Chapter 3
The Neoclassical Theory of International Trade

Before discussing the neoclassical model of international trade (Sect. 3.3 and following), it is advisable to recall from microeconomics some widely-used diagrammatic tools (Sect. 3.2) and to show how the general equilibrium of production and consumption is determined in a simple closed economy (Sect. 3.2), where two final goods (A and B) are produced by the full employment of two *primary* factors of production (K and L). The problems deriving from the presence of *produced* factors of production will be tackled in Sects. 6.4, 6.4.1 and 14.1.

The given data are:

(a) The total amounts of the two factors existing in the economy;
(b) The distribution of these among the members of the economy, namely the amounts of K and L owned by each member;
(c) The tastes of consumers;
(d) The state of technology, represented by well-behaved aggregate production functions (a "well-behaved" production function shows constant returns to scale and has positive but decreasing marginal productivities: see Sect. 19.1.3).

Perfect competition obtains in all markets (commodities and factors).

3.1 The Transformation Curve and the Box Diagram

The tools that we wish to recall are the Haberler-Viner-Lerner-Leontief *product transformation curve* (otherwise known as the *production-possibility curve* or *production-possibility frontier*) and the Edgeworth-Bowley *box diagram* (originally intended to derive the contract curve between two consumers and applied to production problems by Lerner and Stolper-Samuelson); for a detailed analysis of "who was the first" and references, see Savosnick (1958).

The product-transformation curve (henceforth called the *transformation curve*) represents the maximum amount of one commodity obtainable for any given amount

of the other. This requires that the given fixed amounts of productive factors are optimally allocated between the two commodities in accordance with certain marginal productivity conditions which are easily found by using the *box diagram*.

3.1.1   The Box Diagram

In Fig. 3.1, the length of the sides of the box represents the total amounts of the two productive factors existing in the economy, respectively $O_A G = O_B H$ of labour and $O_A H = O_B G$ of capital.

The isoquants concerning the production of commodity A are drawn with reference to the origin O_A, and the isoquants concerning the production of commodity B are drawn with reference to the origin O_B and so appear upside down. In fact, the box can be considered as obtained by first drawing the isoquant maps for the two commodities in the usual way—with the proviso that the lengths of the axes are equal—and then turning one of the two upside down so that the extremes of the axes (points H and G) coincide. Both isoquant maps have the usual properties.

Let us now find the condition of efficiency in production, also called Pareto optimality in the producing sectors. By efficiency in the producing sectors we mean a situation in which—on the assumption of full employment of all factors of production—these factors are allocated between the two commodities in such a way that, given the output of one commodity, the output of the other is maximized. An equivalent definition is that it is not possible, by reallocating the given fixed amounts of productive factors, to increase the output of one commodity without decreasing that of the other. It is clear that if instead it is possible, by means of such a reallocation, to increase the output of one commodity while keeping the output of the other constant, then the situation is inefficient. It can be proved graphically that the condition for efficiency is that the A isoquants and the B isoquants are tangent (for simplicity's sake we neglect possible corner solutions), namely that the marginal rates of technical substitution (MRTS) are equal in the two productive sectors.

For this purpose, consider for example point Q in Fig. 3.1. This point lies at the intersection of isoquant I_A (concerning the production of commodity A) with isoquant I_B (concerning the production of commodity B). The allocation of the productive factors can be read by drawing the coordinates of Q on the sides of the box, which gives $O_A L_A$ of labour to (the production of) commodity A and $L_A G = O_B L_B$ to commodity B, and similarly $O_A K_A$ and $K_A H = O_B K_B$ of capital to commodities A and B respectively. If we connect point Q to the origins by means of two straight line segments, we can read the factor intensities as the slopes of these segments: for example, $\tan \alpha = O_A K_A / O_A L_A$ is the capital/labour ratio in the A sector and $\tan \beta = O_B K_B / O_B L_B$ is the capital/labour ratio in the B sector.

Point Q is not efficient: in fact, by reallocating the productive factors it is possible to move for example to Q' on I'_B while still remaining on I_A; point Q' gives

Fig. 3.1 The box diagram and the efficiency locus
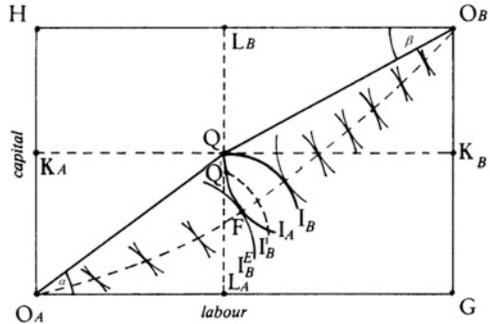
a greater output of commodity B (isoquant I'_B being farther from the origin O_B than I_B, represents a greater output). Continuing in this manner we arrive at the point of tangency F, which corresponds to the highest B isoquant $\left(I_B^E\right)$ achievable, given I_A, namely at F we get the maximum output of B given the output of A.

Further increases in B can be obtained only at the expense of a reduction in A, therefore F is an efficient point. The (optimal) allocation of the productive factors between the two sectors and the corresponding factor intensities can be read as shown above with reference to the non-optimal point Q.

The locus of all such points of tangency is the efficiency locus we are looking for; it is also (improperly) called *contract curve* (this was the original Edgeworth-Bowley denomination, but with reference to consumers' exchange).

3.1.2 The Transformation Curve and Its Properties

The passage from the efficiency locus to the transformation curve is simple: it is sufficient to transfer the indexes attached to each couple of tangent isoquants (these indexes are numbers representing quantities of the two commodities) to the coordinate axes in the *(A,B)* plane. In this way (Fig. 3.2) we obtain a diagram showing the maximum amount of B obtainable for any given amount of A, namely the transformation curve. Since the maximization procedure is perfectly symmetric, the efficiency locus and the transformation curve are the same if we maximize the output of A for any given amount of B. An alternative procedure for deriving the transformation curve from the box diagram is represented in Fig. 3.3 (Savosnick, 1958), which is similar to Fig. 3.1 except that now the right-hand vertical side of the box is used to measure the output of commodity A and the lower horizontal side is used to measure the output of commodity B. For simplicity's sake we assume constant returns to scale in both commodities. In this case, as we know from the properties of production functions homogeneous of the first degree (see Sect. 19.1.3), an isoquant which intersects a straight line through the origin twice as far away as another isoquant will represent twice as large an output. If we take
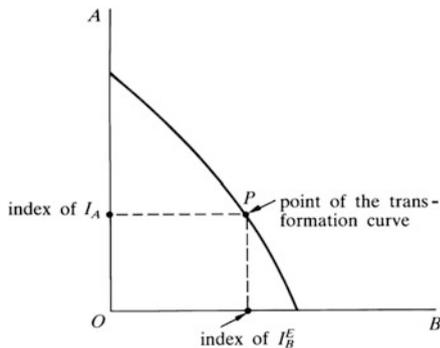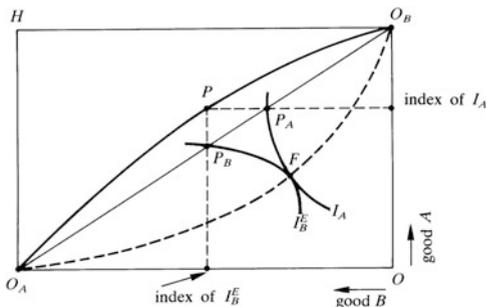
Fig. 3.2 The transformation
curve

Fig. 3.3 The transformation
curve derived from the box
diagram

the diagonal of the box (straight line $O_A O_B$) as such a straight line, we can use it
to project the outputs on the output axes; these projections will correspond exactly
to the relationship just mentioned, and each output axis will have a uniform scale.

Consider for example point F. The isoquant I_A intersects the diagonal at point
P_A, whose projection on the A axis gives the index of I_A; similarly, isoquant I_B^E
intersects the diagonal at point P_B, whose projection on the B axis gives the index
of I_B^E. Therefore point P, which has these projections as coordinates, is a point of
the transformation curve. In this way we obtain the transformation curve $O_A P O_B$
which is the same—apart from scale factors and position—as the curve in Fig. 3.2.
This construction also illustrates the one-to-one correspondence between points
on the efficiency locus and the transformation curve: to every point on this curve
representing an output combination there corresponds a point on the efficiency locus
representing an input combination, and vice versa.

With both production functions exhibiting constant returns to scale, the efficiency
locus must lie on one side of the diagonal of the box diagram and can never
cross it, although locus and diagonal may coincide. In fact, when a point of the
efficiency locus lies on the diagonal, then the whole efficiency locus coincides
with the diagonal itself. This follows from the fact that with constant returns to scale
the marginal rate of technical substitution is constant along a straight line through
the origin. Therefore, if the MRTS of an A isoquant is equal to the MRTS of a B
isoquant at a point on the diagonal of the box, then these MRTS remain the same
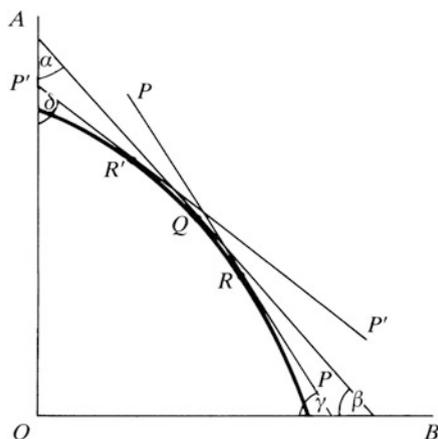
Fig. 3.4 Marginal rate of transformation, opportunity cost, and relative prices

along the diagonal, and if they are equal at one point they must be equal everywhere (in such a special case the capital-labour ratio is the same for both commodities). On the other hand, if the MRTS of an A isoquant and the MRTS of a B isoquant are different at some point on the diagonal, then they must be different at all other points on the diagonal.

The transformation curve can in principle be either convex or concave or both, but with constant returns to scale in both sectors, it will always be strictly concave to the origin if we exclude the particular case of identical capital-labour ratios, just dealt with (in which case it will be linear). This follows immediately from the graphic construction given in Fig. 3.3 and from the property that the efficiency locus must lie on one side of the diagonal of the box diagram (note that if it were all above the diagonal instead than below it, we would measure the output of commodity A on the left-hand vertical side of the box and the output of commodity B on the upper horizontal side, and would obtain a transformation curve concave to the origin now given by point H). Other simple graphic proofs can be found, for example, in Chacholiades (1978, pp. 107–109) and in Findlay (1970, pp. 26–29). In Sect. 19.1.2 we give a general mathematical proof in which we also consider the properties of transformation curves derived from production functions which do not possess constant returns to scale.

The (absolute value of the) slope of the transformation curve (for example $\tan \beta$ in Fig. 3.4) is called the *marginal rate of transformation* or (marginal) *opportunity cost* of B in terms of A, namely the amount of A that the economy has to give up to obtain an additional unit of B. It should be noted that this notion of opportunity cost has general validity, independently of the theoretical frame of reference. For example, in the Ricardian theory treated in Chap. 2, it is possible to identify the opportunity cost with the comparative cost.

It goes without saying that the opportunity cost of A in terms of B is measured by the (absolute value of the) slope of the transformation curve with reference to the A axis, namely $\tan \alpha$ in Fig. 3.4. Note that, since the transformation curve is derived from an optimizing procedure, the amount of A (or of B) given up is the *minimum* possible under the given technical knowledge. The concavity of the transformation curve implies that its slope increases as we move along it to the right, i.e. the opportunity cost of B increases as more of it is produced.

A fundamental proposition is that, under competitive conditions, the economy will always operate on the transformation curve, at a point where the marginal rate of transformation equals the price ratio or relative price of the two commodities p_B/p_A.

To prove the first part of the proposition it is sufficient to show that pure competition will bring producers onto the efficiency locus. Cost minimization requires that the MRTS in each sector is equated to the factor-price ratio, and since with perfect factor mobility the price of a factor is the same everywhere, it follows that the MRTS is the same in both sectors, which is the condition of efficiency.

To prove the second part of the proposition it suffices to show that profit maximization requires equality between opportunity cost and commodity-price ratio. Suppose, for example, that the economy is at point Q while the relative price is indicated by the slope of the line PP. This means that the opportunity cost of producing more B is lower than its price, $\tan \beta < \tan \gamma$ (note that this comparison makes sense because both the opportunity cost and the relative price under consideration are dimensionally homogeneous, being measured in terms of commodity A as *numéraire* or unit of measurement). It follows that producers can increase their profits by increasing the output of B. Only at R are the opportunity cost and the relative price equal and profits maximized. Similarly, if the relative price were given by the slope of the line $P'P'$, the opportunity cost of producing more A ($\tan \alpha$) would be lower than the relative price of A (p_A/p_B is measured by $\tan \delta$), and producers would maximize their profits by moving to point R'.

Another illuminating way of proving the equality between the marginal rate of transformation and the commodity price ratio is to pass through marginal costs. Suppose that we move slightly to the right on the transformation curve, thus increasing the output of commodity B and decreasing that of commodity A. If we consider a small displacement, an amount dK of capital and dL of labour will be transferred from sector A to sector B, and the additional cost in producing B is dC_B, where of course

$$dC_B = p_K dK + p_L dL. \tag{3.1}$$

Since we are moving on the transformation curve and therefore along the efficiency locus, the prices of productive factors must be equal in both sectors. Therefore the additional cost in producing B must be equal to the reduction of the cost in producing A, namely

$$dC_B = -dC_A. \tag{3.2}$$

The marginal costs of the two commodities are defined as

$$MC_A = dC_A/dA, \quad MC_B = dC_B/dB. \tag{3.3}$$

From this and the previous relation we obtain $MC_B = -dC_A/dB$, and if we compute the ratio of the marginal costs we get

$$MC_B/MC_A = (-dC_A/dB)/(dC_A/dA) = -dA/dB. \tag{3.4}$$

Now, $-dA/dB$ is the negative of the slope of the transformation curve (measured with respect to the B axis), namely (since this slope is negative) its absolute value, and it thus measures the marginal rate of transformation. Therefore Eq. (3.4) states that the marginal rate of transformation must be equal to the ratio of the marginal cost of B to the marginal cost of A. This is a general proposition, which is important in itself. To conclude our proof it is sufficient to recall that under competitive conditions in the output markets the price of a commodity equals its marginal cost, $MC_A = p_A$ and $MC_B = p_B$, so that we can rewrite (3.4) as

$$-dA/dB = p_B/p_A, \tag{3.5}$$

which was to be proved.

In Sect. 19.1.1 we give rigorous proofs of the results arrived at intuitively here.

3.2 General Equilibrium in a Simple Closed Economy

3.2.1 The Supply Curves

The first step is to derive from the transformation curve the supply curves of the two commodities as a function of the price ratio or relative price, p_B/p_A. With reference to Fig. 3.5a, suppose that p_B/p_A is equal to $\tan \alpha$: the optimum point on the transformation curve is then H, where the marginal rate of transformation and the relative price are equal. Therefore, quantities OA' of commodity A and OB' of commodity B will be supplied when the relative price is $\tan \alpha$. Similarly, quantities OA_E of commodity A and OB_E of commodity B will be supplied when p_B/p_A is $\tan \beta$. In short, a unique productive combination will correspond to every admissible price ratio.

In Fig. 3.5b we measure the price ratio on the vertical axis and the quantities of the two commodities on the horizontal axis: increasing quantities of A are measured from O to the right and increasing quantities of B from O to the left. Let $OP = \tan \alpha$: to this relative price, therefore, quantities OA' of commodity A and OB' of commodity B will correspond, which are equal to the coordinates of
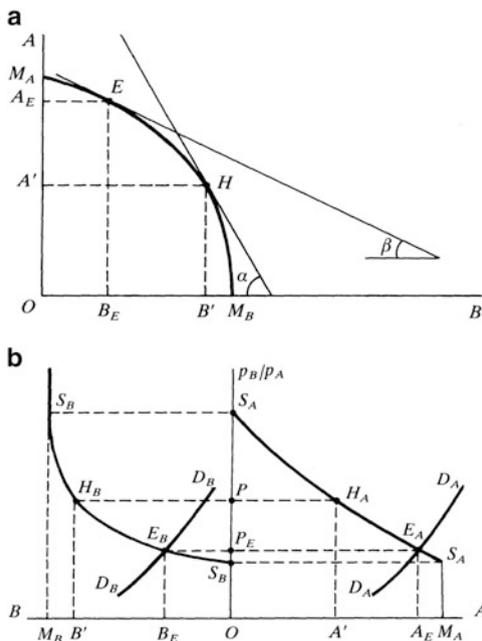
Fig. 3.5 Transformation curve, supply curves, and determination of general equilibrium in a closed economy

point H in Fig. 3.5a. Thus we obtain points H_A and H_B in Fig. 3.5b: since OA' is the quantity of commodity A supplied when the relative price is OP, point H_A will belong to the supply curve of commodity A; similarly, point H_B will belong to the supply curve of commodity B.

Continuing in this manner we obtain the supply curve of commodity A, $S_A S_A$, and the supply curve of commodity B, $S_B S_B$, which are general equilibrium supply curves. Both are increasing with respect to the appropriate relative price: $S_B S_B$ is increasing with respect to p_B/p_A (the relative price of commodity B in terms of commodity A), and $S_A S_A$ is increasing with respect to p_A/p_B. But, since $S_A S_A$ is also drawn as a function p_B/p_A in Fig. 3.5b, it will be monotonically decreasing because p_B/p_A decreases as p_A/p_B increases.

The curve $S_A S_A$ meets the vertical axis at a point which corresponds to that price ratio which causes the optimum point on the transformation curve to coincide with point M_B (see Fig. 3.5a), where the quantity of commodity A is zero and, correspondingly the quantity of commodity B is at its physically possible maximum, namely all the existing productive factors are employed in the production of B (for simplicity's sake we assume that the transformation curve's slope is neither infinite at M_B nor zero at M_A). This is denoted by the vertical stretch of the $S_B S_B$ curve corresponding to OM_B in Fig. 3.5b, to show that it is impossible to produce more of commodity B than this amount.

Similarly, the $S_B S_B$ curve meets the vertical axis at a point which corresponds to that price ratio which causes the optimum point on the transformation curve to coincide with point M_A (see Fig. 3.5a), where the quantity of commodity B is zero and, correspondingly the quantity of commodity A is at its physically possible maximum OM_A.

3.2.2   The Demand Curves

The second step is to derive the demand curves of the two commodities as a function of p_B/p_A. As we have shown in Sect. 3.1, a point on the efficiency locus in the box diagram corresponds to each point on the transformation curve (namely to each price ratio), and so the marginal productivities of capital and labour are determined, for these productivities depend only on the factor ratios when the production functions are homogeneous. We recall from microeconomics that, in competitive equilibrium, the real rewards of the productive factors coincide with their marginal productivities; therefore—since the distribution of these factors is given, as assumed in point (b), Sect. 3.1—the real income of each individual is determined. The fact that a precise real income of each individual corresponds to each given price ratio means that—unlike in partial equilibrium analysis—individual real income cannot be assumed constant as relative prices change (see below). Now, given relative prices and income, each individual, by means of the well-known maximization of a utility index subject to the budget constraint, will determine the quantities of commodity A and of commodity B demanded. Summing these quantities for all individuals, we obtain the overall demands for A and B. If we repeat this procedure for all possible ratios p_B/p_A we obtain the market demand curves for goods A and B as functions of p_B/p_A.

It should be emphasized that these demand curves are different from the usual Marshallian or partial equilibrium demand curves, which express the quantity demanded of a good as a function of its (relative) price, and are obtained on the *ceteris paribus* assumption, namely that everything else—*including* (individual) *income*—is equal. On the contrary, in our derivation *income changes as* p_B/p_A *changes*: in fact, when p_B/p_A is different, we are at a different point on the transformation curve and so at a different point on the efficiency locus in the box diagram; therefore the marginal productivities of the factors will be different and, consequently, each individual's real income will be different. In other words, the demand curves we are dealing with are *general equilibrium demand curves*, which depend on real income as well as on relative prices; but, since real income depends on relative prices alone as shown above, we can express these demand curves as functions of relative prices alone.

For simplicity's sake we assume that these demand curves are decreasing with respect to the appropriate relative price (a rigorous treatment of this topic will be given in Sect. 19.2.3), so that $D_A D_A$—which is decreasing with respect to p_A/p_B—is increasing with respect to p_B/p_A.

3.2.3   General Equilibrium and Walras' Law

The last step is to draw the demand and supply curves on the same diagram, as in Fig. 3.5b. The intersection between the demand and supply curves of a commodity determines the equilibrium point, respectively E_A and E_B for goods A and B; the corresponding equilibrium quantities are OA_E and OB_E, and the equilibrium price ratio is OP_E, equal to $\tan\beta$. The equilibrium point on the transformation curve (Fig. 3.5a) is E; therefore—as we explained above—the allocation of the productive factors between the two sectors is determined, from which the determination of the marginal productivities and hence of the factors' real rewards and of the distribution of income follow. The general equilibrium of the economy has been established.

Last but not least, an important point needs clarification: in Fig. 3.5b we have taken it for granted that the equilibrium price ratio is the same in both markets. This equality is fundamental, since if the two markets were to be in equilibrium at different relative prices, the model would be inconsistent. A simple proof, based on Walras' law, allows us to conclude that if one market is in equilibrium the other must also be in equilibrium, so that the equilibrium price ratio cannot be different in the two markets.

Let p_K and p_L indicate factor rewards, S_A and S_B the quantities of the two commodities supplied, K and L with a subscript A or B the quantities of the two factors allocated in the two sectors. Let us now recall that in each sector total factor rewards equal the value of output. This is true with constant returns to scale (first-degree homogeneous production functions: see Euler's theorem in Sect. 19.1.3), but is also true with any kind of production function provided that free entry and exit of competing firms obtain (see, for example, Mas-Colell, Whinston, & Green, 1995, sect. 10.F). Thus we have

$$p_K K_A + p_L L_A = p_A S_A,$$
$$p_K K_B + p_L L_B = p_B S_B,$$

from which

$$p_K (K_A + K_B) + p_L (L_A + L_B) = p_A S_A + p_B S_B. \tag{3.6}$$

The left-hand side of (3.6) is the total income of all the individuals in the economy (that they obtain by selling the services of the productive factors they own). Since in this model income is entirely spent in buying commodities A and B, we can write

$$p_K (K_A + K_B) + p_L (L_A + L_B) = p_A D_A + p_B D_B, \tag{3.7}$$

where D_A and D_B are the quantities demanded of the two commodities. Equation (3.7) is the aggregate budget constraint. From Eqs. (3.6) and (3.7) it follows that the right-hand sides must be equal, as the left-hand sides are equal. Therefore

$$p_A D_A + p_B D_B = p_A S_A + p_B S_B, \tag{3.8}$$

whence

$$p_A (D_A - S_A) + p_B (D_B - S_B) = 0, \tag{3.9}$$

which is true for any admissible value of p_A and p_B. The form (3.8) states that the sum of the values of the quantities demanded must equal the sum of the values of the quantities supplied; the form (3.9) states that the sum of the values of the excess demands must be equal to zero. This relationship, whichever the form used, is known as *Walras' law*. In general, given n markets linked by a (budget) constraint, Walras' law implies that if $n - 1$ markets are in equilibrium, the nth must also be in equilibrium. In our case there are only two markets, so that if one is in equilibrium the other must also be: for example, if $D_A = S_A$ then Eq. (3.9) implies $D_B = S_B$, and vice versa.

## 3.3 General Equilibrium in Open Economies and International Trade

In this section we extend the previous analysis to the international economy. In addition to those already made, we make the assumptions that only two countries exist, country 1 (the home country) and country 2 (the rest of the world), that transport costs are absent (these will be considered in Chap. 6) and that perfect competition prevails in international markets. Both countries use the same factors, which are internationally immobile, and produce the same goods.

In the absence of international trade, both countries will be in a situation of equilibrium similar to that described in Fig. 3.5b. But, as factor endowments, technology, and tastes are different in each country, it is very unlikely that the equilibrium price ratio will be the same in both. If this were so, there would be no scope for international trade. Let us then assume that the closed economy equilibrium price ratios are different in the two countries; without loss of generality we can assume that this ratio is greater in country 2 than in country 1, as shown in the back-to-back diagram drawn in Fig. 3.6. This diagram was introduced by Cunynghame (1904) and Barone (1908), but in a partial equilibrium framework: see Viner (1937, pp. 589–590).

To avoid confusion with Fig. 3.5b, we stress that in Fig. 3.6 the demand and supply curves refer to the *same commodity in the two countries*: in the right-hand part there are the demand and supply curves for commodity A in country 1, and in the left-hand part there are the demand and supply curves for the same commodity in country 2. As assumed above, the closed-economy equilibrium price-ratio in country 2 (OR_E) is greater than in country 1 (OP_E).

It can be easily shown that when trade is opened up, commercial relations are possible only if the international price ratio or *terms of trade* lies somewhere between the two internal equilibrium price ratios. We first observe that with free trade, perfect competition and no transport costs, the same commodity must have

Fig. 3.6 Determination of
international equilibrium
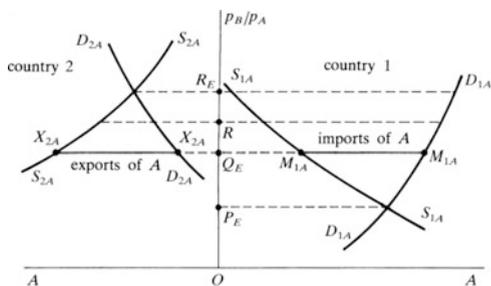
the same price everywhere (the law of one price), so that the international and
the national price ratios are the same. Now, for terms of trade higher than OR_E,
both countries would demand commodity A internationally, because in both of them
there would be an excess demand for this commodity, and no equilibrium would be
possible. Similarly, for terms of trade lower than OP_E, both countries would supply
commodity A internationally, because in both of them there would be an excess
supply of this commodity. Therefore, only intermediate terms of trade are to be
considered, since between OP_E and OR_E country 1 will demand, and country 2 will
supply, commodity A.

International equilibrium will be established at a point where the excess demand
for good A by country 1 (country 1's demand for imports) is exactly matched
by the excess supply of the same commodity by country 2 (country 2's supply
of exports). This point is shown in Fig. 3.6 at the terms of trade OQ_E, where
$M_{1A}M_{1A} = X_{2A}X_{2A}$. It can be shown that this equilibrium is stable under the usual
dynamic behaviour assumption, i.e., that price varies according to excess demand.

Suppose, for example, that we are at point R_E, where country 2 is in internal
equilibrium and so will not demand or supply anything abroad. On the contrary,
country 1 will have an excess demand for commodity A measured by the horizontal
distance between the $D_{1A}D_{1A}$ and $S_{1A}S_{1A}$ curves in correspondence to OR_E.
According to Walras' law—see Eq. (3.9)—an excess supply of commodity B will
also be present in country 1. Therefore this country will supply commodity B (the
exportable commodity) and demand commodity A (the importable commodity) on
the international market. But, since there is no demand for B nor supply of A coming
from country 2, on the international market there will be an excess supply of B and
an excess demand for A. As a consequence the international relative price of B with
respect to A will decrease, for example to OR.

When the terms of trade is OR, in country 1 there is still an excess demand
for commodity A (and so an excess supply for commodity B) though smaller than
before, whereas in country 2 an excess supply of A (and so an excess demand for B)
has appeared. But it is easy to see that the excess demand for A by country 1 is
greater than the excess supply of it by country 2, so that on the international market
an excess demand for A (and thus an excess supply of B) will still be present.
A further decrease in the terms of trade will occur, and this process will go on

until point Q_E is reached, where the excess demand for good A by country 1 is exactly matched by the excess supply of the same good by country 2. On the international market for good A, there is equilibrium between demand and supply at the terms of trade OQ_E (and, as we shall see presently, the international market for good B will also be in equilibrium). Country 1 will import an amount $M_{1A}M_{1A}$ of commodity A, exactly equal to the amount $X_{2A}X_{2A}$ of the same commodity exported by country 2; conversely (see below) country 1 will export, and country 2 will import, commodity B.

We could have arrived at the same point Q_E by starting from a lower price ratio, for example OP_E (internal equilibrium in country 1; excess supply of A and excess demand for B in country 2 and hence on the international market; increase in the relative price of B, etc.).

In Fig. 3.6, the position of the supply and demand curves for A in each country depends, as we know from Sect. 3.2, on factor endowments, technology, and tastes existing in the country. These are the elements that determine, *ceteris paribus*, the relative position of the two sides of the diagram under consideration and, therefore, which commodity will be imported and which exported. In fact, if the above elements were such that OR_E were lower than OP_E, then it would be country 1 which would export, and country 2 which would import, commodity A. This proves the following important conclusion: *in the neoclassical model of international trade, the existence of commercial relations, the pattern and the volume of trade, and the terms of trade, are jointly determined in a general equilibrium setting by factor endowments, technology,and tastes, none of which can be in general said to be an exclusive or predominant causal agent.*

We have stated above that the terms of trade which equate demand and supply in the international market for commodity A must necessarily equate it in the other market. This is a consequence of Walras' law extended to the international economy. In each country, the total value of demands equals the total value of supplies as stated in Eq. (3.8), and if we let the subscripts 1 and 2 refer to countries 1 and 2 respectively, we have

$$
\begin{aligned}
p_A D_{1A} + p_B D_{1B} &= p_A S_{1A} + p_B S_{1B,} \\
p_A D_{2A} + p_B D_{2B} &= p_A S_{2A} + p_B S_{2B}.
\end{aligned}
\tag{3.10}
$$

By addition we obtain

$$
p_A (D_{1A} + D_{2A}) + p_B (D_{1B} + D_{2B}) = p_A (S_{1A} + S_{2A}) + p_B (S_{1B} + S_{2B}),
\tag{3.11}
$$

namely the total value of world demands equals the total value of world supplies. This equation can also be written as

$$
p_A [(D_{1A} - S_{1A}) + (D_{2A} - S_{2A})] + p_B [(D_{1B} - S_{1B}) + (D_{2B} - S_{2B})] = 0,
\tag{3.12}
$$

or

$$p_A \left[(D_{1A} + D_{2A}) - (S_{1A} + S_{2A}) \right] + p_B \left[(D_{1B} + D_{2B}) - (S_{1B} + S_{2B}) \right] = 0,$$
$$(3.13)$$

namely the sum of the values of world's excess demands must equal zero for any admissible value of p_A and p_B.

Suppose now that, at a particular price ratio, the international market for commodity A is in equilibrium, i.e.

$$D_{1A} + D_{2A} = S_{1A} + S_{2A}; \qquad\qquad (3.14)$$

then it follows from Eq. (3.11) that

$$D_{1B} + D_{2B} = S_{1B} + S_{2B}, \qquad\qquad (3.15)$$

namely that the international market for commodity B is also in equilibrium. From (3.14) and (3.15) it also follows that

$$\begin{aligned}
D_{1A} - S_{1A} &= S_{2A} - D_{2A}, \\
S_{1B} - D_{1B} &= D_{2B} - S_{2B},
\end{aligned} \qquad (3.16)$$

which state that excess demand for good A by country 1 (country 1's demand for imports) is equal to excess supply of the same good by country 2 (country 2's supply of exports) and that country 1's supply of exports of good B is equal to country 2's demand for imports of the same good.

It is also worth pointing out that conditions (3.10) imply that no country can be a net importer or exporter of *both* commodities. In fact, if we rewrite these conditions as

$$\begin{aligned}
p_A (D_{1A} - S_{1A}) &= p_B (S_{1B} - D_{1B}), \\
p_A (D_{2A} - S_{2A}) &= p_B (S_{2B} - D_{2B}),
\end{aligned} \qquad (3.17)$$

we see that if $D_{1A} > S_{1A}$ (excess demand for commodity A by country 1, which thus imports this commodity), then $S_{1B} > D_{1B}$ (country 1 exports commodity B) and vice versa. This result is obvious if we think that in the barter model under consideration a country can obtain imports only by paying for them with exports. It should also be noticed that Eqs. (3.17) can be interpreted as the equality, for each country, between the value of its imports and the value of its exports when both are *evaluated at the given international prices*. Therefore, as is typical in the pure theory of international trade, the balance of trade always balances.

Fig. 3.7 Derivation of the offer curve

3.4   Offer Curves, International Equilibrium, and Stability

3.4.1   Derivation of the Offer Curve

An alternative way of determining international equilibrium is to use the Marshallian *reciprocal demand curves* (also called *offer curves* and *demand-and-supply curves*). The notion of reciprocal demand is already present in J.S. Mill, as we saw in Sect. 2.5, but the first complete treatment is to be found in Marshall (1879), who also introduced the graphic apparatus of the offer curves, though he did not show how they are derived from the underlying production and demand conditions.

The offer curve of a country can be defined as the locus of all points which represent the (maximum) quantity of the exported good that the country is willing to give in exchange for a given amount of the imported good (or, if we prefer, the (minimum) quantity of the imported good that the country is willing to accept in exchange for a given amount of the exported good). Equivalently, this curve indicates the various terms of trade at which the country is willing to trade.

There are several ways of obtaining a country's offer curve geometrically; one is Meade's (1952) ingenious geometric technique based on trade indifference curves and the transformation curve. However, the graphic apparatus developed in Sect. 3.2 allows a very simple derivation of the offer curve, and we shall use this instead of Meade's technique.

In Fig. 3.7a the same diagram contained in Fig. 3.5b is drawn. Let us consider an arbitrary price ratio, for example OH. At this relative price, country 1 has an excess demand for good A equal to $H_A H_A$ and an excess supply of good B equal to $H_B H_B$. This country, therefore, is willing to exchange $H_B H_B$ of B for $H_A H_A$ of A on the

international market, namely it is willing to import an amount $H_A H_A$ of commodity A and to export, in exchange for this, an amount $H_B H_B$ of commodity B.

In Fig. 3.7b we draw the amounts of A and B just obtained, measuring the demand for imports on the vertical axis ($OH_A = H_A H_A$) and the supply of exports on the horizontal axis ($OH_B = H_B H_B$); we thus obtain point Q. The *terms of trade* in Fig. 3.7b are represented by OH_A/OH_B, we recall that p_B/p_A expresses the number of units of A for one unit of B, and the same thing is expressed by the ratio (OH_A/OH_B), namely by the slope of OQ, which is $\tan \alpha$; this is equal to OH in Fig. 3.7a.

If we let the price ratio take on all values from OP_E upwards, we obtain other points in a similar way, which give rise to the curve OG_1. For values of the price ratio lower than OP_E the export-import situation of country 1 will be reversed, because there will be an excess supply of commodity A and an excess demand for commodity B. If we adopt the convention of measuring the import demand for B by country 1 on the horizontal axis from O to the left, and the export supply of A by this same country on the vertical axis from O downwards, we obtain the branch OG_1' of the offer curve of country 1. If the price ratio is OP_E in country 1 there will be no excess demand or excess supply, therefore this country's offer curve will pass through the origin; the slope of the $G_1' OG_1$ curve measured at the origin is equal to the internal equilibrium price-ratio OP_E.

To sum up: every point of the OG_1 curve gives the demand by country 1 for imports of commodity A and the corresponding supply of exports of commodity B; every point of the OG_1' curve gives the supply by country 1 of exports of A and the corresponding demand for imports of B. The curve $G_1' OG_1$ is, therefore, the offer curve of country 1. Note that, since the domestic demand and supply curves have been obtained by an optimization procedure (as shown in Sect. 3.2), concerning both the demand and the supply, the excess demands and supplies which give rise to the offer curve, and therefore this curve, have an optimal nature.

In a similar way we can build the offer of country 2, $G_2' OG_2$. Given the assumption made in Fig. 3.6, when the price-ratio is lower than OR_E (which equals the slope at the origin of the $G_2' OG_2$ curve in Fig. 3.7b), country 2 has an excess supply of commodity A (and so an excess demand for commodity B). Then each point of the OG_2 curve gives the supply by country 2 of exports of A and its corresponding demand for imports of B.

This derivation of the offer curve shows the truth of Edgeworth's often quoted statement: "There is more than meets the eye in Professor Marshall's foreign trade curves. As it has been said by one who used this sort of curve, a movement along a supply-and-demand curve of international trade should be considered as attended with rearrangement of internal trade; as the movement of the hand of a clock corresponds to considerable unseen movements of the machinery" (Edgeworth, 1905, p. 70; p. 143 of the reprint. He was actually quoting himself: see Edgeworth, 1894, pp. 424–425; p. 32 of the reprint).

3.4.2   *International Equilibrium and Stability*

We saw above that no international trade is possible when the terms of trade are lower than OP_E or higher than OR_E, and this is reflected in the fact that in the third quadrant in Fig. 3.7b both countries are net suppliers or net demanders of the same commodity. The branches OG'_1 and OG'_2, therefore, are not relevant, and only the first quadrant has to be considered, where country 1 demands A and supplies B, and country 2 supplies A and demands B. The offer curves OG_1 and OG_2 intersect at point E, which is the equilibrium point: country 1 demands OE_A of commodity A, exactly equal to the amount of A supplied by country 2, and supplies OE_B of commodity B, exactly equal to the amount of B demanded by country 2. International trade will take place on the basis of OE_B of B (exported by country 1 and imported by country 2) for OE_A of A (imported by country 1 and exported by country 2); the equilibrium terms of trade are measured by $\tan \beta$ (slope of the ray OE), which is equal to OQ_E in Fig. 3.6.

The offer curves are widely used in international economics not only for determining international equilibrium but also for a number of other purposes, as we shall see in this and in the following chapters. It is therefore important to bear in mind that they are derived from the underlying production and demand conditions, as pointed out in Edgeworth's statement quoted above.

We now put the offer curves to use for examining the *stability* of the equilibrium point E when the adjustment process directly involves quantities rather than the terms of trade. It is well-known that to examine the stability of equilibrium we need *behaviour assumptions* concerning the reaction of the relevant variables to a disequilibrium situation. In Sect. 3.3 we examined the problem of stability by making the assumption that the variable which adjusts itself in the first instance is the terms of trade, reacting to excess demand and supply on the international market. In other words, the adjustment mechanism acted on the relative *price*, and quantities followed. Now—following Marshall (1879, 1923)—we make the assumption that the variables which adjust themselves in the first instance are the *quantities* of the two commodities. There are, however, at least two ways in which this adjustment may take place, namely there are at least two possible behaviour assumptions,[1] that we will now examine.

Behaviour Assumption I Consider any non-equilibrium point P. Owing to the competition between its traders, each country adjusts the quantity of its exports towards that quantity which it would offer at the terms of trade actually prevailing, if such terms remained fixed for all the time needed to complete the adjustment.

With reference to Fig. 3.8, assume that the initial non-equilibrium point is P. Now, OH_1 is the initial quantity of exports of country 1 and OH_2 is the initial

---

[1] See Kemp (1964, chap. 4), who attributes assumption II to Marshall, while leaving assumption I unnamed. Owing to the ambiguity of Marshall's statements (1879, 1923) on this topic, we believe that both assumptions are consistent with what he wrote. See also Samuelson (1947).
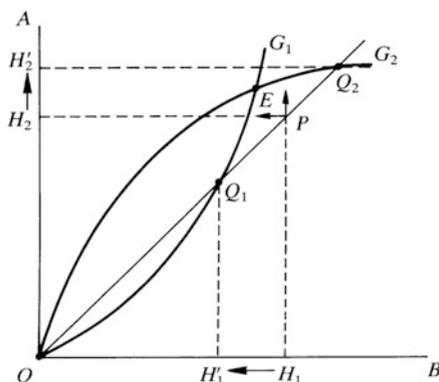
Fig. 3.8 Adjustment of quantities and stability of international equilibrium: Behaviour assumption I

quantity of exports of country 2; the terms of trade are measured by the slope of the straight line passing through P and O. Given these terms of trade the quantity of exports that country 1 wishes to supply is determined, by the very definition of an offer curve, by the abscissa of point Q_1. Therefore, country 1 is inclined to decrease its exports, and adjusts them from OH_1 towards OH_1'. By similar reasoning, it can be seen that country 2 tends to expand its exports, by adjusting them from OH_2 towards OH_2'. Thus point P moves in a direction to be found between the two arrows, tending to point E.

It is perhaps worth pointing out that this method of studying stability by means of pointed arrows representing the forces at play—a method now widely used in international economics as well as in other branches of economics—was first introduced by Marshall (1879) in order to study the stability of international equilibrium. It should however be stressed that the arrows do *not*, by themselves, make it possible to determine the actual trajectory of point P and even less to say whether this point will converge to the equilibrium point, or how. They are useful expository devices, but cannot replace a rigorous formal analysis (for further comments on arrow diagrams, see Gandolfo, 2009, chap. 19, sect. 19.3). This analysis is carried out in Sect. 19.4.2 for behaviour assumptions I and II, to the latter of which we now turn.

Behaviour Assumption II Consider any point P different from the equilibrium point. Each country adjusts its supply of exports towards that quantity of exports which it would offer if the current quantity of imports (corresponding to point P) remained fixed for the whole time needed to complete the adjustment.

In other words, each country moves towards the point on the respective offer curve corresponding to the prevailing quantity of the country's imports. With reference to Fig. 3.9, assume that the initial non-equilibrium point is P. Now, OY is the initial quantity of imports of country 1 and OX is the initial quantity of imports of country 2. The quantity of exports that country 1 wishes to offer in exchange

Fig. 3.9 Adjustment of quantities and stability of international equilibrium: Behaviour assumption II
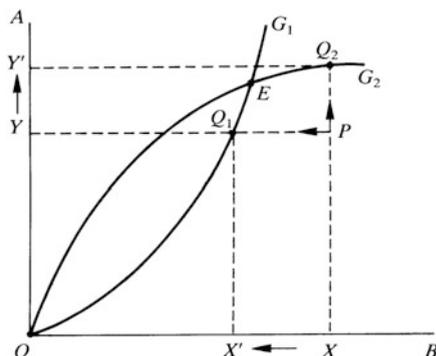
for the current quantity of imports is OX'; consequently, this country adjusts its exports from the current quantity OX towards the desired quantity OX'. Similarly, it can be seen that country 2 adjusts its exports from the current quantity OY towards the desired quantity OY'. Thus point P moves in a direction to be found between the two arrows, tending to point E.

Thus we have seen that the equilibrium point E is stable according to both behaviour assumptions. But this has occurred because we have assumed that the offer curves have the "normal" form, i.e., they are both monotonically increasing and each one is concave to its import axis. But other shapes of the offer curves are admissible, so that cases may arise in which equilibrium is unstable according to both behaviour assumptions, as well as cases in which equilibrium is stable according to one assumption and unstable according to the other (Kemp, 1964, pp. 68–69).

It can be shown (see Sect. 19.4.2) that the local stability conditions can be expressed in terms of the elasticities of the offer curves. These elasticities can be defined in several ways (elasticity of imports with respect to exports, elasticity of exports with respect to imports, etc.). We follow Kemp (1964) in defining the elasticity of an offer curve as the proportional change in (the supply of) exports divided by the proportional change in (the demand for) imports. This implies that, when writing the offer curve as an explicit function, we choose to express (the supply of) exports as a function of (the demand for) imports instead of the other way round. This choice is consistent with the dynamic behaviour assumption just examined, where the variable which adjusts itself is the supply of exports. Formally, let $B^S = G_1(A^D)$ be the offer curve of country 1. The quantity B^S is country 1's supply of exports, which in turn is equal to the domestic excess supply, as shown in Sect. 3.4.1. In symbols, $B^S = S_1^B - D_1^B$. Similar observations hold for A^D, A^S, B^D.

The elasticity of the offer curve—for infinitesimal changes—is defined as

$$e_1 = \frac{\mathrm{d}B^S/B^S}{\mathrm{d}A^D/A^D} = \frac{\mathrm{d}B^S}{\mathrm{d}A^D} \cdot \frac{A^D}{B^S}, \qquad (3.18)$$
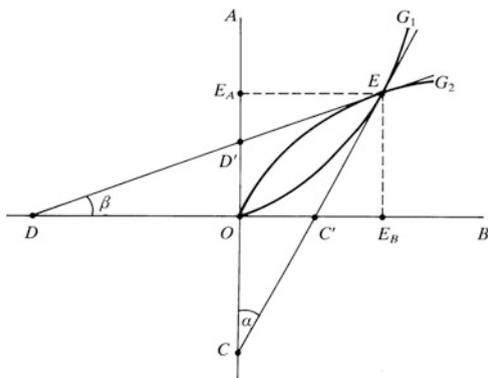
Fig. 3.10 Graphic measurement of offer curve elasticity

where $\mathrm{d}B^S/\mathrm{d}A^D$ is the slope of the OG_1 curve referred to its import axis. Similarly, letting $A^S = G_2\,(B^D)$ be the offer curve of country 2, its elasticity is

$$e_2 = \frac{\mathrm{d}A^S/A^S}{\mathrm{d}B^D/B^D} = \frac{\mathrm{d}A^S}{\mathrm{d}B^D} \cdot \frac{B^D}{A^S}. \tag{3.19}$$

These elasticities can be measured graphically in a simple way. Consider for example point E in Fig. 3.10. The slope of the OG_1 curve with respect to its import axis is $\tan\alpha$. Now, $\tan\alpha = EE_A/E_AC = OE_B/E_AC$; note also that the angle $C'\hat{E}E_B$ is equal to α, so that $\tan\alpha = C'E_B/EE_B$ as well. Furthermore, $A^D = OE_A = EE_B$, and $B^S = OE_B = EE_A$. Therefore

$$e_1 = \frac{OE_B}{E_AC} \cdot \frac{OE_A}{OE_B} = \frac{C'E_B}{EE_B} \cdot \frac{EE_B}{OE_B},$$

from which

$$e_1 = \frac{OE_A}{E_AC} = \frac{C'E_B}{OE_B}. \tag{3.20}$$

In a similar way we obtain

$$e_2 = \frac{OE_B}{E_BD} = \frac{D'E_A}{OE_A}. \tag{3.21}$$

Equations (3.20) and (3.21) are simple and useful expressions for measuring the elasticities of the offer curves graphically. Note that if we defined these elasticities the other way round, their graphic measures would be the reciprocals of the expressions given in Eqs. (3.20) and (3.21).

Fig. 3.11 Multiple equilibria
and stability-instability
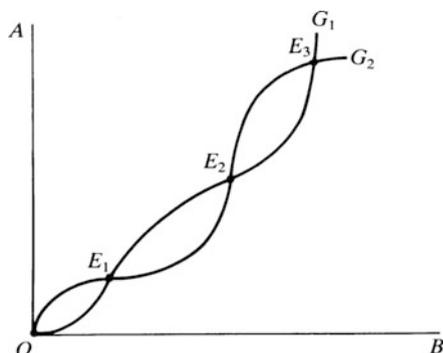
Going back to the stability conditions, it turns out (see Sect. 19.4.2) that necessary and sufficient conditions for local stability are

$$\frac{1 - e_1 e_2}{(1 - e_1)(1 - e_2)} > 0 \tag{3.22}$$

if we adopt behaviour assumption I, and

$$1 - e_1 e_2 > 0 \tag{3.23}$$

if we adopt behaviour assumption II. If both elasticities are positive and smaller than 1, as they are in the cases examined so far, then both (3.22) and (3.23) are satisfied. But in abnormal cases anything may happen, for example contradictory results of the two behaviour assumptions, as already mentioned.

In conclusion, let us note that *multiple equilibria* may occur, as was demonstrated by Marshall (1879, 1923). One of the cases that were treated by him is shown in Fig. 3.11. According to Marshall, point E_2 is unstable, whereas points E_1 and E_3 are stable, thus respecting his proposition XIII (1879; p. 24 of the 1930 reprint) that in the case of multiple equilibria stable and unstable equilibria alternate. Although this proposition is not universally true, it holds in the case of Fig. 3.11, as can be seen either by using the graphic method of arrows or by applying conditions (3.22) and (3.23). It turns out that both e_1 and e_2 are greater than one at point E_2, whereas they are both smaller than one and positive at points E_1 and E_3. Therefore, neither (3.22) nor (3.23) is satisfied at point E_2, whereas both are satisfied at points E_1 and E_3, so that in this case Marshall's proposition holds independently of the behaviour assumption accepted.

3.5   Increasing Returns to Scale

In general, the presence of non-constant (decreasing or increasing) returns to scale has an effect on the curvature of the transformation curve. Since there seems to be a certain amount of imprecision in the literature when this effect is dealt with, we give

a brief summary of the result (for proofs the reader is referred to Herberg (1969); see also Sect. 19.1.3).

In what follows, concavity and convexity are referred to the origin, different factor intensities in the two sectors are assumed, and it is also assumed that increasing (decreasing) returns to scale in a sector can be described by a homogeneous production function of degree higher (lower) than the first.

1. The transformation curve is strictly concave if both sectors have production functions with decreasing returns to scale or, more generally, if no sector produces with increasing returns.
2. Only slightly increasing returns in both sectors will make the transformation curve strictly convex near the coordinate axes and strictly concave somewhere in the intermediate range.
3. The transformation curve is strictly convex everywhere if, and only if, no sector has decreasing and at least one sufficiently strong increasing returns. The amount by which the degree(s) of homogeneity must exceed one is, ceteris paribus, the smaller the less the factor intensities of the commodities differ.
4. The transformation curve has at least one point of inflection if there are increasing returns in one sector and decreasing returns in the other. If the factor intensities happen to be equal in the two sectors, then:
5. Proposition 9.1 remains true if we exclude the case of constant returns in both sectors (in which case, as we know from Sect. 19.1, the transformation curve is linear).
6. The transformation curve is strictly convex if, and only if, one sector has increasing and the other no decreasing returns.

However, increasing returns to scale do not by any means only influence the shape of the transformation curve. As is well known, unlimited increasing returns to scale due to *internal* economies are incompatible with perfect competition; internal economies are however compatible with other market forms, for example monopoly (typical outcome of unlimited increasing returns) or monopolistic competition (these cases will be examined in Chap. 9). The compatibility of increasing returns with perfect competition is however preserved by the introduction of Marshallian *external* economies. On the other hand, when external economies are present, marginal social cost and marginal private cost are no longer the same. As a consequence, it is not certain that the economy produces on the transformation curve (the production point may lie inside this curve) and, even if it does, it is not certain that in equilibrium the price ratio will be equal to the marginal rate of transformation (for details of these problems, see Chipman, 1965b, pp. 736–749). We follow Meade (1952), Kemp (1964, 1969b) and others in assuming away these complications, namely we hypothesize that, notwithstanding the presence of increasing returns, the economy produces on the transformation curve at a point where the price ratio equals the marginal rate of transformation (sufficient conditions for this to be true are given by Kemp (1964, chap. 7, 1969b, chap. 8); for a treatment of the case in which the equality between price ratio and marginal rate of transformation no longer
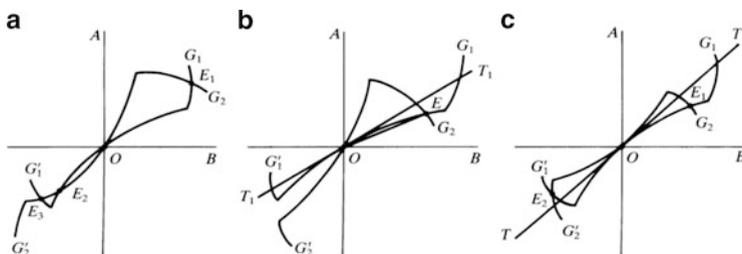
Fig. 3.12 Increasing returns to scale, offer curves, and international equilibrium

holds, see Chacholiades, 1978, chap. 7). We also assume that the transformation curve is strictly convex to the origin.

As a consequence of these simplifying assumptions, the formal analysis of increasing returns to scale does not differ from the analysis of constant returns to scale, for we only have to deal with the fact that the transformation curve is convex, instead of concave, to the origin.

The most interesting results concerning a trading world with increasing returns to scale in both countries are:

(a) In general there are multiple equilibria, and the direction of trade is not univocally determined;
(b) The equilibrium terms of trade may well lie outside the interval defined by the two closed-economy price ratios;
(c) Trade can take place even when the two closed-economy price ratios are equal.

These results can be easily obtained by using the offer curves. It turns out that, under increasing returns to scale, the offer curves have the shape shown in Fig. 3.12 (for their derivation see Chacholiades, 1978; Kemp, 1964, 1969b; Meade, 1952).

Figure 3.12a depicts a situation in which there are three equilibrium points: E_1, E_2, and E_3. Since, in the first quadrant, country 1 wishes to import commodity A and to export commodity B (and vice versa for country 2), whereas in the third quadrant the opposite is true, we see that the direction of trade is indeterminate. In other words, while in the case of constant returns to scale possible multiple equilibria do not alter the direction of trade, in the case under consideration a *normal* consequence of multiple equilibria is that of giving rise to *different* directions of trade. Therefore the direction of trade cannot be predicted on a priori grounds.

Figure 3.12b shows a case in which there is only one equilibrium point, and the equilibrium terms of trade (slope of the straight line segment OE) are lower than the autarkic price ratio in country 1 (the latter is measured, as in Sect. 3.4, by the slope at the origin of the $G_1'OG_1$ offer curve, namely by the slope of the straight line T_1T_1, which is tangent to $G_1'OG_1$ at the origin).

Finally, Fig. 3.12c depicts a situation in which the two autarkic price ratios coincide, for they are both equal to the slope of the straight line TT, which is the common tangent to both offer curves at the origin. Notwithstanding this, trade can

and does take place, as shown by the two equilibrium points E_1 and E_2. We should like to underline this result, which shows that *increasing returns to scale can be a determinant of international trade*.

For a fuller treatment of increasing returns to scale in international trade the reader is referred to Kemp (1964, chaps. 8 and 12, sects. 7–8; 1969b, chaps. 8 and 11, sects. 7–8), Negishi (1972, chaps. 5 and 8), Chacholiades (1978, chap. 7), Helpman (1984b), Vanek (1962), Krauss (1979), Herberg et al. (1982) .

3.6 The Gains from Trade

We saw in the context of the classical theory that international trade is beneficial in so far as it enables a country to obtain a commodity at a lower cost than the domestic production cost or, alternatively, to obtain commodity bundles which were out of reach under autarky. A similar conclusion holds in neoclassical theory.

Consider for example Fig. 3.13 and suppose that the pre-trade closed-economy price ratio is represented by the slope of the straight line PP, whereas the terms of trade (post-trade open-economy price ratio) are represented by the slope of the straight line RR. Before trading started the country produced and consumed a commodity bundle given by the coordinates of point E. When trade is opened up, the country produces the commodity bundle given by the coordinates of point E' (*production* point). But it can now trade along the RR line, thus attaining previously unattainable points, outside its transformation curve. For example, it can move to point E'' (*consumption* point) by trading $H_B E_B'$ of commodity B (exportables) for $H_A E_A'$ of commodity A (importables); point E'' is clearly better (excluding inferior commodities) than the pre-trade point E because the amounts of both commodities are greater at E'' than at E. It can also be seen that—since we have assumed that A is the imported, and B the exported, commodity—the opportunity cost of A in terms of B is greater in the closed economy situation (slope of PP referred to the vertical axis) than in the open economy situation (where the additional amount of B that has to be given up to obtain an additional amount of A is measured by the appropriate terms of trade, namely by the slope of RR referred to the vertical axis).

But what if the post-trade situation is E'''? This point is undoubtedly outside the transformation curve, and thus it could not be reached before trade, but since with respect to E it contains a greater amount of commodity A and a smaller amount of commodity B, it cannot be considered unambiguously better than E. It is however easy to observe that the value of national income at E''' is in any case greater than at E. This is true whether national income is calculated at the closed-economy (pre-trade) prices or at the new (post-trade) prices. Let us first consider the closed-economy prices. The value of national income at E is given by the position of the equal income line (which we call *isoincome*) PP, while at E''' it is given by the position of the isoincome line (not shown in the diagram) parallel to PP and passing through E''', which is clearly more distant from the origin than PP. It follows that national income evaluated at the closed-economy prices is higher at E''' than at E.

Fig. 3.13 The gains from trade
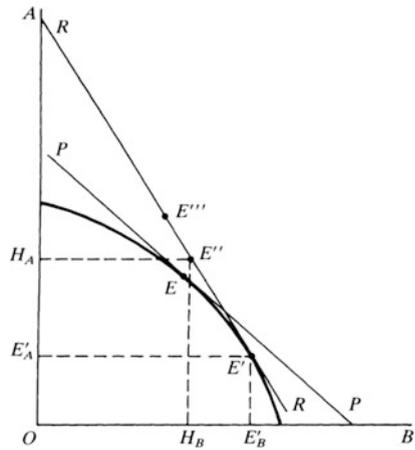
At the post-trade prices, the value of national income at E''' is given by the position of the isoincome line RR, while at E it is given by the isoincome line (not shown in the diagram) parallel to RR and passing through E, which is clearly nearer to the origin than RR (and hence represents a lower income).

It could also be observed that, since trade is *free* and not compulsory the fact that the country chooses point E''', instead of point E'', means that it prefers, in some sense, the former to the latter: we are in the presence of a sort of revealed preference.

The gains from trade can be given a more precise treatment if one is willing to accept the concept of *community* or *social indifference curves*. The problems raised by this concept are among the moot questions in welfare economics (see, for example, Mas-Colell et al., 1995, sect. 4.D; Chacholiades, 1978, chaps. 5 and 16). This notwithstanding, these curves are widely used in international economics and we do not depart from general practice by using them as a helpful expository device, though fully aware of their shortcomings.

In Fig. 3.14a the pre-trade (autarkic equilibrium) situation is depicted; social welfare is maximized at point E, where a social indifference curve is tangent to the transformation curve. In Fig. 3.14b, the terms-of-trade line RR is drawn: the highest indifference curve attainable is that which is tangent to this line, thus determining the consumption point E_C precisely, as well as the imported and the exported commodities and the amounts traded ($H_B E'_B$ of exports for $H_A E'_A$ of imports). The gains from trade are immediately visible, as the social indifference curve tangent at E_C is higher than the curve tangent at E, and so represents a better situation. Ideally, the gains from trade can be subdivided into a *consumption gain* and a *production gain*. The first is due to international exchange only, and can be seen by freezing the production point at the pre-trade point E. In this situation the country can trade along the $R'R'$ terms of trade line, parallel to the RR line; the optimum position is reached at point E'_C. Since the social indifference curve tangent at E'_C is higher than that tangent at E, there is a gain: the *consumption gain*.
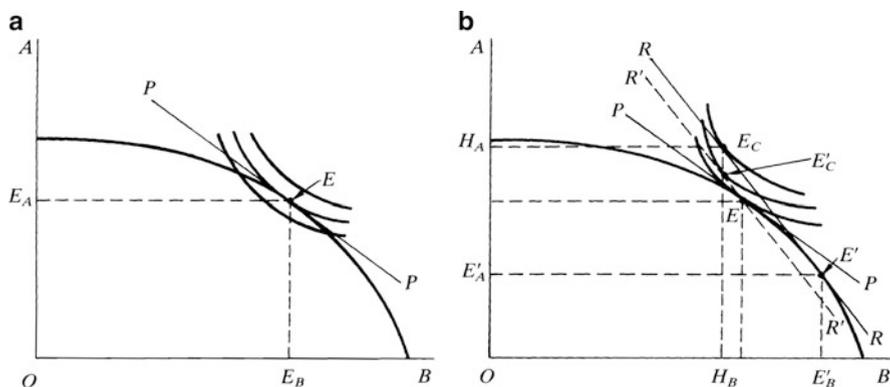
Fig. 3.14 Social indifference curves and the gains from trade: consumption and production gains

The *production gain* is due to specialization, since, as a consequence of the difference between the post-trade and the pre-trade commodity prices, the country changes its pattern of production and specializes (though incompletely) in the production of B, moving from the (now) inefficient production point E to the efficient one E'. This enables the country to reach a still higher indifference curve: the production gain is represented by the movement from E_C' to E_C.

We have stated that specialization is incomplete, as shown in the diagram. In fact, in the neoclassical theory—unlike the classical theory, where *complete* specialization was the necessary outcome of international trade—the specialization is normally incomplete (though complete specialization cannot be excluded: this occurs when the terms-of-trade line is tangent to the transformation curve at one of the points where this curve intersects the axes). The different results are due to the different assumptions concerning opportunity cost. Given a difference between the internal opportunity cost and the terms of trade, the productive combination will be modified in the direction of greater convenience. Now, if these modifications do not alter the opportunity cost (as in the classical theory: linear transformation curve), the inevitable outcome is complete specialization. On the contrary, when they bring about changes in the opportunity cost (as in the neoclassical theory), specialization will stop when opportunity cost becomes equal to the given terms of trade; this will normally occur at a point on the transformation curve somewhere between its two intercepts.

So far, we have considered one country only. What about our two-country world? It can be shown that trade is beneficial to *both* countries. In Fig. 3.15 we have drawn the transformation curves of the two countries together with the pre-trade and post-trade equilibria. The closed-economy equilibrium price ratio p_B/p_A is lower in country 1 (slope of P_1P_1 referred to the horizontal axis) than in country 2 (slope of P_2P_2): we are in a situation similar to that depicted in the back-to-back diagram (Fig. 3.6). The post-trade price ratio will lie between the two pre-trade ratios; country 1 will import commodity A and export commodity

Fig. 3.15 Trade is beneficial
to all countries
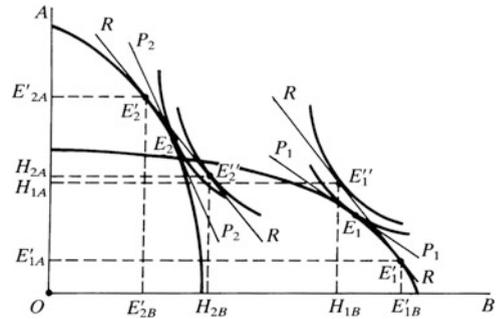
B, whereas the opposite will occur in country 2. This is shown in Fig. 3.15, where the slope of the terms-of-trade line RR measures the post-trade price ratio. Country 1 moves its production pattern from E_1 to E_1' (specializing in B), and country 2 moves its production pattern from E_2 to E_2' (specializing in A). Then country 1 exports $H_{1B}E_{1B}'$ of commodity B (equal to the quantity $E_{2B}'H_{2B}$ imported by country 2) and imports $E_{1A}'H_{1A}$ of commodity A (equal to the quantity $H_{2A}E_{2A}'$ exported by country 2).

As a consequence of these exchanges country 1's consumption point is at E_1'' (which lies on the highest social indifference curve of country 1 attainable given the terms-of-trade line RR) and similarly country 2's consumption point is at E_2'': as we see, both countries are on a higher indifference curve than in the pre-trade situation.

3.7 Generalizations

We have so far worked with the well-known $2 \times 2 \times 2$ model (two countries, two goods, two factors). But what happens when there are many countries, many commodities, and many factors? Among the first attempts to treat this problem formally is Yntema's (1932); 12 years later the problem was again tackled by Mosak (1944). Both of these, however, treated this topic *à la Walras*, namely by writing down equilibrium conditions and then counting equations and unknowns. The equilibrium conditions for the general problem can be written by making a straightforward extension of those holding in the $2 \times 2 \times 2$ model. In fact, application of the optimizing procedures to both the production side and the consumption side of each country makes it possible to derive the supply of and the demand for each commodity in each country as functions of relative prices only. Then world equilibrium requires that for each commodity world demand equals world supply, and, by summing the budget constraints, we find that, if all but one excess demands are equal to zero, then the last must also be.

But, as is well known, the mere counting of equations and unknowns is not a satisfactory procedure for proving the existence of an equilibrium, for in general the equality of the number of equations and of the number of unknowns is neither

a necessary nor a sufficient condition for existence. An adequate proof must therefore rely on the same methods used in mathematical economics to prove the existence of general competitive equilibrium in a closed economy. Among the first modern proofs along these lines is Nikaidô's (1956, 1957); for further details, see Chipman (1965b, sect. 2.6).

There is, however, a price to be paid for this generality, because one must be content with knowing that an equilibrium exists (and with analysing its stability), without being able to find operational propositions allowing one to determine the structure and the volume of international trade, etc., in a simple way. On the other hand, the neoclassical theory can be used to yield simple predictions on the structure of international trade by restricting its generality. As a matter of fact, from the purely analytical point of view, the Heckscher-Ohlin theory (with all its corollaries, such as the factor price equalization theorem, etc.) can be considered as a particular case of the neoclassical theory: see Chap. 4.

The neoclassical theory can be generalized in several other directions, for example, by relaxing the assumption of fixed quantities of factors and introducing variable factor supplies, or by introducing transport costs, non-traded goods, specific factors etc. (see Chap. 6).

3.8   Duality Approach

Duality theory, which studies the dual relations between cost functions and production functions, between direct and indirect utility functions, etc. (for an introduction see Varian, 1992, chap. 6; a more advanced treatment is Diewert's, 1974, 1982) is being increasingly applied to microeconomics and to general equilibrium theory, as it enables us—among other things—to derive in a formally simpler way the comparative statics theorems originally deduced from maximizing behaviour.

Among the first applications of duality theory to international trade is the one by Jones (1965), who showed the dual nature of the Stolper-Samuelson (see below, Sect. 5.3) and Rybczynski (Sect. 5.4) theorems. Indeed, the whole pure theory of international trade can be rewritten by using duality theory: see, for example, Dixit and Norman (1980), Woodland (1982), and Sgro (1986).

However, much of the literature (especially as regards elementary and intermediate international economics textbooks, in some of which the duality approach is not even mentioned) is still based on the conventional approach. One reason may be that the conventional approach more easily lends itself to an intuitive verbal and graphical treatment and hence is more student-friendly. Another may be that the whole doctrinal body of international trade theory, from Ricardo to Heckscher-Ohlin and further, has been constructed and refined through the conventional approach.

Be it as it may, we have adhered to the conventional approach throughout the text, while treating the duality approach in the appendices (see Sect. 19.5 for the basic elements), where we also show in the appropriate places(e.g., Sects. 20.1–20.3, 21.2, 21.3, 22.3, 22.6) how certain formal results can be more easily derived using duality theory instead of the conventional approach.

3.9 Empirical Studies

Surprisingly enough, the neoclassical model of international trade in its general version has received little or no empirical attention, as practically all empirical studies have concentrated on the Heckscher-Ohlin model (see Sect. 4.6), that from the theoretical point of view can be considered as a particular case of the general neoclassical model.

A first step in the direction of filling this gap in the empirical literature was taken by Harrigan (1997), who specified a model of international specialization consistent with the neoclassical explanation. This model, where relative technology levels and factor supplies jointly determine international specialization, gives fairly good empirical results, so that "the neoclassical model comes out looking rather well" (Harrigan, 1997, p. 477).

For further considerations see Sect. 4.6.5.

References

Barone, E. (1908). *Principi di economia politica*.

Chacholiades, M. (1978). *International trade theory and policy* (Pt. III).

Chipman, J. S. (1965b). A survey of the theory of international trade: Part 2, the neoclassical theory.

Cunynghame, H. (1904). *A geometric political economy*.

Diewert, W. E. (1974). Applications of duality theory.

Diewert, W. E. (1982). Duality approaches to microeconomic theory.

Dixit, A. K., & Norman, V. (1980). *Theory of international trade: A dual, general equilibrium approach* (chaps. 1–3).

Edgeworth, F. Y. (1894). The theory of international values, I, II, III.

Edgeworth, F. Y. (1905). Review of H. Cunynghame's book "A Geometrical Political Economy".

Findlay, R. (1970). *Trade and specialization* (chap. 2).

Gandolfo, G. (2009). *Economic dynamics* (Pt. II, chap. 19, sect. 19.3; Pt. III, chap. 21, sect. 21.3.2).

Harrigan, J. (1997). Technology, factor supplies, and international specialization: Examining the neoclassical model.

Helpman, E. (1984b). The factor content of international trade.

Herberg, H. (1969). On the shape of the transformation curve in the case of homogeneous production functions.

Herberg, H., Kemp, M. C., & Tawada, M. (1982). Further implications of variable returns to scale.

Jones, R. W. (1965). The structure of simple general equilibrium models.

Kemp, M. C. (1964). *The pure theory of international trade* (Pt. II).

Kemp, M. C. (1969b). *The pure theory of international trade and investment* (Pt. I).

Krauss, M. B. (1979). *A geometric approach to international trade* (chap. II, 15–24 and 29–32).

Marshall, A. (1879). *The pure theory of foreign trade*.

Marshall, A. (1923). *Money, credit and commerce* (Appendix J).

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*.

Meade, J. E. (1952). *A geometry of international trade* (chaps. I–V).

Mosak, J. L. (1944). *General-equilibrium theory in international trade*.

Negishi, T. (1972). *General equilibrium theory and international trade* (chaps. 5 and 8).

Nikaidô, H. (1956). On the classical multilateral exchange problem.

Nikaidô, H. (1957). A supplementary note to "On the classical multilateral exchange problem".

Samuelson, P. A. (1947). *Foundations of economic analysis* (pp. 266–267).
Savosnick, K. M. (1958). The box diagram and the production possibility curve.
Sgro, P. M. (1986). *The theory of duality and international trade*.
Vanek, J. (1962). *International trade: Theory and economic policy*.
Varian, H. R. (1992). *Microeconomic analysis* (chap. 6).
Viner, J. (1937). *Studies in the theory of international trade*.
Woodland, A. D. (1982). *International trade and resource allocation*.
Yntema, T. O. (1932). *A mathematical reformulation of the general theory of international trade*.