International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

An Interaction Model

Puu, T.

IIASA Collaborative Paper
August 1982

AN INTERACTION MODEL

Tönu Puu*

August 1982
CP-82-50

*Department of Economics
University of Umeå
Box 718
S-90187 Umeå - Sweden

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

Methodologically, the IIASA research program on Regional Development reflects the general attitude of the majority of regional scientists. Among other things, this means that the models developed deal with discrete sets of regions or locations. For planning purposes, this approach is extremely efficient, due to computational advantages. On the other hand, systematic information about regional structures, of the geometric flavor associated with classical location theory, is hard to obtain if one discretizes space from the outset.

To complement this main stream of regional analysis, two scientists currently trying to revive continuous modeling of the space economy, Martin J. Beckmann and Tönu Puu, were invited to IIASA in September 1979. They started writing a comprehensive monograph intended to present the state-of-the-art in the field of continuous regional modeling. The completion of such an extensive work was not possible in the brief period of three weeks.

The authors are currently continuing work on the project. This paper by Tönu Puu constitutes one chapter of the forthcoming monograph. It was completed during his visit to IIASA in August 1982 and follows the chapter circulated as CP-82-11.

Whereas the preceding chapters dealt with commodity trade models with unique patterns of flow, the present one describes a simple interaction model cast in a continuous format. Given a specified need for communication and accommodation, optimal land use (balancing traffic congestion and population crowding) is discussed. In addition to the problems of optimal communication routes, the paper focuses on equilibrium population distributions such that communication and housing costs are in balance.

Laxenburg, August 1982

Boris Issaev
Leader
Regional Development Group

TABLE OF CONTENTS

AN INTERACTION MODEL

Introduction

All the previous chapters have dealt with trade models of various kinds. For each commodity there was a unique flow, representable by a well-behaved vector field. In some cases there were several commodity flows, but their number was always finite, and the flow for each commodity unique.

Uniqueness results from rationality of behaviour in our simplified world of single transportation systems, where backhauling is ruled out. Whenever this is the case it seems superfluous to record the information of the origin and destination of each single piece of commodity. It is immaterial whether such a piece, delivered to consumers at a certain location, has followed a flow line all the way from producers at a distant location, or has been entered at an intermediate location to substitute an identical piece in the orginal flow.

But, what if the pieces "produced" at different locations are all unique? As long as we deal with commodity trade this would seem to be an unnecessary complication, as any real commodity has a sufficient degree of homogeneity to justify the mild abstraction from individual variations. At least we should be able to break the set of commodities down in a more refined, but still finite, set of brands for which the abstraction is justified.

If we, however, deal with general purpose communication (or interaction) between individuals at different locations, rather than with commodity trade, then the "pieces" produced and consumed are all different as soon as either the origins or the destinations differ.

All locations need to communicate with all the other locations, and no such communication could be replaced by an equivalent communication, obtained by changing the origin or the destination. There exist no equivalents!

By this we are in the world of interaction models. If we wish to maintain our continuous paradigm, we have to deal with a non-denumerable infinity of vector fields, each one corresponding to a fixed origin or a fixed destination. These vector fields do not fuse to one resultant field, as they will do in any physical application. They co-exist separately, so that through each location there pass an infinity of flows in different directions, all of them having different origins. The aggregate of the norms of all there flows is a measure of traffic through that location.

Traffic, defined in this way, will be one of the important variables in the following model of regional structure. Ultimately, traffic depends on the demand for communication between various locations, and on the choices of optimal routes for all these communications.

The demand for communications is assumed to depend multiplicatively on population densities at the locations of origin and destination. This is a very simple variant of interaction theory, where the cost-distance dependence is altogether deleted, along with the so called balancing factors.

The reason for the first deletion is that otherwise minimization of communication costs could be attained by the absurd method of making communication so difficult that people abstain from it altogehter. Of course, we could evaluate communication and balance the value of communication against the cost of it. But, it is easier to just stipulate a pattern and volume of communication as a constraint and minimize the costs of realizing it. As we have a limited urban area in mind, this is not too unreasonable.

The skipping of the balancing factors is reasonable as we deal with general-purpose communication. As absurd as it is when dealing with commuting that a doubling of workers and jobs entails a quadrupling of trips, as sensible is it that total communication quadruples in a doubled population. This is so because a growing population also entails an increase in the diversity of activities.

So far we have discussed the demand for communication. In order to determine the (infinity of) flow fields in our model, we also have to consider the choice of routes. This problem will be dealt with as elsewhere in the book, by choosing routes so that the path integrals of local transit costs are minimized along them. The transit cost is again a location-dependent, but direction-independent (isotropic), scalar field. However, we presently do not take it as a given datum, but assume it to depend on congestion measured as the ratio of traffic to road capacity at the location.

These two pieces of theory, the simple interaction model for communication demand, and the optimal routing paradigm, make it possible to derive all the flow fields and hence the traffic distribution. We should note that a complicated feed-back mechanism is involved, as traffic depends on optimal routing, which depends on traffic! So, the resulting traffic distribution is an equilibrium one and may be hard to actually compute.

The given data, resulting in an equilibrium traffic distribution, are: the distributions of population and of road capacity.

Before considering those, and the rest of our model, let us just note that we are in the position of computing the communication costs for each location along with the flow field having its origin there. We know the numbers of communications terminating at each one of the other locations, we know the best routes of communication, and we know the local transit costs along them. We can imagine that a location separated from the main part of the population by some highly congested area will suffer from high communication costs.

Suppose people are free to move from one location to another. What then can make them accept such high communication costs? Low costs of housing is an obvious answer. So, let us stipulate a spatial invariance of the sum of communication and housing costs as a condition for equilibrium in the spatial population distribution.

Like we assumed that each individual needed a certain number of communications with each of the other individuals we also assume that each individual needs a certain living space. If few people live in an area we can house them in one-storey buildings of weak construction, but with a growing population density we have to build higher and higher, at an increasing capital cost per unit of artificial housing space created. The assumption is that it is more expensive to provide an individual with his required living space the higher the ratio of population to the space available for housing.

We are now in the position of closing the model. The natural space available for housing obviously is the part of it not used for the transportation network. For each patch of land we have the option of using it to facilitate housing or communication. If we use it for the first purpose, the result will be decreased crowding of population and residential construction costs, and increased traffic congestion and local transit cost. A balance obviously has to be struck for land-use at each location so that the sum of housing and communication costs is as small as possible.

On the other hand, we also stipulated that this sum should be a spatial invariant. This was the condition for a population distribution in spatial equilibrium. This equilibrium condition removes our last degree of freedom.

Optimal flows

Let us now formalize the model. As usual denote the region studied A and its boundary ∂A. In the present model we deal with pairs of origin and destination locations. Let them be denoted $\xi = (\xi_1,\xi_2)$ and $x = (x_1,x_2)$ respectively. Next, define the population density function

$$p = p(x_1,x_2) \qquad (1)$$

For convenience we abbreviate population $\bar{p} = p(\xi_1,\xi_2)$ at the origin, whereas we let $p = p(x_1,x_2)$ denote population at the destination. This convention is useful as we keep the origins fixed as long as we deal with individual flow fields. (We could have chosen the destinations as fixed instead. This would have worked equally well). Total population is

$$P = \iint_A p \; dx_1 \; dx_2 \tag{2}$$

As already indicated, an individual flow field can be defined uniquely when the origin is fixed. Denote it

$$\phi = (\phi_1(x_1,x_2), \; \phi_2(x_1,x_2)) \tag{3}$$

Of course, the vector field also depends on the location ξ of origin, but keeping it fixed we can delete it as an explicit argument in (3). It should be noted that with a fixed origin all the vector operations, like taking the divergence, are carried out with respect to the variable x-coordinates, not the fixed ξ-coordinates.

According to our assumption on communication demand the number of "communications" originating in ξ and leaving the flow ϕ in the destination x equals the product $\bar{p}p$ of population densities. This is the sink density, and so we obtain

$$\text{div } \phi = - \bar{p}p \tag{4}$$

as our relevant divergence law. To avoid confusion we state once more that the divergence, $\partial\phi_1/\partial x_1 + \partial\phi_2/\partial x_2$, is taken in the x-coordinates.

In order to write down the gradient law we must define the transit cost function. Presently, it is not a given function of location, but depends on the ratio of traffic to capacity as a measure of congestion. Denote traffic by i and capacity by m. Then transit cost is

$$k = k(i/m) \tag{5}$$

Using this, the gradient law, as always, reads

$$k \frac{\phi}{|\phi|} = \text{grad } \lambda \tag{6}$$

Two observations on λ are in order. First, like ϕ, it must depend on the location of origin ξ, and only when we keep the origin fixed can we delete its coordinates as arguments of the scalar field. When we regard ξ as variable we get a double continuum of vector fields ϕ, and likewise a double continuum of scalar fields λ. Second, λ is presently an undetermined Lagrangean multiplier function, associated with the constraint (4). Below we will give an interpretation. But observe that (6) tells nothing more than that the unit flow field is gradient to some, yet undetermined, function whose gradient norm equals local transit cost.

Let us multiply both sides of (6) by the unit vector $\phi/|\phi|$. In the left hand side the unit vectors multiply to scalar unity and so $k(\phi/|\phi|)^2 = k$. In the right hand side we get grad $\lambda \cdot \phi/|\phi| = d\lambda/d\sigma$, where σ is an arc length parameter. This is so because $\phi/|\phi|$ is the unit vector in the direction of the optimal route. Thus

$$k = d\lambda/d\sigma \tag{7}$$

and, integrating along any optimal route having its origin at ξ, we obtain

$$\lambda = \int_0^s k d\sigma \tag{8}$$

because $(d\lambda/d\sigma)d\sigma = d\lambda$ is an exact differential. In equation (8) we have chosen the arbitrary integration constant to be zero. By this, λ becomes the path integral of local transit costs along the most efficient routes of communication. Thus, the potential λ will have zero value at the origin and increases in all directions at the rate of local transit costs. Any positive λ defines a closed curve surrounding the origin ξ, and consists of points as far as possible from ξ when the total amount λ is spent on transportation.

Let us go somewhat deeper into the matter of the determination of the unit flow field $\phi/|\phi|$ from equaiton (6) above to make sure that it does not matter that λ is an unknown Lagrangean. To accomplish this task we will make a little abuse of the terminology from vector analysis and regard curls and cross products as scalar quantities. A cross product of two vectors (in three-space) actually is a vector, perpendicular to the plane spanned by those vectors, and pointing in the direction that forms a right handed set of axes with those two. The norm of the cross product is the area of the parallelogram spanned by the two original vectors.

Likewise, the curl actually is a vector along the axis of rotation in a flow, pointing in the direction which makes the rotation counterclockwise, and having a norm equal to the velocity of revolution.

As we deal with vectors in the plane, both the cross products and the curls always point in directions perpendicular to this plane. Thus, they have only one nonzero component. Our abuse will be to disregard the vectorial character of these two concepts and treat them as if they were identical with the (scalar) values of the single non-zero components. This simplification can cause no confusion. The only remnant of the vectorial character is the sign (or sense), which depends on whether the resultant vectors point outwards or inwards from the plane.

Using this abuse terminology, the formal definition of the cross product of two arbitrary vectors ϕ and ψ is $\phi \times \psi = \phi_1\psi_2 - \phi_2\psi_1$.

Likewise, for an arbitrary vector field ϕ, we define curl $\phi = \partial\phi_2/\partial x_1 - \partial\phi_1/\partial x_2$.

About the cross product we should note the trigonometric formula $\phi x \psi = |\phi||\psi| \sin \alpha$, where α is the angle between the directions of the vectors. As we similarly have $\phi \cdot \psi = |\phi||\psi| \cos \alpha$ we derive the useful relation $(\phi x \psi)/(\phi \cdot \psi) = \tan \alpha$.

After these preliminaries we are prepared to start out with equation (6) by taking the curls of both sides. Now a gradient field is always irrotational and the curl is hence identically zero. So,

$$\text{curl } (k\ \phi/|\phi|) = 0 \tag{9}$$

Expanding this expression we get

$$\text{grad } k \text{ x } \phi/|\phi| + k \text{ curl } (\phi/|\phi|) = 0 \tag{10}$$

(Note the similarity of this expression to the corresponding one for the divergence).

Next, denote the direction of grad k by ω and the direction of $\phi/|\phi|$ by θ. Using our trigonometric relation betweeen cross and dot products, and noting that grad $k \cdot \phi/|\phi| = dk/d\sigma$, we get

$$dk/d\sigma \sin(\theta-\omega) + k \text{ curl}(\phi/|\phi|) \cos(\theta-\omega) = 0 \tag{11}$$

But, $\phi/|\phi| = (\cos \theta, \sin \theta)$ and so, by definition of the curl, and using the chain rule,

$$\text{curl } (\phi/|\phi|) = \cos \theta\ \partial\theta/\partial x_1 + \sin \theta\ \partial\theta/\partial x_2 \tag{12}$$

However, as $(\cos \theta, \sin \theta) = (dx_1/d\sigma, dx_2/d\sigma)$, we immediately transform (12) into

$$\text{curl } (\phi/|\phi|) = d\theta/d\sigma \tag{13}$$

Substituting into (11) we get

$$\sin(\theta-\omega)\, dk/d\sigma + k \cos(\theta-\omega)\, d\theta/d\sigma = 0 \qquad (14)$$

Let us consider (14) a little. The angle ω is defined by the gradient direction to k, the local transit cost, which is known. The variation of this cost in the direction of the route, $dk/d\sigma$, only depends upon the direction. Accordingly, (14) involves as unknowns only the direction of the route, θ, and its rate of change, $d\theta/d\sigma$, as we follow the route. We thus have a differential equation for the route direction with arc length as argument.

This differential equation, in fact, justifies our assertion that (6) would allow us to derive the flow lines, despite the fact that λ is unknown. Its character is most easily understood by some special cases.

First, assume that ω is invariant in space, so that $dk/d\sigma = 0$ and we can drop the first term in (14). What remains can be written

$$d/d\sigma(k \sin(\theta-\omega)) = 0 \qquad (15)$$

which has the first integral

$$k \sin(\theta-\omega) = \text{constant} \qquad (16)$$

Obviously, the sine of angular difference between the directions of maximum transit cost increase and of the route is related reciprocally to transit cost. If the route takes us to locations where transit cost increases, we decrease the angular difference in order to pass the high cost region as fast as possible. If transit cost decreases along the route we increase the difference in order to profit from the low costs during as long a transit as possible.

Equation (16) again reminds us of geometrical optics. In a separation point between two media with refraction indices k_1 and k_2 (and $\omega = 0$ arbitrarily as it is not defined when k_1 and k_2 are sectionally constant) we have

$$k_1/k_2 = \sin \theta_1/\sin \theta_2 \tag{17}$$

This is Snell's law telling that the sines of incidence angles have the same ratio as the refraction indices. It is noteworthy that the corresponding refraction law for transportation, partly on land, partly on sea, with different transit costs (k_1,k_2), was discovered by two economists, Palander (1935) and v Stackelberg (1938).

Second, relax the constraint of a constant k, but assume it to display circular symmetry. So, we can write $k(\rho)$, where $\rho = \sqrt{(x_1^2 + x_2^2)}$. In view of our complete model this, of course, means that the congestion ratio, i/m, itself depends on location x via $\rho = |x|$ only. From this circular symmetry of k, we get

$$\text{grad } k = dk/d\rho(x_1/\rho, x_2/\rho) \tag{18}$$

But, as ω was defined to be the angle of the gradient of k, we can identify the vectors $(x_1/\rho, x_2/\rho)$ and $(\cos \omega, \sin \omega)$. So,

$$x_1 = \rho \cos \omega \tag{19}$$

$$x_2 = \rho \sin \omega \tag{20}$$

and we note that what we have done is to introduce polar coordinates for the cartesian ones.

Let us now differentiate (19)-(20) with respect to the arc length parameter σ, and for convenience denote derivatives with respect to arc length by a dot. So,

$$\dot{x}_1 = \dot{\rho} \cos \omega - \rho\dot{\omega} \sin \omega \tag{21}$$

$$\dot{x}_2 = \dot{\rho} \sin \omega + \rho\dot{\omega} \cos \omega \tag{22}$$

However, θ being the direction of the route, $\dot{x}_1$ and $\dot{x}_2$ denote the direction cosines as differentiation is with respect to arc length. Accordingly

$$\dot{x}_1 = \cos\theta \tag{23}$$

$$\dot{x}_2 = \sin\theta \tag{24}$$

We substitute from (23)-(24) into (21)-(22) and use Cramer's rule to solve for $\dot{\rho}$ and $\rho\dot{\omega}$, which are treated as the two unknowns in the resulting system. In the explicit solutions we make use of the formulas for the cosine and the sine of a difference to obtain

$$\dot{\rho} = \cos(\theta-\omega) \tag{25}$$

$$\rho\dot{\omega} = \sin(\theta-\omega) \tag{26}$$

These trigonometric expressions are now substituted back into our original differential equation (14), which reads

$$\rho\dot{\omega}\dot{k} + k\dot{\rho}\dot{\theta} = 0 \tag{27}$$

If we now differentiate (26) with respect to arc length once more and use (25) for $\cos(\theta-\omega)$, we get $k\dot{\rho}\dot{\theta} = 2k\dot{\rho}\dot{\omega} + k\rho\ddot{\omega}$. Substituting this into (27) and collecting terms (27) turns into

$$d/d\sigma(k\rho\dot{\omega}) + k\dot{\rho}\ddot{\omega} = 0 \tag{28}$$

But this is the same as

$$d/d\sigma(k\rho^2\dot{\omega}) = 0 \tag{29}$$

which has the first integral

$$k\rho^2\dot{\omega} = \text{constant} \tag{30}$$

The last formula is well known from the mechanics of central fields (like planetory motion). To understand the character of this new differential equation, let us denote the constant by c, write out $\dot{\omega} = d\omega/d\sigma$, and note that the arc length element $d\sigma$ equals $\sqrt{(\rho^2 + \rho'^2)}d\omega$ (where $\rho' = d\rho/d\omega$). Thus $\dot{\omega}$ is the reciprocal of $\sqrt{(\rho^2 + \rho'^2)}$ and (30) reads

$$k\rho^2 = c\sqrt{(\rho^2 + \rho'^2)} \tag{31}$$

which is an ordinary differential equation expressed in polar coordinates. It has been much studied in theoretical mechanics, and in fact its solution can always be obtained by integration (if the independent and dependent variables are interchanged).

Explicit solutions are hard or easy to obtain, depending on the character of k. Before giving some illustrations, let us just notice that if we substitute arc length $d\sigma = \sqrt{(\rho^2 + \rho'^2)}d\omega$ into (8) it reads

$$\lambda = \int_0^s k\sqrt{(\rho^2 + \rho'^2)}d\omega \tag{32}$$

If we regard $\rho(\omega)$ as an unknown function, that we have to choose so as to minimize λ, then we get (31) as the proper Euler equation for this variational problem. This corroborates the gradient law as we obviously get the same condition by seeking the optimal routes one by one (as parameterized curves $\rho(\omega)$) so that they minimize transportation costs. We also see that once we have solved for the flow lines, so that we know $\rho(\omega)$, then we can calculate λ. This, of course, is true for the given origin ξ. For another origin we have to run through the whole process again.

Let us finish the section by giving a very simple illustration by power functions

$$k = \rho^{a-1} \tag{33}$$

Unless a is zero, the solution is

$$\rho^{a} = \alpha \sec(a\omega + \beta) \tag{34}$$

This is a two-parameter family of routes, but fixing a point of origin removes one, and so we obtain a set of radiating curves. From (32) we can also calculate λ, using (33)-(34), and obtain

$$a\lambda = \sqrt{(\bar{\rho}^{2a} + \rho^{2a} - 2\bar{\rho}^{a}\rho^{a}\cos(a\omega - a\bar{\omega}))} \tag{35}$$

where $\bar{\rho}$ and $\bar{\omega}$ are the polar coordinates for the fixed point of origin ξ, and ρ and ω are the polar coordinates for the variable point of destination. The calculation of (35) is a bit awkward and therefore not reproduced. The logic is, however, simple evaluation of (32) with substitutions from (31)-(34) being made.

For constant λ, (35) describes a set of concentric transportation cost contours to which the routes defined by (34) are orthogonal. It should be stressed that the sufficency conditions for extremality are fulfilled for (34)-(35) only in a neighbourhood of the origin, more specifically in a wedge with vertex in the origin (of the coordinate system, not the central flow field).

It is easy to recognize the geometrical characters of these solution curves for low integral values of a. The value zero is a special case for which (34) does not hold. It will be dealt with below. The simplest of the remaining cases is when a = 1. Then, transforming (34) back into cartesian coordinates, we get

$$\cos \beta\ x_1 - \sin \beta\ x_2 = \alpha \qquad (a = 1)$$

This is obviously a two-parameter family of straight lines. This character of the routes is intuitively obvious as $k = \rho^0 = 1$ makes transportation cost equal to route length according to (32).

If we let $a = 2$, so that $k = \rho$, we deal with a case where transportation is cheap in the centre and gets more and more expensive towards the periphery. We expect that optimal routes are deflected from the straight line and become convex to the origin. This is verified by the formal solution. Passing again to cartesian coordinates we have

$$\cos \beta(x_1^2 - x_2^2) - \sin \beta(2x_1x_2) = \alpha \qquad (a = 2)$$

This formula represents the family of all hyperbolas that can be arranged symmetrically around the centre (of the coordinate system). By varying α we fill the four sectors, formed by a pair of orthogonal axes through the centre, by rectangular hyperbolas. By varying β, we simply rotate any such set of axes and its corresponding family of hyperbolas.

Next, letting $a = 3$, so that $k = \rho^2$ we note that the transportation advantages in the central parts become greater. We suspect that the convexity of the routes is even more pronounced. This is confirmed by the formal solution

$$\cos \beta(x_1^3 - 3x_1x_2^2) - \sin \beta(3x_1^2x_2 - x_2^3) = \alpha \qquad (a = 3)$$

Now, for any fixed β the space is split in six equal sectors (with vertices in the centre). This is like the previous case where space was split in four sectors. Again, the sectors are filled by hyperbolic curves, now more sharply convex as they are compressed in angles of 60° (instead of 90°). We deal with a so called monkey saddle flow, whereas we dealt with ordinary saddles in the previous case. Changing the value of β again rotates the whole system of solution curves.

From the fact that space is split into sectors (of 60° or 90°) in the two last cases, we can understand that there is no solution curve according to (34) that joins an origin and a destination, separated by a larger acute angle than one of 60° or 90°. So, it becomes intelligible that (34) only provides a local solution as hinted at. Fotunately, for the remaining cases there exists another solution to the optimal routing problem, namely radially from the origin in to the centre and out to the destination again.

As we know sufficiently much about positive values of a, let us now go in the reverse direction. Put a = -1. Then (34) in cartesian coordinates reads

$$\cos\beta\, x_1 - \sin\beta\, x_2 = \alpha(x_1^2 + x_2^2) \qquad (a = -1)$$

This equation represents the set of all circular arcs through the centre of coordinate space. As expected, the shape of the routes is now concave to the origin. As $k = \rho^{-2}$ it is least expensive to travel in the periphery and avoid the centre as much as possible.

Our final case of (34) is with a = -2. We expect the avoidance of the central parts to be even more pronounced. The formal solution

$$\cos\beta(x_1^2 - x_2^2) - \sin\beta(2x_1x_2) = \alpha(x_1^2 + x_2^2)^2 \qquad (a = -2)$$

represents the family of lemniscates through the centre. Again, for each β, space is split in four sectors. These sectors are now elliptic (not hyperbolic). These cases should be enough to help intuition to understand the solution (34) in general.

Let us finally record the special case of (33) where a = 0. Then

$$\ln\rho = \alpha + \beta\omega \qquad (36)$$

is the solution that replaces (34). The value of (32) is obtained according to

$$\lambda = \sqrt{(\ln \rho - \ln \bar{\rho})^2 + (\omega - \bar{\omega})^2} \tag{37}$$

which replaces (35).

Traffic

As indicated in the introduction, we define traffic at the location x by

$$i = \iint_A |\phi| \, d\xi_1 \, d\xi_2 \tag{38}$$

We integrate the norms of all the vector fields passing through x with respect to all possible points of origin ξ. But before being able to integrate according to (38) we must calculate the norms $|\phi|$, which we do not know yet. All the previous discussion concerns the routes of communication, not the volumes.

In the introduction we indicated how the demand for communication, by a gravity type of model, determines sink density and thus flow volumes. As a matter of fact we have already written down the exact mathematical condition for how flow volume changes with sink density in equation (4) above.

Observe that when we know the unit flow field $\phi/|\phi|$ then equation (4) renders a partial differential equation in flow volume alone. As $\phi = |\phi|(\phi/|\phi|)$, we get

$$\operatorname{div} \phi = \operatorname{grad} |\phi| \cdot \phi/|\phi| + |\phi| \operatorname{div}(\phi/|\phi|) \tag{39}$$

where $\phi/|\phi|$ and $\operatorname{div}(\phi/|\phi|)$ are known as soon as we know the flow lines. The only unknowns in (39) are $|\phi|$ and grad $|\phi|$, i e, the flow volume and its partial derivatives. So, (4) indeed supplies a differential equation for $|\phi|$.

Once we know $|\phi|$ for all ξ we can proceed to the integration (38) and calculate traffic.

We will illustrate the procedure by a few examples. First, suppose that transit cost is constant, i e, k = 1 on the whole region. Moreover, suppose we deal with uniform population density, p = 1, everywhere and that the region we deal with is the unit disk, $A = \{(x_1,x_2)|x_1^2 + x_2^2 < 1\}$. This is the simplest imaginable case.

Put a = 1 in (33). Then we know that (34) is a solution. For the present case it reads

$$\rho = \alpha \sec(\omega + \beta) \tag{40}$$

which is the familiar equation of a straight line written in polar coordinates. It is not surprising that the optimal routes with constant transit cost are straight lines, as in all classical location models.

It is more convenient to put the equation of these straight lines in parametric form. Using our familiar notation, where ξ_1,ξ_2 are the coordinates of the point of origin, and θ is the, presently constant, angle of the flow line, we write

$$x_1 = \xi_1 + \sigma \cos \theta \tag{41}$$

$$x_2 = \xi_2 + \sigma \sin \theta \tag{42}$$

As always, σ denotes the arc length parameter. Obviously (41)-(42) is an expression equivalent to (40) when we deal with a set of lines with a common point of intersection.

From (41)-(42) we can easily calculate arc length

$$\sigma = \sqrt{((x_1-\xi_1)^2 + (x_2-\xi_2)^2)} \tag{43}$$

Let us now check the gradient of this arc length measure. Obviously

$$\text{grad } \sigma = ((x_1-\xi_1)/\sigma,\ (x_2-\xi_2)/\sigma)$$

$$= (\cos\theta, \sin\theta) \tag{44}$$

$$= \phi/|\phi|$$

We can thus identify the unit flow field with the gradient of the arc length. As arc length is measured along straight lines, we obviously deal with an enclidean metric. The loci of equal distance then are concentric circles as defined by (43) for any given σ, and the pencil of radials through their common centre obviously is the gradient field to this, the simplest of all, metrics.

Now, using (44),

$$\text{grad } |\phi| \cdot \phi/|\phi| = \text{grad } |\phi| \cdot \text{grad } \sigma \tag{45}$$

and

$$\text{div } (\phi/|\phi|) = \text{div grad } \sigma \tag{46}$$

But the Laplacian div grad $\sigma = \partial^2\sigma/\partial x_1^2 + \partial^2\sigma/\partial x_2^2$ can easily be calculated from (43) to equal $1/\sigma$. On the other hand grad $|\phi| \cdot$ grad σ obviously is the derivative $\partial|\phi|/\partial\sigma$. So, according to (39), and (45)-(46),

$$\text{div } \phi = \partial|\phi|/\partial\sigma + |\phi|/\sigma \tag{47}$$

Using our information that $\bar{p} \equiv p \equiv 1$, equation (4) becomes

$$\partial|\phi|/\partial\sigma + |\phi|/\sigma + 1 = 0 \tag{48}$$

which is quite easy to solve. As neither θ, nor any derivative with respect to it appears in (48), we can treat it as an ordinary differential equation with σ as the only independent variable. The dependence on the angle θ is confined to a variation of the arbitrary integration constant only.

Denoting this constant by S^2, we obtain the solution

$$|\phi| = \frac{1}{2}(S^2-\sigma^2)/\sigma \tag{49}$$

As we deal with communication within the closed disk only, there is no flow crossing the boundary. As, moreover, the routes are straight lines, radiating from interior points of the circular region we see that no route can be tangential to the boundary curve. So, the condition that no flows cross the boundary translates to a condition that all flow volumes are zero on the boundary, i e, $|\phi| = 0$. From (49) we see that $S = \sigma$ on the boundary which means that S can be interpreted as the straight line distance to the boundary ∂A from the point ξ. In another wording, S is the distance from ξ to the boundary in the direction θ. This last formulation indicates how S depends on θ.

Our next task is to evaluate the double integral (38) from (49). But in order to make the integration efficiently we start by changing integration variables from ξ_1,ξ_2 to σ,θ. Now, (41)-(42) tell us that $\xi_1 = x_1 - \sigma \cos \theta$ and $\xi_2 = x_2 - \sigma \cos \theta$. Observe that when we integrate according to (38) we treat the point x_1,x_2 as fixed, thus reversing the roles of ξ and x in comparison to the previous discussion. It is easy to evaluate the Jacobian of the coordinate transformation as

$$\frac{\partial(\xi_1,\xi_2)}{\partial(\sigma,\theta)} = \sigma \tag{50}$$

Accordingly

$$d\xi_1\ d\xi_2 = \sigma\ d\sigma d\theta \tag{51}$$

and, from (38) and (49),

$$i = \frac{1}{2} \iint_A (S^2-\sigma^2) d\sigma d\theta \tag{52}$$

The evaluation of the innermost integral is messy, but straightforward, and yields the result

$$i = \frac{1}{6} \int_0^{2\pi} ((S'+S'')^3 - S'^3 - S''^3) d\theta \tag{53}$$

where

$$S' = \sqrt{(1-\rho^2\ \sin^2\theta)} - \rho\ \cos\ \theta \tag{54}$$

$$S'' = \sqrt{(1-\rho^2\ \sin^2\theta)} + \rho\ \cos\ \theta \tag{55}$$

Note that S' and S" are the lengths of the two segments into which ξ divides a chord of the unit circle in the direction θ. Now,

$$(S' + S'')^3 - S'^3 - S''^3 = 3S'S''(S' + S'') \tag{56}$$

and from (54)-(55)

$$S'S'' = (1-\rho^2) \tag{57}$$

$$(S' + S'') = 2\sqrt{(1 - \rho^2\ \sin^2\theta)} \tag{58}$$

Substituting from (56)-(58) into (53) yields

$$i = (1-\rho^2) \int_0^{2\pi} \sqrt{(1 - \rho^2\ \sin^2\theta)} d\theta \tag{59}$$

where we note that we have been able to move $S'S'' = (1-\rho^2)$ outside the integration sign, as it does not depend on θ. The rest of our expression too is a handy one. We note that the integral of $\sqrt{(1-\rho^2 \sin^2\theta)}$ taken over an angle $\pi/2$ defines the complete elliptic integral of the second kind. As $\sin^2\theta$ has a perfect periodicity over $\pi/2$ our integral is simply four times the elliptic integral, denoted as usually by $E(\rho)$. And so, finally,

$$i(\rho) = 4(1 - \rho^2)E(\rho) \tag{60}$$

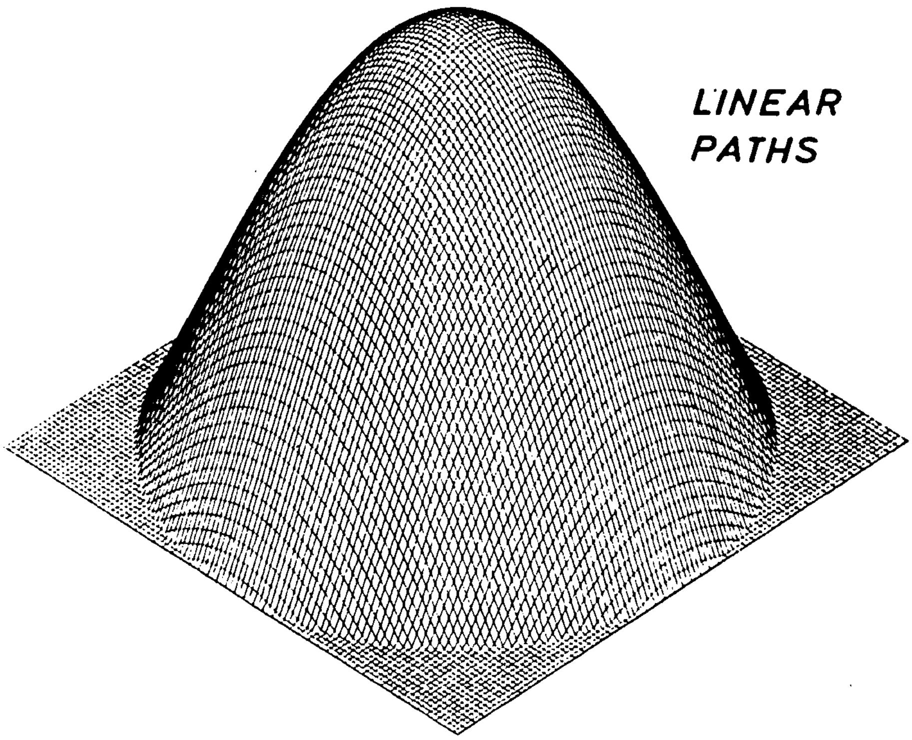
For the convenience of the reader we record the Taylor series for $E(\rho)$, which is the most handy way of computing it. Thus

$$E(\rho) = \frac{\pi}{2}\left(1 - \left(\frac{1}{2}\right)^2 \frac{\rho^2}{1} - \left(\frac{3}{8}\right)^2 \frac{\rho^4}{3} - \left(\frac{15}{48}\right)^2 \frac{\rho^6}{5} - \ldots\right) \tag{61}$$

The resulting traffic distribution is illustrated in Figure 1. Several comments are in order. First, we note that, even though the volumes of each flow, according to (49), did not possess circular symmetry, the traffic distribution has such symmetry. This is reasonable as the whole model is symmetric. The region is a circular disk, population is uniformly distributed, and transit cost is spatially invariant. So, traffic, i, should, according to intuition, have the symmetric property. On the other hand, the origin ξ, associated with the flow volume, $|\phi|$, is in general asymmetrically located in the disk, and so we should not expect any symmetry.

Our second observation is that the traffic distribution was relatively hard to derive despite the fact that we dealt with an extremely simple case. As our second example we will take one that is much easier to treat, but this is an exception. In general, we can expect a lot of computational difficulties. For a detailed discussion of traffic distributions and simulation techniques the reader is referred to Puu (1979).

Figure 1



Now consider our second example. What happens with the solution (34) if the exponent in (33) increases? If we draw the curves (34) for increasing a we see that they become more and more sharply convex to the origin. In the limit, as a goes to infinity, the routes become as convex as they can, i e, they degenerate into pairs of radials joining the points of origin and destination to the centre of the region, which is still the unit disk. So we arrive at the case of radial transportation that is so familiar from von Thünen and the New Urban Economics with its CBD.

Along with the disk-shaped region we retain the assumption of a uniformly dispersed population. As now from each point of origin all communations first go to the centre and then radiate out from there in all directions we conclude that the present flows all are in the direction of grad ρ. So, $\phi/|\phi|$ = grad ρ and all the formulas from (43) to (48) go through with σ replaced by ρ and $\xi_1 = \xi_2 = 0$. The differential equation, equivalent to (48), is now

$$\partial|\phi|/\partial\rho + |\phi|/\rho + 1 = 0 \tag{62}$$

Its solution resembles (49), but is simpler:

$$|\phi| = \tfrac{1}{2}(1-\rho^2)/\rho \tag{63}$$

The main simplicity is due to the fact that the distance to the boundary is presently a unitary constant, independent of θ. Accordingly, integration with respect to all the origins amounts to multiplication of (63) by the area π of our region. This results from the invariance of (63) with regard to ξ. Finally we have to keep in mind that we have only accounted for communication radiating out from the centre. There is as much communication radiating in to the centre, and hence we must double our measure. Thus, we get

$$i(\rho) = \pi(1-\rho^2)/\rho \tag{64}$$

This traffic distribution is illustrated in Figure 2. As traffic becomes infinite in our centre, we have removed the infinite peak at a certain level. It is not surprising that radial transportation leads to a higher degree of traffic concentration at the centre than does linear transportation.

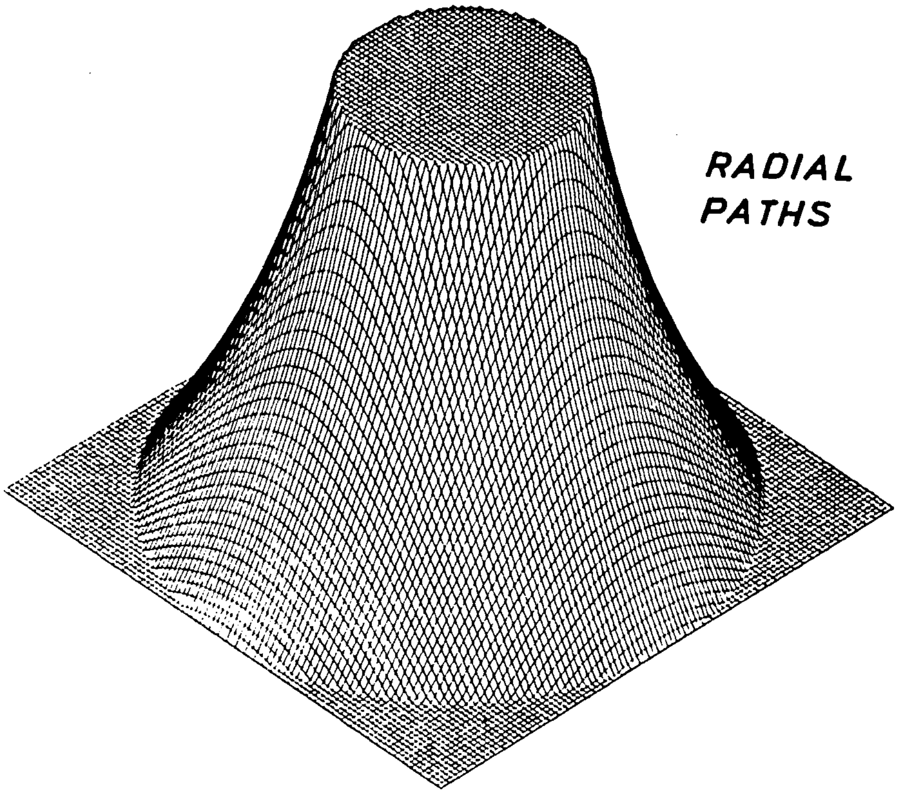
We should note that the fact that traffic is infinite in the centre does not mean that total traffic, the volume under the surface shown (including the infinite peak), is infinite. In fact total traffic $\iint_A i\,dx_1 dx_2$ is an improper integral that converges. Thus

$$\iint_A \pi(1-\rho^2)/\rho \; dx_1 \; dx_2 = 4\pi^2/3 \tag{65}$$

which can be compared to

$$\iint_A 4(1-\rho^2)E(\rho)dx_1 \; dx_2 = 128\pi/45 \tag{66}$$

Figure 2



As $4\pi^2/3 \approx 13.2$ and $128\pi/45 \approx 8.9$ about 50 percent more traffic is created by radial than by linear transportation. As linear transportation should lead to minimum total traffic, because it corresponds to the choice of the shortest route for each communication, the excess created by radial transportation is surprisingly small.

These two examples, expressly chosen to admit analytical treatment, should not give the impression that it is an easy task to derive explicitly all traffic distribution for any case we may wish to treat. On the contrary, the computation is in general very hard. This is particularly unfortunate, because we should deal with the formidable task of deriving an <u>equilibrium</u> traffic distribution when traffic is fed back, via congestion, into local transit cost, which determines the choice of routes and ultimately the traffic distribution itself. We have to conclude that we are not able to actually compute the final equilibrium traffic distribution by analytical methods. Computer simulation could be helpful, but considering the whole model it will be a formidable task.

Communication cost

As we noted above, we are able to calculate, in principle, of course, the communication costs for each point of origin once we are so far that we can calculate traffic.

Let us start by deriving a general relation between various expressions for communication costs. According to our assumption, if $\bar{p}$ people live at ξ and p live at x, then they need the number $\bar{p}p$ of communications. Each of these has a cost of λ, as defined by (8) when optimal routes are chosen in view of the given transit cost function k. So, the most obvious expression for transportation costs is

$$T = \iint_A \bar{p}p \, \lambda \, dx_1 \, dx_2 \tag{67}$$

To be exact T depends on the location of origin ξ. This point is fixed and the integration runs over all points of destination. Observe that this is the reverse of the case when we derived traffic distributions.

Now, equation (2) makes it possible to substitute -div ϕ for the product $\bar{p}p$. So,

$$T = -\iint_A \lambda \, \mathrm{div}\, \phi \, dx_1 \, dx_2 \tag{68}$$

We can transform this expression in a nice way by using Gauss's theorem, but there is one snag in it. The theorem is not applicable to the region A, because the vector field is not regular on it. The troublesome point is just the single location ξ of origin. If there were no net outflow from this singularity there would be no trouble, but we know there is!

So, we will use the artifice of defining a new sort of region with a small hole in it. The hole must contain ξ, but can be as small as we wish. For convenience, as we know that the constant λ contours are concentric closed curves surrounding ξ, we let the boundary of the hole be defined by some λ = constant. Denote this boundary $\partial'A$ and the region with the hole A'. Obviously, we can make the hole as small as we wish by letting $\lambda \to 0$. In other words, we can make A' as equal to A as we wish by this limiting procedure. The important feature of A' is that ϕ is regular on it, which makes Gauss's theorem applicable.

Consider the formula

$$\iint_{A'} \operatorname{div}(\lambda\phi)\, dx_1\, dx_2 = \int_{\partial'A} \lambda(\phi)_n\, d\sigma \qquad (69)$$

stating that the surface integral of the divergence of value flow $\lambda\phi$ equals the curve integral of the normal component of this flow along the boundary. This boundary $\partial'A$ is the inner boundary of the hole. Of course, there is an outer boundary, ∂A, of the whole region, but, as we only study internal communication in the region, we can delete this boundary integral from the outset, $(\phi)_n$ being zero on all of ∂A.

In (69) we can move λ outside the sign of integration, as the curve $\partial'A$ was conveniently defined by a constant λ. Next we use Gauss's theorem once more to transform the remaining curve integral of $(\phi)_n$ to a surface integral of div ϕ. Thus,

$$\int_{\partial'A} \lambda(\phi)_n d\sigma = \lambda\iint_{A'} \operatorname{div}\phi\, dx_1\, dx_2 \qquad (70)$$

But, div $\phi = -\bar{p}p$, where $\bar{p}$ as a constant can be moved outside the integration signs. What then remains in (70) to be integrated is population density. Let us denote total population of A' by P', in analogy to (2). Accordingly,

$$\int_{\partial'A}\lambda(\phi)_n d\sigma = -\lambda\bar{p}P' \qquad (71)$$

By letting λ approach zero $\bar{p}$ remains constant whereas P' goes to P, the population of the whole region A. Formally,

$$\lim_{\lambda \to 0} \int_{\partial' A} \lambda (\phi)_n d\sigma = 0 \qquad (72)$$

because $\bar{p}$ and P are finite, whereas λ goes to zero.

In this limiting process A' goes to A, and so we get from (69) the following relation for our (improper) integral on A

$$\iint_A \operatorname{div}(\lambda\phi)\, dx_1\, dx_2 = 0 \qquad (73)$$

Next, use $\operatorname{div}(\lambda\phi) = \operatorname{grad} \lambda \cdot \phi + \lambda \operatorname{div} \phi$ to get

$$\iint_A \operatorname{grad} \lambda \cdot \phi\, dx_1\, dx_2 = - \iint_A \lambda \operatorname{div} \phi\, dx_1\, dx_2 \qquad (74)$$

We are now prepared for the last step. From (6)

$$\operatorname{grad} \lambda \cdot \phi = k|\phi| \qquad (75)$$

and this substituted into (74) yields

$$- \iint_A \lambda \operatorname{div} \phi\, dx_1\, dx_2 = \iint_A k|\phi|\, dx_1\, dx_2 \qquad (76)$$

But this, according to (68) equals transportation cost and so

$$T = \iint_A k|\phi|\, dx_1\, dx_2 \qquad (77)$$

It is interesting to compare the starting equation (67) with the final derived equation (77). Transportation costs, originally expressed as the aggregate of the number of trips from the origin to other locations multiplied by the cost of each trip, can obviously also be obtained by taking the aggregate of the flow volume at each of the other locations multiplied by the local transit cost. In passing we should note that the equivalence of (67) and (77) applies to all flow fields, not only the optimal (cost minimizing) one, provided λ is defined as accumulated transit cost along the arbitrary flow lines. This is so as we do not need the optimality condition (6) itself, but only its weaker consequence (75).

Our equation (77) is much more useful than (67) both in actual computation and in the general discussion to follow.

Let us now give a simple example of how transportation costs can be derived for the case of $k \equiv 1$, $p \equiv 1$, and $A = \{(x_1,x_2)|x_1^2 + x_2^2 < 1\}$. This is the familiar case of homogeneous space, and hence linear transportation, and uniformly distributed population on the unit disk. We derived traffic distribution for this case.

As indicated it is useful to start from (77). As $k = 1$ we get

$$T = \iint_A |\phi| \, dx_1 \, dx_2 \tag{78}$$

Note the difference between this and the expression (38) above, defining traffic. The integration with respect to destinations x, not the origins ξ, makes a big difference, and the outcome will be different from (60).

However, part of the derivation leading to (60) is still relevant. So, we can use (49) directly. To facilitate integration, we again use the coordinate transformation (41)-(42). The Jacobian is

$$\frac{\partial(x_1,x_2)}{\partial(\sigma,\theta)} = \sigma \tag{79}$$

It happens to be the same as the one in (50) due to the symmetry of x and ξ in the formulas (41)-(42). Accordingly,

$$dx_1\ dx_2 = \sigma\ d\sigma d\theta \tag{80}$$

and so we get from substituting (4) and (80) into (78)

$$T = \frac{1}{2} \iint_A (S^2-\sigma^2)d\sigma d\theta \tag{81}$$

Now, we remember that S denotes the distance from ξ in direction θ to the boundary circle. The chord segments (distances in directions θ and $\theta+\pi$) have been recorded in (54) and (55). However, we only need one of them. So, we can put

$$S = \sqrt{(1-\bar{\rho}^2\sin^2\theta)} - \bar{\rho}\cos\theta \tag{82}$$

Observe that we take $\bar{\rho}$, not ρ, which again has to do with the fact that presently the origin, not the destination, is fixed.

We still have to fix the limits of integration in (81). Obviously, θ has to make a full round of 2π, but as the second half round only repeats the first one, we can let θ range from 0 to π and take twice the integral (81) with the limits for θ thus fixed. As for σ it obviously ranges from 0 to S.

To evaluate the innermost integral is trivial. We just get

$$\int_0^S (S^2-\sigma^2)d\sigma = \frac{2}{3}S^3 \tag{83}$$

Thus (81) becomes

$$T = \frac{2}{3}\int_0^\pi S^3 d\theta \tag{84}$$

with S being defined in (82). This last integral is a bit complicated to evaluate. Expanding the third power of (82) we get four terms, two of which involve $\cos\theta$ and $\cos\theta\sin^2\theta$. Now, the integrals of these from 0 to π are zero. The remaining terms are $(1-\bar{\rho}^2)\surd(1-\bar{\rho}^2\sin^2\theta)$ and $4\bar{\rho}^2\cos^2\theta\ \surd(1-\bar{\rho}^2\sin^2\theta)$ respectively. Both these have a perfect periodicity over $\pi/2$ and so

$$T = \frac{4}{3}(1-\bar{\rho}^2)\int_0^{\pi/2}\surd(1-\bar{\rho}^2\sin^2\theta)d\theta + \frac{16}{3}\int_0^{\pi/2}\bar{\rho}^2\cos^2\theta\ \surd(1-\bar{\rho}^2\sin^2\theta)d\theta \qquad (85)$$

Here we recognize, again, the definition of the complete elliptic integral of the second kind in the first integral. The second can also be evaluated in terms of complete elliptic integrals, but of both the first and second kinds. We already recorded the series expansion of the elliptic integral of the second kind in (61) above. For convenience we write the corresponding expression for the elliptic integral of the first kind

$$F(\rho) = \frac{\pi}{2}(1+(\frac{1}{2})^2\ \rho^2 + (\frac{3}{8})^2\ \rho^4 + (\frac{15}{48})^2\ \rho^6 + \ldots.) \qquad (86)$$

which is similar to (61). In fact the minus signs have been reversed and the denominators of the powers of ρ deleted, but otherwise they are the same.

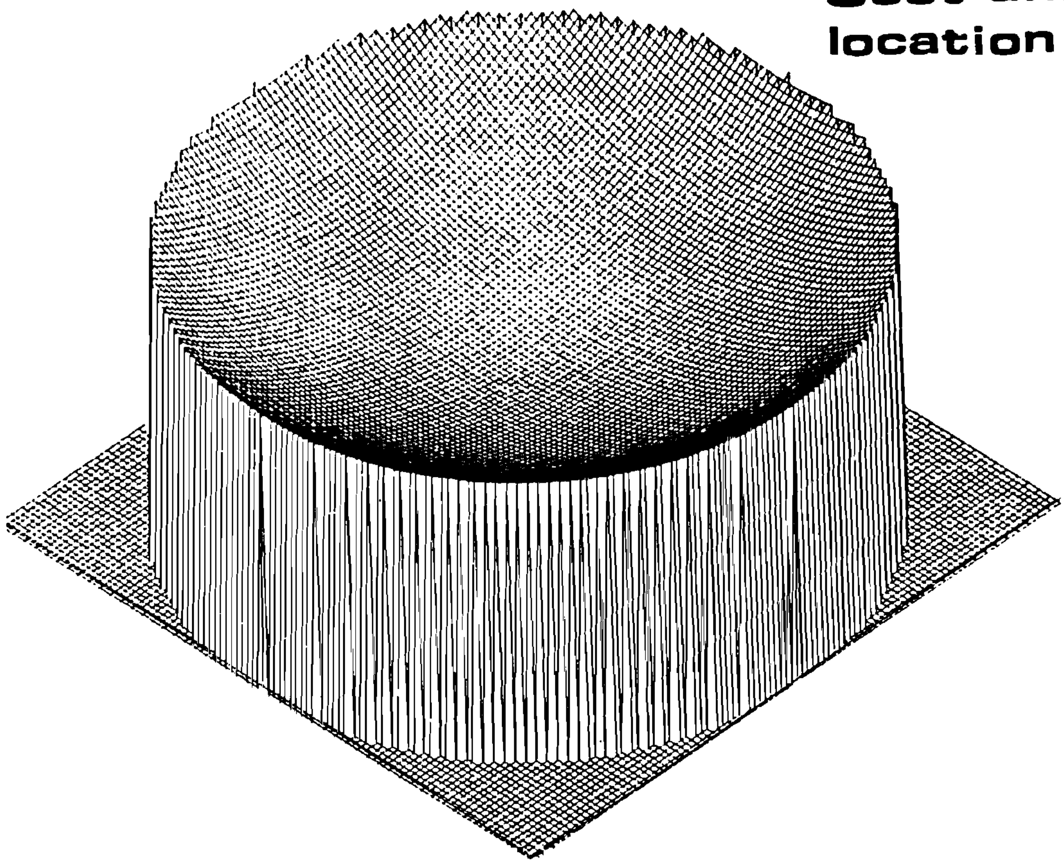
Using these elliptic integrals we finally have

$$T = \frac{4}{3}(1-\bar{\rho}^2)E(\bar{\rho}) + \frac{16}{9}((1+\bar{\rho}^2)E(\bar{\rho}) - (1-\bar{\rho}^2)F(\bar{\rho})) \qquad (87)$$

The distribution of transportation costs is illustrated in Figure 3. Obviously, transportation costs are lowest for those living in the centre and increase monotonically the farther the origin of communications is from the centre. This is appealing to intuition.

We could also compare Figures 1 and 2. Both resulted from integration of the same $|\phi|$, the first with respect to ξ, the second with respect to x.

Figure 3



The difference between the two Figures illustrates the importance of which coordinates we take for integration. However, the volume under the two surfaces is equal. Whether we take the integral of (60) with respect to x or the integral of (87) with respect to ξ, we arrive at

$$\iint_A \iint_A |\phi| \, dx_1 \, dx_2 \, d\xi_1 \, d\xi_2 = 128\pi/45 \tag{88}$$

as the order of integration is immaterial. One interpretation of this integral is total traffic, as we have seen. The other one is that of total communication cost. As local transit cost is unitary, we can equate total communication cost to total communication distance, or total transport work in a more familiar terminology. Hence, total traffic equals total transportation work. This is a conclusion that is not limited to the illustration case, but holds in general.

Land use and equilibrium settlement

Up to now we have considered how the choice of optimum routes in connection with the demand for transportation determined the distribution of traffic on the region considered and the distribution of communication costs for various points of origin. We considered the computational aspects of this in some detail in order to show how complicated an analytical solution can become even in mildly complicated cases. In this process the local transit cost was taken as a given datum. We only noticed that it depended on the congestion ratio of traffic to space available for transportation, and we noted that, due to the feed back mechanism via traffic, we could not regard transit cost as a datum, even if we assumed the fraction of space allocated to transportation as given everywhere.

We must now take in consideration the fact that the quantity of land available for transportation results from a decision concerning the use of land. What determines the quantity allocated to transportation, is the value of the best alternative use of land, which, in the framework of our model, is housing. The value of land use for housing, on the other hand, depends on population. If we, as indicated in the introduction, seek a spatial equilibrium where locations are indifferent, due to exactly balancing costs of housing and communication, then we must consider even the distribution of population as something variable in the model. But let us deal with the problems in order. First, we consider land use, then we proceed to equilibrium settlement.

We have already, in equation (5), defined local transit cost, k, as an increasing function of the traffic to carrying capacity congestion ratio, i/m. We already have a lengthy derivation of i. Let us therefore say that total land avaiable at a location is divided in two fractions, m, used for the transportation network, and, n, used for housing. "Housing" must be understood in a broad meaning, to include construction of buildings for productive purposes, along with residential construction, if our simplified model is to make any sense.

Let us now discuss a bit the dependence of k on i/m. Obviously, numerous empirical, as well as theoretical studies (of e g "follow-the-leader" type), suggest a monotonically increasing relation. This increase is very drastic as there is usually a critical congestion level at which the velocity of traffic flow comes down to zero, and hence its reciprocal, the transit time (a proxy for transit cost), goes to infinity. The general picture is not altered, even if we let k include capital costs for maintenance, as the need of repair due to wear obviously increase with congestion, as do the locomotion costs proper. We can also imagine that the transit cost function takes care of the fact that it is possible to push away the critical congestion ratio to a higher value, by creating <u>artificial</u> space, setting up several storeys of elaborate networks. However, capital costs for such constructions obviously increase with the ratio of traffic to <u>natural</u> space available, and so we can keep our specification.

We have said that a decision for land use has to be reached. Now, the use for transportation has been accounted for, but we still have to formalize the use for housing (in a broad meaning). So, let us suppose that there is a cost function

$$h = h(p/n) \tag{89}$$

for providing each individual with his required living space. This cost increases with the crowding ratio, measured by the quotient of population to natural space availaable for housing. Like the case with land use for transportation, we have in mind a process of creating artificial space at an ever increasing capital cost, the more artificial space, in relation to natural space, has to be constructed. As the need of space was proportionate to population, p/n is the correct argument. Finally, we can state equation (5) once more for convenience

$$k = k(i/m) \tag{90}$$

and the fact that the proportions of land, used for the two purposes included in the model, add up to unity

$$m + n = 1 \tag{91}$$

Now we have to choose the proper expression to optimize by the choice of m and n. In the introduction we argued that it would be reasonable to let the sum of housing and communication costs be minimized. However, it would be a little absurd to do this for each location separately. From an empirical point of view the planning of land use is something taken care of by public agencies planning for whole regions. Also, from the theoretical point of view all the communications, not only those from a certain point of origin, will be affected by changing transit cost there. Therefore, we can expect trouble with the analysis if we put up a local optimum condition for something having global effects.

So, dealing with the region as a whole, the total transportation costs are obtained by integrating (77) with respect to ξ and using the definition (38), as

$$\iint_A k(i/m)\ i\ dx_1\ dx_2 \tag{92}$$

On the other hand total housing costs are

$$\iint_A h(p/n)\ p\ dx_1\ dx_2 \tag{93}$$

Accordingly, we can minimize the sum of housing and transportation costs (92)+(93) with respect to the m and n, subject to (91).

This yields

$$k'(i/m)\left(\frac{i}{m}\right)^2 = h'(p/n)\left(\frac{p}{n}\right)^2 = \mu(x_1,x_2) \tag{94}$$

where μ is a (location dependent) Lagrangean multiplier associated with the constraint. This optimum condition has the nice property that it stipulates a universal relation that must hold everywhere between the local traffic congestion and population crowding ratios. We conclude, supposing second derivatives to be positive (as is reasonable with respect to our discussion above), that a high cost of transit due to congestion is coupled to a high cost of housing due to crowding. As a condition for optimal land use this seems reasonable.

We can also see that (94) and (91) together determine both m and n once i and p are given. The same is true then about k and h. Supposing that we have managed to solve the complicated feedback process of traffic as a determinant for route choice somehow and obtained the equilibrium traffic distribution we see that the single remaining degree of freedom is the spatial distribution of population.

We suppose that there is no incentive to migration if the sum of transportation costs and housing costs is a spatial invariant, i e, if

$$\iint_A k(i/m)\left|\phi\right| \, dx_1 \, dx_2 + h(\bar{p}/\bar{n})\bar{p} = \text{constant} \qquad (95)$$

Observe that the barred symbols again refer to conditions at the fixed point ξ of origin.

The model is now complete. If we limit our discussion to the case of a region with circular symmetry, we conclude that something like the case illustrated in Figures 1 and 3 comes some of the way to an equilibrium solution. Of course, this is only true in a very general sense.

However, the traffic displayed in Figure 1 arose from unit population density and linear communication routes. The latter occurred if transit cost was a spatial constant. Now, transit cost depended on the traffic congestion ratio. As we see from Figure 1, there is a traffic concentration to the centre of the disk. Accordingly, we have to allocate more land to transportation in the central parts in order to arrive at the constant transit cost (i e constant congestion ratio).

On the other hand, little land is available for housing in the central parts as it is used for transportation. So, housing should be expensive in the central parts. As for communication costs, we see from Figure 3 that they are low in the centre and high in the outskirts. It is thus possible that housing and communication costs could balance everywhere.

There is only one qualitative feature in this case that violates our conditions. We saw namely that, for optimal land use, the high population crowding in the centre should be balanced by a high congestion ratio. The latter, however, was a spatial constant.

So it seems that we should either have lower crowding or higher congestion in the centre. A higher congestion ratio would in equilibrium lead to avoidance of the centre, and via the feedback to a reduction of the concentration of traffic there. It could be brought about by allocating less land in the centre to transportation and more to housing. This change would lead to a better balance between crowding and congestion. As, however, the cost of communication would be increased for all having to communicate via the central parts, not only for those living there, whereas the housing costs would be decreased only locally it is likely that such a reallocation of land would make the centre more attractive and make people migrate there.

Finally, we are recognizing the features of reality. Congestion and crowding in the centre, a tendency to avoid the central parts for trips not originating or destined there, but nevertheless a considerable concentration of traffic, more land used for transportation than for housing in the centre, and high costs of housing, offset by centrality of location. It is meaningless to discuss matters in closer detail, because the general case is far too complicated to allow explicit solution.