ON THE GEOMETRY OF EMERGENCY SERVICE MEDICAL PROVISION IN CITIES

G. Hyman
L. Mayhew

April 1982
WP-82-23

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

THE AUTHORS

Leslie Mayhew is an IIASA research scholar working within the Health Care Task of the Human Settlements and Services Area. He is on leave from the Operational Research Unit of the Department of Health and Social Security, UK.

Geoffrey Hyman is an independent consultant in the field of urban and regional planning from the United Kingdom.

FOREWORD

The principal aim of health care research at IIASA has been to develop a family of submodels of national health care systems for use by health service planners. The modeling work is proceeding along the lines proposed in the Institute's current Research Plan. It involves the construction of linked submodels dealing with population, disease prevalence, resource need, resource allocation, and resource supply.

This paper considers a set of methods for the planning and monitoring of emergency medical services in large cities and for the allocation of resources -- vehicles, manpower, and equipment -- to facilities in the system from a geometrical viewpoint. The premise for the approach taken is that variations in weather and traffic conditions, the general complexity of the road systems in cities, and the risk of accidents occurring pose real problems for authorities who seek to provide a consistent and efficient service to the general public, but who find traditional analytic methods of dealing with these problems deficient.

Related publications in the Health Care Systems Task are listed at the end of the paper.

Andrei Rogers
Chairman
Human Settlements
and Services Area

ACKNOWLEDGMENTS

The authors would like to acknowledge the useful meetings and discussions they have had with officials of the London Ambulance Service, members of the Operational Research Services of the Department of Health and Social Security, UK, and scientists in the Human Settlement and Services Area at IIASA. None of these groups is responsible, however, for any views or errors of fact that may occur in the text. The authors would also like to convey their thanks to those IIASA staff involved in the production of the diagrams and the typing and editing of the text.

ABSTRACT

The enormity of many cities today poses special problems for authorities supplying emergency medical services. The scale of the emergency system and variations in the operating conditions rule out some of the traditional methods developed to deal with the complexities involved. In cities the provision of services is affected not only by the daily varying locations of the populations at risk, but also by the prevailing traffic and weather conditions that can hinder ambulance access to the site of an incident or to a treatment facility. The controlling authorities of these services nevertheless like to maintain the highest possible levels of service given the available resources. They are interested in knowing which facilities to open and when, what resources and staffing levels are required, and what the long- and short-term implications of changing operating conditions and of other unforeseen circumstances are on operating standards. This paper presents research into a set of methods that are designed to assist in finding solutions to these problems or, at least, in understanding how to deal with them. The methods are based on the efficiency of movement in cities, particularly the time it takes to access different locations in different traffic conditions, and on the likelihood of incidents occurring at different times of the day. Initial results are presented based on the London area, but the main conclusions are transferrable to many other cities of comparable scale. The distinguishing feature of the methods is that they are based on a type of non-Euclidean geometry that arises from the generalized profiles of the average speeds of traffic flow in cities.

CONTENTS

ON THE GEOMETRY OF EMERGENCY SERVICE MEDICAL PROVISION IN CITIES

1. INTRODUCTION

The enormity of many cities today poses special problems for authorities supplying emergency medical services. These problems are being studied by the Health Care Task at IIASA, and this paper presents some of the intermediate results. Specifically, it describes the development and trial application of a body of geometrical methods for use in long-term planning and routine operation of accident (A) and emergency (E) services in cities. Geometry seems well-suited to analyze such systems. For example, an ambulance is dispatched from a depot (a point) to the site of an accident or emergency (another point), and from there the patient is taken to an emergency treatment centre (a third point). The full journey cycle and its constituent parts needs to be completed within a reasonably short time: to reach the patient, to get the patient to the hospital, and to free the ambulance for response to another call.

The geometry of movement by ambulances in cities is not the simple geometry of Euclid, found in school text-books, where the quickest route between two points is a straight line. It becomes more complicated because of the need to minimize journey times. Travel speeds, however, vary between different locations due to traffic congestion and differences in road patterns, (Raitt 1981; Cantwell et al. 1973). Thus, the directional behavior of ambulances will depend on local conditions in each part of the city.

1.1 The Ambulance Velocity Field

In the geometry used to model the movement of ambulances, however, an important simplification can be made. That is, there is no need to pay close attention to the detailed road networks in the city and the routes that drivers take: rather, we are more concerned with factors such as the location of emergency facilities, the range of operation and geographical coverage of ambulances, typical journey times, and the allocation of resources. It suffices to base our analysis on local average velocities of ambulances, varying across the urban area. Functions describing such velocity variations are called *velocity fields* (Angel and Hyman 1976).

An alternative to the use of velocity fields might be to employ models of the road networks. However, this would impose far greater data processing requirements than geometrical methods, and could only be considered worthwhile at a very local scale of enquiry.*

1.2 Scope of the Paper

In the next section, we discuss the geographical factors affecting the operation of accident and emergency systems in large cities, introduce the idea of time standards for ambulance journeys, and give a checklist of practical problems to which the methods could be addressed.

In section 3, we present preliminary maps of ambulance journey times and show how they can be used to identify possible inefficiencies in emergency coverage and to spot areas where coverage appears to be weak. We then discuss the

*The London region, for example, has over 56,000 listed roads. Even if the network were reduced to include only the major roads, the data requirements would still be very large and many approximations would be needed.

construction of more sophisticated and realistic geometrical models to determine ambulance journey times.

In section 4, the basic geometrical theory is described, with mathematical details being given separately in the appendices.

In section 5, we describe methods for the determination of optimal geographical arrangements for facilities and apply these methods to the Greater London Council (GLC) area.

In section 6, the case study for Greater London is further developed. Density functions are estimated for patient caseloads as they vary over the urban area. The relationship between the ambulance travel time standard and the number of facilities required is derived; the number of facilities required to maintain time standards across the city is determined; the expected levels of patient caseloads at individual facilities are given; and the typical sizes of patient catchment areas are evaluated. Finally methods are described for determining the effects of closing a single facility on time standards and on the patient caseloads of nearby facilities.

2. THE OPERATION OF ACCIDENT AND EMERGENCY SYSTEMS IN CITIES

The functioning of accident and emergency services is affected not only by the patients requiring treatment and the resources available to the service, but also by the level of operating standards and by external conditions. We shall consider the implications of the following criterion for resource allocation. This is to configure treatment facilities geographically, so that practically every location of an accident in a city can be expected to be serviced by an ambulance within an acceptable time. Manpower, vehicles, medical equipment, and other resources are then to be allocated to each facility in accordance with their expected patient caseloads.

The expected patient caseload at individual facilities varies predictably by time of day, on weekends, and by time of year. These variations are partly due to the varying geographical locations of the populations at risk (e.g. commuters, shoppers, schoolchildren), but more importantly to the varying risk of accidents occurring. Figure 1, reproduced from Anderson (1978) shows the average daily variation in emergency calls. It can be seen that during the period 2:30 am. to 6:30 am. the number of calls seems to be about one third the average level over 24 hours. Clearly, such variations need to be recognized in the allocation and scheduling of the resources available at individual emergency facilities.

During the course of the day not only does the risk of accidents vary but also the time it takes ambulance drivers to reach a patient and deliver him to a treatment facility varies, due to changes in prevailing traffic conditions. These factors affect the numbers of facilities that need to be open in order to meet travel time standards and the numbers of ambulances and drivers that need to be available. Unforeseen shortages of staff and equipment further affect the daily operation of the emergency service, which should be sufficiently flexible to cope with these contingencies as well.

2.1 Ambulance Travel Time Standards

Despite the problems of varying operating conditions, the controlling officers of well-organized emergency services will usually have a set of standards that they would like to maintain (e.g., see Groom, 1975). A typical form of such standards would be the times that ambulances take to respond to emergency calls and deliver the patients to suitable nearby places of treatment. The determination of the treatment facility is generally up to the ambulance drivers, but it would usually be the nearest facility, due consideration being given to traffic conditions and other factors. Rigorous time standards may not be imposed in all cases, and medical priorities may be applied when resources are stretched. Nevertheless, it would still be

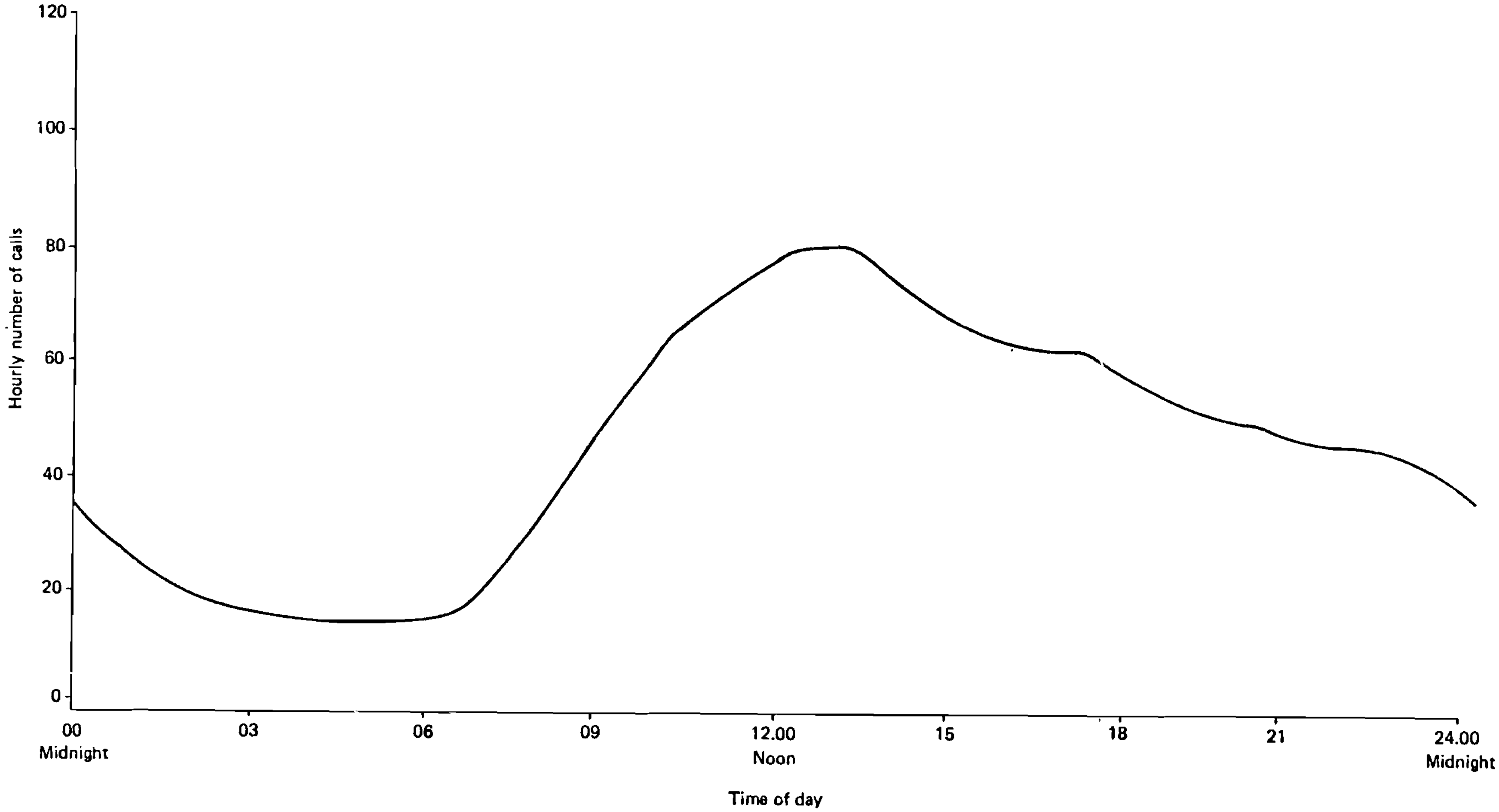


Figure 1. Average daily variations in the number of emergency calls in London for a Monday to Thursday period (Source: Anderson 1978,p.38).

expected that a substantial proportion of all emergency cases would receive treatment within suitable times, related to the experience gained in operating the service.

2.2 A Checklist of Problems for Study

In order to aid the discussion, it is useful to start with a checklist of problems that the proposed techniques might be adapted to solve in practical situations. These problems can be classed according to their significance in the long-term planning of A and E services, and to their day-to-day operation and monitoring applications. The problems posed are:

1. To determine the appropriate subset of facilities that will enable an imposed set of standards to be achieved everywhere in a city for a given set of operating conditions, taking into account size limitations at each facility (long-range planning)

2. To determine the local effects on caseloads and standards of adding or subtracting particular facilities from the set of facilities and to link such effects to staff availability (day-to-day operations and contingency planning)

3. To determine whether services are adequate to deal with major incidents such as fires, plane, or train crashes (long-range and contingency planning)

4. To undertake the routine monitoring of the system in a command and control environment to check that standards are being maintained and that there are sufficient resources available in each location (day-to-day operations)

5. To determine optimum configurations of facilities at given standards to provide benchmarks for evaluating the overall efficiency of existing facility distributions

6. To make recommendations on the approximate number of drivers, ambulances, doctors, and nurses needed on duty at each facility under different operating conditions and standards by day and night (long-range planning)

One important source of information for helping to resolve such problems are the returns filed by ambulance drivers.

These contain data on ambulance journey times and have previously been used in the analysis of the operation of the London Ambulance Service (Anderson 1978).

3. AMBULANCE TRAVEL TIMES AND ISOCHRONE MAPS

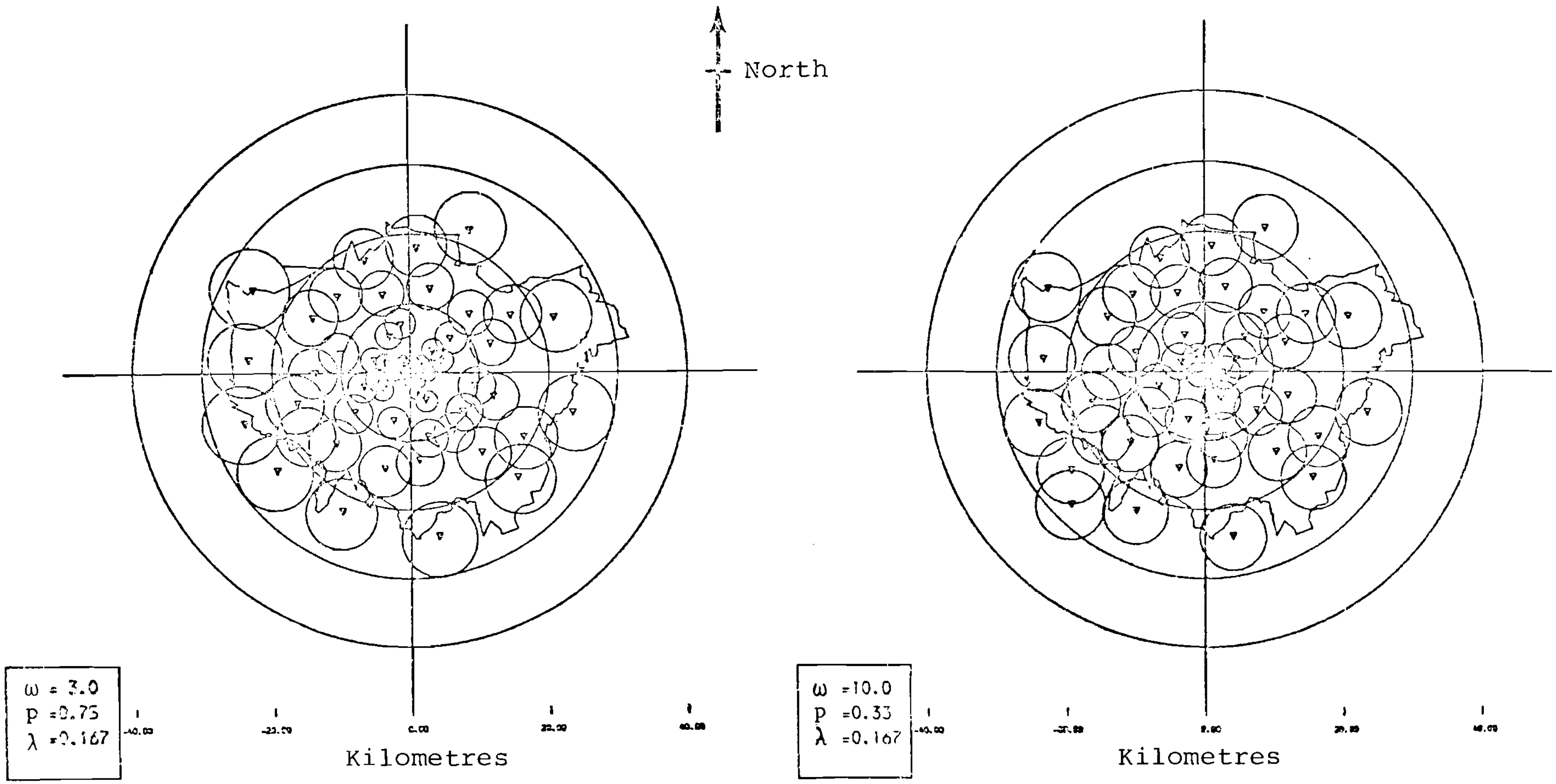
Important graphical aids to decision making are provided by maps based on ambulance travel times. From such maps many inferences relating to the problems in the checklist can be made (see also section 6). First, it is necessary to define an *isochrone*: the locus of points that can just be reached in a given time from a specified point. Such points may be ambulance depots or treatment facilities. If this time is an operating standard, then clearly all points within an isochrone can be reached within this standard. Conversely, points outside such an isochrone cannot be reached within the operating standard.

3.1 The Geographical Coverage of Emergency Facilities

The rings of isochrones show a potential observer the degree of geographical coverage obtained with a set of facilities and a particular time standard. Where two isochrones overlap, there is a duplication of coverage of the enclosed area; where there is no enclosing isochrone the area is not covered within the operating standard. Thus, an observer can easily detect areas where there could be unnecessary duplications of services as well as locations that cannot be served within specified time standards.

In Figure 2 we present four computer-drawn maps of the Greater London Council (GLC) region, with selections of the principal emergency treatment facilities plotted.* Around each

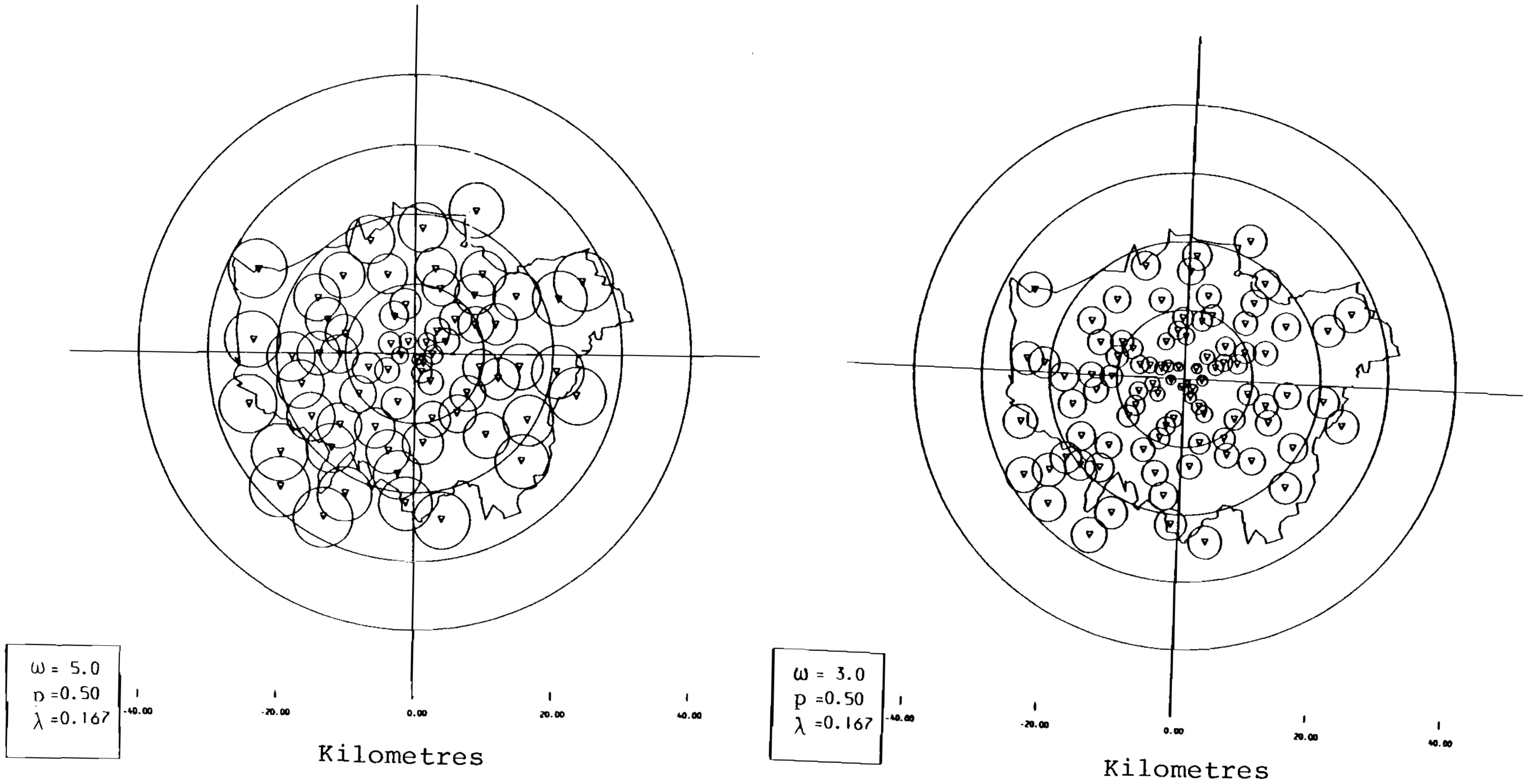
*Facilities are plotted only if they are not located within the isochrones of neighboring facilities (algorithm available). The method provides a general indication of the optimum numbers needed, but *not* exactly which ones.



(a) 53 facilities

(b) 44 facilities

Figure 2. Ten-minute (0.167 hours) isochrones around selected emergency medical centers in London for four different parameter sets.



(c) 66 facilities

(d) 90 facilities

Figure 2 (continued).

facility is drawn an isochrone for a 10-minute travel time interval. A velocity field has been used to create these maps in which V, the average velocity, is presumed to increase with the power p of the distance r from the city centre. That is,

$$V(r) = \omega r^p \qquad (0 < p < 1) \qquad (1)$$

In this equation, the velocity is lowest near the city centre to take account of the higher traffic congestion normally found there. This is why, in each case, the central isochrones enclose smaller areas than the peripheral ones. (For technical details of the derivation of isochrones see mathematical note 1.0 of the Appendix.) The parameters ω and p in equation (1) are the controls in the model that are used to simulate the background operating conditions. They are responsible for the substantial differences in coverage observed in the four examples, despite the same value for the time interval, λ, in each case.

The form of the velocity field also influences the shape of the isochrones. In the cases considered here, they are mostly off-centred circles, flattened on the side nearest the city centre (see Mayhew, 1981). Occasionally, because of the steep decline in velocity in the locality of the city centre, the isochrones in this model will be distorted into other shapes.

3.2. First Estimates of the Numbers of Facilities Required

Figure 2(b), with $V=10r^{1/3}$, and Figure 2(c), with $V=5r^{1/2}$, are both fairly close to more recent estimates of the ambulance velocity field for London. They indicate that there are *more* than sufficient treatment facilities to give a reasonable level of coverage within the ten-minute standard but suggest that patches of poor provision could exist in parts of south and east London. In Figure 2(c), there are 66 facilities; from 1976 data, however, 89 facilities in, and around the edge of, the study region treated more than 1000 new cases, while 111 reported some sort of activity. It would be expected that some facilities be held in reserve for use in adverse operating conditions or for the scheduling of different opening hours, although we have not touched on this problem so far. In sections 5.4 and 6.3, where the more recent estimates of the velocity field are used, it is indicated that 77 facilities, optimally located, would be needed during the daytime and 50 at night, when there is hardly any road congestion. Interestingly, there are just 77 ambulance depots in current operation, many less than the number of treatment facilities.

4. THE GEOMETRICAL THEORY

In this section the general geometrical theory behind the approach is developed. (Readers interested only in the application to emergency services can turn to section 6.) The theory is based on the work by Angel and Hyman (1976) and Mayhew (1979, 1981). To introduce it, we start with the problem of determining journey times between points in a city. To each point is assigned a velocity that is independent of direction and varies continuously from one point to another. The travel time λ on any path between the points A and B is

$$\lambda = \int_A^B \frac{ds}{V} \tag{2}$$

We seek the smallest value of this integral, and hence the quickest journey time. This is a problem in the calculus of variations*, and it gives rise to a function λ of the form,

$$\lambda = \lambda(x_A, y_A, x_B, y_B, a_1, \ldots, a_n) \tag{3}$$

where $x_A y_A$ and $x_B y_B$ are the coordinates of A and B and where $a_1, \ldots, a_n$ are the parameters of $V(x,y)$, the velocity field. If x_B, y_B are allowed to vary while the remaining terms are held constant, equation (3) describes an isochrone around the

---

*In fact the problem being considered is closely related to the famous brachistochrone problem (βραχιστος = shortest, χρόυος = time), solved in the late 17th century, which gave rise to the basic techniques in this field. Koo (1977, p. 172) gives a short account of this.

point (x_A, y_A). By considering A as the location of the facility, λ as the time standard, the isochrone around the facility can be mapped. Empirically, it is necessary to parameterize the function $V(x,y)$ under varying operating conditions by determining the values of $a_1, \ldots, a_n$. While rough estimates could be obtained from general traffic survey data, it is better to use data on the actual journeys reported routinely by the ambulance drivers. These may then be disaggregated into different times of the day or year as desired.

4.1 The City Pattern

The choice of velocity field will depend on the geographical characteristics of the city and its pattern of major roads. Here we give illustrations based on two major city patterns: *radial* and *linear*. The first consists of cities, with a dominant central focus like a business district, that radiate outwards in all directions. Examples in this class besides London could possibly include Paris, Vienna, Rome, Budapest, Warsaw, and Moscow. The second and smaller class of city is the type that develops linearly, instead of radially, along a highway, river, estuary, or coastline. Examples of such cities might be New York, Marseilles, and Brighton and Merseyside (England).

A particularly good example is the city of Genoa, an important, large port in Northern Italy. Points east-west along the coast in Genoa are connected by a modern highway, but inland access is greatly restricted by a patchwork of narrow, twisting streets and alleys that gradually merge into the mountains, which rise directly behind the city.

The distinction between linear and radial cities is, however, to some extent artificial. Most cities contain aspects of both: thus, New York and Marseilles also have strongly developed focal business districts and some radiating suburbs. In practice, therefore, more detailed prior data analysis will be required before deciding on the appropriate way to model a particular case. This will often require that different sections of simple local velocity patterns are combined to give the degree of accuracy desired.

4.2 Radially Symmetric Cities

If velocities in a city are functions of the radius r, the distance from the centre, the city is said to be *radially symmetric*. For these cases, there exist several classes of analytic forms for the equations of isochrones. These forms have been found helpful and realistic in building a velocity field for London, the city studied later in the paper.

4.3 Time Surfaces

A *time surface* can be imagined as a portion of the physical surface of the city that has been transformed into another surface on which shortest paths (geodesics) correspond to quickest paths in the city. Four examples of such surfaces are the plane, cylinder, cone and sphere. It is shown in Angel and Hyman (1976) and in the Appendix mathematical note 1.0 that these surfaces correspond respectively to the following four fields:

Plane: $V(r) = \text{constant}$ (4)

Cylinder: $V(r) = \omega r$ (5)

Cone: $V(r) = \omega r^p \quad (0 < p < 1)$ (6)

Sphere: $V(r) = ar^2 + b$ (7)

In these equations, ω, p, b and a denote fixed parameters. Figure 3 shows how shortest paths on the time surface are transformed into quickest paths, for a point south of the city centre, in the cases of the velocity fields described by equations (4) and (5). For equation (4), straight lines on the time surface, a plane, are simply transformed into straight lines in the city (Figure 3a). The geometry of the time surface is purely Euclidean, and the quickest times are proportional to straight line distances.

For equation (5), shortest paths on the surface of the cylinder between two, non-centrally located points are transformed into curves that spiral around the city centre (Figure 3b) to minimize the delays caused by traffic congestion. The equations of the paths and journey times are derived in mathematical note 1.4.

4.4 Linear Cities

Time surfaces have not yet been constructed in the case of linear cities. However, simple analytic solutions can be obtained for the following two cases

$$\text{Linear:} \qquad V(x) = a - bx \tag{8}$$

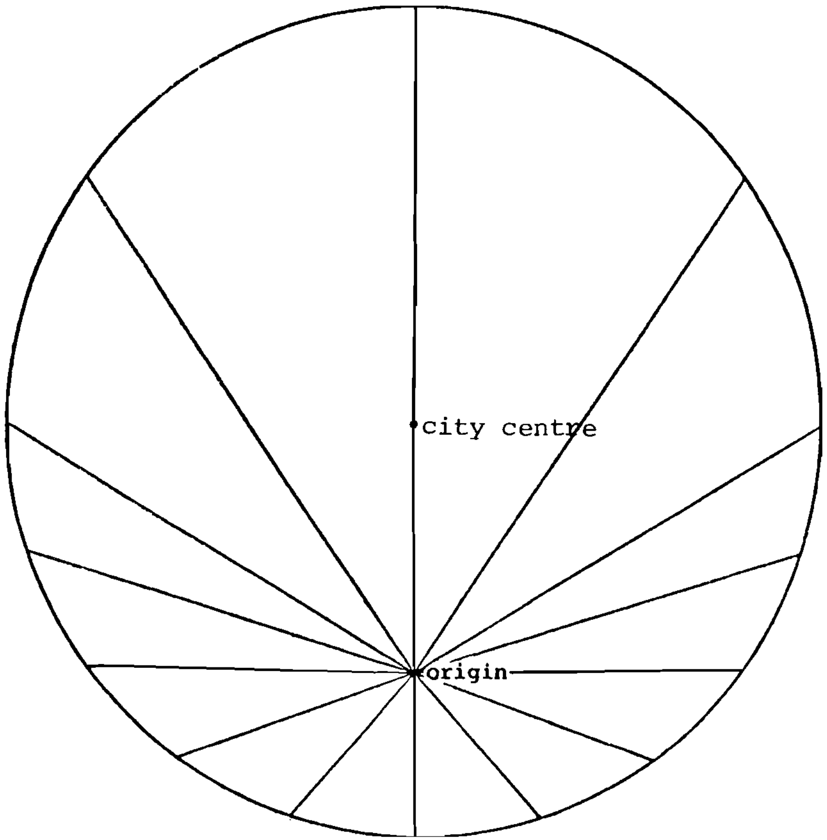
$$\text{Exponential:} \qquad V(x) = ae^{-bx} \tag{9}$$

by using variational techniques to solve the minimization problem in equation 2 (see also mathematical note 2.0). The parameters in these models are a and b.

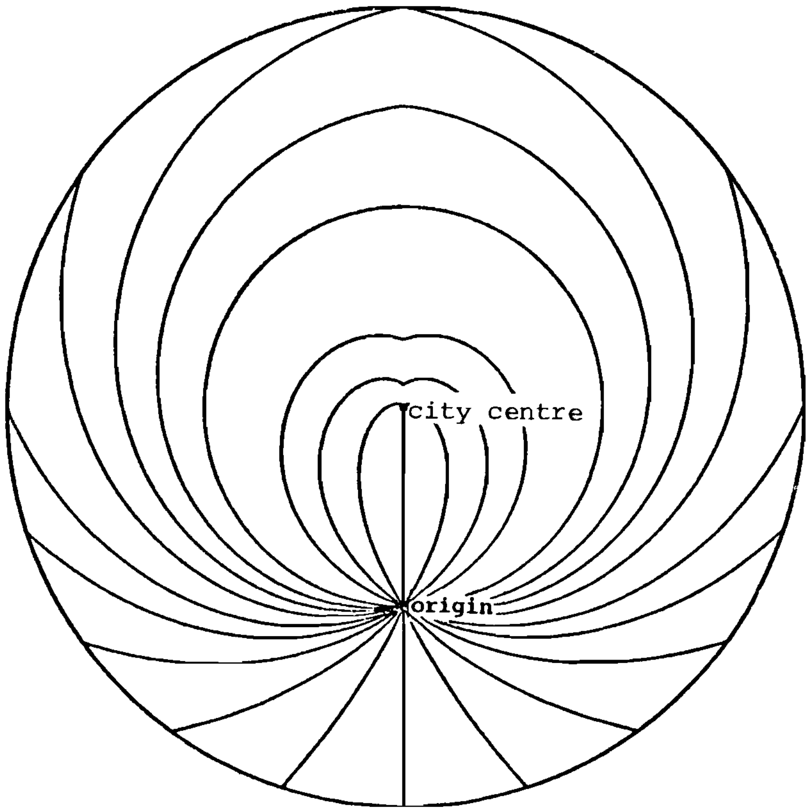
Figure 4a shows the linear velocity profile (equation 8) of a hypothetical coastal city. The coast runs from west to east along the y-axis. On this axis velocities are a maximum, corresponding to $V = a$.* To the north, inland along the x-axis, velocities decrease linearly with distance from the coast. At $x=a/b$, the velocity is zero (see Figure 4b). Quickest paths between nearby points can be constructed with a ruler and compass as follows. First construct the perpendicular bisector between the desired origin A and destination B. The intersection with the line $x = a/b$ (where $V = 0$) is the centre of a circle that passes through the two points. This circle is the required quickest path. For distant points, the circular arcs meet the coast and continue smoothly along it, as seen in the diagram. The isochrones for this field are also easily constructed (see mathematical note 2.0).

---

*To the south the sea produces a discontinuity in the velocity field.

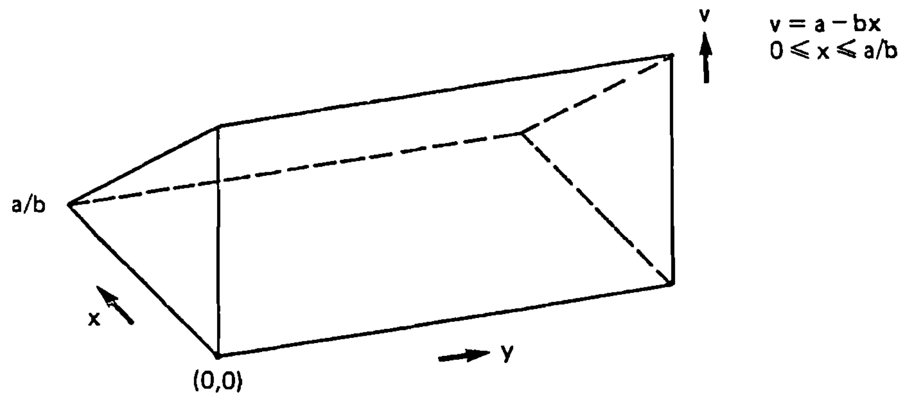


(a) V = Constant

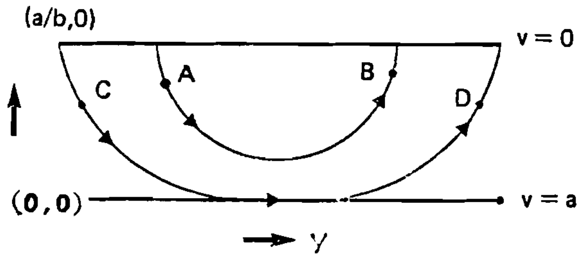


(b) $V = \omega r$

Figure 3. Quickest journey paths in a uniform and non-uniform velocity field from a point south of the city centre.



(a) A three-dimensional profile of a linear velocity field, V = a-bx. The y-axis represents the coastline.



(b) Quickest paths bend in the direction of the coast, which is served by a fast-access highway. Paths like CD join smoothly onto this highway, and then leave it smoothly.

Figure 4. A hypothetical linear city.

4.5 Segmentation of the Ambulance Velocity Field

More general examples of velocity fields combine sections of different fields when one alone cannot provide an adequate description of the velocities of travel in the whole city. A condition for paths crossing two glued sections has been discussed by Zitron (1974) and Braake and Zitron (1980). They required that the gradients of the paths be equal at the point of contact. This can be related to the more general Weirstrass-Erdman condition (Hadley and kemp 1971:40). Figure 5 shows an example of a candidate quickest path for a radially symmetric city in which velocities out to distance R from the centre are constant ($V = V_0$); thereafter, they increase linearly according to $V=\omega r$, such that at $R, V_0=\omega R$. Through the central section, the path is a straight line AB. These points are joined smoothly by paths in the outer section, which are sections of logarithmic spirals (CA and BD).

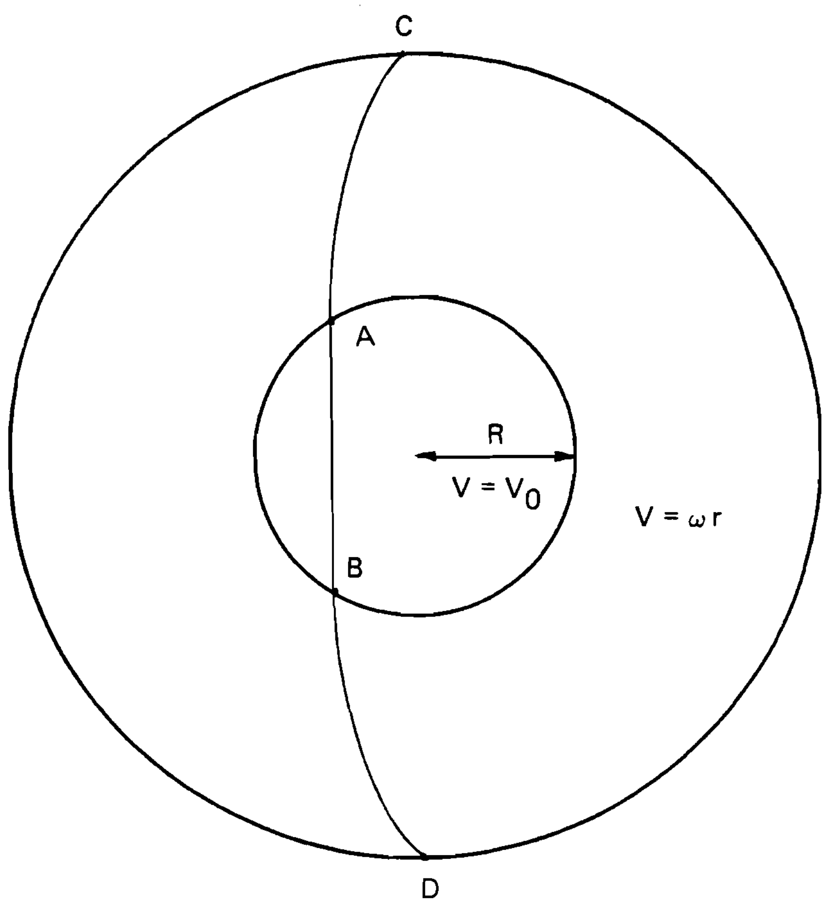


Figure 5. A radial city in which $V=V_0$ $(0<r\leq R)$ and thereafter increases according to $V=\omega r$. The diagram shows a possible quickest path CD across the city (adpated from Zitron, 1974).

4.6 Geographical Obstacles to Ambulance Travel

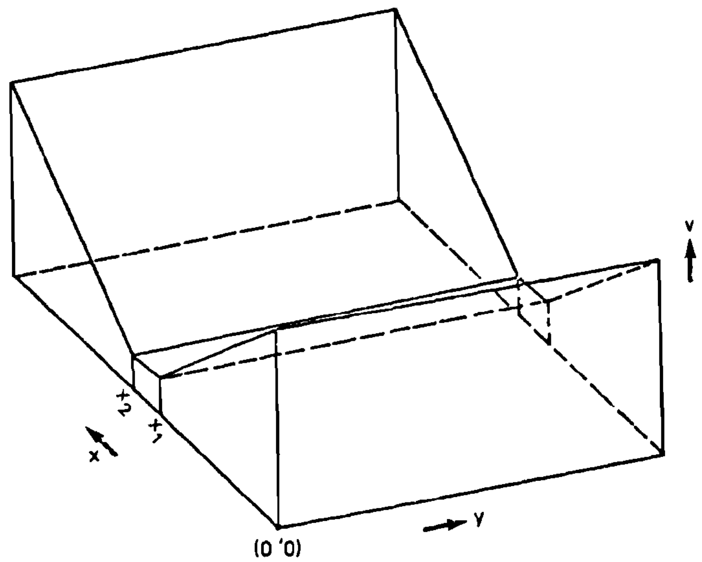
Apart from the case of coastal cities, discontinuous velocity fields also occur when there are obstacles to travel that must be avoided. Here, three types of obstacles are investigated: (i) line barriers, such as rivers; (ii) open expanses, such as parks, water reservoirs, airports, and industrial areas; and (iii) linear obstructions such as railway lines.

The difference between the first and the third type of barrier arises from the density of crossing points. If there are sufficient bridges, then the effect might be represented by a continuous reduction in local velocities as the obstruction is approached.

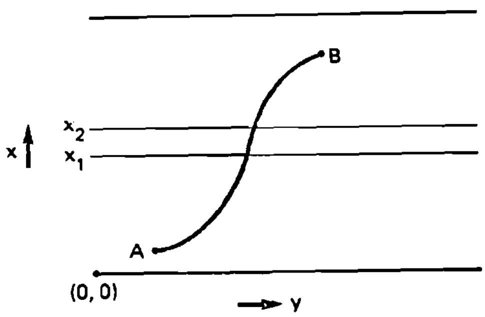
Figure 6, for example, shows two linear fields glued side by side on to a thin strip of constant velocity representing a type (iii) linear obstruction. Paths such as AB are then generated consisting of two circular arcs, meeting along a common tangent. A complete treatment of this example indicates possible bifurcations of routes, so that two optimum paths can exist between two points. These paths meet along a line, forming cusps in the pattern of quickest paths. This pattern is shown in Figure 7, in which the constant strip has been shrunk to have an infinitesimal width.

In contrast to this example, Figures 8 and 9 show two cases of the first two types of barriers, a river and an open expanse, in which path discontinuities are seen to be inevitable. In Figure 8, two routes are shown to cross the river via bridges A or B. This causes the isochrones to contain cusps on the opposite bank. Points to the left of the curve joining the cusps are reached quickest via the bridge A. The curve of cusps is in fact an hyperbola (mathematical note 3.1).

For the open expanse in Figure 9, each vertex A,B,C of the obstruction is the starting point for a change in the pattern of the isochrones for areas that lie in the shadow of the origin, O. The isochrones intersect on the far side of the barrier and a similar curve of cusps to the bridge example is generated.



(a) A profile of the velocity field, where the strip between x_1 and x_2 represents the obstruction.



(b) A quickest path between A and B.

Figure 6. A type (iii) linear obstruction such as a railway line.

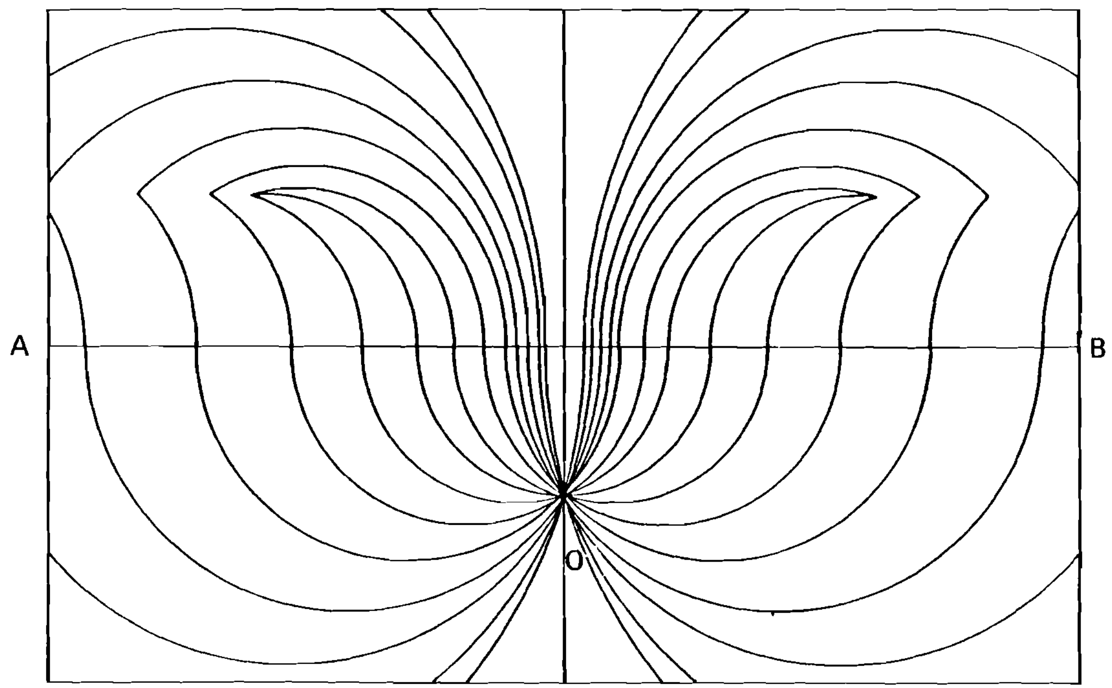


Figure 7. A type (iii) linear obstruction, AB, showing quickest paths from O (see text).

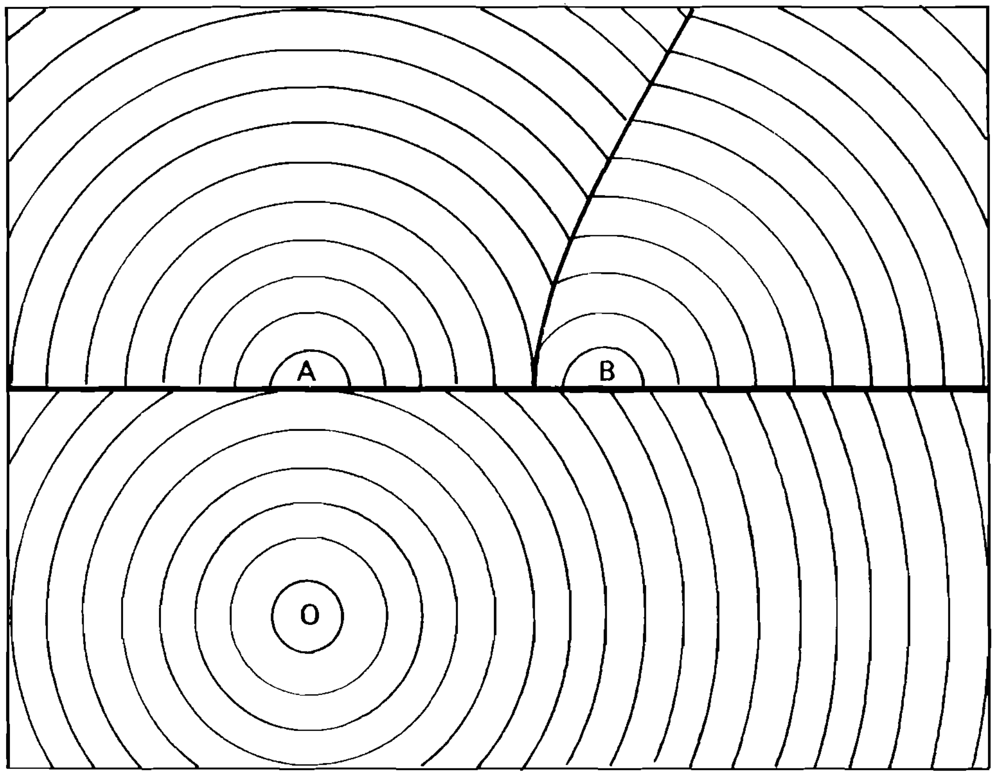


Figure 8. A type (i) river barrier showing isochrones radiating from O, the journey origin. The thick line is the river, and A and B are two crossing points. The line joining the cusps on the opposite bank divides quickest routes to destinations on the left via A or on the right via B.

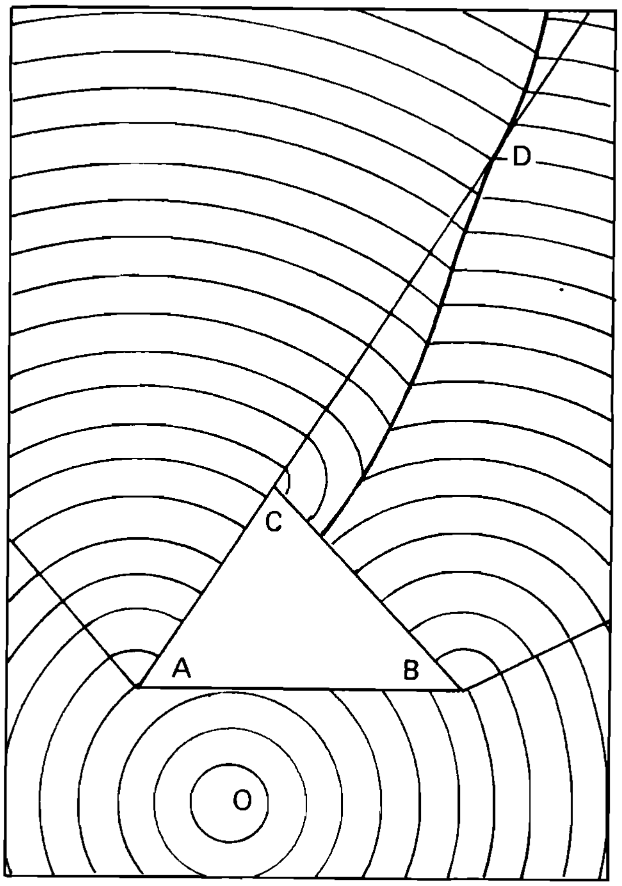


Figure 9. A type (ii) obstacle ABC: The journey origin is O. A curve of cusps through D divides quickest routes as in Figure 8.

5. THE OPTIMUM GEOGRAPHICAL COVERAGE OF FACILITIES

It may be the objective of the operating authority to locate facilities so that, for a set of given conditions, no destination is greater than a fixed travel time standard from a facility. The question then arises: what is the minimum number of facilities required to achieve this goal? We consider a radially symmetric city in which (i) facilities are free to locate anywhere; (ii) there exists a time surface; and (iii) there are no barriers to travel. The area of the time surface is independent of the number and location of facilities. To cover this area with the minimum number of facilities, we want the maximum area to be served by each facility, provided the areas served by different facilities do not overlap. The radius of each such area must not exceed the time standard. Thus, we seek to cover the time surface with service areas that do not overlap and each of which has maximum area for a given radius.

5.1 Tessellations on the Plane

Let us first consider the case in which the velocity of travel is uniform, so that the time surface is a plane. Regular figures maximize the ratio of area to radius (mathematical note 4.1). Thus, the optimum pattern of facilities is a regular tessellation on the plane. There are three possibilities: equilateral triangles, squares, and hexagons (see also mathematical note 4.1). Of these, the hexagon has the maximum ratio of area to radius, and is thus the most efficient.*
For velocities varying continuously the time surface is not a plane but the geometry of the region around a point is locally Euclidean, so that the optimum covering will be locally

---

*Hexagonal market areas also appear for the same reasons in "Die Zentralen Orte in Süddeutschland" (Christaller 1933), which deals with the geographic organization of settlements in a region. See also Tobler (1963) who comments on this, and touches on some of the problems discussed later.

described by a tessellation of regular hexagons. Distortions to this pattern will take place over larger areas as equal-sized hexagons on the time surface correspond to unequal areas of the city.

5.2 Tessellations on the Cylinder, Cone, and Sphere

If the velocity field is $V = \omega r$, the time surface is a cylinder, and the city can be optimally covered with a constant number of facilities around each ring of constant radius (see mathematical note 4.2). If the velocity field is $V = \omega r^p$, the time surface is a cone. Hexagonal tessellations will only fit on the cone for restricted values of p. There are 13 possible values: $0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1, \frac{7}{6}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, \frac{11}{6}, 2$. The cases 0 and 1 correspond to the plane and cylinder already discussed. The cases where $1 < p \leq 2$ correspond to inverses of the cones $0 \leq p < 1$ (see mathematical note 4.3). Unless p is less than unity the number of facilities required diverges as the city centre is approached. In the rest of this section, we confine attention to the cases where p is less than unity. For these values of p, the conical tessellations can be constructed by cutting sectors out of a plane hexagonal tessellation and gluing the cut edges back together.

If the velocity field is $V = ar^2 + b$ the time surface is a sphere. It is impossible to fit a hexagonal tessellation onto the entire sphere (see mathematical note 4.4). In practice only a portion of the sphere would need to be covered, so this is not a serious restriction. However, it is of theoretical interest to investigate tessellations that are efficient covers of the entire sphere. The two best examples are the dodecahedron and the trunctated icosahedron. The former is a regular Platonic solid, which contains 12 pentagonal faces. The latter is a solid with 12 pentagonal faces and 20 hexagonal faces (often used in the construction of soccer balls). However, these could only form solutions to problems of optimum facility coverage if the velocity field $V = ar^2 + b$

and the time standard happen to be scaled in suitable proportions (see mathematical note 4.5).

Figure 10 is a photograph of all the tessellated surfaces described so far. Figure 11 shows the tessellations after they have been transformed to the physical surface of the city. (For details of the transformation, see mathematical note 1.0). The scale of the tessellations depend on the form and parameters of the velocity field, and the time standard λ. Table 1 gives the values of the parameters chosen for these examples.

Table 1. Key to the tessellations.

Figure No.	Velocity field (km/hour)	Time standard (hours)	Surface
11.1	17.0	0.229	plane
11.3	$27.0\ r^{1/6}$	0.096	cone
11.4	$15.0\ r^{1/3}$	0.119	cone
11.5	$8.0\ r^{1/2}$	0.164	cone
11.6	$4.5\ r^{2/3}$	0.242	cone
11.7	$2.5\ r^{5/6}$	0.475	cone
11.2	$1.5\ r$	0.155	cylinder
11.9	$0.77r^2 + 5.0$	0.167	sphere
11.8	$0.31r^2 + 5.0$	0.169	sphere

5.3 Notes Concerning the Tessellations

There are a number of points to note about the previous tessellations.

1. To simplify the construction, the boundaries of the catchment areas are shown as straight lines. While this is an adequate approximation at our scale of inquiry they should more correctly be shown as curves.

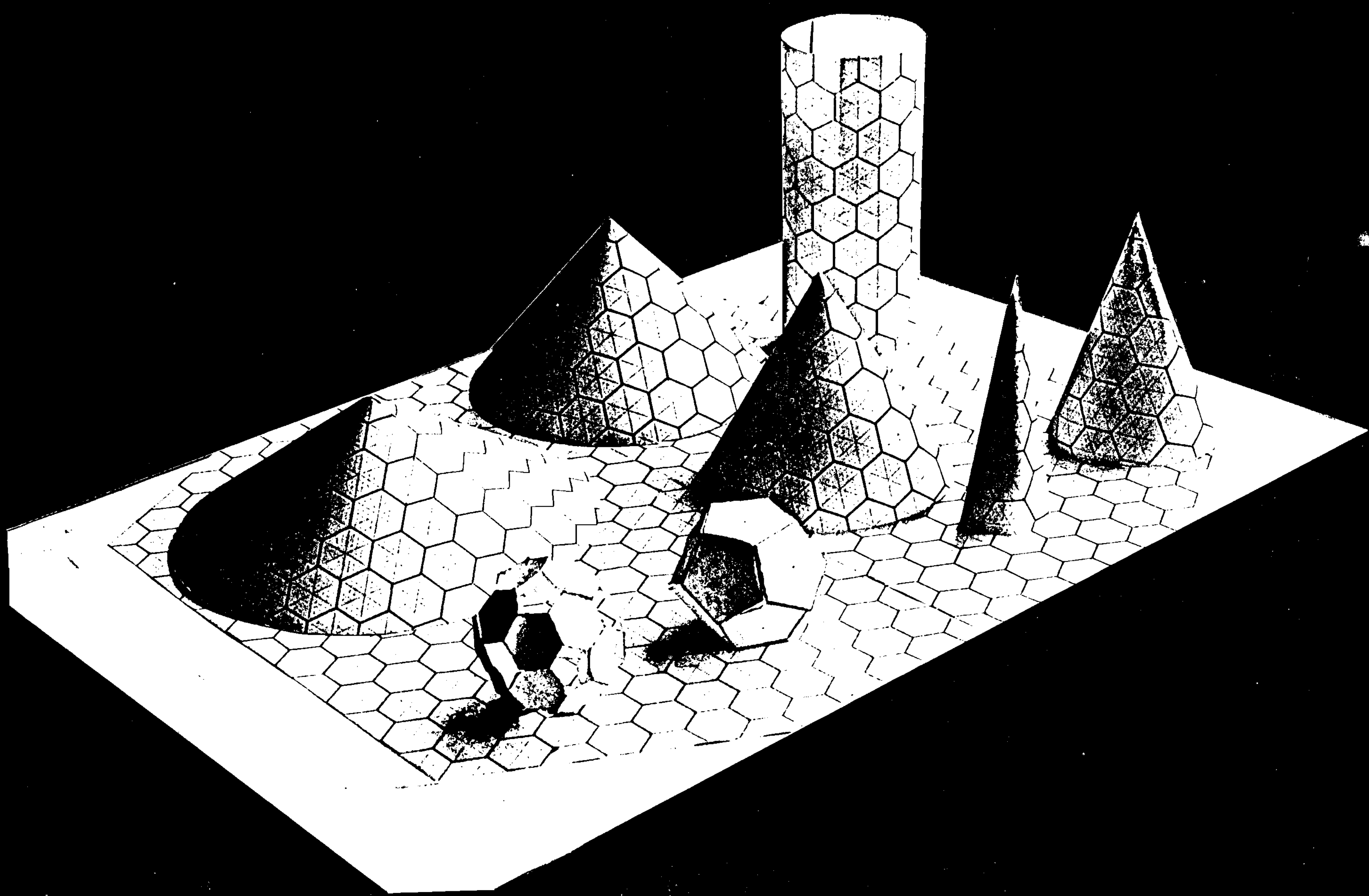


Figure 10. The basic tessellated surfaces.

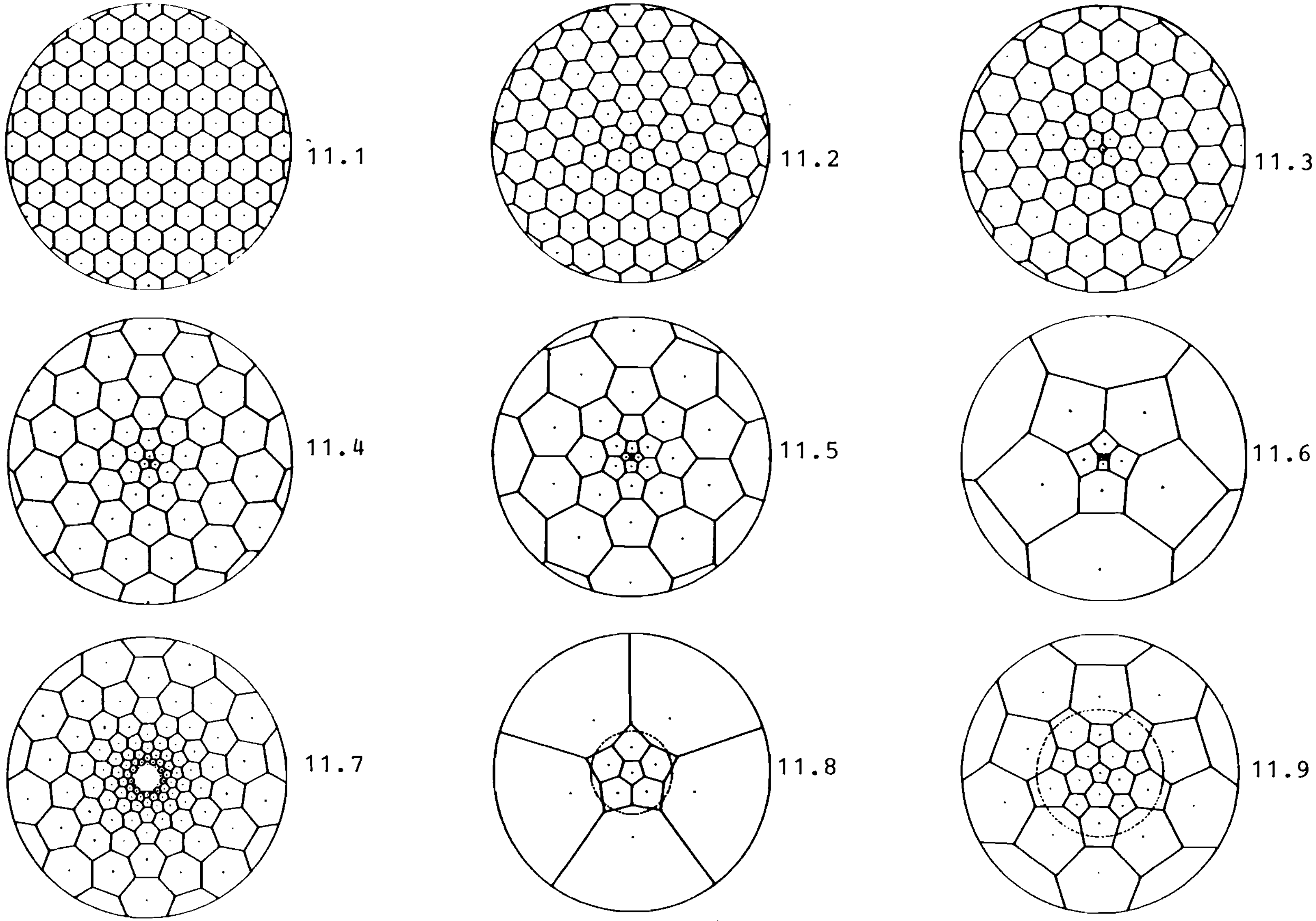


Figure 11. The tessellations transformed to the physical surface of the city.

2. In the cylindrical transformation ($V = \omega r$), the number of facilities is constant in each ring, so that the number of facilities per unit area diverges at the city centre. Since the demand for A and E services is finite it would be impossible to maintain such a pattern near the centre.

3. For the conical transformations ($V = \omega r^p$, including the plane as a special case), the shape of the central facility catchment polygon depends on the value of p. The possible shapes are:

p	central catchment polygon
0	hexagon
1/6	pentagon
1/3	square
1/2	triangle
2/3	biangle*
5/6	monoangle

The number of facilities packed around the city centre increases in successive rings. It varies with p according to $6n(1 - p)$ where n is the ring ($n = 1 \ldots$) and where the first ring is that touching the central polygon.

4. For the two examples shown of the spherical transformation ($V = ar^2 + b$), the tessellations are either all pentagons or a mix of pentagons and hexagons. In all such cases, the facility antipodal to the central facility cannot be shown as it is located at infinity. When the tessellations consist of mixed figures, as in the case of the "soccer ball", time standards λ are not uniformly achieved — $\lambda_{Pentagon}$ being less than $\lambda_{Hexagon}$. The dotted line shown is the image of the equator of the time surface.

These illustrations show that it is possible to construct optimal patterns for the location of emergency facilities to meet a consistent ambulance travel time standard when the ambulance speeds vary across the city. The most flexible class of velocity field was seen to be particular types of

*biangle = 
monoangle = 

power-law variations of velocity with distance from the city centre. These are associated with conical time surfaces and, unlike the examples given for the sphere, are not subject to scale restrictions. Furthermore, the conical time surfaces can be exactly covered with hexagons at all locations apart from the city centre.

To construct a more realistic ambulance velocity field for London, it will be necessary to glue together parts of cones with different power-law variations, as described in the section 5.4.

Later in section 6 the London velocity field will be used to show how the number of facilities required varies with the ambulance travel time standard. In these calculations we will assume that hexagonal areas are an adequate approximation to the local shapes of tessellations on the London time surface.

5.4 The Optimum Geographical Coverage of Facilities in London

A variety of statistical sources were used to estimate the velocity field for London for purposes of presenting the initial results. They were the average distance travelled per patient per year by ambulances in London in 1977; the average time from the scene of an accident or emergency to the treatment facility; the variation in distance separating existing treatment facilities; and general traffic data for London. From these sources two velocity fields were derived: one assumed to operate during the day and evening and the other at night. They are as follows:

1. <u>Day and Evening</u>

$$V(r) = 8.8 \qquad 0 \le r \le 4 \tag{10}$$

$$V(r) = 4.4\, r^{1/2} \qquad 4 < r \le 15 \tag{11}$$

$$V(r) = 1.13\, r \qquad 15 < r \le 25 \tag{12}$$

$$V(r) = 28 \qquad r \ge 25 \tag{13}$$

where V is expressed in kilometres per hour and r in kilometres.

2. Night

$$V(r) = 16.95 \qquad 0 \le r \le 15 \qquad (14)$$

$$V(r) = 1.13\, r \qquad 15 < r \le 25 \qquad (15)$$

$$V(r) = 28.0 \qquad r \ge 25 \qquad (16)$$

The daytime field has a constant velocity for the first 4 km from the centre. Between 4 km and 15 km it rises in proportion to the square root of distance from the centre. From 15 km to 25 km it rises in direct proportion to distance from the city centre. Beyond 25 km the velocity field is constant. Where the local sections join, at 4 km, 15 km, and 25 km the velocity varies continuously.

The nighttime field also has a constant velocity in the central area, but this velocity is greater than the daytime central velocity and the central section reaches out as far as 15 km. Beyond 15 km the nighttime field is the same as the daytime field.

The time surface for the daytime field, apart from the central constant section, was tessellated on a scale that corresponded to a ten-minute maximum journey time. This corresponds approximately to the actual average times achieved in London from the incident sites to the treatment centres as estimated from Anderson (1978). The total number of treatment facilities that we located *within* the Greater London Council area was 77, somewhat larger than the number (64) quoted in Anderson.*

The resultant surface is shown in Figure 12. To align the hexagons exactly the conical section of the surface has been deformed into a 3-sided pyramid and the cylindrical section into a 3-sided prism. On either side of the edges of the prism, the hexagons are unavoidably squashed into

*In fact Anderson excludes several facilities that reported accident and emergency services in 1976, though he notes (p.41) that in all, 101 hospitals were involved to some degree. Our data showed that 89 facilities treated more than 1000 cases in 1976, while 111 reported some sort of activity in, and on the edge of, the Greater London Council region.

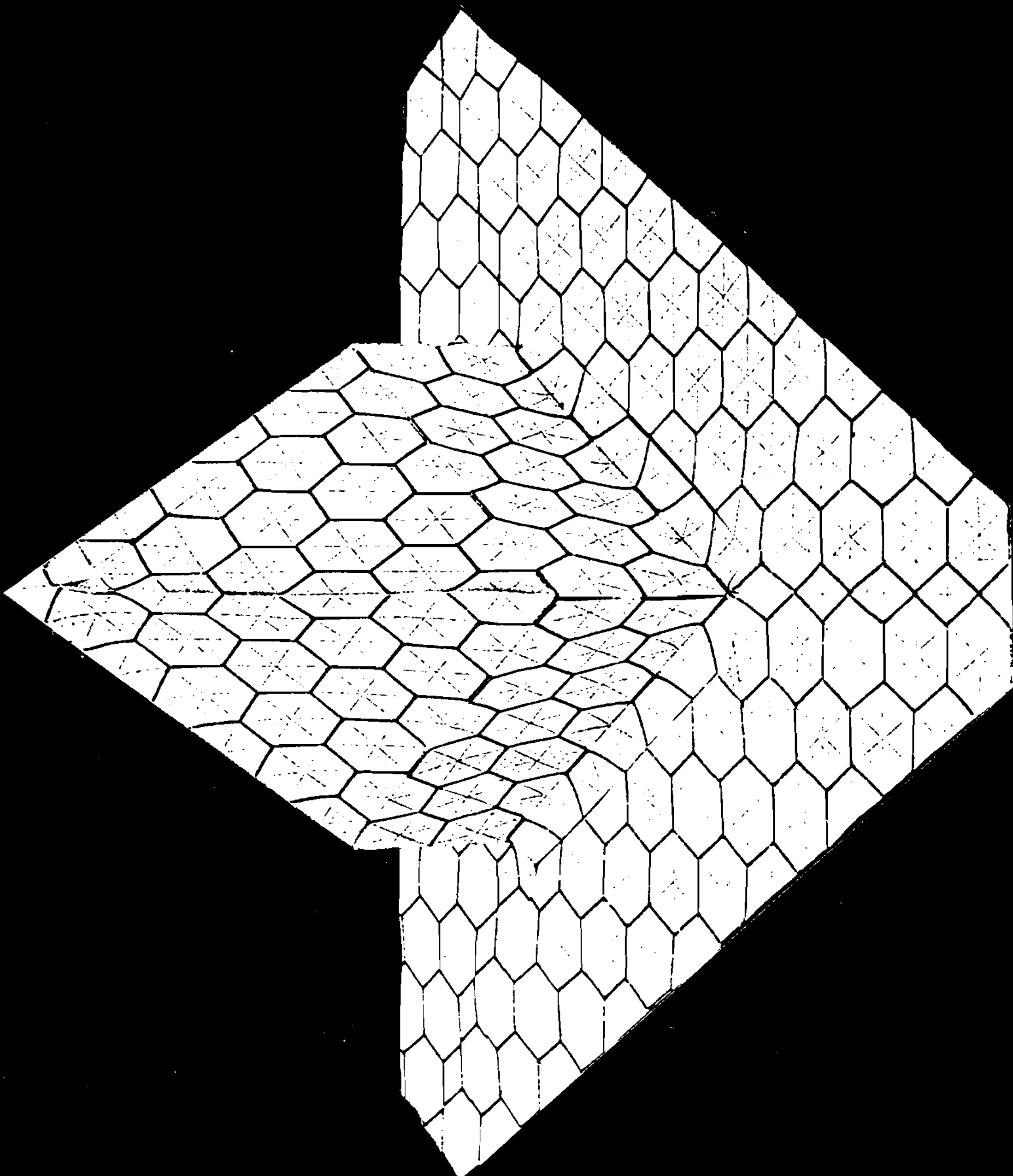


Figure 12. The London time surface.

"siamese twins" as is also apparent from Figure 12. The plane corresponding to the outer constant velocity section of the field, is glued onto the sides of the prism in three separate sections, again to ensure the correct alignment of the hexagons. This induces three spokes of diamond-shaped catchment areas where the plane sections connect. The occurrences of "siamese twins" and "diamond spokes" are examples of unavoidable local inefficiencies in the tessellations that become apparent when sections of different fields are glued together. However, the positions of the "inefficiencies" is in practice indeterminate, since there is no unique alignment procedure. The pattern of facilities obtained for London is shown in Figure 13.

The outer uniform velocity section of the velocity field joins on to the outer ring, but has not been included as it falls outside the GLC boundary. Thus, the diamond-shaped areas are absent. The "siamese twins" effect, however, is observed in the two outermost rings, involving six pairs of facilities. Because in the three-dimensional surface the cone and cylinder had to be distorted into three-sided figures, we observe some local "stretching" of catchments along radials halfway between the "siamese twins".

6. APPLICATIONS OF THE METHODS IN PRACTICE

In this section some illustrative outputs are discussed, broadly pertaining to the checklist of problems in section 2.2. For convenience, attention is focused only on the locations of the treatment facilities, although a similar analysis of ambulance depots or of both types of facility together could also be made.

6.1 Definitions

The first step is to define measures for facilities and caseloads of an A and E system distributed over an urban area.

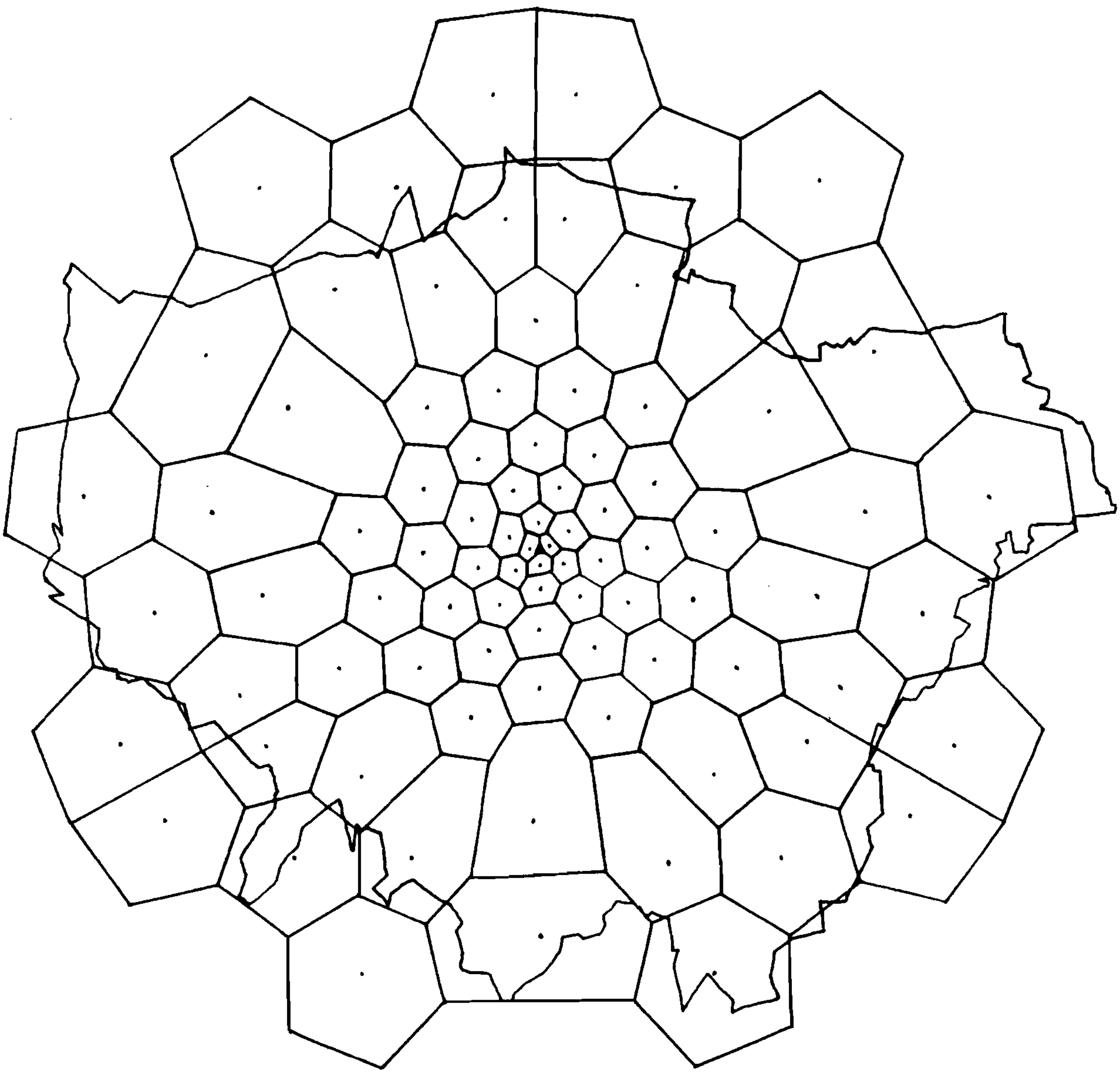


Figure 13. The optimal locations and catchment areas of treatment facilities for the London area given the estimated ambulance velocity field and a time standard of 10 minutes.

These measurements can depend on distance r from the city centre, the time t of the day and on the time standard λ of the service.

1. $D(r,t)$: the density of population
2. $h(r,t)$: the daily probability of generating an accident or emergency per capita
3. $C(r,t)$: the expected number of accidents and emergencies per unit area.

It follows that:

$$C(r,t) = h(r,t)D(r,t) \tag{17}$$

Further define:

4. $A(r)$: the area of a 1-kilometre wide ring (= $2\pi r \times 1$)
5. $N(r,t|\lambda)$: the smoothed number of facilities in a 1-kilometre wide ring given λ, the required time standard
6. $P(r,t)$: the expected number of accidents or emergencies per day in a 1-kilometre wide ring

We also have the identity:

$$P(r,t) = A(r)C(r,t) \tag{18}$$

We also define:

7. $\sigma(r,t|\lambda)$: the catchment area of an optimally located facility

An approximate estimate for σ is given by:

$$\sigma(r,t|\lambda) = V(r,t)^2 \frac{3}{2} \sqrt{3}\lambda^2 \tag{19}$$

where $V(r,t)$ is the local velocity and the remaining terms give the size of a hexagonal catchment area on the time surface.

8. $M(r,t|\lambda)$: the average daily number of patients per facility by time of day, for a given travel time standard

It follows from definitions 4, 5, and 7 that

$$N(r,t|\lambda) = \frac{A(r)}{\sigma(r,t|\lambda)} \tag{20}$$

from 3, 4, and 5 that

$$M(r,t|\lambda) = \frac{A(r)C(r,t)}{N(r,t|\lambda)} \tag{21}$$

and from 3, 4, and 6

$$P(R,t) = \int_0^R A(r)C(r,t)dr \tag{22}$$

where $P(R,t)$ is the total expected number of patients generated out to distance R from the city centre.

6.2 Estimation of the Density of Patient Caseloads

For purposes of presenting the trial illustrations, a crude estimation of caseload density $C(r,t)$ was made for the time periods t corresponding to the velocity field for London described in section 5.2. Annual caseloads at each treatment facility were regressed on distance r and the result divided by 365 to give the average daily rate. A negative exponential model gave the most satisfactory result. It was:

$$C(r,*) = 17.9 \exp(-0.1589r) \tag{23}$$

where * denotes the sum over all time periods during one day.

Figure 1 shows that the nighttime caseloads are approximately one-third of the average level over 24 hours. Thus, equation (23) was partitioned into the following:

$$C(r,t_1) = 5.96 \exp(-0.1589r) \tag{24}$$

and

$$C(r,t_2) = 20.22 \exp(-0.1589r) \tag{25}$$

where t_1 is of 4 hours duration (2:30 a.m. to 6:30 a.m.) and t_2 is of 20 hours duration (6:30 a.m. to 2:30 a.m.). It can be seen that these equations neglect any geographical redistribution of cases between day and night.

6.3. The Trade-off between Time Standards and the Numbers of Facilities

The first illustration given by the approach is a trade-off analysis to ascertain how many facilities in the city are required to meet different time standards λ for the day and night operation of the services. This number is estimated by dividing the area of the London time surface by the area of each hexagon, $\frac{3}{2}\sqrt{3}\lambda^2$. The results of the trade-off analysis are shown in Figure 14. It shows two curves: one for the day and one for night. λ is on the horizontal axis and the number of facilities is on the vertical axis.

Both curves indicate that the number of facilities required varies inversely with the square of the travel time standard. The daytime curve can be described by the equation

$$N = 7700/\lambda^2 \tag{26}$$

where the travel time is quoted in minutes. If the time standard was 10 minutes, then a minimum of 77 accident and emergency facilities need to be open during the day. At night, the number of facilities required is described by the equation

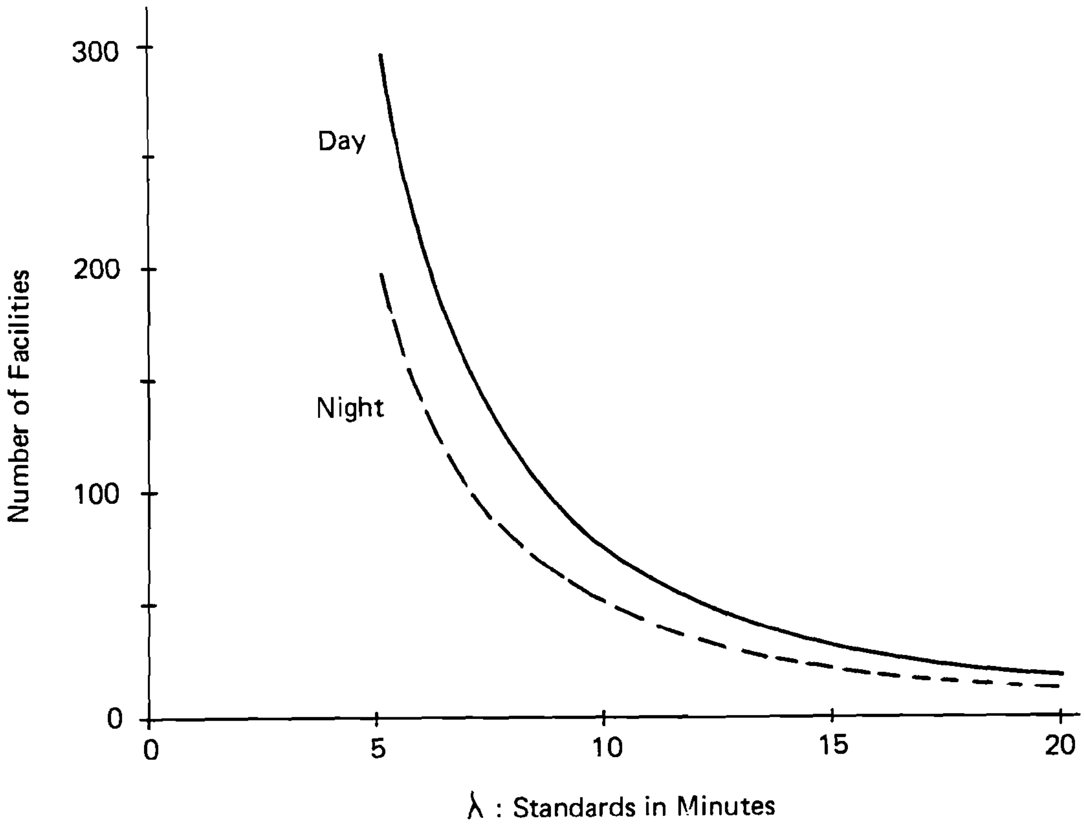


Figure 14. Variations in the number of treatment facilities with operating time standards, λ, by day and night. The curves show the minimum number of facilities necessary to completely cover the London area at a given time standard.

$$N = 5000/\lambda^2 \tag{27}$$

For a travel time standard of 10 minutes the minimum number of facilities is as low as 50.

If road traffic conditions were to continue to worsen, so that, say, ambulance travel speeds would be reduced by 10 percent, there would be either a 2-minute increase in minimum journey times or there would need to be a 20 percent increase in the number of facilities available.

6.4 Variation in the Density of Emergency Facilities by Location

The second illustrative output is a plot of $N(r,t|\lambda)$, the smoothed number of facilities, during the day and night, at distance r from the city centre, for three time standards λ: 9, 10, and 11 minutes. It is seen in Figure 15 that the resultant variation is very lumpy with significant breaks occurring at the gluing joints of the velocity fields. Between 4 kms and 15 kms, the number of facilities is constant during the day (D). This is because this part of the city is covered by the cylindrical section of the London time surface (see section 5.4). A smaller number of facilities is required at night (N), within 15 km of the centre because of the higher average velocities in operation. The consequences of varying standards on different parts of the city are thus indicated by the diagram.

6.5 Variation in Patient Caseload by Location

The third illustration in Figure 16 shows how the expected caseload at each facility, $M(r,t|\lambda)$, varies with distance from the city centre for λ equal to 9, 10, and 11 minutes. Daytime activity (D) everywhere exceeds night activity (N), with the latter at a particularly low level outer locations. It would be interesting to evaluate the effects on standards and caseloads of reducing the nighttime coverage in outer London. A lower time standard reduces the number of facilities required but increases the caseloads at individual facilities.

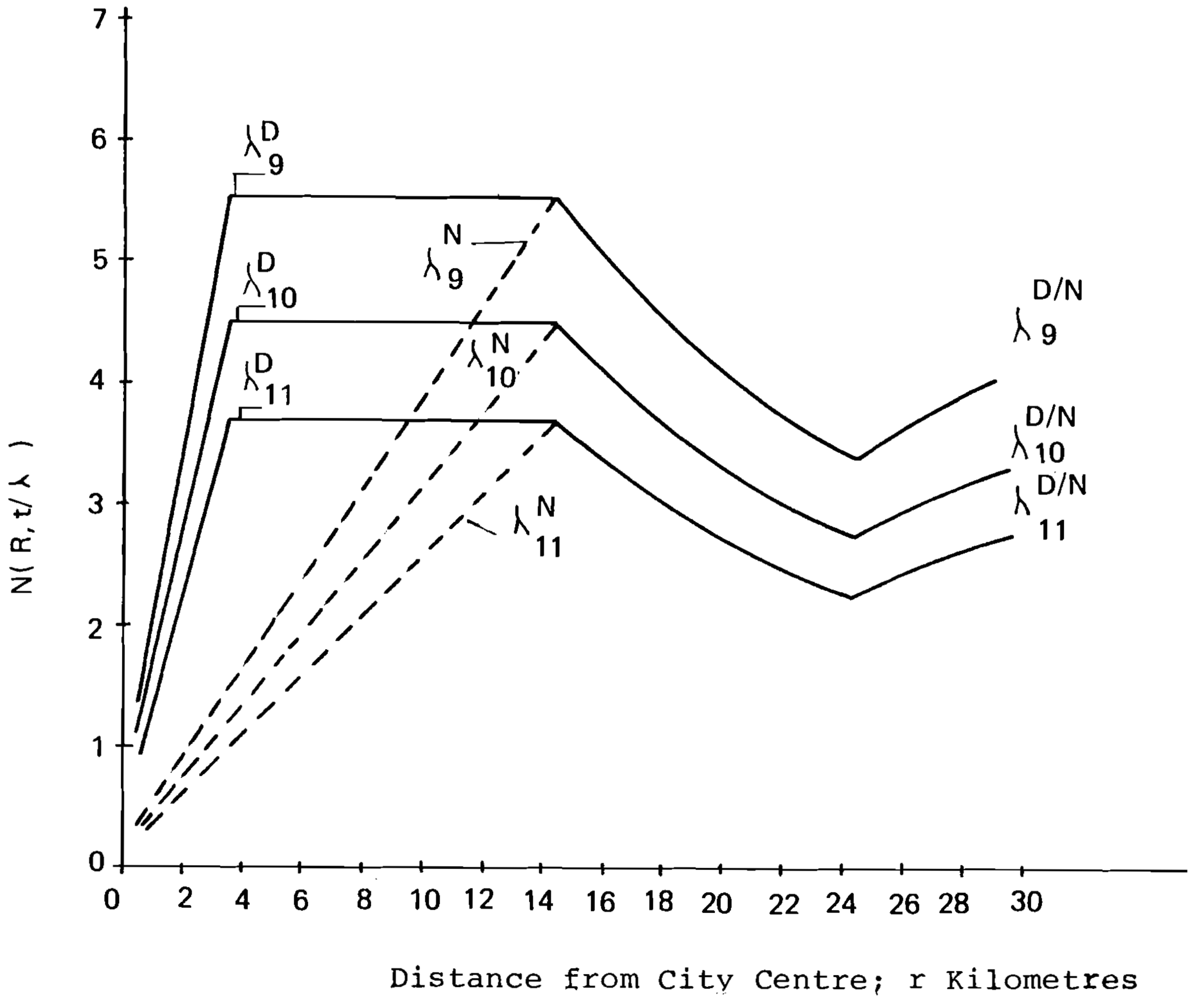


Figure 15. The variation in the minimum number of facilities $N(r,t|\lambda)$ needed at different distances from the city centre by night (N) and day (D) given time standards λ of 9, 10, and 11 minutes.

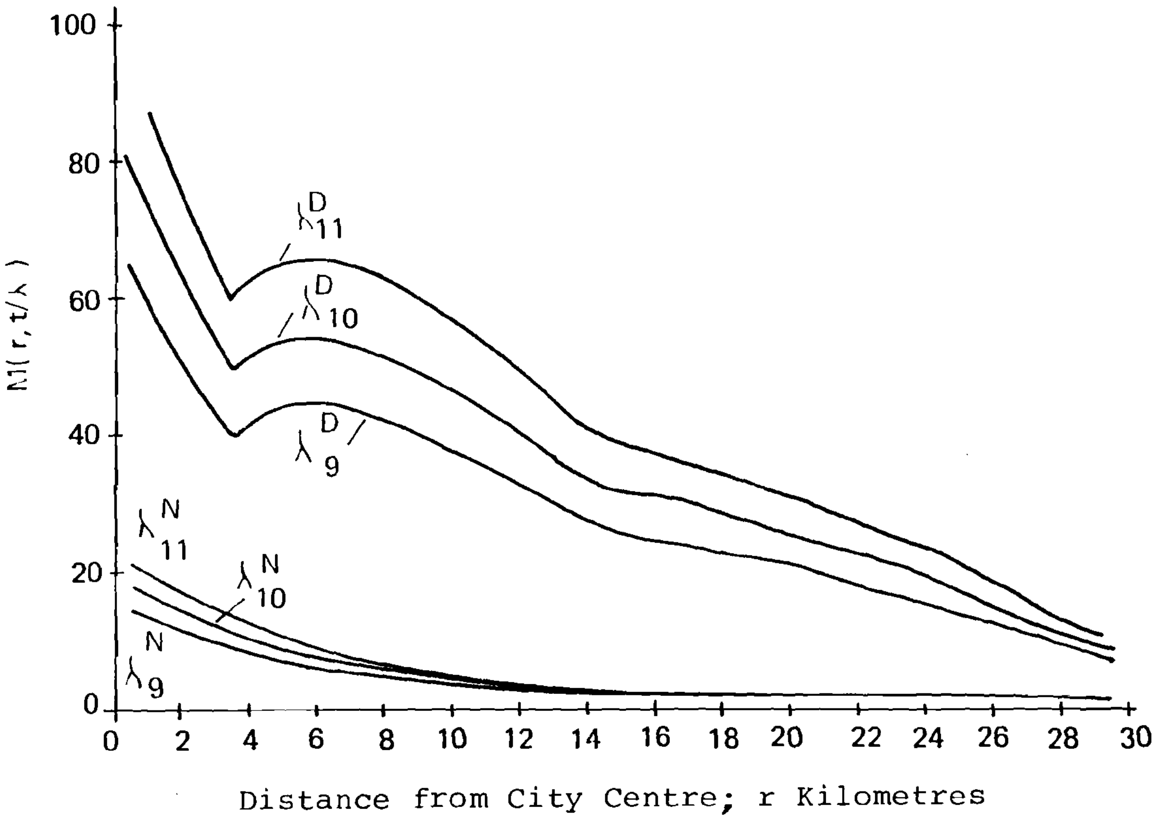


Figure 16. The expected caseload M at a facility optimally located distance r from the city centre during time period t at given time standards λ of 9,10, and 11 minutes (where D=Day and N=night).

6.6 Variation in Catchment Area by Location

The sizes of catchment areas varies with the time standard and with local average velocities. From Figure 17, it is seen that the size of catchment areas,because of the increased velocities, also increases with distance from the city centre. At night central catchments are larger, because of higher central area velocities.

6.7 Geographical Coverage with Fixed Facilities

In spite of the existence (section 5) of optimal solutions to the coverage problem, authorities must work in the short run with a fixed set of facility locations. There is a need to devise computer algorithms that can identify, from the existing set of facilities, the minimum number of centres necessary to cover the city without gaps for a particular time standard. Such algorithms have been investigated by Toregas et al. (1971) and Toregas and Revelle (1973).

6.8. The Effects on Time Standards and Caseloads of the Closure of a Single Facility

In a typical set of schedules, different facilities will be open at different times. Also, facilities sometimes unexpectedly have to close due, for example, to staff shortages. Geometrical methods can be used to simulate the effects in caseloads and time standards by an analysis of a "Dirichlet" region around each facility. A Dirichlet region bounds the area around a point, such that all other points within that area are closer to the point than to any other point (regular hexagons are examples of equi-area Dirichlet regions). Each region, analogous to a catchment area, is generated by joining up the perpendicular bisectors between neighboring points. When a facility is withdrawn from service, the Dirichlet regions are reconstituted, as is shown in Figure 18, and the areas formerly served by the closed facility are reallocated to the adjacent facilities. If travel speeds are assumed to be constant locally, this then yields an estimate of the impact on caseloads

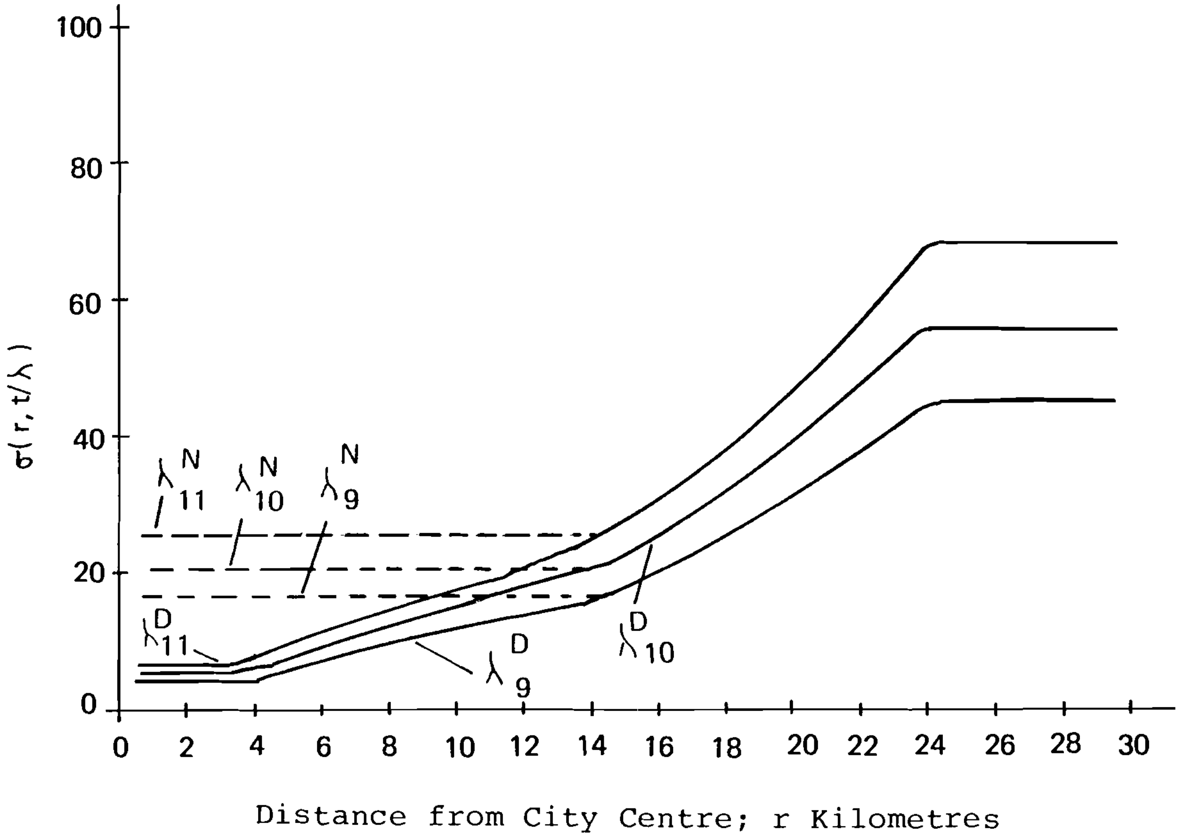


Figure 17. The increase in the catchment area of an optimally located facility with distance r from the city center for two time periods, Night (N) and day (D), given time standards of 9, 10, and 11 minutes.

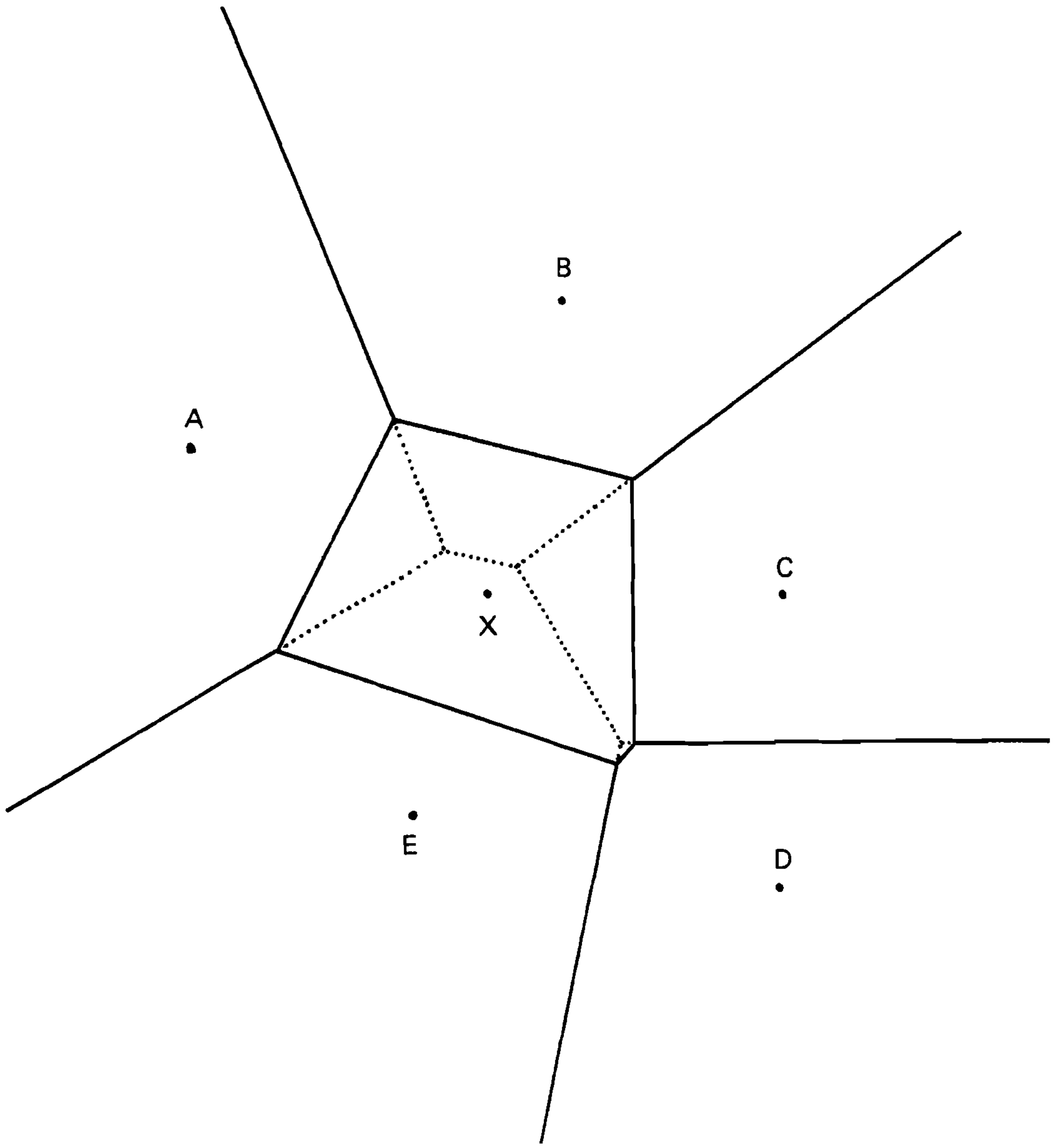


Figure 18. The local effect of withdrawing a facility X on catchment areas of neighboring facilities A,B, C,D, and E.

at the neighboring facilities. These estimates are obtained by multiplying the local caseload density by the reallocated areas. Examination of the positions of the vertices of the newly formed Dirichlet region similarly yields the effect of the reallocation on travel time standards.

7. CONCLUSIONS

In this paper we have outlined a geometrical theory of emergency service medical provision in cities. The theory was developed around the need for ambulance journeys to be made within acceptable time standards given limited facilities for treatment and varying urban traffic conditions. It was shown how simple maps for ambulance journey times can be constructed and used to pinpoint areas where coverage appears to be weak or where there are possible duplications of provision.

From our initial estimates (section 3.1), it was shown that there are probably more than sufficient treatment facilities to cover the London region, at a ten-minute time standard, although poor coverage may occur in one or two places. In sections 5.4 and 6.3, more accurate estimates showed that the *minimum* number of facilities to cover this region was 77 during the day and 50 at night. The current number of ambulance depots, 77 (unlike that number of treatment facilities), seems about right, therefore, though these depots will not, of course, be optimally located: thus, more may be needed. This aspect has still to be studied.

In section 5, a geometrical theory for the optimal location of accident and emergency facilities was developed. This was used to construct benchmarks for determining the minimum numbers of facilities that are required to cover any urban region. These minimum numbers were seen to be strongly dependent on ambulance travel time standards, while their geographical configuration was linked closely to prevailing traffic conditions. This finding highlights the need for controlling authorities to monitor the times actually achieved by ambulances, both on a regular basis and in geographical detail.

At any time the configuration of facilities will be imperfect as it takes a long time to implement decisions to invest in new facilities. Thus inefficiencies are bound to be present, and some areas are bound to be receiving a less favorable provision of emergency facilities than other areas. As traffic conditions in urban areas deteriorate, due to increased road congestion, more and more areas could experience worsening emergency cover. The scope for greater economy in facility provision could thus be eclipsed by the increased risk of deaths occurring because ambulances are unable to get patients to places of treatment in time. It thus seems to be imperative that the best possible use is made of information concerning ambulance journey times, as these provide natural measures for monitoring the changing conditions affecting the services. A regular monitoring of ambulance times will then provide a basis for determining priorities in the allocation of resources within the accident and emergency system.

REFERENCES

Anderson, R.C.T. (1978) The London Ambulance Service—an Overview and Some Details. Occasional Paper. School of Industrial and Business Studies, University of Warwick, England.

Angel, S. and G. Hyman (1976) *Urban Fields*. London: Pion.

Braakke, K.A., and N.R. Zitron (1980) Some Generalizations of the Circular City Model. *Mathematical Modelling* 1:167-179.

Cantwell, J., B. Leneham and J. O'Farrell (1973) Improving the Performance of a Local Authority Ambulance Service. Pages 443-453 in *OR '72*, edited by M. Ross. London: North-Holland Publishing Co.

Christaller, W. (1933) *Die Zentralen Orte in Süddeutschland*, Fischer, Jena. Also *Central Places in Southern Germany*, translated by C.W. Baskin. Prentice-Hall, Englewood Cliffs.

Groom, K.N. (1975) The Estimation of Emergency Ambulance Cover. Report No. 75/16. Royal Institute of Public Administration, National Health Service, Operational Research Group.

Hadley, G., and M.C. Kemp (1971) *Variational Methods in Economics*. North Holland.

Koo, D. (1977) *Elements of Optimization*. New York: Springer-Verlag.

Mayhew, L.D. (1979) *The Theory and Practice of Urban Hospital Location*. Ph.D. thesis, Department of Geography, Birkbeck College, University of London.

Mayhew, L.D. (1981) Automated Isochrones and the Locations of Emergency Medical Facilities in Cities. *Professional Geographer* 33(4):423-428.

Raitt, R. (1981) Ambulance Service Planning. Pages 239-257 in *Operational Research Applied to Health Services*, edited by D. Boldy. London: Croom Helm Ltd.

Tobler, W. (1963) Geographic Area and Map Projections. *Geographical Review* 53:59-78. (Reprinted in above Angel and Hyman reference).

Toregas, C., R. Swain, C. Revelle, and L. Bergman (1971) The Location of Emergency Service Facilities. *Operations Research* 19:1363-1373.

Toregas, C., and C. Revelle (1973) Binary Logic Solutions to a Class of Location Problems. *Geographical Analysis* 5: 145-155.

Zitron, N.R. (1974) A Continuous Model of Optimal-Cost Routes in a Circular City. *Journal of Optimum Theory Applications* 14:291-303.

APPENDIX: MATHEMATICAL NOTES

1.0 Radially Symmetric Fields

We are given a polar coordinate system, (r,θ), for the urban plane in which a velocity field, $V(r)$, has been defined; and a cylindrical coordinate system, (ρ,z,ϕ), for the space in which the time surface is located.

By theorem 3.1 in Angel and Hyman (1976:45), the following changes of variable,

$$\phi = \theta \tag{1.1}$$

$$\rho = \frac{r}{V} \tag{1.2}$$

and

$$z = \int \frac{1}{V^2}\left[2r\,V\frac{dV}{dr} - r^2\left(\frac{dV}{dr}\right)^2\right]^{\frac{1}{2}} dr + C \tag{1.3}$$

define a transformation

$$T_V : (r,\theta) \rightarrow (\rho,z,\phi)$$

This transformation maps travel time on any path P on the urban plane into the length of the image of that path, $T_V(P)$.

COROLLARY 1.1 *The time surface of the velocity field* $V(r) = \omega r$, *where* ω *is a constant, is a cylinder.*

Proof. From (1.2) and (1.3),

$$\rho = \frac{1}{\omega} \tag{1.4}$$

and

$$z = \int \frac{dr}{\omega r} + C = \frac{1}{\omega} \ln\left(\frac{r}{r_0}\right) \tag{1.5}$$

where r_0 is the radius corresponding to $z = 0$. ρ is independent of z, and the time surface is thus a cylinder of radius $\frac{1}{\omega}$ that extends over the complete z axis, from $z = -\infty$ (corresponding to $r = 0$) to $z = +\infty$ $(r = \infty)$.

COROLLARY 1.2 *Journey time* λ *in the field* $V(r) = \omega r$ *is given by*

$$\lambda_{12} = \frac{1}{\omega}\left[\ln\left(\frac{r_2}{r_1}\right)^2 + (\theta_2 - \theta_1)^2\right]^{\frac{1}{2}} \tag{1.6}$$

Proof. Open a cylinder along a generator and lie it flat on the ground. From the definition of a time surface, the minimum travel time between two points is the length of the straight line connecting them. Thus,

$$\lambda_{12} = \left[(z_2 - z_1)^2 + (\phi_2 - \phi_1/\omega)^2\right]^{\frac{1}{2}} \tag{1.7}$$

where $(\frac{1}{\omega}, z_1, \phi_1)$, $(\frac{1}{\omega}, z_2, \phi_2)$ are the coordinates of the points on the cylinder. Substituting (1.1) and (1.5) in (1.7) and simplifying, we obtain:

$$\lambda_{12} = \frac{1}{\omega}\left[\ln\left(\frac{r_2}{r_1}\right)^2 + (\theta_2 - \theta_1)^2\right]^{\frac{1}{2}} \tag{1.6}$$

the required result.

COROLLARY 1.3 *An isochrone in the field $V(r) = \omega r$ is given by*

$$r_2 = r_1 \exp\left(\lambda^2\omega^2 - (\theta_2 - \theta_1)^2\right)^{\frac{1}{2}} \tag{1.8}$$

Proof. To obtain the isochrone, simply rearrange (1.6) so that r_2 becomes the subject. By letting $(r_1\theta_1)$ be the location of the facility and λ, the time standard, and allowing θ_2 to vary, the desired isochrone may be plotted.

COROLLARY 1.4 *The equation of the quickest path in the field $V(r) = \omega r$ is given by*

$$\ln r = m\theta + C \tag{1.9}$$

where m and C are constants.

Proof. The equation of a straight line between two points $(\phi_1/\omega, z_1)$ and $(\phi_2/\omega, z_2)$ on an opened cylinder is

$$z = A + \frac{(z_2 - z_1)}{(\phi_2 - \phi_1)}\,\phi \tag{1.10}$$

where A is a constant. From (1.1) and (1.5)

$$\ln\left(\frac{r}{r_0}\right) = A + \frac{\theta}{(\theta_2 - \theta_1)}\ln\left(\frac{r_2}{r_1}\right) \tag{1.11}$$

Letting $$A + \ln r_0 = C \tag{1.12}$$

and $$\frac{\theta}{(\theta_2 - \theta_1)}\ln\left(\frac{r_2}{r_1}\right) = m \tag{1.13}$$

two constants depending on the origin (r_1, θ_1) and destination (r_2, θ_2), equation (1.12) becomes

$$\ln r = m\theta + C \qquad (1.9)$$

Equation (1.9) describes a logarithmic spiral that radiates outwards from a specified origin. A diagram showing quickest paths in r, θ coordinates is given in Figure 3b in the text.

COROLLARY 1.5 *The time surface of the velocity field* $V(r) = \omega r^p$ $(0 < p < 1)$*, where* ω *is a constant, is a cone.*

Proof. See Angel and Hyman (1976), Corollary 3.2.

COROLLARY 1.6 *Journey time in the field* $V(r) = \omega r^p$ *is given by*

$$\lambda_{12} = \frac{1}{\omega(1-p)} \left\{ r_1^{2-2p} + r_2^{2-2p} - 2r_1^{1-p} r_2^{1-p} \cos[(1-p)\theta_{12}] \right\}^{\frac{1}{2}} \qquad (1.10)$$

Proof. See Angel and Hyman (1976), Corollary 3.3.

COROLLARY 1.7 *An isochrone in the field* $V(r) = \omega r^p$ *based at* (r_1, θ_1)*, is given by*

$$r_2 = \left\{ m \cos[(1 - p)\theta_{12}] \pm \sqrt{\omega^2 \lambda^2 (1 - p) - m^2 \sin^2[(1 - p)\theta_{12}]} \right\}^{\frac{1}{1-p}} \qquad (1.11)$$

where $m = r_1^{(1-p)}$, $\theta_{12} = \theta_2 - \theta_1$.

Proof. Equation (1.11) may be rearranged into the standard form

$$0 = ax^2 + bx + c \qquad (1.12)$$

where $a = 1$

$$b = -2m \cos[(1 - p)\theta_{12}]$$

$$c = m^2 - \omega^2 \lambda^2 (1 - p)^2$$

Now
$$r_2^{1-p} = x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (1.13)$$

Substituting for a, b, and c in (1.13) and simplifying gives (1.11), the required result (see also Mayhew, 1981).

COROLLARY 1.8 *The time surface of the velocity field* $V = ar^2 + b$, *where a and b are positive constants, is a sphere of radius* $\frac{1}{\omega}$ *that satisfies*

$$\left(z - \frac{1}{\omega}\right)^2 + \rho^2 = \frac{1}{\omega^2} \qquad (1.14)$$

where $\omega^2 = 4ab.$ $\qquad (1.15)$

Proof. See Angel and Hyman (1976), Corollary 3.4.

2.0 Linear Fields

For a city where V is a function of x, V(x), it is shown in Angel and Hyman (1976:22) that the differential equation of the minimum time or quickest path is

$$\frac{dy}{dx} = \frac{KV}{\left(1 - K^2V^2\right)^{\frac{1}{2}}} \tag{2.1}$$

where K is a constant of integration.

THEOREM 2.1

$$\frac{d\lambda}{dx} = \frac{1}{V\sqrt{1 - K^2V^2}} \tag{2.2}$$

Proof.

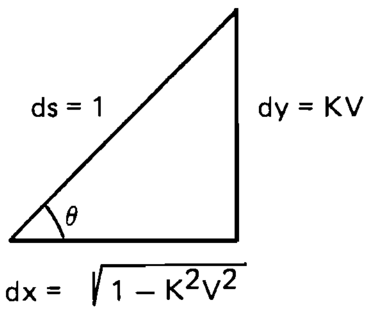


Figure 2.1. Interpretation of equation (2.1) with definition of θ.

$$d\lambda = \frac{ds}{V} \tag{2.3}$$

where λ is the journey time.

From Figure 2.1

$$(ds)^2 = (dx)^2 + (dy)^2 = 1 \tag{2.4}$$

From equation (2.3) and (2.4), therefore

$$\left(\frac{d\lambda}{dx}\right)^2 = \frac{1}{V^2}\left[1 + \left(\frac{dy}{dx}\right)^2\right] \tag{2.5}$$

From equation (2.1), we get

$$\left(\frac{d\lambda}{dx}\right)^2 = \frac{1}{V^2}\left[1 + \frac{K^2V^2}{1 - K^2V^2}\right] \tag{2.6}$$

Hence

$$\frac{d\lambda}{dx} = \frac{1}{V\sqrt{1 - K^2V^2}} \tag{2.2}$$

THEOREM 2.2 *The velocity field*

$$V(x) = a - bx \quad \begin{cases} a > 0, \ b < 0 \\ 0 \leq x < a/b \end{cases} \tag{2.7}$$

has circular quickest paths that satisfy the equation

$$(x - a/b)^2 + (y - C)^2 = \frac{1}{K^2b^2} \tag{2.8}$$

Proof. First observe that

$$\frac{dy}{dx} = \frac{dy}{dV}\,\frac{dV}{dx} = -\,\frac{bdy}{dV} \tag{2.9}$$

Using (2.1) and (2.7), the differential equation of the quickest path is obtained

$$\frac{dy}{dV} = - \frac{KV}{b\sqrt{1 - K^2V^2}} \qquad (2.10)$$

Make the substitution (see Figure 2.1)

$$\sin \theta = KV \qquad (2.11)$$

Then

$$\frac{dy}{d\theta} = \frac{dy}{dV}\frac{dV}{d\theta} = \left(- \frac{\sin \theta}{b \cos \theta}\right)\left(\frac{\cos \theta}{K}\right) = - \frac{\sin \theta}{bK} \qquad (2.12)$$

Integrating, we obtain

$$y = \frac{\cos \theta}{bK} + C \qquad (2.13)$$

$$= \frac{\left(1 - K^2V^2\right)^{\frac{1}{2}}}{bk} + C \qquad (2.14)$$

Thus $$b^2K^2(y - C)^2 = 1 - K^2V^2 \qquad (2.15)$$

Substituting for V from (2.7) and rearranging, we get

$$(x - a/b)^2 + (y - C)^2 = \frac{1}{b^2K^2} \qquad (2.8)$$

where (2.8) describes a circle of radius $\frac{1}{bK}$ with a centre $(a/b, C)$.

COROLLARY 2.2.1. *The minimum path between (x_A, y_A) and (x_B, y_B) is a circular arc which satisfies the equation in Theorem 2.2 and has parameters*

$$C = \frac{x_A^2 + y_A^2 - x_B^2 - y_B^2}{2(x_A + y_A - x_B - y_B)} \tag{2.16}$$

and
$$K = b/\left[(x_A - a/b)^2 + (y_A - C)^2\right]^{\frac{1}{2}} \tag{2.17}$$

Proof. From Theorem 2.2, we see that the minimum path satisfies both the equation

$$(y_A - C)^2 + (x_A - a/b)^2 = \frac{1}{b^2K} \tag{2.18}$$

and the equation

$$(y_B - C)^2 + (x_B - a/b)^2 = \frac{1}{b^2K} \tag{2.19}$$

Eliminating K between these equations gives us

$$(2C - x_A - x_B)(x_A - x_B) = -(2C - y_A - y_B)(y_A - y_B) \tag{2.20}$$

From which we obtain

$$C = \frac{x_A^2 + y_A^2 - x_B^2 - y_B^2}{2(x_A + y_A - x_B - y_B)} \tag{2.16}$$

as required. The equation for K in the statement of the corollary is obvious, completing the proof.

The Set of Minimum Paths

By a suitable choice of origin, the linear velocity field can be written in the form $V(x) = \omega x$. From theorem 2.2, we observe that the set of minimum paths through (x_A, y_A) is the family of circular arcs with centres on the line $x = 0$. If these arcs are extended to form complete circles, they all pass through the image point $(-x_A, y_A)$. Thus, they form co-axial circles. The isochrones are the orthogonal trajectories of these circles, and they also consist of circular arcs but with centres on the line $y = y_A$. To obtain travel times and isochrones for this field we make use of the following theorem.

THEOREM 2.3 *The minimum travel time between (x_A, y_A) and (x_B, y_B) in the field $V(x) = \omega x$ is given by*

$$t = \frac{1}{\omega}\, \mathrm{ch}^{-1}\left[\frac{x_B^2 + x_A^2 + (y_B - y_A)^2}{2x_A x_B}\right] \tag{2.21}$$

Proof.

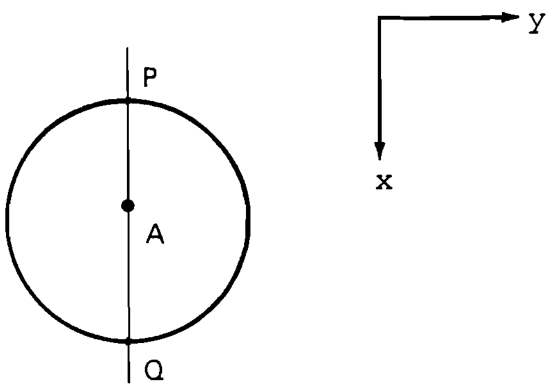


Figure 2.2. An isochrone in the velocity field $V(x) = \omega x$.

First consider travel along $y = y_A$ in Figure 2.2.

$$t = \int_{x_P}^{x_A} \frac{dx}{\omega x} = \frac{1}{\omega} \ln \left(\frac{x_A}{x_P}\right) \tag{2.22}$$

$$t = \int_{x_A}^{x_Q} \frac{dx}{\omega x} = \frac{1}{\omega} \ln \left(\frac{x_Q}{x_A}\right) \tag{2.23}$$

Thus $$x_P = x_A e^{-\omega t} \tag{2.24}$$

and $$x_Q = x_A e^{\omega t} \tag{2.25}$$

The isochrones are circles of the general form

$$(x - x_0)^2 + (y - y_A)^2 = r^2 \tag{2.26}$$

where

$$r = \frac{x_Q - x_P}{2} = x_A \, \mathrm{sh}(\omega t) \tag{2.27}$$

and $$x_0 = \frac{x_P + x_Q}{2} = x_A \, \mathrm{ch}(\omega t) \tag{2.28}$$

Hence,

$$\left(x - x_A \, \mathrm{ch}(\omega t)\right)^2 + (y - y_A)^2 = x_A^2 \, \mathrm{sh}^2(\omega t) \tag{2.29}$$

Thus, rearranging

$$t = \frac{1}{\omega} \mathrm{ch}^{-1}\left[\frac{x^2 + x_A^2 + (y - y_A)^2}{2x_A x}\right] \tag{2.30}$$

as required.

COROLLARY 2.3.1 *The equation of the isochrones in the field $V(x) = \omega x$ is given by*

$$y = y_A + \sqrt{2x_A x \ ch(\omega t) - x_A^2 - x^2} \tag{2.31}$$

Proof. This result is obtained simply by making y the subject of equation (2.30).

THEOREM 2.4 *The minimum path in the field $V(x) = ae^{-bx}$ is given by*

$$y = - \frac{\cos^{-1}(aKe^{-bx})}{b} + C \tag{2.32}$$

Proof. From (2.1)

$$\frac{dy}{dx} = \frac{KV}{\sqrt{1 - K^2V^2}} \tag{2.1}$$

$$\frac{dy}{dV} = \frac{dy}{dx} \Big/ \frac{dV}{dx} = - \frac{dy}{dx} \Big/ bV \tag{2.33}$$

$$= \frac{-K}{b\sqrt{1 - K^2V^2}} \tag{2.34}$$

Let

$$V = \frac{\cos \psi}{K} \tag{2.35}$$

so that $$\frac{dV}{d\psi} = - \frac{\sin \psi}{K} \tag{2.36}$$

$$\frac{dy}{d\psi} = \frac{dy}{dV}\frac{dV}{d\psi} = - \frac{dy}{dV}\frac{\sin \psi}{K} \tag{2.37}$$

which from (2.34) and Figure 2.1 gives

$$\frac{dy}{d\psi} = \frac{1}{b} \tag{2.38}$$

Thus,
$$y = \frac{\psi}{b} + C \tag{2.39}$$

$$= \frac{\cos^{-1}(aKe^{-bx})}{b} + C \tag{2.40}$$

THEOREM 2.5 *The minimum journey time in the field $V(x) = ae^{-bx}$ is given by*

$$\lambda = e^{bx}\sqrt{1 - K^2a^2e^{-2bx}}/ab \tag{2.41}$$

Proof.

$$\frac{d\lambda}{dV} = \frac{d\lambda}{dx} \Big/ \frac{dV}{dx} = - \frac{d\lambda}{dx} \Big/ bV \tag{2.42}$$

From Theorem 2.1, therefore

$$\lambda = - \frac{1}{b} \int \frac{dV}{V^2\sqrt{1 - K^2V^2}} \tag{2.43}$$

Let
$$V = \frac{\cos\psi}{K} \tag{2.44}$$

$$\frac{dV}{d\psi} = - \frac{\sin\psi}{K} \tag{2.45}$$

Therefore,

$$\lambda = \frac{1}{b}\int \frac{\sin\psi/K}{\dfrac{\cos^2\psi}{K^2}\sin\psi}\,d\psi \qquad (2.46)$$

$$= \frac{K}{b}\int \frac{d\psi}{\cos^2\psi} \qquad (2.47)$$

Hence $\lambda = \dfrac{K\tan\psi}{b}$ (2.48)

From Figure 2.1, therefore

$$\lambda = \frac{K\sqrt{1 - K^2V^2}}{bKV} \qquad (2.49)$$

and so $\lambda = e^{bx}\sqrt{1 - K^2a^2e^{-2bx}}/ab$ (2.41)

3.0 Geographical Obstacles to Travel

The River Problem

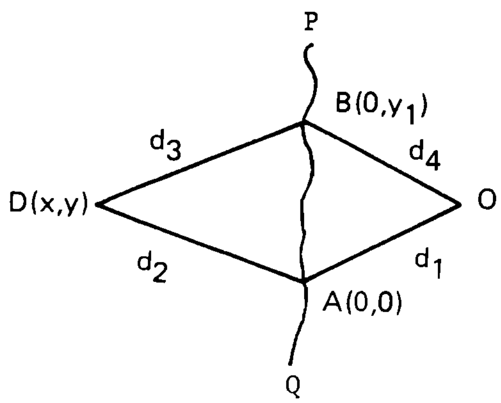


Figure 3.1. Two routes across a river PQ between O and D.

Consider Figure 3.1. PQ is the river, O and D are the origin and destination, and A and B are the bridge locations. Assume a constant speed of travel. Then travel from O to D takes the same time via either bridge when

$$d_1 + d_2 = d_3 + d_4 \tag{3.1}$$

Let $d_4 - d_1 = \varepsilon$, then

$$\varepsilon = d_2 - d_3 \tag{3.2}$$

$$= \sqrt{x^2 + y^2} - \sqrt{x^2 + (y - y_1)^2} \tag{3.3}$$

Thus $$x^2 + y^2 = \varepsilon^2 + x^2 + (y - y_1)^2 + 2\varepsilon\sqrt{x^2 + (y - y_1)^2} \tag{3.4}$$

Simplifying and ordering the terms,

$$4\varepsilon^2 x^2 + 4(\varepsilon^2 - y_1^2)y^2 + 4y_1(y_1^2 - \varepsilon^2)y = (\varepsilon^2 - y_1^2)^2 \tag{3.5}$$

The equation (3.5) is a hyperbola when $|y_1| > \varepsilon$. This inequality will be satisfied unless D lies on either extension of the line segment AB joining the bridges. In this case $y_1 = \varepsilon$ and the equation becomes $x = 0$; $y \leq 0$ or $y \geq y_1$, which describes the two extensions of the line segment AB.

4.0 Tessellations

4.1 Optimal Tessellations

The area of any polygon drawn inside a circle so that the vertices touch the circumference is maximized when all the sides of the polygon are equal. Furthermore, this area increases with n, the number of sides. We are looking, therefore, for a regular polygon that has the maximum number of sides with which to fill a plane without any overlap. The number of polygons needed will then be a minimum for any given radius. To determine the plane tessellations we let k be the number of regular n-gons meeting at a point. The angle made by each n-gon at the join is then $2\pi/k$. The internal angle of a regular n-gon equals $\pi - (2\pi/n)$. Equating, therefore:

$$\frac{2\pi}{k} = \pi - \frac{2\pi}{n} \tag{4.1}$$

we get

$$n = \frac{2k}{(k - 2)} \tag{4.2}$$

The only values of k that give integer values of n are 3, 4, and 6, as is easily verified. These correspond to a hexagon, square and triangle, respectively. Because a hexagon has the most sides, it is, by the first argument, the required regular figure.

4.2 Tessellating the Cylinder

There are two possible orientations of a pattern of regular hexagons that fits exactly around the circumference of a cylinder. If λ is the radius of each hexagon then under one orientation the pattern repeats itself in intervals of $\sqrt{3}\lambda$, in the opposite orientation the interval of repetition is 3λ. In either case the circumference of the cylinder is $2\pi/\omega$. An exact fit for the tessellation will occur when this circumference equals an integer number of interval of repetition. We therefore need to satisfy either the condition

(i) $2\pi\lambda/\sqrt{3}$ is an integer

or the condition

(ii) $2\pi\lambda/3$ is an integer

according to the orientation selected.

When transformed back to the city the circumferences of the cylinder correspond to radii, so that the number of facilities required is the same on each ring around the city centre.

4.3 Tessellations on the Cone

The total number of p-cones that admit exact hexagonal tessellations is *thirteen*, namely:

$$p = 0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1,$$
$$2, \frac{11}{6}, \frac{5}{3}, \frac{3}{2}, \frac{4}{3}, \frac{7}{6}$$

The maps on the urban plane for the field $V = wr^p$ are the *inverses* of maps for the field $V = \omega^* r^{2-p}$. To show this, consider two maps from the same cone. From equation (1.2),

$$r = (\omega\rho)^{\frac{1}{1-p}} \qquad p \neq 1 \tag{4.3}$$

$$r^* = (\omega^*\rho)^{\frac{1}{1-q}} \qquad q \neq 1 \tag{4.4}$$

Then

$$rr^* = \omega^{\frac{1}{1-\rho}} \omega^{*\frac{1}{1-q}} \rho^{\frac{2-(prq)}{(1-p)(1-q)}} \tag{4.5}$$

If $p + q = 2$, we have

$$rr^* = \left(\frac{\omega}{\omega^*}\right)^{\frac{1}{1-p}} \tag{4.6}$$

Hence $$r^* = \left(\frac{\omega}{\omega^*}\right)^{\frac{1}{1-p}} r^{-1} \tag{4.7}$$

We need only to construct the seven basic maps, for $0 \le p \le 1$, and can get the other six from the last equation. The cylinder, $p = 1$, is a special case and is self-inverse.

4.4 Tessellating the Sphere with Pentagons and Hexagons

The sum of the angles meeting at a point cannot exceed 2π radians. The angle of a regular pentagon is $3\pi/5$. Thus, no more than three faces can meet at a vertex. We therefore deduce the following theorem.

THEOREM 4.1 *Any tessellation on the surface of a sphere using only pentagons and hexagons requires at least twelve pentagons.*

Proof. Let p denote the number of pentagons and h the number of hexagons. For a tessellation on the sphere, let f be the number of faces, r the number of vertices and e the number of edges. The total number of faces is given by

$$f = h + p \tag{4.8}$$

Each hexagon has six edges, each pentagon five edges and each edge is shared by two faces, so

$$e = (6h + 5p)/2 \tag{4.9}$$

Each hexagon has six vertices and each pentagon five vertices. At least three faces meet at each vertex, hence

$$v \leq (6h + 5p)/3 \qquad (4.10)$$

The Euler characteristic of a sphere is equal to two, so that

$$f + v - e = 2$$

Hence $$(h + p) + (6h + 5p)/3 - (6h + 5p)/2 \geq 2 \qquad (4.11)$$

Therefore, $p \geq 12$ as required.

4.5 Scaling Restrictions for the Spherical Tessellations

Dodecahedron

The area of a pentagon of radius λ is $5 \sin 54^0 \cos 54^0 \lambda^2 \approx 2.38\lambda^2$. The surface area of the dodecahedron is thus $28.5\lambda^2$. The area of the spherical time surface, for $V = ar^2 + b$, is $4\pi/4ab$. For these to have similar scales we require $\frac{\pi}{ab} \approx 28.5\lambda^2$. This formula was used to calculate the parameters for Figure 11.8 in the main text.

Truncated Icosahedron

The area of a hexagon of radius λ is $\frac{3\sqrt{3}}{2}\lambda^2 \approx 2.60\lambda^2$. The area of an adjacent pentagon of edge length λ is $\frac{5 \tan 54^0}{4}\lambda^2 \approx 1.72\lambda^2$. The surface area of the truncated icosahedron, with twelve pentagons and twenty hexagons, is thus $72.6\lambda^2$. The scaling restriction is thus,

$$\frac{\pi}{ab} \approx 72.6\lambda^2$$

This formula was used to calculate the parameters for Figure 11.9 in the main text.

RECENT PUBLICATIONS IN THE HEALTH CARE SYSTEMS TASK

1. Jean-Marc Rousseau and Richard Gibbs, *A Model to Assist Planning the Provision of Hospital Services*. CP-80-3.

2. Peter Fleissner, Klaus Fuchs-Kittowski, and David Hughes, *A Simple Sick-Leave Model Used for International Comparison*. WP-80-42.

3. Philip Aspden, Richard Gibbs, and Tom Bowen, *DRAM Balances Care*. WP-80-43.

4. Philip Aspden and Martin Rusnak, *The IIASA Health Care Resource Allocation Submodel: Model Calibration for Data from Czechoslovakia*. WP-80-53.

5. Pavel Kitsul, *A Dynamic Approach to the Estimation of Morbidity*. WP-80-71.

6. Evgenii Shigan and Pavel Kitsul, *Alternative Approaches to Modeling Health Care Demand and Supply*. WP-80-80.

7. David Hughes and Andrzej Wierzbicki, *DRAM: A Model of Health Care Resource Allocation*. RR-80-23.

8. Philip Aspden, *The IIASA Health Care Resource Allocation Submodel: DRAM Calibration for Data from the South West Health Region, UK*. WP-80-115.

9. Leslie Mayhew and Ann Taket, *RAMOS: A Model of Health Care Resource Allocation in Space*. WP-80-125.

10. Leslie Mayhew, *The Regional Planning of Health Care Services: RAMOS and RAMOS^{-1}*. WP-80-166.

11. Zenji Nanjo, *A Simple Method of Measuring the Increase of Life Expectancy when a Fixed Percent of Deaths from Certain Causes are Eliminated*. CP-80-35.

12. Mark Pauly, *Adding Demand, Incentives, Disequilibrium, and Disaggregation to Health Care Models*. WP-81-4.

13. Leslie Mayhew, *DRAMOS: A Multi-Category Spatial Resource Allocation Model for Health Service Management and Planning*. WP-81-39.

14. Leslie Mayhew and Ann Taket, *RAMOS: A Model Validation and Sensitivity Analysis*. WP-81-100.

15. Leslie Mayhew and Giorgio Leonardi, *Equity, Efficiency, and Accessibility in Urban and Regional Health Care Systems*. WP-81-102.

16. Leslie Mayhew, *Automated Isochrones and the Location of Emergency Medical Services in Cities: A Note*. WP-81-103.

17. Michał Bojańczyk and Jacek Krawczyk, *Estimation and Evaluation of Some Interdependencies of Environmental Conditions, Welfare Standards, Health Services, and Health Status*. CP-81-29.

18. Michal Bojanczyk and W. Rokicki, *A Concept of Modeling Manpower Educational System*. CP-82-3.

19. Margaret Pelling, *A Multistate Manpower Projection Model*. WP-82-12.

20. Philip Aspden, L. Mayhew, and M. Rusnak, *DRAM: A Model of Health Care Resource Allocation in Czechoslovakia*. RR-82-6. Reprinted from OMEGA: The International Journal of Management Science, 9(5):509-518.