# Methods of Establishing the Shortest Running Distances for Freights on Setting up Transportation Systems

A. L. Lur'e

The volume of goods traffic required to meet all the needs of the national economy depends to a large extent, once the location of industries has been taken into account, on the way in which the actual transport links are established between the areas and points of production and the areas and points of consumption of the various products.

One of the most important requirements which arises when rational schemes of transport links are being set up is the need to establish, all other conditions being equal, the minimum overall running distance for loads (the smallest number of ton-kilometres). Quite obviously setting up transportation systems which satisfy this requirement is by no means the same thing as devising a really rational transport plan which takes into account every factor relevant to the general economy and general transport situation. Nevertheless, the ability to solve this comparatively elementary problem does make other more complicated problems connected with the rationalisation of transport easier to solve.[1]

The methods of calculation proposed below may be applied also in cases where it is possible to utilise data relating to transport costs over separate sections of railway track. For this purpose the costs of conveying a unit of load over the relevant sections must be entered in the diagrams and tables appearing in this article in place of distances between stations.

The methods suggested for establishing systems of load-flow may also be used without modification for establishing systems for routing empty waggons. If this is done, points which have a surplus of empty waggons will take on the

[1] Suggestions on 'How to obtain minimum total mileage' when setting up transportation systems were first put forward by the Soviet economist Tolstoi (see the symposium *Planning of Transportation*, Moscow, 1930; also A. N. Tolstoi, *Methods of eliminating irrational transportation in constructing operational plans*, Moscow, 1941, and Z. N. Pariiskaya, A. N. Tolstoi and A. B. Mots, *Planning Goods Traffic*, Moscow, 1947).

function of dispatch points, and points which have a deficit of these will take on the function of arrival points. Thus the methods of calculation outlined in this article will also offer guidance to reducing the running distances of empty waggons, which is such an important saving in transport.

## 1. The Graphic Method

*Rule 1.* If the railway lines which connect the dispatch and arrival points of any homogeneous load by the shortest routes do not form closed circuits, it is a simple matter to establish a system of transportation which will secure minimal overall running distances by a purely graphic method without recourse to calculating distances. It is necessary only to make certain that

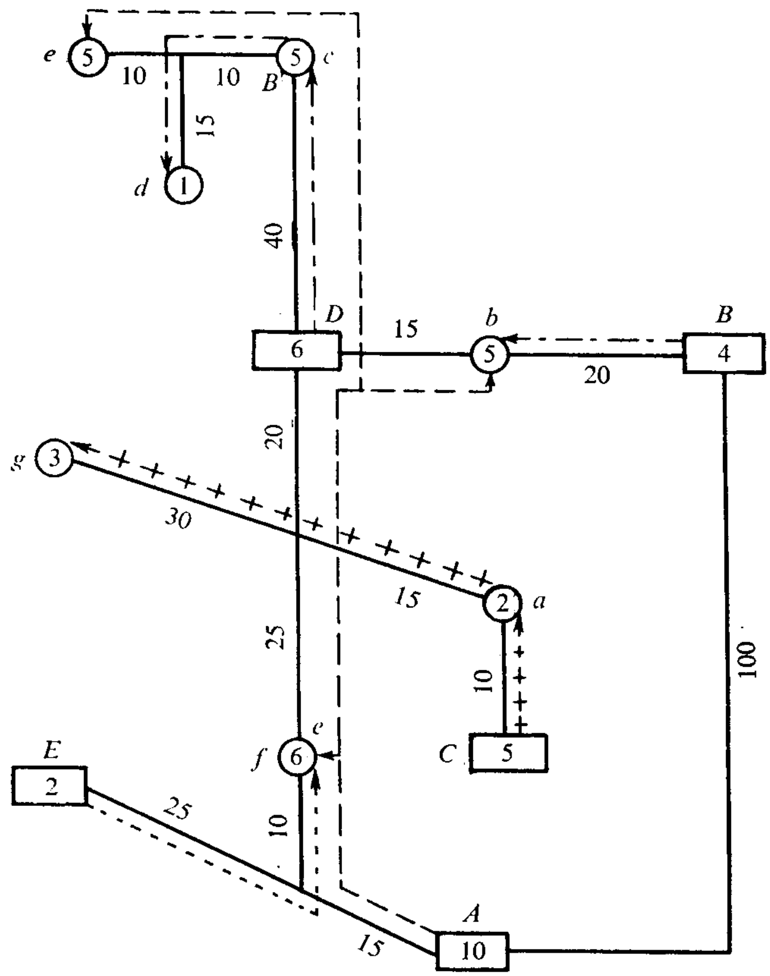


Fig. 1

there are no cross hauls (i.e. the same goods do not travel in opposite directions) when the dispatch and arrival points are being connected. The quantity of goods dispatched from and arriving at each point is presumed to be known.

A case of this sort is shown in Fig. 1. The figures inside the rectangles denote the number of units of load (in thousands of tons, trucks etc.) being dispatched, and the figures inside the circles denote the number of units received. Distances between points are also shown in the drawing. The line *AB*, which forms part of the closed circuit *ABDA*, need not be taken into consideration because the shortest routes from any dispatch point to any arrival point do not pass through it.

It is easy to see that the system in Fig. 1 shown by the various dotted lines which connect the dispatch points *A*, *B*, *C*, *D* and *E* to the arrival points *a*, *b*, *c*, *d*, *e*, *f* and *g* results in the same overall running distances as any other system would, provided that no cross hauls were permitted, and that these distances are minimal in the given conditions. In fact, if we were to link, say, point *g* to *A* instead of *C*, and dispatch the three units of load now surplus at *C* to *e*, while proportionately reducing the loads dispatched from *A* to *e*, there would be no change in the overall running distance. As far as the point of intersection of lines *AD* and *Cg* each unit of load travels the same route as before; beyond that point, three units of load from *A* now travel to *g* instead of *e* as formerly, but at the same time three units from *C*, previously routed to *g*, now go to *e*. Losses exactly counterbalance gains. We should arrive at the same result if we changed, either completely or in part, the pattern of connexions between *c*, *d* and *e* and *A* and *D*, and satisfied part of the demand at *b* or *e* with the two units now surplus at *E*, and so on.

An examination of all these cases confirms the accuracy of our original formulation, and also permits us to draw the following conclusion:

*Rule 2.* If the travel routes of loads from any one of several dispatch points (e.g. *A*, *C* or *E*) to any one of several destinations (e.g. *b*, *c*, *d*, *e* or *g*) pass through at least one common point, the overall running distance does not depend on precisely which dispatch point is connected to which destination point.

## 2. Closed Circuits and the Rule of Continuous Lines

If the railway lines linking dispatch and arrival points by the shortest routes form a closed circuit or several closed circuits (let us call such circuits 'circles'), the purely graphic method for setting up connexions becomes inadequate, and must be supplemented by calculations of the distances involved.

Let us examine Fig. 2. Let point *A* be connected to *b*, *B* to *a* and *C* to *a* and *c*. A system of connexions such as the one shown in the Fig. 2 by a line of dashes does not permit cross hauls yet nonetheless leads to excessive running distances. One may be convinced of this either by comparing the overall totals of ton- or truck-kilometres on this layout with the results obtained by using other systems of connexions, or, less laboriously, by employing the following arguments.

It follows from Rule 2 that if, still employing the same railway lines *Ab* and *Ba*, we dispatch loads from *A* to *a* and from *B* to *b* the overall running distance will be the same as with the previous routing (the travel routes of loads from $A$ and $B$ to $a$ and $b$ pass through a common point $J$). But if we connect $B$ to $b$ it is obvious that loads from $B$ must be sent not by the roundabout route $BJb$ but by the direct route $Bb$.

Thus it clearly emerges that although it was never intended to carry traffic from $B$ to $b$ via $J$, the original system of connexions does in fact lead to

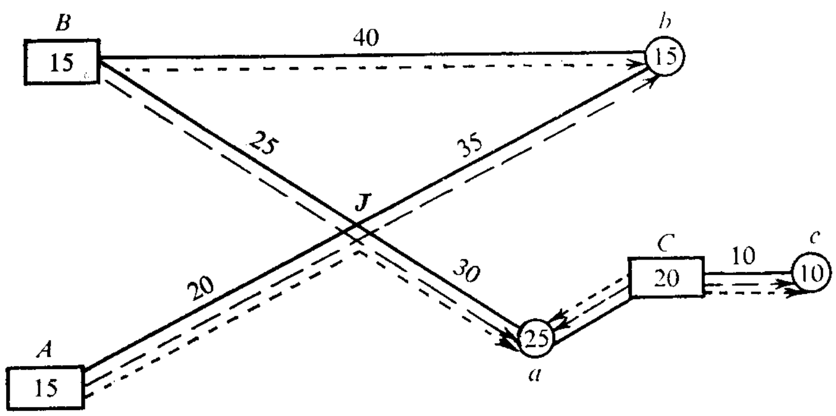


FIG. 2

exactly the same excessive mileage being covered as with any obviously irrational plan. The correct routing is shown by a dotted line.

In order to formulate a general rule based on this example we shall introduce here the following definition: we shall say that two points are joined in any given direction by a *continuous line* of load-flow if consignments encountered at any intermediate point are travelling in the same direction. Thus in the original system of connexions point $B$ was joined by a continuous line of load-flow both to $a$ and to $b$ (along the route $BJb$), but had no continuous communication with point $c$, since although consignments were to be met with on section $aC$ the direction of their travel did not coincide with the direction of load-flow on $Ba$ and $Cc$. The example we have chosen illustrates the 'rule of continuous lines' which follows.

*Rule 3.* If two points on a network are joined by a continuous line of load-flow which is not the shortest route (in our example points $B$ and $b$), a system of connexions incorporating such a line will not yield the shortest overall running distances.

In particular it follows from this that if two points are joined by continuous lines of load-flow in two directions, excess mileage will almost always occur. It is exceptional, in fact, for both routes to be equal in length, and as soon as one ceases to be the 'shortest route' the system of connexions will no longer yield the shortest overall running distances in accordance with Rule 3.

Numbers of defects in projected systems of transportation or routing of empty waggons[1] are the result of ignorance of the scientific methods of estab-

[1] Of course, in some cases this rule may be deliberately broken in the interests of some local factor, e.g. limits imposed by carrying capacity. If Rule 3 is broken, however (and

lishing connexions. The 'rule of continuous lines' often enables them to be picked out at a glance.

In the majority of cases where closed circuits are involved neither the absence of cross hauls nor the observance of the conditions following from Rule 3 are sufficient to ensure that the shortest overall running distances will be obtained. For example, the transportation system shown in Fig. 3 by a line of dashes does not break any of the general rules, but a comparison with the dotted line system shows that the latter produces shorter running distances.

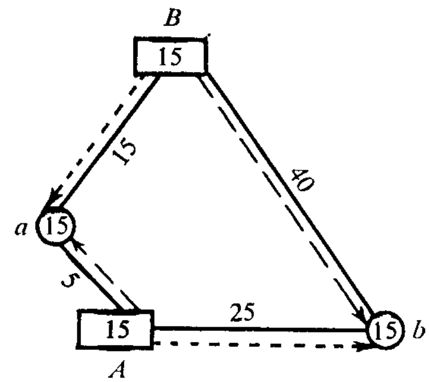


FIG. 3

Where the number of dispatch and arrival points is considerable a direct comparison of all possible alternatives to establish the shortest overall running distance is a process extremely laborious to apply in practice and may even be impossible if the number of alternatives is exceptionally great. A special method is needed, therefore, which will make it possible to arrive at the most efficient system of connexions by combining graphic representation with relatively simple calculations. This will remove the need to establish total running distances for each alternative. Such a graphico-analytical method is the method of *circle differences.*[1]

### 3. The Circle Differences Method

If an analysis is made of the systems of connexions shown in Fig. 3 it is easy to see that selection of the most efficient alternative may be made without recourse to calculations of total ton- or truck-kilometres. It is sufficient to compare 'gains' (+) and 'losses' (−), and these we obtain by changing the direction of travel of any unit of load in relation to the original version of the system. Any conclusion which is true for one unit of load (tons or trucks) will also be true for all subsequent values.

By re-routing a unit of load from $A$ to $b$ instead of $a$ as in the original version (shown by the line of dashes) we 'gain' 5 kilometres from $A$ to $a$, which the unit in question no longer has to cover, and 'lose' 25 kilometres, which is the distance this unit must now be sent. But as point $b$ will now be receiving a unit of load from $A$, the unit previously sent from $B$ to $b$ becomes superfluous, and this can (and must) be sent to $a$ in exchange for the unit which has changed direction. To do this will result in a 'gain' of 40 kilometres ($Bb$) and a 'loss' of 15 kilometres ($Ba$). Adding, we obtain

indeed if any increase at all is made in the overall running distances), it should be for a definite reason.

[1] Tolstoi applies the term 'graphico-analytical' to the method outlined by us in Section 1.

$+5-25+40-15 = +5$. The gain exceeds the loss by 5 kilometres, and therefore our modification of the system has been advantageous.

In this way we have been able to show that the system represented by the dotted line is the preferable one simply by comparing the distances between *A*, *B*, *a* and *b*, without having had to calculate the total running distances involved.

Calculation of the differences between 'losses' and 'gains' obtained by changing the direction of travel of a unit of load also forms the basis of the 'circle differences' method.[1]

We shall demonstrate the application of this method with an actual example (see Fig. 4).

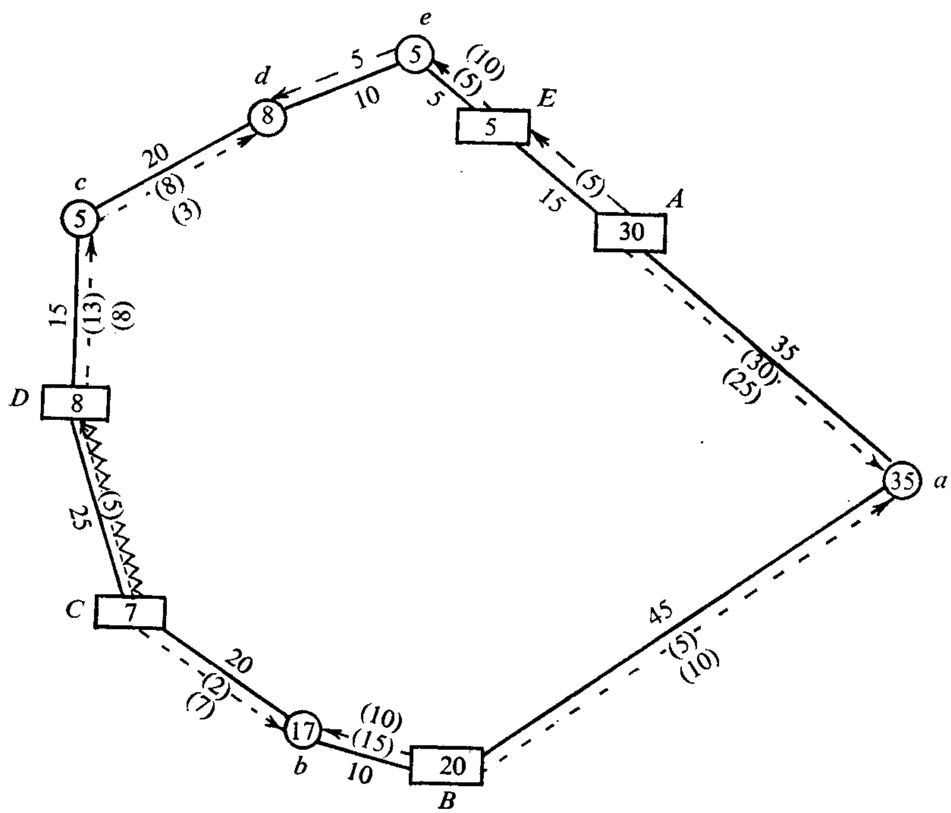


FIG. 4

The work of setting up an efficient system of connexions is begun by connecting dispatch points and arrival points in an arbitrary fashion, taking care only that the elementary rules (concerning convergent routes and continuous lines) are not infringed. It is best to begin with the largest dispatch or arrival point, e.g. point *a*. This we connect to the nearest dispatch points,

[1] Both in its original form, as described in the work of Tolstoi, and in the version outlined below. This same comparison of 'losses' and 'gains' also forms the basis of the methods of Kantorovich and Gavurin (see the symposium *Problems of raising transport efficiency*, USSR Academy of Sciences, 1949) as well as of a number of methods of resolving 'the transportation problem', in foreign literature.

routing 30 units of load for *a* from *A* and 5 from *B*. The remaining 15 units available at *B* must evidently be sent to *b*, otherwise we shall find ourselves with converging load-flows. Continuing to connect all the dispatch and arrival points one after the other in such a way as to prevent load-flows converging, we arrive at the system indicated in Fig. 4 by the dotted lines.[1]

This system leaves certain details unsettled. It does not show whether part of the traffic is sent from *B* to *d*, or whether *d*'s demands are entirely satisfied by *D* and *c*'s by *C*. Such precision, however, is not needed for our purposes, because Rule 2 shows that these factors would have no effect on the total running distances anyway.[2]

Does the system we have obtained yield the best results by guaranteeing the smallest amount of ton-kilometre work? If not, how may it be modified in order to do so?

Let us agree to term any change in the system of connexions an *anti-clockwise advance* if it has the effect of reversing the direction of loads originally travelling in a clockwise direction. Any change of the opposite sort we shall term a *clockwise advance*. It is easily seen that the 'advance' of any unit of load is inevitably accompanied by changes in the numbers of consignments on all sections of the circle.

For example, suppose we re-route to *b* a unit of load originally travelling from *B* to *a* (clockwise advance). It is obvious that this will make available at *C* a unit of load which was previously sent to *b*. This will have to be sent in a clockwise direction until it reaches the first point receiving consignments from the opposite direction, i.e. to *e*. The unit which *e* used to receive from *E* will now have to be sent to *a*, i.e. to the consumption point from which we began the 'advance'. On all the sections over which in the original version of the system loads were travelling in a direction opposite to the advance (in this case an anti-clockwise direction) the load-flow will be reduced by one unit, and on all the remaining sections of the circle it will be increased by one unit.

The example of advancement we have chosen leads us to the following conclusion. The general 'gain' derived from advancing a unit of load (the saving in ton- or truck-kilometres) is equal to the length of the sections over which loads were travelling in the direction opposite to the advance in the original system of connexions, while the 'loss' is equal to the length of the remaining portions of the circle, i.e. to the length of those sections over which loads were travelling in the same direction as the advance, and of those sections over which no loads were travelling at all (*'free' sections*).

[1] Besides distances, figures showing traffic density are also entered on the diagram in brackets interrupting the dotted lines (e.g. 30 on section *Aa*, 5 on *Ba*, etc.). The significance of the line of dashes and zigzag line will be explained below.

[2] Once the general outline of the system is established, the suitability of connexions which are not going to have any bearing on the overall running distance can be decided on other considerations, such as the principle of concentrating connexions, which facilitates a wide application of dispatch routing.

If the sections over which loads travel in a clockwise direction are longer than those over which they travel in the opposite direction (as in our example), we call the clockwise direction *predominant*. In the converse case the anti-clockwise direction will be predominant.

From all that has been said above it is possible to deduce the following rule for connecting dispatch and arrival points which are situated on the same closed circuit ('circle'):

*Rule 4.* If the load travel routes are on a closed circuit, and the difference between the length of the sections carrying loads in one direction and the length of the sections carrying loads in the opposite direction is less than the length of the 'free' sections; or, in other words, if the length of the sections carrying loads in the predominant direction is less than the length of the rest of the circle, then the corresponding system of connexions will yield the shortest overall running distance. If this condition is not present, an advance must be made in the direction opposite to the predominant direction.

In our example the length of the sections carrying loads in a clockwise direction is equal to 105 kilometres; in the opposite direction, 70 kilometres; and not carrying loads at all, 25 kilometres. Since 105 is greater than $70+25$ it follows that it is expedient to make an advance in an anti-clockwise direction.

But if it is expedient to advance one unit of load in an anti-clockwise direction, it seems reasonable by the same token to move two or more. How many units, then, should we advance? Evidently the expediency of advancing units will not be called in doubt while the original balance of losses and gains continues to hold good.

The losses and gains in our example will be the same for units 2, 3, 4 and 5; but by the time unit 6 is reached the balance of pluses and minuses will have altered, since all the loads which had been travelling previously from point $C$ in a clockwise direction (5 units) will now prove to be travelling in the opposite direction, and section $CD$ will have become free.

A general rule to establish the quantity of loads which should be advanced may be stated in the following terms:

*Rule 5.* The number of units of load which should be advanced on the basis of Rule 4 is equal to the lowest density of traffic on any section carrying traffic in the predominant direction of the original system.

This 'limiting' section may be at once identified as $CD$ in our drawing.

Once the direction and extent of the advance needed are ascertained, it only remains to put it into effect, i.e. to plot on the scheme the new system of connexions resulting from the relevant modifications of the original system. The solution of this problem presents no difficulties. One can begin at any dispatch or arrival point. Let us take point $a$, say. If we advance 5 units in an anti-clockwise direction, this will means that we must send 10 units from $B$ to $a$ instead of 5.[1] It follows that out of the 30 units previously received

[1] The new density figures are set out in Fig. 4 beside the old.

by $a$ from $A$, 5 are now travelling in an anti-clockwise direction to the nearest point which had been receiving loads in the predominant direction (opposite to the advance), i.e. to point $d$. The release of 5 units, which had previously been coming to $d$ from the direction of $C$ (whether directly from $C$ or $D$ is immaterial), makes it possible to send 5 units from $C$ to $b$ in exchange for the ones which we sent from $B$ to $a$ instead of $b$.

The new transportation system is represented in Fig. 4 in the following manner: sections which were free in the original version but which are now carrying traffic, are denoted by a line of dashes; dotted lines are retained for sections over which traffic has not changed direction, and the new indices of density are marked; and for sections where it has been possible to reduce traffic the dotted lines are cancelled by a zigzag line.

When checking the new system of connexions, we obtain the following results: the length of sections carrying loads in a clockwise direction is 80 kilometres, and in the reverse direction 95 kilometres, while the length of free sections comes to 25 kilometres. The system we have obtained guarantees minimum overall running distances in accordance with Rule 4.

It is possible to simplify the testing of the new system. For this purpose one must take double the length of the free sections resulting from the advance, and compare this figure with the result calculated for the original version, since this is exactly the amount by which the difference between the sections of the predominant direction and the remainder of the circle will have changed.

In our example this section is $CD$. Twice its length is equal to 50 kilometres, and the difference between the length of the sections carrying traffic in the predominant direction and the remainder of the circle was $105-70-25 = 10$ kilometres. It is clear that this difference will cease to be positive in the new system of connexions (since 50 is greater than 10) and that the condition of Rule 4 will be satisfied as a consequence.

If a test of the second version of the plan showed the necessity for a new advance, it would be necessary to construct a third version, and once again to test its suitability by means of the same simple calculation. If we adopt this procedure we shall eventually arrive at a system of connexions which will ensure that the shortest overall running distance is obtained, wherever the dispatch and arrival points happen to be situated on the closed circuit.

## 4. General Application of the Graphico-Analytical Method

The method of circle differences permits a system to be constructed which will link dispatch and arrival points in such a way as to ensure that minimal total running distances are obtained whatever the layout and the situation of the dispatch and arrival points.

The normal procedure for establishing such a system is as follows. First we take those sections of the transport network for which the system of connexions can be set up by making use of the graphic method (Rules 1 and 2). Once these have been disposed of we can set about establishing preliminary links between the dispatch and arrival points which lie on closed circuits (circles).

These preliminary links can be established by rule of thumb; travel lines must not be permitted to converge, of course, and the rule of continuous line must be observed. Then all the closed circuits must be tested and corrected one after the other by the circle differences method. In the course of this it may happen that changes made at a later stage impair the efficiency of systems of connexions adopted for other circuits which have already been

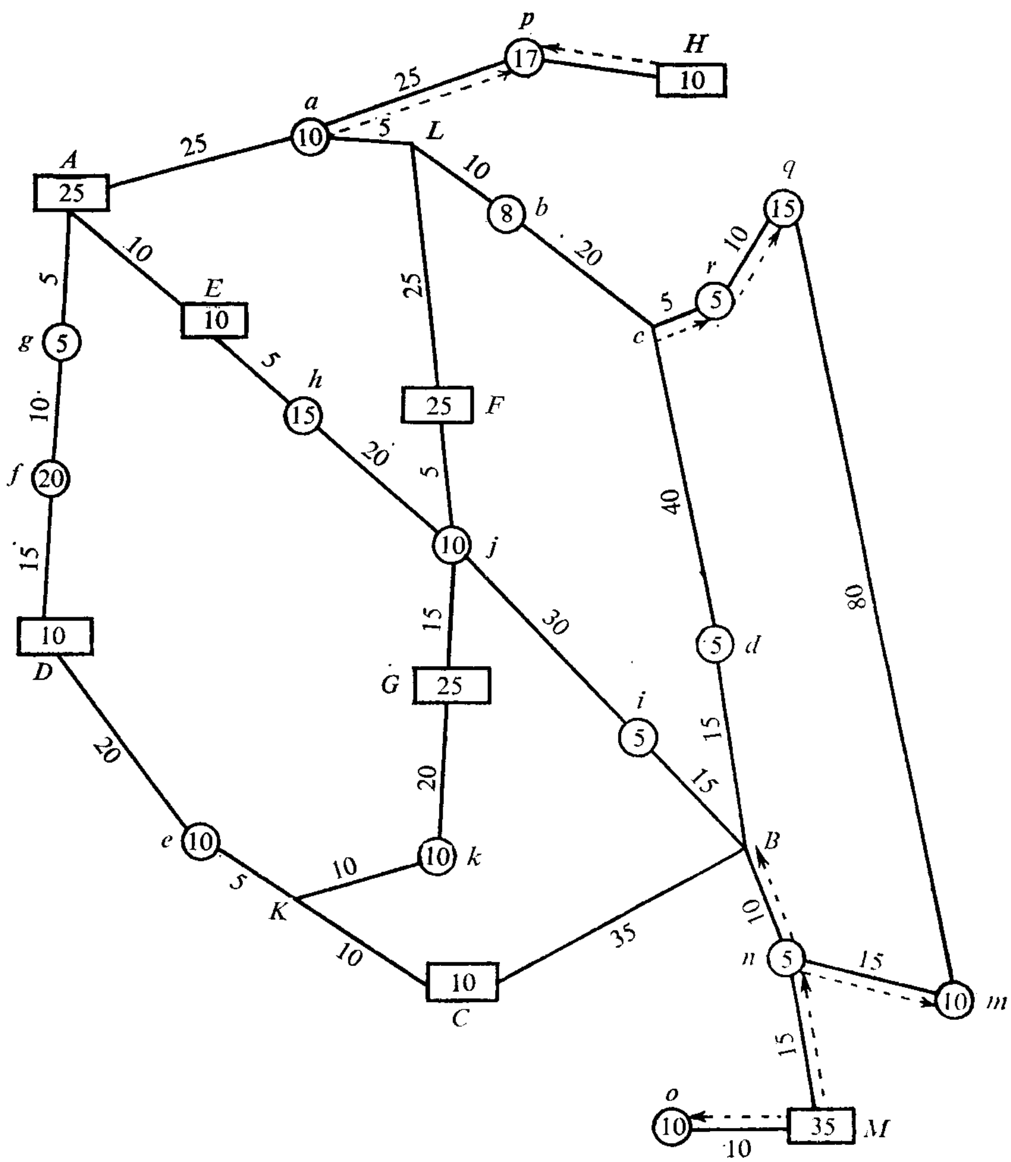


FIG. 5

tested and corrected. In this event, these systems will have to be re-tested. We proceed in this way until every circle satisfies the conditions of Rule 4.

Let us take an example to show the successive steps needed in the normal application of the methods described above in setting up a system of connexions (see Fig. 5).

We note first of all that the extreme right-hand line *mq* does not form part of a closed circuit because the shortest routes for goods from any dispatch to any arrival point do not pass through it. Therefore we apply the graphic method, not only to stations *p* and *H*, but also to *r* and *q*, and to points *o*, *M*, *n* and *m*. Consequently we can plot the load-flows shown in the drawing by dotted lines, without any calculations.

For example, it is obviously necessary to dispatch from *M* 10 units of load to *o*, 5 to *n*, and 10 to *m*. The remaining 10 units at *M* will be dispatched to other points, and these will have to pass through *B* (we have already established that there would be no point in using the section *mq* with the present arrangement of dispatch and arrival points). The distance loads must run from *M* to *B* does not depend on what stations beyond *B* are connected

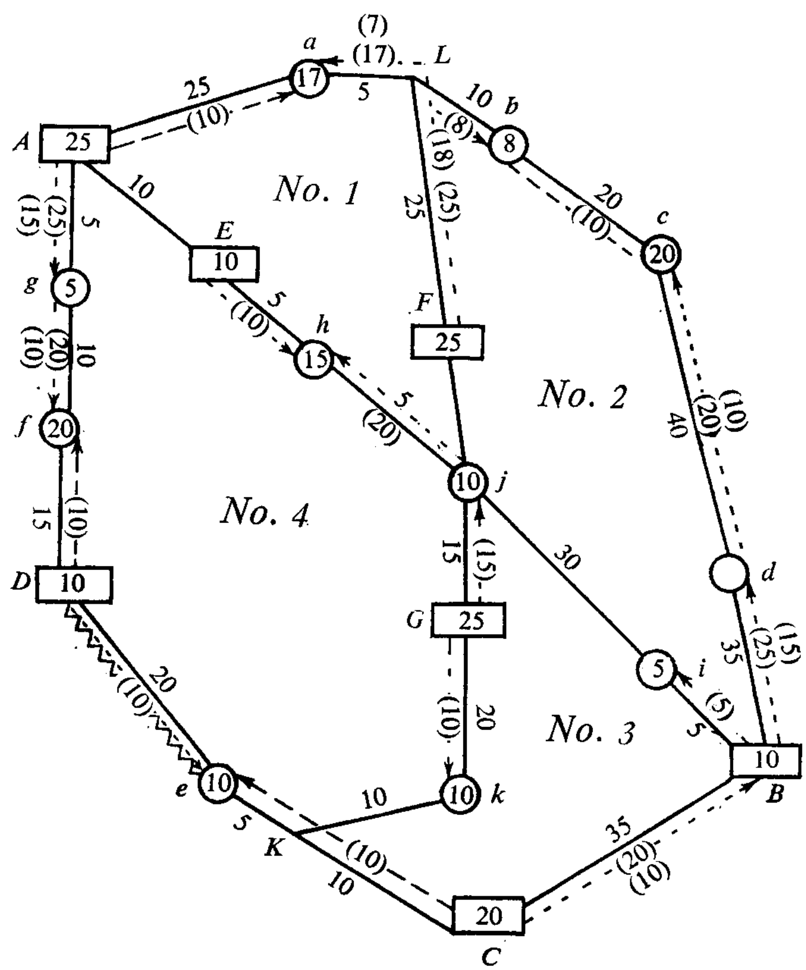


FIG. 6

to *M*. In all subsequent calculations, therefore, we may ignore line *MB* and regard the junction *B* as the source of 10 units of load.

Similarly point *c* will be regarded henceforth as a junction receiving the 20 units of load which are in practice intended for *r* and *q*, and point *a* as an arrival point for 17 units, since apart from the 10 units needed to satisfy *a*'s own requirements, 7 units for *p* must also inevitably pass through this point. Thus we can substitute junctions for those parts of the network to which the graphic method is applicable.

Subsequent steps in developing the system of connexions are shown in Fig. 6. *B*, *c* and *a* are shown as dispatch or arrival points for calculated quantities of load (10, 20 and 17 units respectively) in accordance with our substitution of these junctions for the three groups of stations *o*, *M*, *n* and *m*, *r* and *q*, and *p* and *H* respectively. Those stations of the network which did not form parts of closed circuits have been omitted entirely.

We make a first draft of the system at random. Let us say, for example, that, having begun with the major dispatch points *A*, *F* and *G*, we have arrived at the system represented in the diagram by a dotted line We set about testing the circles.

To do this let us regard the points where the circuit we are testing intersects with the other circles as dispatch points or arrival points depending on whether loads, departing from these points or approaching them from other parts of the network, enter sections of the circle being tested, or whether on the other hand loads arriving at these points from the circle being tested are either unloaded there or depart to neighbouring sections of other circuits. Thus, for example, in the given system of connexions *L* will be, for the 'large' circuit *ALcBCDA*, a dispatch point for the 25 units of load travelling to it from *F*, whereas for the circle *ALjA* the same point *L* will be a point of arrival for 8 units, since through it 8 units leave for the neighbouring section *Lb*. We can refer to the circuits we are testing in a briefer manner by putting in brackets the numbers of the segments which each encompasses. Thus, for example, the large circle *ALcBCDA* may be denoted by (1, 2, 3, 4), the circle *ALjA* mentioned above by (1), the circle *ALcBjA* by (1, 2), etc. We begin our test with (1, 2, 3, 4). The total length of sections carrying traffic in a clockwise direction (let us call such sections *positive*) adds up to 10 kilometres (section *Lb*), and the total length of sections carrying traffic in the opposite direction (let us call these *negative*) adds up to 130 kilometres. The length of the free sections is 75 kilometres. The length of the negative sections exceeds the rest of the circle by 45 kilometres. A clockwise advance is therefore necessary (Rule 4). 10 units of load should be advanced, according to Rule 5 (the limiting section is *De*). The results of this on the circuit (1, 2, 3, 4) are shown in the diagram thus: where the direction of traffic has not changed, the dotted line has been left and the new load densities inserted, e.g. (15) instead of (25) on *Ag*, (10) instead of (20) on *cd*, etc.; traffic on previously free sections is denoted by a line of dashes (*Aa*, *bc*, *Ce*, and *Df*); and where traffic has ceased (on section *De*) the dotted line is cancelled by a zigzag line.

Testing the new system we obtain the following lengths: positive sections, 85 kilometres; negative sections, 110 kilometres; free sections (only *De* in fact), 20 kilometres. Since 110 is greater than 85+20, a further advance is necessary.

If we advance another 7 units of load in the same direction (the limiting section is now *La*) we obtain a system of connexions which differs from the second version in having traffic (7 units) on *eD* and none on *La*. This system is shown by dotted lines in Fig. 7.

On being tested a third time the circle (1, 2, 3, 4) is seen to satisfy Rule 4, since twice the length of the now free section *La* (10 kilometres) is greater than the difference found in the previous test (5 kilometres).

The system shown by dotted lines in the Fig. 7 can be arrived at by an even shorter method, and for this it is necessary to modify Rule 5 in such a way as to ensure that no more than one advance is needed on any closed circuit to obtain a system of connexions which will satisfy Rule 4. Such a modification of Rule 5 is quite feasible. As we have seen, the difference

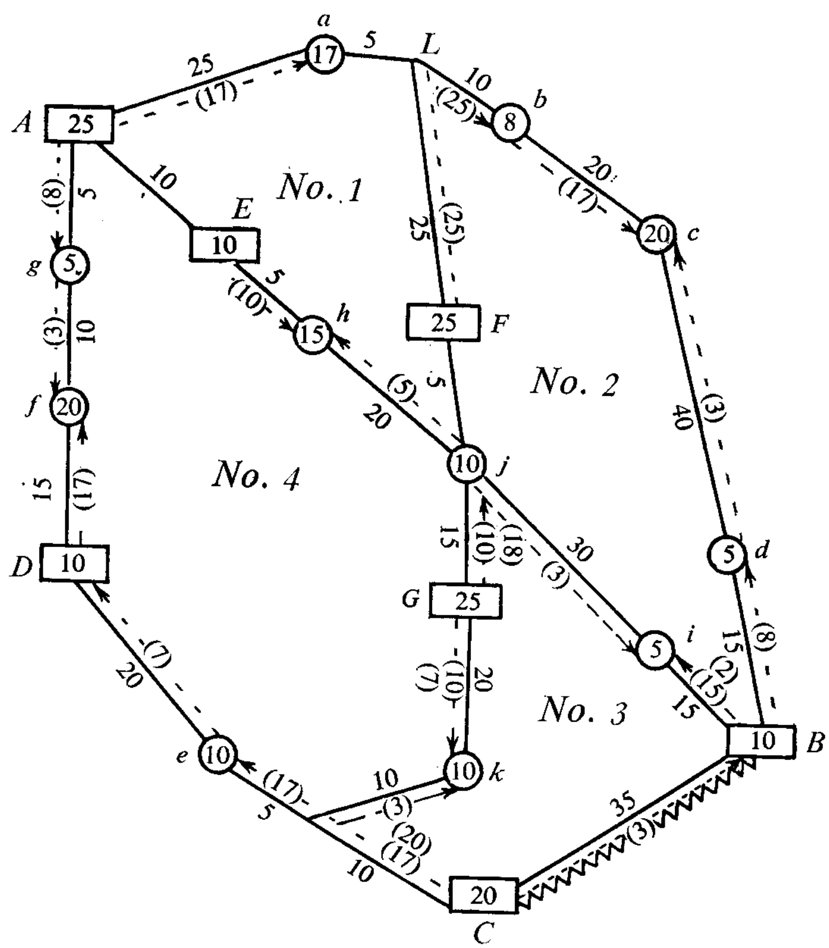


FIG. 7

between the length of the sections carrying traffic in the predominant direction and the rest of the circuit, which shows a gain when a unit of load is advanced, is reduced every time one of the sections carrying traffic in the predominant direction (i.e. the section where traffic is least dense) becomes free, and reduced, moreover, by twice the length of that section. We shall therefore find it useful to proceed in accordance with this rule:

*Rule 5a.* To establish how many units of load should be advanced, twice the length of the section carrying the least traffic in the predominant direction must be subtracted from the difference between the length of the sections carrying traffic in the predominant direction and the rest of the circuit. If this difference is greater than zero, we subtract from it twice the length of the section carrying the next least volume of traffic in the predominant direction, and we proceed in this way until our subtraction results in a difference less than zero. The number of units of load which should be advanced is equal to the density of traffic on the last section which we have to treat in this way.

Making use of this rule, we calculate the difference between the lengths of negative sections and the remainder of the circuit (1, 2, 3, 4) as 45 kilometres (see p. 334), and proceed as follows: We subtract twice the length of the section carrying the least volume of traffic in the anti-clockwise direction (i.e. *De*, see Fig. 6) from 45. Since $45-40 = 5 > 0$, we subtract from 5 twice the length of section *La*, which is next in order of traffic density, and we obtain $5-10 = -5 < 0$. We therefore conclude that 17 units of load, equal to the traffic density on *La* should be advanced. We may now put into effect a clockwise advance of 17 units, commencing, let us say, by re-routing to *a* 17 of the 20 units which went from *A* to *f* in the original version. The resulting system is shown in Fig. 7.

It will be found that Rule 5a considerably reduces the work involved, since it obviates the necessity of producing intermediate versions of the system of connexions.

Let us now test circuit (1). Calculations of the lengths of positive, negative and free sections give respectively $+45$, $-30$, 20, and so no advance is necessary. We reach the same conclusion for circle (2): $+70$, $-55$, 35. A test of (3) gives $+25$, $-70$, 40, and since the length of the negative sections exceeds the sum of the lengths of the positive and free sections by 5 $(70-40-25)$, a clockwise advance is called for. The number of units to be advanced is 3, namely the traffic density on *CB*, since we obtain a negative quantity if we subtract twice the length of this section from 5. A line of dashes indicates where these changes have taken place, and a zigzag cancelling the dotted line from *C* to *B* shows that this section has now become free.

Let us now see whether the alterations made to traffic on circle (3) call for any corrections to the traffic systems on circuits already tested. On any circle an increase in the difference between the length of sections carrying traffic in the predominant direction and the remaining sections may cause traffic to appear on sections which were previously free, or may transform negative sections into positive or positive into negative. The appearance of a new free section, however, cannot lead to an increase in the difference with which we are concerned, and for this reason the unoccupied section which our alterations to circuit (3) have produced on the large circle (1, 2, 3, 4) does not require further test calculations.

The advance which we have made on circle (3), therefore, necessitates a second test only of circuit (2), which has a common section *ji* with circuit (3). (Circuit (4) has not been tested at all as yet.) This new test of circuit (2) gives us +70, −85, 5, and therefore a clockwise advance is necessary. If we advance 3 units (the limiting sections are *ji* and *cd*), we arrive at a new version of the system of connexions. This is shown in Fig. 8 by the dotted lines.

Traffic has appeared on section *jF*, and sections *ji* and *dc* have become free. Twice the length of sections *ji* and *dc* is greater than 5. It follows that it is not necessary to advance more than 3 units. But this has produced a change in the situation or circuit (2), since traffic has now appeared on section *jF*, and this requires a second test of circle (1). The result of this test

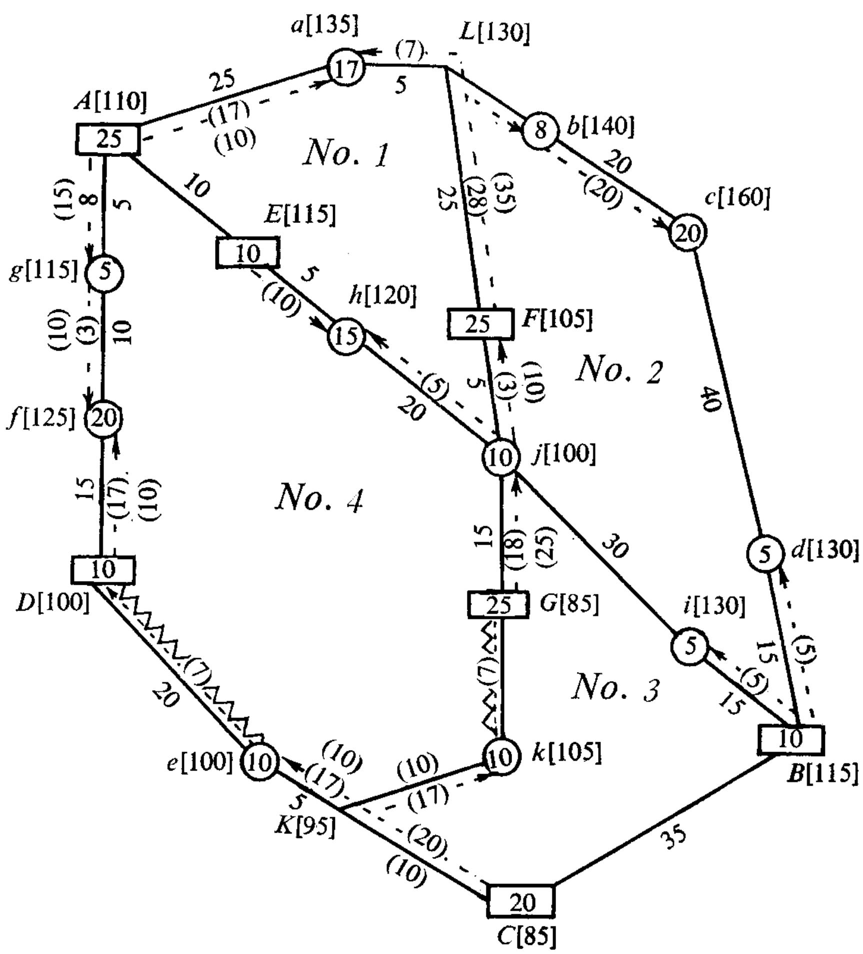


FIG. 8

will be obvious. The length of the negative sections will be increased at the expense of the free sections by only 5 kilometres, and therefore the relationship between the length of the sections carrying traffic in the predominant direction (in this case positive) and the rest of the circle will remain undisturbed; circle (1) will continue to satisfy the conditions of Rule 4. In this way we have examined all the consequences of our modifications of circuit (3).

Continuing the test of the circuits, we note that no change is required on circle (4), since the relevant calculations show +65, −60, 10. Will it be necessary to re-connect the points which are situated on circle (1, 2)? A simple rule enables us to decide this question without resorting to further calculations. Before formulating this rule let us agree to call those sections which form part of two circuits their *common line*, and the remaining sections of the same circuits their edges.

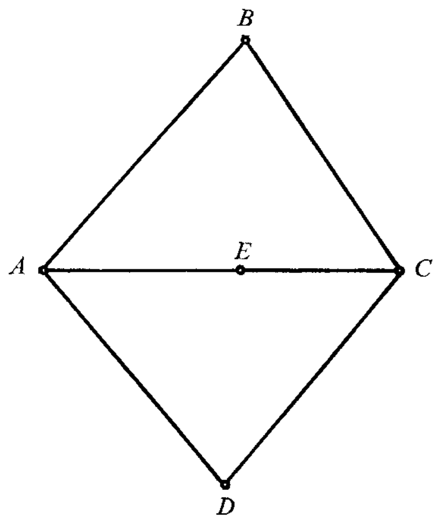

FIG. 9

For example, line *AEC* in Fig. 9 is a common line for circuits *ABCEA* and *AECDA*; *ABC* and *CDA* are the edges of these two circuits. At the same time *ABC* is a common line for the circuits *ABCEA* and *ABCDA* and the lines *AEC* and *CDA* are their external parts. Similarly, *CDA* is the common line, and *ABC* and *AEC* are the edges of the circuits *ABCDA* and *AECDA*.

*Rule 6.* If on a line common to two circuits (e.g. *AEC*) no free sections occur, and both the circuits (in our case *ABCEA* and *AECDA*) satisfy the conditions of Rule 4, then any circuit formed from their edges (e.g. *ABCDA*) will also satisfy these conditions.

NOTE. It is not necessary for all the areas encircled by the two original circuits to lie within the circle formed by their edges. For example, if there were no free sections on *ABC* and tests showed that the circuits *ABCEA* and *ABCDA* met the requirements of Rule 4, testing the circle *AECDA* would be superfluous.

Now let us return to Fig. 8. After the changes we have made *jFL*, the line common to circuits (1) and (2), has no free sections. Consequently, by virtue of Rule 6 circle (1, 2), formed by the edges of circles (1) and (2), does not call for any further alterations to the system of connexions.[1]

Calculations of positive, negative and free sections are required for circuit (2, 3), and the results of these (+95, −35, 75) show that this circuit satisfies Rule 4.

[1] If we check circle (1, 2) by the method outlined earlier we see that Rule 6 is correct for this case, for we obtain +80, −20, 85.

No calculations are required for circuit (3, 4) because this circuit consists of the edges of (3) and (4) which have already been tested, and because there are no free sections on their common line *Kj*.

A test of circle (1, 4) shows that an anti-clockwise advance is necessary (the figures here are +85, −70, 5). The number of units to be advanced is indicated here by sections *Gk* and *eD*, with a traffic density of 7 units. We show the consequences of this advance of 7 units of load by cancelling the dotted lines running from *e* to *D* and from *G* to *k* by zigzag lines, and showing by a line of dashes where traffic has appeared on section *La*.

The appearance of traffic on *La* calls for a second test of circuits (1, 2, 3, 4) and (1). For (1, 2, 3, 4) it is sufficient to establish that the total length of negative section (which has increased by 5 kilometres) is now 25 kilometres, and therefore clearly less than the rest of the circle. We obtain the same result for circle (1). It follows that no changes are needed in the system connecting the points on these two circuits.

There remain to be tested now the circuits (1, 2, 3), (2, 3, 4), (3, 4, 1) and (4, 1, 2). Calculations are needed only in the case of circuit (3, 4, 1), and the results of these (+55, −65, 85) show that this circuit answers the requirements of Rule 4. Rule 6 shows that no changes are required in the system connecting the points situated on the remaining circuits. To save reiterating the same argument for every circuit, let us take circle (2, 3, 4) as a representative example. This circle is made up of the edges of circuits (1, 2, 3, 4) and (1) (i.e. of *LcBCKDA* and *LjA*), and their common line *AaL* has no free sections on it. It follows from this that (2, 3, 4) does satisfy the conditions of Rule 4.

We have now tested and corrected all the circuits shown in Figs. 6, 7 and 8. The dotted lines in Fig. 8 which are not cancelled by zigzag lines show the system guaranteeing the shortest overall running distance in its final form.[1]

## 5. The Use of Potentials in Testing Transportation Systems

The volume of work involved in the application of the graphico-analytical method quickly grows with the increase in the number of closed circuits which need to be tested and corrected by the calculation of circuit differences. Moreover, where the layout of the transportation network is very complicated the possibility of overlooking one or two circuits is not to be excluded. It is therefore recommended that transportation systems set up by means of the methods described above be submitted to a final check by the calculation of potentials.[2]

[1] In practice all the work may be done on one diagram by successively deleting lines showing load-flow on sections which become free and adding lines on sections where our modifications have caused traffic to appear.

[2] See L. V. Kantorovich and M. K. Gavurin, op. cit. In our opinion it is not always desirable to calculate potentials while actually setting up transportation systems and to enter them on the diagram for each intermediate version. More work may be involved in setting up a traffic plan than in using the circuit differences method.

Let us briefly explain the concept of potentials. Let each point on a network $i, k, l \ldots$ ($i, k, l \ldots$ may be either dispatch or arrival points, or junctions) be allotted a value $P$ ($P_i, P_k, P_l \ldots$).

If the traffic system ensures the shortest overall running distance for loads, the values of $P$ may be selected in such a way as to satisfy two requirements:

I. If on section $ik$ traffic is travelling from $i$ to $k$, the difference $P_k - P_i$ must be positive, and equal to the length of the section $ik$.

II. If section $ik$ is free, the absolute value of the difference $P_k - P_i$ must not be greater than the length of section $ik$.

Values of $P$ satisfying these requirements are called *potentials*. The possibility of constructing a system of such potentials where any traffic plan is being considered is not only a necessary but also a sufficient condition of the establishment of the shortest overall running distances for loads.[1]

Still using our previous example we shall demonstrate the use of potentials in testing the adequacy of networks. Let us determine the potentials for the final version of the traffic plan shown in Fig. 8. To one of the points on the network, to $j$ let us say, we attribute an arbitrary potential of 100.[2] We obtain the potentials of the neighbouring points linked to $j$ by traffic routes either by adding the lengths of the relevant sections to 100 if the traffic passes over them away from $j$, or by subtracting these lengths if the traffic is moving in the opposite direction. This procedure ensures that requirement I (above) is observed, and gives as a result a potential of 105 for $F$, 120 for $h$, and 85 for $G$ (see the numbers in square brackets in Fig. 8). By adding the distance from $F$ to $L$ to the potential of $F$ we obtain the potential of the junction $L$ (130), and it is now a simple matter to assign potentials to $a$ (135) and ($b$) (140). If we continue our calculations in this fashion we shall obtain potentials for all the points on the network which are linked to $j$ by traffic routes.

For stations $B$, $i$ and $d$, and stations $C$, $K$, $k$ and $e$, which are not linked by traffic routes to any points on the remainder of the network, we can obtain potentials in the following way. To determine the potentials of the group of stations $B$, $i$ and $d$ we commence with the station which is nearest to points which have potentials already, in this case $i$. It is obvious that $i$'s potential may not be greater than 130 (100, the potential of $j$, plus 30, the length of section $ji$) or less than $100-30 = 70$, otherwise requirement II will not be observed. If we try out a potential of 130, we obtain 115 ($130-15$) for $B$ and 130 ($+115+15$) for $d$. As may be seen from Fig. 8, a potential of 130 satisfies requirement II, since $160-130 = 30 < 40$.

[1] The relevant theorem was proved by Kantorovich in 1942 (see *Reports of the USSR Academy of Sciences*, Vol. 37, Nos. 7-8, pp. 227-9, 1942). It is easy to show that the possibility of constructing systems of potentials is equivalent to satisfying the requirements of Rule 4.

[2] It is convenient to take a positive number large enough to ensure that only positive numbers are obtained in the successive calculations.

If requirement II had not been satisfied with respect to section *cd*, we should either have had to try out different potentials for *i* within the range of values permissible (70 to 130) until we hit upon one which did satisfy requirement II, or, if that proved impossible, recognise that the system we had constructed was not the 'best', and that it would itself have to be changed. For example, if the potential of point *c* had been 80 instead of 160, the difference ($130-80 = 50$) would have been greater than the length of the section (40). In this case it would have been possible to reduce *i*'s potential by 10, which would also have meant the reduction of *d*'s potential by 10 (120 instead of 130), which would *ipso facto* have ensured that requirement II was met. But if *c*'s potential had been less than 30 or greater than 170, no amount of adjustment could have brought about a simultaneous observation of requirement II with respect to both sections *ij* and *cd*. The plan would have had to be rejected as not securing minimal overall running distances.[1]

Let us assign potentials to the group of stations *C*, *K*, *k* and *e*. Let us attribute 105 ($85+20$) to *k*; we obtain thereby the following potentials: *K* 95, *e* 100, *C* 85. It is easy to see that no adjustment of these figures is necessary, since requirement II is satisfied for both sections *BC* and *De*. At this point we may conclude our calculations of potentials. It follows from the nature of these calculations that requirement I is satisfied, and a comparison by means of Fig. 8 of the length of any free section with its two terminal points will show that requirement II is also satisfied. If the system we were testing did not secure the miminum overall running distance for loads, the simultaneous satisfaction of both requirement I and II would not have been possible.

### 6. Analytical (Tabular) Methods for Establishing the Shortest Running Distances for Loads when Setting up Transportation Systems

A. N. Tolstoi has suggested a remarkably simple and convenient method of drawing up traffic plans which will secure the shortest overall running distance for loads in cases where only two dispatch points are involved: the method of *successive differences*.[2] This method has great practical importance, since quite complicated transportation systems may frequently be reduced, either wholly or in part, to problems involving two dispatch points or two arrival points, particularly if junctions are substituted for groups of stations. Let us demonstrate this method.

[1] For greater detail on the construction of potentials see L. V. Kantorovich and M. K. Gavurin, op. cit.

[2] For this method see also T. S. Khachaturov, 'Fundamentals of railway transport economics, pt. 1', *Transzheldorizdat*, 1946, pp. 363-6; Y. I. Koldomasov, 'Fundamentals of transportation planning on railways', *Transzheldorizdat*, 1949, chap. 6; E. D. Khanukov, 'Transport and the location of production', *Transzheldorizdat*, 1955, chap. 9; S. K. Danilov, ed., 'Transport economics', *Transzheldorizdat*, 1957, chap. 7.

Let $A$ dispatch 70 units of load, and $B$ 80. The arrival points $a$, $b$, $c$, $d$, $e$ and $f$ require respectively 10, 20, 30, 15, 35 and 40 units of load. Table 1 shows the distances between these points and $A$ and $B$. Whatever the layout of the railway track on which these points are situated, the connexions to be made between $A$ and $B$ and $a$, $b$, $c$, $d$, $e$, $f$ may be ascertained in the following way:

TABLE 1

| Dispatch points | Distances between dispatch and arrival points | | | | | |
|---|---|---|---|---|---|---|
| | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
| $A$ | 30 | 20 | 40 | 10 | 50 | 15 |
| $B$ | 50 | 30 | 20 | 25 | 5 | 10 |
| Gain' (+), 'loss' (−) on being connected to $A$ | +20 | +10 | −20 | +15 | −45 | −5 |

For each arrival point $a$, $b$, $c$, $d$, $e$ and $f$ we subtract its distance from $A$ from its distance from $B$. The differences will indicate the 'gains' (+) and 'losses' (−) which we should obtain for each unit of load if we connected these arrival points to $A$ in preference to $B$. Let us arrange the arrival points in the order in which they should be connected to $A$ to obtain maximum advantage, and then join them to $A$ one after another until all the loads dispatched from $A$ (70 units) are accounted for. This order is: $a$, $d$, $b$, $f$, $c$, $e$. Taking into consideration the consumption requirements of each point we are now brought to our final conclusions: $a$, $d$ and $b$ must be entirely supplied from $A$; $f$ must receive 25 units from $A$ and 15 from $B$; the remaining arrival points ($c$ and $e$) must be connected to $B$.

The method of successive differences is a purely analytical or tabular method. Its application does not involve plotting load-flows on a diagram of the track layout, and the only initial data needed are the distances between the dispatch and arrival points. When dispatch and arrival points are situated in a way that would involve the examination of a number of circles by the graphico-analytical method the method of successive differences offers a considerable simplification.

The question will suggest itself whether this purely analytical method may be applied more generally to situations where there are more than two dispatch or arrival points. The method put forward by L. V. Kantorovich for the solution of a number of technological, organisational and planning problems with the help of *resolving multipliers* or *ratings* makes it possible to give an affirmative answer. We have already met a specific instance of the use of ratings in the solution of the problem of drawing up a traffic plan by means of potentials. It is possible to draw up a plan which will guarantee the shortest overall running distances by constructing a number of consecutive alternative plans in the form of 'chess tables' which are tested and corrected by exactly the same calculations of potentials as we described above.[1]

[1] Very close to this method of drawing up a transportation plan and solving this problem are the various modifications of the *simplex method* put forward in foreign literature.

In this article, however, we put forward another means of solving these problems by a tabular method which derives immediately from the process using *resolving multipliers* evolved by Kantorovich in 1939. With this method of establishing the 'best' system of transportation there is no need to multiply the quantities which stand in the place of the *resolving multipliers* (as we shall see below, they are simply added to the distances between dispatch and arrival points), and for this reason we shall in future call these quantities *resolving addends*.[1]

The essence of the resolving addends method may be expressed as follows. Let us renumber all the dispatch and arrival points, and let us denote the numbers of the dispatch stations by the letter $i$, and the numbers of the arrival stations by the letter $k$. If $n$ is the number of dispatch points and $m$ the number of arrival points, then $i$ will represent the values $1, 2, \ldots, n$, and $k$ the values $1, 2, \ldots, m$. The distances between dispatch and arrival points may be denoted as $l_{ik}$, i.e. $l_{11}$ will be the distance from dispatch station to arrival station 1, $l_{12}$ the distance from the same dispatch station to arrival station 2, etc.

Is it then possible to arrive at the least total ton-kilometre work by connecting each unloading point ($k$) to the dispatch station ($i$) to which the distance ($l_{ik}$) is shortest? This simple solution of the problem will not work. The consumption requirements of a number of stations $k$ would remain unsatisfied, and no use would be made of loads at a number of points $i$ because as a rule the volume of loading and unloading at the dispatch and arrival stations nearest to each other would not match.

The idea of the method proposed here is to select values for $\lambda_1, \lambda_2, \ldots, \lambda_n$ in such a way that if we connect up our points according to the shortest distance principle (each destination to the nearest dispatch point), but instead of actual distances use some assumed distances (arrived at by adding to the actual distances $l_{ik}$ the corresponding numbers $\lambda_i$, i.e. taking $l_{11}+\lambda_1$ instead of $l_{11}$, $l_{12}+\lambda_1$ instead of $l_{12}$, $l_{21}+\lambda_2$ instead of $l_{21}$, etc.), then the requirements of each destination will be fully met by the dispatch points to which it is connected.

[1] If we apply the connotations adopted by Kantorovich (see pp. 228-9) of the present work) we can express the conditions of our problem as follows:

$$(1)\ h_{ik}>0 \qquad (2)\ \sum_{k=1}^{m} h_{ik} = 1 \quad (i = 1, 2, \ldots, n)$$

$$(3)\ \sum_{i=1}^{n} \alpha_i h_{ik} = z_k \qquad (4)\ \sum_{1}^{n} \alpha_i = \sum_{1}^{m} z_k.$$

It is necessary with $\alpha_i$ and $z_k$ given to find $h_{ik}$ which secures the minimum value of:

$$\sum_{i=1}^{n} \sum_{k=1}^{m} c_{ik}\alpha_i h_{ik}$$

($c_{ik}$ are given). Because the values of $\alpha_i$ are independent of $k$ there is no need to carry out multiplication to arrive at a solution of this problem.

If values could be found for $\lambda_1, \lambda_2, \ldots, \lambda_n$ which would satisfy this condition (let us call them *resolving addends*), the problem of establishing the shortest overall running distance would be solved. Connexions made over the shortest assumed distances $(l_{ik}+\lambda_i)$ would be the most efficient, i.e., they would ensure minimal ton-kilometre work.

Indeed, if we denote the number of loads to be sent from $i$ to $k$ by $x_{ik}$ $(k = 1, 2, \ldots, m)$, the overall running distance, if we reckon with actual distances, may be expressed as $\sum_1^n \sum_1^m l_{ik}x_{ik}$, and if we reckon with assumed distances, as

$$\sum_1^n \sum_1^m (l_{ik}+\lambda_i)\, x_{ik}$$

It is easy to see that if we connect the arrival and dispatch points over the shortest assumed distances (in this case we take the values of $x_{ik}$ which do not correspond to the shortest assumed distances as equal to zero) the following inequality occurs:

$$\sum_{i=1}^{n} \sum_{k=1}^{m} (l_{ik}+\lambda_i)\, x_{ik} \leqq \sum_{i=1}^{n} \sum_{k=1}^{m} (l_{ik}+\lambda_i)\, x'_{ik} \qquad \ldots \; [1]$$

where $x'_{ik}$ is the quantity of loads sent from $i$ to $k$ on any other version of the system of connexions. Since, whatever the version of the plan used, all the loads at every dispatch point must be distributed, it follows that

$$\sum_{k=1}^{m} x_{ik} = \sum_{k=1}^{m} x'_{ik} = \alpha_i \quad (i = 1, 2 \ldots n) \qquad \ldots \; [2]$$

where $\alpha$ is the quantity of loads at the dispatch points. It follows further that the sums

$$\sum_{i=1}^{n} \sum_{k=1}^{m} \lambda_i x_{ik} = \sum_{i=1} \lambda_i \sum_{k=1}^{m} x_{ik}$$

and $\ldots \; [2]$

$$\sum_{i=1}^{n} \sum_{k=1}^{m} \lambda_i x'_{ik} = \sum_{i=1}^{n} \lambda_i \sum_{k=1}^{m} x'_{ik}$$

are equal, since each is equal to $\sum_1^n \lambda_i \alpha_i$. If we subtract these sums from the left and right sides of the inequality (1), we obtain:

$$\sum_{i=1}^{n} \sum_{k=1}^{m} l_{ik}x_{ik} \leq \sum_{i=1}^{n} \sum_{k=1}^{m} l_{ik}x'_{ik} \qquad \ldots \; [3]$$

In other words, when arrival points are connected to dispatch points over the shortest assumed distances ($x_{ik}$), the overall running distance cannot be greater than it would be in any other alternative plan ($x'_{ik}$).[1]

We have thus been able to reduce the problem of establishing the shortest overall running distance for loads when drawing up a transportation plan linking dispatch points and destinations to the problem of determining the values of the resolving addends ($\lambda_1, \lambda_2, \ldots, \lambda_n$)[2] and of drawing up a plan based on the shortest assumed distances. The resolving addends, and the system of connexions giving the shortest running distances, may be found by means of successive approximations.

### 7. Calculating Procedure for the Method of Resolving Addends

Let us take an example. The conditions are set out in Table 2.[3]

Table 2

$d_1 = 5$

| Dispatch points | Arrival points: *a* 4 | *b* 30 | *c* 16 | *d* 27 | *e* 18 | *f* 5 | 'Surplus' or 'deficit' of loads (+, −) |
|---|---|---|---|---|---|---|---|
| *A* 15 | **10** (4) | 75 | 80 | 50 | 52 | 40 | +11 |
| *B* 20 | 100 | **22** (20) | **16** | 48 | 70 | **29** | −31 |
| *C* 25 | 25 | 27 | 31 | 90 | 66 | 45 | +25 |
| *D* 40 | 44 | 26 | 38 | **37** (27) | **33** (13) | 50 | − 5 |
| Differences in distances | | 5 | 15 | 13 | 19 | 11 | |

In this table letters denote dispatch points (*A*, *B*, *C* and *D*) and arrival points (*a*, *b*, *c*, *d*, *e* and *f*). The figures beside these letters denote the quantity of loads dispatched or received, and the figures in the respective columns

[1] The existence of resolving addends is not only a sufficient but also a necessary condition for ensuring that the transportation plan meets the requirements for producing minimal overall running distances for loads. This proposition can be considered as a consequence of the potentials theorem. It is, in fact, easy to see that the resolving addends and shortest assumed distances are the respective potentials of the dispatch and arrival points of loads.

[2] Dispatch and arrival points have an exactly analogous role in setting up networks. Where there are fewer arrival points than dispatch points it is more convenient to find resolving addends for the arrival points instead of the dispatch points: $\lambda', \lambda'_2, \ldots, \lambda'_m$. In this case the assumed distances will be expressed by the values $l_{ik} + \lambda'_i$ and not $l_{ik} + \lambda_i$ as in the text. Similar changes will be made in all subsequent considerations.

[3] The example shown and the form of table used were worked out in 1948 by A. M. Dubinsky, at that time a student at the Moscow Institute of Railway Transport Engineers, and put forward by him at a students' scientific conference.

indicate the distances between the dispatch and arrival points (e.g. from $A$ to $a$ is 10 kilometres, from $A$ to $b$ 75, etc.).

Let us first examine what would be the result of connecting each arrival point to the nearest dispatch point. In that case we should have to connect $a$ to $A$; $b$, $c$ and $f$ to $B$; and $d$ and $e$ to $D$, and the distances (the smallest number in each column) would be as shown in Table 2 in heavy type. With this network $A$ would meet in full the requirements of $a$, but 11 units of load out of the 15 at $A$'s disposal would not be utilised.

A similar 'surplus' would occur at $C$, which is not the 'nearest dispatch point' for any arrival point at all. None of $C$'s 25 units would be utilised. We shall call such stations *surplus points.* At the same time the 20 units available at $B$ would prove insufficient to meet all the requirements of the points connected to it ($b$, $c$ and $f$), since $30+16+5 = 51$; 31 units would be 'lacking'. Another 'deficit' would be revealed at $D$, which with only 40 units to dispose of is connected to $d$ and $e$ which together require 45 units. Stations in this position we shall call *deficit points.*

The 'surpluses' (+) and 'deficits' (−) obtained are set out in the last column of Table 2. The totals of the positive and negative figures must balance.

To obtain a clearer picture of the transportation possibilities between the various stations with the present scheme, we have set out in Table 2 figures in brackets to show the number of loads which may be received by each destination from its respective source. To some extent these figures are arbitrary. For instance, if we reduced the supply from $B$ to $b$ shown here, we could route some of the loads at $B$ to other points connected to this station, such as $c$ or $f$. Such changes would make no difference to our argument, however, since the amounts of 'surplus' or 'deficit' would remain the same.

The next part of our procedure is to calculate the differences between ($a$) the distances between deficit dispatch stations and the arrival points connected to them, and ($b$) the distances between these same arrival points and the surplus stations nearest them. To the deficit station $B$ are connected points $b$, $c$ and $f$. We can find the distances between these arrival points and their nearest surplus stations in the relevant vertical columns: 27, 31 and 40. The differences in distance which concern us are $27-22 = 5$, $31-16 = 15$ and $40-29 = 11$. Similarly for $d$ and $e$, which are connected to another deficit station, $D$, we obtain $50-37 = 13$ and $52-33 = 19$. These differences are set out in the last row of Table 2.

Let us represent the smallest of these differences by $d_1$ and enter it in the top right-hand corner of Table 2. This difference we may take as a first approximation to our resolving addend for $B$ and $D$ (which are 'deficit stations'), that is to say, we shall use this number to obtain a first variant of our 'assumed distances'. To do this we increase all the distances between $B$ and $D$ and the destination points (all the figures in the rows $B$ and $D$) by $d_1$ ( $= 5$ in our example). The results obtained are shown in Table 3.

TABLE 3

$d_2 = 8$

| Dispatch points | $a$ 4 | $b$ 30 | $c$ 16 | $d$ 27 | $e$ 18 | $f$ 5 | 'Surplus' or 'deficit' of loads (+, −) |
|---|---|---|---|---|---|---|---|
| | Arrival points | | | | | | |
| $A$ 15 | **10** (4) | 75 | 80 | 50 | 52 | 40 | +11 |
| $B$ 20 | 105 | **27** (5) | **21** (15) | 53 | 75 | **34** | − 6 |
| $C$ 25 | 25 | 27 (25) | 31 | 90 | 66 | 45 | − 0 |
| $D$ 40 | 49 | 31 | 48 | **42** (27) | **38** (13) | 55 | − 5 |
| Differences in distances | | 48 | 59 | 8 | 14 | 6 | |

Making use this time of the distances shown in Table 3, we once again connect each destination point to its nearest dispatch point (as in Table 2 the relevant figures are printed in heavy type) and assess the quantity of supplies which the dispatch stations can now send to each destination (the figures in brackets), always bearing in mind the need to make the best use of the loads available at the dispatch stations and to satisfy as far as possible the requirements of the arrival points. The changes in the distances make it possible to reduce the amounts 'surplus' and 'deficit'. Now only 6 units are 'lacking' from the destinations connected to $B$ instead of 35, and there is no longer any 'surplus' of loads at $C$.

The question arises how we should regard $C$ henceforth: is it a 'surplus' or a 'deficit' point? The goods available at this station are now utilised in full, as in the case of deficit stations; at the same time the requirements of point $b$ connected to it are also fully satisfied, as in the case of destinations connected to surplus stations. To cover cases like these we shall adhere to the following rule. If the requirements of the destination points connected to a given dispatch point are satisfied in full, but if even one of these points is being simultaneously supplied from another station which is a deficit point, then the dispatch station we are considering must also be regarded as a deficit point. In the converse case a station which is able to satisfy all the requirements of the destinations connected to it must still be regarded as a surplus point, even if in fact the 'surplus' at its disposal is equal to zero.

From this it follows that $C$ is a deficit point, since $B$ participates in the supply to $b$. We therefore prefix the '0' in the last column of the Table, row $C$, with the minus sign.

We now proceed as we did with Table 2, finding the differences between ($a$) the distances between deficit dispatch stations and the arrival points connected to them, and ($b$) the distances between these arrival points and their nearest surplus stations. By referring to the columns under the arrival points which are connected to deficit stations, we obtain the figures shown in the bottom row of Table 3. The smallest difference is 6.

However, if we increase all the distances between arrival points and deficit stations by 6 (in the same way as we increased them by 5 when we made the transition from Table 2 to Table 3) it soon becomes evident that in the new network there will be no change in the classification of dispatch points as surplus and deficit stations. The only change, in fact, will be that alongside *B* the surplus station *A* will also have to be connected to *f* because the distance between *f* and *A* and *f* and *B* will become the same. This will ensure that *f*'s requirements are satisfied, but *A* will still remain a surplus station and *B* and *C* deficit stations. To make the transition to the next variant of the network in such cases as these it is best to use the second smallest difference (i.e. the second smallest figure in the last row of the relevant table) instead of the smallest or, if this fails to alter the balance of surplus and deficit stations, the third smallest, and so on.

In Table 3 the smallest figure after 6 is 8 (the difference between the distances *d* to *D* and *d* to *A*), and so we enter $d_2 = 8$ in the top right-hand corner of the Table. By adding 8 to the figures in the rows of the deficit stations we obtain our new assumed distances, and these are set out in Table 4. Just as with Tables 2 and 3 we now set up a new network, calculate the load 'surpluses' and 'deficits', and enter the differences in the distances which concern us in the bottom row of the table.

TABLE 4

$d_3 = 4$

| Dispatch points | Arrival points *a* 4 | *b* 30 | *c* 16 | *d* 27 | *e* 18 | *f* 5 | 'Surplus' or 'deficit' of loads (+, −) |
|---|---|---|---|---|---|---|---|
| *A* 15 | **10** (4) | 45 | 80 | **50** (6) | 52 | **40** (5) | +0 |
| *B* 20 | 113 | **35** (5) | **29** (15) | 61 | 83 | 42 | −1 |
| *C* 25 | 33 | **35** (25) | 39 | 98 | 74 | 53 | −0 |
| *D* 40 | 57 | 39 | 51 | **50** (21) | **46** (18) | 63 | +1 |
| Differences in distances | | 4 | 22 | | | | |

We note that $d_3 = 4$, and then set up the next variant of the network exactly the same way as when making the transitions from Table 2 to Table 3 and Table 3 to Table 4. As is shown by Table 5, the new assumed distances are now such as to permit the full supply of every arrival point from its 'nearest' dispatch points. Table 5, therefore, offers the final solution to this particular problem.[1] *A* supplies *a*, *d* and *f* (4, 6 and 5 units), *B* supplies *b* and *c* (4 and 16), *C* sends all its loads to *b*, and *D* has a share in the supply of *b* (1), *d* (21) and *e* (18).

TABLE 5

| Dispatch points | Arrival points: *a* 4 | *b* 30 | *c* 16 | *d* 27 | *e* 18 | *f* 5 | 'Surplus' or 'deficit' of loads (+, −) |
|---|---|---|---|---|---|---|---|
| *A* 15 | **10** (4) | 75 | 80 | **50** (6) | 52 | **40** (5) | 0 |
| *B* 20 | 117 | **39** (4) | **33** (16) | 65 | 87 | 46 | 0 |
| *C* 25 | 37 | **39** (25) | 43 | 102 | 78 | 57 | 0 |
| *D* 40 | 57 | **39** (1) | 51 | **50** (21) | **46** (18) | 63 | 0 |

It will be readily observed that the distances in each row of Table 5 differ from the corresponding figures in the same row of Table 2 by the same amount. These amounts are: row *A*, 0 (the figures aie the same in both tables); row *B*, 17 $(d_1+d_2+d_3)$; row *C*, 12 $(d_1+d_3)$; and row *D*, 13 $(d_1+d_2)$. In each case they are also the same as the resolving addends $(\lambda_A, \lambda_B, \lambda_C, \lambda_D)$ which were discussed in Section 6. As we can see, for the purposes of setting up a transportation system there is no point in working out their ultimate values because our network has already emerged at an earlier stage of our calculations. The actual process of establishing the system requires only the auxiliary numbers $d_1, d_2, d_3, \ldots, d_n$. The method just outlined for establishing networks, however, is based on the fact that there exist systems of numbers which will satisfy the definition of resolving addends and that transportation schemes based on these systems will minimise overall running distances.

It is advisable to calculate resolving addends for the purpose of checking

[1] In this example three 'steps' are needed in the transition from the original to the final version of the plan. It is possible to show that the number of 'steps' needed to solve the problem is finite. This may be proved as follows. Let the squares of any table of assumed distances from arrival points to their nearest dispatch points correspond exactly to the same squares of a second table. These tables we shall call 'equivalent' (the same versions of the system of connexions correspond to them). It is easy to see that the gradual transition according to the rules stated above from one table of assumed distances to another cannot produce 'equivalent' tables. It is similarly easy to show that the number of 'non-equivalent' tables distinguished from each other by the assumed distances set out in the corresponding squares is finite. Hence it follows that a finite number of 'steps' must lead to a solution of the problem. It must be borne in mind that when ordinary practical problems are being tackled instead of specially selected examples, the number of 'steps' will as a rule remain below $n+m$ and, in fact, only in exceptional cases exceed half that sum $\left(\frac{n+m}{2}\right)$.

TABLE 6*

| Arrival points / Dispatch points | $a$ 4 | | | | $b$ 30 | | | | $c$ 16 | | | | $d$ 27 | | | | $e$ 18 | | | | $f$ 5 | | | | $d_1=5$ | $d_2=8$ | $d_3=4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | | | | | | | | | | | 'Surplus' or 'deficit' of loads (+, −) | | |
| $A$ 15 | **10** (4) | **10** (4) | **10** (4) | **10** (4) | 75 | 75 | 75 | 75 | 80 | 80 | 80 | 80 | 50 | 50 | **50** (6) | **50** (6) | 52 | 52 | 52 | 52 | 40 | 40 | **40** (5) | **40** (5) | +11 | +11 | +0 0 |
| $B$ 20 | 105 | 105 | 113 | 117 | **22** (20) | **27** (5) | **35** (5) | **39** (4) | **16** | **21** (15) | **29** (15) | **33** (16) | 48 | 53 | 61 | 65 | 70 | 75 | 33 | 87 | **29** | **34** | 42 | 46 | −31 | −6 | −1 0 |
| $C$ 25 | 25 | 25 | 33 | 37 | 27 | **27** (25) | **35** (25) | **39** (25) | 31 | 31 | 39 | 43 | 90 | 90 | 98 | 102 | 66 | 66 | 74 | 78 | 45 | 45 | 53 | 57 | +25 | −0 | −0 0 |
| $D$ 40 | 44 | 49 | 57 | 57 | 26 | 31 | 39 | **39** (1) | 38 | 43 | 51 | 51 | **37** (27) | **42** (27) | **50** (21) | **50** (21) | **33** (13) | **38** (13) | **46** (18) | **46** (18) | 50 | 55 | 63 | 63 | −5 | −5 | +1 0 |
| | | | | | 5 | | | | 15 | | | | 13 | | | | 19 | | | | 11 | | | | | | |
| Differences in distances | | | | | | 18 | | | | 59 | | | | 8 | | | | 11 | | | | 6 | | | | | |
| | | | | | | | 4 | | | | 22 | | | | | | | | | | | | | | | | |

* At each stage in the calculations only the last figures set out on the right are to be taken into consideration. All the rest may be deleted.

transportation systems which have been established by the method just described. All that need be done is to check (as we did above) that the figures in each row in the final table have indeed increased by the same amount (the resolving addend) when compared with the figures in the corresponding rows in the original table, and that the system does satisfy the principle of the shortest assumed distances.

For each successive approximation to our ideal network we made use of a separate table. We did this for clarity of exposition: in practice, so long as the number of stations is not too great, all the calculations may be set out in one large working table. Table 6 shows how such a table would look for the example we selected.

If we take as starting data the distances between stations and the quantities of goods dispatched from and unloaded at separate points in accordance with the network shown in Fig. 6, the final appearance of the working table after the application of the method of resolving addends will be as shown by Table 7. The network derived from this table is, as may have been anticipated, identical with the one produced by the use of the circuit differences method (see Fig. 8).

## 8. Comparison of the Features of Different Methods of Setting up Transportation Systems

The possibility of successfully solving problems connected with transportation systems by purely analytical methods by no means invalidates the graphic or graphico-analytical methods. The graphic method (see Section 1) remains, of course, the simplest and most convenient in all cases where closed circuits are not involved. The first step in making connexions between dispatch and arrival points when railway or inland waterway traffic is being planned should therefore always be to plot the load data on a diagram of the layout of the track (or waterways) and to apply the graphic method to every section possible.

It is only when the load-flows which are discernible as efficient by a direct examination of the diagram have been drawn in and relevant junctions substituted for groups of stations (see Section 4) that the question of making use of the graphico-analytical or resolving addends methods arises. (Other methods have not been examined in this paper.) To assess the relative merits and defects of the two latter methods is a more complicated problem, and the selection of one rather than the other must depend on the actual conditions of the problem involved.

The comparative laboriousness of the two methods varies in accordance with the number of stations to be linked up and the number of closed circuits formed by the routes linking the dispatch and arrival points. The resolving addends method grows rapidly more laborious with an increase in the number of dispatch and arrival points, but is not directly affected by the configurations of the transport network. On the other hand the amount of work

TABLE 7

| Arrival points / Dispatch points | *a* 17 | | | | *b* 8 | | | | *c* 20 | | | | *d* 5 | | | | *e* 10 | | | | *f* 20 | | | | *g* 5 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *A* 25 | **25** (17) | **45** (17) | 50 (10) | 50 (10) | 40 | 60 | 65 | 65 | 60 | 80 | 85 | 85 | 95 | 115 | 120 | 120 | 50 | 70 | 75 | 75 | **15** (8) | **35** (8) | **40** (10) | **40** (10) | 5 | 25 | 30 (5) | 30 (5) |
| *B* 10 | 80 | 100 | 100 | 100 | 75 | 95 | 95 | 95 | **55** | **75** | **75** | **75** | **15** (5) | **35** (5) | **35** (5) | **35** (5) | 50 | 70 | 70 | 70 | 85 | 105 | 105 | 105 | 85 | 105 | 105 | 105 |
| *C* 20 | 90 | 90 | 90 | 90 | 95 | 95 | 95 | 95 | 90 | 90 | 90 | 90 | 50 | 50 | 50 | 50 | **15** (10) | **15** (10) | **15** (10) | **15** (10) | 50 | 50 | 50 | 50 | 60 | 60 | 60 | 60 |
| *D* 10 | 55 | 75 | 80 | 80 | 70 | 90 | 95 | 95 | 90 | 110 | 115 | 115 | 85 | 105 | 110 | 110 | 20 | 40 | 45 | 45 | **15** (10) | **35** (10) | **40** (10) | **40** (10) | 25 | 45 | 50 | 50 |
| *E* 10 | 35 | 55 | 60 | 65 | 50 | 70 | 75 | 80 | 70 | 90 | 95 | 100 | 85 | 105 | 110 | 115 | 60 | 80 | 85 | 90 | 25 | 45 | 50 | 55 | 15 | 35 | 40 | 45 |
| *F* 25 | 30 | 50 | **50** | **50** | **35** (8) | **55** (8) | **55** (8) | **55** (8) | **55** (7) | **75** (17) | **75** (17) | **75** (17) | 65 | 85 | 85 | 85 | 55 | 75 | 75 | 75 | 55 | 75 | 75 | 75 | 45 | 65 | 65 | 65 |
| *G* 25 | 50 | 50 | **50** (7) | **50** (7) | 55 | **55** | **55** | **55** | 75 | **75** (3) | **75** (3) | **75** (3) | 75 | 75 | 75 | 75 | 35 | 35 | 35 | 35 | 65 | 65 | 65 | 65 | 55 | 55 | 55 | 55 |
| | 25 | | | | 20 | | | | 20 | | | | 35 | | | | | | | | 35 | | | | 50 | | | |
| Difference in distances | | 5 | | | | | | | | | | | | | | | | | | | | 15 | | | | 30 | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TABLE 7 *continued*

| Arrival points / Dispatch points | *h* 15 | | | | *i* 5 | | | | *j* 10 | | | | *k* 10 | | | | $d_1$ 20* | $d_2$ 5 | $d_3$ 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *A* 25 | 15 | 35 | 40 | 40 | 65 | 85 | 90 | 90 | 35 | 55 | 60 | 60 | 65 | 85 | 90 | 90 | −7 | −7 | +0 | 0 |
| *B* 10 | 65 | 85 | 85 | 85 | **15** (5) | **35** (5) | **35** (5) | **35** (5) | 45 | 65 | 65 | 65 | 55 | 75 | 75 | 75 | −13 | +0 | +0 | 0 |
| *C* 20 | 75 | 75 | 75 | 75 | 50 | 50 | 50 | 50 | 55 | 55 | 55 | 55 | **20** (10) | **20** (10) | **20** (10) | **20** (10) | +0 | +0 | +0 | 0 |
| *D* 10 | 45 | 65 | 70 | 70 | 85 | 105 | 110 | 110 | 65 | 85 | 90 | 90 | 35 | 55 | 60 | 50 | −0† | −0† | +0 | 0 |
| *E* 10 | **5** (10) | **25** (10) | **30** (10) | **35** (10) | 55 | 75 | 80 | 85 | 25 | 45 | 50 | 55 | 60 | 80 | 85 | 90 | −5 | −5 | −5 | 0 |
| *F* 25 | 25 | 45 | 45 | 45 | 35 | 55 | 55 | 55 | **5** (10) | 25 | 25 | 25 | 40 | 60 | 60 | 60 | −0† | +0 | +0 | 0 |
| *G* 25 | 35 | 35 | 35 | **35** (5) | 45 | 45 | 45 | 45 | 15 | **15** (10) | **15** (10) | **15** (10) | **20** | **20** | **20** | **20** | +25 | +12 | +5 | 0 |
| Differences in distances | 30 | | | | 30 | | | | 10 | | | | | | | | | | | |
| | | 10 | | | | | | | | | | | | | | | | | | |
| | | | 5 | | | | | | | | | | | | | | | | | |

* If $d_1$ is taken as 10 the division of the dispatch points into surplus and deficit stations in the second variant of the plan will remain unchanged in comparison with the original version.

† Arrival point *f* (connected to *D*) is two units short. But since these two units were reckoned in *A*'s deficit (and *f* is also connected to *A*) *D*'s deficit is taken as nil to avoid a double reckoning. Alternatively *A*'s deficit could be reckoned as 5 instead of 7 (attributing only the under-supplying of *G* to *A*) and the figures for *D* entered on the graph as −2 instead of −0. The only essential is to avoid double counting. Similarly the shortage at *c* (13 units) could be ascribed to a deficit either at *B* (as in Table 9) or at *F*, the two stations responsible for supplying *c*; or, finally, it could be split up between both (e.g. 7 units at *B* and 6 at *F*).

involved with the graphico-analytical method depends mainly on the number of 'circles' which must be checked, and only to a small extent on the number of stations involved.

Thus, for example, constructing a network for the example dealt with in Section 3 (see Fig. 4) with the aid of the circuit differences method involves almost no work at all; whereas the use of this method to solve the problem posed in Section 4 (see Figs. 5, 6, 7 and 8) would necessitate a vast expenditure of labour, since here we would be dealing with thirteen closed circuits instead of only one. However, the expenditure of labour would differ but slightly, if either of these problems were solved by the resolving addends method (5 dispatch and 5 arrival points as against 7 and 11). In many cases it may be best to combine different methods, including the resolving addends and circuit differences methods. It is possible, for example, where there is a large number of dispatch and arrival points, to consider the loading and unloading at minor stations as functions of neighbouring major stations, and having thus reduced the number of stations to be examined, to apply the resolving addends method, and then to plot the results obtained on the diagram and to check doubtful points with the aid of the graphico-analytical method.

Where it is desired to establish minimum transport costs instead of running distances, expressed either in money or natural units (roubles, truck-hours, tons of fuel, etc.) it is important not to overlook one limitation of the graphico-analytical method. It can be used only if the transport costs from $A$ to $C$ via $B$ are equal to the sum of the transport costs from $A$ to $B$ and from $B$ to $C$.

This proviso does not prevent the setting up of systems to minimise expenditure envisaged on the basis of actual or planned transportation costs on separate sections. This may be done simply by substituting costs for distances in the diagram and using these costs when calculating circuit differences or potentials. This method of calculation need hardly be changed even where the costs of running laden trucks and empty ones are to be kept apart. In this case two figures ('forward' and 'return') must be entered for each section, and slight modifications will have to be made to the rule formulated above.

Situations may arise, however, where the proviso as to the balance of costs on adjacent sections cannot be observed. This is the case, for example, where a plan is required to minimise transport costs arising from the system of freight rates. The combined costs from $A$ to $B$ and from $B$ to $C$ will not, as a rule, be equal to that from $A$ to $C$ (on account of freight reductions for distance and the existence of special rates). In cases like these the graphico-analytical method is not applicable (nor is the purely graphical method), whereas calculations in accordance with the resolving addends method may still be made exactly as outlined above. The resolving addends method can thus be applied more widely than the graphico-analytical method.

Definite conclusions as to the merits and demerits (especially the laboriousness) of the various methods of drawing up transportation plans to minimise running distances and costs must await the accumulation of practical experience. It should be observed, however, that the application of any of the methods described in this paper on a massive scale to an entire transport network with a great number of widely scattered dispatch and arrival stations will involve a formidable amount of labour and time, and for this reason it is advisable to make use of the latest computer techniques. If fast modern computers are used, any transportation system can be set up without any difficulty in a minimum of time.

The methods described here have already given practical proof of their efficacy. For example, the Institute of Complex Transportation Problems of the Academy of Sciences used the resolving addends method to draw up a plan for the most efficient road haulage of sand in Moscow. The orders for sand issued by the construction enterprise Mosstroisbyt during a ten-day period in June 1958 were used as starting data. To establish the most efficient plan distances had to be calculated from each of the eight wharves where sand was picked up to each of 209 building sites. All calculations were carried out with the help of the *Strela* electronic digital computer in one hour thirty-five minutes (including fifty minutes spent on preparation of input data).

Comparison of the optimum plan obtained with the lorry journeys which would have been necessary under the old system of ordering showed a reduction of 189,000 ton-kilometres, i.e. a saving of 11·4%. Approximate calculations show that the introduction of the optimum plan in Moscow would mean a saving of more than two million roubles a year, solely for the transportation of sand.