilla, John V., and Otto Eckstein: "Multiple Purpose River Development: Studies in Applied Economic Analysis" (Baltimore: The Johns Hopkins Press, 1958).

Leftwich, Richard H.: "The Price System and Resource Allocation" (New York: Holt, Rinehart and Winston, Inc., 1962).

Lipsey, Richard, and Kevin Lancaster: The General Theory of the Second Best, *Rev. Econ. Studies*, vol. 24 (December, 1956).

Maass, Arthur, et al.: "Design of Water-resource Systems" (Cambridge, Mass.: Harvard University Press, 1962), pp. 88–117.

Samuelson, Paul A.: "Economics: An Introductory Analysis," 5th ed. (New York: McGraw-Hill Book Company, 1962).

PROBLEMS

4-1 The total cost of producing a given output varies with output in the manner indicated in the following table:

Total cost	*Output*	*Total cost*	*Output*
0	0	385	50
200	10	410	60
300	20	460	70
350	30	560	80
375	40	760	90

a Plot the total-cost curve.
b Plot the average-cost curve.
c Plot the marginal-cost curve.
d If under conditions of pure competition the price of the good is 5, what output will the firm produce in the short run?
e What will be the net profit or loss at this price?
f At what price will the firm just break even?

4-2 For a particular production enterprise, marginal benefits can be related to output by the equation $B + Y = 20$. Marginal costs can be related to output by the equation $C = 2 + 1.5Y$. If the output were produced in a public works project, what would be the:

a Price
b Quantity of output produced
c Gross receipts
d Production cost
e Net receipts (producer's surplus)
f Gross benefits
g Net benefits
h Consumer's surplus

If the output were produced by a private monopoly, what would be the values of each of the above eight items?

4-3 The marginal cost of water supply is expressed by the equation $5P + 5Q = 30$, where P is the cost and Q is the amount produced. The price received for water is expressed by the equation $3P + 6Q = 30$.
a How much water would a private monopoly produce?
b What is the cost of producing the amount of water in part *a*?
c How much revenue would result from selling this much water?

d How much water would a public works project produce?
e What is the cost of producing the amount of water in part *d*?
f How much revenue would result from selling this much water?

4-4 The demand curve for a particular good is expressed by the equation $2P + 3Y = 15$, where P is the price and Y is the quantity of the good. The curve for another good is expressed by the equation $4P + 3Z = 12$.
a At what price could two units of the first good be sold?
b What is the elasticity of this good at this price?
c At what price is the elasticity of this good equal to 1?
d Plot the aggregate-demand curve for the two goods if they are both market goods.
e Plot the aggregate-demand curve for the two goods if they are both collective goods (Sec. 5-9).

4-5 The demand curve for gadgets is expressed by the equation $2P + 5G = 20$. The number of G available for sale is 1.5.
a What will the price be?
b What will the total revenue be?
c What will the consumer's surplus be?
d What will the aggregate value in use be?

4-6 Use a Lagrange multiplier to find the maximum value of $y = 10xw - 4w^2$ subject to the constraint $x + w = 15$. What would the maximum value be were the constraint $x + w = 16$?

4-7 Gambies are used in the production of whoozits. A factory can sell a whoozit for \$10 and buy a gamby for \$4. The number of whoozits produced can be increased by using more gambies as follows:

Whoozits:	0	10	20	30	40	50	60	70	80
Gambies:	0	2	4	8	16	32	64	128	256

a What should be the marginal physical product of gambies in producing whoozits? (Fractional answers are legal, and graphical solution is quicker and more accurate for this and the two following problems.)
b How many whoozits should the factory produce?
c How many gambies should be used in their production?
d What will the profit be?

4-8 Xyphos are also used in the production of whoozits. The factory can buy a xypho for \$2. The relationship between gambies and xyphos in producing a fixed number of whoozits is as follows:

Xyphos:	15	10	6	3	1
Gambies:	1	3	6	10	15

a What should be the marginal rate of substitution of gambies for xyphos?
b How many gambies should be used?
c How many xyphos should be used?
d What is the total cost of gambies and xyphos used in production?

4-9 Gambies and xyphos can be used to produce whidgits as well as whoozits. Whidgits can be sold for $6. The relationship between the number of whidgits and the number of whoozits which can be produced from a fixed input is as follows:

Whidgits:	20	19	17	14	10	5	0
Whoozits:	0	5	9	12	14	16	17

a What should be the marginal rate of transformation of whoozits for whidgits?
b How many whoozits should be produced?
c How many whidgits should be produced?
d What is the total revenue from sales of whoozits and whidgits?

4-10 The production function for a two-input two-output process is $Y_a^2 + Y_b^3 = X_1^{1/2}X_2^{1/3}$.
a Determine the optimum values for both inputs and both outputs if $MB_a = MB_b = MC_1 = MC_2 = 1$.
b Determine the optimum values for both inputs and both outputs if $MB_a = MB_b = 10$ and $MC_1 = MC_2 = 1$.